

BIOCHEMICAL ANALYSIS OF EVOLUTIONARY DIVERGENT VERTEBRATE
LARP6 PROTEINS

by

Jose Miguel Castro, 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 2017

Committee Members:

Karen A. Lewis, Chair

Lisa R. Warner

Ronald B. Walter

Steve T. Whitten

COPYRIGHT

by

Jose Miguel Castro

2017

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

This work is protected by the Copyright Laws of the United States (Public Law 94-553, section 107). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgement. Use of this material for financial gain without the author's express written permission is not allowed.

Duplication Permission

As the copyright holder of this work I, Jose Miguel Castro, authorize duplication of this work, in whole or in part, for educational or scholarly purposes only.

DEDICATION

Para mi familia.

ACKNOWLEDGEMENTS

When I started this journey, in my head, I had the idea that this would be a one person ride. However, since day one, I knew it was just a thought. There have been several people behind me, supporting me, which I'd like to acknowledge. First, I would like to mention that even though my direct family was not with me each and every day, their love, and support was felt through the entire process. I would like to thank my first PI, Dr. Linette Watkins, for allowing me to join her lab as an undergraduate with no research or biochemical experience whatsoever. She believed in me and taught me the foundations of my future career. Furthermore, I would like to thank my current mentor and PI, Dr. Karen A. Lewis. When I interviewed with her to decide if the lab was a good fit for me, her curiosity and drive for learning was the spark that encouraged me to join her lab. We had our ups and downs the first year. However, as time passed and we started to know each other, I knew I had made one of the best decisions in my life. I am very thankful for everything she has taught me. She is such an inspiration. Every moment I spent talking with her left me in amazement of her well rounded knowledge. Yes Dr. Lewis, I still want to be like you when I grow up. I promise to keep being myself in the process.

I am very grateful for my lab family. I have learned from every single one of them. I felt like their brother (perhaps some of them saw me as a dad). Eliana

Peña is a very interesting friend (we need to keep in touch). She is the one that I spent gossiping with and would talk about life in general. She has the right combination of brains and sassiness, which I love. You will do great girl. Next, Eliseo Salas. When I first met him he seemed so quiet. Little did I know he is not. He became such a close friend (like a brother). We both learned a lot from each other. While I helped him with filters, he taught me to be myself and not care a lot about what people may think about you. I wish him the best on his new journey. Another lab mate to mention is Leticia (Lety) Gonzalez. She might be the new lab technician, however, in such a short time I felt very comfortable around her. Those trips to the vending machine in which we got to loosen up a little were great. Will miss her.

My Laredo-San Marcos Family. Thank you for making this journey less difficult. Carolina, Emma, Jorge, Karina, Martha, Valeria and Rudy, I am glad we spent time helping each other out in this town. We made it our new hometown.

I would like to specially thank Cally Moore, Elizabeth McIver and Leona Martin for all the chatting and time we spent together. The four of us would get together for coffee once in a while and catch up. I am very grateful for their friendship. I am proud for each of them and hope we keep in touch.

Lastly, I will like to thank a very special person, a “new” friend and someone who helped me through the struggle and pushed me to become a

better person, Edgar Tovar. I am very thankful for all the time he spent with me in this process. From when I started my upper education in community college to my years here at San Marcos, he was always there for me. Sometimes his method of trying to encourage me was not the adequate one, but I knew that he wanted me to succeed.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	v
LIST OF TABLES	ix
LIST OF FIGURES	x
ABSTRACT	xiv
CHAPTER	
I. INTRODUCTION	1
II. MATERIALS AND METHODS	22
III. RESULTS	42
IV. DISCUSSION	94
REFERENCES	103

LIST OF TABLES

Table	page
1. Size Exclusion Chromatography S200 Standards	22
2. Size Exclusion Chromatography S75 Standards	23
3. Primers used in this thesis work	25
4. Plasmids used in this thesis work	25
5. Cell lines and respective antibiotics	26
6. List of purification buffers	34
7. Selection of additives for filter assay	38

LIST OF FIGURES

Figure	Page
1. Common RNA binding domains	4
2. Schematic representation of the LARPs.....	7
3. Sequence alignment of La Modules in human LARPs.....	8
4. Phylogenetic comparison between La motifs and RRM.....	10
5. Human collagen type I mRNA 5' untranslated region (UTR)	11
6. LARP6 morpholino knockdown effect in zebrafish embryogenesis	12
7. Structure of Human LARP6 La motif	13
8. Structure of Human LARP6 RRM	15
9. Sequence alignment of <i>H. sapiens</i> , <i>X. maculatus</i> , and <i>D. rerio</i> LARP6 proteins	19
10. Solubility screening overview.....	36
11. Expression trials of <i>HsLARP6</i>	42
12. Full length <i>HsLARP6</i> Affinity chromatography fractions	44
13. S200 size exclusion chromatogram of <i>HsLARP6</i>	45
14. SDS-PAGE analysis of fractions from size exclusion chromatography of full-length <i>HsLARP6</i>	46
15. Verification of <i>HsLARP6</i> stability at -70 °C	47
16. Affinity purification of His ₆ -UlpI	48
17. S75 size exclusion chromatogram of His ₆ -UlpI.....	49
18. SDS-PAGE analysis of size exclusion chromatography of His ₆ -UlpI	50
19. SDS-PAGE analysis of initial UlpI cleavage of His ₆ -SUMO- <i>DrLARP6a</i>	51
20. SDS-PAGE analysis of second UlpI cleavage of His ₆ -SUMO- <i>DrLARP6a</i>	52
21. <i>XmLARP6</i> amplification from cDNA and confirmation of successful double digest	53

22. <i>DrLARP6a</i> amplification from synthetic DNA and pET28SUMO- <i>DrLARP6a</i>	54
23. <i>DrLARP6b</i> amplification from synthetic DNA and pET28SUMO- <i>DrLARP6b</i> double digest	54
24. Expression trials of <i>XmLARP6</i> in Rosetta™ (DE3) pLysS competent cells ..	55
25. Expression trials of <i>DrLARP6a</i> in Rosetta™ (DE3) pLysS competent cells .	56
26. Expression trials of <i>DrLARP6b</i> in Rosetta™ (DE3) pLysS competent cells .	56
27. His ₆ -SUMO- <i>XmLARP6</i> affinity chromatography fractions profile	58
28. S200 size exclusion chromatography of SUMO-tagged <i>XmLARP6</i>	60
29. SDS-PAGE analysis of fractions from size exclusion chromatography of SUMO-tagged <i>XmLARP6</i>	60
30. Initial screening of purification buffer additives to increase protein solubility	63
31. Second iteration of the additive screen, Part 1	64
32. Second iteration of the additive screen, Part 2	65
33. Effect of glycerol in combination with urea and NaI	67
34. Effect of glucose in combination with urea and NaI	68
35. His ₆ -SUMO- <i>XmLARP6</i> affinity chromatography fractions using optimized buffer conditions	69
36. S200 size exclusion chromatogram of His ₆ -SUMO- <i>XmLARP6</i> , with optimized buffer conditions.....	71
37. SDS-PAGE analysis of fractions from size exclusion chromatography of His ₆ -SUMO- <i>XmLARP6</i> with optimized buffer	71
38. Ulpl cleavage of His ₆ -SUMO- <i>XmLARP6</i>	72
39. S200 size exclusion chromatogram of <i>XmLARP6</i> after SUMO cleavage.....	74
40. SDS-PAGE analysis of fractions from size exclusion chromatography of <i>XmLARP6</i> post-SUMO cleavage	74

41. His ₆ -SUMO- <i>Dr</i> LARP6a affinity chromatography fractions using optimized buffer conditions	75
42. S200 size exclusion chromatogram of His ₆ -SUMO- <i>Dr</i> LARP6a with optimized buffer conditions	76
43. SDS-PAGE analysis of fractions from size exclusion chromatography of His ₆ -SUMO- <i>Dr</i> LARP6a with optimized buffer	77
44. Ulpl cleavage of His ₆ -SUMO- <i>Dr</i> LARP6a after size exclusion chromatography.....	78
45. S200 size exclusion chromatogram of <i>Dr</i> LARP6a after SUMO cleavage.....	79
46. SDS-PAGE analysis of fractions from size exclusion chromatography of <i>Dr</i> LARP6a post SUMO cleavage	80
47. His ₆ - <i>Xm</i> LARP6 La Module affinity chromatography fractions.....	81
48. S200 size exclusion chromatogram of His ₆ - <i>Xm</i> LARP6-LAM	82
49. SDS-PAGE analysis of fractions from size exclusion chromatography of His ₆ - <i>Xm</i> LARP6-LAM	83
50. Temperature trials of limited proteolysis of <i>Hs</i> LARP6	85
51. Topology studies of full-length LARP6 proteins by limited proteolysis	87
52. Topology studies of <i>X. maculatus</i> La Module by limited proteolysis	88
53. Limited trypsinolysis of <i>Xm</i> LARP6 and <i>Dr</i> LARP6a and gel extraction for coupled liquid chromatography-mass spectroscopy	89
54. Histograms of peptide species found by LC-MS from the 40 kDa trypsinolysis products from <i>Xm</i> LARP6 and <i>Dr</i> LARP6a.....	90
55. Binding activity of LARP6 proteins for human collagen 1A1 mRNA stem loop	92
56. RNA binding activity against biotinylated <i>Hs</i> COL1A1	93
57. Disorder prediction of LARP6 proteins	97

58. His ₆ -SUMO- <i>Dr</i> LARP6b affinity chromatography fractions using optimized buffer conditions	98
59. Disorder prediction of <i>Dr</i> LARP6b protein	99
60. Human LARP6 chimera constructs scheme	102

ABSTRACT

The La-related proteins (LARPs) contain a conserved bipartite RNA binding domain, called the La module, which consists of the La motif and an RNA recognition motif (RRM). Across the LARP subfamilies, the La module interacts with a variety of RNAs, allowing this protein superfamily to perform roles as chaperone proteins, post-transcriptional regulatory factors, regulators of RNA metabolism. The ligand specificity of the LARP6 subfamily is still unknown. The only endogenous ligand for LARP6 is found in the 5' untranslated region of collagen type I mRNA. Animal studies confirmed that the binding of a particular RNA structure by LARP6 was critical for collagen production and muscle development. In this study, we are using the LARP6 proteins from zebrafish (*Danio rerio*), platyfish (*Xiphophorus maculatus*) and human to carry out detailed biochemical analyses of vertebrate LARP6 RNA binding activity. The purpose is to determine if there is a species-specific protein-RNA interaction due to co-evolution of the ligand-binding surface within the LARP6 RRM and the RNA ligand sequence and/or structure. Recombinant fish and human LARP6 proteins were expressed in *E. coli* with an N-terminal SUMO tag. A filtration screening assay was used to identify additives that increase solubility and stability to the fish proteins. To robustly compare direct RNA binding affinity of the fish proteins to the human LARP6, the SUMO tag was cleaved by Ulp1. Limited proteolysis studies were used to compare the domain topology of the fish LARP6 proteins to

the human homolog. The RNA binding activity of all LARP6 proteins for the COL1A1 stem loop RNA from the appropriate species was measured by electrophoretic mobility shift assays (EMSAs).

I. INTRODUCTION

RNA Binding Proteins

The “central dogma” of molecular biology states that biological information flows from DNA to RNA to protein. Proteins can function as antibodies, enzymes, messengers, structural components, transporters, just to name a few. There is a distinct group of proteins that interact with RNA called RNA binding proteins. There are many structural and sequence level motifs that allow proteins interact with a variety of RNAs.¹ Since RNA structure is less stable and more dynamic than its DNA counterpart.² As a result, the protein motifs that bind to RNA may be more complicated when compared to the ones that bind DNA.²

In contrast to the DNA-protein interactions, in which backbone interactions are dominant to stabilize the complex, RNA-protein complexes are stabilized by protein contacts with both the RNA backbone and RNA bases. However, because of the complicated tertiary structure some RNAs can have, the packing density of the complex is not as favorable as the ones seen from DNA binding proteins. The structural mechanisms of the interactions can be used to generally divide the RNA binding motifs into two classes: (1) groove-binding proteins, which primarily use α -helices to bind the RNA, and (2) β -sheet-binding proteins, which use β -sheets to coordinate functional groups on unpaired bases.² Some domains commonly found in RNA binding proteins are the RNA recognition motif (RRM), K-homology (KH), Piwi/Argonaute/Zwille (PAZ) domain, RNA binding zinc fingers (ZnF), and S1 domains (Figure 1).³

The RNA Recognition Motif (RRM) is the most abundant RNA binding domain (Figure 1A)⁴. Even though there is a large number of variants that have been found, the canonical structure is composed of four antiparallel β -sheets and two helices in the $\beta\alpha\beta\beta\alpha\beta$ fashion⁴. RNA binding by the RRM is usually managed by conserved arginines and/or lysines. The β -strands that contain the Arg or Lys interact with the phosphate backbone of the RNA by forming salt bridges. In addition, canonical phenylalanine residues located in the β -strands contribute to nucleic acid stability by inducing π - π stacking with the RNA bases.⁵

The PAZ domain is an RNA binding domain that interacts with the 3' end of siRNAs and miRNAs, named after the PIWI, Argonaute, and Zwiille proteins. Its overall structure is composed of a β -barrel against an α - β (Figure 1B). In tandem, the motifs work together to bind RNA.⁶ This interaction happens with no specificity in sequence. One of the primary purposes is to cleave or inhibit mRNA translation.⁷

The K Homology (KH) was identified primarily in the human heterogeneous nuclear ribonucleoprotein K (hnRNPK). The KH domain can identify and bind to RNA. The phosphate backbone coordinates and binds to KH domain in a pocket formed by α -helices 1 and 2 and the conserved sequence GXXG (Figure 1C). In contrast to the typical π -stacking interaction seen by several RNA binding proteins, KH domains recognize ssRNA by hydrophobic interactions between non-aromatic, nonpolar residues and the ribose sugars.⁸

Zinc fingers are classified based on the residues used to coordinate the zinc ion. As an example, the CCHH zinc fingers use 2 cysteines and 2 histidines.

In addition, there are CCCH and CCHC zinc fingers. Of these, CCCH zinc fingers have been found to interact with ssRNA; most zinc fingers bind to DNA. The interaction happens by sequence-specific RNA recognition to the protein backbone by hydrogen bonding. The protein shape plays an important role in the interaction (Figure 1D). The flexibility and rigidity of the structure allows the zinc fingers to generate a unique hydrogen bonding matrix that recognizes specific sequences.⁹

The S1 domain is composed of a β -barrel formed by 5 antiparallel β strands, and a right handed 3_{10} helix (Figure 1E).¹⁰ When binding, the S1 domain identifies the RNA by aromatic residues stacking interactions between the nucleotides and two out of the five β -strands. The loops present in the S1 domain serve as structural elements to provide proper orientation of the protein and the RNA.¹¹

Individual domains can interact and bind to different or specific RNAs by tuning both specificity and affinity. These features could be achieved by using an expanded RNA binding surface. A larger surface area can be buried in a protein/RNA interaction, often by lengthening or ordering loops adjacent to secondary structures. Such augmentation of the binding site would increase the points of contact with the ligand, which could in turn increase both binding affinity and specificity.¹²

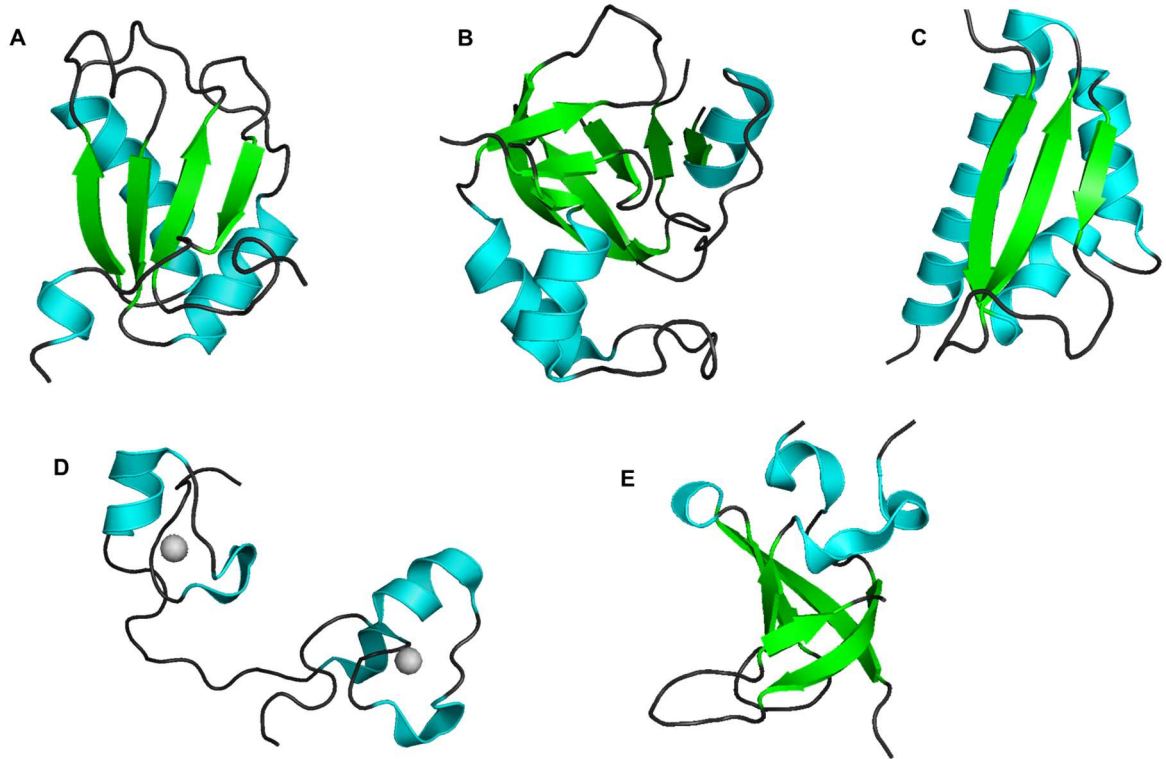


Figure 1. Common RNA binding domains. RNA binding domains have distinctive structural features that provide necessary surface area to interact with different specificity to their ligands. A) RNA Recognition Motif (RRM, PDB: 1URN). B) The PIWI, Argonaute and Zwiille domain (PAZ, PDB: 1SI3). C) The K Homology domain (KH, PDB: 1EC6). D) Zinc Finger (ZnF-CCCH, PDB: 1RGO). E) S1 domain (PDB: 1SN8). All structures made using PyMOL.

In contrast to single binding domains, multiple domain RNA binding proteins display higher RNA specificity. With tandem domains, there is both an increased surface area and potential for more specific interactions to occur. This increases the ability of a protein to recognize longer ligands in a sequence- and structure- specific manner. Additionally, the orientation of the protein domains can modulate binding activity. The flexibility and mobility of the domains relative to each other are usually moderated by interdomain linkers. The linker can either be rigid, in which it holds the domains in a specific orientation, or flexible, which would allow the domains to interact in a dynamic manner that can alter the surface area as needed. By linking two RNA binding domains together, the

overall affinity for RNA is increased. When an RNA binding domain is bound to the RNA ligand, the linker will localize the other domain closer to the ligand, which lowers the energy required for binding to occur by the second domain.¹⁴ The length of the linker could also contribute to a modulation of RNA affinity, as well as preventing or allowing the protein to bind to more than one RNA molecule.¹⁵

La and La-Related Proteins

Genuine La Protein (also called SSB and LARP3) – During genetic studies of the autoimmune diseases lupus erythematosus and Sjörger's syndrome, in which the immune system attacks normal healthy tissue, researchers discovered a protein that they named "La protein" or "SS-B". The name of the patient in whose tissue the La antibody was detected gave the La protein its name, while "SS-B" stands for Sjörger's syndrome antigen B.¹⁶

Experiments that used the anti-La antibodies for immunoprecipitation revealed that the La protein associates with a variety of small RNAs that are transcripts from RNA polymerase III.¹⁶ *Homo sapiens* La (*HsLa*) binds to these transcription products by interacting with the 3'-UUU_{OH} sequence that is characteristic of RNA Pol III transcripts. The terminal uridylate tail is removed during the RNA maturation process.¹⁷ Together, these observations indicated that *HsLa* protein plays an important role during the maturation, folding, and the distribution of these RNAs.

Since the discovery of this "genuine La" protein, several closely-related proteins have been identified and named "La-related proteins" (LARPs). In addition to the La family, four LARP families have been distinguished: LARPs 1, 4, 6, and 7.¹⁸ Because of the additional members, genuine La is now known as LARP3.¹⁹

All LARPs contain a conserved RNA binding domain called the La module (LaM) located roughly in the N-terminal part of the protein. This module consists of two important subdomains: the N-terminal La motif and an RRM (Figure 2).

These two domains are connected by a small linker sequence. A multiple sequence alignment of LARP La Modules shows divergence in both length and chemical characteristics (Figure 3). Outside of the La Module, individual LARP subfamilies are distinguished by other structural and functional domains (Figure 2).

Although theLARPs were first characterized as human proteins, genome sequencing has revealed homologs in a wide variety of eukaryotes (Figure 4). The structural and evolutionary relationship between the La-related proteins can be organized based on sequence similarity (Figure 4).

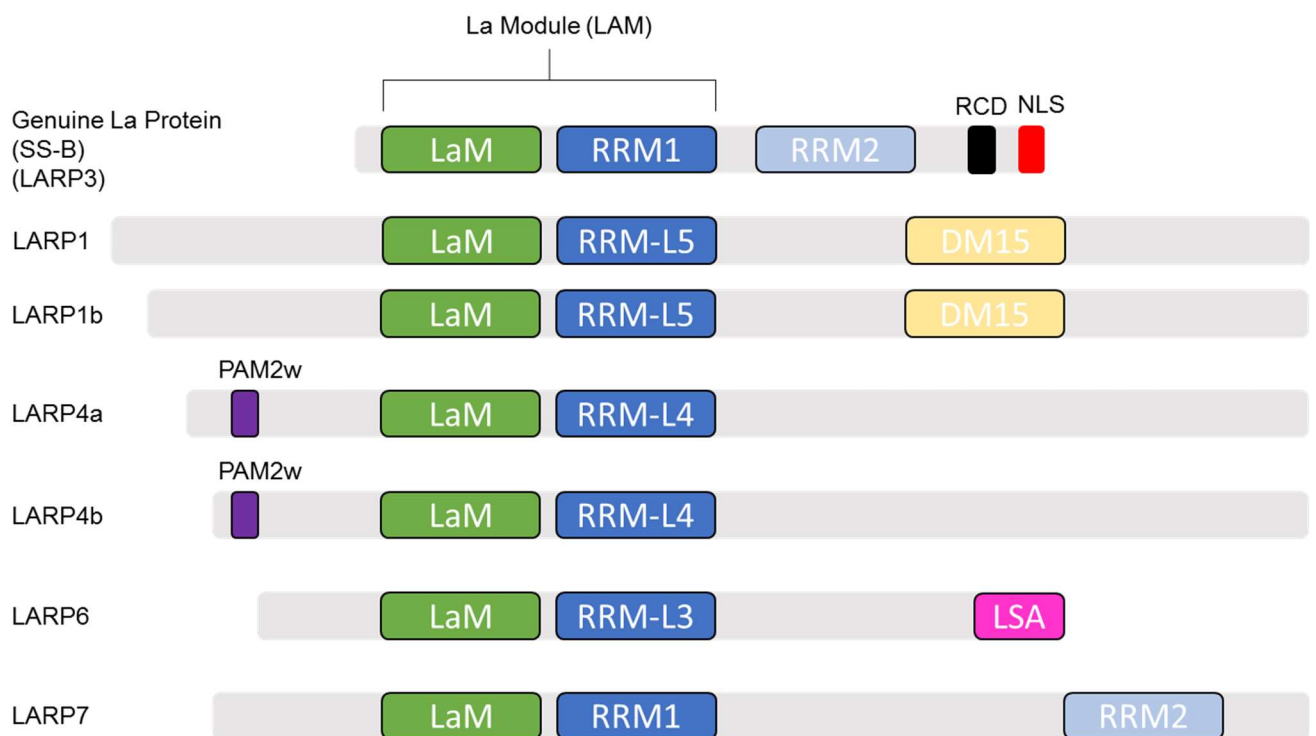


Figure 2. Schematic representation of the LARPs.

All LARPs share a La Module composed of a La motif (LaM) and a RNA Recognition Motif (RRM). In addition, each LARP has other domains which allow them to generate different functions. Abbreviations: DM15-repeat containing region (DM15); nuclear localization signal (NLS); an atypical Poly(A) Binding Protein interaction motif (PAM2w); RNA chaperon domain (RCD); RNA recognition-like motif (RRM-L); La and S1 associated motif (LSA). (Modified from Stavra and Blagden, 2015.)¹⁹

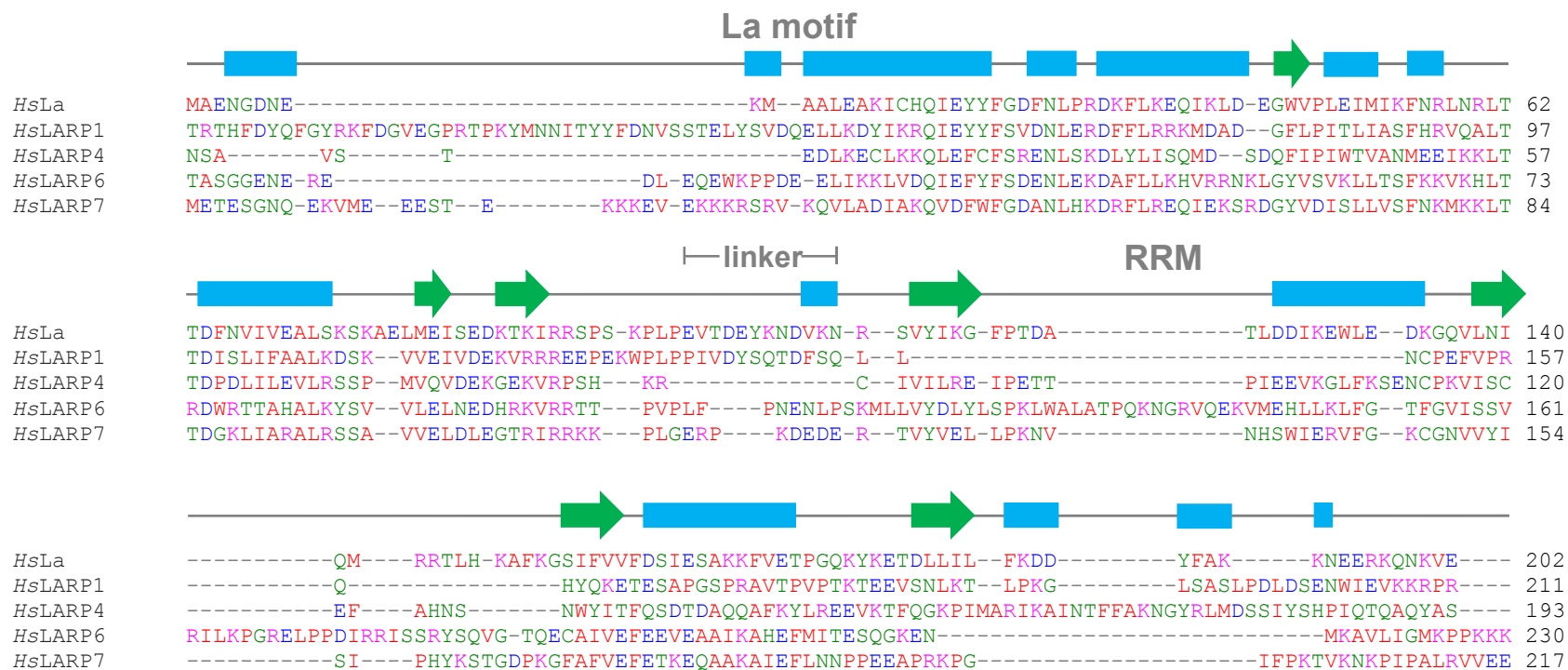


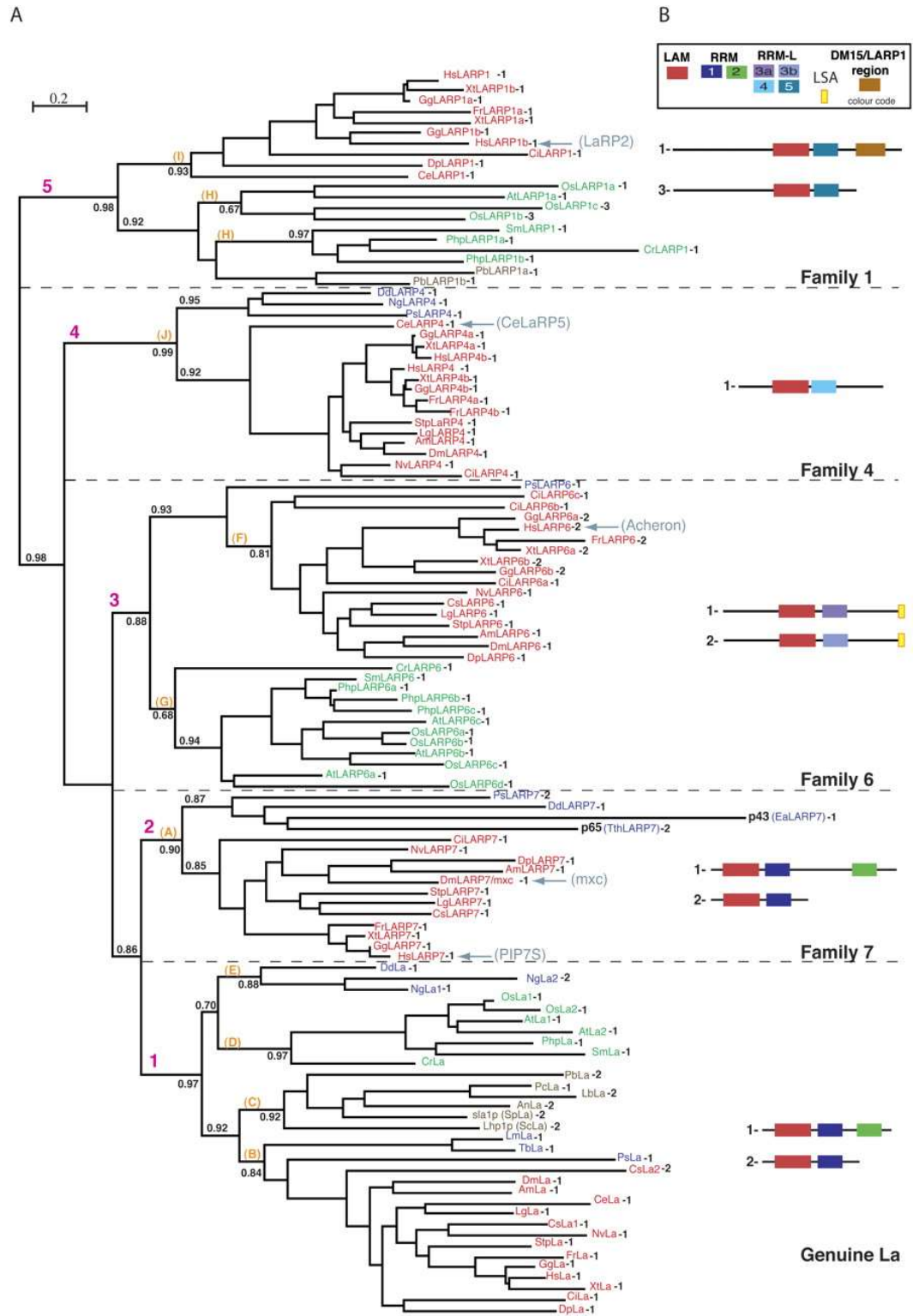
Figure 3. Sequence alignment of La Modules in human LARPs. Human genuine La (LARP3) La Module secondary structure scheme is used to visualize the variability between LARPs La modules sequences. Sequence alignment generated by Clustal Omega. Figure adapted from Martino, L. *et al* (2015).²⁰

LARP7 (also called PIP7S, and HDCMA18) — The LaM in the HsLARP7 family is most closely related to genuine HsLa, conserving most of the amino acids required for the 3'-UUU_{OH} interaction. This allows LARP7 to have similar target recognition as genuine La.¹⁸ Similar to genuine la, LARP7 subfamily members contain a second RRM located in the C-terminus. LARP7 may be a tumor-suppressor as deletion of the LARP7 gene caused an increase in tumor development in gastric tissue.²¹

LARP1 and LARP4 —LARPs 1 and 4 are involved in mRNA metabolism and translation within the cytoplasm.²² Both have been found to interact with the 3'- polyA tails in pre-mature mRNA. LARP4 does this by a motif called Positive Allosteric Modulator 2 (PAM2) located N-terminal to the LaM, which interacts with Poly-A Binding Proteins (PABP).²³ In addition, the DM15 domain of LARP1 binds to the 5' methyl guanosine cap of mature mRNAs.²⁴ Together, these categorize LARP1 and LARP4 as post-transcriptional regulatory factors in the cytosol.²⁵

(following page) Figure 4. Phylogenetic comparison between La motifs and RRMs.

LARP sequences were organized into a phylogenetic tree to understand their evolutionary relationship. Blue: proteins from protists; green: proteins from plants; brown: proteins from fungi; red: proteins from animals. Reproduced with permission under Creative Commons License from (14). Courtesy of the Cold Spring Harbor Laboratory Press for the RNA Society.



LARP6 (also called Acheron) — In contrast to LARP7 and genuine La, and LARP1 and 4, the LARP6 La Module differ more in structure, amino acid content, and orientation.²⁰ Previous work done by the Stefanovic Lab identified that LARP6 interacts with the 5'- untranslated region (UTR) stem loop of collagen type I mRNA (Figure 5).²⁶

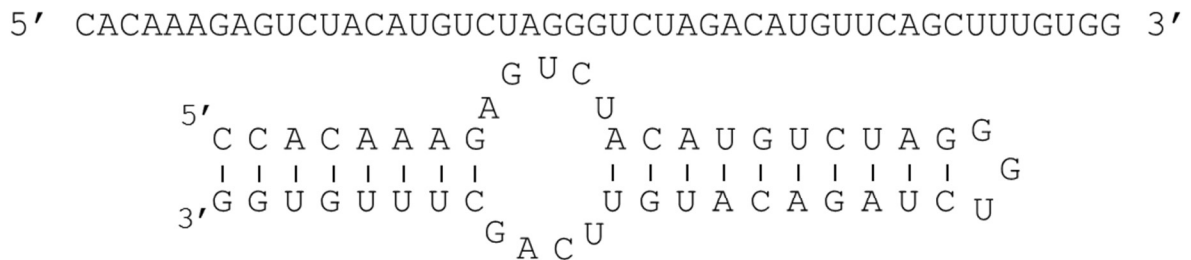


Figure 5. Human collagen type I mRNA 5' untranslated region (UTR). (Top) The mRNA that encodes *HsCOL1A1* contains a sequence in the 5' UTR to which *HsLARP6* has been found to bind. (Bottom) This sequence is predicted to form a stem loop secondary structure.

Studies in fish and mammals report that LARP6 interacts and works with transcriptional factor MyoD, a key protein for muscle differentiation, thereby controlling muscle development.²⁷ A study of a LARP6 knockdown in zebrafish reported that embryos were unable to develop properly, suggesting that LARP6 may be a key factor controlling myogenesis (Figure 6).²⁸ When the LARP6 gene was knocked down by antisense oligonucleotide morpholinos, the embryo did not develop proper somites due to smaller and more disorganized muscle fibers. In contrast, when a post-fertilized embryo was injected with LARP6 mRNA to stimulate overproduction of LARP6, longer muscle fibers and bigger somites were produced. Another study performed in mice disrupted the 5'-UTR stem loop in the collagen type I mRNA, and the loss of this LARP6 binding site caused a

collagen deficiency upon injury.²⁹ These studies demonstrate that LARP6 is a critical posttranscriptional factor needed for collagen synthesis and proper muscle development.

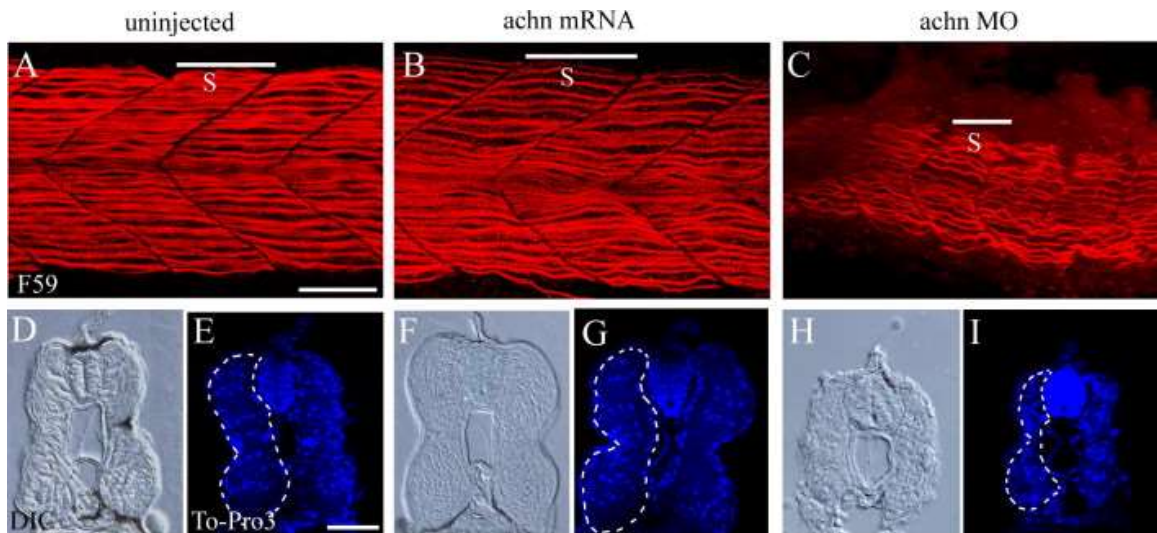


Figure 6. LARP6 morpholino knockdown effect in zebrafish embryogenesis. Panels A to C represent lateral views of zebrafish embryos 24 hours post-fertilization labeled with the F59 antibody (reveal slow twitch muscle fibers). Somites represented by (S). A) Control embryo muscle fiber. B) LARP6 mRNA injected embryos. C) Embryo injected with LARP6 morpholino. Panels D, F and H are cross-section images of a trunk of 48-hour post-fertilization (D, control; F LARP6 mRNA injected embryo; H, LARP6 morpholino injected embryo). Panels E, G, and I are To-Pro3 stained confocal images of the 48-hour embryo cross-sections. E) Control To-Pro3 with left somite outlined. G) LARP6 mRNA injected somite. I) LARP6 morpholino injected embryo. Picture reproduced with permission from Wang *et al* (2009).

Evolutionary Divergence of LARP6 La module

A molecular structure has not been solved for any of the full-length LARPs. However, solution NMR structures have been determined for the genuine La and LARP6 La motifs (Figure 7) and the RRM (Figure 8).²⁰ Generally, the structures are very similar. The La motifs both have a winged-helix topology. There are some minor differences including the orientation and length of some helices. For example, helix $\alpha 1$ is generally in the same position on both La motifs, where it packs against helix $\alpha 5$ and the beta strands. In *HsLARP6*, $\alpha 1$ is overall shorter than the $\alpha 1$ in genuine La protein.

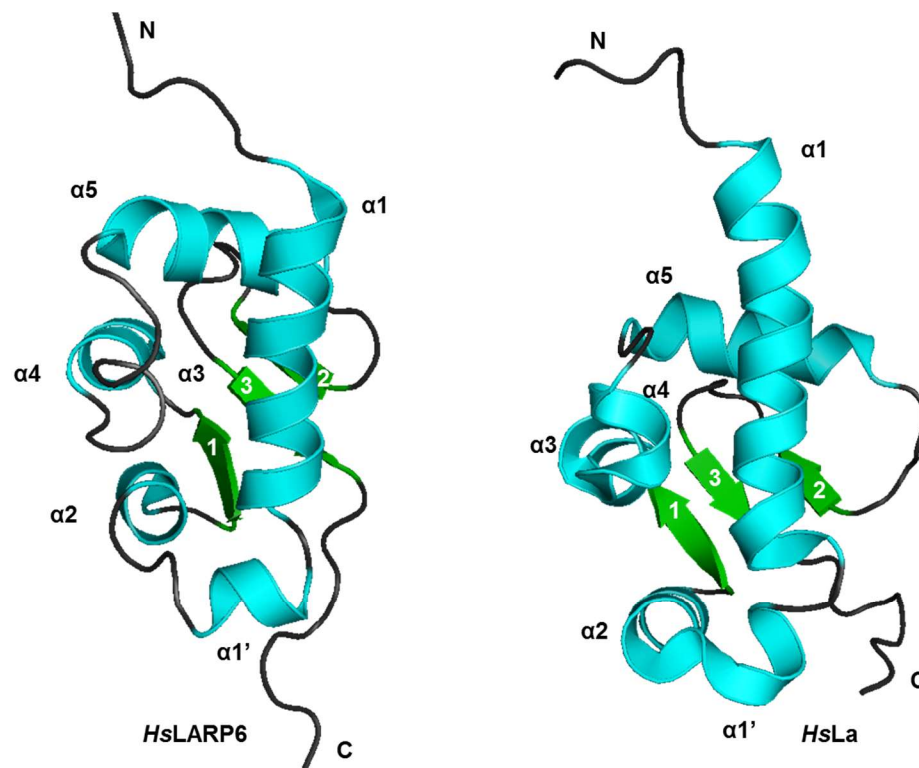


Figure 7. Structure of Human LARP6 La motif. *HsLARP6* La motif NMR structure (left, PDB: 2MTF) compared to the *HsLa* La motif (right, PDB: 1S7A). The overall structure has similar orientation and components. However, *HsLa* La motif seems to have a more determined secondary structure than human LARP6 La motif (contains longer loops between helices). Structures generated with PyMOL.

Additionally, *HsLARP6* helix $\alpha 1'$ is more defined compared to $\alpha 1'$ in *HsLa*, which seems to be more closely connected to helix $\alpha 2$; a similar phenomenon is present in the $\alpha 4$ helices. Another minor difference can also be seen in $\beta 3$, which in *HsLARP6* looks longer compared to the same structure in *HsLa*. Overall, the La motif does not show major differences in structure other than the secondary structures being somewhat longer.

In contrast, the RRM in *HsLARP6* shows major differences in secondary structure elements and orientation compared to the structure of the genuine La RRM. First, the *HsLa*-RRM contains a one-turn helix, $\alpha 0$, coming from the N-terminus into $\beta 1$. This structure is not present in the *HsLARP6*-RRM. Second, there is a longer loop between $\beta 1$ and $\alpha 1$ in *HsLARP6* that contains a small one-turn helix, designated $\alpha 0'$.²¹ An interesting feature appears in *HsLARP6* between $\beta 2$ and $\beta 3$. A longer loop between $\beta 2$ and $\beta 3$ contains an α helix ($\alpha 1'$), which is not present *HsLa*. This helix conceals two β strands of the antiparallel β -sheet core domain of the RRM. While its general location is somewhat similar to helix $\alpha 3$ in *HsLa*, the LARP6 $\alpha 1'$ helix appears to fully block the canonical RNP that is known to interact with RNA. Helix $\alpha 2$ appears to be longer in human LARP6 as compared to the genuine La protein. The other major difference is in the C-terminal $\alpha 3$ helix that is present in *HsLa*. As previously mentioned, $\alpha 1'$ might have a similar function based on its location in the RRM. Overall, the RRMs of the La-related proteins exhibit more structural differences than the La motifs.

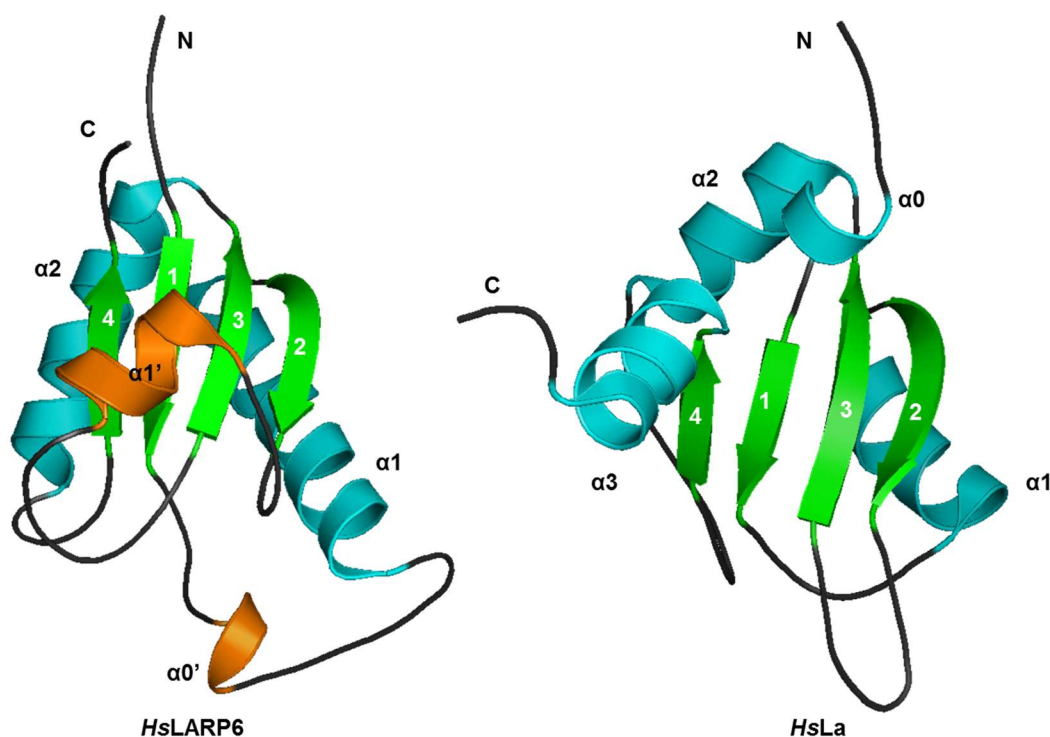


Figure 8. Structure of Human LARP6 RRM. *HsLARP6* RRM NMR structure (left, PDB: 2MTG) compared to the *HsLa* RRM (right, PDB: 1S79). *HsLARP6* RRM is missing the $\alpha 0$ from *HsLa*. In contrast, it contains an $\alpha 0'$ between $\beta 1$ and $\alpha 1$. In addition, $\alpha 1'$ (between $\beta 2$ and $\beta 3$) may be analogous to $\alpha 3$ in *HsLa* RRM. However, the orientation and sequence in which $\alpha 1'$ takes place is different from genuine La. Structures generated with PyMOL.

In addition to differences between the LARPs visible in the multiple sequence alignment, a recent study exchanged human LARP6 sequence sections with the equivalent domains from human La (LARP3) to determine if the sequence exchange is deleterious.²⁰ A chimera in which the linker sequence between the La motif and RRM of *HsLa* (TDEYKNDVKNR, residues 101-111) was inserted between residues 178-180 in *HsLARP6*. Another construct was generated in which the loop between strand $\beta 1$ and helix $\alpha 1$ of *HsLARP6* RRM (loop1, residues 190-208) was replaced with equivalent loop from *HsLa* (GFPTDATL, residues 117-124).

The RNA binding activity of the two mutant proteins was measured by isothermal titration calorimetry, using the 48 nt stem loop 5'-UTR of collagen type I mRNA for human. The interlinker mutant bound to the ligand with a $K_{d,app}$ of 667 nM, compared to the *HsLARP6* La Module binding affinity of 48 nM. These data demonstrate that the interlinker length as well as the chemical characteristics are important for full binding activity. In contrast, the loop mutant exhibited a $K_{d, app}$ of 18 nM. The same study also exchanged *HsLARP6* RRM with the *HsLa* RRM. This exchange of the entire RRM domain displayed no binding activity at all.²⁰ Together, these experiments demonstrate the inability to swap sequences across subfamilies, which indicates that genuine La and LARP6 have evolved different mechanisms for binding to RNA substrates. However, it is still unknown whether these mechanisms are conserved within each LARP subfamily. If they are, then protein sequences or domains of a LARP from two different species should be able to be exchanged with minimal effect on structure and biochemical activity.

Evolutionary divergence within LARP6 proteins

To date, the LARP6 proteins from human and plants have been studied in detail. Often, animal models are used to study both the biochemistry and physiology of proteins. By studying the structure and function of LARP6 from fish, which are non-mammalian vertebrates, the specificity of the interaction between LARP6 and target mRNAs could be screened for a possible application in fibrotic diseases. Zebrafish (*Danio rerio*) is a widely used genetic model system, but due to an evolutionary genome duplication event, it expresses two LARP6 homologs that might complicate the evaluation of LARP6 cellular function *in vivo*. Another teleost genetic model organism, *Xiphophorus maculatus*, expresses only one LARP6 protein, which could be more useful for identifying natural regulatory targets.

With regard to size, human LARP6 is the longest of the four proteins (~57 kDa) compared to the fish LARP6 (~52 to ~53 kDa). Figure 9 shows the sequence alignment between human LARP6, *D. rerio* LARP6a and LARP6b and *X. maculatus* LARP6. The known secondary structure of LARP6 Module is represented by rectangles (α helices) and arrows (β sheets) as well as the start of the La motif, the RRM and the LSA motif.

Between the four sequences, there is a high degree of sequence conservation within the La motif α helices. The conservation may be required for proper function and orientation within the La motif and intramolecular interactions

with the RRM and the LSA. In contrast, the RRM contain a higher sequence variability.

Despite the structural differences from the canonical RRM, there are several chemical characteristics of LARP6 RRM that are similar to the mechanisms used by other RRM to bind their ssRNA ligands. Most noticeable is the presence of positively charged residues (lysine and arginine) that can interact with the negatively charged phosphate backbone of the RNA. Additionally, LARP6 RRM contains several hydrophobic groups, such as alanine, methionine, and valine that can interact with the bases in the RNA.

(following page) **Figure 9. Sequence alignment of *H. sapiens*, *X. maculatus*, and *D. rerio* LARP6 proteins.** The known secondary structures of HsLARP6 La Motif and RRM are aligned to the sequences. Underlined sequence represent the interlinker. Residues in the La motif that when mutated to alanine completely inhibit RNA binding are highlighted in green. When the aspartate that is highlighted in pink is mutated to alanine, RNA binding affinity merely reduced. The residues in the RRM that are highlighted in yellow increase binding affinity when mutated to alanine. All conserved prolines are highlighted in orange. Their conservation strongly suggests that they serve as structural elements for the vertebrate LARP6 proteins.



Previous studies mutated a set of highly conserved residues.²⁰ In the La motif, single mutation to alanine at position 99 (Q99A), 103 (Y103A), 114 (F114A) or 135 (F135A) each completely prevented *HsLARP6* from binding the 48-nt stem loop of the 5'-UTR of collagen type I mRNA (Figure 9, green highlight). Similarly, mutation of aspartic acid at position 112 to alanine (D112A; Figure 9, pink highlight) did not prevent binding but did significantly weaken the interaction between the *HsLARP6* and the RNA ligand. The negative charge of this amino acid might contribute repulsive forces against the phosphate backbone or the orientation of the La motif. Because these residues are also conserved in fish LARP6 proteins, they could also be critical for the fish proteins to bind to their natural ligand.

In contrast, mutation of either lysine 187 or glutamic acid 262 to alanine in the RRM (Figure 9, yellow highlight) was found to increase the binding affinity to the 48-nt stem loop. This may be caused by simply removing potentially repulsive charges and replacing them with a nonpolar alanine. Alternatively, exchanging the wildtype charged residues for alanine could change the surface shape for RNA binding or other interactions that could lower the dissociation constant. Even so, it is important to note that tighter binding does not directly correlate to efficiency as a regulator of gene expression at the level of translation. The project undertaken in this thesis takes advantage of the evolutionary diversity between the human and fish proteins to probe the role of these and other residues in the RNA binding mechanism of LARP6.

This study had several goals. First, we sought to clone *X. maculatus* and *D. rerio* LARP6 genes into pET28 plasmids. Second, we wanted to recombinantly express and purify human, *X. maculatus* and *D. rerio* LARP6 proteins. These experiments required the optimization of buffer conditions to ensure stability and solubility of the fish proteins. Furthermore, the recombinant proteins needed to be biochemically characterized by measuring binding affinity to the human LARP6 collagen type I mRNA ligand. Finally, domain topology studies were performed by limited proteolysis studies. These advances lay the foundation for future studies to test if LARP6 binding mechanisms are conserved within subfamily members.

II. MATERIALS AND METHODS







Calibration of size exclusion chromatography columns

To ensure proper isolation of all of the proteins used in this work, the calibration of both size exclusion columns was performed. The Sephadex S200 size exclusion column was previously calibrated.³⁰ The Sephadex S75 size exclusion column was calibrated for this work, by eluting a set of standard proteins over the column. The elution volume of each protein was correlated with the known globular molecular weight. Both calibration processes demonstrated that the columns were efficiently packed and useful for protein purification. The sets of standards used for the S200 and S75 columns are listed in Table 1 and Table 2, respectively. The color markers for each standard protein will be presented in every chromatogram from here on.

Table 1. Size Exclusion Chromatography S200 Standards.

	Standard	MW (kDa)	Elution Volume (mL)
●	Blue Dextran	2,000	45.66
●	Ferritin	440	59.9
●	Aldolase	158	73.3
●	Conalbumin	75	80.3
●	Ovalbumin	44	86.1
●	Carbonic Anhydrase	29	92.4
●	Aprotinin	6.5	107.2

Table 2. Size Exclusion Chromatography S75 Standards.

	Standard	MW (kDa)	Elution Volume (mL)
	Blue Dextran	2,000	45.03
	Conalbumin	75	53.29
	Ovalbumin	44	58.75
	Carbonic Anhydrase	29	67.16
	RNAse A	13.7	79.72
	Aprotinin	6.5	92.87

For each protein, K_{av} (the “available value” of the distribution coefficient K_D), was calculated as follows:

$$K_{av} = (V_e - V_o) / (V_c - V_o)$$

where V_e is the protein in question elution volume, V_o is the column void volume (determine by Blue Dextran), and V_c is the column total volume (120 mL).

From the K_{av} values and elution volume data in Table 1, linear regression analysis was performed to generate an equation to yield the apparent experimental molecular weight:

$$y = -0.3454X + 2.1528$$

That process was repeated for the data in Table 2:

$$y = -0.5105X + 2.5757$$

where x is the \log_{10} of the molecular weight of the standard (Da) and y is K_{av} .

Cloning and expression constructs

The human LARP6 (*HsLARP6*) expression plasmid was a kind gift from Dr. Mark Bayfield (York University, Canada).²² *Xiphophorus maculatus* and *Danio rerio* cDNA were a kind gift from Dr. Ron Walter (Texas State University). The *DrLARP6a* and *DrLARP6b* coding sequences were first commercially synthesized as inserts in pcDNA3.1 (GenScript). The expression plasmids for the SUMO fusion system (pET28-SUMO and pET28-Ulp1) were kind gifts of Dr. Christopher Lima (Rockefeller University) and Cornell University.

Amplification of *XmLARP6* was performed using *X. maculatus* cDNA as template with primers containing the *Bam*HI and *Xho*I sites (Table 3) and inserted into *Bam*HI/*Xho*I-digested pET28-SUMO.³¹ The commercial pcDNA3.1 constructs described above were used for amplification of the *Danio rerio* LARP6 coding sequences using primers containing *Bam*HI and *Xho*I sites (Table 3). The amplified product was then inserted into *Bam*HI/*Xho*I-digested pET28-SUMO (Table 4). All cloned expression constructs were screened by PCR and double restriction digest and then confirmed by sequencing. Upon sequencing the pET28-SUMO-*DrLARP6a* construct, a 2-nucleotide insertion was found that caused a deleterious frameshift. Site directed mutagenesis was subsequently performed to delete the extraneous two nucleotides and restore the coding sequence (Table 3).

Table 3. Primers used in this thesis work.

Protein	Sequence	Description
DrLARP6a	AAAAGGATCCTTATGAGCAGCGAGCAGCCGCCGCGG	PCR fwd.: (<i>Bam</i> HI)
	CCGGCTCGAGTCATTATAGACAGGTCTGTGTGGACGCAGTTTTGCCGAGCGG	PCR rev.: (<i>Xho</i> I)
DrLARP6a SDM	GAACAGATTGGTGGATCCATGAGCAGCGAGCAG	SDM fwd.
	CTGCTCGCTGCTCATGGATCCACCAATCTGTTC	SDM rev.
DrLARP6b	CATGGATCCATGTCCTCACCCGGTTTTACTGACATCTTCCATTG	PCR fwd.: (<i>Bam</i> HI)
	GAATCTCGAGTCATTGGGGTAGCAGCACTTTCCTCGAC	PCR rev.: (<i>Xho</i> I)
XmLARP6	CATGGATCCATGGGGACGGTGACGATCACCGTGGCGATTGAGGC	PCR fwd.: (<i>Bam</i> HI)
	CGTACTCGAGTCATTACTGTGTGGCAGCAGTCTTCCCCTCTCGGCAACTG	PCR rev.: (<i>Xho</i> I)
XmLARP6- LAM	CATGGATCCGGCACCAGTGGAGGGGAGTTGGAGG	PCR fwd.: (<i>Bam</i> HI)
	GCAACTCGAGTCATTACCCCCCTCCTCGCGAGGTCTG	PCR rev.: (<i>Xho</i> I)

Table 4. Plasmids used in this thesis work.

Protein	Plasmid	Antibiotics
DrLARP6a	pcDNA3.1 pET28-SUMO	Ampicillin Kanamycin
DrLARP6b	pcDNA3.1 pET28-SUMO	Ampicillin Kanamycin
HsLARP6	pET28	Kanamycin
Ulp1	pET28-SUMO	Kanamycin
XmLARP6	pET28-SUMO	Kanamycin
XmLARP6-LAM	pET28-SUMO	Kanamycin

Cell transformations

For cloning, *Escherichia coli* DH5 α competent cells were used. Cells and DNA were incubated for 30 min on ice, heat shocked for 90 s at 37 °C, incubated for 2 min on ice and recovered with 900 μ L of autoclave-sterile Luria Broth (LB) at 37 °C for 1 h with shaking. Recovered cells were plated onto a LB-agar plate containing the appropriate antibiotic and incubated at 37 °C overnight. Plasmid DNA was collected using QIAprep Spin Miniprep Kit from 3 mL of suspension culture grown from a single colony from the plate. The manufacturer's protocol was followed, and DNA eluted in 40 μ L of nuclease-free H₂O (Integrated DNA Technologies).

For protein expression, *E. coli* BL21 (DE3) pLysS (a kind gift of Dr. Robert McLean, Texas State University department of Biology) and Rosetta™ (DE3) pLysS competent cells (EMD Millipore) were transformed as follows: cells and DNA were incubated on ice for 30 min. Samples were then heat shocked for 45 s at 42 °C, followed by incubation on ice for 2 min. 900 μ L of autoclave-sterile Luria Broth (LB) was added, and samples were recovered at 37 °C for 1 h with shaking. After recovery, samples were plated onto LB agar plates containing appropriate antibiotic for each plasmid/cell line combination (Table 5) and incubated overnight at 37 °C.

Table 5. Cell lines and respective antibiotics

Cell line	Antibiotic
DH5α	N/A
<i>Escherichia coli</i> BL21 (DE3) pLysS	Chloramphenicol
<i>Escherichia coli</i> Rosetta™ (DE3) pLysS	Chloramphenicol

Sequences of expressed proteins

>His₆-HsLARP6

MGSSHHHHHHSSGLVPRGSHMAQSGGEARPGPKTAVQIRVAIQEAEDVDELEDEEEGAETRGAGDPARYLSPGWGSASEEEPSRGHSGT
TASGGENEREDLEQEWKPPDEELIKKLVDQIEFYFSDENLEKDAFLLKHVRRNKLGYVSVKLLTSFKKVKHLTRDWRTTAHALKYSVVL
ELNEDHRKVRRTTPVPLFPNENLPSKMLLVYDLYLSPKLWALATPQKNGRVQEKVMEHLLKLFGTFTGVISSVRILKPGRELPRDIRRIS
SRYSQVGTQECAIVEFEEVEAAIKAHEFMITESQGKENMKAVLIGMKPPKKKPAKDKNHDEEPTASIHNLKSLNKRVEELQYMGDESSA
NSSSDPESNPTSPMAGRRAATNKLSPSGHQNLFLSPNASPCTSPWSSPLAQKGVSRKSPLAEEGRLNCSTSPEIFRKCMDYSSDSSV
TPSGSPWVRRRRQAEMGTQEKSPGTSPLLSRKMQTADGLPVGVLRLPRGPDNTRGFHGHHERSRACV

> His₆-SUMO-DrLARP6a

MGHHHHHHHSSGHIEGRHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFlyDG
IRIQADQTPEDLDMEDNDIIEAHREQIGGSMSSSEQPPREISAPVTITVAIQAAEEDDEPDEEPSCNTIELQTGSGSEDELGRHDKSSGA
GTSGGELEEEESWQPPDPELIQKLVAQIEYYLSDENLEHDAFLLKHVRRNKLGFVSVKLLTSFKKVKHLTRDWRTTAYALRHSNLELND
EGRKVRRRSTVPVFASESLPSRMLLLSELKRWPELGIALGGDSNNGSGPTQQERLMELLLKAFGNYGPIASVRVLKPGKDLPADLKKLS
GRYSQLGTEECAIVEFEEVEAAMKAHEAVGGEGGNRGLGLKVVLIGTKPPKKKVPKDRPRDEGIGGMRKSRSLNSRVRELQYHGDDSA
ASSSETESNPTSPRLARKSRSCNKLSPTSAGPNHLSPPVSPRSSPWSSPRASPCTQRKTHPSGKSPLASEGRLSPEPGRRWADYSSDSS
LTPSGSPWVQRRKQVASQESSPVGSPMLARKIQNADGLPPGVVRLPRGPDGTRGFHCPPLGKTASTQTCL

> His₆-SUMO-DrLARP6b

MGHHHHHHHSSGHIEGRHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFlyDG
IRIQADQTPEDLDMEDNDIIEAHREQIGGSMSSPGFTDIFHCEDTQQSKGLFSYTDGLNDSLDGSSTDLDVFEEDFCEGELWI PPND
DVTQQIATQLENYLSDENLSDDAFLLKHVQRNKMGYVSLKLLTSFKKIRDLDTRDWRTTLAAARTSPQLEVNMGTKVRRRTTPVPDWLLC
IPTSKLLLLAWNFLDGAGPVKEKFESPGVEQLGIMEAAMRVFSPYGTISSLRILRPGKEIPAELEKRYTKKHLELGRKVCAYVEYEYLEGA
RKAFEALKVEEQQGGRGICVLLGSRGTRKPGCSQGLVDEEQEDCIDIDVLKRPDRKARQFVYSLEDSAVCSSSESDFAPASPRPNRRV
SRPQALYGSPLAIPRVSTFRSDPYRNPLGSPVGSPLQPRKLFPCSHVTSPLATHPFSSTPSGAGTSFKYKASGELSPDDLGTSSPWVQ
RRKNAAQVIQPEMASSMSPSQLIKSLSVLVVRQPVGPDGTKGFHNCIGRGKVLLPQ

> His₆-SUMO-XmLARP6

MGHHHHHHSSGHIEGRHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFKRQ GKEMDSLRFlyDG
IRIQADQTPEDLDMEDNDIIEAHREQIGGSMGTVTITVAIQAAEDEEPEEEHHPGNVECLRGSCSEDELGRHDKSRHSGAGTSGGELEEE
SWQPPDTELIQKLVTQIEFYLSDENLEHDAFLLKHVRRNKLGFVSVKLLTSFKKVKHLTRDWRTTAYALKH SKILELNDEGRKVRRKSA
VPVFASESLPSRM LLLSDLQKWPELAALTKDNGSNEGGATQQEQLMKLLLKA FGTYGAIASVRVLKPGKDL PADLKRLSGRYAQLGNEE
CAIVEFEEVEAAVKANEAVGGEDGGTGS LGLKVVLIGTKPPKKKVLKERPREEGGMRKSRSLNSRVRELQYHGDD SACSSSETESTPTS
PRLARKSQSCNKLSPTTAGISFQNNHLSPGISPRNSPWSSPRASPCPQRKAPHSHKSPLLGDGRLSPEAGR RWADYSSDSSLTPSGSPW
VQRRKQVASQESSPVGSPMLGRKIQNADGLPPGVMRLPRGPDGTRGFHGVTV AERGKTAATQ

> His₆-XmLARP6-LAM

28

AGTSGGELEEEESWQPPDTELIQKLVTQIEFYLSDENLEHDAFLLKHVRRNKLGFVSVKLLTSFKKVKHLTRDWRTTAYALKH SKILELN
DEGRKVRRKSAVPVFASESLPSRM LLLSDLQKWPELAALTKDNGSNEGGATQQEQLMKLLLKA FGTYGAIASVRVLKPGKDL PADLKRL
SGRYAQLGNEECAIVEFEEVEAAVKANEAVGGEDGGTGS LGLKVVLIGTKPPKKKVLKERPREE

Protein expression time trials

LB plates containing transformed BL21 (DE3) pLysS or Rosetta™ (DE3) pLysS were stored at 4 °C for up to 6 weeks. An overnight culture was made by inoculating a single colony into LB containing the antibiotics and growing at 37 °C for 12—16 h with shaking until stationary phase was reached. 1 mL aliquots of the overnight cultures were transferred to autoclaved-sterile 250 mL LB with appropriate antibiotics. Cultures were allowed to grow to an OD₆₀₀ between 0.5-0.7 (mid-log phase) at 37 °C with shaking at 225—250 rpm. When desired density was reached, the cultures were placed on ice for 5—10 mins with occasional shaking. 1 mL of culture was removed and cells pelleted at 16,000 xg for 1 min for a basal expression sample. Protein expression was induced with 0.1 M IPTG and placed on a shaker at 18 °C for 16 h. During the expression period, 1 mL sample of culture was removed after 2 h, 4 h, 6 h, 8 h, and overnight and the cells pelleted as above before storage at -20 °C. Each cell pellet was resuspended in 500 µL 1X SDS sample buffer (5X buffer: 0.5 M Tris [pH 7.5 at 4 °C], 4 mM β-ME, 0.4 M SDS). Once resuspended, samples were incubated at 90 °C for 5 min before immediate loading onto duplicate 10% (37.5:1 acrylamide: bis-acrylamide) polyacrylamide, 0.75 mm thick minigels. The gels were electrophoresed in 1X Tris-Glycine running buffer (50 mM Tris and 0.5 M glycine, 0.4 M SDS) at 200 V until the dye front reached the bottom of the gel (~ 1 h). One gel was analyzed by Coomassie blue staining, and the other was analyzed by anti-His Western blot.

Coomassie blue staining

Total protein content was detected using Coomassie Brilliant Blue staining solution [0.05% (w/v) Coomassie Brilliant blue, 40% (v/v) methanol, 10% (v/v) glacial acetic acid, and 50% (v/v) with Milli-Q polished deionized H₂O] for 15 min at room temperature with shaking. Gels were destained in destaining solution (40% (v/v) methanol, 10% (v/v) glacial acetic acid, and 50% (v/v) MQ H₂O) for 20–30 min with Kimwipes added to soak up released Coomassie molecules. Destained gels were imaged using ChemiDoc XRS+ molecular imager following Coomassie stain settings.

Silver staining

Sensitive detection of total protein content was performed using silver staining. In a glass container, gels were fixed in 50% ethanol for 15-30 min at room temperature with shaking. After fixing, gels were incubated in staining solution (7.56% NaOH, 1.5% (v/v) NH₄OH, 4.7 M AgNO₃) for 15 - 30 mins with shaking. Gels were washed 3 times with 200 mL MilliQ-polished H₂O. Then, developing solution (2.5% citric acid, 37% formaldehyde) was added and mixed by hand until the desired intensity was reached. Immediately, 100-200 mL of kill solution was added (45% (v/v) methanol and 2% (v/v) acetic acid). Gels were imaged using ChemiDoc XRS+ molecular imager using the silver stain setting.

Anti-His Western blot

Anti-His Western blots were used to determine protein expression efficiency. The Bio-Rad Trans-Blot® Turbo™ Transfer System and kit was used to transfer proteins from the SDS gel to a nitrocellulose membrane using the

Mixed Molecular Weights program (1.3 A, 25 V, 7 min). Once transferred, the membrane was incubated for 1 h in blocking solution (10% bovine serum albumin (BSA) in 1x Tris-buffered saline containing 0.05% Tween-20 (TBS-T)). After blocking, the membrane was incubated for 1 h with a 1:5000 dilution of anti-His probe (ThermoFisher) in TBS-T. Membrane was washed twice with 1X TBS-T for 10 min each followed by a final 10 min wash with TBS (no Tween-20). The membrane was then incubated with 20 mL of homemade enhanced chemiluminescence reagent (*a.k.a.* “The Juice”: 0.1 M Tris [pH 8.8 at 25 °C], 1.25 mM luminol in DMSO, 2 mM 4-IPBA in DMSO) and 12 μ L of 30% H₂O₂.³² The nitrocellulose membrane was then imaged using the ChemiDoc XRS+ molecular imager under the instructions of the chemiluminescence protocol.

Anti-LARP6 Western Blot

Anti-LARP6 Western blots were used to probe for full-length LARP6, as well as protein stability. The same procedure for anti-His was followed. Briefly, SDS gel proteins were transferred to a nitrocellulose membrane using the Mixed Molecular Weights program (1.3 A, 25 V, 7 min). Once transferred, the membrane was incubated for 1 h in blocking solution. Subsequently, the membrane was incubated for 1 h with rabbit anti-*Hs*LARP6 (C-terminus) antibody (1:1000, v/v in blocking buffer) (Abcam). Membrane was washed three times with 1X TBS-T for 10 min each. Membrane was then incubated with goat anti-rabbit (1:20,000 v/v in blocking buffer) for 1 hour. Membrane was washed twice with 1X TBS-T for 10 min each followed by a final 10 min wash with TBS (no Tween-20).

Membrane was then incubated with chemiluminescence reagent and imaged as previously described.

Protein expression and purification

Rosetta™ cells were used for all large-scale purifications. 1 L of LB with appropriate antibiotics was inoculated with overnight cultures of transformed Rosetta™ cells with a target OD₆₀₀ of 0.5—0.7. Cells were induced with 0.1 M IPTG and incubated overnight with shaking at 18 °C. Cells were then collected by centrifugation at 5,000 *xg*, 4 °C for 10 min. Cells were resuspended using Lysis/Wash #1 buffer (Table 6) and a dissolved protease inhibitor tablet (ThermoFisher). Cells were lysed open by 8 sonication cycles of 20 s each with rest intervals of 30 s at 50% amplitude in an ice water bath. The supernatant was collected by centrifugation at 18,000 *xg*, 4 °C for 15 min.

Nickel affinity chromatography

Ni²⁺-agarose beads (4 – 5 mL of a 50% slurry per 1 L original culture volume) were washed with MilliQ-polished H₂O and equilibrated with Lysis/Wash#1 buffer containing dissolved protease inhibitor tablet. The equilibrated beads were added to the cleared cell lysate and incubated for 1 h at 4 °C with shaking. The beads/lysate mixture was transferred to a glass column and allowed to settle. Once the resin bed was uniformly settled, the flow-through was collected. The beads were then treated with Lysis/Wash #1 buffer without protease inhibitors and was collected. The resin was then treated with wash #2 Buffer and two fractions were collected. After the second wash, Elution Buffer was added and 5 fractions were collected. The pellet and fractions collected were

analyzed via SDS-PAGE and Coomassie blue staining as described above.

Size exclusion chromatography

The size exclusion column (Sephadex 75 or Sephadex 200, as appropriate) was equilibrated with storage buffer which had been filtered through a 0.2 μm nitrocellulose membrane to remove particles. Elution fractions from the Ni^{2+} affinity column were pooled based on SDS-PAGE band intensity and concentrated to 2.5—3 mL using a pre-rinsed Vivaspin™ centrifugal concentrator (MWCO 10,000, Sartorius) by centrifugation at 4,000 $\times g$, 4 °C for 10—15 min cycles. The concentrated protein was then filtered with a 0.2 μm syringe filter before loading into the ÄKTA Pure Fast Protein Liquid Chromatography apparatus. The column was set to elute a total volume of 180 mL at a flow rate of 1 mL/min. Once fractions were collected, the chromatogram of A_{280} that was monitored by the UNICORN software was used to determine which fractions to analyze by SDS-PAGE and Coomassie staining as described above. Fractions were pooled based on band intensity and purity. The pooled fractions were either subjected to Ulp1 cleavage (see below) followed by a second size exclusion column as described above, or prepared for long-term storage by being distributed into small aliquots that were first snap-frozen in liquid N_2 and then stored at -70 °C.

Table 6. List of purification buffers.

Proteins	Lysis/Wash buffer #1	Wash buffer #2	Elution buffer	SEC buffer
<i>DrLARP6a</i> <i>DrLARP6b</i> <i>XmLARP6</i> <i>XmLARP6-LAM</i>	50 mM NaH ₂ PO ₄ / Na ₂ HPO ₄ [pH 8.0] 200 mM NaCl 0.5 M glucose, 300 mM NaI, 10 mM imidazole [pH 8.0]	50 mM NaH ₂ PO ₄ / Na ₂ HPO ₄ [pH 8.0] 200 mM NaCl 0.5 M glucose 300 mM NaI 30 mM imidazole [pH 8.0]	50 mM NaH ₂ PO ₄ / Na ₂ HPO ₄ [pH 8.0] 200 mM NaCl 0.5 M Glucose 300 mM NaI 30 mM imidazole [pH 8.0]	50 mM NaH ₂ PO ₄ / Na ₂ HPO ₄ [pH 8.0] 200 mM NaCl 0.5 M Glucose 300 mM NaI
<i>HsLARP6</i>	50 mM NaH ₂ PO ₄ / Na ₂ HPO ₄ [pH 7.5] 300 mM NaCl 10 mM imidazole [pH 8.0]	50 mM NaH ₂ PO ₄ / Na ₂ HPO ₄ [pH 7.5] 300 mM NaCl 30 mM imidazole [pH 8.0]	50 mM NaH ₂ PO ₄ / Na ₂ HPO ₄ [pH 7.5] 300 mM NaCl 300 mM imidazole [pH 8.0]	50 mM Tris-HCl [pH 7.5] 100 mM NaCl 1 mM DTT 5% glycerol
<i>Ulp1</i>	50 mM Tris-HCl [pH 8.0 at 4 °C] 350 mM NaCl 3 mM βME 10 mM imidazole [pH 8.0] 10% glycerol	50 mM Tris-HCl [pH 8.0 at 4 °C] 350 mM NaCl 3 mM βME 30 mM imidazole [pH 8.0] 10% glycerol	50 mM Tris-HCl [pH 8.0 at 4 °C] 350 mM NaCl 3 mM βME 250 mM imidazole [pH 8.0] 10% glycerol	50 mM Tris-HCl [pH 8.0 at 4 °C] 200 mM NaCl 1 mM βME 10% glycerol

Solubility screening assay

In order to increase stability and solubility of the recombinant fish proteins, an iterative screen of different chaotropes, osmolytes, and kosmotropes was done following the protocol described by Dr. Kelly Churion and Dr. Sarah Bondos (Table 7).³³

Figure 10 shows a general representation of the solubility assay. All stock solutions of additives and Standard Buffer (50 mM NaH₂PO₄/ Na₂HPO₄ [pH 8.0], 200 mM NaCl) were prepared at a 10X concentration. A cell pellet of expressed protein from a 250 mL expression culture was resuspended with Standard Buffer containing a dissolved protease inhibitor tablet (Fisher Scientific). Sonication and lysate collection were performed as previously described. For testing the first set of additives as well as a negative control, the following reactions were carried out in 0.5 mL autoclave-sterile microcentrifuge tubes on ice:

For the negative control, 100 µL 1X solution was prepared:

10 µL 10x standard buffer
90 µL MilliQ water
<hr/>
100 µL total

For each 10X additive, 100 µL 1X solutions were prepared:

10 µL 10x standard buffer
10 µL 10X additive
80 µL MilliQ water
<hr/>
100 total

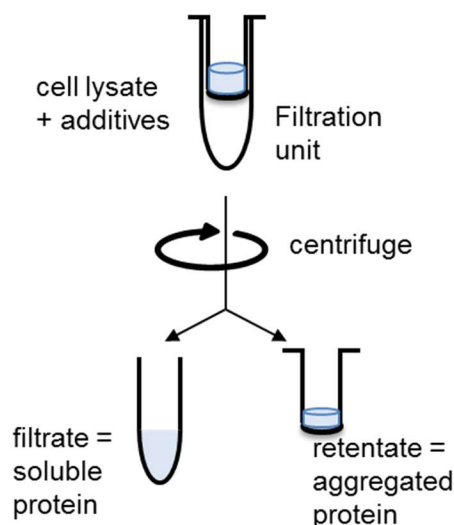


Figure 10. Solubility screening overview.

Cell lysate and additives being tested were left to reach equilibrium in a filtration unit. Then, subjected to centrifugation at 15,000 xg , 4 °C for 20 min. Then the filtrate was prepared for SDS-PAGE analysis as previously described. The retentate and membrane were washed with 30 μL of nanopure water to resuspend any aggregated retained protein. Resuspended samples were also combined with SDS sample buffer. Samples were analyzed by gel electrophoresis and coomassie stain.

To each tube, including the negative control, 10 μL of cleared lysate was added. The reactions were incubated for 2 hours on ice before transferring them into Durapore PVDF 0.1 μm centrifugal units (Merck Millipore). Reactions were then subjected to centrifugation at 15,000 xg , 4 °C for 20 min. In preparation for SDS-PAGE analysis, 20 μL of filtrate was combined with 5 μL 5X SDS sample buffer. The filter membrane was washed with 30 μL of nanopure water to resuspend any aggregated retained protein. Resuspended samples were also combined with SDS sample buffer. Gel electrophoresis was performed using a 10% polyacrylamide gel, and then transferred to nitrocellulose using the

TransBlot Turbo Mixed Molecular Weight setting. The N-terminal His₆ tag on the recombinant proteins was detected with an anti-His Western blot. Based on the outcome of each iteration, additional additives, concentrations, and combinations were then performed following the same procedure.

Table 7. Selection of additives for filter assay. Adapted from K.A Churion and S. E. Bondos.³³ Reproduced with copyright permission from Elsevier.

Additives for Assay Iteration 1

1. 300 mM NaCl (500mM NaCl total concentration)
2. 10% Glycerol
3. 0.5 M Urea
4. 500 mM Arginine
5. 0.5 M TMAO
6. 1% Nonident p40

Additives for Assay Iterations 2 and 3: optimization and concentration

Additive	Initial concentration	Recommended concentration range
----------	-----------------------	---------------------------------

If NaCl works, try other weak kosmotropes (1,2) or strong (4-6):

1. NaCl	200 mM	0-1 M
2. KCl	200 mM	0-1 M
3. MgSO ₄	100 mM	0-0.4 M
4. (NH ₄)SO ₄	50 mM	0-0.2 M
5. Na ₂ SO ₄	50 mM	0-0.2 M
6. Cs ₂ SO ₄	50 mM	0-0.2 M

If glycerol or TAMO worked, try following osmolytes:

1. Glycerol	10 %	0-40 % v/v
2. TMAO	20 mM	0-1 M
3. Glucose	1 M	0-2 M
4. Sucrose	500 mM	0-1 M
5. Ethylene glycol	10 %	0-60%
6. Trehalose	500 mM	0-1 M

If urea worked, try mild chaotropes:

1. Urea	0.5 M	0-1.5 M
2. NaI	0.2 M	0-0.4 M
3. CaCl ₂	0.1 M	0-0.2 M
4. MgCl ₂	0.1 M	0-0.2 M

If arginine worked, try other amino acids derivatives, as well as osmolytes:

1. Glycine		0.5-2%
2. Arginine	200 mM	0-2 M
3. Arginine ethylester	200 mM	0-500 mM
4. Proline	200 mM	0-1 M
5. Potassium glutamate	200 mM	0-500 mM

If Nonidet P40 worked, try other detergents:

1. Nonidet P40	0.2%	0-1 %
2. Tween 80	0.01%	0.02 % (v/v)

Trials of UlpI cleavage of SUMO tag

Fish LARP6 proteins contained an N-terminal SUMO tag to increase solubility and facilitate purification. In order to produce tag-free protein, the SUMO tag was cleaved by UlpI digest. To optimize cleavage conditions, 300 μ L of the stored fish protein were treated with 100 μ L of 25 μ M UlpI and incubated at 16 °C overnight. An untreated 20 μ L sample of the protein was saved as negative control with 5X SDS sample buffer and stored at -20 °C. 20 μ L sample of reaction was removed after 2 h, 4 h, 6 h, 8 h, and overnight and combined with 5X SDS sample buffer before storage at -20 °C. Samples were then analyzed by gel electrophoresis and Coomassie stain.

Limited proteolysis

To probe the structural topology of the LARP6 proteins, limited proteolysis analysis with trypsin was performed. A dilution of stored LARP6 protein was prepared, and a sample removed and combined with 5X SDS sample buffer and stored at -20 °C as a “0” timepoint sample. Then, trypsin (0.1 mg/mL stock concentration) was added to the remaining protein to a final concentration of ~1.6 μ g/mL and the reactions were incubated at various temperatures. Small aliquots were removed at indicated timepoints and treated with 1x SDS sample buffer as before. Samples were analyzed by denaturing gel electrophoresis and silver stain.

Electrophoretic Mobility Shift Assays (EMSAs)

In order to measure binding interaction between LARP6 and the putative RNA ligands, EMSAs were performed. A 10X binding buffer stock (100 mM Tris-HCl (pH 7.46 at 4 °C), 200 mM KCl, and 10 mM MgCl₂) was made and stored at -20 °C. The final binding buffer (1X binding buffer + 15% glycerol) was prepared from the stock. Biotinylated RNA was thawed on ice before heating at 80 °C for 2 minutes and then placed back on ice. Serial protein dilutions were prepared in 0.5 mL autoclave-sterile microcentrifuge tubes at twice the protein concentration of the final binding reaction.

In new tubes, the biotinylated RNA stock was serially diluted to twice the RNA concentration of the final binding reaction in 0.2 µm-filtered Tris-EDTA (TE). The binding reactions were prepared by a 1:1 mixture of the 2X protein dilutions and the 2X RNA dilution, and incubated on ice for 1 h to reach equilibrium. Reactions were loaded onto a 5.5% polyacrylamide (29:1, acrylamide: bisacrylamide) tris-borate-EDTA (TBE) native 1.5 mm thick minigel containing 5% glycerol. The gels were electrophoresed in cold 1X TBE Buffer at 200 V for 15 min with -70 °C ice blocks to absorb excess heat. A 50% sucrose/bromophenol blue solution was used to track gel progress. The protein/RNA complexes were transferred to a Hybond N+ membrane (GE Healthcare) with the Transblot Turbo (25 V, 1.0 A, 30 min), then crosslinked with a UV oven for 45 s at 120 mJ/cm² and the membrane was allowed to dry overnight. The biotinylated RNA was detected using the LightShift RNA EMSA Chemiluminescence Nucleic Acid Detection Module (ThermoScientific). The

dried membrane was rehydrated with the Nucleic Acid Detection Blocking Buffer from the commercial detection module. Then blocked with 20 mL of blocking buffer for 15 min with gentle shaking with the same buffer. Then, the membrane was incubated with 20 mL of blocking buffer containing 66.7 μ L of streptavidin-horseradish peroxidase for 15 min with gentle shaking. Following 4 washes of 5 min each with 20 mL of 1X washing solution, the membrane was treated with 30 mL of Substrate Equilibration buffer for 5 min with gentle shaking. The membrane was transferred out of the incubation tray and carefully drained dry by blotting a corner onto a paper towel. The membrane was then placed RNA side up on clean plastic wrap and covered with a small amount of a 1:1 mixture of luminol enhancer solution and peroxidase solution for 5 min. The membrane was carefully drained again and was placed RNA-side down on a fresh piece of plastic wrap. The membranes were then imaged using the chemiluminescence protocol of the ChemiDoc XRS+ molecular imager.

III. RESULTS

***HsLARP6* expression and purification**

E. coli BL21 (DE3) pLysS and Rosetta™ (DE3) pLysS competent cells were transformed with pET28a-*HsLARP6* to determine best expression cell line. Briefly, after mid-log phase was reached, a sample for basal level of expression was taken from both. Then, protein expression was induced by 1 mM of IPTG and cell samples were collected at various timepoints. Samples were analyzed by gel electrophoresis and western blot (Figure 11). We found that by Coomassie stain, Rosetta™ (DE3) pLysS showed the most robust yield, with a strong band at around 60 kDa (*HsLARP6* construct MW is 57 kDa) increasing over expression time. The anti-His Western blot confirmed the presence of the expressed His₆-*HsLARP6* protein in both cell lines.

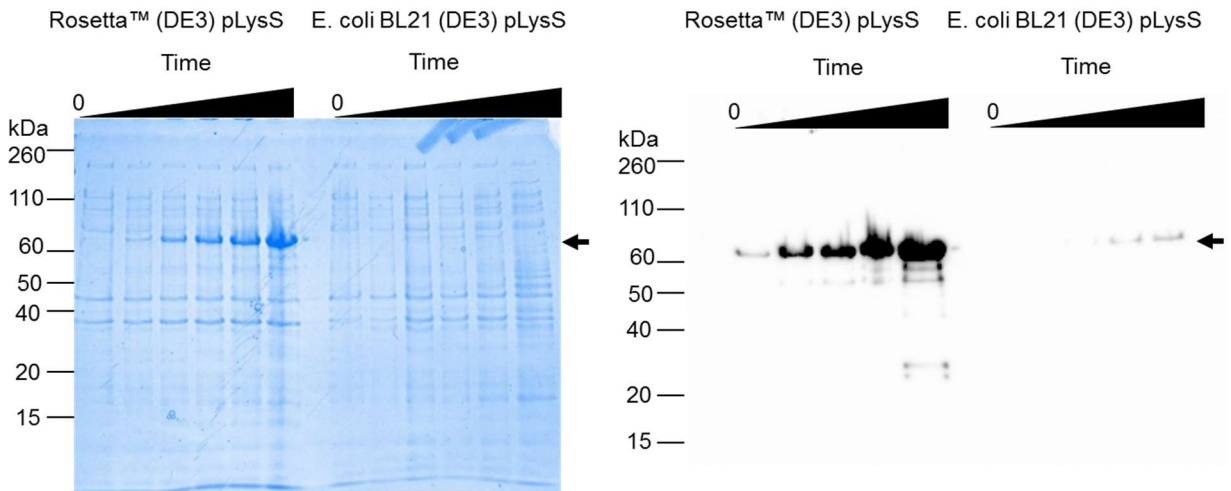


Figure 11. Expression trials of *HsLARP6*. pET28 plasmid containing the coding sequence for *H. sapiens* LARP6 was obtained from Dr. Bayfield and transformed into *E. coli* BL21 (DE3) pLysS and Rosetta™ (DE3) pLysS competent cells. Following induction with 1 mM IPTG, expression was monitored at regular intervals of 2,4,6,8, and overnight. Cell pellets were lysed in 1X SDS-PAGE sample buffer and subjected to gel electrophoresis, followed by either (A) Coomassie blue stain or (B) anti-His western blot. Rosetta cells show expression of *HsLARP6* in a robust manner compared to BL21 (DE3) (arrow around 60 kDa).

However, the strong intensity of the band in the Rosetta™ (DE3) pLysS cells demonstrated that they are the best option to express the human LARP6 in larger quantities compared to *E. coli* BL21 (DE3) pLysS. This result was verified by analyzing the human LARP6 protein sequence with the Graphical Codon Usage Analyser.³⁴ This analysis found that the coding sequence contains anticodons that are underrepresented in the bacterial complement of tRNAs. Therefore, the supplemental tRNA plasmid in Rosettas is able to support the expression of these codons.

The overnight induction sample showed the highest protein production of all the time points (Figure 11A). However, due to more degradation products seen in the anti-His western blot (Figure 11B), we decided to induce for 8 hours. This was chosen to improve the ratio of full-length protein to degradation products.

Rosetta cell pellets expressing His₆-HsLARP6 were lysed by sonication and subjected to purification by nickel affinity chromatography. The elution fractions were analyzed by gel electrophoresis (Figure 12). A cell pellet sample (“p”) was used to show the purification process from start to finish. As expected, the elution fractions showed a band around 60 kDa, characteristic of full-length His₆-HsLARP6 (actual molecular weight is 57 kDa) and possible degradation products shown by lower bands in gel. Elution fractions showing highest yield were collected and prepared for size exclusion chromatography.

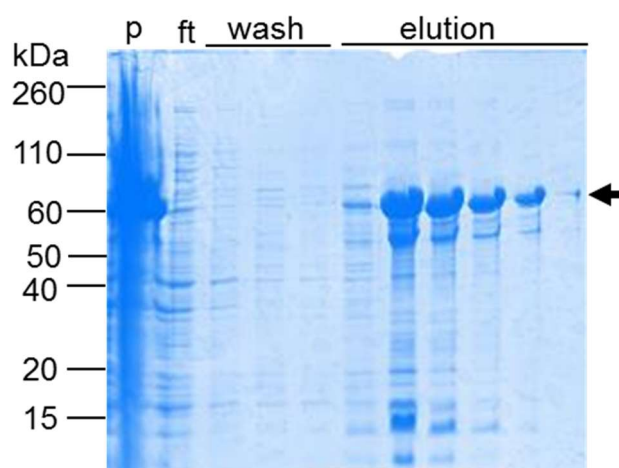


Figure 12. Full length *HsLARP6* Affinity chromatography fractions.

The cell lysate was purified using nickel-affinity chromatography ("p", expression culture pellet, "ft", flowthrough). Aliquots were separated by gel electrophoresis and stained with Coomassie. Elution fractions 2—5 contained full-length *HsLARP6* (arrow) were pooled for further purification.

The elution profile obtained from the FPLC (Figure 13) displayed the separation of the full-length His₆-*HsLARP6* and the other bacterial proteins over the S200 Sephadex as monitored by UV absorbance at 280 nm. The elution profile displayed two major peaks: one at ~71 mL and one at ~130 mL. Based on the standards and the molecular weight calculation from the regression, it was suspected that the peak at 71 mL contained full-length His₆-*HsLARP6*.

When back calculating molecular weight by elution profile, the protein collected have an apparent molecular weight of 177.6 kDa. The protein seems elute as a trimer (comparing predicted molecular weight of 57 to the experimental size of 177.6 kDa).

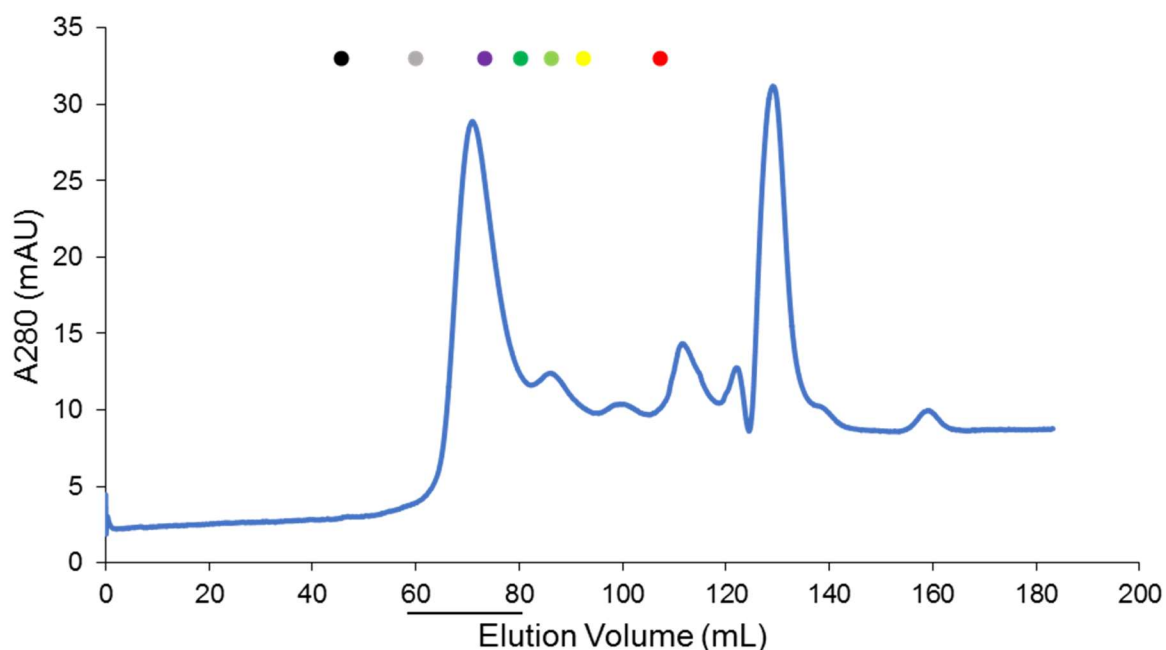


Figure 13. S200 size exclusion chromatography of *HsLARP6*.

HsLARP6 that eluted from affinity chromatography was loaded to the FPLC in which was monitored by UV absorbance at 280nm at a 1mL/min flowrate. Fractions were collected in 2 mL aliquots automatically.

To analyze and decide which fractions to collect, the fractions under the peak were analyzed by gel electrophoresis (Figure 14). A sample of the protein mixture that was loaded onto the FPLC was used to compare the purification quality acquired by size separation. From the fractions analyzed, some showed the expected 60 kDa band corresponding to full-length His₆-*HsLARP6* protein. In addition, those fractions contained fewer degradation products were visible due to the size separation performed by the FPLC. Fractions 22—28 were pooled and stored as 50 μ L aliquots at -70 °C for further experiments.

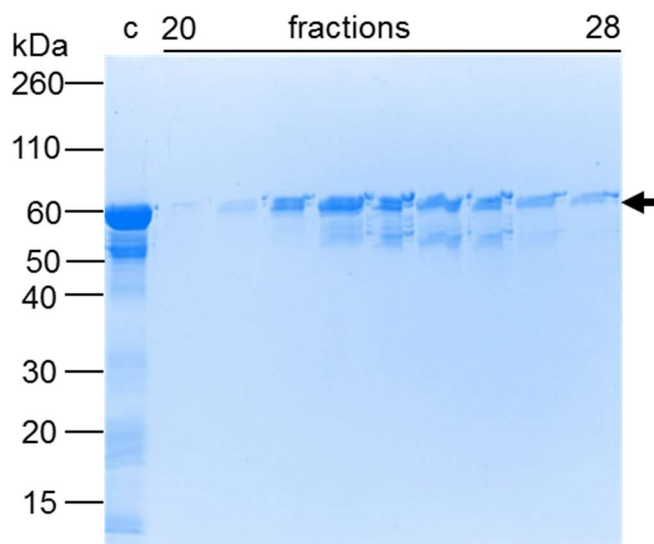


Figure 14. SDS-PAGE analysis of fractions from size exclusion chromatography of full-length *HsLARP6*. The pooled elution from affinity chromatography was filtered through a 0.2 μ m filter (lane “C”) and separated on an S200 Sephadex column. Fractions 20—28 were analyzed by SDS-PAGE and Coomassie blue staining. Fractions 22—28, corresponding to elution volumes 68—80 mL, contained full-length *HsLARP6* protein. Fractions were stored in 50 μ L aliquots at -70°C for further experiments.

The SDS-PAGE analysis confirmed that fractions 22—28, which corresponded to the peak at 75 mL, contained full-length protein. To ensure that His₆-*HsLARP6* was not degrading under these storage conditions, an aliquot was thawed and analyzed by both anti-His and anti-*HsLARP6* western blot (Figure 15). The His₆ tag attached at the N-terminus served as a mechanism to check for degradation products that contain this section. Similarly, the anti-*HsLARP6* antibody is specific to a sequence of amino acids located close to the C-terminus.

The use of these two antibodies covers both ends of the protein, allowing for a fuller picture of protein degradation and stability. The anti-His blot showed three bands, which is similar to the Coomassie-stained SDS-PAGE gel from the S200. In contrast, the anti-*HsLARP6* blot only showed a single intense band around the expected MW. Since both, the anti-His and the anti-*HsLARP6* C-terminal antibodies can detect *HsLARP6* at the expected molecular weight (~60 kDa), it demonstrates that full-length *HsLARP6* was successfully purified as well as stable when stored at -70 °C.

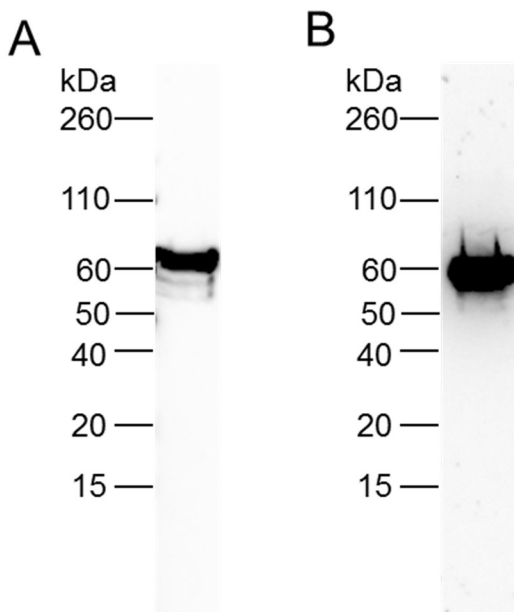


Figure 15. Verification of *HsLARP6* stability at -70 °C. A stored aliquot of full length-*HsLARP6* was analyzed by SDS-PAGE and western blotting. **A)** Western blotting with His-α probe (Pierce). **B)** Western blotting with an antibody against the C-terminus of *HsLARP6* (Abcam). Both blots confirmed that *HsLARP6* (band around the 60 kDa MW) is stable at -70 °C.

Purification of Ulpl protease and optimization of SUMO cleavage

The expression platform for the fish homologs of LARP6 is an N-terminal fusion to the Small Ubiquitin Modifier protein (SUMO) from *Saccharomyces cerevisiae*.³¹ The SUMO protein functions as a solubility and stability enhancer.³⁰ Additionally, an N-terminal His₆ tag allows for affinity purification. To characterize the fish proteins, the SUMO tag must to be removed. The ubiquitin-like specific protease 1 (Ulpl) is a yeast protease that identifies SUMO and cleaves the Gly-Gly-Ser linker.³⁵ The pET28-Ulpl plasmid was obtained from Dr. Lima, courtesy of Cornell University. His₆-tagged Ulpl was expressed and purified by affinity chromatography. All fractions were analyzed by SDS-PAGE and Coomassie stain (Figure 16).

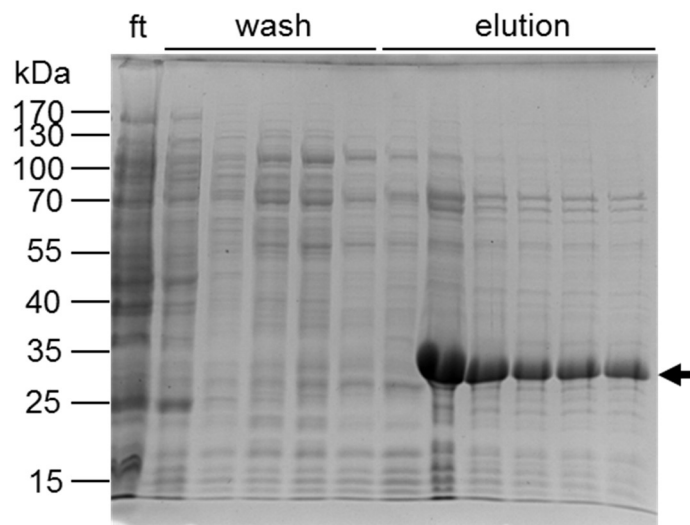


Figure 16. Affinity purification of His₆-Ulpl. Ulpl is required to remove the SUMO tag from the purified proteins for further studies. Cell pellets of BL21 (DE3) pLysS cells expressing Ulpl were resuspended with proper buffer, sonicated as stated and centrifuged. The cell lysate was purified using nickel-affinity chromatography ("ft", flowthrough). Aliquots were separated by gel electrophoresis and stained with Coomassie blue. Elution fractions 2—6 (arrow) were pooled for further purification.

The Ulp1 protein is marked by an arrow, showing that the tagged enzyme eluted in fractions 2—6. The many other bands in the gel electrophoresis lanes of the nickel affinity purification fractions show that there are other proteins co-eluting with Ulp1.

Therefore, the protein was concentrated to approximately 3 mL, and subjected to size exclusion chromatography using an S75 Sephadex column. Since the S75 resin has a smaller pore size compared to S200, it is more appropriate for the molecular weight of Ulp1 (~30 kDa). Protein elution was analyzed by A_{280} and collected in 2 mL fractions. Figure 17 shows the elution profile. However, to verify both peaks were analyzed by SDS-PAGE and Coomassie stain (Figure 18). The first lane is a sample of the concentrated total of protein prior to SEC. Nearly all the proteins with an apparent molecular weight larger than Ulp1 eluted first (in fractions 11—16).

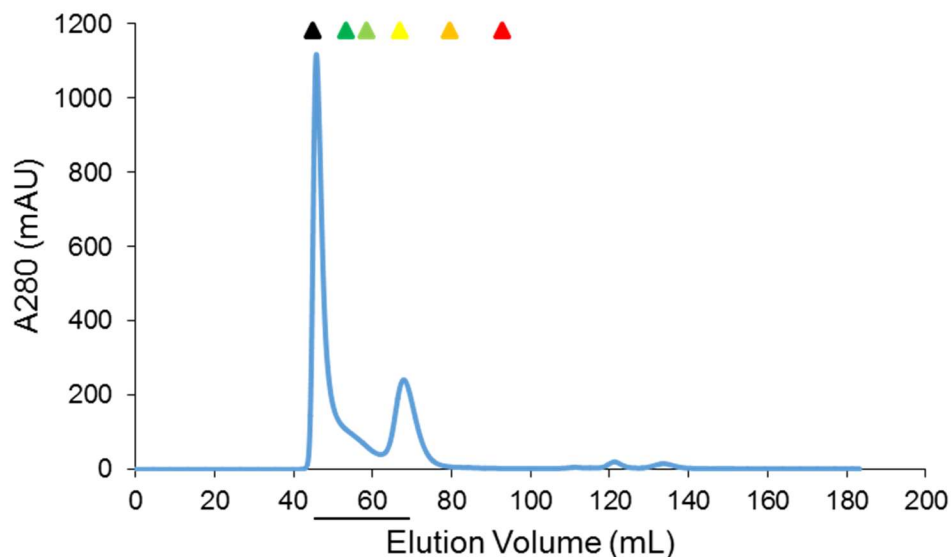


Figure 17. S75 size exclusion chromatogram of His₆-Ulp1. The His₆-Ulp1 protein that was eluted from affinity chromatography was subjected to size exclusion chromatography over an S75 Sephadex column.

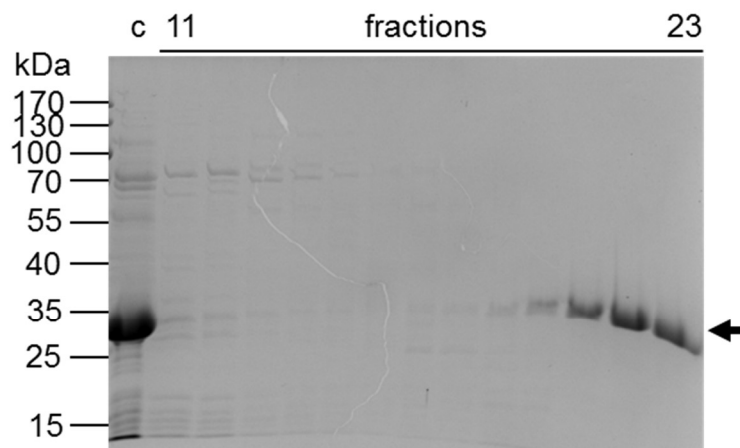


Figure 18. SDS-PAGE analysis of size exclusion chromatography of His₆-Ulp1. The pooled elution from affinity chromatography were filtered through a 0.2 μ m filter (lane “C”) and separated on an S75 Sephadex column. Fractions 11—23 were analyzed by SDS-PAGE and Coomassie blue staining. Fractions 21—24, corresponding to elution volumes 65—73 mL, were pooled out. Fractions were mixed and added 500 μ L of glycerol. Ulp1 was aliquot into 500 μ L units and stored at -20 $^{\circ}$ C.

Fractions containing Ulp1 were pooled and brought to a final concentration of 15% glycerol before freezing 500 μ L aliquots and storage at -20 $^{\circ}$ C.

To test Ulp1 enzymatic activity, a reaction mix was prepared with a 1000:1 molar ratio of His₆-SUMO-DrLARP6a (purified as described below) to His₆-Ulp1. The reaction was incubated at 16 $^{\circ}$ C and monitored over time. Samples were then analyzed by SDS-PAGE and Coomassie stain (Figure 19). Over time it is noticeable that Ulp1 is active as two defined bands, which can be seen after 2 hours. Uncleaved His₆-SUMO-DrLARP6a has a molecular weight of 67 kDa (in the gel, it runs around the 70 kDa mark). As the reaction was incubated for longer times, the band around 60 kDa seems to increase in intensity as the enzymatic product increases (DrLARP6a). However, based on band intensity not even a 50% cleavage rate, the last time point recorded is 19 hours, was achieved.

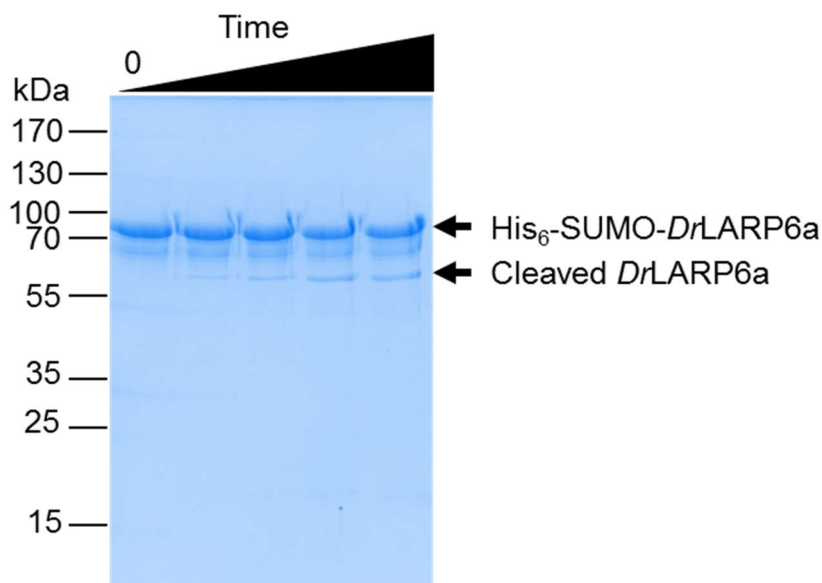


Figure 19. SDS-PAGE analysis of initial Ulp1 cleavage of His₆-SUMO-*DrlARP6a*. 250 μ L of His₆-SUMO-*DrlARP6a* were treated with 0.3 μ L of purified His₆-Ulp1 for SUMO cleavage (1000:1) mol/mol ratio. Reaction was held at 16 °C and monitored every 0, 2, 4, 16 and 19 hours. Samples were analyzed by gel electrophoresis and Coomassie stain. His₆-SUMO-*DrlARP6a* MW is 67 kDa. As reaction took place, lower bands show up meaning enzymatic cleavage was happening. However, less than 50% cleavage rate was obtained after 19 hours of incubation.

Two new reactions, defined as 10X and a 50X, were set with a higher molar ratio to optimize Ulp1-mediated cleavage of the SUMO tag (Figure 20). The 10X reaction after 2 hours of incubation gave the same outcome that the original reaction did after 19 hours. With the 10X molar ratio, the overnight timepoint shows almost complete cleaved *DrlARP6a* with few degradation products. The 67 kDa full-length band characteristic of His₆-SUMO-*DrlARP6a* can be seen disappearing as reaction takes place over time. In addition, three new bands were noticed after the 4 hour timepoint: the cleaved *DrlARP6a* band (~55 kDa), the Ulp1 band (~30 kDa) and SUMO (~22 kDa). However, the 50X reaction gave an almost 100% cleavage product within the first 2 hours of incubation. These conditions were applied for the purifications of the fish LARP6 proteins.

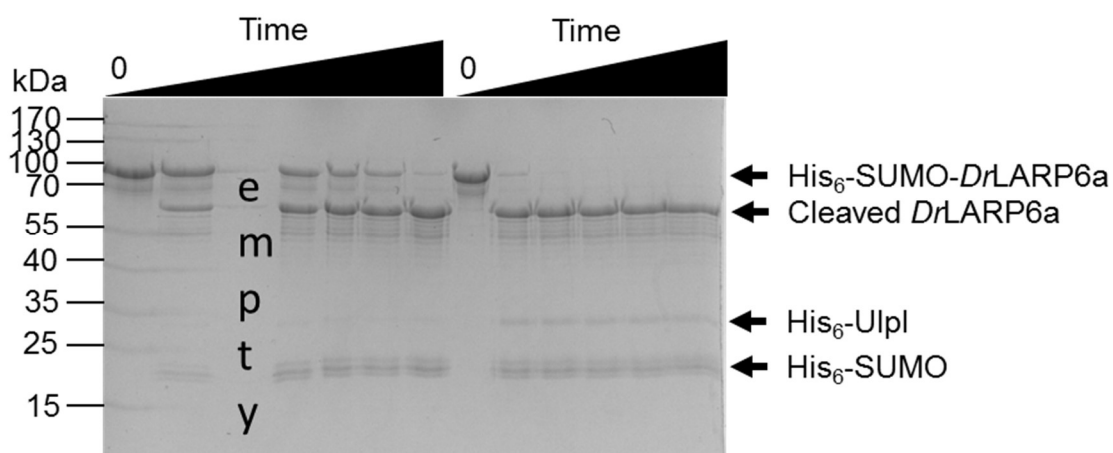


Figure 20. SDS-PAGE analysis of second Ulp1 cleavage of His₆-SUMO-*DrlARP6a*.

Previous digest was done with 250 μ L of His₆-SUMO-*DrlARP6a* to 0.3 μ L of Ulp1 (1000:1) mol/mol ratio. Two new reactions were set up ("10 X", left, "50X", right) and analyzed as before. Reaction was hold at 16 °C and monitored every 0,2,4,6, 8 hours and overnight. Samples were analyzed by gel electrophoresis and Coomassie stain. The 67 kDa full-length band, as reaction takes place over time, starts disappearing. In addition, 3 new bands can be seen: the cleaved *DrlARP6a* band (~55 kDa), the Ulp1 band (~30 kDa) and SUMO (~22 kDa)

Cloning of Fish LARP6 Expression Constructs

Xiphophorus maculatus cDNA was used as the template for a PCR reaction to amplify the coding sequence for the *X. maculatus* LARP6 protein. Figure 21A shows the amplified product on a 0.7% agarose gel as well as the double digest (BamHI/XhoI) product from a miniprep (pET28-SUMO-*XmLARP6*). Plasmid was then sent for commercial Sanger sequencing to confirm the insert.

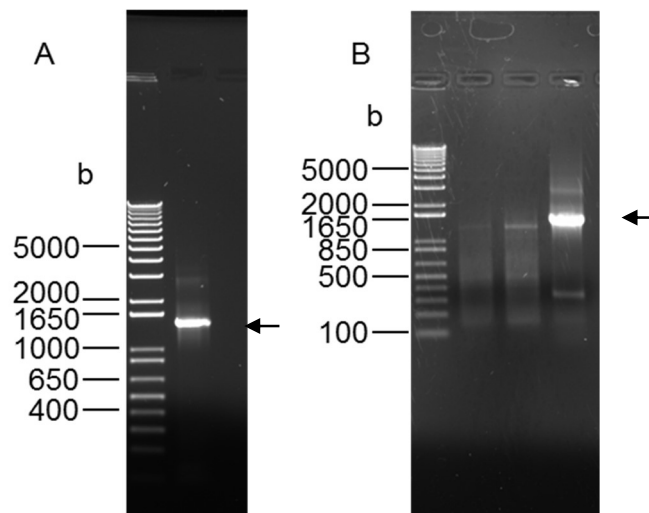


Figure 21. *XmLARP6* amplification from cDNA and confirmation of successful double digest. A) *XmLARP6* was amplified from *X. maculatus* cDNA using primers from table 3. A band ~1400 base (arrow) correspond to the size of *XmLARP6* DNA. B) Three minipreps from pET28a SUMO *XmLARP6* were subjected to double digest by BamHI and XhoI. The expected band ~1400 was obtained from sample #3. Sample then was send to sequencing.

Danio rerio LARP6a and LARP6b were amplified using synthetic gene sequences, produced by GenScript. Both were inserted into pET28-SUMO plasmid. Figures 22 and 23 show the amplification and double digest confirmation of the amplified and inserted product. The proper DNA fragment was cloned in frame and confirmed by commercial Sanger sequencing.

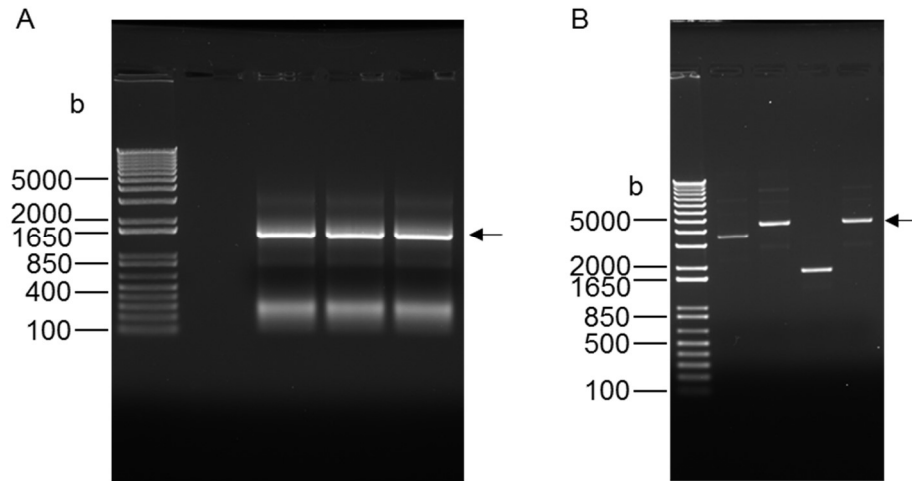


Figure 22. *DrLARP6a* amplification from synthetic DNA and pET28SUMO-*DrLARP6a*.

A) *DrLARP6a* was amplified from synthetic cDNA using primers from table 3. A band ~1400 base (arrow) corresponding to the size of *DrLARP6a* DNA. B) Lane two shows pET28, Lane 3 to 5 three different miniprep ligation products (pET28-*DrLARP6a*, arrow). Ligation products from lane 3 and 5 were sent for sequencing.

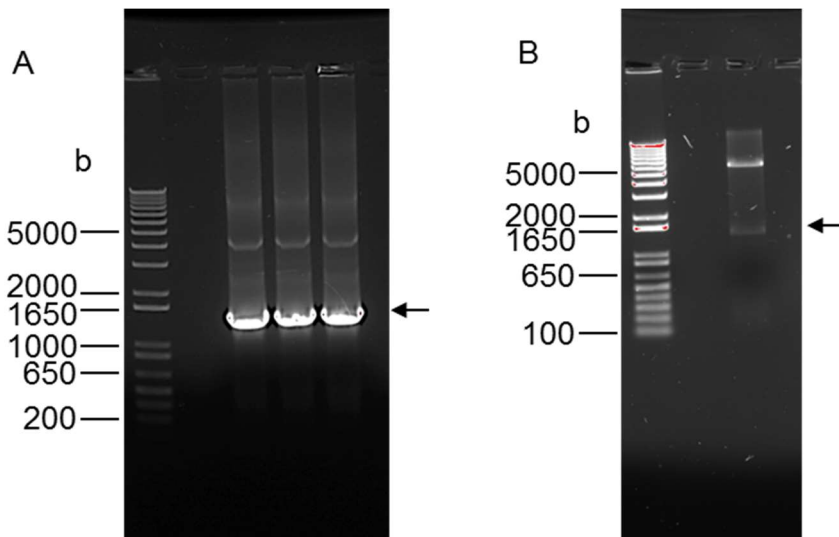


Figure 23. *DrLARP6b* amplification from synthetic DNA and pET28SUMO-*DrLARP6b* double digest.

A) *DrLARP6b* was amplified from synthetic cDNA using primers from table 3. A band ~1400 base (arrow) corresponding to the size of *DrLARP6b* DNA. B) Lane 3 shows double digest (BamHI and XhoI) of pET28-*DrLARP6b* from a miniprep ligation products. Arrows point to the characteristic *DrLARP6b* band. Plasmid was sent for sequencing to confirm insert and correct coding frame.

Expression of Fish LARP6 Proteins

Based on the robust yield of human LARP6 protein obtained using the Rosetta™ (DE3) pLysS competent cells, this strain was also used for the fish expression trials. As previously described for the human homolog, the SUMO-tagged fish constructs were each analyzed for protein expression (Figures 24, 25, and 26).

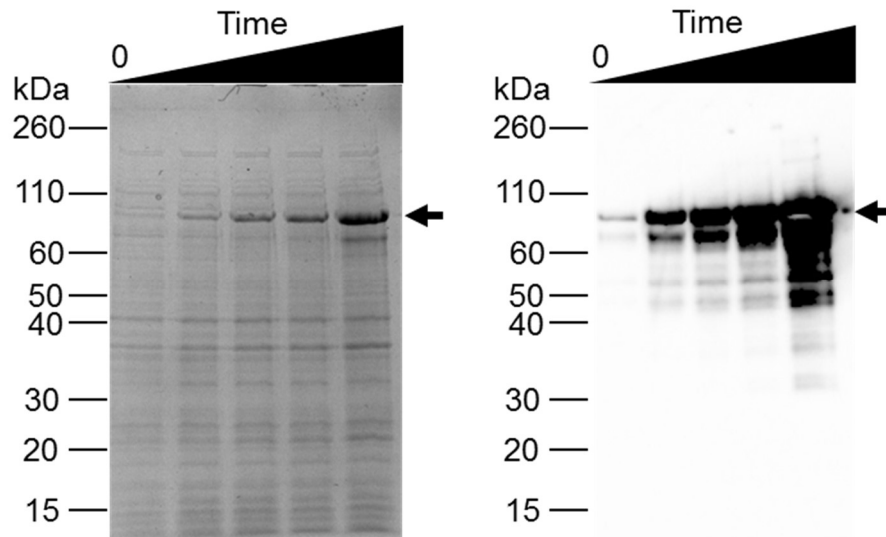


Figure 24. Expression trials of *XmLARP6* in Rosetta™ (DE3) pLysS competent cells. *X. maculatus* LARP6 was cloned from cDNA and recombinantly expressed as a His₆-tagged SUMO fusion protein in Rosetta (DE3) *E. coli*. Following induction with 1 mM IPTG, cells were placed to incubate at 18 °C. Expression was monitored at regular intervals of 2, 4, and 6 hours, and after overnight expression. Cell pellets were lysed in 1X SDS-PAGE sample buffer and analyzed by SDS-PAGE, followed by either Coomassie stain (left) or anti-His western blot (right). Between samples, a band starts to appear around the 70 kDa (arrow). This band increases intensity as incubation time increases meaning His₆-SUMO-*XmLARP6* (66 kDa) is being expressed.

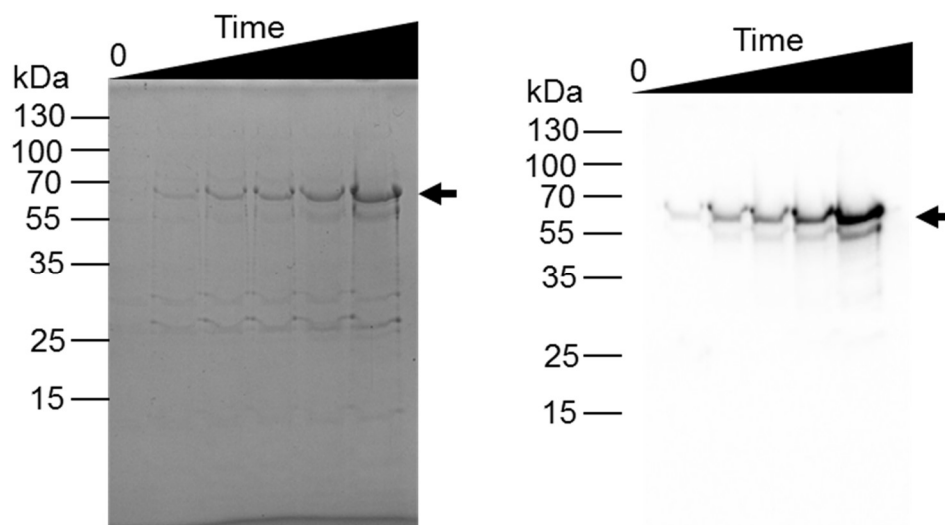


Figure 25. Expression trials of *DrLARP6a* in Rosetta™ (DE3) pLysS competent cells. The *D. rerio* LARP6a gene commercially synthesized and subcloned for recombinant expression as a His₆-tagged SUMO fusion in Rosetta (DE3) *E. coli*. Following induction with 1 mM IPTG, cells were placed to incubate at 18 °C. Expression was monitored at 2, 4, 6, and 8 hours, and after overnight expression. Cell pellets were lysed in 1X SDS-PAGE sample buffer and analyzed by SDS-PAGE, followed by either Coomassie stain (left) or anti-His western blot (right). Between samples, a band starts to appear around the 70 kDa (arrow). This band increases intensity as incubation time increases confirming His₆-SUMO-*DrLARP6a* (67 kDa) expression.

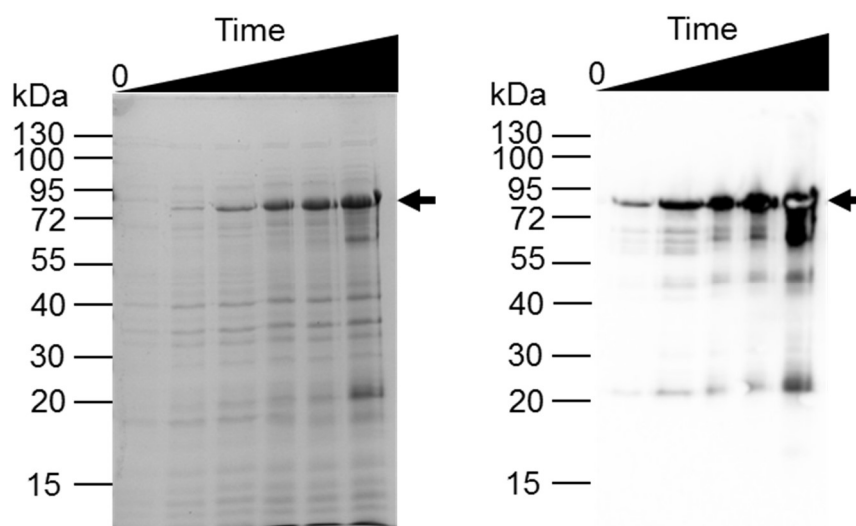


Figure 26. Expression trials of *DrLARP6b* in Rosetta™ (DE3) pLysS competent cells. The *D. rerio* LARP6b gene commercially synthesized and subcloned for recombinant expression as a His₆-tagged SUMO fusion in Rosetta (DE3) *E. coli*. Following induction with 1 mM IPTG, cells were placed to incubate at 18 °C. Expression was monitored at 2, 4, 6, and 8 hours, and after overnight expression. Cell pellets were lysed in 1X SDS-PAGE sample buffer and analyzed by SDS-PAGE, followed by either Coomassie stain (left) or anti-His western blot (right). As cells are incubating, a band around the 70 kDa mark shows-up (arrow). The band intensity increases as incubation proceeds giving an insight in His₆-SUMO-*DrLARP6b* (66 kDa) expression.

Each of the three fish proteins is predicted to have an apparent molecular weight similar to *HsLARP6* (~52 kDa for all three fish LARP6). However, since the fish expression constructs contain a SUMO tag to increase protein stability, the expressed fusion proteins migrate at a higher apparent molecular weight in the SDS-PAGE (~66-67 kDa). All three proteins are robustly expressed, as demonstrated by the time-dependent increase in a protein band around 70–75 kDa (Figures 24, 25, 26; left panels). The fusion proteins of interest were confirmed by anti-His western blot, which reacts with the N-terminal His₆ tag. Although overnight induction produces a large amount of full-length protein, there are a great number of degradation products, seen as smaller bands in both the Coomassie stain and western blot images. Therefore, the 8-hour expression time was chosen for large-scale purification.

Initial Purification of His₆-SUMO-XmLARP6

The cell pellet from a 1 L culture expressing His₆-SUMO-XmLARP6 was lysed open and subjected to nickel affinity chromatography, as previously described for the human LARP6 protein. The fractions from the affinity chromatography were analyzed by gel electrophoresis and Coomassie stain (Figure 27). A pellet sample was used to show the purification process from start to finish. Elution fractions showed bands ~60 kDa (band reflects the SUMO tag fused to the XmLARP6 protein) and possible degradation products shown by lower bands in gel. Elution fractions 2 through 6 contained the most protein, and were collected and prepared for size exclusion chromatography over the Sephadex S200 column.

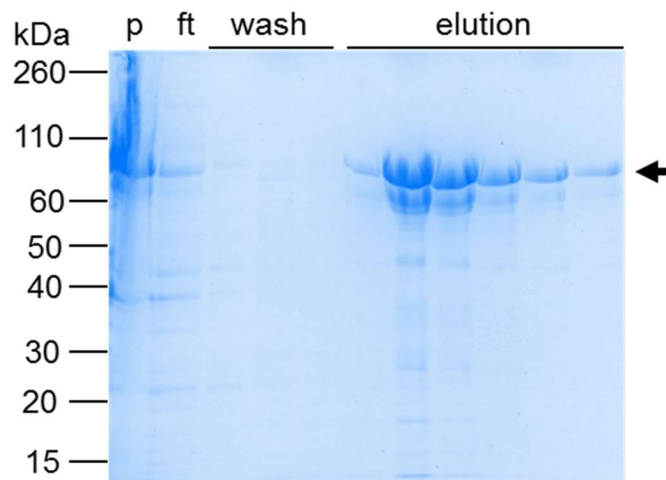


Figure 27. His₆-SUMO-XmLARP6 affinity chromatography fractions profile. The cell lysate was purified using nickel-affinity chromatography (“p”, pellet, “ft”, flowthrough). Aliquots were separated by gel electrophoresis and stained with Coomassie blue. Elution fractions 2—6 (arrow) were pooled for further purification. The typical 65—70 kDa band characteristic of the His₆-SUMO-XmLARP6 construct ensures that the desired protein is being eluted.

Figure 28 represents the elution profile of His₆-SUMO-*Xm*LARP6 and degradation products from the S200 Sephadex size exclusion column as monitored by UV absorbance at 280 nm. The profile shows that the protein started eluting around where Blue Dextran elutes (Table 1), meaning that some protein aggregation occurred. In addition, a broad set of four overlapping peaks is seen around the volume at which His₆-SUMO-*Xm*LARP6 protein is expected to elute (81.93 mL). Fractions within the overlapping peaks were analyzed by SDS-PAGE and Coomassie stained to decide which ones to pool, aliquot, and store (Figure 29). The peaks have an apparent MW of 1117.41, 178.23 and 137.71 kDa respectively. Based on the elution profile displayed by *Hs*LARP6, it is possible that His₆-SUMO-*Xm*LARP6 is eluting with an apparent higher molecular weight from the column.

Based on the gel, the entire peak is composed of His₆-SUMO-*Xm*LARP6. It is notable that protein concentration is not constant; it increases from left to right. This indicates that some protein aggregated and eluted as a higher molecular weight complex. As can be seen on the right side of the gel, the concentration of SUMO-*Xm*LARP6 in the fractions increases and some degradation products begin to appear. Fractions 21—24 were pooled and stored for later use. However, based on the elution profile, being unable to differentiate peaks by the denaturing gel and the ratio between full-length and degradation products, we were unable to trust this purification. To ensure fish full length protein purification, protein stability and less degradation products, a buffer optimization approach followed.

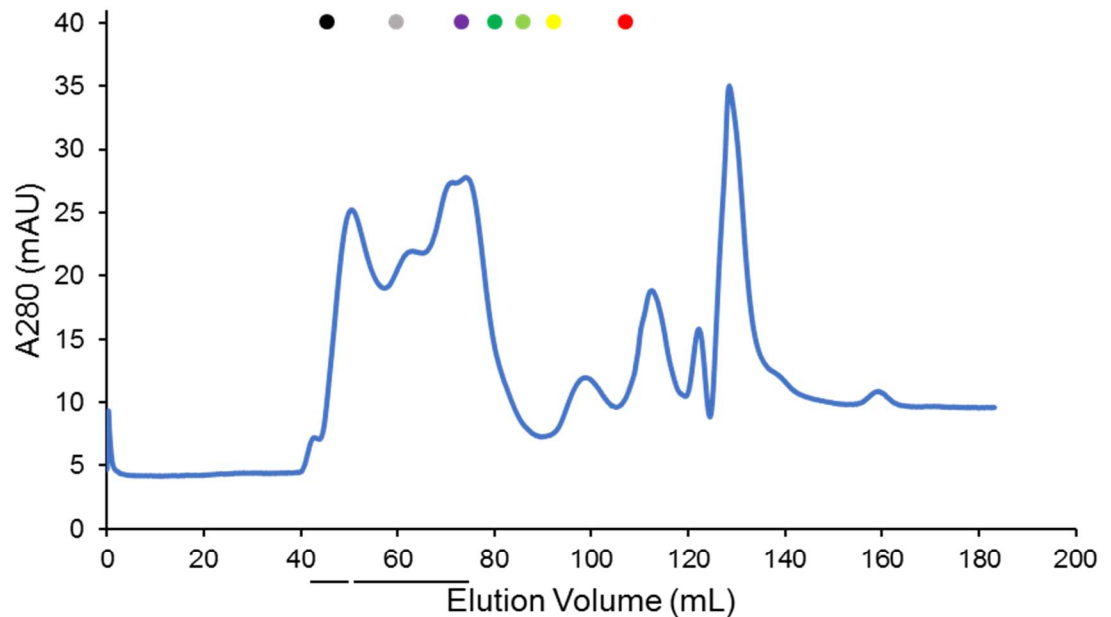


Figure 28. S200 size exclusion chromatography of SUMO-tagged *XmLARP6*. Elution fractions selected from affinity chromatography were concentrated to 2—2.5 mL against a molecular weight cutoff (MWCO) of 10 kDa by centrifugation at 5,000 x g, and filtered with a 0.2 μ m filter. Filtered protein was then loaded to the FPLC in which was monitored by UV absorbance at 280nm at a 1mL/min flowrate. Fractions were collected in 2 mL aliquots automatically. In contrast to a single eluting peak, there are three peaks that eluted between the expected elution volume (81.9 mL). The peaks have an apparent MW of 1117.41, 178.23 and 137.71 kDa, respectively.

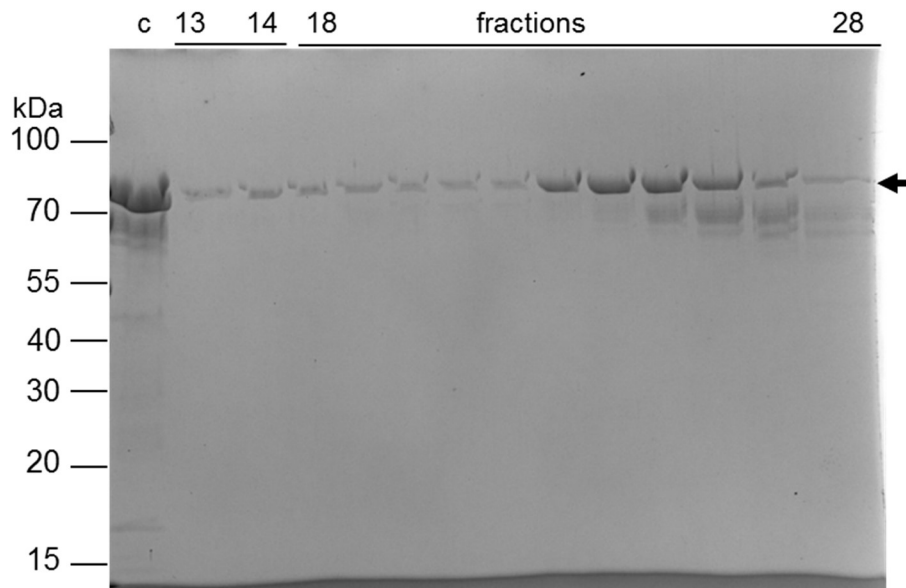


Figure 29. SDS-PAGE analysis of fractions from size exclusion chromatography of SUMO tagged *XmLARP6*. The pooled elution from affinity chromatography was filtered through a 0.2 μ m filter (lane “C”) and separated on an S200 Sephadex column. Fractions 13, 14, 18—28 were analyzed by SDS-PAGE followed by Coomassie blue staining. Fractions 21—24, corresponding to elution volumes 66—72 mL, were pooled out. Fractions were stored in 50 μ L aliquots at -70°C for further experiments.

Solubility Screen Assay

Expression trials of His₆-SUMO-*Xm*LARP6 demonstrated robust expression in *E. coli*. However, while purifying with the same buffer conditions as used for *Hs*LARP6, a significant amount of degradation products and aggregates were observed. Even though the elution fractions from affinity chromatography contained a good yield of protein, when subjected to SEC, a large set of overlapping peaks suggested that His₆-SUMO-*Xm*LARP6 was present in different forms (aggregated, full-length and degradation products). To prevent aggregation and improve the solubility and stability of His₆-SUMO-*Xm*LARP6, a filter-based additive screening assay was performed.³³ The preliminary round of screening served two purposes. First, it was used to identify which centrifugal units to use, either the Durapore PVDF 0.1 µm or the Amicon Ultra 0.5 mL 100,000 NMWL. Second, the initial screen also identified which of the general chemical groups (kosmotrope, osmolyte, chaotrope, amino acid, and detergent) most improves His₆-SUMO-*Xm*LARP6 solubility.

From the two filter units, the Amicon Ultra 100,000 NMWL unit retained nearly all the protein on the filter, as a strong band at ~70 kDa can be seen in the retentate fraction on both the short exposure and over-exposure of the western blot (Figure 30, left panels). In contrast, most of the His₆-SUMO-*Xm*LARP6 protein flowed through the Durapore PVDF 0.1 µm unit (Figure 30, right panels). These data demonstrate that the Amicon filter units were too small for the tagged LARP6 protein to flow through. For further screening studies, the PVDF 0.1 µm Durapore filtration units were used.

However, most the protein flowed through the larger pores, and very small amounts were detected as retained on the filter. Therefore, two complementary strategies were used to identify solubilizing additives. First, the presence of smaller molecular weight species was monitored. These bands represent C-terminal truncations of the fusion protein, which may be contributing to protein aggregation.³³ Second, the amount of His₆-SUMO-*XmLARP6* that was in the retentate fractions relative to the flowthrough fractions was considered. Often, the retentate fraction was only visible in overexposed images.

Taking these two parameters of protein solubility together, all five categories of additives that were tested improve the stability of His₆-SUMO-*XmLARP6*. This is observed by comparing the additive-treated samples to the untreated lysate. All additives increase the amount of soluble degradation products in the flowthrough, and simultaneously decrease the amount of insoluble full-length His₆-SUMO-*XmLARP6* in the retentate.

Based on that analysis, the sets of additives that most significantly improved His₆-SUMO-*XmLARP6* solubility were determined to be osmolytes (arginine, glycerol, and TMAO) and a chaotrope (urea).

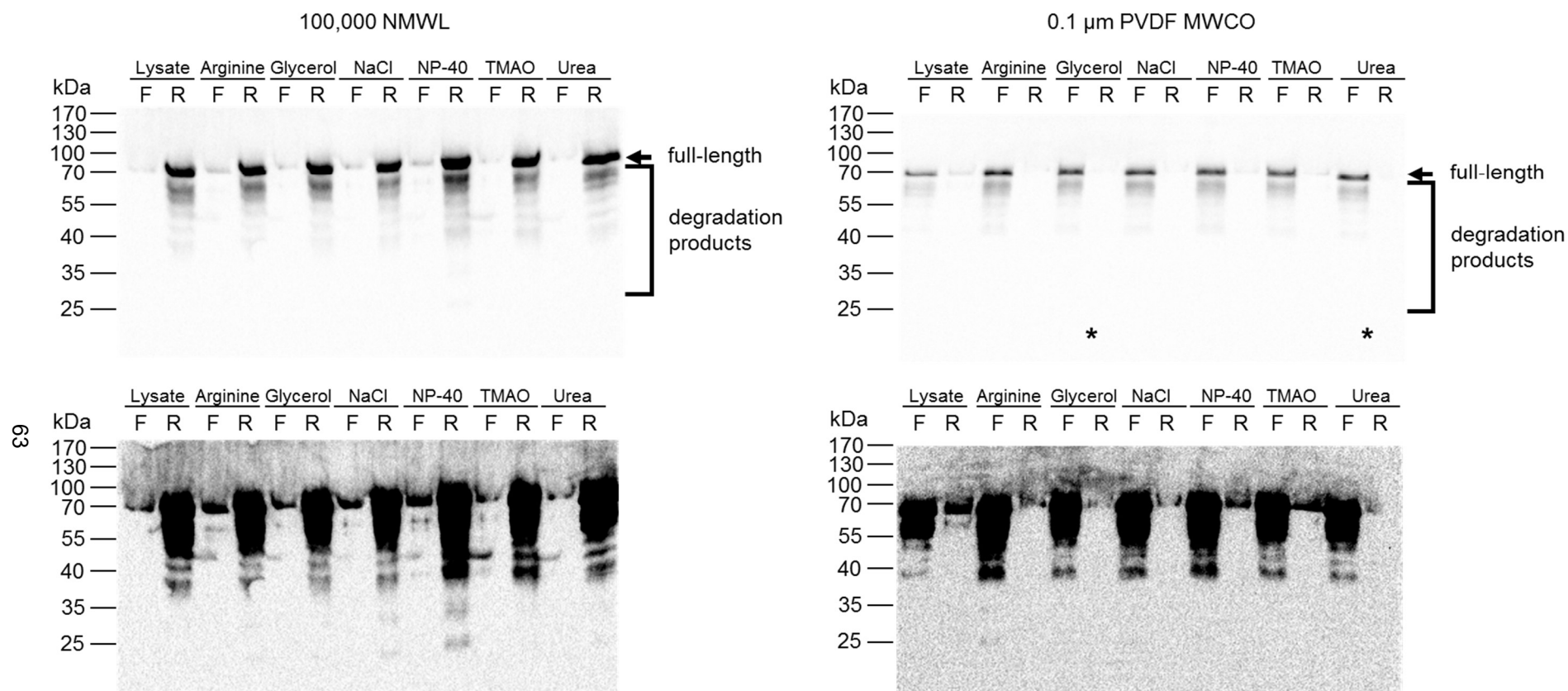


Figure 30. Initial screening of purification buffer additives to increase protein solubility. The initial screening tested different chemical characteristics of potential additives (chaotrope, kosmotrope, osmolytes, and amino acid). Cell pellets expressing His₆-SUMO-XmLARP6 were resuspended with 50 mM NaH₂PO₄/Na₂HPO₄, pH 8.0, and 200 mM NaCl, sonicated to lyse, and centrifuged at 20,000 $\times g$. Screening reactions were composed of standard buffer, clarified supernatant, and additive as indicated. Reactions were equilibrated at 4 °C for 2 hours. Two kinds of centrifugal filter units were tested (Amicon Ultra 0.5 mL 100,000 NMWL (left panels), and Durapore 0.1 µm PVDF (right panels). Both filter units were pre-rinsed with Milli-Q polished deionized H₂O. Reactions were then applied to the filters and centrifuged for 20 min at 15,000 $\times g$ at 4°C. The flowthrough was recovered as the filtrate (“F”). The aggregated retentate (“R”) was re-suspended from the filter surface with 5X SDS-PAGE sample buffer. Reactions were subjected to gel electrophoresis and analyzed by anti-His western blot. Shown are short exposures (top panels) and long exposures (bottom panels). The 0.1µm PVDF centrifugation unit showed better separation of soluble and aggregated protein.

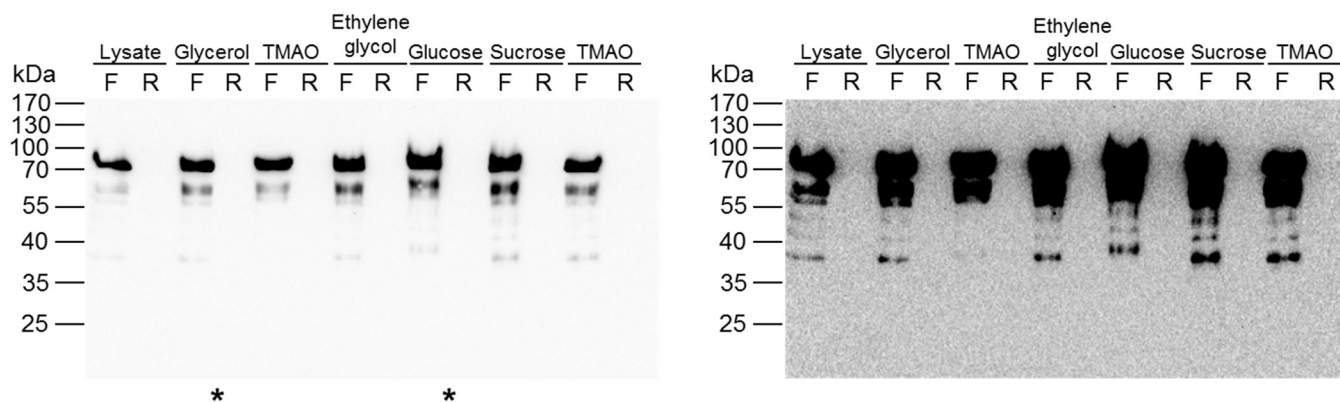


Figure 31. Second iteration of the additive screen, Part 1. The first screening showed that His₆-SUMO-*XmLARP6* stability was improved by osmolytes (arginine, glycerol, and TMAO) and the chaotrope urea. The second screening consisted of testing different osmolytes and chaotropes. For the osmolytes, the assay called for glycerol, TMAO, ethylene glycol, glucose and sucrose. Screening was performed as described on the first round but using the 0.1 μ m PVDF filtration units. The flowthrough was recovered as the filtrate ("F"). The aggregated retentate ("R") was re-suspended from the filter surface with 5X SDS-PAGE sample buffer. Based on the anti-His western blot, all additives improved solubility. Glycerol was picked as a possible buffer additive due to reducing truncated products. Glucose was chosen as possible buffer additive because it increases full-length (*).

The second round of screening consisted of testing other additives within the broad categories of osmolytes and chaotropes.³³ Six osmolytes were tested for effects on protein solubility: 10 % glycerol, 20 and 50 mM TMAO, 10 % ethylene glycol, 1 M glucose, 500 mM sucrose (Figure 31). Of these tested osmolytes, 10 % glycerol and 1 M glucose displayed a considerable increase in both the amount of full-length protein and fragments in the filtrate. Four chaotropes were also screened: 0.5 M urea, 0.2 M calcium chloride, 0.1 M magnesium chloride, and 0.2 M sodium iodide. Of these, the 0.5 M urea and 0.2 M sodium iodide were equally effective at increasing solubility of fragments. In contrast, when the chloride salts were added to the clarified lysate, they induced immediate precipitation, observed as fluffy white particles. This prevented any protein from filtering through the membrane, and no protein was detected as retentate on the filter membrane (Figure 32).

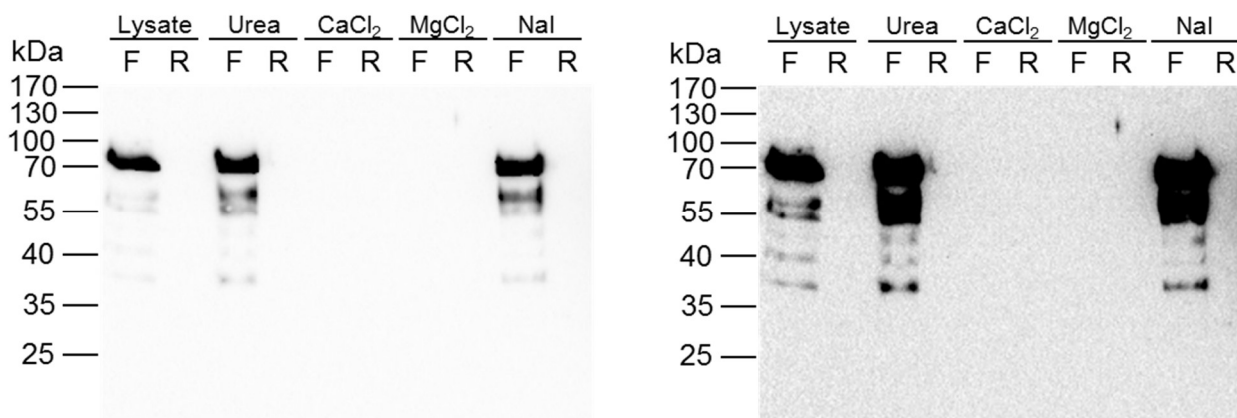


Figure 32. Second iteration of the additive screen, Part 2. The first screening showed that His₆-SUMO-XmLARP6 stability was improved by osmolytes (arginine, glycerol, and TMAO) and the chaotrope urea. The second screening consisted of testing different osmolytes and chaotropes. For the chaotropes, the assay called for 0.5 M urea, 0.1 M CaCl₂, 0.1 M MgCl₂, and 0.2 M NaI. Screening was performed as described on the first round but using the 0.1 μ m PVDF filtration units. The flowthrough was recovered as the filtrate ("F"). The aggregated retentate ("R") was re-suspended from the filter surface with 5X SDS-PAGE sample buffer. Samples containing CaCl₂, MgCl₂ when mixing with supernatant generated a cloudy and precipitation occurred. Urea and NaI were selected for further screening.

The last round of screens was carried out to determine both the optimal concentrations of the selected osmolytes and chaotropes as well as the possible benefit to combining them in one buffer. Figure 33 displays the solubility and stability effect when mixing different concentrations of the optimal osmolytes (glycerol and glucose) with varying concentrations of the two best chaotropes (sodium iodide and urea). Based on the overexposed western blots, 5% glycerol appears to stabilize truncated products, with the overall effect to increase protein solubility. As the glycerol concentration was increased to 10% and 15%, the amount of aggregates in the flowthrough decreased as compared to the cell lysate. The contribution of sodium iodide and urea to the 15% glycerol is mixed, as some concentrations of the chaotropes are beneficial while others are not. For example: 0.5 M urea seems to be increasing full-length and degradation products

compared to the 0.2 M concentration. However, there are some aggregated products on the retentate that are lacking on the lower urea retentate. In addition, an intermediate sodium iodide concentration resembles the same outcome. 0.1 and 0.3 M sodium iodide lack any denature products as compared to the 0.2 M aggregated products. Overall, the 5% glycerol with either 0.1 M or 0.3 M sodium iodide seem to yield the best improvement in protein solubility.

Different glucose concentrations were also tested against urea and NaI (Figure 34). Protein solubility appears to be improved with glucose alone. This can be point out by comparing the cell lysate full-length protein to degradation ratio and the visible aggregated protein on the retentate to the three glucose concentrations in question. Furthermore, 0.5 M glucose showed the best full-length to degradation ratio and partition between the filtrate and retentate between the three concentrations. As for the contribution gained on solubility and stability by the addition of urea to the 0.5 M glucose, it is prominent that higher concentrations increase protein stability. However, it also increases the amount of aggregated protein as can be seen on the retentate. For sodium iodide, the one that overall helped the most was 0.3 M.

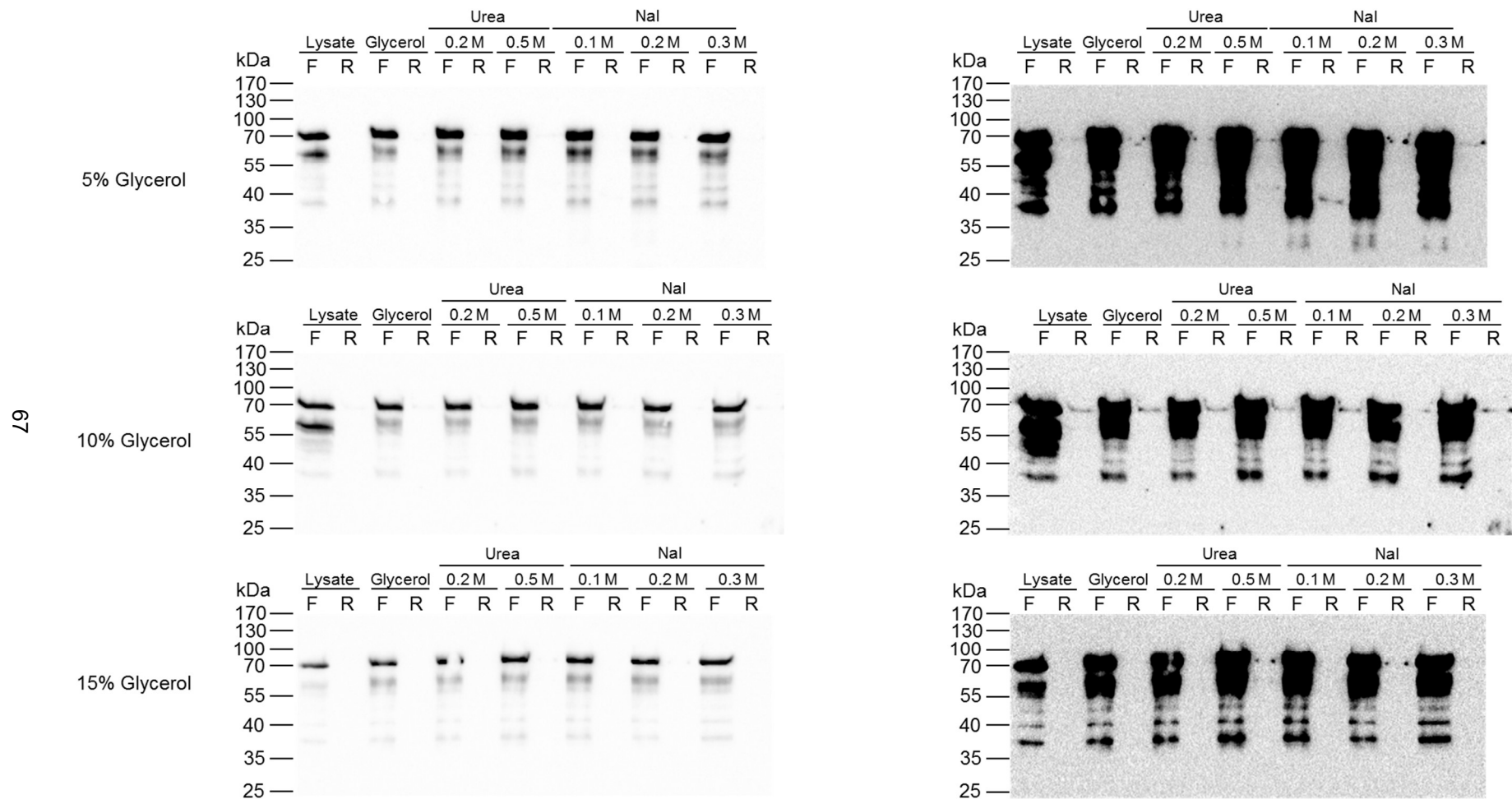


Figure 33. Effect of glycerol in combination with urea and NaI. The osmolyte glycerol was evaluated in different concentration with the chaotropes urea and NaI at different concentrations to identify optimal conditions that promoted solubility of His₆-SUMO-XmLARP6. Screening was performed as described before. The flowthrough was recovered as the filtrate ("F"). The aggregated retentate ("R") was re-suspended from the filter surface with 5X SDS-PAGE sample buffer. 5% glycerol increased stability of truncated products more compared to full length. 10% glycerol showed more insoluble protein between the tested glycerol concentrations. From the combinations, 5% glycerol and 0.2M NaI were considered as possible additives.

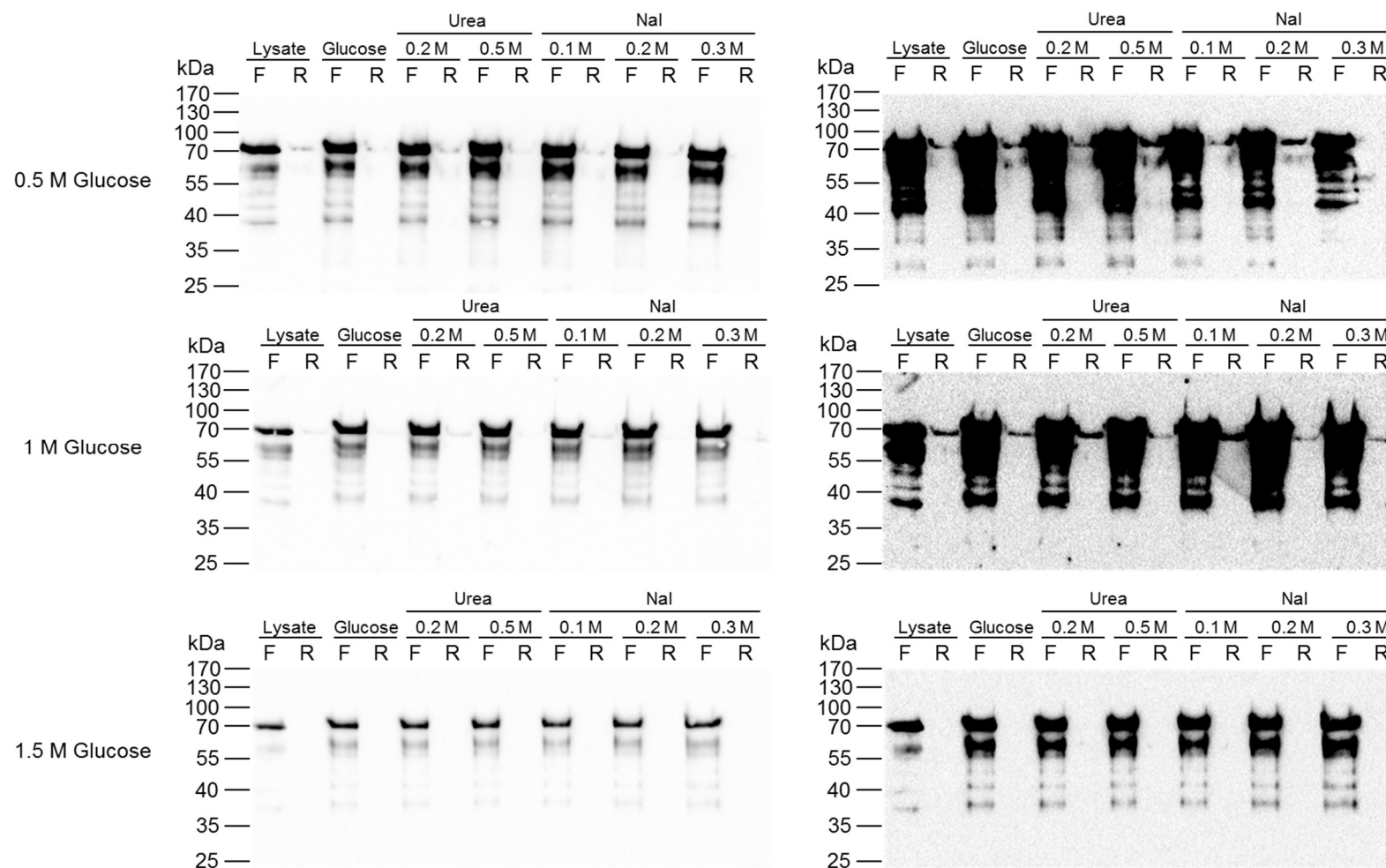


Figure 34. Effect of glucose in combination with urea and NaI. The osmolytes glucose in different concentration was evaluated with the chaotropes urea and NaI at different concentrations to identify optimal conditions that promoted solubility of His₆-SUMO-XmLARP6. Screening was performed as described before. The flowthrough was recovered as the filtrate ("F"). The aggregated retentate ("R") was re-suspended from the filter surface with 5X SDS-PAGE sample buffer. All glucose concentrations evaluated showed improvement on full length solubility. Compared to the glycerol trial, it seems that NaI in higher concentration stabilizes full length. 0.5 M glucose was selected to be tested out in full protein purification process.

The combination of 0.5 M glucose and 0.3 M NaI increased solubility of both full-length and shorter fragments while minimizing the amount of aggregated protein in the retentate. We decided to use this additives at the mentioned concentrations with the standard buffer to proceed with the purification of His₆-SUMO-XmLARP6.

His₆-SUMO-XmLARP6 purification with optimized buffer

To verify that the optimized buffer conditions were adequate for His₆-SUMO-XmLARP6 purification, Rosetta cell pellets expressing desired protein were lysed by sonication in the optimized buffer, 50 mM NaH₂PO₄/ Na₂HPO₄ [pH 8.0], 200 mM NaCl, 0.5 M glucose, and 300 mM NaI, and, subjected to affinity purification, and analyzed by gel electrophoresis.

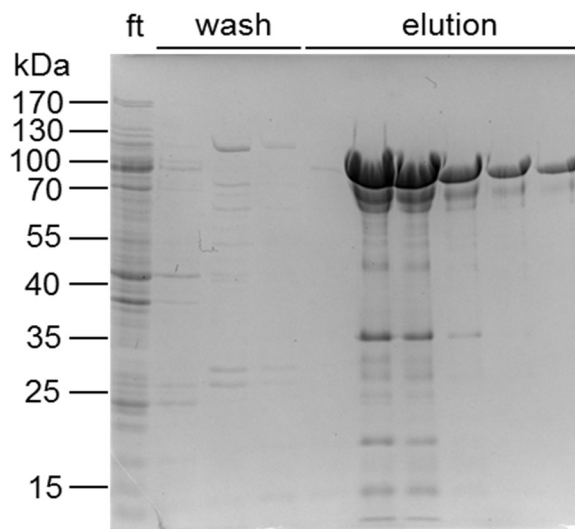


Figure 35. His₆-SUMO-XmLARP6 affinity chromatography fractions using optimized buffer conditions. Cells expressing of His₆-SUMO-XmLARP6 were resuspended with optimized buffer from screenings (50 mM NaH₂PO₄/ Na₂HPO₄ [pH 8.0], 200 mM NaCl, 0.5 M glucose, 300 mM NaI, 10 mM imidazole [pH 8.0]) sonicated and centrifuged. The cell lysate was purified using nickel-affinity chromatography (“p”, pellet, “ft”, flowthrough). Aliquots were separated by gel electrophoresis and stained with Coomassie blue. Elution fractions 2—6 were pooled for further purification. Comparing the elution profile to the one with previous buffer, purification yield increased significantly.

The elution fractions showed the expected band at ~60 kDa, consistent with previous results (Figure 35). There is no obvious increase in overall protein yield between un-optimized (Figure 27) and optimized conditions (Figure 35). As before, elution fractions 2—6, were collected and prepared for size exclusion chromatography.

When the concentrated His₆-SUMO-*XmLARP6* with new buffer was analyzed and separated by the SEC, a very prominent peak is seen at an elution volume of 66.6 mL, with a small preceding shoulder at 53.5 mL (Figure 36). Based on the standards, those peaks display an apparent molecular weight of 560.8 and 253.64 kDa respectively. The elution profile indicates that His₆-SUMO-*XmLARP6* has a bigger hydrodynamic radius than expected. It could be that it displays non-globular features or intermolecular interactions generate tetramers while eluting. By comparing the old buffer condition chromatograph (orange dash) to the new one (blue solid), it is very clear that the screen identified additives that improved His₆-SUMO-*XmLARP6* solubility and stability. As described previously for the human protein, the elution fractions were analyzed by denaturing gel electrophoresis and Coomassie blue staining (Figure 37). The concentrated column load sample (“C”) displays a much higher amount of protein than the previous one under the initial buffer conditions (compare Figure 29 to Figure 37). In addition, the elution fractions from the new buffer conditions have notably more protein content than the initial SEC fractions (compare Figure 28 to Figure 36).

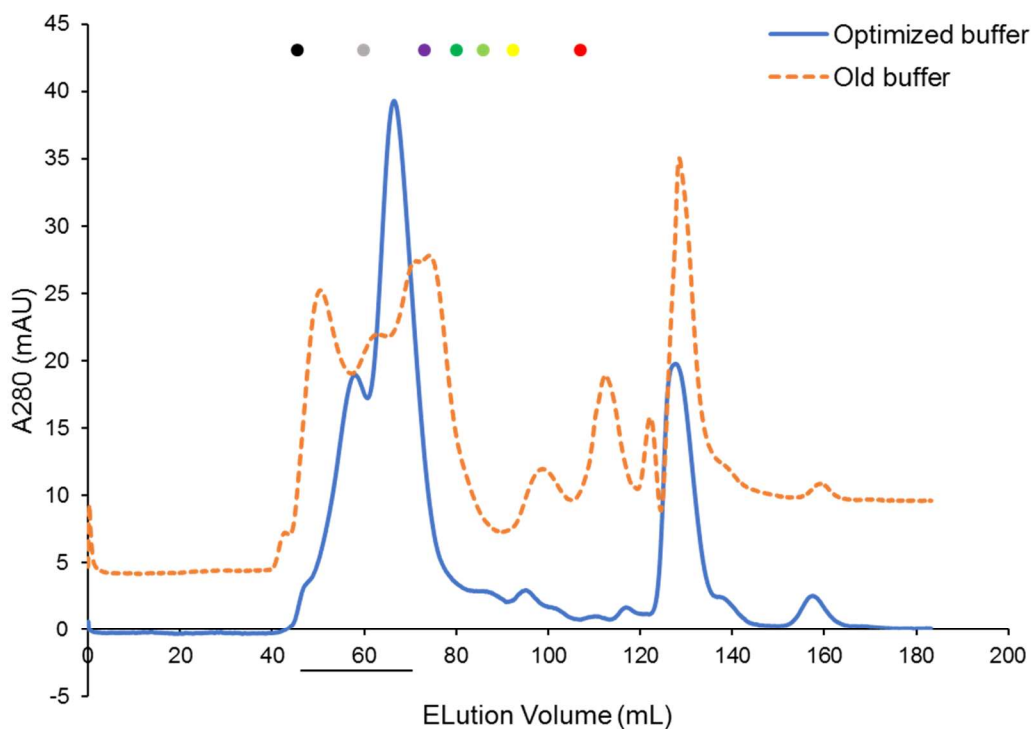


Figure 36. S200 size exclusion chromatogram of His₆-SUMO-XmLARP6, with optimized buffer conditions. Elution fractions selected from affinity chromatography were treated as before. Filtered protein was then loaded to the fPLC in which was monitored by UV absorbance at 280nm at a 1mL/min flowrate as before. Fractions were collected in 2 mL aliquots automatically. Comparing the previous purification scheme (orange dash) to the one with the new buffer (blue solid), it is noticeable that the additive stabilized His₆-SUMO-XmLARP6 giving a higher yield.

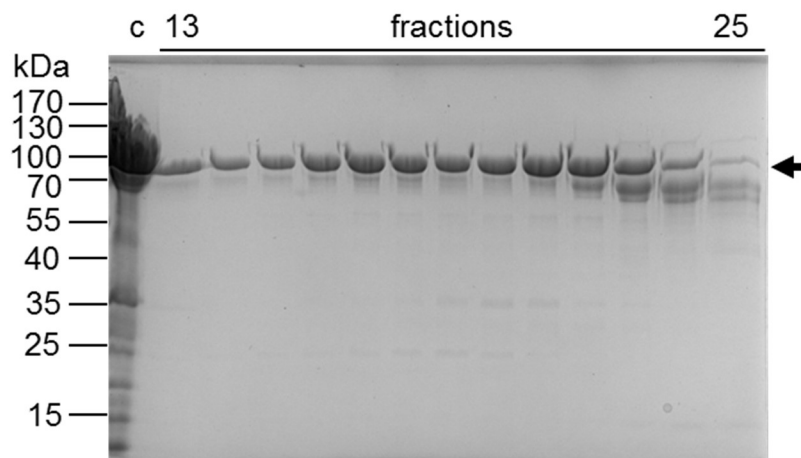


Figure 37. SDS-PAGE analysis of fractions from size exclusion chromatography of His₆-SUMO-XmLARP6 with optimized buffer. The pooled elution from affinity chromatography was filtered through a 0.2 μ m filter (lane "C") and separated on an S200 Sephadex column. Fractions 13—25 were analyzed by SDS-PAGE and Coomassie blue staining. New buffer conditions yield very robust amount of protein compared to old buffer used. Additives screened for helped with protein stability and solubility. Fractions 16—22, corresponding to elution volumes 55—68 mL, were pooled out. Fractions were stored in 50 μ L aliquots at -70 $^{\circ}$ C for further experiments.

Ulp1 cleavage and purification of XmLARP6

After the His₆-SUMO-XmLARP6 fractions were collected from the FPLC SEC, they were pooled in a conical vial. The appropriate amount of Ulp1 was added to an approximate molar ratio of 1000:1 LARP6:Ulp1, and incubated at 16 °C for 2 hours. It was expected that by cleaving the SUMO tag and then running the reaction products over the nickel column, the cleaved proteins will elute in the flowthrough and the His₆-SUMO portion will be left behind interacting with the beads. However, that result was not achieved, as shown in the SDS-PAGE analysis of the affinity chromatography (Figure 38).

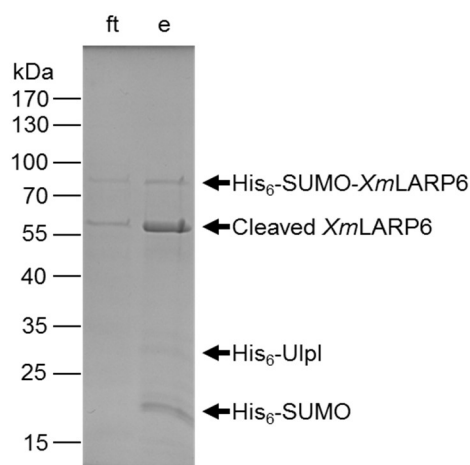


Figure 38. Ulp1 cleavage of His₆-SUMO-XmLARP6. The His₆-SUMO-tagged LARP6 was eluted from the SEC and pooled. The pooled protein was incubated with 1000:1 ratio of XmLARP6:His₆-Ulp1 for 2 hours at 16 °C. Affinity chromatography was performed to separate cleaved protein from the fractions containing the Ulp1 and SUMO proteins, which are both histidine-tagged (“ft”, flowthrough, “e”, elution). The elution buffer contained 300 mM imidazole. Samples were subjected to denaturing gel electrophoresis and Coomassie blue stain. Reaction did not go to completion since uncleaved protein (~70 kDa) is present. In addition, cleaved product (~60 kDa) interact with the nickel beads. Therefore, the flowthrough and elution were used for further purification.

Even though there was some *XmLARP6* containing the SUMO tag (~70 kDa), a large amount of the cleaved product (~60 kDa) interacted with the nickel beads. Since a 300 mM imidazole concentration was used to elute the protein that was left behind interacting with the column, His₆-SUMO, (~20 kDa) and His₆-UlpI (~30 kDa) co-eluted with the cleaved *XmLARP6*. The SEC chromatogram of *XmLARP6* displays all the species found on the SDS-PAGE (Figure 39) prior to the run. However, since un-cleaved His₆-SUMO-*XmLARP6* was still present as a co-elution with cleaved *XmLARP6*, fractions under the highest peaks were analyzed by gel electrophoresis (Figure 40). The SEC load sample (lane "C") demonstrates the difference in species concentration between cleaved and un-cleaved *XmLARP6*. Despite the small amount of un-cleaved *XmLARP6* elution overlaps with cleaved protein, UlpI and SUMO co-elution products acquired from the affinity chromatography, were separated based on size thanks to the SEC. Fractions 18–24, corresponding to elution volumes 58–72 mL, were collected, concentrated, and stored at -70 °C for further characterization.

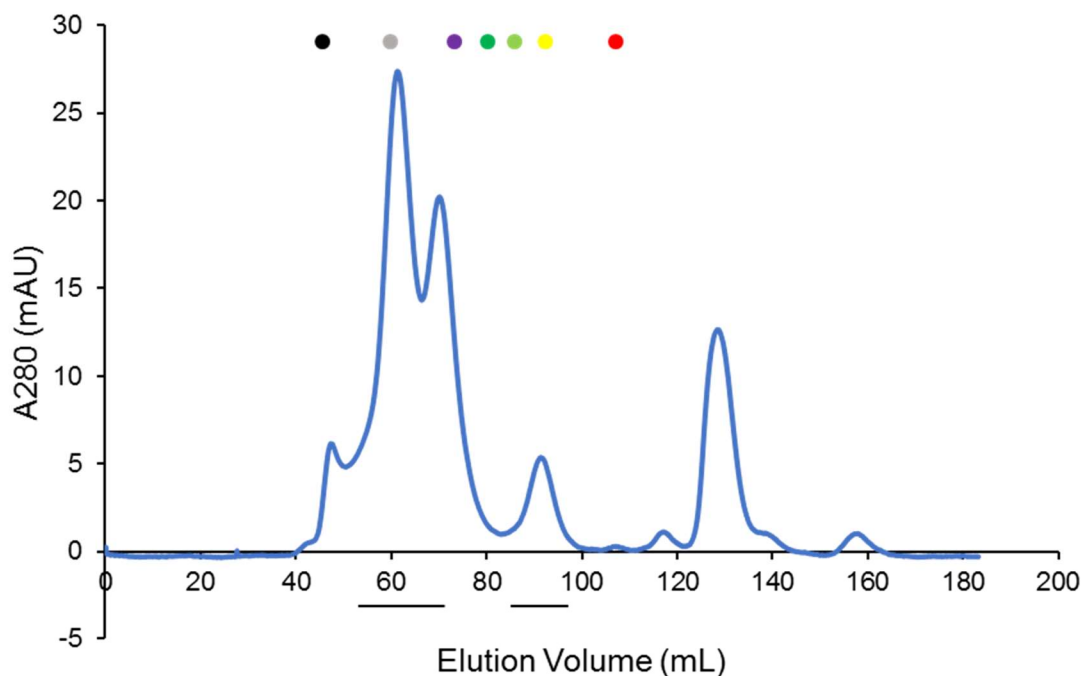


Figure 39. S200 size exclusion chromatogram of *XmLARP6* after SUMO cleavage. The flowthrough and elution from the Ulp1 digest affinity chromatography were concentrated to 2—2.5 mL against a molecular weight cutoff (MWCO) of 10 kDa by centrifugation at 5,000 x g, and filtered with a 0.2 μ m filter as before. Protein was susceptible to same procedure described for His₆-SUMO-*XmLARP6* and monitored with same parameters.

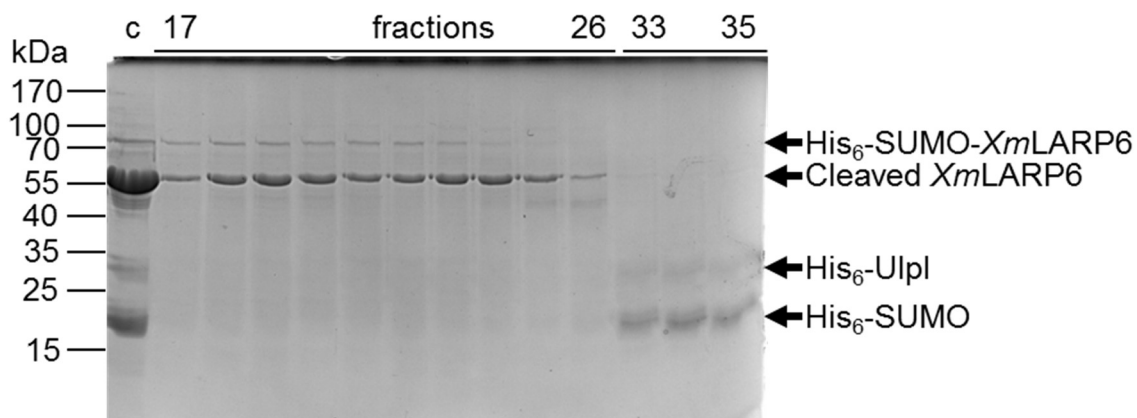


Figure 40. SDS-PAGE analysis of fractions from size exclusion chromatography of *XmLARP6* post-SUMO cleavage. The flowthrough and elution from Ulp1 digest were through a 0.2 μ m filter (lane "C") and separated on an S200 Sephadex column. Fractions 17-26 and 33-35 were analyzed by SDS-PAGE and Coomassie blue staining. Between the two peaks all 4 products can be seen. The characteristic His₆-SUMO-*XmLARP6* band at ~70 kDa, *XmLARP6* cleaved product (~55 kDa), the His₆-SUMO product (~20 kDa) and His₆-Ulp1 (~30 kDa). Fractions 18—24, corresponding to elution volumes 58—72 mL, were pooled out. Fractions were then concentrated to 2—3 mL as before. Protein was then aliquoted into 50 μ L fractions and stored at -70 °C for future experiments.

His₆-SUMO-DrLARP6a purification with optimized buffer

Since the additive screening showed to extremely increase the yield of SUMO tagged *Xm*LARP6, the same buffer conditions obtained were used to purify His₆-SUMO-*Dr*LARP6a. Rosetta cell pellets expressing SUMO tagged *Dr*LARP6a susceptible to same treatment for affinity chromatography and analyzed by SDS-PAGE Coomassie stained as previously done (Figure 41). The flowthrough sample shows a great amount of protein around the same molecular weight that elution fractions display for SUMO-tagged *Dr*LARP6a protein. Probably, there was a lot of protein for the nickel beads to bind to. When analyzed by SDS-PAGE, the nickel elution fractions showed a strong band at ~75 kDa. This is somewhat higher than the expected molecular weight of the SUMO-tagged construct (~67 kDa).

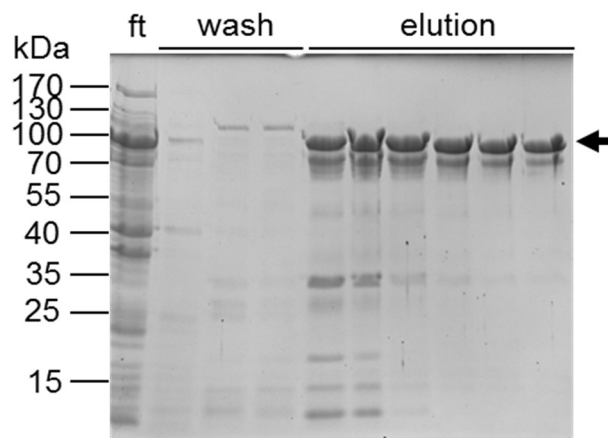


Figure 41. His₆-SUMO-*Dr*LARP6a affinity chromatography fractions using optimized buffer conditions. Decided to test the new buffer with zebrafish LARP6a. Cells expressing His₆-SUMO-*Dr*LARP6a were treated the same as *Xm*. The cell lysate was purified using nickel-affinity chromatography ("ft", flowthrough). Aliquots were separated by gel electrophoresis and stained with Coomassie blue. All elution fractions contained the band of interest (~75 kDa) characteristic of the fish LARP6 constructs. In addition, other protein species co-eluted from the nickel bead. These other proteins could be other *E. coli* proteins that are able to interact with nickel. The six elution fractions (arrow) were pooled for further purification.

Additionally, there are other smaller species that co-elute with the full-length protein.

These could be degradation products from the protein of interest, or other *E. coli* proteins that could be interacting with the nickel beads comparable to those seen during the purification of *XmLARP6*. All elution fractions were prepared for size exclusion chromatography by concentration and filtration. Figure 42 is the chromatogram obtained by the FPLC from the concentrated SUMO-tagged *DrLARP6a* and the degradation products (analyzed as before). The profile shows a broad peak that starts next to what is considered the column void volume.

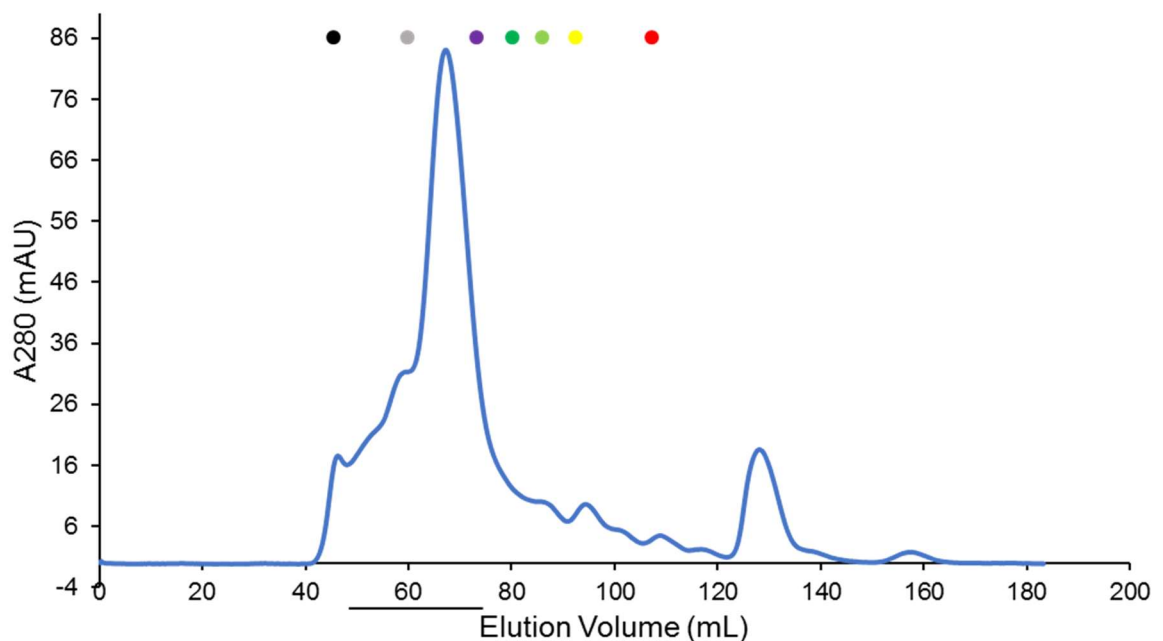


Figure 42. S200 size exclusion chromatogram of His₆-SUMO-*DrLARP6a* with optimized buffer conditions. Elution fractions selected from affinity chromatography were treated as before. Protein was susceptible to same procedure described for His₆-SUMO-*XmLARP6* and monitored with same parameters. A left-skewed peak is observed. Based on this peak elution volume relative to the standards, it is concluded that some protein aggregated. In addition, the protein is showing an apparent higher molecular weight compared to the expected one.

Based on the presence of other protein species, it does not necessarily mean that some His₆-SUMO-DrLARP6a aggregated, yet, it does not discard the possibility that some of the protein aggregated.

However, most of it eluted, as can be seen by the peak. The protein is eluting at a lower volume with an apparent molecular weight of 471.33 kDa compared to the expected 67 kDa. Based on the standards, it was expected to elute between 80—86 mL. Fractions within the peak were analyzed by SDS-PAGE and Coomassie stained to decide which ones to collect for continued purification (Figure 43).

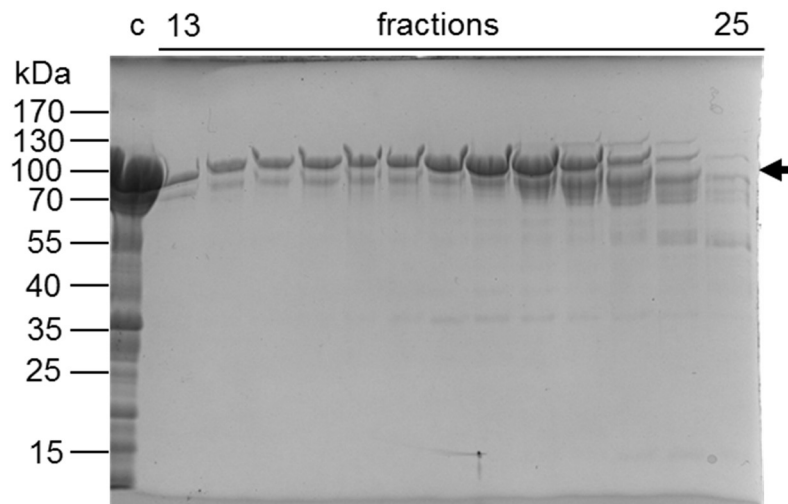


Figure 43. SDS-PAGE analysis of fractions from size exclusion chromatography of His₆-SUMO-DrLARP6a with optimized buffer. The pooled elution from affinity chromatography was filtered through a 0.2 µm filter (lane “C”) and separated on an S200 Sephadex column. Fractions 13—25 were analyzed by SDS-PAGE and Coomassie blue staining. Again, the characteristic band at ~70 kDa can be seen throughout the elution profile. Fractions 20—22, corresponding to elution volumes 65–71 mL, were pooled. Fractions were stored in 50 µL aliquots at -70°C for further experiments.

The gel analysis shows that the concentrated load sample contains a similar amount of His₆-SUMO-DrLARP6 protein to the amount of SUMO-tagged

XmLARP6. Based on the gel, the entire peak is composed of His₆-SUMO-*DrLARP6a*. However, the majority of the protein that eluted at later volumes, indicated by the band intensity increases from left to right. The earlier elution samples confirmed that some protein aggregated. Higher eluting volume fractions displayed an increase of protein concentration and some protein degradation as protein eluted from column. Fractions 20—22 were pooled, and stored as previously described.

Ulp1 cleavage and purification of DrLARP6a

The Ulp1 cleavage reaction of the His₆-SUMO-*DrLARP6a* went to completion, unlike the *XmLARP6* cleavage reaction. However, as seen for *XmLARP6*, a significant portion of the cleaved product (~60kDa) still bound to the nickel beads during affinity purification (Figure 44). The flowthrough and the elution were pooled and subjected to size exclusion chromatography.

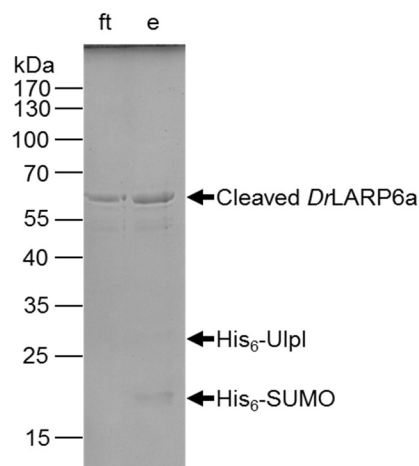


Figure 44. Ulp1 cleavage of His₆-SUMO-*DrLARP6a* after size exclusion chromatography. Once His₆-SUMO tagged fish proteins eluted fractions from SEC were pooled out, they were collected in conical vials and incubated with 1000:1 ratio of *DrLARP6a*:His₆-Ulp1. Reactions were incubated at 16 °C for 2 hours. Then, affinity chromatography was performed to separate cleaved protein from the fractions containing the histidine track ("ft", flowthrough, "e", elution). The 300 mM imidazole buffer was used for the elution. Samples were susceptible to gel electrophoresis and Coomassie blue stain. In both cases the flowthrough and elution were used for further purification.

The elution profile of *Dr*LARP6a from the FPLC SEC presents a distinctive peak related to the single cleaved species seen on the SDS-PAGE (Figure 45). Fractions were analyzed by SDS-PAGE and Coomassie blue staining as before (Figure 46). A higher band can be seen from the control prior to SEC presumably SUMO tagged *Dr*LARP6a. In contrast to *Xm*LARP6, less tagged *Dr*LARP6 species is visible. Fractions were pooled and treated as before and stored for future studies.

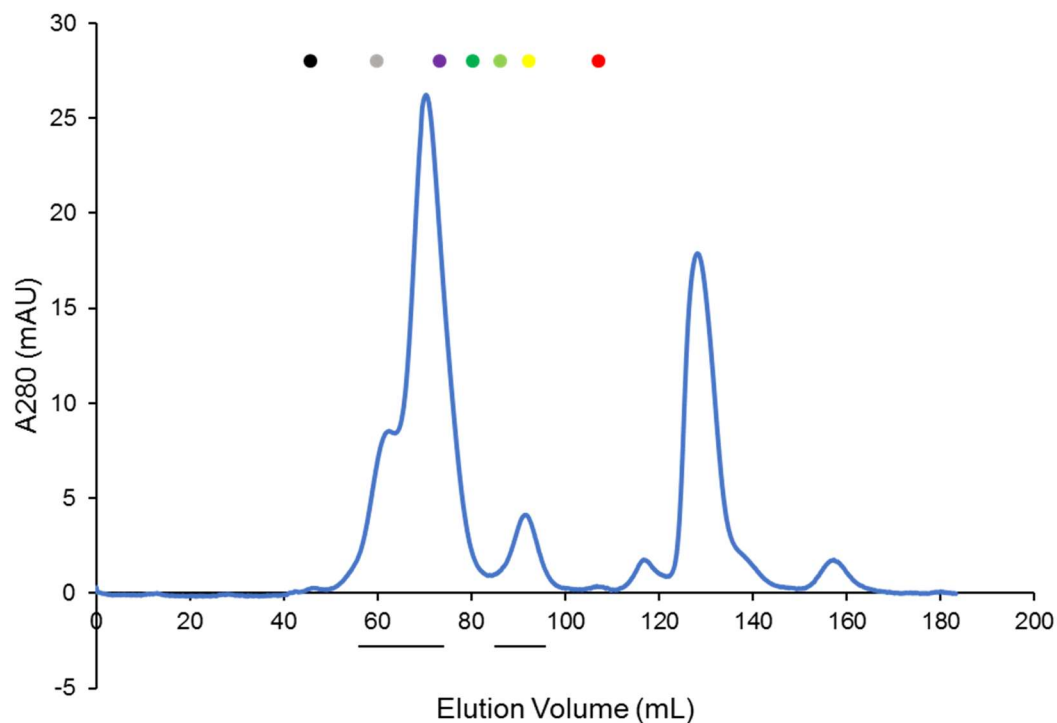


Figure 45. S200 size exclusion chromatogram of *Dr*LARP6a after SUMO cleavage. The flowthrough and elution from the Ulpl digest affinity chromatography were concentrated to 2—2.5 mL against a molecular weight cutoff (MWCO) of 10 kDa by centrifugation at 5,000 xg, and filtered with a 0.2 μ m filter as before. Protein was susceptible to same procedure described for His₆-SUMO-*Xm*LARP6 and monitored with same parameters.

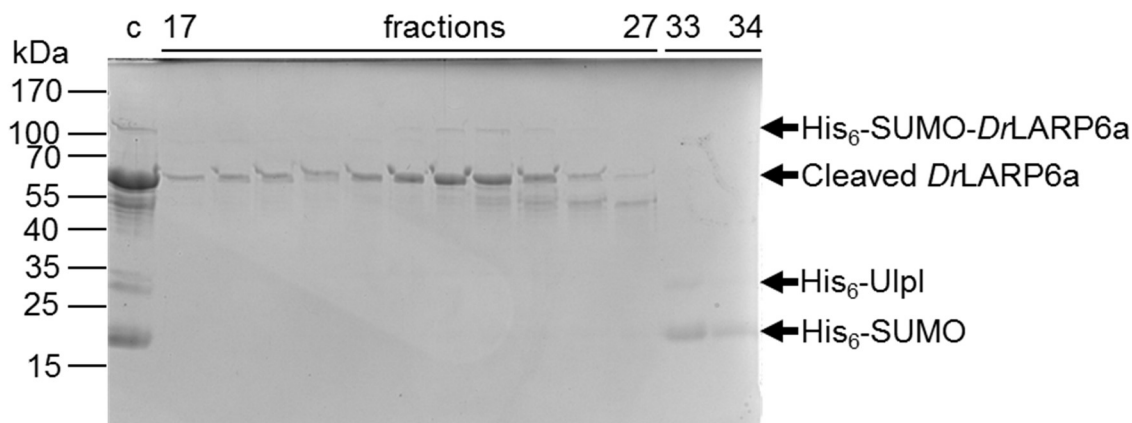


Figure 46. SDS-PAGE analysis of fractions from size exclusion chromatography of *DrLARP6a* post SUMO cleavage. The flowthrough and elution from Ulp1 digest were through a 0.2 μ m filter (lane “C”) and separated on an S200 Sephadex column. Fractions 17—27 and 33—34 were analyzed by SDS-PAGE and Coomassie blue staining. Fractions 22—24, corresponding to elution volumes 65—70 mL, were pooled out. Fractions were then concentrated to 2—3 mL as before. Protein was then aliquoted into 50 μ L fractions and stored at -70 °C for future experiments

Purification of Xiphophorus maculatus La Module

Previous work analyzing the structure and function of LARP6 has been carried out using the RNA binding domain, also called the La Module. The La motif and RRM of human LARP6 are known to have a folded conformation and RNA binding activity.²¹ Therefore, we also recombinantly expressed and purified the La Module of *XmLARP6* for biochemical analysis.

The domain boundaries of the *X. maculatus* LARP6 La Module (“*XmLARP6*-LAM”) were determined by multiple sequence alignment of the fish and human proteins (Figure 9). The La Module coding sequence was amplified by PCR, using the full-length *XmLARP6* coding sequence as the template, and cloned into the BamHI/XhoI sites of pET28 (Table 3). Protein expression was carried out essentially as described above for the full-length protein.

XmLARP6-LAM was purified using the same buffer conditions used for the fish full-length proteins. Rosetta cell pellets expressing His₆-*XmLARP6*-LAM were lysed by sonication and subjected to affinity purification as described above for the full-length protein (Figure 47). The expected molecular weight of the construct is ~35—38 kDa. There is also a lower-intensity band of a slightly higher molecular weight that co-elutes with *XmLARP6*-LAM. Elution fractions 1—3, which contained the largest amount of protein, were collected and prepared for size exclusion chromatography as previously described for the full-length proteins.

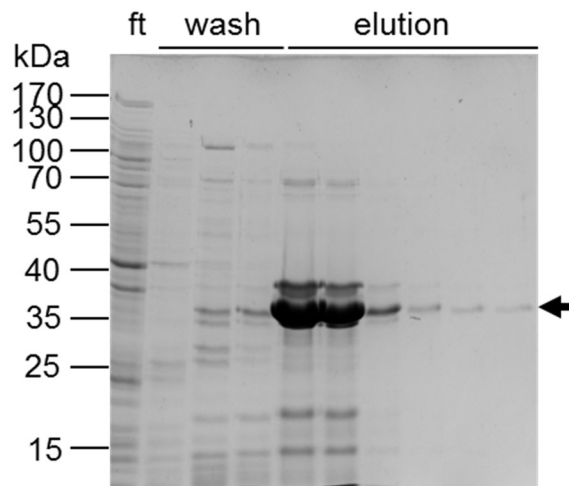


Figure 47. His₆-*XmLARP6* La Module affinity chromatography fractions. Cells expressing His₆-*XmLARP6* La Module (LAM) were resuspended and treated the same as Full-length His₆-SUMO-*XmLARP6*. The cell lysate was purified using nickel-affinity chromatography (“p”, pellet, “ft”, flowthrough). Aliquots were separated by gel electrophoresis and stained with Coomassie blue. Elution fractions 1—3 (arrow) were pooled for further purification.

Affinity-pure His₆-*XmLARP6*-LAM was then purified by size exclusion chromatography (Figure 48). The chromatogram obtained from the SEC displayed one strong, symmetrical peak at 80 mL, suggesting that only one species of His₆-*XmLARP6*-LAM eluted from the column. However, the apparent

molecular weight in which the construct is eluting is 63.8 kDa, almost double the expected size.

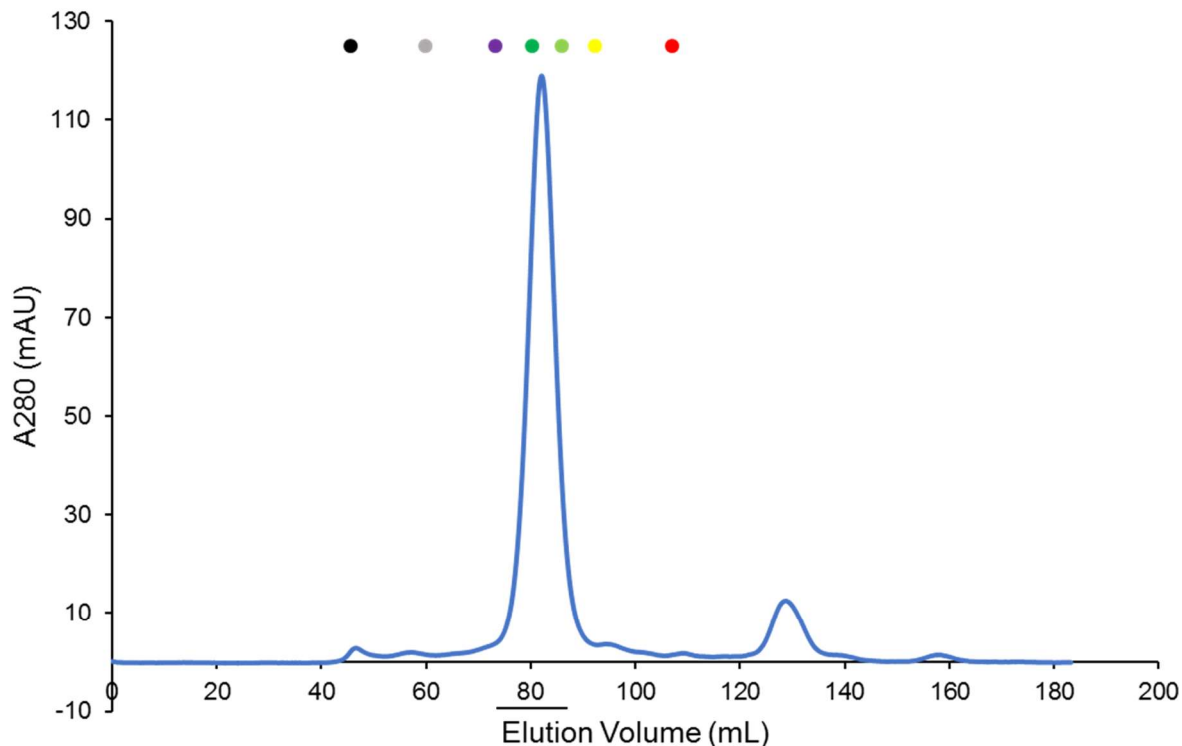


Figure 48. S200 size exclusion chromatogram of His₆-XmLARP6-LAM. Elution fractions selected from affinity chromatography were concentrated to 2—2.5 mL against a molecular weight cutoff (MWCO) of 10 kDa by centrifugation at 5,000 xg, and filtered with a 0.2 µm filter. Filtered His₆-XmLARP6-LAM was then loaded to the FPLC in which was monitored by UV absorbance at 280 nm at a 1mL/min flowrate. Fractions were collected in 2 mL aliquots automatically.

As before, fractions were analyzed by SDS-PAGE and Coomassie stain (Figure 49). Comparing the elution fractions to the concentrated load sample (lane “C”), it was clear that the higher molecular weight band was still present after size exclusion separation. As seen in the previous S200 chromatograms, a smaller peak around 130 mL elution was not analyzed by SDS-PAGE. This peak is predicted to contain imidazole, based on previous purifications showed no protein by SDS-PAGE and because the peak is eluting at a volume larger than

the column volume (120 mL), at which point the volume of the column would have completely exchanged. Fractions 28-31, which contained the largest amount protein as determined by band intensity, were pooled, concentrated, and stored in aliquots at -70 °C.

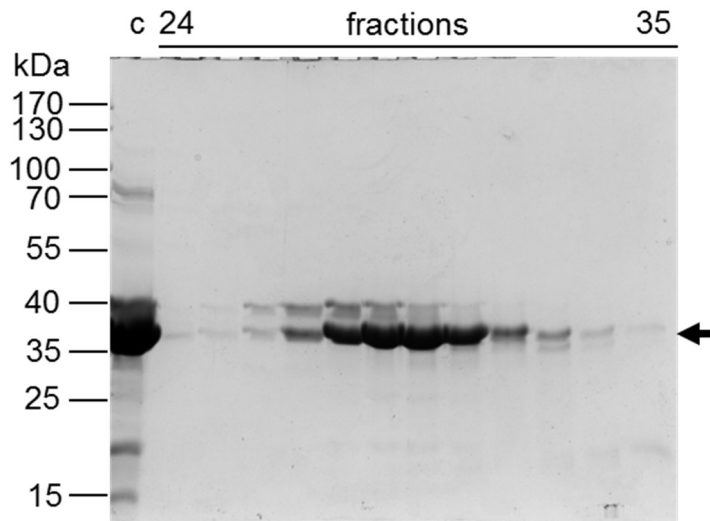


Figure 49. SDS-PAGE analysis of fractions from size exclusion chromatography of His₆-XmLARP6-LAM. The pooled elution from affinity chromatography was filtered through a 0.2 µm filter (lane “C”) and separated on an S200 Sephadex column. Fractions 24—35 were analyzed by SDS-PAGE and Coomassie blue staining. Fractions 28—31, corresponding to elution volumes 78—83 mL, contained full-length His₆-XmLARP6-LAM protein. Fractions were stored in 50 µL aliquots at -70 °C for further experiments.

Limited proteolysis to probe protein topology

Since there is no crystal structure of full-length human LARP6, limited proteolysis was performed to compare the fish proteins to human LARP6 by gross topology. Temperature studies were first performed to identify the ideal conditions under which to obtain the most useful results. Human LARP6 was subjected to trypsin digest and monitored over time. Samples were analyzed by denaturing gel electrophoresis and silver stain (Figure 50).

The same reaction was monitored at 4 °C, 16 °C, room temperature (~22 °C), and 37 °C. At 37 °C, it can be seen that *HsLARP6* digest occurs relatively fast. There are some domains that can be seen through the timecourse. However, protein stability due to heat could unfold LARP6 allowing trypsin to digest it faster. At room temperature, trypsin enzymatic activity seems to be at its highest possible performance according to the digestion profile displayed. However, the increased degradation with temperature is likely the result of thermal protein unfolding or flexibility, which allowed trypsin to access more cleavage sites in human LARP6. At 16 °C, it was notable that trypsin digested *HsLARP6*. However, the fragments generated seem more stable than at room temperature and 37 °C. Furthermore, the 4 °C reactions gave the same digest profile that 16 °C, however, the domains generated by trypsin digest showed more stability.

At all temperatures, a ~40 kDa band appears (marked by an asterisk in Figure 50). However, the effect of temperature compared to the rate of appearance this species to appear varies. On the 4 °C, the band appears after 2 minutes and is maintained over the whole timecourse. In the 16 °C and room temperature reactions, the same complex takes a little longer to appear (after 5—10 minutes) and displays lower stability (in other words, it degrades more quickly over time). In contrast, the 37 °C reaction shows significantly less stability overall, and a noticeable reduction in the production and retention of the ~40 kDa species. Although the 4 °C and 16 °C reactions generally display the same species, the 4 °C reaction displays a more stable digestion profile based on band

stability over time and overall number of digestion products. Therefore, the 4 °C temperature was selected to carry out the trypsinolysis analysis of the fish LARP6 proteins for comparison to the human LARP6.

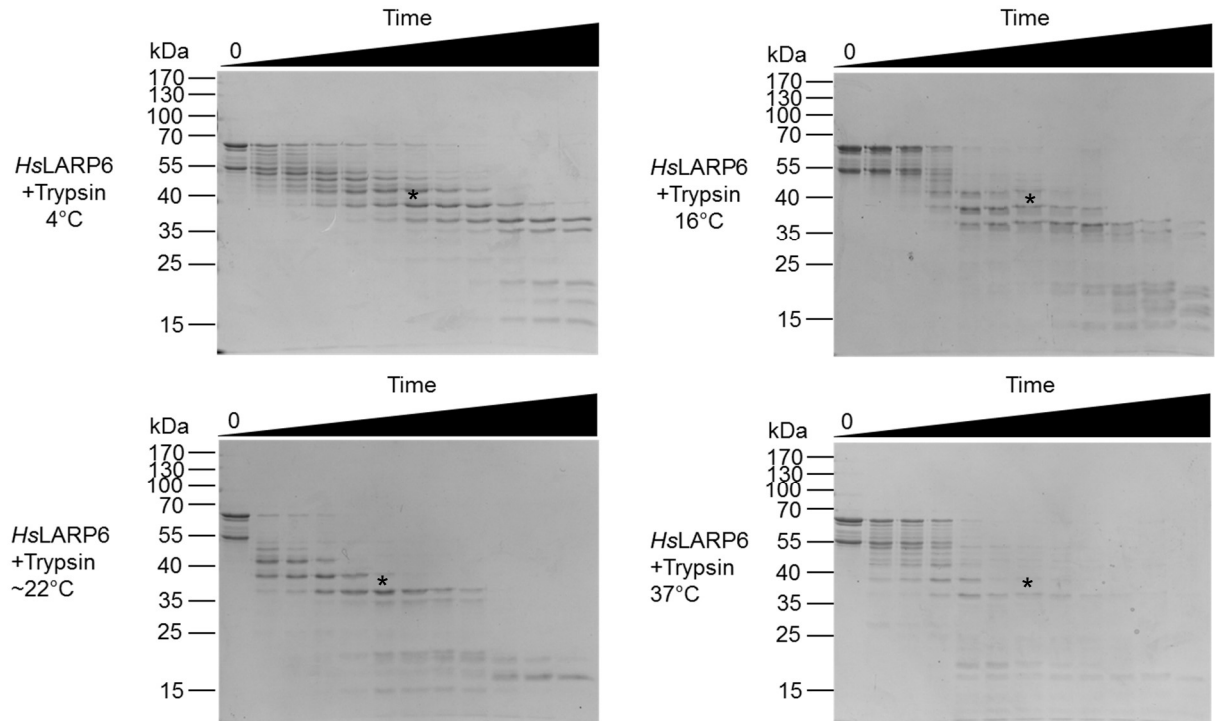


Figure 50. Temperature trials of limited proteolysis of HsLARP6. Trypsinolysis was used to compare the domain topology and stability of human LARP6 at four temperatures. After removing a sample for a zero timepoint, trypsin was added at a mass ratio of 60:1 (protein to protease) and the reaction incubated at the indicated temperatures. Aliquots were removed at 1, 2, 5, 10, 20, 30, 45, 60, 120, 180, and 240 min. Samples were analyzed by denaturing gel electrophoresis and silver stain. The 4 °C was selected for future studies because it showed the most detailed and persistent digestion pattern.

The optimized purification buffer used for the fish LARP6 proteins contains large concentrations of stabilizing salts and osmolytes, which were selected for contributions to protein structure or stability. However, the human LARP6 protein was purified in the absence of these salts. Therefore, to control for any buffer effects on structure and stability, the human and fish LARP6 proteins were diluted

into a standard buffer to normalize all buffer components (150 mM NaI, 250 mM glucose, and 250 mM NaCl). In addition to denaturing gel electrophoresis and silver stain, the trypsinolysis samples were analyzed by Western blotting with an antibody that identifies a conserved sequence at the C-terminus of vertebrate LARP6 (Figure 51).

The silver stained gels show a domain of approximately 40 kDa that is very resistant to trypsin digest present on both fish proteins. A similarly-sized domain can be seen in the trypsinolysis of *HsLARP6*. However, the human domain displays a rate of formation lower than the fish. In addition, the domain seems to have less protein stability since it does not persist through the timecourse of the reaction as compared to the fish ~40 kDa domains. The anti-LARP6 antibody was used to determine which of these fragments retains the C-terminus of the protein. All three proteins show strong signal at the start of the reaction, but as trypsin digests the proteins over time, the signal from the antibody starts to dissipate. For *HsLARP6*, the signal for full-length protein lasts for almost one hour of the protein being exposed to the protease. In contrast, for both *XmLARP6* and *DrLARP6a*, the full-length signal is not visible after the first 10 to 20 minutes of the reaction, indicating that the C-terminus is rapidly exposed to proteolytic cleavage. Since the antibody does not recognize the stable ~ 40 kDa domain seen in the silver stained gels, it eliminates the possibility that the domain includes the C-terminus. More likely, the metastable domain is composed of amino acids that are located at the N-terminal or internal sections of the protein.

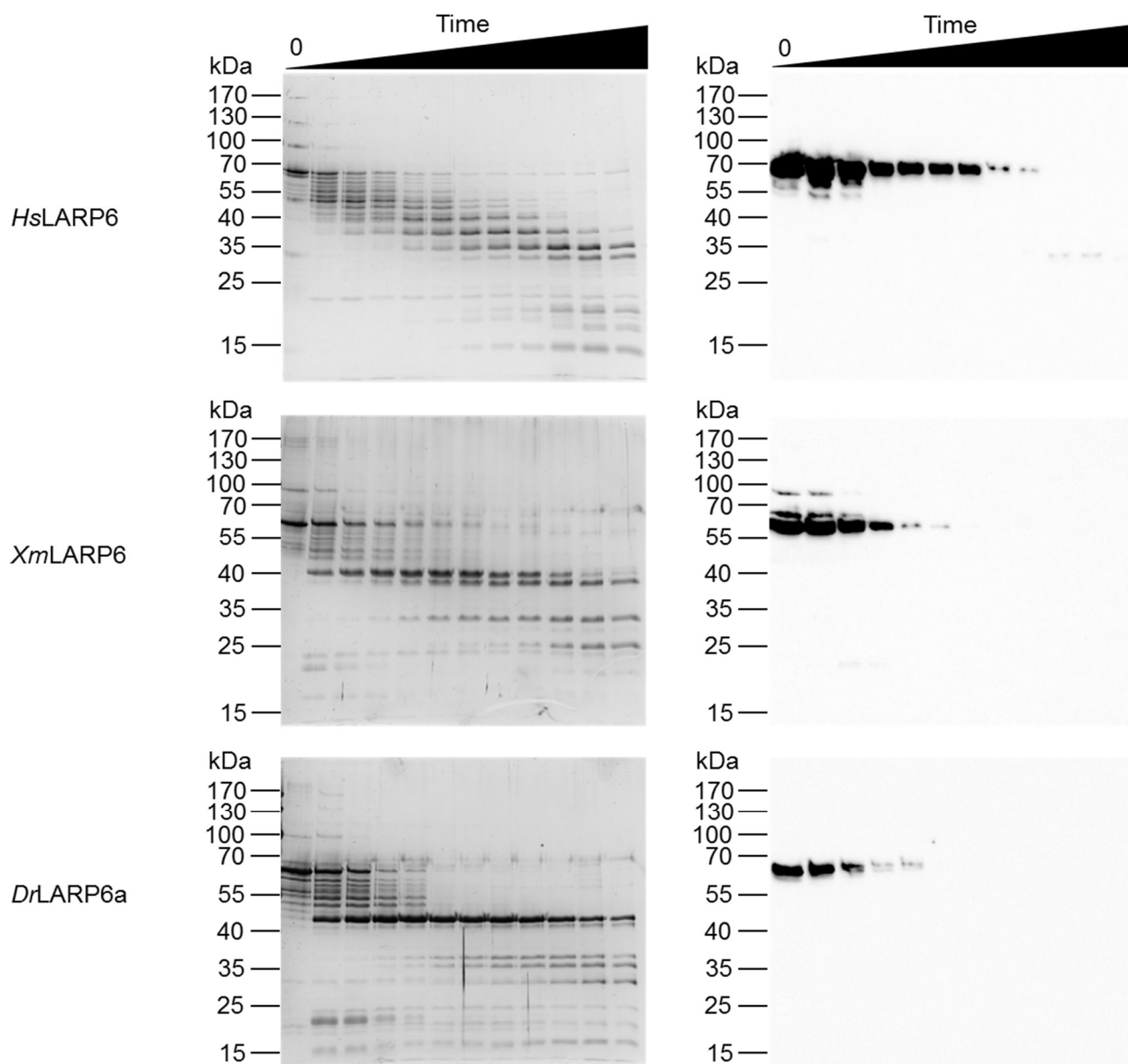


Figure 51. Topology studies of full-length LARP6 proteins by limited proteolysis. Limited trypsinolysis of the human (*Hs*), platyfish (*Xm*), and zebrafish (*Dr*) LARP6 proteins. All three LARP6 proteins were diluted into a standard reaction buffer so that all salts and additives were at identical concentrations (150 mM NaI, 250 mM glucose, and 250 mM NaCl). Proteins were incubated with a protein: trypsin ratio, and incubated on ice (4 °C). Aliquots were removed at 0, 1, 2, 5, 10, 20, 30, 45, 60, 120, 180, and 240 min. Samples were subjected to gel electrophoresis, and either silver stained (left) or western blotted with an antibody against the C-terminus of LARP6.

Trypsinolysis of His₆-XmLARP6-LAM

To determine the identity of the ~40 kDa band, His₆-XmLARP6-LAM was subjected to limited trypsin digestion under the same conditions done for full-length. Figure 52 shows the results of the limited trypsinolysis, run alongside the 120-minute digestion timepoint of full-length XmLARP6. The lane labeled “X” on the gel shows the stable 40 kDa band that was derived from *X. maculatus* full length LARP6. As the His₆-XmLARP6-LAM was subjected to trypsin digest, the *Xm* La module individual band produces two bands after 10 minutes. Additionally, the direct comparison of the 120-minute time point from the full-length digestion to the His₆-XmLARP6-LAM full length and trypsin digest products shows that the LAM has a lower molecular weight compared to the stable 40 kDa domain. This means that the stable band is composed of a longer sequence than just *X. maculatus* LARP6 La Module.

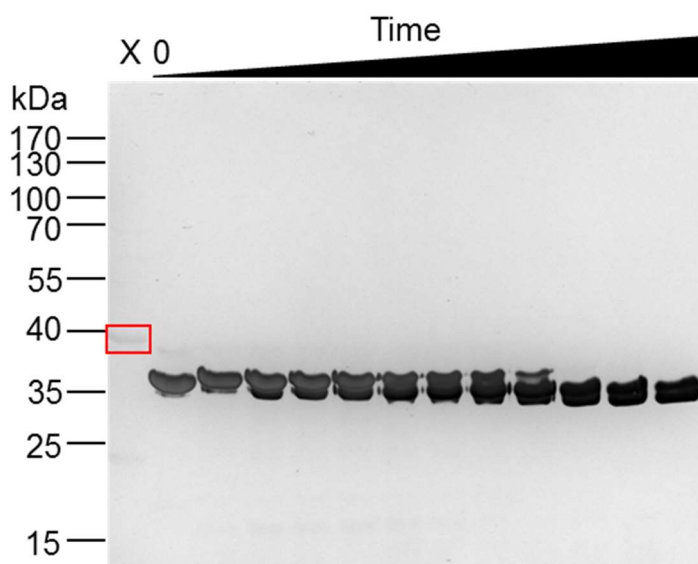


Figure 52. Topology studies of *X. maculatus* La Module by limited proteolysis. Limited trypsinolysis of the XmLARP6-LAM vs the 120 min full-length XmLARP6 trypsinolysis (“X”). Trypsinolysis of XmLARP6-LAM was performed and analyzed as described for the full-length protein.

Together, the limited proteolysis studies provide insight into the identity of the ~40 kDa stable domain. To conclusively identify the boundaries of this species that is within the 40 kDa band, *Xm*LARP6 and *Dra*LARP6a trypsinolysis was performed once again. The 10 and 45 min digest products of both full-length fish LARP6 proteins were analyzed by SDS-PAGE and silver stain. Bands of the ~40 kDa were excised and sent to Dr. Shin Pu at Boise State University Biomolecular Research Center for LC-MS analysis (Figure 53).

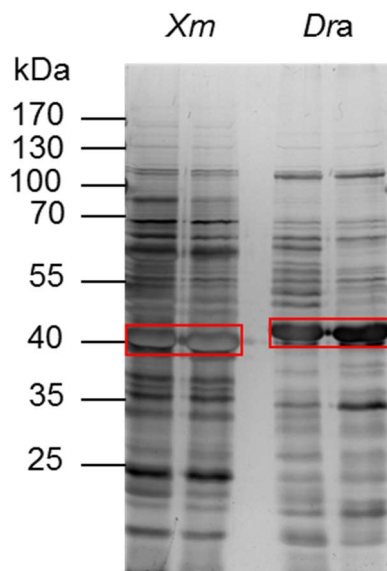


Figure 53. Limited trypsinolysis of *Xm*LARP6 and *Dra*LARP6a and gel extraction for coupled liquid chromatography-mass spectroscopy. Full-length fish proteins (*Xm*: *Xm*LARP6, *Dra*: *Dra*LARP6a) were subjected to trypsin digest. The 10 and 45 min timepoints were analyzed by SDS-PAGE and silver stain. The ~40 kDa bands were excised, destained, subjected to in-gel tryptic digest followed by LC-MS (collaboration with to Dr. Shin Pu and the Boise State University Biomolecular Research Center). The boxes enclose the bands that were sent for MS analysis.

Dr. Pu performed a complete trypsin digest of the excised bands and then subjected the liquid product to LC-MS. Based on the data obtained (Figure 54), the trypsin digest site to be more present in *Xm*LARP6 is located at the 280

amino acid position which is right at the end of the RRM. In addition, from *DrLARP6a* 40 kDa band, the 290 amino acid position was reported to be one of the most abundant trypsin digest cut (at the end of the RRM).

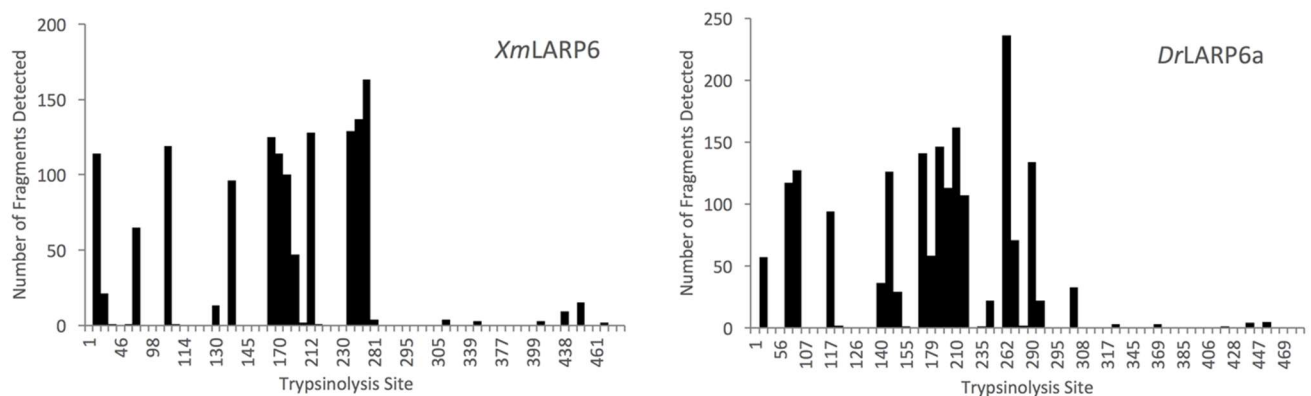


Figure 54. Histograms of peptide species found by LC-MS from the 40 kDa trypsinolysis products of *XmLARP6* and *DrLARP6a*. The data collected by Dr. Shin Pu by LC-MS from the excised bands (after full trypsin digest) shows a range of different cut sites. However, from *XmLARP6* the most abundant cut site is at amino acid 280 right at the end from the RRM. In addition, from *DrLARP6a* one of the most prevalent cut site is at position 290, which is also at the end of the RRM.

Biochemical characterization of LARP6

To measure RNA binding affinity of the LARP6 proteins, electrophoretic mobility shift assays of *HsLARP6*, *XmLARP6* and *DrLARP6a* were performed using the stem-loop sequence from the 5'UTR of *HsCOL1A1* mRNA. Briefly, protein dilution was equilibrated on ice for 1 h with biotinylated *HsCOL1A1* stem-loop RNA. Then, the RNA:protein complex was separated from free RNA on a 5.5% native polyacrylamide gels with ice packs to reduce heat. Then, the RNAs were transferred to a Hybond N+ membrane and UV-crosslinked, followed by detection of the biotin moiety as described in Chapter II.

Figure 55 shows a representative EMSA for the *Hs*LARP6, *Xm*LARP6, and *Dr*LARP6a against *Hs*COL1A1. As can be seen, the bands that appear at a higher apparent molecular weight correspond to LARP6 interacting with the mRNA. The lower apparent molecular weight bands are unbound RNA. Using the ChemiDoc XRS+ molecular imager volume tools, pixel intensity was measured and then used to determine fractional saturation. Fractional saturation data were plotted against protein concentration, and then fit with the binding isotherm:

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

where $[L]_T$ is the total concentration of the ligand (*Hs*COL1A1), $[P]$ the free protein (LARP6), $[PL]/[L]_T$ represents the fraction of bound RNA, S is a saturation offset, $K_{D, app}$ is the calculated apparent dissociation constant, and O is a background offset.³⁶

Figure 56 shows the nonlinear least squares regression for all three sets of RNA binding data. Previous studies of LARP6 RNA binding activity measured dissociation constants using the same sequence of *Hs*COL1A1 mRNA and only the La Module of human LARP6. They reported a K_d of 48 nM. The full-length experiments that we performed, *Hs*LARP6 shows to have an apparent $K_{d,app}$ of 12 ± 2 nM to the same 48 nt ligand. Surprisingly, both fish LARP6 proteins show a higher affinity compared to human; *Xm*LARP6 was found to have a $K_{d,app}$ of 0.6 ± 0.1 nM and the *Dr*LARP6a $K_{d,app}$ was determined to be 1.4 ± 0.3 nM. Both fish LARP6 proteins appear to be interacting with higher affinity to the human mRNA sequence that is the only known ligand for LARP6.

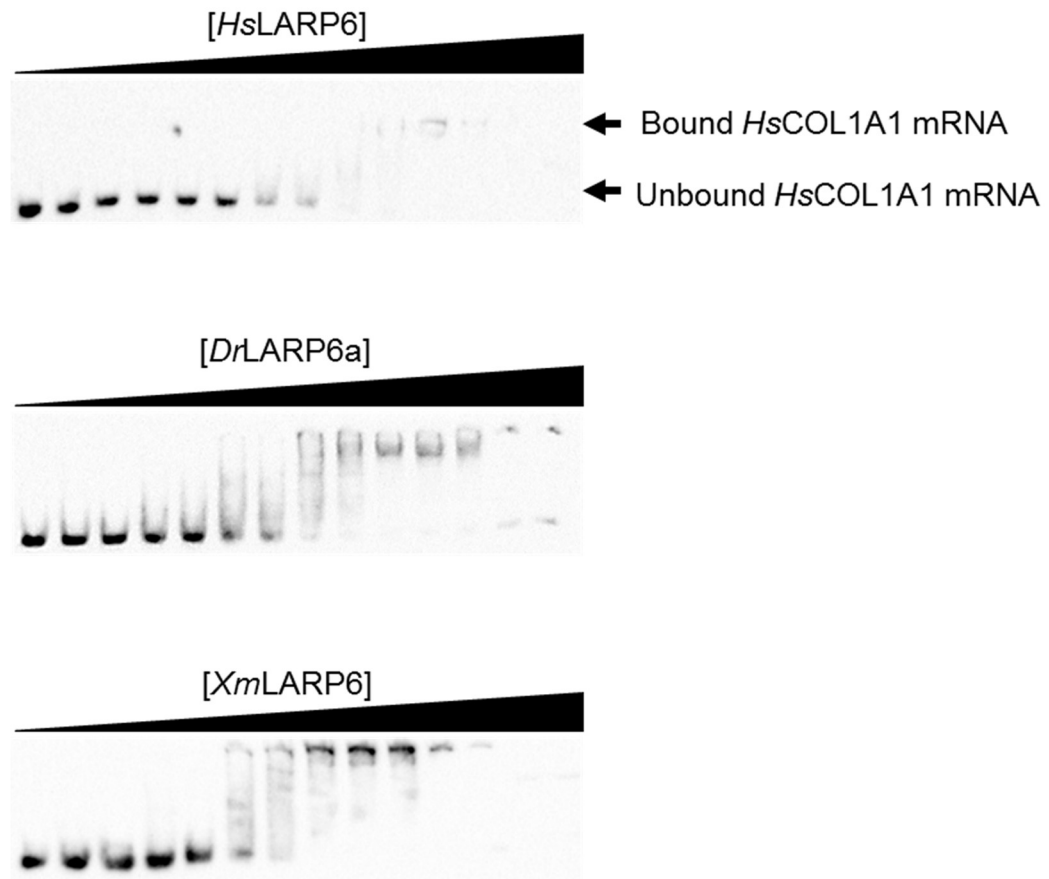


Figure 55. Binding activity of LARP6 proteins for human collagen 1A1 mRNA stem loop. In order to measure the apparent binding activities of the LARP6 proteins against the only known ligand (*HsCOL1A1* mRNA), EMSAs were performed. Different protein concentrations were set to reach equilibrium with a defined RNA concentration. Complexes were subjected to gel electrophoresis by native gel, transblotted to Hybond N+ membrane and cross-linked by UV.

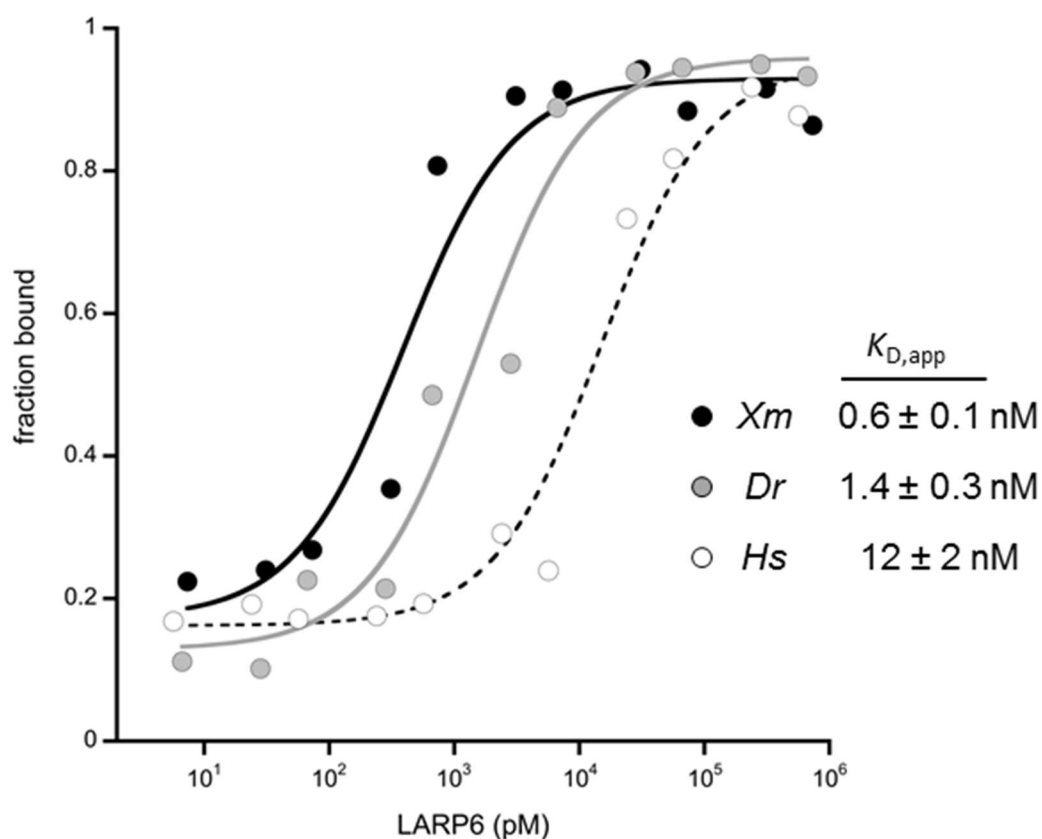


Figure 56. RNA binding activity against biotinylated *HsCOL1A1*. To confirm proper fold and biochemical activity of the full-length LARP6 proteins, the known RNA ligand for human LARP6 was used for assaying RNA binding activity. The recombinant, purified human LARP6 protein was used as a positive control, which exhibited a $K_{D,app} = 12 \pm 2$ nM, comparable to previously published reports.²¹ As expected for homologous proteins, RNA binding activity was also observed for both recombinant fish proteins. Intriguingly, the fish proteins show a tighter binding affinity for the human RNA ligand than the human protein (*Dr*LARP6a, $K_{D,app} = 1.4 \pm 0.3$ nM; *Xm*LARP6, $K_{D,app} = 0.6 \pm 0.1$ nM).

IV. DISCUSSION

Expression and purification of human LARP6

In this study, human LARP6 was transformed into Rosetta cells, expressed and purified as described in previous studies.^{20, 22} *HsLARP6* is predicted by sequence to have a molecular weight of 57 kDa, which is consistent with its migration in SDS-PAGE. However, it elutes from a sizing column at an apparent molecular weight (MW_{app}) of 177.6 kDa. This difference between the expected and observed molecular weight can be explained in several ways: 1) intermolecular interactions generate a homotrimer and therefore it elutes as a bigger protein, 2) *HsLARP6* is not perfectly globular, as the La Module (La motif and RRM) are connected by a flexible linker, and that extended structure influences the apparent molecular weight, or 3) Protein might be interacting with the column matrix by non-specific interactions. Since both domains can display individual movement and the linker can help rotate the whole protein, the hydrodynamic radius that *HsLARP6* experiences can be the same one of a non-globular and/or bigger protein.

Expression of fish LARP6 proteins

In addition, *X. maculatus* LARP6, *D. rerio* LARP6a, and *D. rerio* LARP6b were cloned into the pET28-SUMO plasmid, transformed into Rosetta cells, and expressed. However, when trying to purify SUMO-tagged *XmLARP6* in the same buffer conditions as for human LARP6 (Table 6), a broad elution peak was obtained from the size exclusion column. The protein obtained from this peak is not reliable because the elution had a range of almost 40 mL, suggesting a

polydisperse mixture of structures (Figure 28). As can be seen from the SDS-PAGE in which the fractions within the peak were analyzed (Figure 29), the entire area under the curve contains to SUMO-tagged *XmLARP6*. Additionally, almost 6 mL of the protein eluted in the column's void volume, which strongly indicates that the protein is aggregated.

To effectively purify high-quality preparations of the fish LARP6 proteins, it was necessary to optimize the buffer components. Dr. Kelly Churion and Dr. Sarah Bondos developed an additive screening assay to increase the solubility and stability of intrinsically disordered proteins.³³ The first screening round identified that all six additive categories increased both protein solubility, as full length His₆-SUMO-*XmLARP6* was collected in the filtrate, and protein stability, as few degradation products were found to precipitate (Figure 30).

The optimization stage (iteration two) screened a broader set of additives based on initial stage outcome.³³ This stage determined that the two osmolytes that yielded the most promising results were glucose and glycerol (Figure 31). Both contain hydrophobic and hydrophilic regions that can either coordinate water molecules around the protein shell or help stabilize the protein by interacting directly with the residues. During the chaotropes screening, as reported in Chapter III, the calcium chloride and magnesium chloride solutions induced white precipitates. When analyzed by gel electrophoresis (Figure 32), protein was not detectable in these solutions either as filtrate or retentate. The divalent cations in these solutions interact with the screening assay's base phosphate buffer, becoming insoluble calcium phosphate and magnesium

phosphate compounds. The newly formed salt precipitates could have induced protein precipitation. This would explain the absence of protein in both the filtrate, but the absence in retentate remains unexplained.

The third stage of screening was performed to test glucose and glycerol (osmolytes) in different concentrations in combinations with the chaotropes.³³ From figure 33, 5% glycerol seemed promising. However, since previous purifications of human LARP6 purification buffer contain the same amount of glycerol and gave poor results in the initial *XmLARP6* purification, glucose was chosen as a more viable alternative. The optimized buffer conditions were 50 mM NaH₂PO₄/ Na₂HPO₄ (pH 8.0), 200 mM NaCl, 0.5 M glucose, and 300 mM NaI.

The purified *XmLARP6* and *DrlARP6a* exhibited a larger molecular weight by SEC than predicted by sequence, consistent with the behavior of the human LARP6 protein. To test the hypothesis that extended regions may exist in the proteins that would increase their apparent hydrodynamic radius, the LARP6 protein sequences were analyzed by PONDR (Predictors of Natural Disordered Regions).³⁷ Figure 57 contains graphs obtained from three different disorder algorithms provided by PONDR: VLXT (blue), VL3 (red), and VSL2 (green). Each algorithm classified the human and fish LARP6 proteins to be $\geq 50\%$ disorder. Because a lack of globular features increases hydrodynamic radius, a protein with disordered regions would appear to be a larger protein or protein complex by SEC. This phenomenon could explain the smaller elution volume obtained on the SEC.

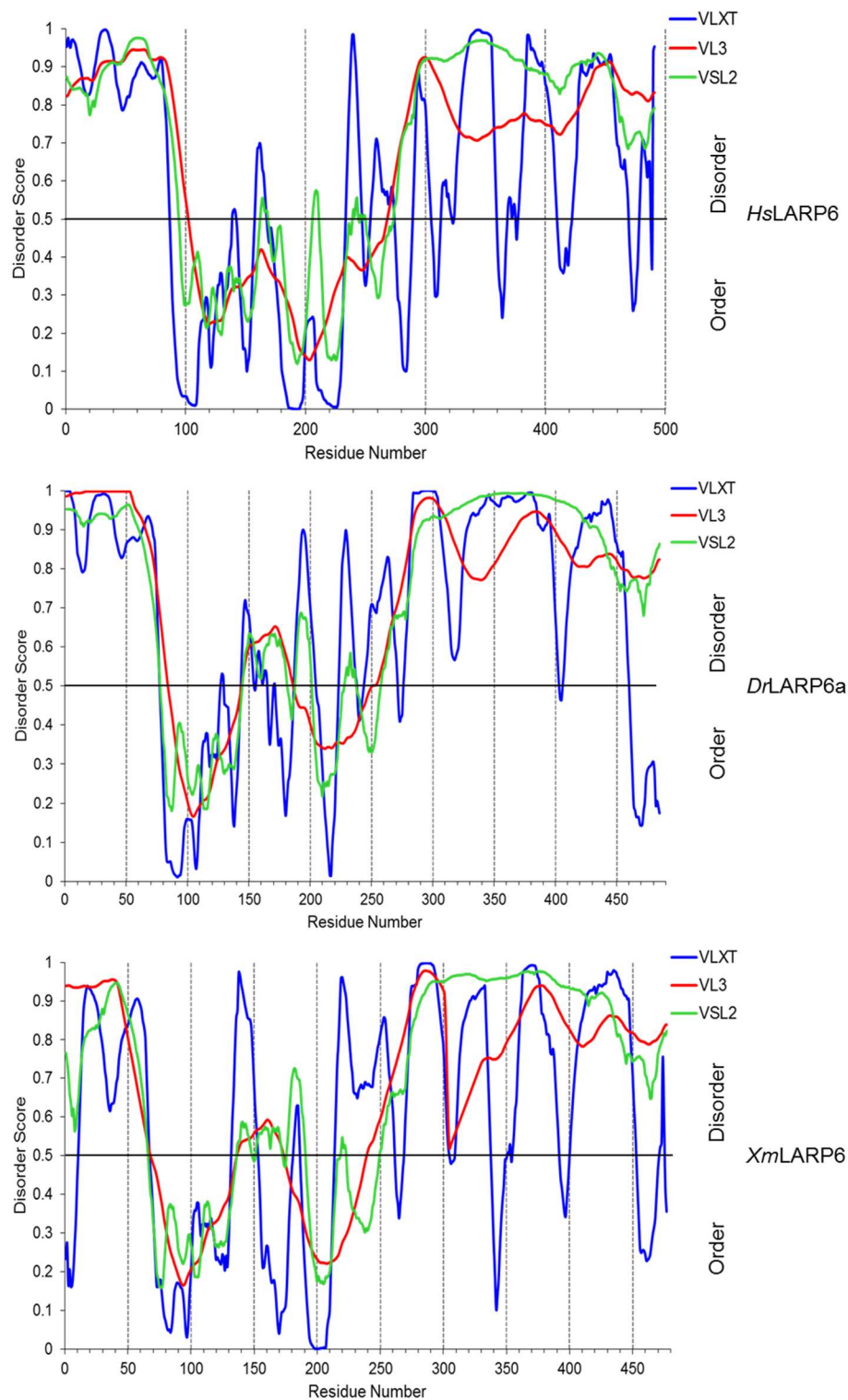


Figure 57. Disorder prediction of LARP6 proteins. The Predictor of Natural Disordered Regions (PONDR), web server was used to determined disorder regions within each LARP6.³⁷ The three algorithms available at the PONDR website were used to predict disorder: VLXT (blue), VL3 (red), and VSL2 (green). All three algorithms classified each LARP6 as $\geq 50\%$ disorder proteins.

As previously mentioned, the *D. rerio* genome contains two LARP6 proteins (LARP6a and LARP6b). SUMO-tagged *Dr*LARP6b was purified in the same buffer conditions obtained from the solubility screen assay for *Dr*LARP6a. Figure 58 shows the SDS-PAGE from the affinity chromatography of His₆-SUMO-*Dr*LARP6b. The ratio of full-length protein to small molecular weight species was low. Roughly, less than 50% of the protein eluted as full-length, with a large amount of degradation product. Consequently, purification was put on hold. The *Dr*LARP6b sequence was then analyzed by PONDR and the three different disorder prediction algorithms previously described.

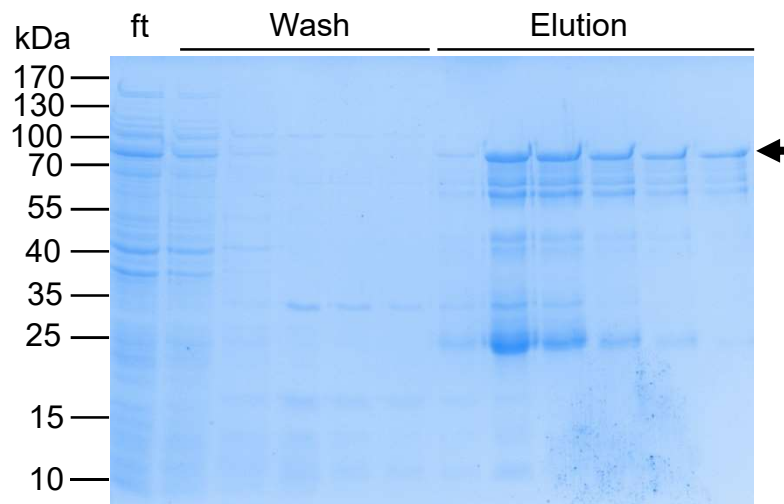


Figure 58. His₆-SUMO-*Dr*LARP6b affinity chromatography fractions using optimized buffer conditions. Decided to test the new buffer with zebrafish LARP6b. Cells expressing His6-SUMO-*Dr*LARP6a were treated the same as *Xm*LARP6. The cell lysate was purified using nickel-affinity chromatography ("ft", flowthrough). Aliquots were separated by gel electrophoresis and stained with Coomassie blue. The five elution fractions (arrow) showed a great amount of degradation products compared to full length.

Figure 59 predicts the level of disorder in *DrLARP6b* protein. As can be seen from the graphs, this protein likely contains more extensive disordered regions compared to the other three LARP6 proteins discussed in this work. This larger amount of disorder could make His₆-SUMO-*DrLARP6* more susceptible to degradation by proteases at the start of the protein purification, even though protease inhibitors are dissolved in the lysis buffer before sonication. Because sufficient quantities of the *XmLARP6* and *DrLARP6a* protein were obtained, the decision was made to delay optimization of *DrLARP6b* purification for a later date. In the future, further screening assays should be performed to develop a buffer that will promote the solubility and stability of *DrLARP6b*.

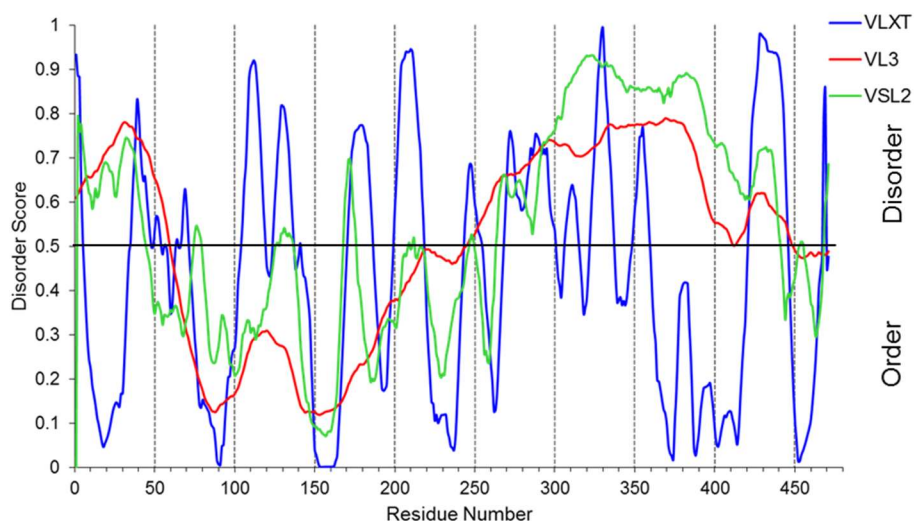


Figure 59. Disorder prediction of *DrLARP6b* protein . The Predictor of Natural Disordered Regions (PONDR), web server was used to determined disorder regions within *DrLARP6b*.³⁷ Three algorithms to predict disorder available in PONDR website were used: VLXT (blue), VL3 (red), and VSL2 (green). All three algorithms classified each LARP6 as $\geq 60\%$ disorder proteins.

Biochemical characterization of fish proteins

To test for functional protein, RNA binding studies were done in which the 48 nt 5'UTR from *HsCOL1A1* was biotinylated and bound to *Hs*, *Xm*, or *Dr* LARP6 protein (Figure 5). EMSAs were performed to measure apparent dissociation constants. Surprisingly, both *Xm*LARP6 and *Dr*LARP6a bound to the stem-loop of *HsCOL1A1* with a higher affinity than human LARP6. The elucidation of the molecular mechanism(s) that increase the binding affinity of the fish proteins should be pursued. Some of the factors that could yield this result are differences in amino acid composition of the three proteins, and possibly differences in the tertiary structure of the fish proteins compared to the human protein.

Physical characterization of fish proteins

X. maculatus LARP6 and *D. rerio* LARP6a were subjected to limited trypsin digest to probe the general domain topology. A stable ~40 kDa band was generated from both proteins. These band were excised from the gel and sent for mass spectroscopy analysis. The prevalent species found in those bands was a digested segment of the full-length proteins, consisting of the extreme N-terminus through the end of the RRM. Since the band was generated between the first minutes the reactions started and it lasted for the four hours the reaction was monitored, this result indicates that the polypeptide chain forms a well-folded, protease-resistant structure. The discovery that the N-terminus of LARP6 may

interact with the RNA binding domain (La Module) presents interesting possibilities for intra-molecular regulation of LARP6 RNA binding activity.

Future directions

To further characterize and understand the species specificity between LARP6 and the natural ligand, human and fish proteins should be analyzed with RNA binding studies using biotinylated fish collagen type 1 mRNAs 5'UTR sequences. In addition, the predominant abundant species acquired from the 40 kDa bands from *X. maculatus* and *D. rerio* LARP6a was a full N-terminal product. These constructs should be cloned, expressed, and biochemically characterized as done with the full-length proteins to determine what role, if any, the N-terminal domain plays in RNA binding activity. These experiments should be performed alongside measurements of binding activity of the full-length and isolated La Module to their respective co-evolved collagen ligands.

Finally, to answer one of the original questions posed in Chapter I, whether if domain swapping within a same protein subfamily maintains proper functionality with natural ligand, some chimeras could be generated for recombinant expression, purification, and characterization. Figure 60 shows a schematic representation of the chimeras that would be most useful for these proposed studies. Human LARP6 will exchange either the La motif, the RRM or the whole La Module with the respective domains from *X. maculatus* LARP6. RNA binding activity will be measured against the 5'UTR stem-loop sequence from collagen type 1 mRNA in both human and *X. maculatus* to determine if the domain swaps affect binding specificity or affinity.

A collaboration with Dr. Lisa Warner from Boise State University is in progress to determine the NMR structures of the La motif and RRM domains from both *Xm*LARP6 and *Drl*LARP6a. Determining the structure of the fish domains will help visualize the surface area that might be interacting with their respective ligands. Structural comparison studies can then be done between the human and the fish domains to identify regions that could be the source of the differential binding activity.



Figure 60. Human LARP6 chimera constructs scheme. To understand if LARP6 domains are interchangeable within the same superfamily, chimeras will be constructed. Individual domain swaps (*Xm*LARP6-La motif or RRM (blue) will be substituting the same domain in Human LARP6) or the entire *Xm*LARP6-La Module will be exchange with the human one. Chimeras will be subjected to biochemical characterization with human and *X. maculatus* collagen type 1 mRNA as well as limited proteolysis.

Since other LARPs have multiple RNA ligands, I propose a reverse crosslink study with either platyfish or zebrafish tissues, in which LARP6 would be crosslinked to natural RNA ligands. Then, LARP6 would be immunoprecipitated, and the co-precipitated RNAs would be sequenced to identify other possible ligands or proteins that could be interacting with LARP6. Another experiment that could be done would be in collaboration with Dr. Walter, to study changes in LARP6 mRNA expression levels under different environmental conditions. A gene network map could be generated that may provide clues to the cellular pathways in which LARP6 might be participating.

REFERENCES

1. Bandziulis, R. J., Swanson, M. S., and Dreyfuss, G. (1989) RNA-binding proteins as developmental regulators. *Genes & Development* **3**, 431—437.
2. Blackburn, G. M., Gait, M. J., Loakes, D., and Williams, D. M. (2006) “*Nucleic Acids in Chemistry and Biology*” 3rd edition.
3. Glisovi, T., Bachorik, J., *et al* (2008) RNA-binding proteins and post-transcriptional gene regulation. *FEBS letters* **582**, 1977—1986.
4. Oubridge, C., Ito, N., Evans, P. R., Hiang Teo, C., and Nagai, K. (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372**, 432–438.
5. Cléry, A., Blatter, M., and Allain, F. H.-T. (2008) RNA recognition motifs: boring? Not quite. *Struct. Biol.* **18**, 290–298.
6. Yan, K. S., Yan, S., Farooq, A., Zeng, L., and Zhou, M. M., (2003) Structure and conserved RNA binding of the PAZ domain. *Nature* **426(6965)**, 468–474.
7. Tang, G (2005) siRNA and miRNA: an insight into RISCs. *Trends Biochem. Sci.* **30** (2), 106–14.
8. Valverde, R., Edwards, L., and Regan, L. (2008) Structure and function of KH domains. *The FEBS J.* **275(11)**, 2712—2726.
9. Stefl, R., Skrisovska, L., and Allain, F. H.-T. (2005). RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Reports. Nature Publishing Group.* **6** (1), 33–38.
10. Bycroft, M., Hubbard, T. J., Proctor, M., Freund, S. M., and Murzin, A. G. (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell* **88**, 235–242.
11. Schubert, M., Edge, R. E., Lario, P., Cook, M. A., Strynadka, N. C., *et al.* (2004) Structural characterization of the RNase E S1 domain and identification of its oligonucleotide-binding and dimerization interfaces. *J. Mol. Biol.* **341**, 37–54.
12. Auweter, S. D., Oberstrass, F. C. and Allain, F. H. (2006) Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic Acids Res.* **34(17)**, 4943–4959.
13. Maris, C., Dominguez, C. and Allain, F. H. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* **272**, 2118–2131.
14. Shamoo, Y., Abdul-Manan, N., Patten, A. M., Crawford, J. K. *et al* (1994) Both RNA-binding domains in heterogenous nuclear ribonucleoprotein A1 contribute toward single-stranded-RNA binding. *Biochemistry* **33**, 8272–8281.

15. Shamoo, Y., Abdul-Manan, N. and Williams, K. R. (1995) Multiple RNA binding domains (RBDs) just don't add up. *Nucleic Acids Res.* **23**, 725–728.
16. Wolin, S. L., and Cedervall, T. (2002) The La protein. *Annu. Rev. Biochem.* **71**, 375–403.
17. Francoeur, A. M., Gritzmacher, C. A., Peebles, C. L., Reese, R. T., and Tan E. M. (1985) Synthesis of small nuclear ribonucleoprotein particles by the malarial parasite *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **82** (11), 3635–3639.
18. Naeeni, A., Conte, M. R., and Bayfield, M. A. (2012) RNA Chaperon Activity of Human La Protein is Mediated by Variant RNA Recognition Motif. *J. Biol. Chem.* **287**, 5472–5482.
19. Stavraka, C. and Blagden, S. (2015) The La-Related Proteins, a Family with Connections to Cancer. *Biomolecules* **5**(4), 2701–2722.
20. Martino, L., Pennell, S., Kelly, G., Busi, B., Brown, P., Atkinson, R. A., and Conte, M. R., et al. (2015) 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 *Nuc. Ac. Res.* **43**(1), 645–660.
21. Cheng, Y., Jin, Z., Agarwal, R., Ma, K., Yang, J., Ibrahim, S., and Mori, Y., et al. (2012) LARP7 is a potential tumor suppressor gene in gastric cancer. *Lab Invest. A Journal Tech. Meth. Path.* **92** (7), 1013–1019.
22. Hussain, R. H., Zawawi, M., and Bayfield, M. A. (2013) Conservation of RNA chaperone activity of the human La-related proteins 4, 6 and 7. *Nucleic Acids Res.* **41** (18), 8715–8725.
23. Yang, R., Gaidamakov, S. A., Xie, J., Lee, J., Martino, L., Kozlov, G., and Conte, M. R. (2011) La-related protein 4 binds poly(A), interacts with the poly(A)-binding protein MLLE domain via a variant PAM2w motif, and can promote mRNA stability. *Mol. Cell. Biol.* **31**(3), 542–556.
24. Lahr, R. M., Fonseca, B. D., Ciotti, G. E., Al-Ashtal, H. A., Jia, J.-J., et al (2017) La-related protein 1 (LARP1) binds the mRNA cap, blocking eIF4F assembly on TOP mRNAs. *eLIFE*
25. Aoki, K., Adachi S., Homoto, M., Kusano, H., Koike, K., and Natsume, T. (2013) LARP1 specifically recognizes the 3' terminus of poly (A) mRNA. *FEBS Lett.* **587**(14), 2173–2178.
26. Cai, L., Fritz, D., Stefanovic, L., and Stefanovic, B. (2010) Binding of LARP6 to the Conserved 5' Stem-Loop Regulates Translation of mRNAs Encoding Type I Collagen. *J. Mol. Bio.* **395**(2), 309–326.
27. Bayfield, M. A., Ruiqin, Y., and Richard J. M. (2010) Conserved and divergent features of the structure and function of La and La-related proteins (LARPs) *Biochem Biophys Acta.* **1799**(5-6), 365-378.
28. Wang, Z., Glenn, H., Brown, C., Valavanis, C., Liu, J.-X., Seth, A., Thomas, J. E., Karlstrom, R., and Schwartz, L. M. (2009) Regulation of

- muscle differentiation and survival by Acheron. *Mech. Dev.* **126**(8-9), 700—709.
29. Bousquet-Antonelli C., and Deragon, J. M. (2009) A comprehensive analysis of the La-motif protein superfamily *RNA*. **15**(5), 750-764.
 30. Betancourt, F. (2017) “Recombinant Expression and Purification of LARP6 Proteins from *Arabidopsis thaliana*.” (Master’s Thesis). In press, Texas State University Library.
 31. Mossessova, E. and Lima, C. D. (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell.* **5** (5), 865—876.
 32. Haan, C. and Behrmann (2007) A cost effective non-commercial ECL-solution for Western blot detections yielding strong signals and low background. *Jour. Of Immuno. Methods* **318**(1-2), 11—19.
 33. Churion, K. A. and Bondos, S. E. (2012) Identifying Solubility-Promoting Buffers for Intrinsically Disordered Proteins Prior to Purification. *Mth. Mol. Biol.* **896**,415—427.
 34. Fuhrmann M, Hausherr A, Ferbitz L, Schödl T, Heitzer M, Hegemann P. (2004) Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant Mol Biol.* **55**(6), 869—81.
 35. Li, S-J., and Hochstrasser, M. (1999) A new protease required for cell-cycle progression in yeast. *Nature*. **398**(6724), 246—251.
 36. Altschuler, S. E., Lewis, K. A. and Wuttke, D. S. (2013) Practical strategies for the evaluation of high-affinity protein/nucleic acid interactions. *J. Nucleic Acids Investig.* **4**(1), 19—28.
 37. Romero, P., Li, X., Dunker, A. K., Obradovic, Z., Garner, E., Radiovajoc, P., Peng, K., (2007) *PONDR: Predictors of Natural Disordered Regions software*.