MECHANISMS OF HUMAN LARP6 NUCLEAR EXPORT

by

Samantha Kathleen Zepeda, 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 2019

Committee Members:

Karen A. Lewis, Chair

Tania Betancourt

Wendi David

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

ACKNOWLEDGEMENTS

Science is not an independent pursuit and for this reason, I would like to acknowledge the many people in my life who have assisted me on this journey. In my personal life, I would like to thank my friends and family for their support. In particular, I would like to thank my husband for his support, patience and endless coffee runs. Additionally, my siblings, Jay and Kaitlynn Griesenbeck, who have always given me motivation to continue reaching for the next challenge.

Professionally and personally, my mentor, Dr. Karen A. Lewis, has been a wonderful guide on this journey and I would like to thank her for all of the time, support and mentorship that she has provided. In her lab, I have had the opportunity to fall in love with research and grow into a more knowledgeable and proficient researcher. Her lab has changed my career path and set me on the road to a Ph.D. program. I originally believed that the path I am now on, was unattainable, and I will forever be grateful for the role Dr. Lewis has provided in helping me develop the skills and confidence to continue on.

My lab mates past and present have been invaluable resources to me throughout my time here. In particular I would like to thank Ms. Leticia Gonzalez for her contribution of Ran-GTP for this project and the initial cloning and purification attempts for CRM1 and SUMO full-length LARP6. Her presence, advice, and preparation of reagents and maintenance of cell culture lines has been invaluable.

I would also like to thank Melissa Carrizales for her kindness and feedback. She has been my sounding board throughout this journey and I am forever grateful for her attitude, intelligence, and friendship. Julia Roberts and Chelsea Toner have provided endless feedback and support. I would also like to thank José Castro and Eliana Pena for their training and guidance when I started this journey.

In addition, I would also like to thank Dr. Joe Koke, Dr. Raquel Salinas, and Dr. Tania Betancourt for their advice and support as we set up a cell culture room for my project. Without their guidance I would truly be lost. Lastly, I would like to thank the Texas State University Analytical Research Service Center (ARSC) and its staff for their technical expertise, training, and guidance in cellular imaging. I would also like to thank Alissa Savage in particular for her endless patience, guidance, and support in cellular imaging and cell culture.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Description
LARP6	La Related Protein 6
NES	Nuclear export sequence
GST	Glutathione S-transferase
RRM	RNA recognition motif
LARP	La Related Protein
TOP	Terminal oligopyrimidine
LSA	La and SUZ-C-associate domain
NLS	Nuclear localization sequences
XPO1	Exportin 1
HIVrev	HIV-1 Rev
PKIα	Protein kinase inhibitor
LocNES	Locating Nuclear Export Signals or NESs
LMB	Leptomycin B
IDT	Integrated DNA technologies
SDS	Sodium Dodecyl Sulfate
OD_{600}	Optical density at 600 nm
IPTG	Isopropyl b-D-1-thiogalactopyranoside
DTT	Dithiothreitol
PCR	Polymerase Chain Reaction
OE-PCR	Overlap extension PCR
CGM	Complete growth media
PFA	Paraformaldehyde
BCA	Bicinchoninic assay
ICC	Immunocytochemical
RT-qPCR	Real time quantitative PCR
SEC	Size exclusion chromatography

ABSTRACT

The La Related Protein 6 (LARP6) regulates the expression of type I collagen synthesis by binding to the encoding mRNA. However, the molecular mechanism of ligand binding and cellular localization is not fully understood. Understanding the mechanisms of LARP6 regulation requires knowledge of where and how LARP6 functions inside the cell. Previous work suggested that LARP6 undergoes nucleocytoplasmic shuttling, supported by the identification of a putative nuclear export sequence (NES) and to localization in both the nucleus and the cytoplasm. Subsequent high-resolution structures of the RNA binding domain of LARP6 showed that the NES is located in a core part of the RNA Recognition Motif, where it may contribute to RNA binding and/or be a core structural element. The goals of this thesis were to combine these cellular and structural models of LARP6 localization into a cohesive model of LARP6 function. First, to test if LARP6 export is mediated by the nuclear exportin Chromosome Region Maintenance 1 (CRM1) *in vivo*, we performed inhibition studies in cell lines that endogenously expressed LARP6. Second, we developed two related approaches that leveraged the recombinant LARP6 protein approaches that had been established in the lab. To determine whether computationally-identified sequences are responsible for nuclear export, we recombinantly expressed putative NES motifs as glutathione S-transferase (GST) fusion proteins for use in *in vitro* CRM1 pulldown assays. We identified three additional putative NES using LocNES, a more recent and robust NES identification algorithm. We have also designed NES deletion mutants, which will be used to examine the effects on structural stability caused by the deletion of all putative NES motifs by introducing those mutations into the recombinant, full-length protein construct that is expressed in *E. coli*.

I. INTRODUCTION

The central dogma of biology describes a pathway in which genetic information flows from DNA to RNA and finally to proteins.[1] This simple model describes a biological process that is in fact much more complicated. A variety of regulatory systems broaden the repertoire of the cell and are further augmented by eukaryotic cell structure and the separation of biological processes into subcellular compartments.[2, 3, 4] These systems act to expand control over biological processes, allowing a shift from dependence on concentration and diffusion to precise control of many biological functions, including the regulation of gene expression.[2]

RNA binding proteins exert a wide array of functions, including controlling the availability of RNA through degradation, acting as chaperones to help RNA fold into a functional structure, assisting in the nuclear export of mRNA after transcription, or to chemically modify RNA molecules to regulate their activity.[4, 5] To correctly identify their target, RNA-binding proteins must associate with RNA in a specific manner. They utilize one or more of several common structural motifs to bind the RNA, including the RNA recognition motif (RRM), zinc finger motif, cold shock domain, PUF domain, and the K homology domain.[6] The RRM is one of the most common and best-characterized domains in RNA binding proteins.[6, 7] Collectively, these proteins offer the ability to tightly control gene expression at the post-transcriptional level by binding to the mRNA message itself.

La-Related Proteins

The members of the La Related Protein (LARP) superfamily are RNA binding proteins that are distinguished by a shared structural feature called the La Module (Figure 1). The La Module is comprised of two globular domains, the La motif and an RRM, which are connected by a flexible inter-domain linker. All three components are required to bind to RNA.[8] The La motif forms a novel winged helix-turn-helix,[9, 10] while the RRM1 adopts a variation of the "canonical" RRM that contains a fourth beta strand (creating a $\beta1\alpha1\beta2\beta3\alpha2\beta4$ fold).[10, 11] The wing of the helix-turn-helix domain is the presence of structured loops on exterior of the traditionally defined helix-turn-helix motif. Collectively, the LARP families bind a variety of RNA targets, including pre-tRNA and mature mRNA, despite the sequence and structural conservation within the La Module.[12] Of the two components of the La Module, the La Motif has the highest degree of conservation.[11] The RRM is more divergent, and therefore is predicted to be a large driver of the variety of RNA ligands in the LARP family repertoire.

LARP members have different domains that are only associated with one LARP family and have specific functions that expand the functionality of that particular LARP beyond the La Module. For example, the DM15 domain in the C-terminus of LARP1 has been recently associated with the translation of mRNAs containing a terminal oligopyrimidine (TOP) motif, through binding of the mRNA cap and blocking of translation initiation factors.[14]

Dysregulation of members of this superfamily is widely implicated in cancers, demonstrating the importance to human health of their varied regulatory

roles in the cell.[13] The RNA ligands for the “genuine La” protein (which is now also called LARP3) include pre-tRNA and mRNA. LARP7 is most closely associated with LARP3 in sequence and structure, and also binds RNA Polymerase III transcripts (specifically 7SK RNA, which is involved in mRNA regulation).[13] LARP1 has been demonstrated to bind 5’ TOP mRNAs, which encode for ribosomal proteins and translation elongation factors.[14] LARP4 associates with the polyA-binding proteins involved in mRNA regulation and stabilization.[15] While LARP6 is the only member of the family with no known general cellular function, it is known to interact with a conserved structure in the 5’ untranslated region of type I collagen mRNA.[16]

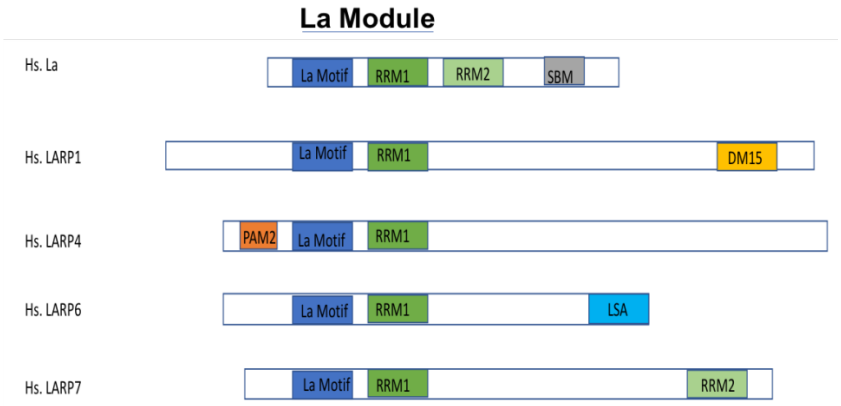

Figure 1: Topology of the La Related Protein Superfamily. Topology of human La protein (LARP3) and LARP members 1, 4, 6, and 7 are shown. The common structural motifs of the La Motif and RRM1 are shared by all members of this superfamily. Additional domains in the N- and C-termini distinguish individual families.

LARP6

LARP6 is found in all multicellular eukaryote genomes that have been sequenced to date.[11] In vertebrates, the LARP6 family post-transcriptionally regulates the expression of collagen type I through an interaction with a stem

loop in the 5' untranslated region of collagen type I mRNA.[16] In 2010, the same group reported that collagen mRNA is bound by LARP6 as the mRNA was translocated into the cytoplasm from the nucleus.[16] Subsequently, LARP6 has been implicated in interactions with trafficking machinery to localize collagen type I mRNA to the rough ER for translation.

In addition to its role in regulation of collagen synthesis, LARP6 has been implicated in muscle differentiation, apoptosis, and in some cancers.[17, 18] This observation suggests that LARP6 may exert regulatory functions within the nucleus, perhaps by cooperating in a complex with these transcription factors. In addition, aberrant LARP6 expression has been associated with some breast cancers, although the mechanism of this relationship is unknown.[19] Together, these data demonstrate the need for a full understanding of the molecular mechanisms of LARP6 cellular function.

LARP6 Structure

Structurally, LARP6 shares the characteristic La Motif and the RRM with the other members of the LARP family. These motifs were structurally characterized using solution NMR (Figure 2).[20] The RRM of LARP6 contains additional structural elements when compared to other members of the LARP family, including an extended loop connecting the β strands and additional α helices (α0' and α1'). The α1' helix packs against the beta sheet, which is the RNA binding surface in canonical RRMs. However, the NMR structures showed

that there is very little dynamic movement occurring within the structure, suggesting that this helix does not move upon RNA binding.[20]

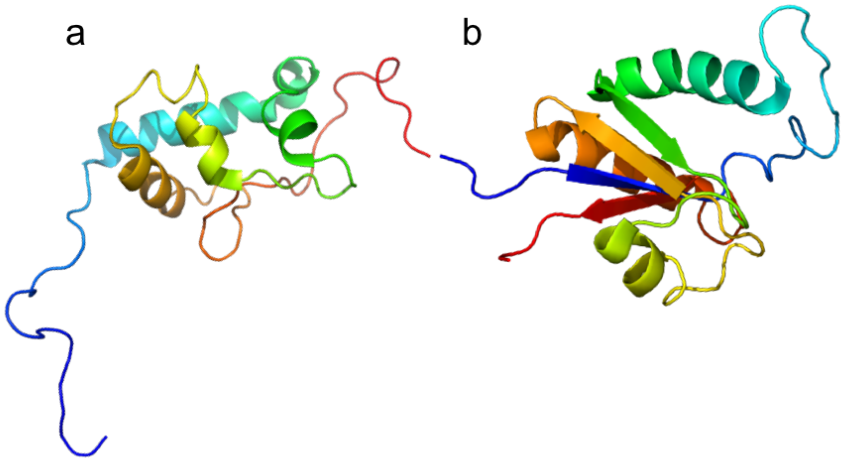

Figure 2: Solution NMR structures of human LARP6. (a) La Motif (PDB: 1MTF) and (b) RRM (PDB: 1MTG).[20] The N-terminus is labeled in blue and the C-terminus is labeled in red. Images created in PyMol (Schrodinger, Inc.). Note that the orientation of the La Motif and RRM between (a) and (b) is a prediction, based on the termini of each domain as determined in the NMR structural models.

In addition to the core La Module, LARP6 contains an additional domain, called the La- and SUZ-C-associated domain (LSA). While the LSA in LARP6 has no known RNA binding activity, it is predicted to be important in protein-protein interactions, particularly with the Serine/Threonine Receptor-Associated Protein.[21]

Previous work using cell culture has shown that LARP6 can be found both in the nucleus and the cytoplasm.[16, 19] This localization pattern suggests nucleocytoplasmic shuttling. However, both the phenomenon of shuttling as well

as the specific transport mechanism have yet to be fully elucidated, including whether export to the cytosol is dependent on binding an RNA ligand.

Nuclear Localization

In eukaryotes, the nuclear envelope physically separates the processes of transcription and translation.[22, 25] Macromolecules cross the nuclear membrane through nuclear pore complexes, which allow for bi-directional transport of proteins and RNAs.[23] The pores allow for passive diffusion of molecules smaller than ~ 40kDa but require active transport for macromolecules of larger sizes.[24] This active transport requires both a carrier protein and an energy source. Nuclear import and nuclear export are mediated by transport proteins of the karyopherin family.[23, 36] The association of these molecules and their cargos is on the basis of an interaction with a sequence or "signal" on the cargo protein. These signals are called "nuclear localization sequences" (NLS) for nuclear import and "nuclear export sequences" (NES) for nuclear export.[25]

Exportins bind their cargo in the nucleus and release it in the cytosol.[23] The molecular basis of this interaction is the specific recognition of the NES by a binding site in the exportin. The predominant transport protein for nuclear export is Chromosome Region Maintenance 1 (CRM1) or Exportin 1 (XPO1), a member of the β-karyopherin family.[26, 27] Among the exportins, CRM1 is the most generalized and prevalent, with the other family members executing highly specialized cargo transport.[26, 27] CRM1-mediated nuclear export is a form of active transport, energy is provided in the form of the Ran protein that binds to

GTP (Ran-GTP). CRM1 binding to Ran-GTP is required for passage out of the nucleus through the nuclear pore complex.[22] The directionality of nuclear export is maintained through the maintenance of the Ran-GNP complex, by compartmentalizing its own regulators to either the nucleus or cytosol.[22] The Ran regulatory molecules control whether Ran is bound to GTP or GDP (Figure 3). The guanosine nucleotide exchange factor, Ran-GEF, is found in the nucleus and chaperones the exchange of GDP for GTP to make the Ran-GTP complex.[22] The GTPase activating protein, Ran-GAP, is found in the cytoplasm and promotes the GTPase activity of Ran to convert GTP into GDP to make Ran-GDP.[22] Successful nuclear export of cargo through the nuclear pore complex is dependent on the formation of a complex comprised of CRM1, Ran-GTP, and the cargo protein. In the cytosol, CRM1 releases the cargo, and Ran-GDP and CRM1 are then recycled back into the nucleus as nuclear import cargo.[22]

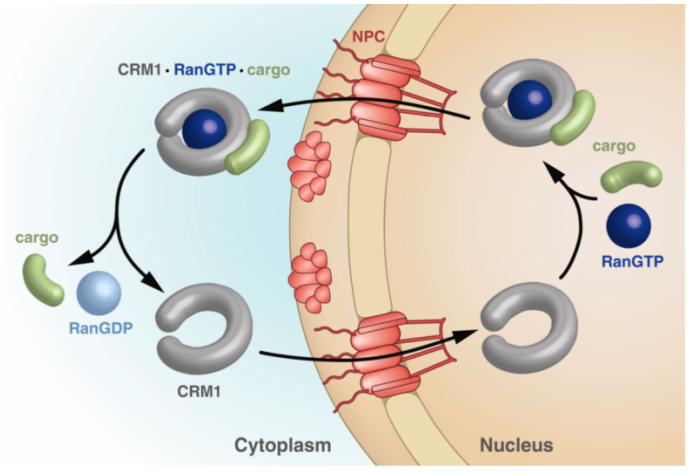

Figure 3: Mechanism of CRM1 mediated nuclear export. In the nucleus, a NES on a CRM1 cargo protein acts as bait for CRM1 binding. When in complex with Ran-GTP, CRM1 is capable of transporting its cargo protein through the nuclear pore complex and releasing it in the cytosol. CRM1 can then be recycled into the nucleus. Image reproduced under Creative Commons copyright.[28]

Cargo Recognition by CRM1

CRM1 contains a binding pocket on its convex side which contains five hydrophobic amino acid residues. Hydrophobic interactions with its cargo protein occur within this binding pocket.[37] The segment of the protein that binds to the CRM1 binding pocket is called the Nuclear Export Sequence/Signal. NESs are short peptide sequences comprised of 8-15 amino acids. These are found in cargo proteins and direct transport from the nucleus to the cytoplasm. Blocking of the interaction of CRM1 and its cargo protein is one avenue of treatment for cancer.[42] The first identified NES were found in the proteins HIV-1 Rev (HIVrev), and cyclic AMP-dependent protein kinase inhibitor (PKIα), which contain NES that were rich in leucine residues.[29, 30] While many NES motifs follow this pattern, it was later discovered that leucine residues were not a prerequisite for nuclear export.[29, 31] In fact, several other hydrophobic amino acids can function in NES motifs in place of leucine. As more NESs have been functionally identified, the field's understanding of what sequences constitute an NES has shifted.[27, 29 ,30, 31, 38, 39, 40, 41] The defining characteristic of these sequences continue to be defined intervals of hydrophobic amino acids, which are preferentially leucine, isoleucine, valine, methionine, and phenylalanine.[31, 33] The prevalence and positioning of hydrophobic residues is a direct consequence of the convex crescent shape and hydrophobic character of CRM1's binding pocket (Figure 4).[34]

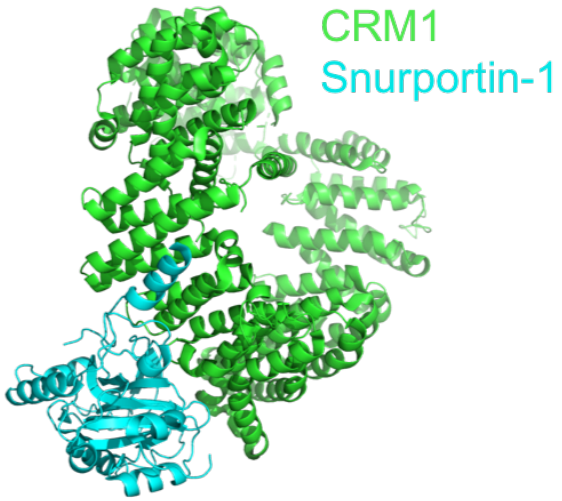

Figure 4: Crystal structure of CRM1/Snurportin-1 Complex. Crystal Structure of experimentally confirmed cargo protein Snurportin-1 (cyan) in complex with CRM1 (green). This interaction is based on a NES within Snurportin-1 binding to hydrophobic residues within the CRM1 binding pocket (PDB: 3GB8). Images created in PyMol (Schrodinger, Inc.).[32]

The "Kosugi consensus sequence", named for one of the principal founders of the field of NES prediction, is one of several NES-naming conventions used as the nuclear-export field developed and often used to identify novel NES motifs in cellular proteins.[31] Notably, the Kosugi consensus sequence was used by previous LARP6 researchers to identify the putative NES that we now know to be located within the RRM.[19] As the NES field matured, the Kosugi consensus sequence pattern became organized according to their amino acid sequence into three classes, with the first class further divided into four subcategories (Figure 5).[33, 34] A fourth class was recently identified by an independent group and bears similarities to the Class 3 Kosugi NES sequence (Figure 5). Experimentally confirmed examples of the first three classes of NES include PKIα (class I), HIVrev (class II), and Mad1 (class III).

Class 1a	ϕXXXϕXX-ϕX--ϕ
Class 1b	ϕXX-ϕXX-ϕX--ϕ
Class 1c	ϕXXXϕXXXϕX--ϕ
Class 1d	ϕXX-ϕXXXϕX--ϕ
Class 2	ϕX--ϕXX-ϕX--ϕ
Class 3	ϕXX-ϕXXXϕXX-ϕ
Class 4	ϕX--ϕXXXϕXXXϕ

ϕ- Leu, Val, Ile, Phe, Met
X- any amino acid

Figure 5: Nuclear Export Motifs: Consensus Sequences. Sequence alignment of Kosugi consensus patterns. These are defined intervals of hydrophobic residues with a preference for Leu, Val, Ile, Phe, or Met are shown as (ϕ) with any amino acid (X) separating the hydrophobic residues. Additionally, the (-) indicated no amino acid and is for alignment purposes only. Classes 1-3 are termed the "Kosugi consensus sequences"; Class 4 was recently reported by an independent group and bears the closest similarity to Kosugi Class 3.

While adherence to a general consensus sequence is a necessary feature of a NES, it is an insufficient predictor of functional nuclear export.[41] The shape of the putative cargo protein, its biological function(s), and the accessibility of predicted NES are all important elements that contribute to successful prediction of a biologically functional NES. Protein folding often buries hydrophobic residues in the core of the protein. This is a structure-stabilizing mechanism, maintaining the tertiary structure of the protein. While transient masking of the NES is a well-established mechanism to regulate transport, the burial of hydrophobic residues deep in the core of a protein fold is unlikely to be recognized by nuclear transport carriers.[35] The high incidence of hydrophobic residues in NESs can result in the false positive identification of putative NES that lie within buried elements of the tertiary protein structure or in other protein-protein interaction surfaces that rely on solvent-exclusion to mediate binding.

However, when sequence content is combined with experimentally determined structural data, a more accurate prediction of NES is possible. In particular, prediction algorithms that incorporate both experimentally confirmed NES consensus sequences and structural data, including factors such as disorder propensity and the predicted accessibility of the hydrophobic residues, are more robust computational predictors of new NES sequences.[36, 37] Several research groups have independently developed such prediction algorithms, including NetNES, NESsential, and Wregex.[38, 39, 40] Each of these programs employs different prioritization schemes that variably weight the adherence to consensus sequences, reliance on machine-learning algorithms, and scoring of either known or predicted biophysical features. Recently, one such prediction software, Locating Nuclear Export Signals or NESs (LocNES), was demonstrated to improve the accuracy and precision of NES identification.[41] This software uses a position-specific scoring matrix that evaluates the sequence content, consensus pattern, solvent accessibility, secondary structure, and predicted intrinsic disorder of a protein sequence to robustly identify NES motifs.[41] In addition to these features, LocNES algorithms are trained on experimentally confirmed NES to help reduce the number of false positive sequences identified.

Leptomycin B

Cancers have been demonstrated to have increased nuclear export of some proteins. One treatment option has been Leptomycin B (LMB). LMB functions as a therapeutic agent in cancer treatment by blocking nuclear export

of proteins.[28] Its clinical uses have been limited by its limited efficacy and toxicity, however, it is currently being explored as part of a combination therapy in the treatment of some lung cancers.[42] This is accomplished through a specific and selective inhibition of CRM1. This is a form of competitive inhibition and binding of LMB occurs within same binding pocket as the NES. LMB binds covalently to Cys-529 of CRM1, located within the hydrophobic cargo-binding pocket.[43] This covalent conjugation is dependent on the hydrolysis of LMB's lactone rings.[44] This specific and selective inhibition of CRM1 allows LMB to be used in cell culture experiments as a way to screen for proteins involved in nuclear export through CRM1-mediated pathways.

Identification of proteins involved in nuclear export is important because these proteins often have important regulatory functions in gene expression, signal transduction, and other critical regulatory processes. Confirmation of nucleocytoplasmic shuttling can assist in identification of the mechanism of a protein's activity. The basis of nuclear export is through interaction between CRM1 and the NES on the cargo protein. The identification of these sequences also gives important information about the proteins structure and function. Due to the regulatory role of LARP6 in the synthesis of type I collagen and its interaction with type I collagen mRNA, confirmation of nucleocytoplasmic shuttling behavior would provide significant insight into its function in the cell. Specifically, understanding the mechanism of cellular localization could potentially allow for drug targeting of localization as a form of treatment for diseases with excessive type I collagen expression. To probe the potential role that CRM1 might have in

the nuclear export of LARP6, LMB inhibition experiments were selected as the best way to test the hypothesis that CRM1 mediates the nuclear export of LARP6.

LARP6 Cellular Localization

Previous work identified both a putative NLS and a putative NES within human LARP6.[19] In this study, the putative NLS was defined as residues 271-305, and the putative NES was defined as residues 186-218. Localization of LARP6 was monitored by transfecting fluorescently-tagged recombinant wild-type LARP6 into a cell line that did not endogenously express LARP6 and viewing localization microscopically. They attempted to identify the NES by deleting their proposed NES (residues 186-218) and proposed NLS (residues 271-305) from the fluorescently-tagged recombinant protein and observing the cellular localization of the mutants. Wild-type LARP6 was predominantly observed in the cytoplasm, with diffuse signal in the nucleus (Figure 6). A deletion mutant that lacked the proposed NLS resulted in LARP6 accumulation in the cytosol. In contrast, the deletion of the proposed NES resulted in accumulation of LARP6 in the nucleus (Figure 6). These results suggested that LARP6 experiences nucleocytoplasmic shuttling. Cai *et al.* in 2010 showed that when LARP6 is bound to type I collagen mRNA in the cytoplasm, the translation of collagen type I is blocked; the same study also showed that LARP6 is bound to mRNA in nuclear extracts.[16] This is an indicator that LARP6 may be shuttling

in complex with type I collagen mRNA, though this has not been experimentally confirmed.

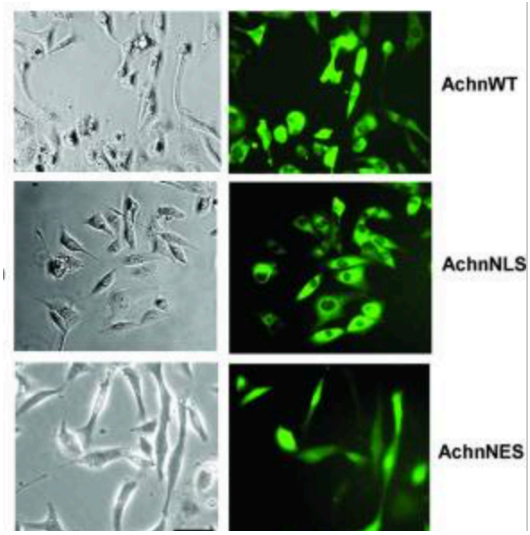

Figure 6: Altered localization of GFP-LARP6 mutants *in vivo*. LARP6-eGFP fusion proteins were recombinantly expressed in MDA-MB-231 cells. C-terminal GFP tagged LARP6 WT ("AchnWT"), LARP6 ΔNLS ("AchnNLS"), and LARP6 ΔNES ("AchnNES") were visualized by fluorescence microscopy. (Adapted from the *International Journal of Cancer*, reproduced with permission). [19]

At the time of this study, high-resolution structures were not available for the LARP6 La Module, and so the deletion mutants were constructed on the basis of sequence alone. In 2015, Martino *et al.* solved the solution NMR structures of the LARP6 La Motif and RRM.[20] The location of these deletions within the protein structure raises questions about the accessibility of the previously identified NES. The putative NES, residues 186-218, is located in the core of the RRM (Figure 7c). Specifically, this sequence constitutes the β1 strand and follows through the α0' helix through to the α1 helix (Figure 7a). The β1 strand is a central element in a β-sheet and its deletion could disrupt the folding and assembly of the β4 strand and the β3 strand that flank this element in the

native structure. Additionally, the β1 strand maybe structurally inaccessible to the binding pocket of CRM1 based on its positioning in the tertiary structure (Figure 7a). As described above, it is unlikely that the α helix that lies over the β1 strand would move to allow CRM1 binding.[20] The NLS deletion mutant faces similar stability concerns due to the large segment of protein deleted from the structure comprising the α2 helix through the β4 strand (Figure 7b).

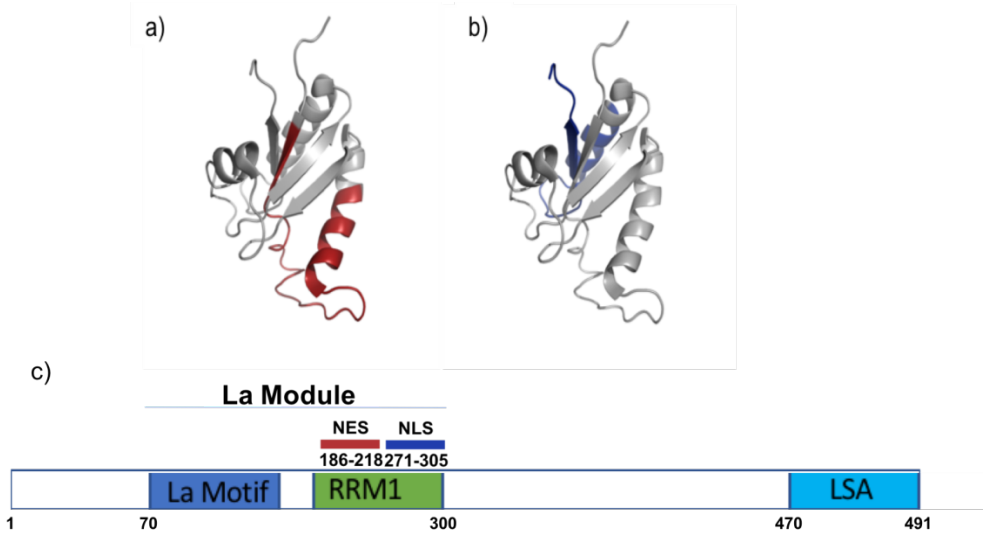

Figure 7: Location of putative localization motifs within the LARP6 RRM. The solution NMR structure of the LARP6 RRM (PDB: 2MTG) is used to map the putative localization motifs. (A) The putative NES identified by cell biological studies (residues 186-218) is marked in red. This sequence comprises the core beta strand, loop L1, and majority of helix a0'. (B) The putative NLS identified by cell biological work (residues 271-305) is highlighted in blue and comprises a beta strand on the edge of the beta sheet as well as part of a helix and the connecting loop L3. C) Topology of LARP6 and location of putative NLS/NES within the RRM. The putative NES is shown in red and the putative NLS is shown in blue. Figures (A) and (B) created in PyMol (Schrodinger, Inc.).[32]

Further validation of these putative localization elements is required using methods that are informed by the high-resolution structures. In particular, the deletion of ~30 amino acids that constitute the core of the RRM could be

structurally destructive, and as such the cellular localization results might need to be reconsidered in light of the high-resolution structural data. Alternatively, the location of an NES and an NLS within the RRM could indicate that nucleocytoplasmic shuttling may be dependent on RNA binding. Additionally, computational advances in NES prediction since the Shao studies allow for the refinement of NES motif prediction within human LARP6.

The identification of the NES in the previous work used CRM1 cargo NES consensus sequences as the predictive algorithm. However, there is no conclusive evidence for the participation of CRM1 in the localization of LARP6 *in vivo*. If LARP6 undergoes CRM1-mediated nuclear export, the previously identified NES should be bound by CRM1 *in vitr*o like other CRM1-recognized NESs. Similarly, inhibition of CRM1 *in vivo* should alter LARP6 localization. Finally, in light of the structural data informing where the putative NES is located in the RRM structure, the *in vivo* deletion mutants may have altered protein structure and/or RNA binding activity.

Confirmation of CRM1-mediated nuclear export can be performed using both *in vitro* reconstituted systems and in cell culture, using human cell lines that endogenously express LARP6. When combined with fluorescent labelling of LARP6 through the use of commercial antibodies, cellular localization and disruption of localization can be monitored using confocal microscopy. The overall aims of this project are to confirm the pathway of nuclear export of LARP6, experimentally confirm the NES involved in nuclear export of human LARP6, and to test deletion mutants for stability of the folded protein.

Aim 1. Is LARP6 an *in vivo* cargo for CRM1? We propose to test whether LARP6 is cargo for CRM1-mediated nuclear export by inhibiting CRM1 function using Leptomycin B treatment. Competitive inhibition of CRM1 using Leptomycin B results in nuclear accumulation of CRM1 cargo proteins. Therefore, if this exportin is responsible for cycling out of the nucleus, LMB treatment allows for confirmation or exclusion of the CRM1 mediated nuclear export pathway for LARP6 localization. Endogenous LARP6 protein localization will be observed in HeLa and HEK-293 cells, using immunocytochemistry with commercial antibodies, in the absence and presence of Leptomycin B treatment.

Aim 2. Test LocNES-predicted putative NES motifs in LARP6 for CRM1 binding *in vitro*. Identification of alternative NES motifs will be performed using LocNES, an NES prediction software that informs its algorithms on experimentally confirmed NES databases. Experimental tests of the LocNES-predicted motifs will be performed in several ways. First*, in vitro* binding assays of recombinantly expressed CRM1 and Ran-GTP will be performed using recombinantly expressed, GST-tagged fusion proteins of putative NES. This method will then be further used to identify point mutations of hydrophobic residues in the predicted NESs that disrupt binding to CRM1. With these NES-disrupting mutations, one can either test for CRM1 binding using *in vitro* pull-down assays with full-length mutant LARP6 or transfect into cell culture and observe localization.

Aim 3. Assess impact of *in vivo* deletion mutants on LARP6 structural stability. Additionally, we hypothesize that the structural distortion caused by the deletion of LARP6(186-218) and LARP6(271-305) may be responsible for the previously observed disruption of nuclear localization and export. Therefore, we will measure the effect of deleting the putative NESs within full-length LARP6 using two *in vitro* assays that probe structure. First, we will use overlap-extension PCR to construct deletion mutants of all predicted nuclear export sequences in full-length LARP6, recombinantly express and purify them. If the proteins are able to be purified, they will then be evaluated for structural stability using limited proteolysis and differential scanning fluorimetry assays to determine if the deletions compromise the structural integrity of the full-length protein relative to wildtype LARP6.

II. MATERIALS AND METHODS

Oligonucleotides and PCR Primers

All oligonucleotides (Appendix I, Table 1) and primers (Appendix I, Table 2) were designed in I for this project and ordered from Integrated DNA technologies (IDT). All of the oligonucleotides were designed for insertion into the pGEX4T3-TEV plasmid between a N-terminal EcoRI endonuclease recognition site and a C-terminal XhoI endonuclease recognition site. Oligonucleotides were reconstituted in nuclease-free H_2O (IDT) to a final concentration of 100 μM. Primers were diluted to a working stock concentration of 20 μM with IDT nuclease free H_2O. All DNA stock solutions were stored at -20 °C.

General Protocols

Transformation into cloning lines of E. coli – Transformations were carried out using 1-7 μL ligation reaction mixed with 50 μL of homemade DH5a ultracompetent *E. coli* cells (prepared per reference 45). This mixture was incubated on ice for 30 minutes, heat shocked at 37 °C for 1 minute 30 seconds, then incubated on ice for 2 minutes. To recover transformed cells, 500 – 1000 μL of sterile Lennox Broth was added and the sample was incubated at 37 °C for 1 hour while shaking. 50 μL of the recovering culture was plated on half of a Lennox Broth agar plate with antibodies indicated by the expression vector and cell line (Tables 1 and 2). The remaining cells were pelleted by centrifugation at max speed for 1 min at room temperature. The pelleted cells were resuspended in 100 μL of reserved supernatant and were plated on the other half of the

Lennox Broth agar plate. The plated cells were incubated at 37 ºC overnight, and the plates were wrapped in Parafilm for short-term storage at 4 ºC. An overnight culture was performed by using a single colony to inoculate 5 mL of Lennox broth and growing it overnight at 37 °C in the presence of appropriate antibodies (Table 1 and 2). The plasmids were then isolated using a commercial miniprep kit (QIAGEN, 27106).

Table 1: Plasmids and Antibiotic Resistance

Vector	Antibiotic Resistance
pGEX4T3-TEV	Ampicillin (50 µg/mL)
pET28-SUMO	Kanamycin (50 µg/mL)

Table 2: *E. coli* Cell Lines and Antibiotic Resistance

Cell Line	Antibiotic Resistance
DH5α	N/A
BL-21(pLys) (DE3)	Chloramphenicol (34 µg/mL)
BL21	N/A
Rosetta (DE3)	Chloramphenicol (34 µg/mL)

Transformation into protein expression lines of E. coli – Transformations were carried out using 1-5 µL miniprepped DNA (30 – 120 µg/mL) mixed with 50 µL competent cells (prepared as per reference 46). This mixture was incubated on ice for 30 minutes, heat shocked at 37 °C for 45 seconds, then incubated on ice for 2 minutes. To recover transformed cells, 500 – 1000 µL of sterile Lennox BROTH was added and the sample was incubated at 37 °C for 1 hour while shaking. 50 µL was plated on half of a Lennox broth agar plate with antibodies indicated by the expression vector and cell line (Tables 1 and 2). The remaining cells were pelleted by centrifugation at max speed for 1 min at room temperature.

The pelleted cells were resuspended in 100 μL of reserved supernatant and were plated on the other half of the Lennox Broth agar plate. The plated cells were incubated at 37 ºC overnight, and the plates were wrapped in Parafilm for short-term storage at 4 ºC.

SDS-PAGE Electrophoresis – Samples were prepared in 1X Sodium Dodecyl Sulfate (SDS) sample buffer (5X: 0.5 M Tris [pH 7.5 at 4 °C], 4 mM β-mercaptoethanol, 0.4 M SDS). Gel electrophoresis was performed using 10% SDS-PAGE gels (37.5:1 acrylamide:bis-acrylamide; ProtoGel, National Diagnostics), run in 1X Tris-Gly Running buffer (50 mM Tris and 0.5 M glycine, 0.4 M SDS) at 200 V for approximately 1 hour. PageRuler Pre-stained Ladder (Thermo Scientific) was used to determine apparent molecular weight. Unless otherwise noted, lanes contained either 5 μL of ladder or 10 μL of each sample.

Coomassie Staining – After gel electrophoresis, some SDS-PAGE gels were Coomassie stained to detect total protein. The gel was placed in Coomassie Blue Stain (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 H_2O) and incubated at room temperature for 20-30 minutes with shaking. The stain was then decanted and replaced with Coomassie Destain (40% (v/v) methanol, 10% (v/v) glacial acetic acid, and 50% (v/v) MQ H_2O) and allowed to incubate with shaking until visible bands were detected. Kim-wipes (Kimberly-Clark) were added to the container during the de-stain process to

accelerate the de-staining. This is accomplished by their ability to retain the stain lowering its concentration in the detaining solution and increasing the rate of diffusion out of the gel. Imaging of gels were performed on a Chemi-Doc XRS using the “Coomassie stain” settings.

Anti-GST Western Blot – To detect GST tags, proteins were transferred from the completed SDS-PAGE to 0.2 µm nitrocellulose membranes on the Trans-Blot module (Bio-Rad) using the “Mixed Molecular Weight” protocol (25 V, 1.3 A, 7 min). Unless otherwise noted, the anti-GST blot was performed as follows. The nitrocellulose membrane was blocked in a solution of 5% dry milk (w/v) in 1X TBS-T (1X TBS: 50 mM Tris, 150 mM NaCl, pH 7.5, with 0.05% Tween 20) for 1 hour at room temperature. Following a brief wash with TBS-T, the membrane was incubated with a 1:2000 dilution of anti-GST monoclonal antibody (Thermo Fisher Scientific, MA4-004, BB242221) in TBS-T at 4 ºC overnight with shaking or at room temperature for 2 hours. This incubation was followed by three washes in TBS-T were performed for 5 min each, at room temperature. The membrane was then incubated with a 1:20,000 dilution of goat anti-mouse secondary antibody conjugated with horse radish peroxidase (Jackson ImmunoResearch, 115-035-003) in TBS-T for 1 hour at room temperature. Three washes in 1X TBS-T were then performed for 5 min each, shaking at room temperature. This was followed by two washes in 1X TBS for 5 min each, shaking at room temperature. HRP was detected by incubating for 2 min at room temperature with chemiluminescence reagent (100 mM Tris-HCl pH

8.8, 1.25 mM luminol, 2 mM 4-Iodophenylboronic acid) that was activated with 30% H_2O_2.[47] The nitrocellulose membrane was drained before being imaged on the Chemi Doc XRS. Using the "Chemi Blot" setting, the first image was taken at 5 seconds, and the last image recorded at 125 seconds for a total of 8 images. The image with the strongest signal to background ratio without saturating any pixels was chosen for retention and analysis.

Anti-LARP6 Western Blot – To detect LARP6, proteins were transferred from the completed SDS-PAGE to 0.2 µm nitrocellulose membranes on the Trans-Blot module (Bio-Rad) using the "Mixed Molecular Weight" protocol (25 V, 1.3 A, 7 min). Unless otherwise noted, the anti-LARP6 blot was performed as follows. The nitrocellulose membrane was blocked in a solution of 5% dry milk (w/v) in 1X TBS-T for 1 hour at room temperature. Following a brief wash with TBS-T, the membrane was incubated with a 1:1000 dilution of mouse anti-*LARP6* polyclonal antibody (Novus, H00055323-B01P, H8211) in TBS-T at room temperature for 1 hour with shaking. This incubation was followed by three washes in TBS-T were performed for 5 min each, at room temperature. The membrane was then incubated with a 1:20,000 dilution of goat anti-mouse secondary antibody conjugated with horse radish peroxidase (Jackson ImmunoResearch, 115-035-003) in TBS-T for 1 hour at room temperature. Three washes in 1X TBS-T were then performed for 5 min each, shaking at room temperature. This was followed by two washes in 1X TBS for 5 min each, shaking at room temperature. HRP was detected by incubating for 2 min at room

temperature with chemiluminescence reagent that was activated with 30% H_2O_2 as previously described.[47] The nitrocellulose membrane was drained before being imaged on the ChemiDoc XRS as described above.

Anti-His Blot- To detect His tagged proteins, proteins were transferred from the completed SDS-PAGE to 0.2 μm nitrocellulose membranes on the Trans-Blot module (Bio-Rad) using the "Mixed Molecular Weight" protocol (25 V, 1.3 A, 7 min). The nitrocellulose membrane was blocked in a solution of 5% dry milk (w/v) in 1X TBS-T for 1 hour at room temperature. The membrane was incubated with a 1:20000 dilution of anti-His probe (Pierce, 15165, lot no) for 1 hour at room temperature with shaking. This incubation was followed by three washes in TBS-T performed for 5 min each at room temperature. This was followed by two washes in 1X TBS for 5 min each, shaking at room temperature. HRP was detected as described above for the other blots.

Calibration of size exclusion chromatography columns – To ensure proper isolation and characterization of all proteins used in this work, the calibration of both the Sephadex S200 size exclusion column and the Sephadex S75 size exclusion column was performed. The S200 column was previously calibrated. The S75 size was calibrated specifically for this work. This was performed through loading mixes of protein standards and recording the volumes at which they eluted (V_e). The blue dextran elution, as the largest protein in the mix and therefore the first to elute was established as the void volume (V_0). The total

volume of the column (V_t) was 120 mL. The partition coefficient (K_{av}) was determined using the relationship $K_{av} = (V_e - V_0) / (V_t - V_0)$. This was plotted against the log molecular weight of the protein to find the relationship between the K_{aV} (y) and molecular weight (x). This relationship can be described as y=-0.5193x+2.6091 for the S75 column and y=-0.3788x+2.2685 for the S200 column (Figures 8 and 9). The R^2 values for the plotted molecular weights and K_{av} for the S75 and S200 columns were respectively 0.9929 and 0.9973 indicating the columns are both in good working condition for protein purification and characterization. The sets of protein standards for each column are listed in Table 3 and 4.

Table 3: Size Exclusion Chromatography S75 Standards symbols

	MW	Elution Volume
Blue Dextran	2,000 kDa	46.25 mL
Conalbumin	75 kDa	53.75 mL
Ovalbumin	44 kDa	59.37 mL
Carbonic Anhydrase	29 kDa	67.16 mL
RNAse A	13.7 kDa	80.30 mL
Aprotinin	6.5 kDa	93.27 mL

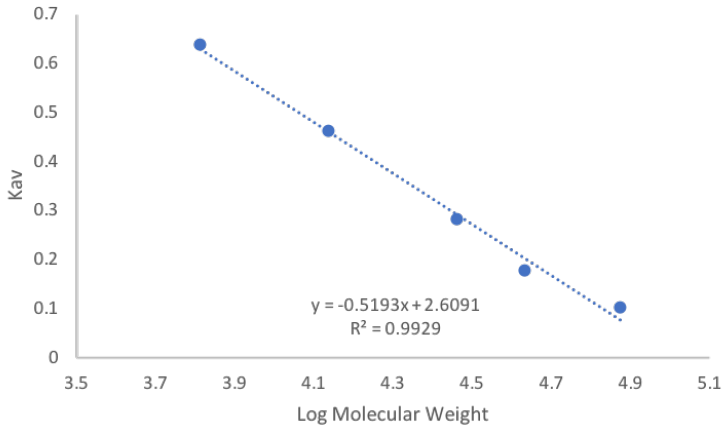

Figure 8: Size Exclusion Chromatography S75 Standardization Curve. Protein standards were loaded onto the FPLC and run on the S75 column. Their elution volumes were compared to the void volume (blue dextran elution) and the total volume of the column to determine the K_{av}. The K_{av} showed a strong linear relationship with the log of the protein standards molecular weight as evidenced by the R^2 value of 0.9929. The relationship of K_{av} to the logarithm of the molecular weight is fitted with the linear regression y=-0.5193x+2.6091.

Table 4: Size Exclusion Chromatography S200 Standards symbols

	MW	Elution Volume
Blue Dextran	2,000 kDa	51.48 mL
Ferritin	440 kDa	60.1 mL
Aldolase	158 kDa	72.85 mL
Conalbumin	75 kDa	80.1 mL
Ovalbumin	44 kDa	85.43 mL
Carbonic Anhydrase	29 kDa	91.98 mL
RNase A	13.7 kDa	99.38 mL

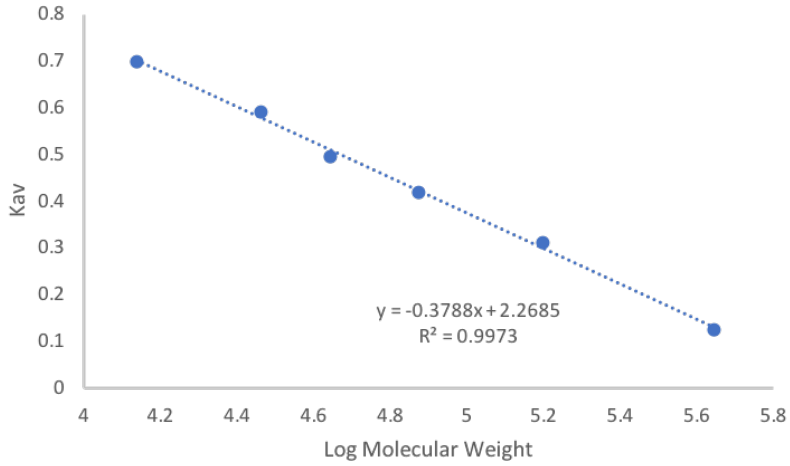

Figure 9: Size Exclusion Chromatography S200 Standardization Curve. Protein standards were loaded onto the FPLC and run on the S200 column. Their elution volumes were compared to the void volume (blue dextran elution) and the total volume of the column to determine the K_{av}. The K_{av} showed a strong linear relationship with the log of the protein standards molecular weight as evidenced by the R^2 value of 0.9973. The relationship of K_{av} to the logarithm of the molecular weight is fitted with the linear regression y = -0.3788x + 2.2685.

Cloning of GST-Peptide Fusions

Cloning of LocNES-identified and control NES sequences – Oligonucleotides were designed using I software with *Eco*RI and *Xho*I compatible overhangs were ordered from IDT for annealing and direct ligation into the pGEX4T3-TEV vector. Oligonucleotides were reconstituted at a concentration of 20 μM in nuclease-free H_2O (IDT) then mixed at a 1:1 ratio. Annealing was performed by heating the mixtures to 90 °C for 2 min, then slow-cooled to room temperature for one hour. Twenty micrograms of pGEX4T3-TEV vector was prepared for direct ligation of the putative NES constructs by double digest performed at 37 °C for 1 hour with 1.5 units of each of the restriction endonucleases *Eco*RI-HF and *Xho*I. The vector was then treated with 5 units of Antarctic Phosphatase (NEB) for one hour at 37 °C. After phosphatase treatment, the double digest was confirmed by agarose gel and then gel purified. Gel purification of digested DNAs was performed in 0.8-0.9% TAE-agarose gels containing 0.01% (v/v) ethidium bromide and electrophoresed at 90 V for 60 min in 1X TAE buffer (40 mM Tris, 20 mM Acetic acid, 1 mM EDTA). Bands were visualized on the Chemi-Doc XRS using the “Ethidium Bromide” settings, excised with a clean razor blade, and the DNA extracted from gel matrix using E.N.Z.A Gel Extraction kit (Omega Bio-tek). Ligation was performed by mixing the purified/digested DNAs (ranging from molar ratios of vector: insert 1:1 to 1:9) and an equal volume of 2X T4 Quick Ligase Buffer (NEB), adding T4 Quick Ligase (NEB), and the reaction incubated at room temperature for 5 minutes.

Cloning of putative NES LARP6(186-218) for ligation into pGEX4T3-TEV vector – Primers were designed using Snapgene software to flank the coding sequence for amino acids 186-218, with *Eco*RI and *Xho*I restriction sites engineered into the primers. These were ordered from IDT for PCR amplification from the full-length pET28:*LARP6* template. PCR reaction conditions were screened by gel electrophoresis. Successful PCR reactions were gel purified, double digested with *Eco*RI-HF and *Xho*I, and again gel-purified as previously described.

Expression and Purification of GST-Peptide Fusions

Expression Trials – A single colony from a transformation of sequence confirmed GST-putative NES/ NES was picked from the Lennox broth-agar plate and used to inoculate 5 mL Lennox Broth with 50 µg/mL ampicillin and 34 µg/mL chloramphenicol. The culture was incubated overnight (16-18 hours) at 37 °C while shaking. One milliliter of overnight culture was added to a 100 mL Lennox Broth/Amp/Chlor and incubated at 37 ºC with shaking. The optical density at 600 nm (OD_{600}) was measured until the cultures were between 0.5 and 0.6, at which time they were cold shocked by 10 min incubation on ice. Induction of expression was performed using 1 mM final volume isopropyl b-D-1-thiogalactopyranoside (IPTG). Aliquots were taken at indicated timepoints during the induction period and separated on a 10% SDS-PAGE gel. A Coomassie stain and anti-GST western blot were performed for each trial expression trial to detect expressed protein.

Preparative Expression Cultures – Overnight cultures were prepared using 50 mL Lennox Broth/Amp/Chlor by selecting a single colony from a plate of transformed Rosetta cells and incubated overnight at 37 ºC with shaking. A 1 L expression culture containing antibiotics was inoculated with 10 mL from the overnight culture, and the OD_{600} monitored until the cultures reached 0.5-0.6. Following cold shock on ice for 10 min, expression was induced with 1 mM final concentration IPTG. Expression was carried out at 30 °C for 5 hours. Cells were harvested by centrifugation (5000 x*g*, 8 min, 4°C), transferred to conical vials, and stored at -20 °C.

Lysis and glutathione affinity chromatography – Cell pellets from the 1 L preparative expressions were resuspended in GSH Lysis Buffer (Table 5) containing protease inhibitors (Pierce, PIA32965). Cells were lysed by sonication in an ice-water bath at 30% amplitude for 1 pulse per second for 20 seconds with 30 second rest periods for a total of 6 cycles (2 minutes active sonication time). Cell debris was pelleted by centrifugation (5000 x*g*, 8 min, 4 °C). The clarified supernatant was then incubated with glutathione resin at room temperature for 30 min. Flow through was collected after centrifugation at 500 x*g* for 5 min at room temperature. A series of three washes was then performed by resuspending the beads in 25 mL of Wash Buffer (5:1 buffer:50%slurry) (Table 5), followed by centrifugation at 500 x*g* for 5 min. Elution was performed by resuspending the beads in Elution Buffer (Table 5) by centrifugation at 500 x*g* for 5 min each for a total of three elution fractions. All of the purification buffers

contained dithiothreitol (DTT), a reducing agent that protects against the oxidation of thiol groups. All fractions were analyzed by SDS-PAGE and either Coomassie staining or GST blot. The elution fractions with the highest protein content were pooled and prepared for size exclusion chromatography.

Table 5: Buffers for batch purification and SEC of GST-NES peptides

Buffer	Content
Cell Lysis	1x phosphate buffered saline, pH 7.3 with 1 Pierce protease inhibitor tablet
Binding Buffer	140 mM sodium chloride, 2.7 mM potassium chloride, 10 mM sodium phosphate dibasic, 1.8 mM potassium phosphate monobasic, 2 mM DTT, pH 7.3
Elution Buffer	50 mM Tris-HCl, 10 mM reduced glutathione, 5 mM DTT, pH 8.0

Table 6: Buffers for batch purification and SEC of GST-NES peptides with 5% Glycerol

Buffer	Content
Cell Lysis	1x phosphate buffered saline, pH 7.3 with (1 tab/ 50 mL) Pierce protease inhibitor tablet
Binding Buffer	140 mM sodium chloride, 2.7 mM potassium chloride, 10 mM sodium phosphate dibasic, 1.8 mM potassium phosphate monobasic 5% glycerol, 2 mM DTT, pH 7.3
Elution Buffer	50 mM Tris-HCl, 10 mM reduced glutathione, 5% glycerol, 5 mM DTT pH 8.0

Size exclusion chromatography – All samples were concentrated by centrifugation (5000 x*g,* 4 °C) in 5 min intervals using a Viva-Spin 5,000 MWCO centrifugal concentrator (Sartorius). Care was taken to gently mix the retentate between runs via aspiration to minimize protein concentration gradients and to monitor protein aggregation. Samples were then filtered using a 0.2 μm syringe filter before loading onto a 16/300 S75 Superdex column (GE Life Sciences), using the ÄKTA Pure L fast performance liquid chromatography system. The flow

rate was 1 mL/min. Samples under the peak were analyzed using SDS-PAGE and Coomassie stained to confirm the presence of the target protein and to select appropriate fractions. Fractions selected were pooled and snap frozen in liquid nitrogen then stored at -70 °C.

Cloning of GST-LARP6

Cloning- LARP6 was cloned by PCR using the pET28:*LARP6* as the template. Confirmation of successful PCR by gel electrophoresis on a 1.2% Agarose gel stained by ethidium bromide and imaged as previously described. Gel purification was performed to isolate the target DNA and a double digest was performed with 1.5 units each *Sal*I and *Not*I at 37 °C for 1 hr. A gel purification was performed as previously described. The 20 μg of the pGEX4T3-TEV plasmid was double digested as described above and a phosphatase treatment was performed as previously described. Ligation reactions, and transformations into DH5α cells were performed as previously described. PCR was used to confirm the presence of the wild-type LARP6 in the pGEX4T3-TEV plasmid and analyzed by gel electrophoresis before samples were sent for commercial Sanger sequencing(Genewiz).

Expression and Purification of GST-TEV-*Hs*CRM1

Cloning -The plasmid containing *Hs*CRM1 behind an N-terminal GST tag with an intervening TEV cleavage site was a generous gift from Dr. Yuh Min Chook (University of Texas – Southwestern Medical Center at Dallas). The

plasmid containing the *Hs*CRM1 gene was confirmed by commercial Sanger sequencing (Genewiz). This plasmid was then transformed into DH5α, BL-21, BL-21(pLys), and Rosetta cell lines as previously described for amplification and/or expression.

Expression trials – Multiple expression trials were performed to examine the effects of cell line, pH/salt content of the media broth, cold shock, and aeration on expression. In all cases an overnight culture was obtained from a fresh transformation plate for approximately 18-20 hours in Lennox Broth. One mL of the overnight culture was used to inoculate 100 mL Lennox Broth with the appropriate antibiotics (Tables 1 and 2) and grown at 37 ºC. When OD_{600} reached 0.8-0.9. The cells were then cold shocked unless otherwise noted and induction of expression with 1mM IPTG was performed. Cells were then grown at 25 °C for 10 hours.

Preparative expression cultures – Fresh transformations in to BL21 (pLys) E. coli cells were performed before every expression these were done on Lennox Broth Agar plates containing ampicillin and chloramphenicol (Table 1 and 2) and grown for 15-18 hours at 37 ºC with shaking. A colony was selected from the plate and used to inoculate 100 mL Lennox Broth with ampicillin and chloramphenicol and grown at 37 °C for 16-18 hours. In early trials, 1L expression cultures were made with capsulated Luria Broth, in later trials this was modified to only use granulated Lennox Broth, as discussed in the Results

section. Cells were grown at 37 °C until OD_{600} readings were ~ 0.5-0.6. A cold shock was performed on ice for 10-20 min and expression was induced with 1 mM final concentration IPTG. Expression was carried out at 25 °C for 10 hours. Cells were harvested by centrifugation (5000 x*g*, 10 min, 4 °C).

Lysis and glutathione affinity chromatography – Cell pellets from the 1 L preparative expressions were resuspended in *Hs*CRM1 Cell Lysis Buffer (Table 7) containing a protease inhibitor tablet (Pierce, PIA32965). Cells were lysed by sonication in an ice-water bath at 30% amplitude for 1 pulse per second for 20 seconds with 30 second rest periods for a total of 6 cycles (2 minutes active sonication time). Cell debris was pelleted by centrifugation (5000 x*g*, 8 min, 4 °C).

The clarified supernatant was then incubated with glutathione resin at room temperature for 30 min while shaking. The unbound protein was collected as the supernatant after centrifugation at (500 x*g*, 5 min, 4 °C). A series of three washes was then performed by resuspending the beads in 25 mL of Wash Buffer (5:1 v:v ratio of buffer to 50% bead slurry) (Table 7), followed by centrifugation at 500 x*g* for 5 min. Elution was performed by resuspending the beads in Elution Buffer (Table 7) by centrifugation (500 x*g,* 5 min, 4 °C) and collecting the supernatant and repeating 2x for a total of three elution fractions. All fractions were analyzed by SDS-PAGE and either Coomassie staining or anti-GST western blot. The elution fractions with the highest protein content were pooled and prepared for size exclusion chromatography.

Size exclusion chromatography – All samples were concentrated by centrifugation at 5000 x*g* in 5 min intervals using a Viva-Spin 10,000 MWCO centrifugal concentrator (Sartorius). Care was taken to gently mix the retentate between runs to minimize protein concentration gradients and to monitor protein aggregation. Samples were then filtered using a 0.2 μm syringe filter before loading onto a 16/300 S200 Superdex column (4thGE Life Sciences), using the ÄKTA Pure L fast performance liquid chromatography system. The flow rate was 1 mL/min. Samples under the peak were run on an SDS-PAGE gel and Coomassie stained to confirm the presence of the target protein and to select appropriate fractions. Fractions selected were pooled and snap frozen in liquid nitrogen then stored at -70 °C.

Table 7: *Hs*CRM1 Purification Buffers

Buffer	Content
Cell Lysis	40 mM HEPES (pH 7.5) 2 mM magnesium acetate, 200 mM sodium chloride, 10% glycerol, 10 mM DTT, (1 tab/ 50 mL) pierce protease inhibitor tablet
GF (10 mM HEPES)	10 mM HEPES (pH 7.5) 5 mM magnesium acetate, 100 mM sodium chloride, 10% glycerol, 2 mM DTT
GF (10 mM HEPES + 200 mM NaCl)	10 mM HEPES (pH 7.5) 5 mM magnesium acetate, 200 mM sodium chloride, 10% glycerol, 2 mM DTT
GF (10 mM HEPES + 1 mM DTT)	10 mM HEPES (pH 7.5) 5 mM magnesium acetate, 100 mM sodium chloride, 10% glycerol, 1 mM DTT
ATP Wash	20 mM HEPES (pH 7.5) 5 mM magnesium acetate, 100 mM sodium chloride, 20% glycerol, 2 mM DTT, 5mM ATP, (1 tab/ 50 mL) pierce protease inhibitor tablet
GF	20 mM HEPES (pH 7.5) 5 mM magnesium acetate, 100 mM sodium chloride, 10% glycerol, 2 mM DTT

Expression and Purification of Ran-GTP

Cloning -The pET21d GSP1 Q71L (Ran-GTP) plasmid was a generous gift from Dr. Yuh Min Chook (UT-Southwestern). This plasmid was transformed into Rosetta cells as described above.

Preparative expression cultures – A colony was selected from a Rosetta transformation plate and used to inoculate overnight cultures of Luria Broth with ampicillin and chloramphenicol, which was then grown to saturation at 37 °C overnight. The inoculum cultures were grown at 37 °C until OD_{600} readings were ~ 0.5-0.6. A cold shock was performed on ice for 15 min and expression was induced with 0.5 mM IPTG. Expression was performed at 20 °C for 11 hours. Cells were harvested by centrifugation at 5000 x*g* for 10 min at 4 °C.

Lysis and glutathione affinity chromatography – All 4 of the cell pellets from the 4- 1 L preparative expressions were resuspended 25 mL of the Ran-GTP Cell Lysis Buffer (Table 8) containing protease inhibitors (Pierce, PIA32965). Cells were lysed by sonication in an ice-water bath at 30% amplitude for 1 pulse per second for 30 seconds with 30 second rest periods for a total of 4 cycles (2 minutes active sonication time). After sonication the lysate from the four 1 L cell pellets was combined and was purified together. Cell debris was pelleted by centrifugation (20,000 x*g*, 60 min, 4 °C).

The clarified supernatant was then incubated with nickel beads at 4 °C for 1 hour with shaking. The nickel resin, and lysate mixture was decanted into the flex column and the resin was allowed to settle. The first wash was performed with 50 mL of the High Salt Wash 1 Buffer (Table 8) and collected in 25 mL fractions. A second wash was performed with 50 mL of Low Salt Wash 2 Buffer (Table 8) and collected in 25 mL fractions. A third wash was performed with 50 mL of the 20 mM Imidazole Wash 3 Buffer (Table 8) and collected in 25 mL fractions. The first set of elutions were performed with 50 mL of the Elution 1 Buffer (Table 8) and collected in approximately 15 mL fractions. A second set of elutions was performed with 75 mL of the Elution 2 Buffer (Table 8) and collected in 25 mL fractions. All fractions were analyzed by SDS-PAGE and either Coomassie staining or anti GST blot. The elution fractions with the highest protein content were pooled and prepared for size exclusion chromatography.

Size exclusion chromatography – All samples were concentrated by centrifugation (4000 x*g*, 4 °C) in 5 min intervals using a Viva-Spin 10,000 MWCO centrifugal concentrator (Sartorius). Care was taken to gently mix the retentate via aspiration between runs to minimize protein concentration gradients and to monitor protein aggregation. Conversion from Ran-GDP to Ran-GTP was performed in two steps. Incubation with the Ran-GTP Conversion Buffer 1 (Table 8) for 30 min on ice followed by a 30 min incubation on ice with the Ran-GTP Conversion Buffer 2 (Table 8). Samples were then filtered using a 0.2 μm syringe filter and diluted with 30 mL of the SPA buffer (Table 8) before injecting onto the Superloop and run through the HiTrap SP FF Sepharose column (GE Life Sciences), using the ÄKTA Pure L fast performance liquid chromatography system. System was equilibrated with the SPA Buffer (Table 8) and run with the SPB buffer (Table 8). Samples under the peak were run on an SDS-PAGE gel and Coomassie stained to confirm the presence of the target protein and to select appropriate fractions. Fractions selected were pooled, and then repeatedly concentrated and diluted with Buffer Exchange Buffer (Table 8) to exchange into storage buffer before being snap frozen in liquid nitrogen, then stored at -70 °C.

Table 8: Ran-GTP Purification Buffers

Buffer	Content
Cell Lysis	50 mM HEPES (pH 7.4) 2 mM magnesium acetate, 500 mM sodium chloride, 10% glycerol, 2 mM DTT, 5 mM Imidazole, (1 tab/ 50 mL) Pierce protease inhibitor tablet
High salt	50 mM HEPES (pH 7.4) 2 mM magnesium acetate, 50 mM sodium chloride, 10% glycerol, 2 mM DTT, 5 mM Imidazole
Low Salt	50 mM HEPES (pH 7.4) 2 mM magnesium acetate, 50 mM sodium chloride, 10% glycerol, 2 mM DTT, 5 mM Imidazole
20 mM Imidazole	50 mM HEPES (pH 7.4) 2 mM magnesium acetate, 50 mM sodium chloride, 10% glycerol, 2 mM DTT, 20 mM Imidazole
Elution 1	50 mM HEPES (pH 7.4) 2 mM magnesium acetate, 50 mM sodium chloride, 20% glycerol, 2 mM DTT, 5mM ATP, 200 mM Imidazole
Elution 2	50 mM HEPES (pH 7.4) 2 mM magnesium acetate, 50 mM sodium chloride, 10% glycerol, 2 mM DTT, 500 mM Imidazole
RanGTP Conversion Buffer 1	0.65 mM EDTA
RanGTP Conversion Buffer 2	0.65 mM EDTA, 20 mM magnesium acetate
SEC Buffer SPA	20 mM HEPES (pH 7.4), 4 mM magnesium acetate, 10% glycerol, 1 mM DTT,
SEC Buffer SPB	20 mM HEPES (pH 7.4), 4 mM magnesium acetate, 1mM sodium chloride, 10% glycerol, 1 mM DTT,
Exchange Buffer	20 mM HEPES (pH 7.4), 4 mM magnesium acetate, 100 mM sodium chloride, 10% glycerol, 1 mM DTT,

NES deletion mutants

Cloning – Primers for overlap extension polymerase chain reaction (OE-PCR) were created using the computer program SnapGene to generate deletion mutants of the previously identified putative NES (LARP 186-218) and deletion mutants of each of the LocNES-identified putative NES: LARP6(179-193), LARP6(210-224), LARP6(457-471) (Table 1). For each of these NES deletion constructs, an initial PCR reaction was carried out using the pET28-*LARP6*(Δ*Bam*HI) construct as a template. This plasmid contains a silent mutation in the coding sequence for the N-terminal region to ablate an endogenous *Bam*HI site, allowing for the use of *Bam*HI in subsequent cloning. An initial PCR reaction was carried out to create two fragments of the LARP6 DNA encompassing the area on either side of the deletion. These were confirmed by gel electrophoresis and gel-purified on a 0.8% agarose-TAE gel, using a commercial gel extraction kit to recover the fragments (Qiagen). The fragments created in the first PCR reaction were then used as megaprimers/templates for the OE-PCR. This PCR product was then purified using a PCR clean-up kit (Qiagen). Another round of PCR was used to amplify the OE-PCR products. This was confirmed by gel electrophoresis and gel-purified. A final round of PCR was used to add sequence for restriction enzyme-based cloning into pET28-SUMO vector, including a 5' *Bam*HI cut site and a 3' *Not*I cut site.

Trial Expression of SUMO-LARP6 Proteins – Trial expressions were performed in 100 mL of Lennox Broth with kanamycin and chloramphenicol

(Table 1 and 2). Inoculation with 1 mL of overnight culture was performed as previously described. Cells were grown at 37 °C until OD_{600} readings were ~ 0.5-0.6. A cold shock was performed on ice for 15 min and expression was induced with 1 mM final concentration IPTG. Expression was performed at 16 °C and 18 °C for 24 hours with time point aliquots taken throughout. Timepoint aliquots were analyzed by SDS-PAGE electrophoresis and either Coomassie Stain or anti-His western blots.

Preparative Expression of SUMO-LARP6 Proteins – Preparative expressions were performed in 1 L of Luria Broth with kanamycin and chloramphenicol (Table 1 and 2). Inoculation with 1 mL of overnight culture was performed as previously described. Cells were grown at 37 °C until OD_{600} readings were ~ 0.5-0.6. A cold shock was performed on ice for 15 min and expression was induced with 1 mM IPTG. Expression was performed at 16 °C for 21 hours with time point aliquots taken before induction and before cell harvest. Timepoint aliquots were analyzed by SDS-PAGE electrophoresis and either Coomassie staining or anti-His western blots. Cells were harvested by centrifugation (5000 x*g*, 10 min, 4 °C).

Lysis and Nickel Affinity Chromatography of SUMO-LARP6 Proteins – Cell pellets from the 1 L preparative expressions were resuspended in Cell Lysis Buffer (Table 10) containing protease inhibitors (Pierce, PIA32965). Cells were lysed by sonication in an ice-water bath at 30% amplitude for 1 pulse per second

for 20 seconds with 30 second rest periods for a total of 6 cycles (2 minutes active sonication time). Cell debris was pelleted by centrifugation (18,000 x*g*, 15 min, 4 °C). The clarified supernatant was then incubated with Ni^{2+}- NTA beads at 4 °C for 1 hour. The cell lysate resin mixture was then applied to a flex column and allowed to settle at 4 °C. Flow through was collected in two 15 mL fractions. Wash 1 buffer (40 mL, Table 10) was applied to the column. Wash 1 was collected in 4-10 mL fractions. Wash 2 buffer (40 mL, Table 10) was applied to the column. Wash 2 was collected in 4-10 mL fractions. Elution Buffer (24 mL, Table 10) was applied to the column and collected in 6- 4 mL fractions. All fractions were analyzed by SDS-PAGE and either Coomassie blue or anti-His western blot. The elution fractions with the highest and most pure protein content were pooled and prepared for size exclusion chromatography.

Table 9: Buffers for Nickel Affinity Purification and SEC of SUMO-LARP6

Buffer	Content
Cell Lysis	50 mM sodium phosphate (pH 7.4), 300 mM sodium chloride, 10 mM imidazole, (1 tab/ 50 mL) protease inhibitor tablet
Wash 1	50 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole, 2 mM β-mercaptoethanol
Wash 2	50 mM sodium phosphate, 300 mM sodium chloride, 30 mM imidazole, 2 mM β-mercaptoethanol
Elution Buffer	50 mM sodium phosphate, 300 mM sodium chloride, 300 mM imidazole, 2 mM β-mercaptoethanol
Storage Buffer	50 mM Tris-HCl (pH 7.5 at 4 °C), 150 mM sodium chloride, 5% (v/v) glycerol, 2 mM β-mercaptoethanol
Cleavage Buffer	50 mM Tris-HCl (pH 7.5 at 4 °C), 150 mM sodium chloride, 5% (v/v) glycerol, 2 mM β-mercaptoethanol, 30 mM imidazole

Size Exclusion Chromatography – All samples to be separated by size exclusion chromatography were first concentrated by centrifugation (4000 x*g*, 4 °C) in 5 min intervals using a Viva-Spin 10,000 MWCO centrifugal concentrator (Sartorius). As described above, the retentate was gently mixed via aspiration between runs to minimize protein concentration gradients and to monitor protein aggregation. Samples were then filtered using a 0.2 μm syringe filter and 2 mL were loaded onto the FPLC. Prior to loading, the system was equilibrated with filtered and degassed Storage Buffer (Table 10). Samples were run at a flow rate of 1 mL/min on a 16/300 S200 Superdex column (GE Life Sciences). Samples under the peak were analyzed by SDS-PAGE and Coomassie stained. Fractions containing the target protein were pooled to prepare for second nickel affinity column and ULP1 cleavage of the N-terminal SUMO tag.

ULP1 Cleavage and Second Nickel Affinity Chromatography – ULP1 concentrations for SUMO cleavage were calculated for a 1:200 molar rations of SUMO-fusion proteins to ULP1 and performed by incubation for 2 hours at 16 °C. A second nickel column purification was performed by incubating the cleavage reaction with Ni^{2+}-NTA beads and incubating at 4 °C for 1 hour while shaking. The protein/bead mixture was then applied to a column and allowed to settle. The unbound proteins were collected as the flow through, and then and Wash 2 buffer (Table 12) was applied and the eluate collected. The elution buffer was applied (Table 12) and collected. ULP1 cleavage and the affinity column were analyzed by SDS-PAGE and Coomassie stained.

Size Exclusion Chromatography – The flow through was concentrated by centrifugation (4000 x*g*, 4 °C) in 5 min intervals using a Viva-Spin 10,000 MWCO centrifugal concentrator (Sartorius), again aspirating the retentate between runs to minimize protein concentration gradients and to monitor protein aggregation. Samples were then filtered using a 0.2 μm syringe filter and 2 mL was loaded onto the run on a 16/300 S200 Superdex column (GE Life Sciences) at a flow rate of 1 mL/min. Samples under the peak were analyzed by SDS-PAGE and Coomassie stain to confirm the presence of the target protein and to select appropriate fractions. Fractions selected were pooled, concentrated and snap frozen.

Human Cell Line Culture

HeLa, HEK-293, MDA-MB-231, and MCF7 cell lines were selected for cellular localization studies on the basis of previous work in the lab indicating endogenous LARP6 expression. The HeLa, HEK-293, and MCF-7 cell lines were obtained from ATCC. The MDA-MB-231 cells were a generous gift from Dr. Tania Betancourt (Texas State University). Cells were grown in the appropriate complete growth media (CGM) (Table 10) at 37 °C with 5% CO_2.

Table 10: Human Cell Culture Lines

Cell Line	Source/ Catalog #	Complete Growth Media	BSL
HeLa	ATCC	EMEM, 10% FBS	2
HEK-293	ATCC/ CRL-1573	EMEM, 10% FBS	2
MCF-7	ATCC	EMEM, 10% FBS, 0.01 mg/mL human recombinant insulin	1
MDA-MB-231	Betancourt Lab	L15, 10% FBS	1

Immunocytochemistry

Seeding of Cells on Coverslips – After using a hemocytometer to count cells, the cells were seeded at 125,000- 250,000 cells/mL onto poly-D-lysine coverslips (12 mm diameter, Neuvitro, NC0343705), placed in the center of each well of a 24-well plate. Cells were grown on the coverslips until ~50-60% density. Once the appropriate density was reached, cells were treated with leptomycin B if indicated, and fixed using one of two approaches: 4% paraformaldehyde (PFA) or methanol fixation, as described in detail below. The 4% paraformaldehyde fixation was ultimately used for the majority of the immunofluorescence imaging.

Methanol Fixation, Permeabilization, and Blocking – Media was removed from the cells. Then cells were rinsed with 1X PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO, pH 7.4). Cells were treated with ice cold 100% methanol for 5 min at -20 °C. The methanol was drawn off of the cells and a 1X PBS rinse was performed. After fixation cells were permeabilized in 0.5% Triton-X 100 in 1X PBS for 5 min at room temperature. Cells were then gently washed with 1X PBS for 5 min at room temperature prior to antibody treatment. Blocking was performed in 1% BSA in 1X PBS for 1 hour at room temperature.

4% PFA Fixation, Permeabilization and Blocking – Media was removed from the cells, and the cells were rinsed with 1X PBS. Cells were treated with 4% PFA for 10 min at room temperature. The paraformaldehyde was removed and

the cells rinsed with 1X PBS. After fixation, cells were permeabilized in 0.5% Triton-X 100 in 1X PBS for 5 min at room temperature. Cells were then gently washed with 1X PBS for 5 min at room temperature prior to antibody treatment and blocked in 1% BSA in 1X PBS.

Antibody Staining – Cells were covered in a primary antibody solution mouse anti-LARP6 (Abnova, H00055323-B01P, H8211) in 1X PBS at the indicated concentrations and incubated at 4 °C overnight. The cells were washed with 1mL of 1X PBS for 5 min at room temperature and repeated for a total of three washes. The laboratory room was darkened by turning off overhead fluorescent lights and working by the light of a single lamp with an LED bulb. The cells were treated with the anti-mouse AlexaFluor488-conjugated secondary antibody (Jackson ImmunoResearch, 715-545-150, 132588) at varying concentrations and incubated at room temperature for 1 hour. The plate containing the coverslips was wrapped in foil to minimize light exposure. After incubation with the secondary antibody the cells were washed gently with 1X PBS at room temperature for 5 min and repeated for a total of three washes. The coverslips were then mounted to microscope slides using Prolong Diamond Antifade Mountant with DAPI (Fisher Scientific, P36966). Slides were stored at 4 °C in the dark for no more than 48 hours before imaging to minimize opportunities for photobleaching.

Confocal Imaging, Image Processing – All confocal imaging was performed using the Fluoview 1000 (Olympus) confocal system in the Texas State University Analytical Research Service Center (ARSC). Except where otherwise noted, triplicate images were taken of each slide. Images were processed using FIJI software.[48]

Cell Lysate Screening with anti-LARP6 Antibodies

Method 1: Cell Lysis by Trypsin Detachment and Centrifugation – Cells were grown in T25 flask until approximately 70% confluency. Cells rinsed with 1X DPBS (ATCC), 2.5 mL trypsin-EDTA was added to the plate and incubated at 37 °C 5% CO_2 for 3-5 min. Three milliliters of CGM (Table 10) was added to the flask. Cells were transferred to conical tubes and centrifuged at 130 x*g* for 5 min at room temperature. The supernatant was discarded, and cells were resuspended in 1 mL CGM. Cell count was performed with hemocytometer. Cells were pelleted by centrifugation at 1000 x*g* for 2 min at 4 °C. After thawing on ice, the cell pellet was washed with 1 mL ice-cold 1X PBS and cells were again pelleted as before. The PBS wash was repeated for a total of three washes before being stored overnight at -20 °C. Cell pellet was thawed on ice and resuspended in RIPA cell lysis buffer (0.5 mM PMSF and protease inhibitor cocktail added) (1 μL per 20,0000 cells). Cell debris was pelleted by centrifugation (max speed, 10 min, 4 °C). Supernatant collected and stored on ice, cell pellet was discarded. Commercial bicinchoninic assay (BCA) (Pierce) was performed to calculate total protein content. Analysis of cell lysis was

performed by loading equalized or maximal quantities of protein onto 10% SDS-PAGE gels and anti-LARP6 western blots were performed. Recombinant LARP6 protein was used as positive control for western blot.

Method 2: Cell Lysis by Freeze Thaw, Scraping and Centrifugation – Cells were grown in round bottom cell culture plates with initial cell densities equalized at the time of seeding. Cells were grown at 37 °C 5% CO_2 until confluency was approximately 70% then rinsed with 1X DPBS (ATCC) four times. Due to cells in different round bottom cell culture plates initial seeding at same density and growth under identical conditions it was assumed if visual inspection demonstrated similar confluency that plates can same cell count. This allowed for one plate to be counted while the others of the same cell type were lysed. The plate used for cell counting, was drained of all 1X DPBS and 10 mL of 1X DPBS was added. Cells were manually detached by a cell scrapper and a cell count was performed as previously described. For the treatment plate, after 4 rinses with 1X DPBS, the plate was drained of all liquid and frozen at -20 °C for 5 min. Cells were thawed on ice and RIPA lysis buffer (0.5 mM PMSF and protease inhibitor cocktail added) was applied (400 μL per 10,000,000 cells). Cells were scraped from the plate and centrifuged (14,000 x*g*, 4 °C, 5 min) to pellet cell debris. Cell pellet was resuspended in the supernatant before being centrifuged (14,000 x*g*, 4 °C, 5 min) again. The supernatant was collected and stored on ice as the cell lysate. A commercial BCA assay was performed to calculate total protein content of the cell lysate. Analysis of cell lysis was performed by loading

equalized or maximal quantities of protein onto an SDS-PAGE and an Anti-LARP6 western blot was performed. Recombinant LARP6 protein was used as positive control for western blot.

III. IS LARP6 AN *IN VIVO* TARGET FOR CRM1 MEDIATED NUCLEAR EXPORT?

While the previous work on LARP6 demonstrated altered localization when GFP-tagged LARP6 deletion mutants of putative NES/NLS were transfected into stem cell lines, these data do not give information as to the pathway by which this is occurring. While there are many exportins in the cell, many have very narrowly defined and specialized functions. CRM1 is the predominant exportin in the cell. The predictions of NES/NLS in the primary literature are based on NES consensus sequences in CRM1 cargo proteins. One aim of this project was to identify if LARP6 is an *in vivo* target for CRM1-mediated nuclear export.

The foundational work in this field relied on the overexpression of recombinant LARP6. When expression levels are too low to visualize endogenous levels of a protein, overexpression can allow for visualization. However, overexpression comes with the risk that the high levels of the protein in the cell will cause negative effects and shift cellular processes from their standard state due to the complex, interrelated, and delicate balance of cellular processes. One goal of this study was to examine cellular localization at endogenous levels to minimize such effects.

A previous graduate student the Lewis lab, Elizabeth McIvor, screened several cell lines using commercially obtained anti-*LARP6* antibodies through western blots of cell lysates (Figure 10). Through the use of Method 1 for cell

lysate screening, as described in Methods, and the commercial anti-LARP6 antibody (page 46), a western blot was performed to assess cell lines for endogenous LARP6 expression. Cell lines MDA-MB-231, HEK-293, SKOV, CHO, COS, and HCT-116 all demonstrated bands at the expected molecular weight of approximately 70 kDa (Figure 10). On this basis, cell lines HEK-293 and MDA-MB-231 were selected as LARP6-positive cell lines for further work. The MCF-7 line was selected due to its apparent lack of endogenous LARP6, both by mRNA expression (Human Genome Atlas) and western blot (Figure 10). In addition, HeLa cells were selected as a final cell line to be tested. Three cell lines (HeLa, HEK-293 and MCF-7) were ordered from ATCC, while MDA-MB-231 cells were a kind gift of the Betancourt Lab at Texas State University. The mouse anti-*LARP6* antibody (Abnova, catalog number H00055323-B01P) was selected as it was reported by the manufacturer as being appropriate for both western blot and immunocytochemical (ICC) studies and its apparent success in this initial cell lysate screening.

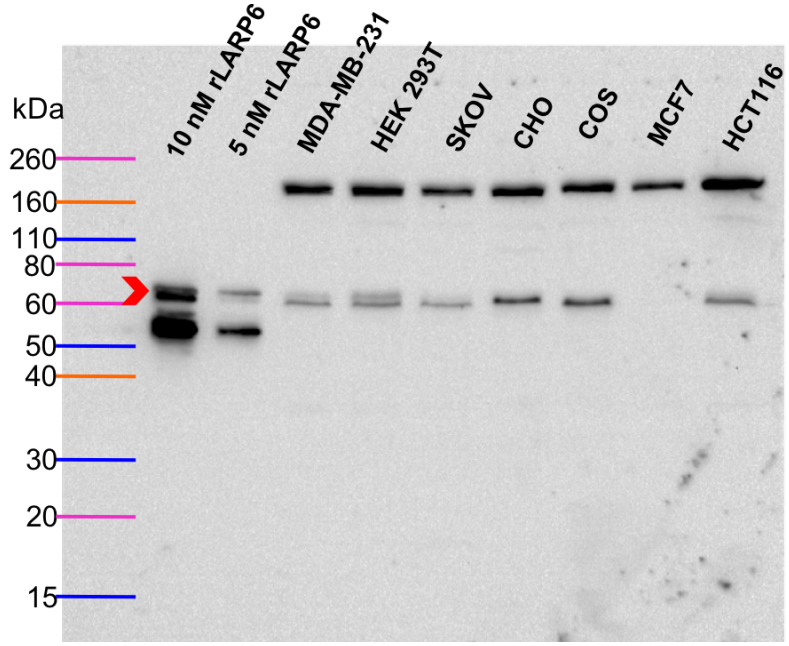

Figure 10: Cell Lysate Screening with commercial anti-LARP6 antibody. Prior graduate student, Liz McIvor performed cell lysate screening. Recombinant LARP6 expression and all human cell lines except MCF-7 demonstrate band at the expected molecular weight of LARP6, approximately 70 kDa. The band at ~200 kDa is a nonspecific interaction.

Adherence and seeding density screening

In order to determine if HEK-293 cells were capable of adhering to poly-D-lysine coverslips, a cell-adherence and seeding-density experiment was carried out. Cells were counted and seeded onto coverslips as previously described at densities ranging from 10,000 -100,000 cells per well to determine the density at which cells should be seeded for ICC. Additionally, cells were seeded into wells lacking coverslips to test for comparative adherence without a coverslip. Incubation occurred for 24-72 hours at 37 °C, 5% CO_2. Images were obtained on Dr. Betancourt's EVOS FL Cell Imaging system (Figure 11). It was determined

that adherence was improved with poly-D-lysine coverslips as opposed to the growth in the well alone. Additionally, seeding density at approximately 50,000 cells per well was the ideal density to obtain the ideal confluence for ICC imaging (Figure 11). Seeding density is important in immunofluorescent experiments not only because overcrowding of cells can change cellular expression but also because densely grown cells result in signal from one cell bleeding into the signal of another.

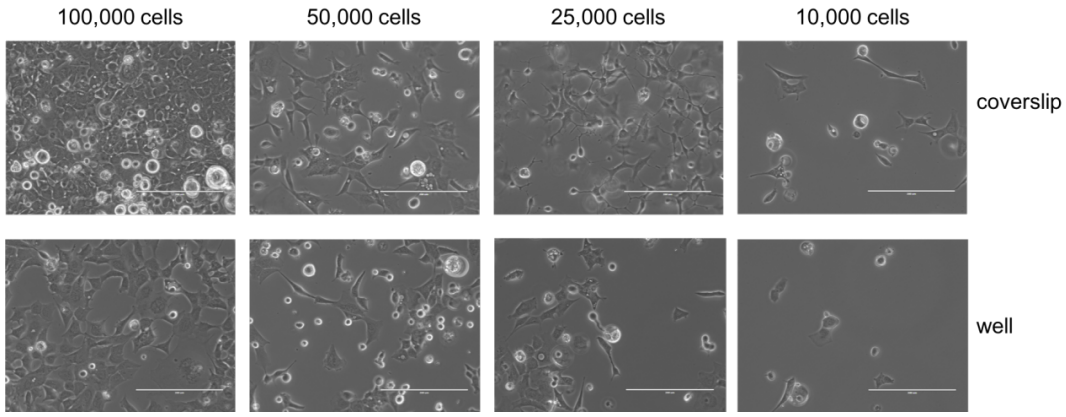

Figure 11: HEK-293 Seeding Density and Adherence on Poly-D-Lysine coverslips. HEK-293 cells were counted using a hemocytometer and cell density per mL was calculated. Cells were seeded either onto poly-D-lysine coverslips within a 24-well plate or seeding into an empty well of the same plate. All wells containing cells were brought a working volume of 1 mL with complete growth media and incubated at 37 °C 5% CO_2. Images taken with EVOS FL Cell Imaging system. Scale bar is 200 μm.

Determining fixation and antibody concentrations

After a seeding density and adherence protocol was established, the appropriate fixation method and optimal antibody concentration of the commercially obtained anti-*LARP6* antibody needed to be determined. Multiple methods of fixation were attempted. Paraformaldehyde (PFA) fixation, which occurs on the chemical basis of covalent crosslinking, can stabilize biomolecules through causing covalent interactions with one another. This was the utilized in previous LARP6 localization studies. However, this method has some drawbacks, particularly the toxicity of PFA to the user requiring careful avoidance of contact, particularly of inhalation. For this reason, methanol fixation was also selected to assess its ability to fix cells. The chemical basis of this fixation method is that methanol dehydrates the cell. Both methanol and 4% paraformaldehyde (PFA) were selected as the initial screening fixation methods for HEK-293 cells. Methanol fixation resulted in detachment of some cells from the coverslip, which was visually observed when large portions of the fixed cells came off of the coverslip during the post-methanol rinse. Additionally, the signal of the DAPI stain in the nucleus appeared to be reduced throughout the nuclei as evidenced by the lack of blue signal in the center of the nuclei (Figure 12, compare panels [A-H] to panels [I-P]). For these reasons subsequent experiments utilized 4% PFA fixation.

Screening of antibody concentrations of Abnova mouse anti-LARP6 antibody – To identify appropriate antibody concentrations for robust detection of

LARP6, a range of 1:500 to 1:2000 of primary anti-LARP6 was evaluated using either 1:500 or 1:1000 dilution of secondary (Figure 12). Unfortunately, during the mounting of the coverslips to the microscope slides, the coverslip for methanol fixation 1:500 primary/ 1:500 secondary was broken (Figure 12, panel I). For this reason, no images were taken of this coverslip. Only one image per coverslip was taken on the confocal microscope. The remaining images showed signal in the cytosol, as expected for LARP6. Moreover, there was no secondary antibody background signal with dilutions of 1:500 or 1:1000 in the absence of primary antibody (Figure 12, panels D, H, L, P). To be wholly confident in these concentrations, these antibody dilutions will be screened in a cell line that does not express LARP6 (e.g., MCF-7; see below).

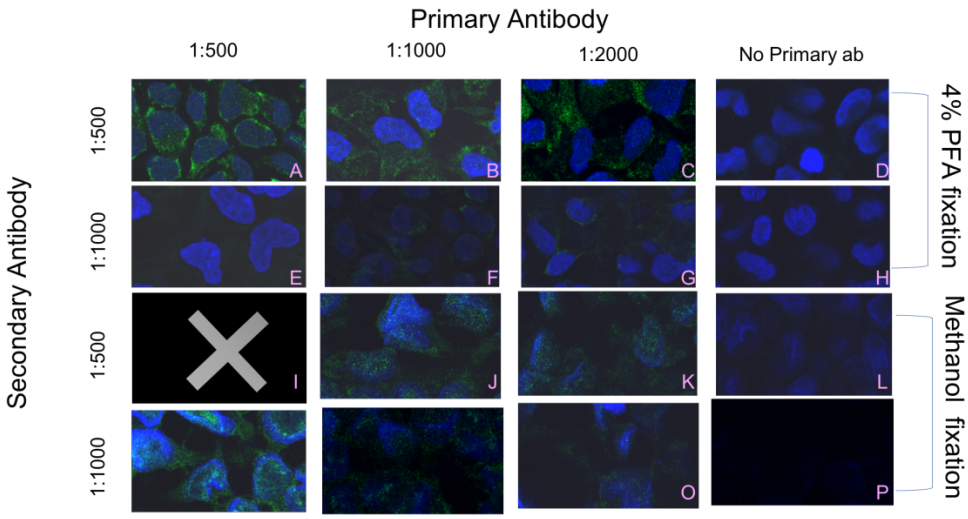

Figure 12: Fixation and antibody dilution screening HEK-293 P16. HEK-293 cells were counted using a hemocytometer and cell density per mL was calculated. Cells were seeded onto poly-D-lysine coverslips and fixed with either 100% ice cold methanol for 5 min at -20 °C or 4% PFA for 10 min at room temperature. Permeabilization was performed with 0.5% Triton-X 100 in 1x PBS for 5 min at room temperature. Cells were then treated with mouse anti-*LARP6* primary antibody with dilutions ranging from 1:500 to 1:2000 and anti-mouse Alexa Fluor 488 secondary antibody concentrations ranging from 1:500 to 1:1000. The last column was in the absence of primary antibody. DAPI stain shown in blue and Alexa 488 in green.

As expected, the highest level of primary antibody (1:500) produced the highest intensity of signal across both conditions (secondary antibody concentration and fixation method). In the paraformaldehyde fixation, the localization of anti-LARP6 signal is clearly observed around the nucleus. This localization is not unexpected, as LARP6 is known to chaperone the type I collagen mRNA to the rough endoplasmic reticulum for protein synthesis. To confirm this the dye, $DioC_6(3)$ can be used in future studies to stain the ER during live cell imaging.[49] This pattern of localization is less apparent in lower primary antibody concentrations, but can be generally observed across PFA

fixation slides (Figure 12 panels A-C). The methanol fixation resulted in diffuse signal for the DAPI nuclear stain, but signal for the LARP6 was observed both within the nucleus and in the area surrounding it (Figure 12, panels J,K, M-O).

The lack of background signal of the secondary antibody at dilutions of 1:500 in the initial localization, fixation, and antibody concentration screening was used to select a primary antibody concentration of 1:2000 and a secondary antibody concentration of 1:500 for future studies. The comparatively diffuse signal throughout the nucleus under methanol fixation was not desired and this method was ultimately discarded in favor of PFA fixation.

Leptomycin B inhibition

LMB, an inhibitor of CRM1 mediated nuclear export is an established method for confirmation of CRM1 cargo proteins. It functions as a competitive inhibitor of nuclear export and through its ability to shut down nuclear export of a cargo protein, results in the accumulation of that protein in the nucleus. Through the use of fluorescently labeled putative cargo proteins, such nuclear accumulation can be visualized allowing for the confirmation or exclusion of CRM1 mediated nuclear export.

For the initial LMB screening, Leptomycin B treatment was performed as previously described and LARP6 was imaged on the Fluoview Confocal microscope. These data appear to show little to no background signal with

secondary antibody concentrations at a 1:500 dilution factor, confirming the earlier results (Figures 12 and 13). The data showed apparent nuclear accumulation with LMB treatment. However, these results also need to be confirmed using a LARP6-null cell line (Figure 13).

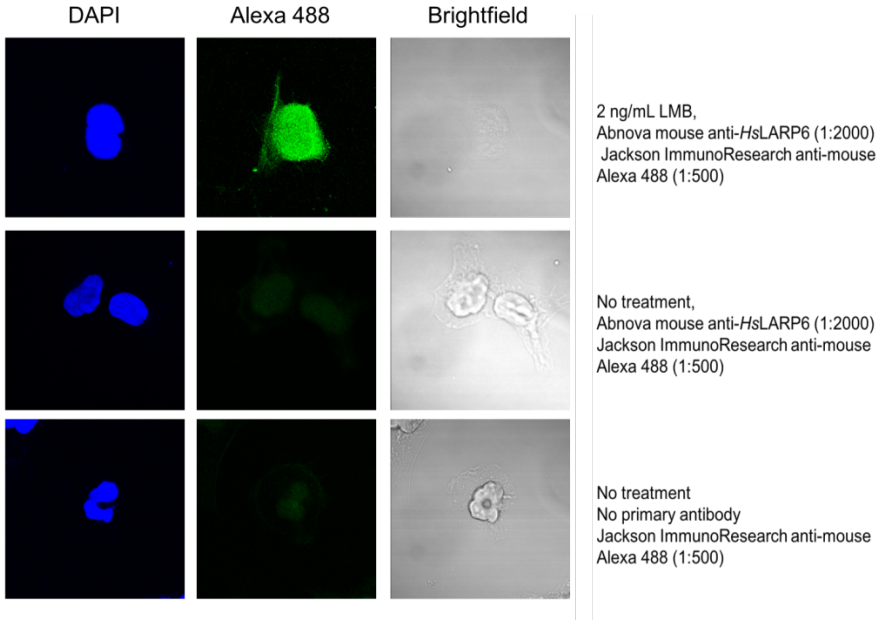

Figure 13: LMB inhibition of CRM1 screening in HEK-293 cells. HEK-293 cells (passage 17) were seeded onto poly-D-lysine coverslips and either treated with 2 ng/mL of LMB for 3 hours or experienced no treatment. Fixation and permeabilization was performed with 4% PFA and 0.5% Triton-X 100 in 1x PBS respectively. Cells were then treated with mouse anti-*LARP6* primary antibody (1:2000) and anti-mouse Alexa Fluor 488 secondary antibody (1:500). DAPI stain shown in blue and Alexa 488 in green.

In addition to the lack of LARP6-null cell line, several other points of optimization were required as only one image of each coverslip/condition was taken in the initial LMB screening (Figure 13). In addition to this, no image was taken in the absence of primary and secondary antibody to check for autofluorescence, which should be performed in future experiments. In the

following experiments, passage 29, passage 34, and 37, three images were taken across various points of each coverslip. For all of these experiments (Figure 13-15) controls of no primary antibody (i.e., only secondary antibody) were performed. For passage number 29 and 34, a control to detect autofluorescence, no primary or secondary antibody was also performed (Figure 14 and 15). The autofluorescence control for passage 37 was unable to be images due to the coverslip being dropped face down during fixation. The controls for auto fluorescence for passage 29 and passage 34 show no autofluorescence. While in future experiments this must be done for every trial, the lack of autofluorescence is promising especially passage 17 and on had identical confocal microscopy setting. This should not be overinterpreted and future experiments must have these controls in place for every experiment.

In these subsequent experiments with later passage number, secondary background signal was observed (Figure 14). The only difference between these experiments was the passage number, with the initial two experiments being performed at passage number 16 and 17 respectively and second LMB trial being at passage number 29. Secondary antibody signal was also observed in passage 34 and 37 (Figure 15). In these later passage numbers, a smaller secondary antibody concentration (1:1000) was also tested. Changes in morphology and gene expression can occur in later passage numbers and might explain the observation of nonspecific secondary signal in later passage number cells. However, unhealthy cells due to stress or infection can also have alter morphology. Additional replications of this experiment will be necessary to test

either smaller secondary antibody concentrations (e.g., 1:2000, 1:5000) and/or keep the passage number low (below 20-25 passages).

The bright field images of HEK-293 passage 17 (Figure 13), and representative images of passage 29 (Figure 14), 34 and 37 (Figure 15) show unusual morphology, which may indicate that these cells are sick, either from infection or from stress (starvation, over growth, poor adherance to CO_2). While poor fixation can also cause unusual appearance of morpgology the increasing sensitivity to secondary antibody non-specific binding may indicate poor heath of the cells that is getting worse with further passaging of the cells. During future replications, attention must be given to the morphology of the cells to ensure they are healthy throughout the study as unhealthy cells can have significant impacts on results.

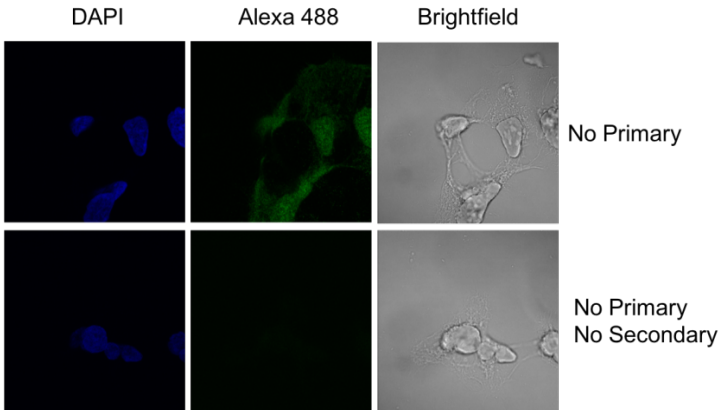

Figure 14: Nonspecific antibody signal in HEK-293 P29 cells at secondary antibody dilution of 1:500. HEK-293 cells P29 were counted using a hemocytometer and cell density per mL was calculated. Cells were seeded onto poly-D-lysine coverslips and experienced no treatment. Fixation and permeabilization was performed with 4% PFA and 0.5% Triton-X 100 in 1x PBS respectively. Cells were then treated with no primary antibody and anti-mouse Alexa Fluor 488 secondary antibody (1:500) or no primary and no secondary antibody. DAPI stain shown in blue and Alexa 488 in green. Representative images are shown.

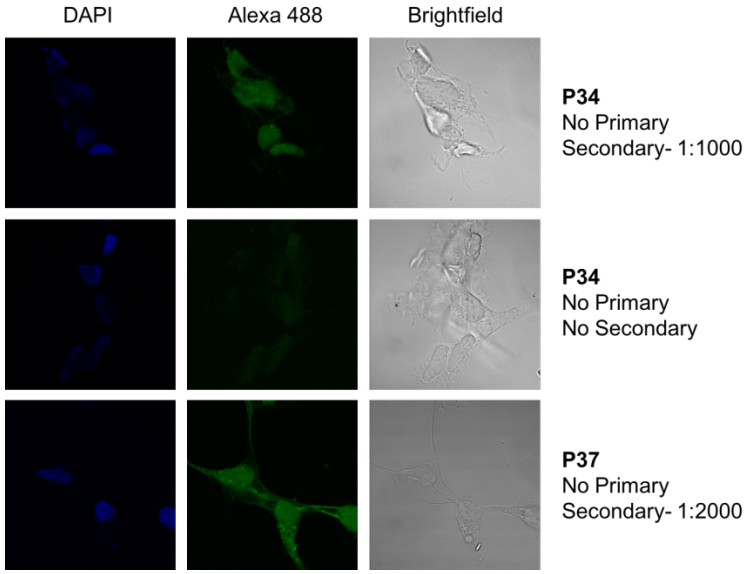

Figure 15: Nonspecific antibody signal in HEK-293 cells at secondary antibody dilution of 1:1000. HEK-293 cells (P34 and P37) were counted using a hemocytometer and cell density per mL was calculated. Cells were seeded onto poly-D-lysine coverslips and experienced no treatment. Fixation and permeabilization was performed with 4% PFA and 0.5% Triton-X 100 in 1x PBS respectively. Cells were then treated with no primary antibody and anti-mouse Alexa Fluor 488 secondary antibody (1:1000). DAPI stain shown in blue and Alexa 488 in green. Representative images shown.

As previously noted, the ICC results cannot be fully interpreted without confirmation of antibody concentrations in a cell line lacking expression of LARP6. To confirm the antibody concentrations used in earlier experiments were appropriate the LMB treatment experiments were replicated in MCF-7 cells. Unfortunately, these data showed background signal with lower secondary antibody concentrations than used in HEK-293 localization screenings (Figure 16). The nonspecific signal does not account for all the signal observed in the presence of primary antibody. This was unexpected as this cell line was meant to be a negative control for endogenous LARP6. This was performed at low passage number (P9), which rules out passage number as a causative factor of this signal. This secondary signal could be a result of unspecific primary antibody binding, but this goes against the initial findings of a prior student in the lab (Figure 10). An additional observation is the morphlogy of these cells is not as expected, this could be caused by stress, infection, or inappropriate fixation methodology. This methodolgy was never never optimized for this cell line.

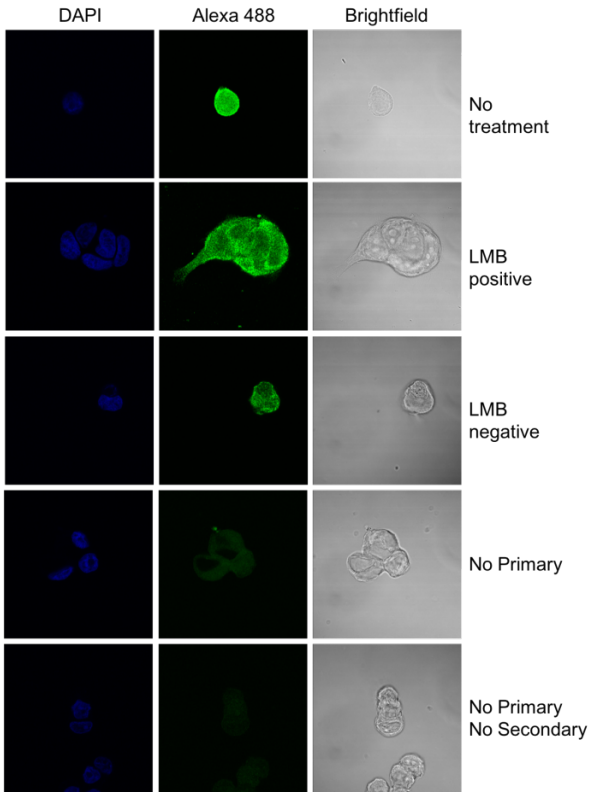

Figure 16: LMB inhibition of CRM1 screening in MCF-7 cells. Imaging performed on Fluoview Imager on MCF7 P9 cells under various conditions no treatment, (complete growth media alone), LMB positive (2 ng/mL LMB in complete growth media), LMB negative (ethanol in complete growth media) and no primary (no treatment and no primary antibody). Demonstrates both primary and secondary signal that was unexpected for a negative cell line.

The unexpected results from the localization studies resulted in a closer examination of the prior cell lysate screening. Despite several attempts made to reproduce the cell lysate screening data obtained by a prior student, no cell lysate screens with the mouse anti-LARP6 antibody was successful. An initial cell lysate screening trial was performed using the method piloted by a former student ("cell lysis method 1"). However, when the recombinant *LARP6* protein was equalized with the protein concentrations calculated from cell lysate by BCA

assay, no signal was observed with any of the antibodies tested including the Abnova mouse anti-LARP6 (data not shown). Due to the lack of signal for the positive control, two possibilities were considered. First, the western blot may have been failed due to some reagent; second, the concentrations of the recombinant LARP6 and the cell lysates were too low to be seen when equalized.

Due to these difficulties, and many possible variables in the previous cell lysate blot, the following experiment was performed with a 1:10 dilution of recombinant LARP6 and maximal concentrations of the cell lysates. Maximal concentrations of cell lysate were used because it was hypothesized that the lack of signal in the previous attempt was a result of low levels of endogenous LARP6 in the cell. These data showed signal for the recombinant LARP6 that was smeared across the membrane (Figure 17). It was hypothesized that the preparation of recombinant LARP6 had degraded in the -70 °C and that at a 1/10 dilution the recombinant protein still was significantly more concentrated than the cell lysate. Additionally, abnormalities/ smearing of the ladder on the membrane make these data uninterpretable. Even so, it appears the cell lysate band was migrating significantly higher than the expected molecular weight (Figure 17).

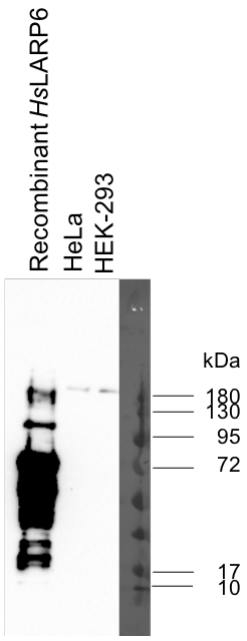

Figure 17: Cell Lysate Screening maximal protein concentrations. Maximal concentrations of cell lysates prepared with Method 1 Cell Lysate Screening were analyzed by SDS-PAGE and anti-LARP6 western blot. Smearing of the recombinant protein, indicative of degradation after long term storage in the freezer was observed. Problems with the ladder smearing and the unexpectedly high bands for HeLa and HEK-293 cells make this difficult to interpret.

Other members of my lab group independently worked to replicate the earlier cell lysate data but were also unsuccessful (data not shown). These early cell lysate screening were performed on early passage numbers of HeLa/HEK-293 cell though exact passage numbers were not tracked. During later passage numbers, an alternative method of cell lysis was used, as method 1 had not been successful at replicating the prior student's results. When protein concentrations of cell lysates were equalized with a newer prep of recombinant LARP6 protein than was used in previous studies, no signal was observed for HeLa, HEK-293,

MDA-MB231 or MCF7 cells (Figure 18, right). The recombinant LARP6 protein showed a band at the expected molecular weights for this antibody. While a secondary higher running band is present, this has been seen before in the Lewis Lab with LARP6 proteins and is also observed on the website of the manufacturer of this commercial antibody. This western blot showed high background signal suggesting optimization of the western blot protocol is required. When maximal quantities of the recombinant protein and the cell lysates were run on a separate SDS-PAGE gel and analyzed by western blot, the recombinant *LARP6* protein showed stronger bands than in the equalized western blot (Figure 18, left). However, there was still no apparent signal for the cell lysates and there was high background. These results combined with the cellular localization studies suggest there may be a problem with the cell lines and/or methodology selected for studying LARP6 *in vivo.*

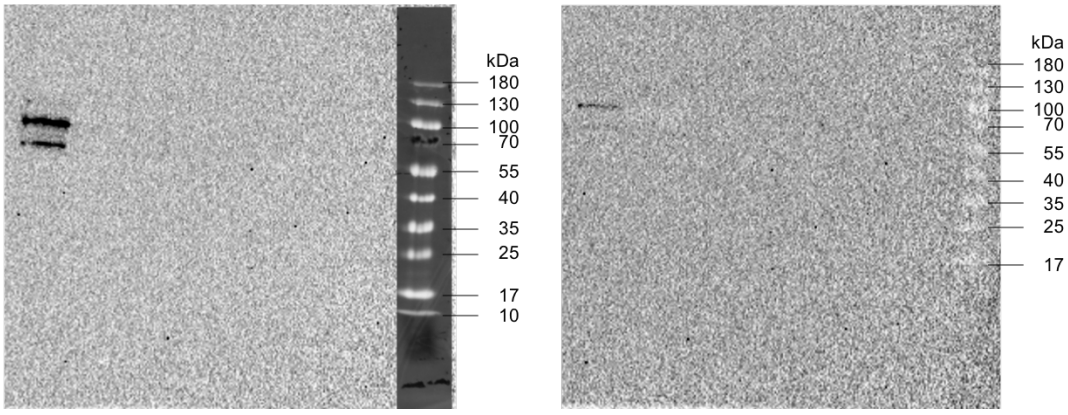

Figure 18: Cell lysate screening with Cell lysate Method 2. Cell lysate screening of HEK-293, HeLa, MDA-MB-231 and MCF-7 cells by anti-LARP6 western blot with Abova anti-*LARP6* antibody. HeLa and HEK-293 cells were at later passage numbers (>30) MCF-7 and MDA-MB-231 cell were at lower passage numbers (<20). Maximal quantities (left) and normalized (right) of recombinant protein and cell lysates were analyzed by anti-LARP6 western blot.

A closer look into the basis of the cell line selection was performed by examining RNA expression data available online. There was very little to be found about RNA expression of LARP6 in the primary literature. However, the human protein atlas database characterization of LARP6 RNA expression may offer some insight as to the aberrant localization observed (Figure 19).[50] While moderate expression of LARP6 RNA is present in HEK-293 cells, HeLa cells demonstrated very small expression levels. But what is more concerning is the low-level expression in MCF-7, which was expected to be LARP6-null. While the presence of RNA does not directly signify the presence of the protein, the absence of endogenous LARP6 in MCF-7 cells should be confirmed. If this cell line in fact does express LARP6, that would explain the IHC signal we observed in MCF-7 cells. This means that the higher antibody concentration where LMB treatment appeared to altered localization may or may not be within the acceptable range of antibody concentrations. Even if the expression levels HEK-293 and HeLa cells is too low for localization studies using IHC methods, the U-2197 cell line should be screened for LARP6 protein expression via western blot. Without primary literature or RT-qPCR data the data from the human protein atlas should be questioned but it does offer some possible explanation for the results observed. The ten cell lines with the highest concentrations as identified on the human protein atlas (U2197, U-87 MG, BJ, HHSteC, PC3, U-251 MG, hTERT-HME1, BJ hTERT24-B, HEK-293) and a cell line identified as having no LARP6 expression (U-937) should be screened using RT-qPCR to confirm their expression levels.

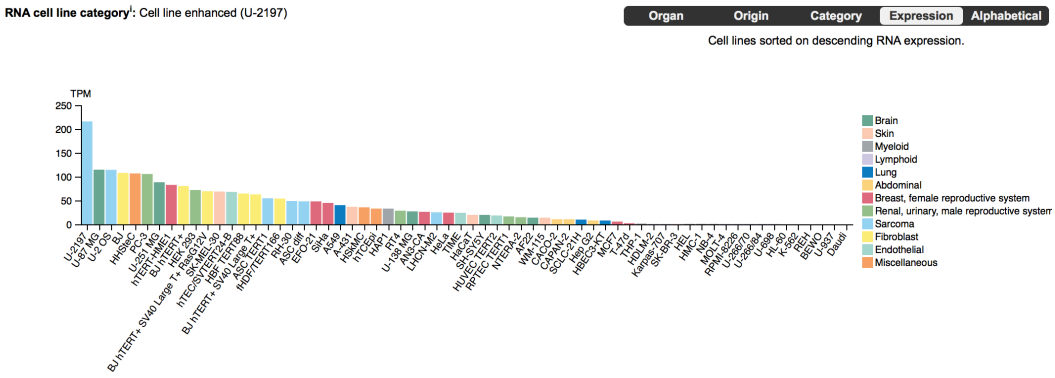

Figure 19: Human Protein Atlas LARP6 RNA expression by cell line. Relative expression levels of LARP6 in human cell lines. This was obtained from the Human Protein Atlas.

Real Time Quantitative PCR (RT-qPCR) data should be obtained on any other cell lines being considered for endogenous LARP6 localization studies in the future and new cell lines should be selected that may have higher LARP6 expression. A new cell line for a negative control should also be confirmed by RT-qPCR that lacks RNA for LARP6. New cell lysate data should also be obtained and multiple anti-*LARP6* antibodies should be screened. Due to the potential differences in cellular expression at different passages, RT-qPCR data should be confirmed across the range of passage numbers being used in future studies. However, caution should be taken as the presence of RNA does not directly equate the presence of a protein. This requires combining this information with cell lysate screening with a several of anti-LARP6 antibodies to ensure appropriate cell line and antibody selection.

Our attempts to detect endogenous levels of LARP6 encountered two major challenges that must be overcome to be successful. The potential

unreliability of the commercially obtained anti-human LARP6 antibody will require significant screening before any results could be trusted. Additionally, the potential of low levels of endogenous LARP6 is a complicating factor. While these methods would preserve our initial aims to explore the localization of endogenous LARP6, it is advisable to first test the CRM1-mediated nuclear export of LARP6 though LMB treatment by using GFP-tagged LARP6 transfected into a cell line system. This will reduce the number of variables that need to be considered and most closely mirror previous localization studies. To probe the localization of endogenous LARP6, future cell line choices to be screened by RT-qPCR and western blot screening of cell lysates before localization studies are conducted. It is therefore preferable to replicate GFP-tagged LARP6 transfections used in the previously published localization studies as a proof of concept before attempting to tackle endogenous LARP6 localization.

There are other tagging mechanisms that would remove some of the potential drawback that arise from using a large tag such as GFP. One alternative is to tag LARP6 with immunoreactive tags such as HA or Myc. This approach, however, will require transfecting tagged proteins into human cell lines. If performed with approaches that preserve endogenous expression levels, this strategy would likely be able to address the original aims of identifying CRM1-dependence of nuclear localization of LARP6.

IV. *IN VITRO* INTERACTION BETWEEN LARP6 AND CRM1

As described above, previous work in mammalian cells identified a putative NES within the LARP6 RRM.[19] However, these results preceded high-resolution structural data. The solution NMR structure showed that the putative NES is also a central element of the core β-sheet in the RRM.[20] The deletion of this sequence could cause structural distortion and therefore cause disruption of the cellular localization. Additionally, since the original identification of the LARP6 NES, the ability of the field to computationally predict NES has advanced in both accuracy and precision.[41] LocNES is an algorithm that is trained on the basis of experimentally-confirmed NES and incorporates structural data. This allows for alternative NES to be identified by considering consensus sequence adherence and predicted structural information for CRM1 accessibility.

LocNES-Identified NES Fusion Constructs

When human LARP6 sequence was analyzed by LocNES, several new putative NES motifs were identified. While other alternative candidates were also identified, there was a clear drop in LocNES scoring and we set our cutoff at the three top scoring NES (Table 11). The three top-scoring putative NES are, in order of scoring matrix rank: residues 457-471, residues 210-224, and residues 179-193 (Table 11). The 457-471 motif is located in the C-terminal LSA motif of LARP6, for which we have no structural information. While the previously identified Shao NES (residues 186-218) was not strictly identified by LocNES, the

179-193 and 210-224 NES sequences flank and overlap the Shao NES. It is noteworthy that these predicted NESs are located in either the RRM or the LSA.

Table 11: LocNES-identified NES

Putative NES	Putative NES Sequence	LocNES score	Predicted NES Class
LARP6(457-471)	KMQTADGLPVGVLRL	0.357	II
LARP6(210-224)	QEKVMEHLLKLFGTF	0.280	III
LARP6(179-193)	ENLPSKMLLVYDLYL	0.264	I
LARP6(452-466)	PLLSRKMQTADGLPV	0.194	I
LARP6(219-233)	KLFGTFGVISSVRIL	0.185	I/II/III
LARP6(213-227)	VMEHLLKLFGTFGV	0.185	I/III

Because the solution NMR structure has been solved, the putative NESs in the RRM can be evaluated in a structural context. LARP6(179-193) is comprised of the β1 strand into the α0' helix (Figure 20a). Of the three LocNES-predicted motifs, this is lowest-ranked. That score may be a result of the likely inaccessibility of the β1 strand to the CRM1 binding pocket. In contrast, LARP6(210-224) falls within the α1 helix, which is on the outside of the RRM and may be more accessible to CRM1 (Figure 20b). As the RRM structure was determined to be relatively rigid,[20] it is expected the β1 strand will remain shielded by the α1' helix.

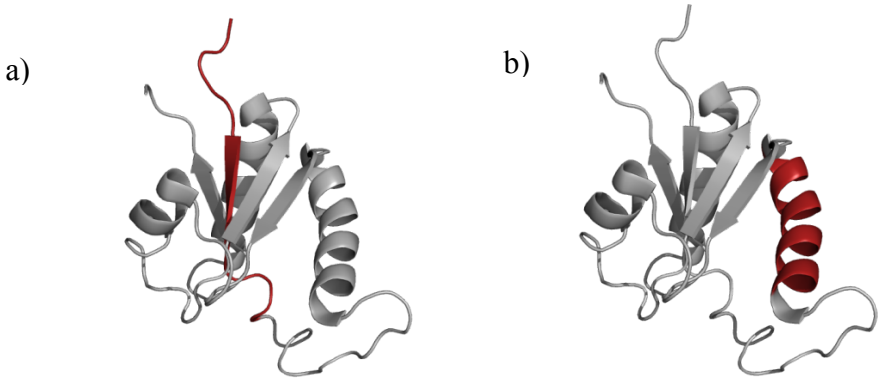

Figure 20: LocNES-identified RRM putative NES. The solution NMR structure of the LARP6 RRM (PDB: 2MTG) is used to map the putative NES in the RRM. (a) LARP6 residues 179-193 is marked in red. This sequence comprises the core beta strand. (b) LARP6 residues 210-224 is marked in red. This sequence comprises the majority of helix a0'. Figures created in PyMol (Schrodinger, Inc.).[32]

However, the most interesting prediction is the strongest predicted NES, LARP6(457-471), located within the LSA. Even though there is no structural data for the LSA, the literature suggests that the LSA motif can contribute to protein localization and protein-protein regulatory interactions.[51] This suggests it is accessibility for protein-protein interactions such as the one necessary for CRM1 mediated nuclear export. Together, both the prior *in vivo* experiments and the results of the LocNES predictions strongly suggest that LARP6 may be a bona fide cargo for CRM1.

Approach for Determining CRM1 Binding to LARP6

Pull down assays allow the detection of protein-protein interactions *in vitro* using recombinantly expressed proteins. This method has been used with success in the identification of NES.[41] In this system, the putative NES with a GST tag acts as "bait" for the CRM1 Ran-GTP complex (Figure 21). If a putative

NES is capable of being bound by CRM1, then CRM1 and Ran-GTP will be retained and can be detected by SDS-PAGE and Coomassie Staining. The basis of this method is on the interaction of the exportin, CRM1 with NES, which is Ran-GTP dependent. [41] Using GST-tagged putative NES as bait and immobilizing them on a column before incubating the with excess molar ratios of CRM1 and Ran-GTP allows determination of the ability of CRM1 to bind a putative NES. This method requires the recombinant expression of all three components of this interaction: Ran-GTP, CRM1, and the "bait" – GST tagged putative NES sequences. As a negative control GST tag was expressed by itself. Additionally, to characterize CRM1's ability to interact with the full-length LARP6, recombinant expression of GST tagged wild-type LARP6 was also performed.

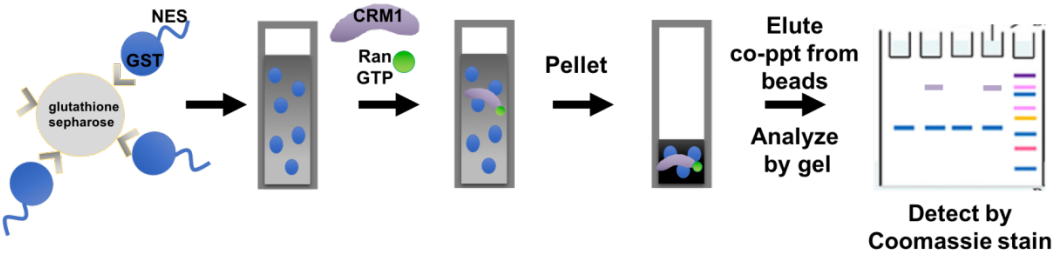

Figure 21: Cartoon depiction of NES bait interaction/pull down mechanism. GST tagged putative NES "bait" (blue) bind the GSH resin. When incubated with excess molar quantities of the exportin CRM1 (purple) and the energy molecule Ran-GTP (green) are pelleted by centrifugation. If CRM1 binds the putative NES it will be retained. Alternatively, if it is incapable of binding it will be washed away. Bound CRM1, Ran-GTP and the bait NES will be eluted from the resin and analyzed on SDS-PAGE gels. Binding can be determined by Coomassie/silver staining.

Cloning of GST-tagged NES Peptides

We have identified 3 putative NES using LocNES LARP6(179-193), LARP6(210-224), and LARP6(457-471). Due to the small size of these constructs, they were ordered as forward and reverse compliment oligonucleotides with *Eco*RI- and *Xho*I-compatible ends for annealing and direct ligation into the pGEX4T3-TEV plasmid. (Appendix 1, Table 1). Successful annealing was determined by a shift in size when separated on agarose gel electrophoresis (Figure 22). This same method was used for cloning three control GST-NES constructs, using one known NES for each of three classes of NES: PKIα, HIVrev, and Mad1 (Figure 23).

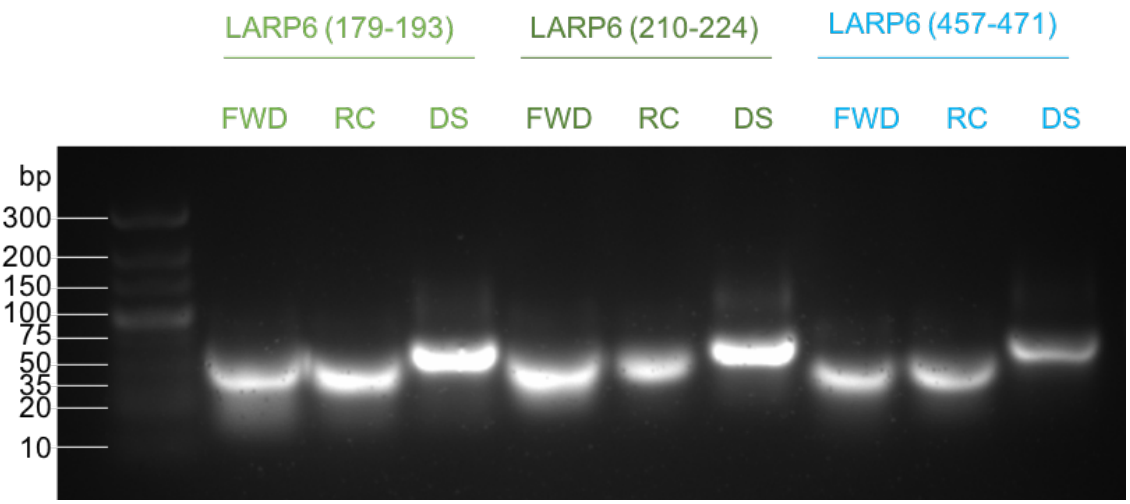

Figure 22: Annealing of Putative NES Oligonucleotides. Single stranded forward (FWD) oligonucleotides and reverse complement (RC) oligonucleotides and double stranded (DS) annealed oligonucleotides were run on a 3.5% agarose gel stained with ethidium bromide. The shift in molecular weight for the double stranded DNA indicates successful annealing.

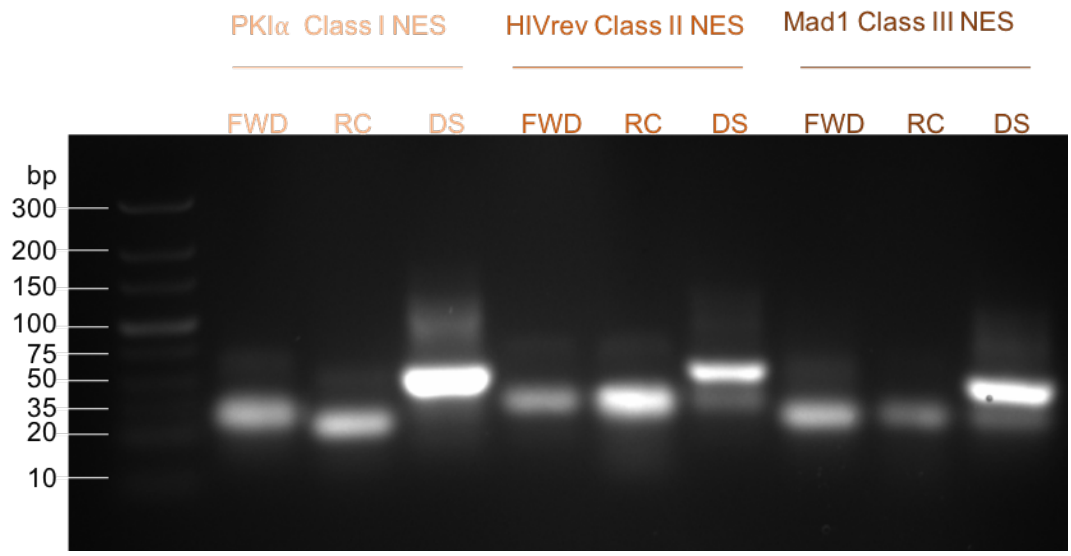

Figure 23: Annealing of Experimentally confirmed NES Oligonucleotides. Single stranded forward (FWD) oligonucleotides and reverse complement (RC) oligonucleotides and double stranded (DS) annealed oligonucleotides were run on a 3.5% agarose gel stained with ethidium bromide. The shift in molecular weight for the double stranded DNA indicates successful annealing.

In addition to cloning the LocNES-predicted motifs as GST fusions, the Shao NES was also cloned as a GST fusion. Because this was a longer sequence, the insert was amplified by PCR for conventional restriction enzyme-based cloning into pGEX4T3-TEV. Successful PCR was determined by analysis by gel electrophoresis (data not shown). The PCR product was double-digested with EcoRI and XhoI and then gel-purified. Initial extraction of this digested insert resulted in the loss of the DNA, because its size was below the minimum recoverable size from the standard QIAGEN Gel extraction kit (70 bp). Subsequent gel extractions should use a gel extraction kit designed for small DNA fragments.

The host plasmid, pGEX4T3-TEV, was also double-digested and treated with phosphatase to prevent re-circularization. Confirmation of the successful

digest of the plasmid was also performed by gel electrophoresis (Figure 24). The digested, purified plasmid was extracted from the gel, and then combined with each NES fragment for ligation. The ligated DNA was transformed into DH5α *E. coli* cells. From the resultant plated colonies, an overnight culture was inoculated to amplify the plasmid. The plasmid was extracted using a commercial miniprep kit, and the isolated DNA was analyzed by commercial Sanger sequencing (Genewiz). Sequencing results were analyzed by aligning to the intended product using the program Snapgene.

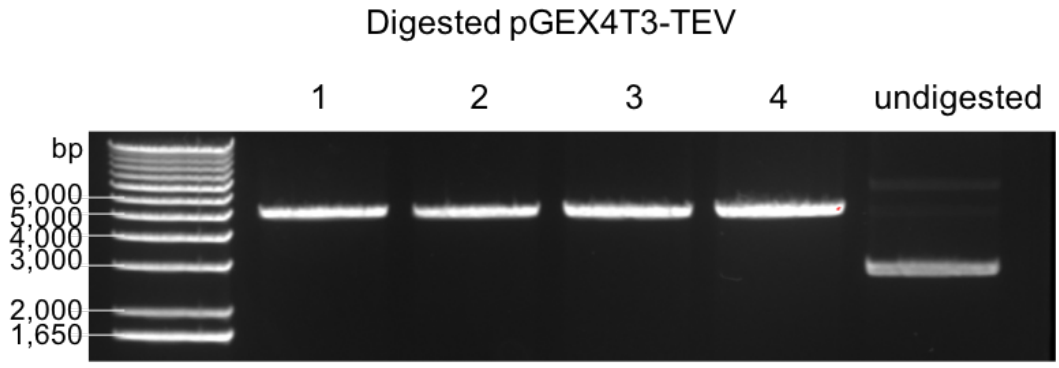

Figure 24: Double Digest of pGEX4T3-TEV plasmid with EcoRI and XhoI endonucleases. Double digest of the 20 μg of pGEX4T3-TEV plasmid with 1.5 units each of the endonucleases EcoRI and XhoI at 37 C for 1 hour. Treated with 5 units of Antarctic Phosphatase at 37 C for 1 hour before beginning run on 0.9% Agarose gel and stained with ethidium bromide. Undigested pGEX4T3-TEV plasmid was also run as control. Digested pGEX4T3-TEV ran at expected molecular weight, approximately 5000 bp.

Trial Expressions of GST-tagged NES peptides

Sequence-confirmed DNA was transformed into Rosetta *E. coli* cells for protein expression. A single colony was used to inoculate an overnight culture that was grown to saturation. The next day, an aliquot of that culture was then used to inoculate expression cultures. Following the published protocol for GST-NES fusions, protein expression was induced with 1 mM IPTG at OD_{600} of 0.5-

0.6 and then cultured at 30 °C for 6 hours. Expression levels of the GST-fusions did not change significantly between hours 5 and 6 and protein expression was strong at 5 hours post-induction (Data not shown). Therefore, the expression protocol for preparative cultures (1 L) was designed for 5-hour expression times.

Preparative Expression of GST-tagged NES Peptides

One-liter preparative expressions were performed as previously described. Throughout the expression time course aliquots were taken to confirm expression of the target protein over time. GST-LARP6(179-193) was expressed and time aliquots were taken prior to induction (time 0) and every hour after until cells were harvested at 5 hours post induction (time 5). Time course aliquots were analyzed by SDS-PAGE and Coomassie stain (Figure 25). At time point 0 there is only a faint band at the expected molecular weight of ~30 kDa, but by time point 1 and for every successive time point there is an increase in band intensity (Figure 25). This is indicative of successful purification and at time point 5 cells were harvested as previously described and stored at -20 °C until GSH affinity purification.

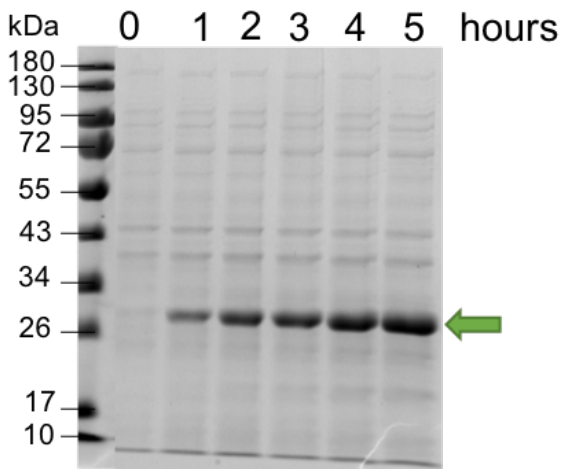

Figure 25: Preparative Expression of GST-LARP6(179-193). Hourly time aliquots of GST-LARP6(179-193) expression in 1 L Luria Broth were run on 10% SDS-PAGE gel at 200 V for 1 hr in 1X SDS Tris Gly Running Buffer. The increasing band intensity at 30 kDa over time is indicative of successful expression (green arrow).

The same process as describe above was used in the preparative expression of GST-LARP6(210-224). Analysis was performed by SDS-PAGE and Coomassie stain (Figure 26). The hourly time point aliquots taken prior to induction of expression (0) until 5 hours post induction (5) demonstrate increasing band intensity at the expected molecular weight of approximately 30 kDa (Figure 26). This is evidence of successful expression and at timepoint 5 cells were harvested by centrifugation. The cell pellet was stored at -20 °C until affinity purification.

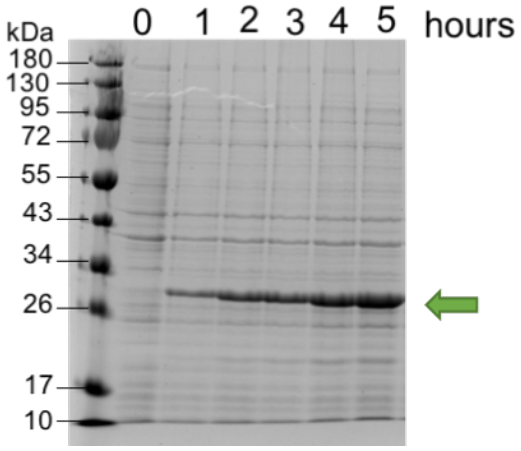

Figure 26: Preparative Expression of GST-LARP6(210-224). Hourly time aliquots of GST *LARP6* 210-224 expression in 1 L Luria Broth. Aliquots pelleted by centrifugation and resuspended in 1x SDS-PAGE loading buffer. Run on 10% SDS-PAGE gel at 200 V for 1 hour in 1x SDS Tris Gly Running Buffer. The increasing band intensity at 30 kDa over time is indicative of successful expression (green arrow).

GST-LARP6(457-471) was expressed as described, and analysis was performed by SDS-PAGE and Coomassie stain (Figure 27). The hourly time point aliquots taken prior to induction of expression (0) until 5 hours post induction (5) demonstrate increasing band intensity at the expected molecular weight of approximately 30 kDa (Figure 27). This is evidence of successful expression and at timepoint 5 cells were harvested by centrifugation. The cell pellet was stored at -20 °C until affinity purification.

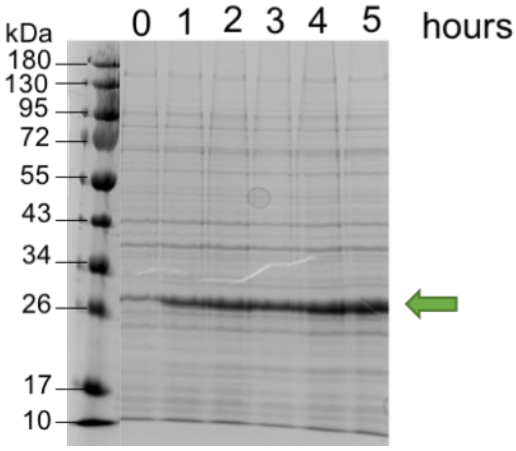

Figure 27: Preparative Expression of GST-LARP6(457-471). Hourly time aliquots of GST *LARP6* 457-471 expression in 1 L Luria Broth. Aliquots pelleted by centrifugation and resuspended in 1x SDS-PAGE loading buffer. Run on 10% SDS-PAGE gel at 200 V for 1 hour in 1x SDS Tris Gly Running Buffer. The increasing band intensity at 30 kDa over time is indicative of successful expression (green arrow).

Similar work was carried out to produce preparative expressions of positive-control NES peptides. First, the class II positive control GST-HIVrev was expressed and analyzed by SDS-PAGE and Coomassie stain (Figure 28). The hourly time point aliquots taken prior to induction of expression (0) until 5 hours post induction (5) demonstrate increasing band intensity at the expected molecular weight of approximately 30 kDa (Figure 28). This is evidence of successful expression and at timepoint 5 cells were harvested by centrifugation. The cell pellet was stored at -20 °C until affinity purification.

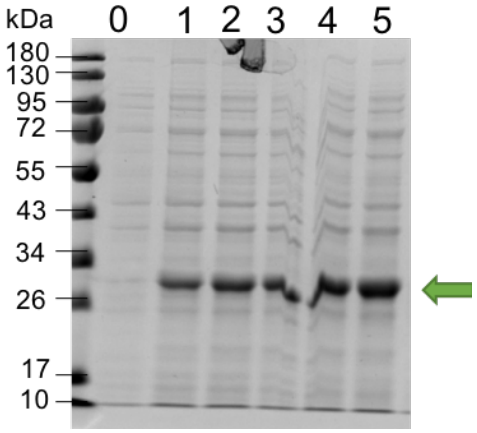

Figure 28: Preparative Expression of GST-HIVrev (Class II NES). Hourly time aliquots of GST HIVrev NES expression in 1 L Luria Broth. Aliquots pelleted by centrifugation and resuspended in 1x SDS-PAGE loading buffer. Run on 10% SDS-PAGE gel at 200 V for 1 hour in 1x SDS Tris Gly Running Buffer. The increasing band intensity at 30 kDa over time is indicative of successful expression

Finally, a preparative expression of the NES negative control GST tag was performed as previously described. Analysis was performed by SDS-PAGE and Coomassie stain (Figure 29). The time point aliquots taken prior to induction of expression (0) until 5 hours post induction (5) demonstrate increasing band intensity at the expected molecular weight of approximately 30 kDa (Figure 29). This is evidence of successful expression and at timepoint 5 cells were harvested by centrifugation. The cell pellet was stored at -20 °C until affinity purification.

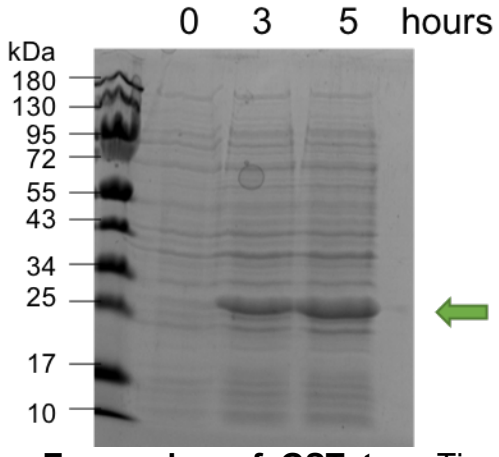

Figure 29: Preparative Expression of GST tag. Time aliquots of GST tag expression in 1 L Luria Broth were taken. Aliquots were pelleted by centrifugation and resuspended in 1x SDS-PAGE loading buffer. Run on 10% SDS-PAGE gel at 200 V for 1 hr in 1x SDS Tris Gly Running Buffer. The increasing band intensity at 30 kDa over time is indicative of successful expression.

Cell Lysis and Batch Affinity Purification of GST-tagged NES Peptides

The frozen cell pellet for all constructs were thawed and lysed as previously described in preparation for affinity purification. Clarified cell lysates were incubated with the GSH resin for 30 min prior to batch affinity purification. Batch Affinity purification was performed by centrifugation and analyzed by SDS-PAGE and Coomassie stain.

Three GST proteins were readily purified: the negative-control isolated GST (Figure 30), the positive-control GST-HIVrev (Figure 31), and the experimental GST-LARP6(457-471) (Figure 32). Successful affinity purification was evidenced by the ~30 kDa band present in the elution fractions. For each of those purifications, the elution fractions containing the protein of interest were pooled together and concentrated in preparation for size exclusion chromatography (SEC).

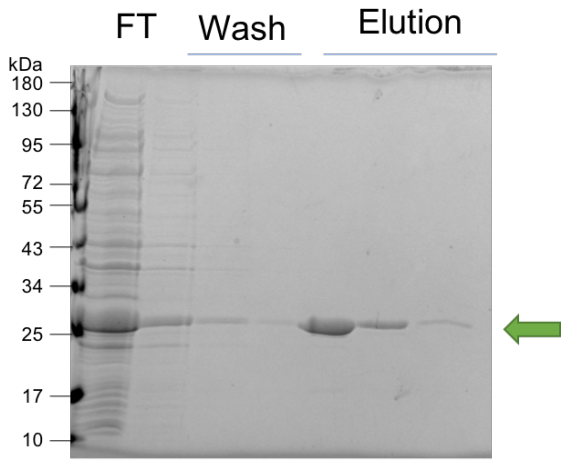

Figure 30: Batch Purification of GST tag. Batch affinity purification was performed with the GSH resin and GST tag cell lysate after a 30 min incubation. Elution fractions 1-2 containing protein at the expected molecular weight, 30 kDa, were collected, pooled and concentrated for SEC.

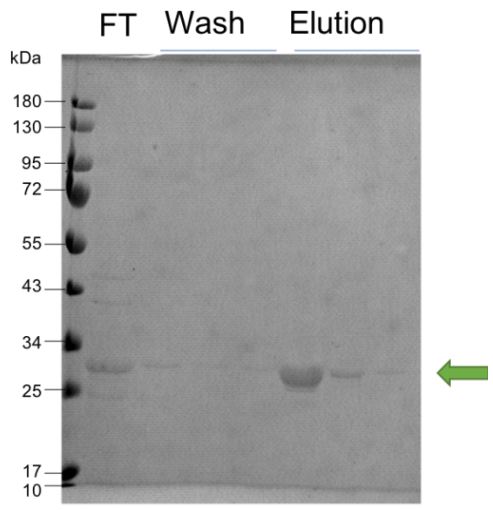

Figure 31: Batch Purification of GST-HIVrev. Batch affinity purification was performed with the GSH resin and GST-HIVrev cell lysate after a 30 min incubation. Elution fractions 1-2 containing protein at the expected molecular weight, 30 kDa, were collected, pooled and concentrated for SEC.

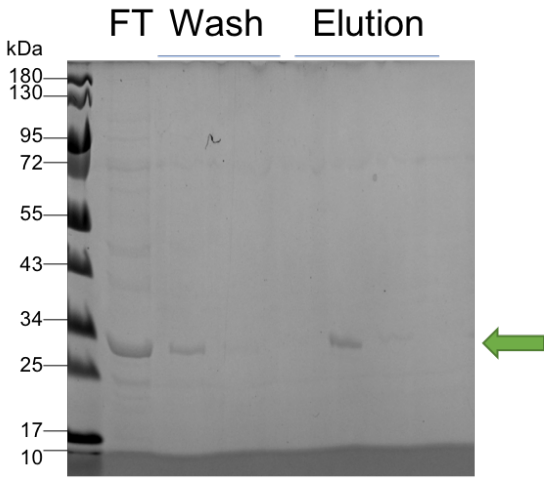

Figure 32: Batch Purification of GST tagged *LARP6*(457-471). Batch affinity purification was performed with the GSH resin and GST-LARP6(457-471) cell lysate after a 30 min incubation. Elution fraction 1 containing protein at the expected molecular weight, 30 kDa, were collected, pooled and concentrated for SEC.

In contrast, the GST-LARP6(179-193) and GST-LARP6(210-224) underwent multiple failed attempts at purification. Representative results from the affinity purifications are shown in Figures 33 and 34, respectively.

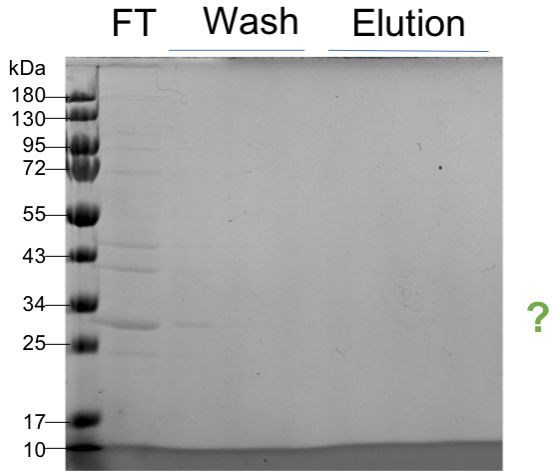

Figure 33: Batch Purification of GST tagged *LARP6*(179-193). Batch affinity purification was performed with the GSH resin and GST-LARP6(179-193) cell lysate after a 30 min incubation. Bands at the expected molecular weight of 30 kDa are only present in the FT and the wash suggesting interference of successful interaction of GST tag with GSH resin.

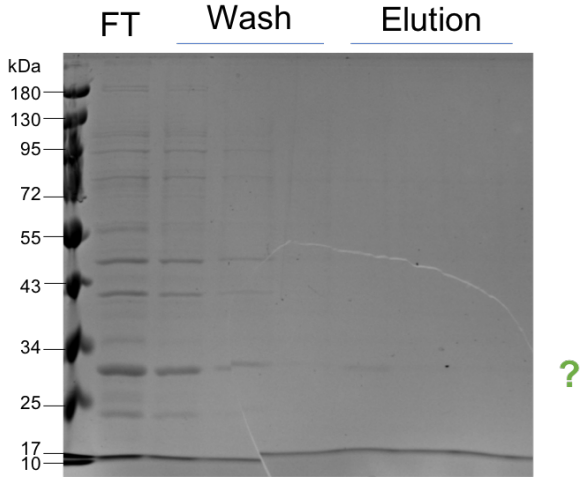

Figure 34: Batch Purification of GST tagged *LARP6*(210-224). Batch affinity purification was performed with the GSH resin and GST-LARP6(179-193) cell lysate after a 30 min incubation. Strong bands at the expected molecular weight of 30 kDa are only present in the FT and the wash suggesting interference of successful interaction of GST with GSH resin.

Three hypotheses were generated as possible explanations of the inability to affinity-purify these constructs. The first hypothesis was that the incubation of clarified cell lysate with the resin needed to be optimized to increase interaction between the GST-fusion protein and the beads. Therefore, the affinity purification protocol was screened for a variety of times and temperatures for the incubation with GST-LARP6(179-193) (Figure 35). Because these experiments showed no improvement in the recovery of tagged protein, no permanent adjustments in the affinity purification protocol were made.

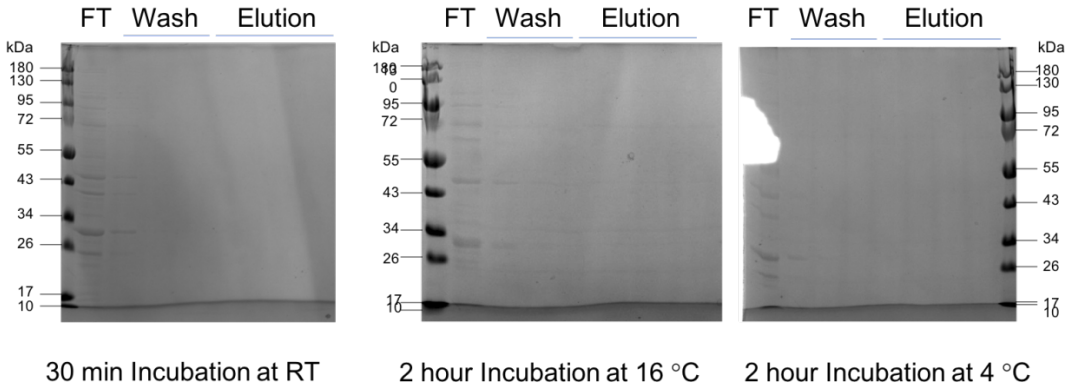

Figure 35: Incubation screening of GSH resin and GST-LARP6(179-193). Incubation conditions were tested between the GSH resin and the clarified cell lysate. The standard protocol of a 30-minute incubation at room temperature was tested (left) with no protein visible in the elution as detected by Coomassie stain. Additionally, two-hour incubations at 16 °C (middle) and 4 °C (right) were also attempted.

Another potential explanation for the failed affinity purifications was based on the inherent hydrophobicity of the putative NES motifs. This is an inherent feature of NES sequences, because they rely on hydrophobic interactions with the binding pocket of CRM1 to be exported from the nucleus. Alternatively, sequence-specific interactions could have interfered with the interactions between the GST tag and the glutathione sepharose affinity resin. Two methods were employed to test these hypotheses. To test the role of sequence specificity, oligonucleotides were designed to randomize the sequence of the NES peptides while retaining sequence content. Secondly, to increase solubility of the hydrophobic tags, the osmolyte glycerol was added to the purification buffers.

Sequence-Specific Effects in Failed Purifications

One approach to determining why the purification of LARP6(179-193), and LARP6(210-224) failed was to identify if any sequence specific interactions were responsible. To test this hypothesis, the protein sequences of the NES peptides were entered into RandSeq Tool on ExPASy. This required a minimum of 20 amino acids to be input but LARP6(179-193) and LARP6(210-224) were only 15 amino acids in length. Since neither sequence contained alanine, five alanines were entered into the sequence generator and deleted from the resultant sequence (Table 12).

Table 12: Random Sequence Generation

	Original Sequence	Randomized Sequence
LARP6(179-193)	ENLPSKMLLVYDLYL	LYLKNLMLPSEDLYV
LARP6(210-224)	QEKVMEHLLKLFGTF	EHKLVFGFLKLTMQE

Using the randomized sequences, oligonucleotides were designed to maintain nucleic acid sequence composition of the original oligonucleotides. These were then ordered from IDT as forward and reverse complements for annealing and direct ligation into pGEX4T3-TEV plasmid between the EcoRI and XhoI cut sites as previously described. Annealing was confirmed by gel electrophoresis (Figure 36). Ligation of the annealed oligonucleotides into the pGEX4T3-TEV plasmid was performed as described previously. Ligated DNA was transformation into DH5α cells and DNA minipreps were performed to isolate the plasmid DNA. Sanger sequencing through Genewiz confirmed the successful cloning of these randomized sequences behind the GST tag of the pGEX4T3-TEV plasmid. The sequence confirmed DNA was then transformed into Rosetta *E. coli* cells and prepared for expression. These constructs have been expressed in preparative quantities as previously described (data not shown). Affinity purification must be carried out to determine if sequence specific interactions have played a role in the difficulties purifying LARP6(179-193) and LARP6(210-224).

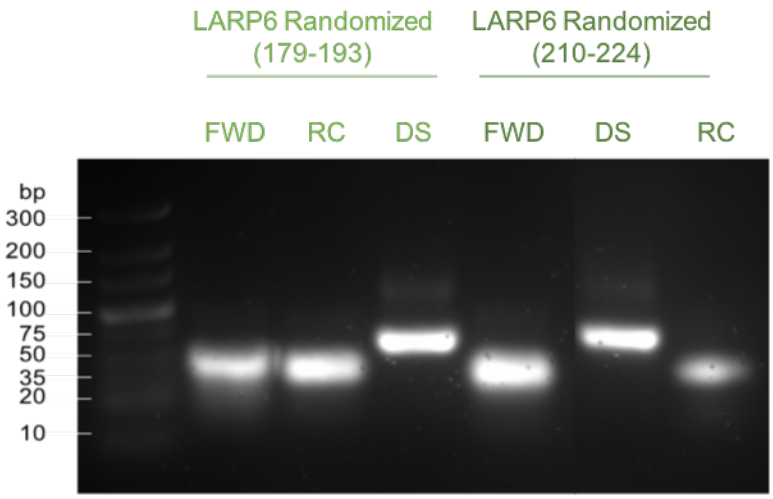

Figure 36: Annealed randomized putative NES. Single stranded forward (FWD) oligonucleotides and reverse complement (RC) oligonucleotides and double stranded (DS) annealed oligonucleotides were run on a 3.5% agarose gel stained with ethidium bromide. The shift in molecular weight for the double stranded DNA indicates successful annealing.

Optimization of GST-tagged NES Peptides Affinity Purification Buffers

If the inherent hydrophobicity of the NES is responsible for the failed purifications the addition of the surfactant glycerol may help overcome this challenge. An addition of 5% glycerol to the purification buffers (Table 8) and subsequent rounds of batch purification by centrifugation were performed as previously described. As with the previous attempts at batch purification, analysis of the purification success was performed by SDS-PAGE gel electrophoresis and stained by Coomassie staining.LARP6(179-193) was successfully purified using this technique (Figure 37). LARP6(210-224) was also performed and successfully purified (Figure 38). Elution fractions containing the protein of interest were pooled together and concentrated in preparation for size exclusion chromatography.

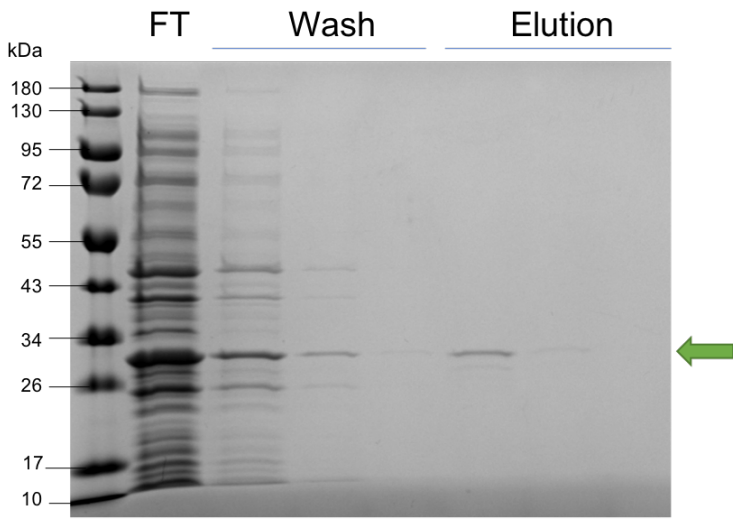

Figure 37: Affinity Purification of GSTLARP6(179-193) with addition of 5% glycerol. Batch affinity purification with 5% glycerol in the purification buffers was performed with the GSH resin and GST-LARP6(179-193) cell lysate after a 30 min incubation. Strong band at the expected molecular weight of 30 kDa are present in the elution indicating successful purification.

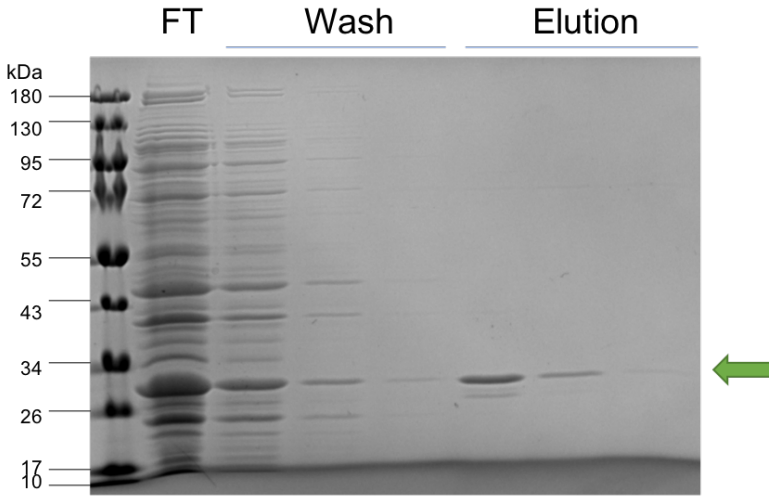

Figure 38: Affinity Purification of GST *LARP6*(210-224) with addition of 5% glycerol. Batch affinity purification with 5% glycerol in the purification buffers was performed with the GSH resin and GST-LARP6(210-224) cell lysate after a 30 min incubation. Strong band at the expected molecular weight of 30 kDa are present in the elution indicating successful purification.

Size Exclusion Chromatography of GST-NES Peptides

With all GST-NES proteins affinity-purified, they were subjected to size exclusion chromatography on a Superdex 75 (S75) column to assess polydispersity and execute a buffer-exchange into suitable storage buffer. Fractions falling under the A280 peaks were analyzed by gel electrophoresis using SDS-PAGE gels as previously described and were stained in Coomassie Stain. Fractions identified as having the appropriate protein as determined by molecular weight were pooled together and snap frozen as previously described before being stored at -70 °C.

The NES negative control, GST tag, was readily purified by SEC with a single sharp peak being observed at the expected molecular weight of dimerized GST tags ~60 kDa (Figure 39). The fractions falling under the peak were analyzed by SDS-PAGE and Coomassie stain (Figure 40).

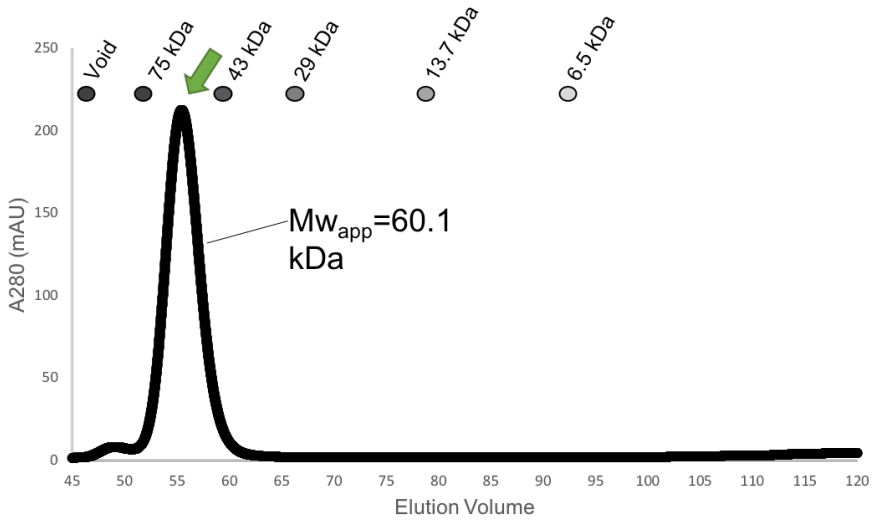

Figure 39: SEC of GST tag. SEC analyzed by A280 reading from GST tag batch purification with no glycerol added to purification buffers. Eluted as dimer (~60 kDa). Green arrow indicates peak of interest.

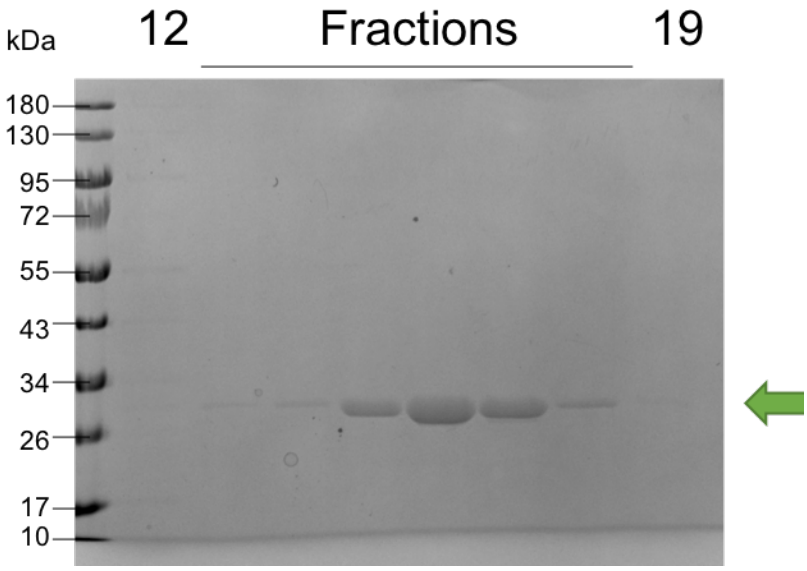

Figure 40: Coomassie Stain of SEC fractions of GST tag. SEC peak analyzed by SDS-PAGE and Coomassie stain. Fractions 15-18 containing isolated protein at expected molecular weight (~30 kDa) were pooled and snap frozen.

The NES class II positive control, GST-HIVrev, was readily purified by SEC with a single sharp peak observed at the expected molecular weight of dimerized GST (~65 kDa)(Figure 41). The fractions under the peak were analyzed by SDS-PAGE and Coomassie stain before being pooled and frozen (Figure 42).

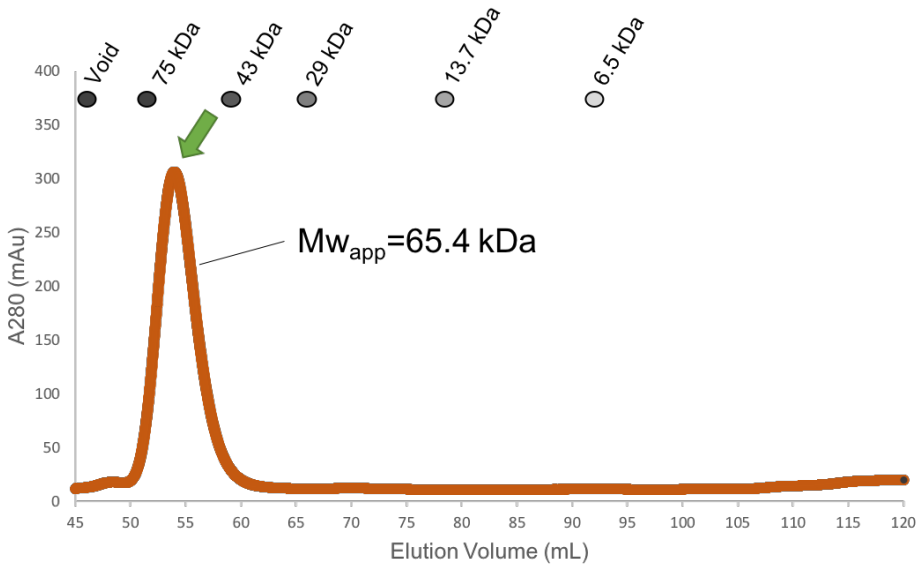

Figure 41: SEC of GST HIVrev. SEC analyzed by A280 reading from GST HIVrev batch purification with no glycerol added to purification buffers. The protein appears to have eluted as dimer.

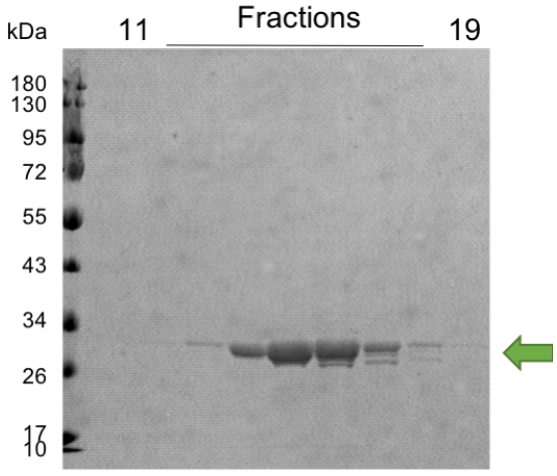

Figure 42: Coomassie Stain of SEC fractions of GST HIVrev. SEC peak analyzed by SDS-PAGE and Coomassie stain. Fractions 13-15, containing the isolated protein at expected molecular weight (~30 kDa) were pooled and stored.

The GST-LARP6(457-471) was readily purified by SEC with a single sharp peak observed at ~66 kDa, near the expected molecular weight of dimerized GST (~60 kDa) (Figure 43). The fractions under the peak were analyzed by SDS-PAGE and Coomassie stain before being pooled and frozen(Figure 44).

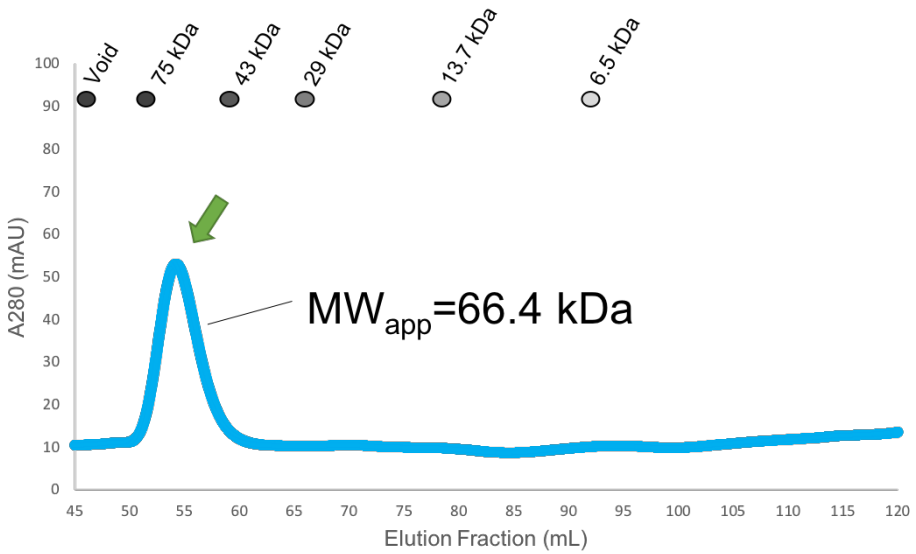

Figure 43: FPLC Purification of GST *LARP6*(457-471). SEC analyzed by A280 reading from GST LARP6 457-471 batch purification with no glycerol added to purification buffer. Again, the protein eluted as an apparent dimer.

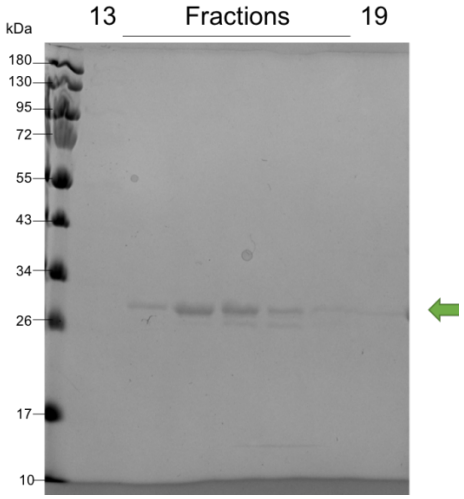

Figure 44: Coomassie Stain of SEC fractions of GST-LARP6(457-471). SEC peak analyzed by SDS-PAGE and Coomassie stain. Fractions 14-17, containing the isolated protein at expected molecular weight (~30 kDa) were pooled and stored.

The GST LARP6(179-193) was purified by SEC in the presence of glycerol with a sharp peak being observed at the expected molecular weight of dimerized GST tags ~60 kDa (Figure 45). The fractions falling under the peak were analyzed by SDS-PAGE and Coomassie stain before being pooled and frozen for future use (Figure 46). The first peak in this chromatogram has a apparent molecular weight of approximately 101 kDa and is likely a result of aggregation of this protein. This is not an unexpected result considering the difficulties in purifying this protein in the absence of glycerol. This aggregated product eluted at a volume sufficiently removed from the dimer, and was successfully excluded when selecting fractions to concentrate for storage and later use.

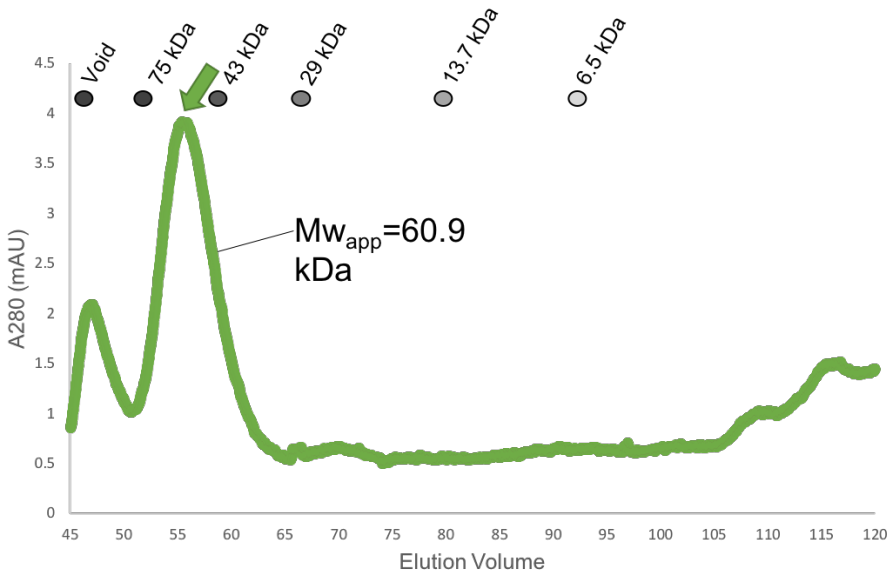

Figure 45: SEC of GSTLARP6(179-193). SEC analyzed by A280 reading fromLARP6(179-193) in purification buffers with 5% glycerol added. The protein appears to have eluted as an aggregate and as a dimer.

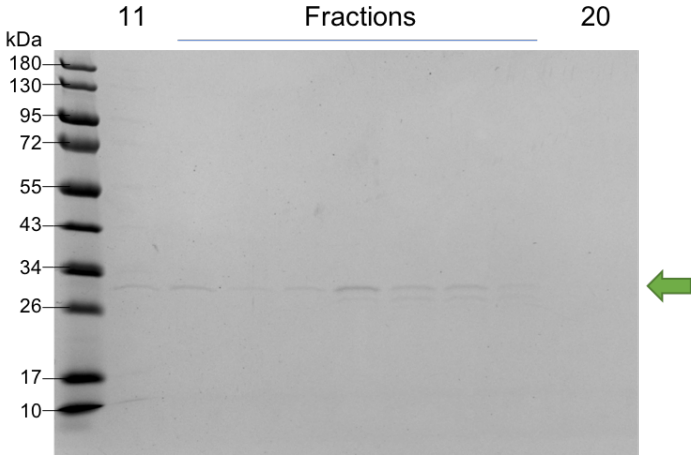

Figure 46: Coomassie Stain of SEC fractions of GST-LARP6(179-193). SEC peak analyzed by SDS-PAGE and Coomassie stain. Fractions containing the isolated protein at expected molecular weight (~30 kDa) were pooled for storage.

The GST LARP6(210-224) was not successfully purified by SEC in the presence of glycerol. There was no sharp peak observed in the expected elution fractions. A later peak of absorbance at 280 nm was observed (Figure 47), but there were no proteins detected by Coomassie stain (Figure 49). Abnormalities in the concentration of the protein before loading demonstrated why this may have failed. The pre-concentrated sample, and the concentrators retentate and flow through all appear to have equal amounts of protein indicating the concentration of this protein failed (Figure 48). This procedure should be repeated with a fresh batch of concentrators as there were abnormalities in the rate of flow for the concentration, and these results indicate that the concentrator was very likely faulty.

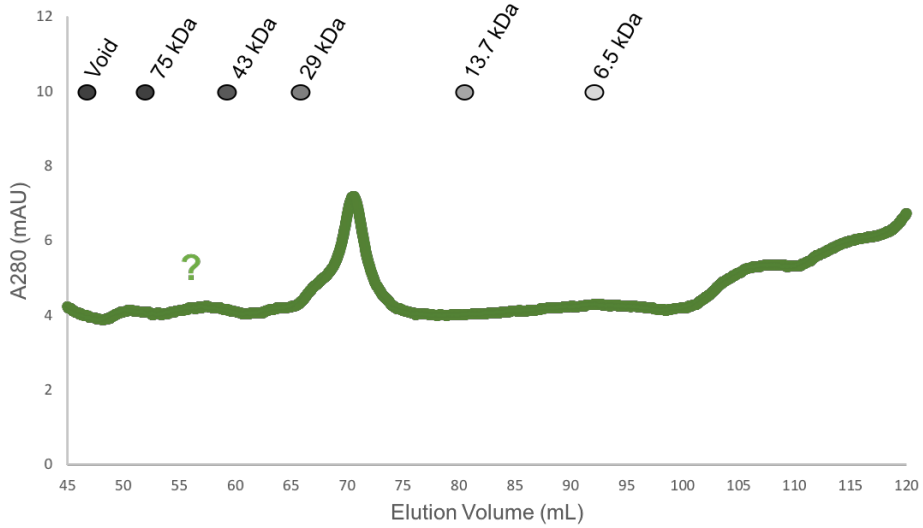

Figure 47: SEC Purification of GST LARP6(210-224). SEC analyzed by A280 reading from LARP6 210-224 in purification buffers with 5% glycerol added. The peak was shifted from the expected elution fractions by approximately 5 mL and samples under the peak were no visible by Coomassie stain. This purification was ultimately discarded.

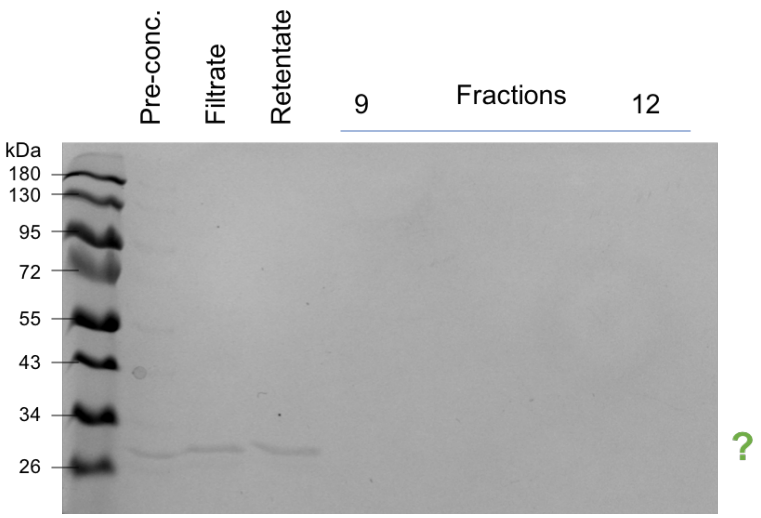

Figure 48: Coomassie Stain of SEC fractions of GST-LARP6(210-224). SEC peak analyzed by SDS-PAGE and Coomassie stain. Difficulties in concentrating the protein before loading may account for the failed purification.

Cloning of GST-LARP6

The full length LARP6 was also cloned behind the GST tag. Miniprepped pET28-LARP6 plasmid was used as the template in a PCR reaction to amplify the wildtype LARP6 with NotI and SalI cut sites. (Appendix 1, Table 1). Agarose gel electrophoresis was used to confirm successful cloning and gel-purify the product. (Figure 49). The presence of a band at ~1500 bp indicated a successful PCR. Double digest with *Not*I and *Sal*I endonucleases was performed to clone the PCR product into the pGEX4T3-TEV vector.

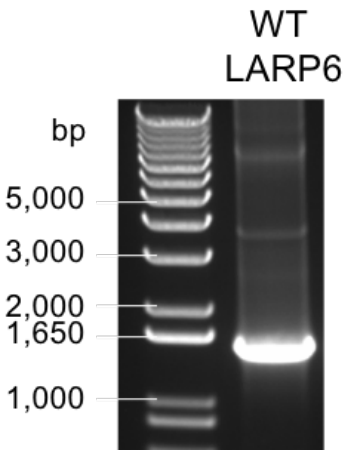

Figure 49: PCR of WT LARP6 for cloning into the pGEX4T3-TEV plasmid. PCR products were observed at the expected molecular weight (~1500 bp) indicative of successful PCR.

The recovered, ligated plasmids were confirmed by commercial Sanger sequencing and analyzed with SnapGene as described for the GST-peptide fusions. Initial expression trials were performed, but due to the low level of expression the time and temperature of expression need to be optimized (data not shown). Therefore, further optimization of expression is needed before preparative expression and purification can be performed. Once purified, this construct can be used to assess the ability of CRM1 to bind the full-length LARP6 protein *in vitro*, in the absence and presence of RNA ligand.

Ran-GTP Expression and Purification

The plasmid containing Ran-GTP was successfully purified using nickel-affinity chromatography and size exclusion chromatography. The SEC fractions were analyzed by SDS-PAGE and Coomassie stain (Figure 50). Fractions 21-24 were pooled, aliquoted, snap-frozen and at stored at -70 °C.

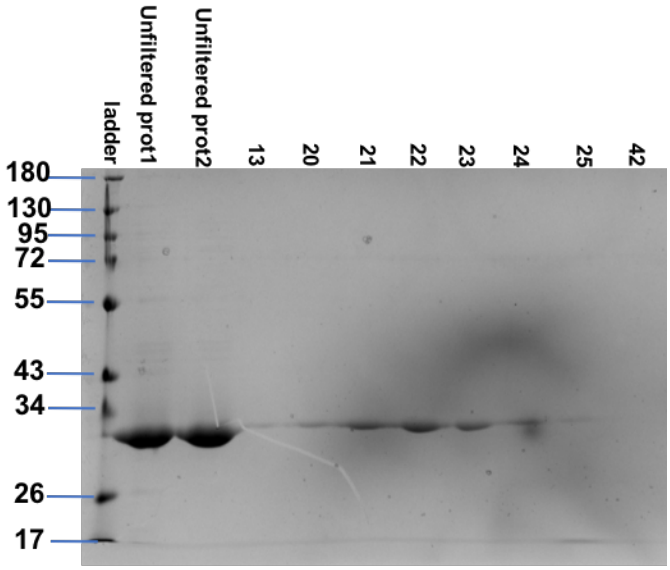

Figure 50: SEC-purified Ran-GTP (Leticia Gonzalez). SEC analyzed by A280 reading were used to select fractions to be analyzed by SDS-PAGE and Coomassie Stain. (Credit to Ms. Leticia Gonzalez for this work)

CRM1 Expression and Purification

The human CRM1 is recombinantly expressed as a GST fusion from the pGEX4T3-TEV plasmid, which was a generous gift from Dr. Yuh-Min Chook (UT-Southwestern). Several attempts at large-scale expression and purification were made by previous lab members, most notably Ms. Leticia Gonzalez but were unable to be successfully purified (Data not shown). This protein was modestly expressed in small-volume cultures (Figure 51). However, large-scale expressions demonstrated consistently low expression yields and no purified final product was able to be isolated (Data not shown). In an attempt to optimize expression, several factors were tested including cold shock (Figure 51).

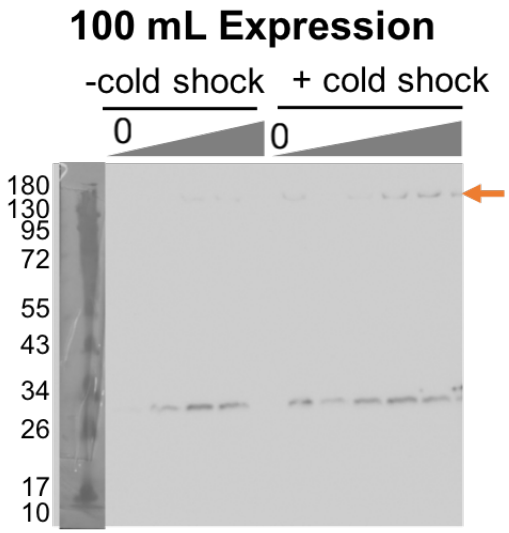

Figure 51: Small Scale 100 mL expression of *Hs*CRM1. Representative Small-Scale Expression with and without cold shock in BL21 cells. Analyzed by SDS-PAGE and anti-GST western blot. Last cold shock negative and first cold shock positive aliquot were swapped when loading.

First, expression was screened using multiple strains of E. coli. Expression was tested in BL-21(pLys), BL21(DE3), and Rosetta(DE3) both with and without a cold shock. These data showed improved expression when a cold shock was performed in BL-21(pLys) (Figure 51) compared to the other cell lines (Data not shown). However large-scale expression attempts still demonstrated modest expression compared to small scale expressions (Figure 52).

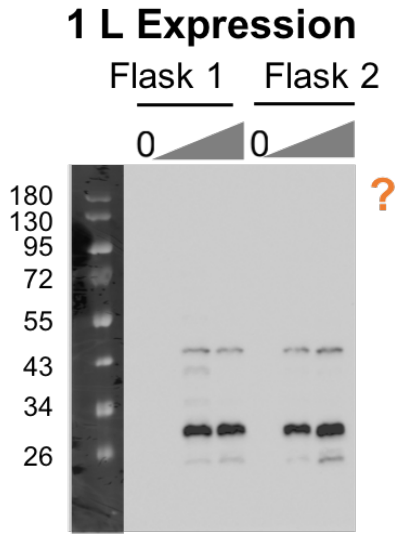

Figure 52: Large-Scale 1000 mL expression of *Hs*CRM1. Representative large-scale Expression in BL-21(pLys) cells.

In an attempt to understand the differences between large- and small-scale expression, the granulated Lennox Broth used for small scale expression and overnight cultures was compared with the capsuled Luria Broth used for 1 L expressions (Table 13).

Table 13: Broth Composition Differences for Scaled Expression.

	Capsulated Luria Broth (g/L) (1 L Expressions)	Granulated Lennox Broth (g/L) (100 mL Expressions)
Casein	10	10
Yeast extract	5	5
NaCl	10	5
Tris HCl	1.5	0

While most proteins in our laboratory's experience express better under the higher salt and pH-buffering provided by the capsulated Luria Broth, GST-CRM1 showed higher and more stable expression when grown in the lower-salt and unbuffered granulated Lennox Broth (Figure 53).

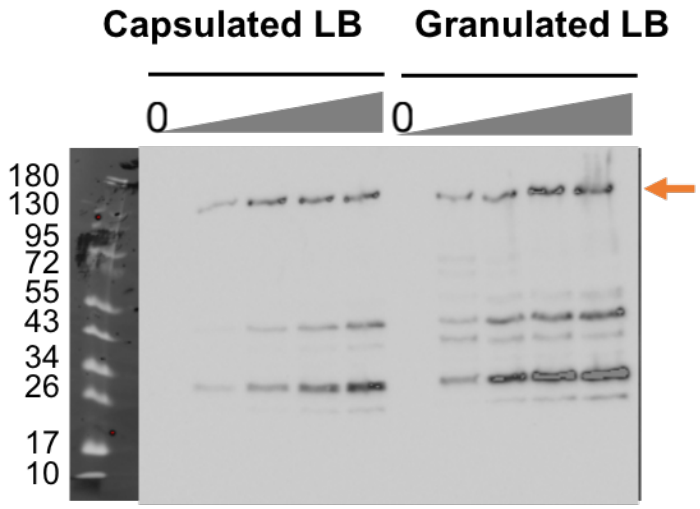

Figure 53: Effects of Broth Composition on Expression. Small-scale expression trial in BL-21(pLys) cells in different broth conditions. Cold shock was performed before induction with 1 mM IPTG. The orange arrow indicates expression at the expected molecular weight.

Therefore, a large-scale expression was performed using the granulated Lennox Broth. During this time, a western blot was also repeated and demonstrated that CRM1 degrades upon freezing (Data not shown). Further experiments were carried out by expression followed by purification immediately avoiding the freezing of the cell pellet. A subsequent trial expression of GST-CRM1 in the BL-21(pLys) cell line was carried out using granulated Lennox broth in 750 mL Lennox broth in a 2 L baffled expression flask (Figure 54). This expression showed improved results from earlier large-scale expressions (Figure 52); however, it failed to reach the same levels of expression observed in small scale (Figure 53). A small volume of CRM1 was able to be isolated from this preparation (Figure 55 -56). However, it was of such concentration that it could not be detected by Coomassie from the fractions without first concentrating the fractions (Data not shown). While concentrating made it visible by Coomassie it was of such low volume and concentration (Data not shown) as to prevent effective downstream use, which requires molar excess quantities of CRM1 to the putative NES bait.

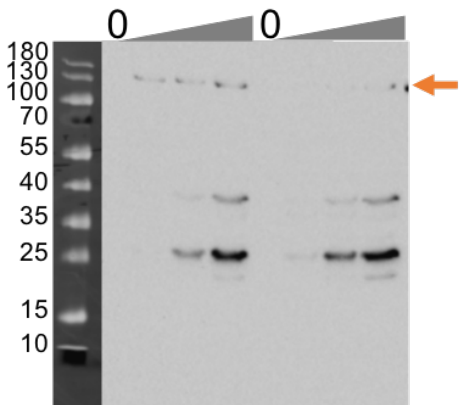

Figure 54: Large-Scale expression of CRM1 in Lennox Broth. Expression was performed in BL-21 (pLys) in a 750 mL Lennox Broth in 2L baffled flask. Time aliquots were taken throughout the expression until the time of cell harvest by centrifugation. The orange arrow indicates bands of increasing intensity at the expected molecular weight.

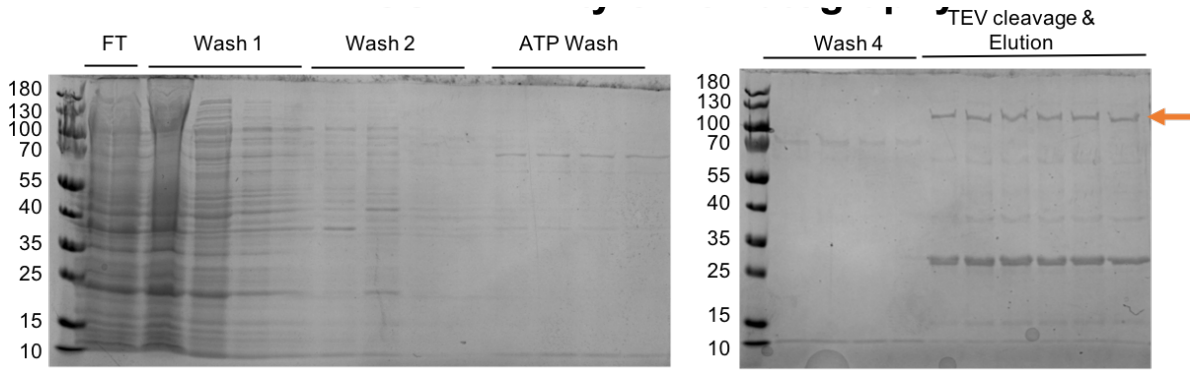

Figure 55: GSH affinity purification and TEV cleavage of CRM1. Harvested cells were lysed and incubated with the GSH resin before affinity purification and TEV cleavage. TEV cleavage and elution fractions containing the band of interest (orange arrow) were pooled and concentrated in preparation for SEC.

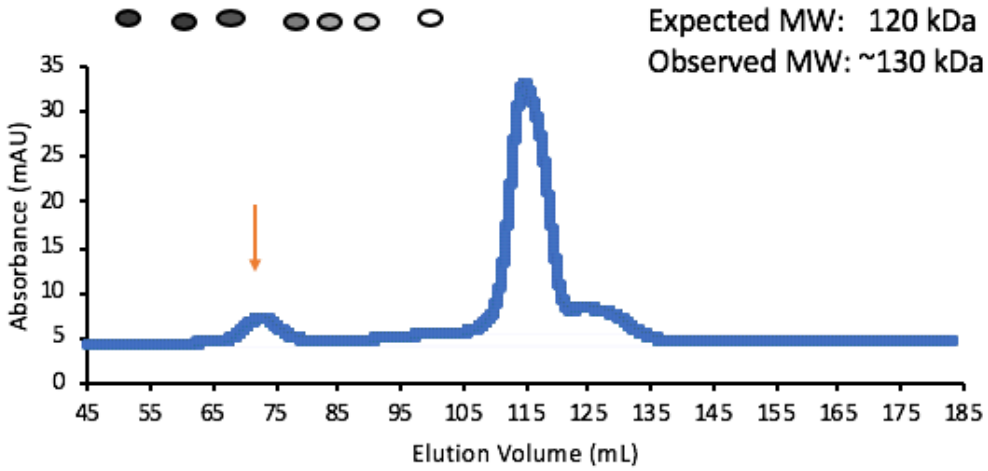

Figure 56: SEC chromatogram of recombinant CRM1. Elution was monitored by A280 readings. The expected elution peak at ~72 mL was confirmed by SDS-PAGE and Coomassie Stain after the fractions were pooled and concentrated.

On the basis of this successful purification, a 3 L expression was performed to increase the final quantity of purified protein. Unfortunately, this attempt failed despite tight adherence to the prior 1 L purification (Data not shown). The variability of the expression and purification raised many questions, so an additional expression was attempted to more closely replicate the conditions of small-scale expression through altering the flask volume to broth volume which can affect aeration. Due to the many uncertainties about expression and purification variability, aliquots were taken at every step of purification, including pre-lysis, post-sonication, and post-cell debris pelleting. These aliquots were analyzed by anti-GST western blot. Remarkably, this investigation demonstrated that sonication was causing the destruction of the GST CRM1 protein (Figure 57). For this trial, increasing bands at the expected molecular weight of 150 kDa demonstrated successful expression, although

other products were also present. This band was retained when the cell pellet from the expression was resuspend in cell lysis buffer. It disappeared completely post sonication only showing bands at lower molecular weights indicative of destruction of the protein during sonication. This explains our difficulties is purification quite well. We suspect that in all previous purification attempts, most if not all of the CRM1 present at the end of expression was destroyed during this step.

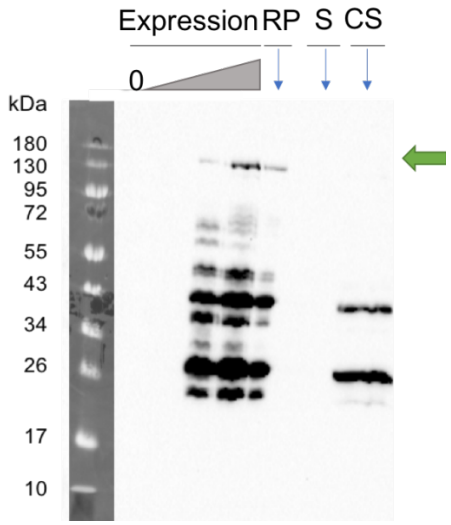

Figure 57: Destruction of CRM1 by sonication. Anti-GST western blot of cell lysis aliquots of Expression time aliquots, resuspended pellet (RP), the sonicate (S), and the Clarified Sonicate (CS). Expression increased over time at the expected MW ~150 kDa but this band was not retained post-sonication.

Delving into the original protocols for CRM1 purification from the Chook group at UT-Southwestern, it was discovered that their method utilized a French press in cell lysis. Therefore, we now suspect that recombinant human CRM1 is not stable enough to survive cell lysis by sonication, and that mechanical disruption of cells by French press is an important factor in successful purification. However, Texas State does not have a French press instrument for lysis. As a result, the next steps in this project are to attempt chemical cell lysis methods that will be gentler to the apparently delicate structure of recombinant GST-CRM1.

V. DELETING PUTATIVE NES IN RECOMBINANT LARP6

The final goal of this project was to assess the stability of putative NES deletion mutants like those used in the initial LARP6 localization studies. The rationale behind this was to determine if the NES deletion mutant used is previous work was structurally destabilizing as predicted based on the NMR solution structure.[20] To delete these large segments of coding sequence, overlap-extension polymerase chain reaction(OE-PCR) was used to create putative NES deletion mutants within the full-length LARP6. The plasmid pET28:LARP6(Δ*Bam*HI) was used as the primary template. The initial PCR reaction generates the fragments that are then used for the overlap extension step. For ease in describing these fragments, each putative NES was assigned a numerical identifier to signify which deletion construct it was associated with and whether it was associated with the N-terminal fragment or the C-terminal fragment (Table 14).

Table 14: OE-PCR fragment identifier and expected molecular weight.

Construct	Fragment	Identifier	Expected MW (bp)
LARP6 Δ(179-193)	1-178	1.1	550
LARP6 Δ(179-193)	194-491	1.2	915
LARP6 Δ(186-218)	1-185	2.1	575
LARP6 Δ(186-218)	219-491	2.2	841
LARP6 Δ(210-224)	1-209	3.1	646
LARP6 Δ(210-224)	225-491	3.2	823
LARP6 Δ(457-479)	1-456	4.1	1379
LARP6 Δ(457-479)	480-591	4.2	160

The initial PCR reaction was carried out as previously described and analyzed by gel electrophoresis (Figure 58). The fragments were excised from the gel and gel purified as previously described.

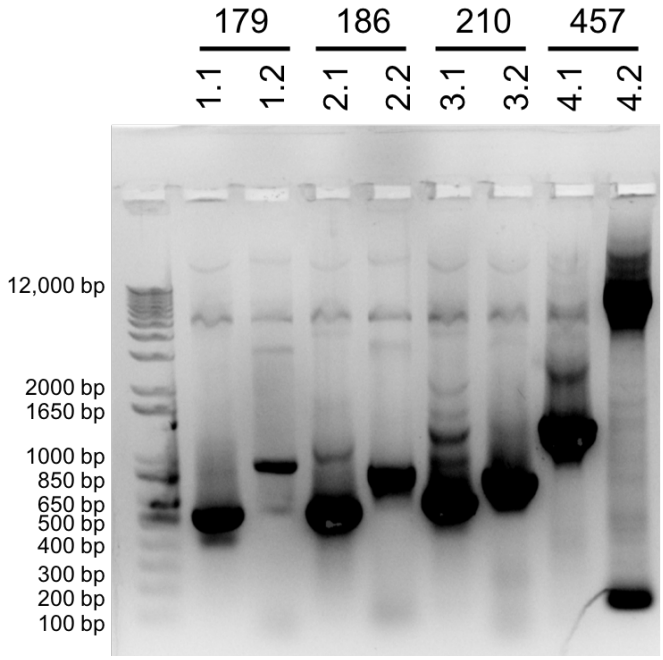

Figure 58: Putative NES deletion mutants OE-PCR fragments. PCR products of the fragments to be used as megaprimers for subsequent reactions were run on a 0.9% agarose gel and stained with ethidium bromide before being isolated by gel purification. Bands were observed at expected molecular weights for all fragments (Table 1).

Gel purified fragments were then used in the OE-PCR reaction, in which they serve as both megaprimers and templates. This “overlap extension” step elongates each of the fragments and results in a coding sequence that contains deletion mutants. Following the overlap extension PCR reaction, the number of copies of the deletion mutant product are too low to be seen by gel electrophoresis. Therefore, a PCR cleanup kit was used to isolate the PCR products. The isolated PCR products were then used as a template in an PCR amplification reaction to generate products that are both visible on an agarose

gel. The products were analyzed by gel electrophoresis and gel purified (Data not shown). As of this thesis, only the LARP6(Δ457-479) has been successfully cloned.

After confirmation of the successful OE-PCR, a final PCR reaction was performed to add the appropriate cut sites for ligation into either the pET28-SUMO or the pGEX4T3-TEV expression vectors (Table 1). This final PCR product was analyzed and purified by gel electrophoresis (Data not shown). These products were then double digested as previously described and gel purified (Figure 59).

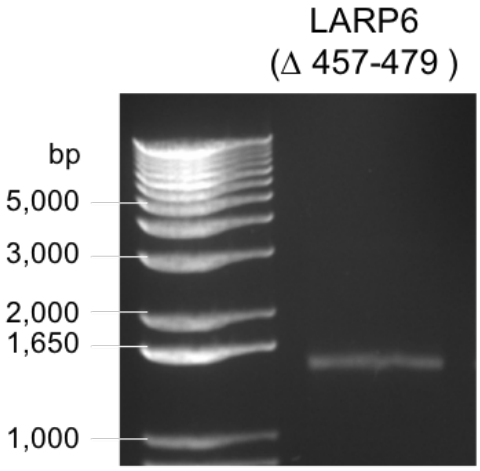

Figure 59: Double Digest LARP6(Δ457-479) with *Bam*HI and *Not*I. Double digested LARP6(Δ457-479) was run on 1.2% Agarose gel. A band at the expected molecular weight of ~1500 was used as

Colony PCR reactions were performed to confirm the presence of the ligated insert before miniprepped DNA was sent out for commercial Sanger sequencing (Data not shown). The sequence-verified SUMO LARP6(Δ457-471) was transformed into Rosetta cells for expression. Initial trial expressions have been successful as determined by Coomassie and anti-his western blot (Data not

shown). Several purification attempts by Ms. Leticia Gonzalez were made to obtain SUMO-LARP6 (Data not shown). Unfortunately, the purification of this construct needs to be optimized as it has, thus far, demonstrated an unexpected secondary band. As this will act as the control for the stability assays, its purification must be optimized before proceeding with the purification of the deletion mutant. Additionally, the other three NES deletion mutants must be cloned.

VI. DISCUSSION

Together, this thesis serves as the first set of studies into the molecular mechanisms of nuclear export of LARP6. It demonstrates the challenges associated in the complex processes that are involve in the nuclear export of proteins. The identification and characterization of these complex processes requires careful consideration of the cellular pathways involved, the intermolecular interactions between its components, and the stability of the component proteins.

The aim of identifying endogenous LARP6 localization and exploring the CRM1/LARP6 interactions *in vivo* demonstrated the importance of reproducibility and the unique challenges of working with human cell lines. The potential changes of cell line expression over time must be considered in the design of future experiments of endogenous LARP6 expression *in vivo.* In addition to this there are a great number of challenges associated with the use of commercial antibodies, which should be validated before any future experiments occur. This should be performed with cells during early on in passaging. Cell lines should also be stringently chosen based on endogenous LARP6 expression and RT-qPCR will allow for the best selection of cell line for this purpose. Our collaborator, Dr. Lisa Warner at Boise State University has seen strong and specific signal in the cell line PC12, which may make this cell line amenable to this type of study. In addition, the top 10 cell lines identified on the Human Protein Atlas with endogenous LARP6 RNA expression are good cell lines to screen.

Lastly, it must be considered that not all proteins are present in cells at level appropriate for endogenous level detection by antibodies, leaving two potential avenues of future directions. The first is transfection of GFP-tagged LARP6 into human cell lines. The second is using a CRISPR CAS9 based system to express GFP tagged LARP6 at endogenous levels. It is recommended the next step in this study due to its minimization of mitigating factors is to use a GFP tagged LARP6 transfected for over expression in the cell. While this does have some potential consequences, in altering a potentially delicate balance of complex and interrelated cellular processes, it is the ideal next step to attempt LMB inhibition of CRM1. This is because it most closely mirrors the previously published studies and minimizes the number of considerations (cell line, antibody selection, method of transfection, under/overexpression, etc.) required. If such results mirror those observed in the previous localization study which utilized NES deletion mutants, it would be a strong indicator that nuclear export is occurring by the CRM1 mediated nuclear export pathway.

As an alternative to GFP tags, the same method of transfection can be used with HA or other alternatively tagged LARP6. This could be detected with a robustly confirmed commercial antibody for that particular tag. This would remove concerns about the specificity of the primary antibody. However, will still require extensive screening of antibody and fixative methods to optimize. This method combined with a transfection method that preserves endogenous LARP6 expression would best address our aims to remove the potential interference of the bulky GFP in cellular processes and examine endogenous levels of LARP6.

Finally, actin is a protein cargo for CRM1 mediated nuclear export and should be used as a positive control for LMB activity in future LMB studies.

The second aim of this study was to identify if other putative NES candidates exist and which if any are capable of being bound by CRM1. This system demonstrates the complexity of nuclear export which not only requires a very large exporting (~120 kDa) but also the energy molecule Ran-GTP and the NES “bait”. While the expression Ran-GTP and many of the bait proteins proved successful, we have evidence that the high levels of hydrophobicity that are required for NES interaction with the binding pocket on CRM1 can complicate its purification. The success of affinity purification with glycerol added for LARP6(179-193) and LARP6(210-224) demonstrate that their inherent hydrophobicity played a role in the difficulties in purification. Large proteins such as CRM1 are extremely difficult to express in large quantities in *E. coli* cells. In addition, they appear sensitive to cell lysis by sonication. This requires further studies to focus on alternative methods of purification such as chemical lysis. Additionally, all GST tagged NES must be validated by mass spectrometry to ensure the NES is intact behind the GST tag. Finally, pull-down assays performed in the presence and absence of Ran-GTP and LMB should be performed to test for specific interactions between the full length LARP6, its putative NES and CRM1.

The final aim of this study was to determine NES deletion mutants like those used in previous studies significantly alter the stability of the LARP6 protein. While one NES deletion mutant has been successfully cloned and

expressed, difficulties in the purification of SUMO WT LARP6 must first be overcome before testing of this hypothesis can occur. In addition, the other three NES deletion mutants must have the OE-PCR reaction conditions optimized and be expressed and purified before we can fully address this aim.

Together these findings demonstrate the challenges involved in identification of nuclear export mechanisms. This work articulates several new avenues of study for identifying the localization mechanisms of LARP6 and have developed critical tools to carry out these studies. It lays a strong foundation for future work to develop a complete understanding of the role of LARP6 in cellular physiology and the regulation of gene expression.

APPENDIX SECTION

Table 1: Experimentally confirmed, Putative and Randomized NES Oligonucleotides

Protein	Sequence	Description
GST LARP6 179-193	AATTCCGAAAACCTGCCGAGCAAAATGTTGCTGGTGTATGATCTCTACCTGC	FWD LARP6(179-193)
	TCGAGCAGGTAGAGATCATACACCAGCAACATTTTGCTCGGCAGGTTTTCGG	RC LARP6(179-193)
GST LARP6 210-224	AATTCCCAGGAAAAAGTGATGGAGCATCTGTTGAAACTGTTCGGCACCTTTC	FWD LARP6 210-224
	TCGAGAAAGGTGCCGAACAGTTTCAACAGATGCTCCATCACTTTTTCCTGGG	RC LARP6 210-224
GST LARP6 457-471	AATTCCAAAATGCAGACCGCGGATGGCCTGCCGGTGGGTGTTTTGCGCCTGC	FWD LARP6 457-471
	TCGAGCAGGCGCAAAACACCCACCGGCAGGCCATCCGCGGTCTGCATTTTGG	RC LARP6 457-471
GST HIVrev NES	AATTCCCCGCTGCAGTTGCCACCGCTGGAACGCTTAACCCTGC	FWD HIVrev NES
	TCGAGCAGGGTTAAGCGTTCCAGCGGTGGCAACTGCAGCGGGG	RC HIVrev NES
GST PKIα NES	AATTCCAACGAACTGGCGTTGAAACTGGCGGGCTTAGATATTC	FWD PKIα NES
	TCGAGAATATCTAAGCCCGCCAGTTTCAACGCCAGTTCGTTGG	RC PKIα NES
GST Mad1 NES	AATTCCGCGCAGACCACGATTCAGCTGTTGCAAGAAAAATTAGAAAAACTGC	FWD Mad1 NES
	TCGAGCAGTTTTTCTAATTTTTCTTGCAACAGCTGAATCGTGGTCTGCGCGG	RC Mad1 NES
GST Randomized *LARP6* 179 -193	AATTCCCTGTATTTGAAAAACCTGATGCTCCCGAGCGAAGATCTGTACGTGC	FWD Rand. LARP6(179-193)
	TCGAGCACGTACAGATCTTCGCTCGGGAGCATCAGGTTTTTCAAATACAGGG	RC Rand. LARP6(179-193)
GST Randomized *LARP6* 210-224	AATTCCGAACATAAACTGGTGTTTGGCTTCTTGAAACTGACCATGCAGGAGC	FWD Rand. LARP6 210-224
	TCGAGCTCCTGCATGGTCAGTTTCAAGAAGCCAAACACCAGTTTATGTTCGG	RC Rand. LARP6 210-224

Table 2: PCR Primers

Protein	Sequence	Description
GST LARP6 186-218	CCAGCGAATTCACTCCTGGTCTATGATCTCTACTTGTCTCCTAAGC	PCR fwd: LARP6(186) into pGEX-4T3 (EcoRI)
	GCCACTCGAGGAGCAGGTGTTCCATCACCTTCTCT	PCR rev: LARP6(218) into pGEX-4T3 (XhoI)
GST negative control SDM	AATTCCCGGGTCAGACTCGAGCTGCAGCG	SDM fwd: pGEX-4T3-TEVeng, push C-term GST tail into frame for pull down neg ctrl
	CGCTGCAGCTCGAGTCTGACCCGGGAATT	SDM rev: pGEX-4T3-TEVeng, push C-term GST tail into frame for pull down neg ctrl
GST LARP6 WT	GATGTCGACATGGCCCAGTCCGGCGGG	PCR fwd: LARP6 into pGEX4T3 (SalI)
	CGCGGCCGCTTATACACAGGCCC	PCR rev: LARP6 into pGEX4T3 and/or pET28-SUMO (NotI)
SUMO LARP6	GATGGATCCATGGCCCAGTCCGGCGGG	PCR fwd: LARP6 into pET28-SUMO (BamHI)
	CGCGGCCGCTTATACACAGGCCC	PCR rev: LARP6 into pGEX4T3 and/or pET28-SUMO (NotI)
SUMO LARP6 (∆179-193)	ATGGCCCAGTCCGGCGGGGA	OE-PCR: LARP6(∆179-193)Frag1.FOR
	CCCACAGCTTAGGAGAGTTGGGGAACAGTGGGACGGGGGT	OE-PCR: LARP6(∆179-193)Frag1.REV
	CCCACTGTTCCCCAACTCTCCTAAGCTGTGGGCTCTGGCCACCC	OE-PCR: LARP6(∆179-193)Frag2.FOR
	TTATACACAGGCCCTGCTCCTCTCATGGCCATGA	OE-PCR: LARP6(∆179-193)Frag2.REV
SUMO LARP6 (∆186-218)	ATGGCCCAGTCCGGCGGGGA	OE-PCR: LARP6(∆186-218)Frag1.FOR
	CCAAAAGTCCCAAAAAGCTTCATCTTGCTGGGGAGGTTCTCGTTGGGGA	OE-PCR: LARP6(∆186-218)Frag1.REV
	GAACCTCCCCAGCAAGATGAAGCTTTTTGGGACTTTTGGAGTCATCTCATCAGTGCGGA	OE-PCR: LARP6(∆186-218)Frag2.FOR
	TTATACACAGGCCCTGCTCCTCTCATGGCCATGA	OE-PCR: LARP6(∆186-218)Frag2.REV
SUMO LARP6 (∆210-224)	ATGGCCCAGTCCGGCGGGGA	OE-PCR: LARP6(∆210-224)Frag1.FOR
	GCACTGATGAGATGACTCCCACCCTTCCATTCTTCTGGGGGGTGGC	OE-PCR: LARP6(∆210-224)Frag1.REV
	CCAGAAGAATGGAAGGGTGGGAGTCATCTCATCAGTGCGGATCCTCAAACCTGG	OE-PCR: LARP6(∆210-224)Frag2.FOR

	TTATACACAGGCCCTGCTCCTCTCATGGCCATGA	OE-PCR: LARP6(∆210-224)Frag2.REV
SUMO LARP6 (∆457-471)	ATGGCCCAGTCCGGCGGGGA	OE-PCR: LARP6(∆457-471)Frag1.FOR
	GGACCCCTGGGCCGGGAGAGCAGGGGACTCGTACC	OE-PCR: LARP6(∆457-471)Frag1.REV
	GCTCTCCCGGCCCAGGGGTCCTGACAACACCAGAG	OE-PCR: LARP6(∆457-471)Frag2.FOR
	GGTGGCAGCAGCCAACTCAGCTTCCT	OE-PCR: LARP6(∆457-471)Frag2.REV

Table 3: Protein sequences

GST *LARP6*			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSENLY FQGGSPNSRV DMAQSGGEAR PGPKTAVQIR VAIQEAEDVD ELEDEEEGAE TRGAGDPARY LSPGWGSASE EEPSRGHSGT TASGGENERE DLEQEWKPPD EELIKKLVDQ IEFYFSDENL EKDAFLLKHV RRNKLGYVSV KLLTSFKKVK HLTRDWRTTA HALKYSVVLE LNEDHRKVRR TTPVPLFPNE NLPSKMLLVY DLYLSPKLWA LATPQKNGRV QEKVMEHLLK LFGTFGVISS VRILKPGREL PPDIRRISSR YSQVGTQECA IVEFEEVEAA IKAHEFMITE SQGKENMKAV LIGMKPPKKK PAKDKNHDEE PTASIHLNKS LNKRVEELQY MGDESSANSS SDPESNPTSP MAGRRHAATN KLSPSGHQNL FLSPNASPCT SPWSSPLAQR KGVSRKSPLA EEGRLNCSTS PEIFRKCMDY SSDSSVTPSG SPWVRRRRQA EMGTQEKSPG TSPLLSRKMQ TADGLPVGVL RLPRGPDNTR GFHGHERSRA CV	Theoretical PI	7.04
		Expected MW (kDa)	82.6
		Ext. Coefficient	91260
		Ext. Coefficient (Reduced Cysteines)	90760
GST LARP6(179-193)			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSME TAIIRYIADK HNMETLGGCP KERAEISMET LEGAVLDIRY GVSRIAYSKD FETLKVDFLS KLPEMETLKM ETFEDRLCHK TYLNGDHVTH PDFMETLYDA LDVVLYMETD PMETCLDAFP KLVCFKKRIE AIPQIDKYLK SSKYIAWPLQ GWQATFGGGD HPPKSDLVPR GSENLYFQGG SPNSENLPSK METLLVYDLY LLELQRPHRD	Theoretical PI	5.03
		Expected MW (kDa)	32.6
		Ext. Coefficient	47580
		Ext. Coefficient (Reduced Cysteines)	47330

GST LARP6 210-224			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSME TAIIRYIADK HNMETLGGCP KERAEISMET LEGAVLDIRY GVSRIAYSKD FETLKVDFLS KLPEMETLKM ETFEDRLCHK TYLNGDHVTH PDFMETLYDA LDVVLYMETD PMETCLDAFP KLVCFKKRIE AIPQIDKYLK SSKYIAWPLQ GWQATFGGGD HPPKSDLVPR GSENLYFQGG SPNSQEKVME TEHLLKLFGT FLELQRPHRD	Theoretical PI	5.16
		Expected MW (kDa)	32.6
		Ext. Coefficient	44600
		Ext. Coefficient (Reduced Cysteines)	44350
GST LARP6 457-471			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSME TAIIRYIADK HNMETLGGCP KERAEISMET LEGAVLDIRY GVSRIAYSKD FETLKVDFLS KLPEMETLKM ETFEDRLCHK TYLNGDHVTH PDFMETLYDA LDVVLYMETD PMETCLDAFP KLVCFKKRIE AIPQIDKYLK SSKYIAWPLQ GWQATFGGGD HPPKSDLVPR GSENLYFQGG SPNSKMETQT ADGLPVGVLR LLELQRPHRD	Theoretical PI	5.16
		Expected MW (kDa)	32.9
		Ext. Coefficient	44600
		Ext. Coefficient (Reduced Cysteines)	44350
GST Randomized LARP6179-193)			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSME TAIIRYIADK HNMETLGGCP KERAEISMET LEGAVLDIRY GVSRIAYSKD FETLKVDFLS KLPEMETLKM ETFEDRLCHK TYLNGDHVTH PDFMETLYDA LDVVLYMETD PMETCLDAFP KLVCFKKRIE AIPQIDKYLK SSKYIAWPLQ GWQATFGGGD HPPKSDLVPR GSENLYFQGG SPNSLYLKNL METLPSEDLY VLELQRPHRD	Theoretical PI	5.03
		Expected MW (kDa)	32.6
		Ext. Coefficient	47580
		Ext. Coefficient (Reduced Cysteines)	47330
GST Randomized LARP6(210-224)			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSME TAIIRYIADK HNMETLGGCPK ERAEISMET LEGAVLDIRY GVSRIAYSKD FETLKVDFLS KLPEMETLKM ETFEDRLCHK TYLNGDHVTH PDFMETLYDA LDVVLYMETD PMETCLDAFP KLVCFKKRIE AIPQIDKYLK SSKYIAWPLQ GWQATFGGGD HPPKSDLVPR GSENLYFQGG SPNSEHKLVF GFLKLTMETQ ELELQRPHRD	Theoretical PI	5.11
		Expected MW (kDa)	32.6
		Ext. Coefficient	44600
		Ext. Coefficient (Reduced Cysteines)	44350

GST PKIα NES			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSENLY FQGGSPNSNE LALKLAGLDI LELQRPHRD	Theoretical PI	5.85
		Expected MW (kDa)	30.0
		Ext. Coefficient	44600
		Ext. Coefficient (Reduced Cysteines)	44350
GST Mad1 NES			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSENLY FQGGSPNSAQ TTIQLLQEKL EKLLELQRPH RD	Theoretical PI	6.01
		Expected MW (kDa)	30.5
		Ext. Coefficient	44600
		Ext. Coefficient (Reduced Cysteines)	44350
GST HIVrev NES			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSENLY FQGGSPNSPL QLPPLERLTL LELQRPHRD	Theoretical PI	6.01
		Expected MW (kDa)	30.1
		Ext. Coefficient	44600
		Ext. Coefficient (Reduced Cysteines)	44350
GST tag			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSME TAIIRYIADK HNMETLGGCP KERAEISMET LEGAVLDIRY GVSRIAYSKD FETLKVDFLS KLPEMETLKM ETFEDRLCHK TYLNGDHVTH PDFMETLYDA LDVVLYMETD PMETCLDAFP KLVCFKKRIE AIPQIDKYLK SSKYIAWPLQ GWQATFGGGD HPPKSDLVPR GSENLYFQGG SPNSRVDSSC SGRIVTD	Theoretical PI	5.07
		Expected MW (kDa)	30.8
		Ext. Coefficient	44600
		Ext. Coefficient Reduced Cysteines)	44350

GST *LARP6* Δ179-193			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSME TAIIRYIADK HNMETLGGCP KERAEISMET LEGAVLDIRY GVSRIAYSKD FETLKVDFLS KLPEMETLKM ETFEDRLCHK TYLNGDHVTH PDFMETLYDA LDVVLYMETD PMETCLDAFP KLVCFKKRIE AIPQIDKYLK SSKYIAWPLQ GWQATFGGGD HPPKSDLVPR GSENLYFQGG SPNSRVDMET AQSGGEARPG PKTAVQIRVA IQEAEDVDEL EDEEEGAETR GAGDPARYLS PGWGSASEEE PSRGHSGTTA SGGENEREDL EQEWKPPDEE LIKKLVDQIE FYFSDENLEK DAFLLKHVRR NKLGYVSVKL LTSFKKVKHL TRDWRTTAHA LKYSVVLELN EDHRKVRRTT PVPLFPNSPK LWALATPQKN GRVQEKVMET EHLLKLFGTF GVISSVRILK PGRELPPDIR RISSRYSQVG TQECAIVEFE EVEAAIKAHE FMETITESQG KENMETKAVL IGMETKPPKK KPAKDKNHDE EPTASIHLNK SLNKRVEELQ YMETGDESSA NSSSDPESNP TSPMETAGRR HAATNKLSPS GHQNLFLSPN ASPCTSPWSS PLAQRKGVSR KSPLAEEGRL NCSTSPEIFR KCMETDYSSD SSVTPSGSPW VRRRRQAEME TGTQEKSPGT SPLLSRKMET QTADGLPVGV LRLPRGPDNT RGFHGHERSR ACV	Theoretical PI	5.68
		Expected MW (kDa)	85.0
		Ext. Coefficient	88280
		Ext. Coefficient (Reduced Cysteines)	87780
GST LARP6 Δ186-218			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSENLY FQGGSPNSRV DMAQSGGEAR PGPKTAVQIR VAIQEAEDVD ELEDEEEGAE TRGAGDPARY LSPGWGSASE EEPSRGHSGT TASGGENERE DLEQEWKPPD EELIKKLVDQ IEFYFSDENL EKDAFLLKHV RRNKLGYVSV KLLTSFKKVK HLTRDWRTTA HALKYSVVLE LNEDHRKVRR TTPVPLFPNE NLPSKMKLFG TFGVISSVRI LKPGRELPPD IRRISSRYSQ VGTQECAIVE FEEVEAAIKA HEFMITESQG KENMKAVLIG MKPPKKKPAK DKNHDEEPTA SIHLNKSLNK RVEELQYMGD ESSANSSSDP ESNPTSPMAG RRHAATNKLS PSGHQNLFLS PNASPCTSPW SSPLAQRKGV SRKSPLAEEG RLNCSTSPEI FRKCMDYSSD SSVTPSGSPW VRRRRQAEMG TQEKSPGTSP LLSRKMQTAD GLPVGVLRLP RGPDNTRGFH GHERSRACV	Theoretical PI	6.84
		Expected MW (kDa)	78.8
		Ext. Coefficient	82780
		Ext. Coefficient (Reduced Cysteines)	82280
GST LARP6 Δ210-224			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSME TAIIRYIADK HNMETLGGCP KERAEISMET LEGAVLDIRY GVSRIAYSKD FETLKVDFLS KLPEMETLKM ETFEDRLCHK TYLNGDHVTH PDFMETLYDA LDVVLYMETD PMETCLDAFP KLVCFKKRIE	Theoretical PI	5.61
		Expected MW (kDa)	85.0

	AIPQIDKYLK SSKYIAWPLQ GWQATFGGGD HPPKSDLVPR GSENLYFQGG SPNSRVDMET AQSGGEARPG PKTAVQIRVA IQEAEDVDEL EDEEEGAETR GAGDPARYLS PGWGSASEEE PSRGHSGTTA SGGENEREDL EQEWKPPDEE LIKKLVDQIE FYFSDENLEK DAFLLKHVRR NKLGYVSVKL LTSFKKVKHL TRDWRTTAHA LKYSVVLELN EDHRKVRRTT PVPLFPNENL PSKMETLLVY DLYLSPKLWA LATPQKNGRV GVISSVRILK PGRELPPDIR RISSRYSQVG TQECAIVEFE EVEAAIKAHE FMETITESQG KENMETKAVL IGMETKPPKK KPAKDKNHDE EPTASIHLNK SLNKRVEELQ YMETGDESSA NSSSDPESNP TSPMETAGRR HAATNKLSPS GHQNLFLSPN ASPCTSPWSS PLAQRKGVSR KSPLAEEGRL NCSTSPEIFR KCMETDYSSD SSVTPSGSPW VRRRRQAEME TGTQEKSPGT SPLLSRKMET QTADGLPVGV LRLPRGPDNT RGFHGHERSR ACV	Ext. Coefficient	91260
		Ext. Coefficient (Reduced Cysteines)	90760
GST LARP6 Δ457-471			
	MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSME TAIIRYIADK HNMETLGGCP KERAEISMET LEGAVLDIRY GVSRIAYSKD FETLKVDFLS KLPEMETLKM ETFEDRLCHK TYLNGDHVTH PDFMETLYDA LDVVLYMETD PMETCLDAFP KLVCFKKRIE AIPQIDKYLK SSKYIAWPLQ GWQATFGGGD HPPKSDLVPR GSENLYFQGG SPNSRVDMET AQSGGEARPG PKTAVQIRVA IQEAEDVDEL EDEEEGAETR GAGDPARYLS PGWGSASEEE PSRGHSGTTA SGGENEREDL EQEWKPPDEE LIKKLVDQIE FYFSDENLEK DAFLLKHVRR NKLGYVSVKL LTSFKKVKHL TRDWRTTAHA LKYSVVLELN EDHRKVRRTT PVPLFPNENL PSKMETLLVY DLYLSPKLWA LATPQKNGRV QEKVMETEHL LKLFGTFGVI SSVRILKPGR ELPPDIRRIS SRYSQVGTQE CAIVEFEEVE AAIKAHEFME TITESQGKEN METKAVLIGM ETKPPKKKPA KDKNHDEEPT ASIHLNKSLN KRVEELQYME TGDESSANSS SDPESNPTSP METAGRRHAA TNKLSPSGHQ NLFLSPNASP CTSPWSSPLA QRKGVSRKSP LAEEGRLNCS TSPEIFRKCM ETDYSSDSSV TPSGSPWVRR RRQAEMETGT QEKSPGTSPL LSRPRGPDNT RGFHGHERSR ACV	Theoretical PI	5.60
		Expected MW (kDa)	85.2
		Ext. Coefficient	91260
		Ext. Coefficient (Reduced Cysteines)	90760
SUMO *LARP6* WT			
	HHHHHHHHHH SSGHIEGRHM ASMSDSEVNQ EAKPEVKPEV KPETHINLKV SDGSSEIFFK IKKTTPLRRL MEAFAKRQGK EMDSLRFLYD GIRIQADQTP EDLDMEDNDI IEAHREQIGG SMAQSGGEAR PGPKTAVQIR VAIQEAEDVD ELEDEEEGAE TRGAGDPARY LSPGWGSASE EEPSRGHSGT TASGGENERE DLEQEWKPPD EELIKKLVDQ IEFYFSDENL EKDAFLLKHV RRNKLGYVSV KLLTSFKKVK HLTRDWRTTA HALKYSVVLE LNEDHRKVRR TTPVPLFPNE NLPSKMLLVY DLYLSPKLWA LATPQKNGRV QEKVMEHLLK LFGTFGVISS VRILKPGREL PPDIRRISSR YSQVGTQECA IVEFEEVEAA IKAHEFMITE SQGKENMKAV LIGMKPPKKK PAKDKNHDEE PTASIHLNKS LNKRVEELQY MGDESSANSS SDPESNPTSP MAGRRHAATN KLSPSGHQNL FLSPNASPCT	Theoretical PI	6.88
		Expected MW (kDa)	68.6
		Ext. Coefficient	48150
		Ext. Coefficient (Reduced Cysteines)	47900

	SPWSSPLAQR KGVSRKSPLA EEGRLNCSTS PEIFRKCMDY SSDSSVTPSG SPWVRRRRQA EMGTQEKSPG TSPLLSRKMQ TADGLPVGVL RLPRGPDNTR GFHGHERSRA CV		
SUMO *LARP6* Δ179-193			
	HHHHHHHHHH SSGHIEGRHM ASMSDSEVNQ EAKPEVKPEV KPETHINLKV SDGSSEIFFK IKKTTPLRRL MEAFAKRQGK EMDSLRFLYD GIRIQADQTP EDLDMEDNDI IEAHREQIGG SMAQSGGEAR PGPKTAVQIR VAIQEAEDVD ELEDEEEGAE TRGAGDPARY LSPGWGSASE EEPSRGHSGT TASGGENERE DLEQEWKPPD EELIKKLVDQ IEFYFSDENL EKDAFLLKHV RRNKLGYVSV KLLTSFKKVK HLTRDWRTTA HALKYSVVLE LNEDHRKVRR TTPVPLFPNS PKLWALATPQ KNGRVQEKVM EHLLKLFGTF GVISSVRILK PGRELPPDIR RISSRYSQVG TQECAIVEFE EVEAAIKAHE FMITESQGKE NMKAVLIGMK PPKKKPAKDK NHDEEPTASI HLNKSLNKRV EELQYMGDES SANSSSDPES NPTSPMAGRR HAATNKLSPS GHQNLFLSPN ASPCTSPWSS PLAQRKGVSR KSPLAEEGRL NCSTSPEIFR KCMDYSSDSS VTPSGSPWVR RRRQAEMGTQ EKSPGTSPLL SRKMQTADGL PVGVLRLPRG PDNTRGFHGH ERSRACV	Theoretical PI	7.04
		Expected MW (kDa)	66.9
		Ext. Coefficient	45170
		Ext. Coefficient (Reduced Cysteines)	44920
SUMO *LARP6* Δ457-471			
	HHHHHHHHHH SSGHIEGRHM ASMSDSEVNQ EAKPEVKPEV KPETHINLKV SDGSSEIFFK IKKTTPLRRL MEAFAKRQGK EMDSLRFLYD GIRIQADQTP EDLDMEDNDI IEAHREQIGG SMAQSGGEAR PGPKTAVQIR VAIQEAEDVD ELEDEEEGAE TRGAGDPARY LSPGWGSASE EEPSRGHSGT TASGGENERE DLEQEWKPPD EELIKKLVDQ IEFYFSDENL EKDAFLLKHV RRNKLGYVSV KLLTSFKKVK HLTRDWRTTA HALKYSVVLE LNEDHRKVRR TTPVPLFPNE NLPSKMLLVY DLYLSPKLWA LATPQKNGRV QEKVMEHLLK LFGTFGVISS VRILKPGREL PPDIRRISSR YSQVGTQECA IVEFEEVEAA IKAHEFMITE SQGKENMKAV LIGMKPPKKK PAKDKNHDEE PTASIHLNKS LNKRVEELQY MGDESSANSS SDPESNPTSP MAGRRHAATN KLSPSGHQNL FLSPNASPCT SPWSSPLAQR KGVSRKSPLA EEGRLNCSTS PEIFRKCMDY SSDSSVTPSG SPWVRRRRQA EMGTQEKSPG TSPLLSRPRG PDNTRGFHGH ERSRACV	Theoretical PI	6.76
		Expected MW (kDa)	67.1
		Ext. Coefficient	48150
		Ext. Coefficient (Reduced Cysteines)	47900

REFERENCES

[1] Crick, F. On protein synthesis. *Sym Soc Exp Biol.* 12, 138-162 (1958)

[2] Turpin, P., Ossareh-Nazari, B., and Dargemont, C. Nuclear transport and transcriptional regulation. *FEBS Lett.* 452, 82-86 (1999)

[3] Corbett, A. Post-transcriptional regulation of gene expression and human disease. *Current Opinion in Cell Biology.* 52, 96-104 (2018)

[4] Brinegar, A., and Cooper, T. Roles for RNA-binding proteins in development and disease. *Brain Res.* 1647, 1-8 doi10.1016/j.brainres.2016.02.050 (2016)
[5] Vargas, D., Mechanism of mRNA transport in the nucleus. *PNAS*. 102, 17008-17013. (2005)

[6] Giudice, G. *et al.* ATtRACT – a database pf RNA-binding proteins and associated motifs. *Database*. 1-9 doi: 10.1093/database/baw035 (2016)

[7] Lunde, B., Moore, C., and Varani, G. RNA binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol.* 8(6), 479-490 (2007)

[8] Maraia, R. et al. The La and related RNA- binding proteins (LARPs): structures, functions and evolving perspectives. *WIREs RNA*. doi: 10.1002/wrna.1430 (2017)

[9] Dong et al. Structure of the La Motif: a wingef helix domain mediates RNA binding with a conserve aromatic patch. *EMBO J.* 23, 1000-1007 (2004)
[10] Alfano *et al*. Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nautre structural and molecular biology.* 11:4, 323-329. (2004)

[11] Bousquet-Antonelli, C. and Deragon, J. A comprehensive analysis of the La Motif protein superfamily. *RNA*. 15, 750-765 (2009)

[12] Hussain, R. H., Zawawi, M. & Bayfield, M. A. Conservation of RNA chaperone activity of the human La-related proteins 4, 6 and 7. *Nucl. Acids Res.* gkt649 (2013).

[13] Stavraka, C. and Blagden, S. The La Related proteins, a family with connections to cancer. *Biomolecules.* 5, 2701-2722 (2015)

[14] Lahr, R. *et al.* La-related protein 1 (LARP1) binds the mRNA cap, blocking eIF4F assembly on TOP mRNAs. *eLife.* DOI: 10.7554/eLife.24146 (2017)

[15] Yang, R. *et al.* 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 (2011)

[16] Cai, L., Fritz, D., Stefanovic, L. & Stefanovic, B. Binding of LARP6 to the Conserved 5′ Stem–Loop Regulates Translation of mRNAs Encoding Type I Collagen. *Journal of Molecular Biology* 395, 309–326 (2010)

[17] Wang, Z. *et al.* Regulation of muscle differentiation and survival by Acheron. *Mechanisms of Development.* 126, 700-709 (2009)

[18] Valavanis *et al.* Acheron, a novel member of the Lupus antigen family in induced during the programmed cell death of skeletal muscles in the moth *Manduca sexta. Gene.* 393, 101-109 (2007)

[19] Shao et al. The novel lupus antigen protein Acheron enhances the development of human breast cancer. *International Journal of Cancer.* 130, 544-554 (2012)

[20] Martino, L. *et al.* Synergic interplay of the La motif, RRM1 and the interdomain linker of LARP6 in the recognition of collagen mRNA expands the RNA binding repertoire of the La module. *Nucl. Acids Res.* 43, 645–660 (2014).

[21] Zhang, Y. and Stefanovic, B. LARP6 meets Collagen mRNA: Specific Regulation of Type I collagen Expression. *J. Int. Mol. Sci.* 17(3), 419. doi:10.3390/ijms17030419. (2016)

[22] Kim, Y., Han, M., and Oh, S. The molecular mechanism for nuclear transport and its application. *Anatomy and Cell biology.* 50, 77-85 (2017)

[23] Chook, Y.M. and Fung H.Y. Atomic basis of CRM1-cargo recognition, release and inhibition. *Semin Cancer Biol. 0,* 52-61 (2014)

[24] Schmidt, H.B. and Gorlich, D. Transport selectivity of nuclear pores, phase separation, and membraneless organelles. *Trends Biochem. Sci.* 41, 46-61 (2016)

[25] Cautain, B. *et al.* Components and regulation of nuclear transport processes. *FEBS J.* 282, 445-6463 (2015)

[26] Fung, H.Y. and Chook, YM. Atomic basis of CRM1-cargo recognition, release and inhibition. *Semin Cancer Biol.* doi: 10.1016/j.semcancer.2014.03.002. (2014)

[27] Fornerod, M. *et al.* CRM1 is an export receptor for leucine rich nuclear export signals. *Cell.* 90, 1051-1060 (1997)

[28] Dickmanns, A., Monecke, T., and Ficner, R. Structural Basis of Targeting the Exportin CRM1 in Cancer. *Cells*. 4, 538-568 (2015)

[29] Wen *et al.* Identification of a signal for rapid export of proteins from the nucleus. *Cell*. 82, 463-473 (1995)

[30] Fischer, U. *et al.* The HIV-1 REV activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell*. 82, 475-483 (1995)

[31] Kosugi, S. *et al.* Nuclear export signal consensus defined by using a location-based yeast selection system. *Traffic*. 9, 2053-2062 (2008)

[32] The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.
[33] Kosugi, S. *et al.* NESmapper: Accurate prediction of Leucine rich nuclear export signals using activity-based profiles. *PLoS Comput. Biol.* 10(9), e1003841 (2014)

[34] Fung, H.Y., Fu, S.C., and Chook Y.M. Nuclear export receptor CRM1 recognized diverse conformations in nuclear export signals. *eLIFE.* E23961 (2017)

[35] Behrens, R. *et al.* NES masking regulates HIV-1 Rev trafficking and viral RNA nuclear export. *J. Virol.* 92(10), doi:10.1128/JVI.02107-16 (2016)

[36] Marfori et al. Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochimica et Biophysica Acta* 1813, 1562-1577 (2011)

[37] Xu, D., Farmer, A., Collett, G., Grishin, N.V., and Chook, Y. Sequence and structural analyses of export signals in the NESdb. *Molecular Biology of the cell.* 23, 3677-3693 (2012)

[38] Fu, S.C. *et al.* Prediction of leucine rich nuclear export signal containing proteins with nessential. *Nucleic Acids Res.* 39, e111(2011)
[39] Prieto, G. *et al.* Prediction of nuclear export signals using weighted regular expressions (Wregex). *Bioinformatics.* 30, 1220-1227 (2014)

[40] la Cour, T. *et al.* Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng. Des. Sel.* 17, 527-536 (2004)

[41] Xu et al. LocNES: a computational tool for location classical NESs in CRM1 cargo proteins. *Bioinformatics* 31, 1357-1365 (2015)

[42] Cromie, M. and Gao, W. Epigallocatechin-3-Gallate Enhances the Theraeutic effects of leptoycin B on the Human Lung Cancer A549 Cells. *Oxid. Med. Cell Longev.* 2015 (2015)

[43] Kudo, N. *et al*. Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proc Natl Acad Sci U S A.* 96(16) 9112-9117(1999)

[44] Sun, Q. *et al*. Nuclear export inhibition through the covalent conjugation and hydrolysis of Leptomycin B by CRM1. *PNAS.* 110(4) 1303-1308 (2013)

[45] Inoue H., Nojima H., and Okayama H. High efficiency transformation of Escherichia coli with plasmids. *Gene*. 96: 23–28 (1990)

[46] Hanahan,D. Studies on Transformation of *Escherichia coli* with Plasmids. *J. Mol. Biol.* 166, 557-580. (1983)

[47] Haan, C. and Behrmann, I. A cost effective non-commercial ECL-solution for Western blot detections yielding strong signals and low background. *J. Immunological Methods*. 318, 11-19. (2007)

[48] Schindelin, J.; Arganda-Carreras, I. & Frise, E. et al. "Fiji: an open-source platform for biological-image analysis", *Nature methods* 9(7): 676-682, PMID 22743772, doi:10.1038/nmeth.2019(on Google Scholar). (2012)

[49] Sabnis, R.W. *et al*. $DiOC_6$: a Usefule Dye for staining the Endoplasmic Reticulum. *Biotechnic & Histochemistry*. 72(5) 253-258 (1997)

[50] (n.d.) Retrieved from https://www.proteinatlas.org/ENSG00000166173-LARP6/cell.

[51] Song, M. *et al.* The conserved protein SZY-20 opposed the Plk-4-related kinase ZYG-1 to limit centromere size. *Dev Cell.* 15(6), 901-912 (2008)