Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Ayushi Sharma April 10, 2018

Purification and phosphatidylcholine specificity of StAR related lipid transfer domain containing 7

by

Ayushi Sharma

Dr. Eric Ortlund

Adviser

Department of Biology

Matthew Tillman, MSP Candidate

Adviser

Dr. William Kelly

Committee Member

Dr. Llewellyn

Committee Member

2018

Purification and phosphatidylcholine specificity of StAR related lipid transfer domain containing 7

By

Ayushi Sharma

Dr. Eric Ortlund

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Biology

2018

Abstract

Purification and phosphatidylcholine specificity of StAR related lipid transfer domain containing 7

By Ayushi Sharma

STAR related lipid transport domain containing 7, or StarD7, is a lipid binding protein that binds and shuttles phosphatidylcholine from the endoplasmic reticulum and Golgi apparatus. StarD7 is important in maintaining the mitochondria's membrane composition, which is important for oxidative phosphorylation and cellular homeostasis. StarD7 is upregulated in cancer cells, leading to a higher frequency of metastasis, suggesting that overexpression of StarD7 may also lead to cancer phenotypes [1]. Mice that have both StarD7 alleles knocked out are nonviable, while the heterozygous StarD7 variant experienced significant increases in pulmonary inflammation, mucous cell metaplasia, and airway hyperresponsiveness [2]. Knockdown HEPA-1 cells also experienced reduced oxygen consumption rates and lowered activity of respiratory enzymes. This data emphasizes the importance of StarD7's role in maintaining the mitochondrial phospholipid membrane composition and cellular homeostasis. Elucidating the structure of StarD7 using X-ray crystallography would allow us to better understand these biological mechanisms.

Determining the structure of the PC binding pocket of StarD7 would help us better understand the mechanism to which PC's bind StarD7. Additionally, knowing the specific structure of the binding pocket would lead to better predictions as to which types of PCs StarD7 has an affinity for. We hypothesize that StarD7 binds to specific PCs, which will activate mitochondria localization. To confirm our hypothesis, we must determine the molecular structure of the StarD7 protein.

The purification and isolation of StarD7 proved to be incredibly difficult in both bacterial and mammalian cell systems, as much of the protein remained insoluble, and when the gene was altered and tagged to improve solubility, the protein coordinated with heat shock proteins. In the future, we could

create a PCTP and StarD7 chimera gene to help shield the hydrophobic regions of StarD7 or potentially use different animal models of the StarD7 gene.

Purification and phosphatidylcholine specificity of StAR related lipid transfer domain containing 7

By

Ayushi Sharma

Dr. Eric Ortlund

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Biology

2018

Acknowledgements

I would like to acknowledge Dr. Eric A. Ortlund, an Associate Professor and head of the Ortlund Lab in the Department of Biochemistry at the Emory University Medical School, who has been my advisor during the course of my honors thesis. I would also like to acknowledge Matthew C. Tillman, an MSP candidate, who served as an invaluable mentor during the course of my honors thesis.

TABLE OF CONTENTS

List of Figures

Figure 22. Final StarD7 purified protein concentrated.

Figure 23. The orange indicates the N-terminus and C-terminus of the long transcript, while the orange indicates the N-terminus and C-terminus of the short transcript.

Figure 24. The graph on the right shows the crystallization potential of the truncated StarD7 transcript, versus the crystallization potential of the long version of our transcript on the left, after removing the disordered regions of the StarD7 gene.

Figure 25: First-pass purification of MBP-tagged StarD7.

Figure 26. SDS-PAGE confirmed the presence of StarD7, particularly concentrated in the highest peak (band 2).

Figure 27: Size-Exclusion Chromatography (SEC).

Figure 28. SDS-PAGE was conducted using multiple sample of the main peak.

Figure 29. First-pass purification of Lic-tagged StarD7.

Figure 30. PLVX vector containing truncated, His-tagged StarD7, LTR regions, an IRES site, and a ZS-Green tag for fluorescent detection.

Figure 31. A structural prediction of the truncated StarD7.

Figure 32. Swiss model makes a structural model of your protein based of the sequence and known crystal structures of similar proteins.

Figure 33. SWISS, a homology modeling server model predictions show a 28.8% similarity between PCTP and StarD7, and much of the conserved region lies in the binding pocket.

Figure 34. This structural prediction shows StarD7 and PCTP overlaid on the left.

Chapter 1

Introduction

Background:

The mitochondrion is an organelle composed of a double-membrane system that consists of inner and outer mitochondrial phospholipid membranes. The composition of the phospholipid membranes are carefully maintained and important for regular mitochondrial function [1]. Phosphatidylcholine (PC) is the most abundant phospholipid in mammalian cells, and is highly enriched within mitochondrial membranes [1]. Phosphatidylcholine is synthesized by biosynthetic enzymes in the endoplasmic reticulum (ER), and must be transported to the mitochondria to maintain proper mitochondrial function [2]. One of the primary ways that lipids are transported between the ER and the mitochondria is via mitochondria associated membranes, or MAMs. The MAM region of the ER has high levels of biosynthetic enzyme activity, and many of these enzymes can be reversibly tethered to the mitochondria. The import of phosphatidylserine into the mitochondria, for example, uses a contact site between the mitochondria and a MAM [3]. Most PC, conversely, is not transported through MAMs, but is localized to the mitochondria via lipid shuttle proteins such as StarD7 [4].

The START domain Proteins: An Overview

START domain proteins all share one common feature of their crystal structures: an α/β fold built around a U-shaped incomplete β-barrel as well as an internal lipid-binding pocket [5]. There are six major groups of START domain proteins (Fig. 1). The STARD1/StAR group serves as intracellular cholesterol carriers. The STARD4 group includes STARD4, STARD5, and STARD6, and their function has not been clearly elucidated yet. It is hypothesized that the STARD4 and STARD5 proteins are cholesterol or sterol-specific binding proteins [6]; additionally, the STARD6 is an androgen-binding protein [7]. The RhoGAP START group stimulate GTPase activity, while the thioesterase START group act as hydrolases. The lipid-specificity of the STARD9 group is currently unknown. The STARD2/PCTP group is more varied in its function and lipid specificity. STARD11/CERT binds and

phosphorylates Goodpasture antigen, or the C-terminus region on collagen-IV. STARD10 binds and shuttles phosphatidylcholine as well as phosphatidylethanolamine, while STARD2/PCTP and STARD7 bind and shuttle PC to the mitochondrial membrane [6].

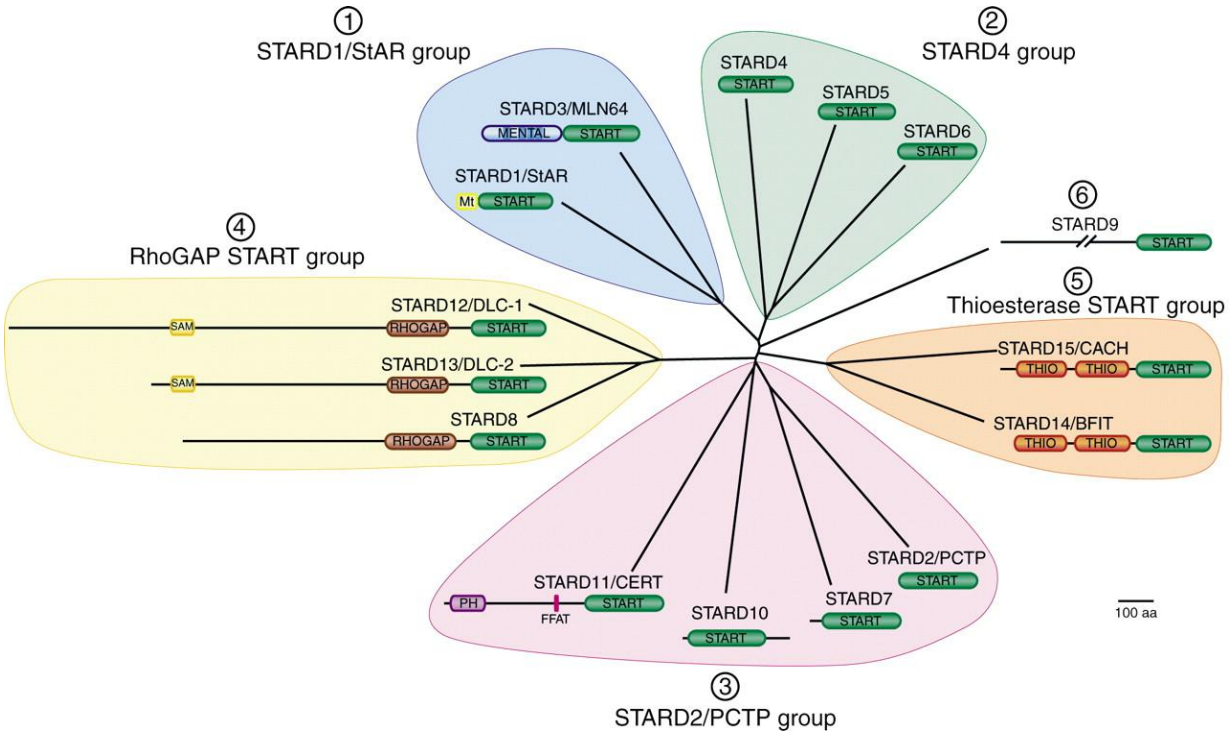

Figure 1. A phylogenetic tree showing the six groups of START domain proteins [6].

StarD7:

StarD7 is a lipid binding protein that specifically binds and transports phosphatidylcholine, or PC, to the mitochondrion [8]. Structurally, the StarD7 sequence has a mitochondria targeting sequence (MTS) that assists with localization to the mitochondria. Additionally, there is a transmembrane domain C-terminal to the MTS that tethers the protein to the mitochondrial membrane, while the START domain primarily assists with lipid shuttling.

Figure 2. StarD7 consists of the mitochondria targeting sequence, or MTS, at the N-terminus, followed by a transmembrane domain, and then the START domain.

There are two possible mechanisms of action hypothesized for StarD7. The first suggests that the transmembrane domain is tethered to the outer mitochondrial membrane and the START domain extends into the cytoplasm [9]. StarD7 exchanges PC between the outer leaflet of the ER and the outer leaflet of the mitochondrial membrane. A protease cleaves the mitochondrial localization sequence to produce the mature form of the StarD7 protein [9] (Fig. 3).

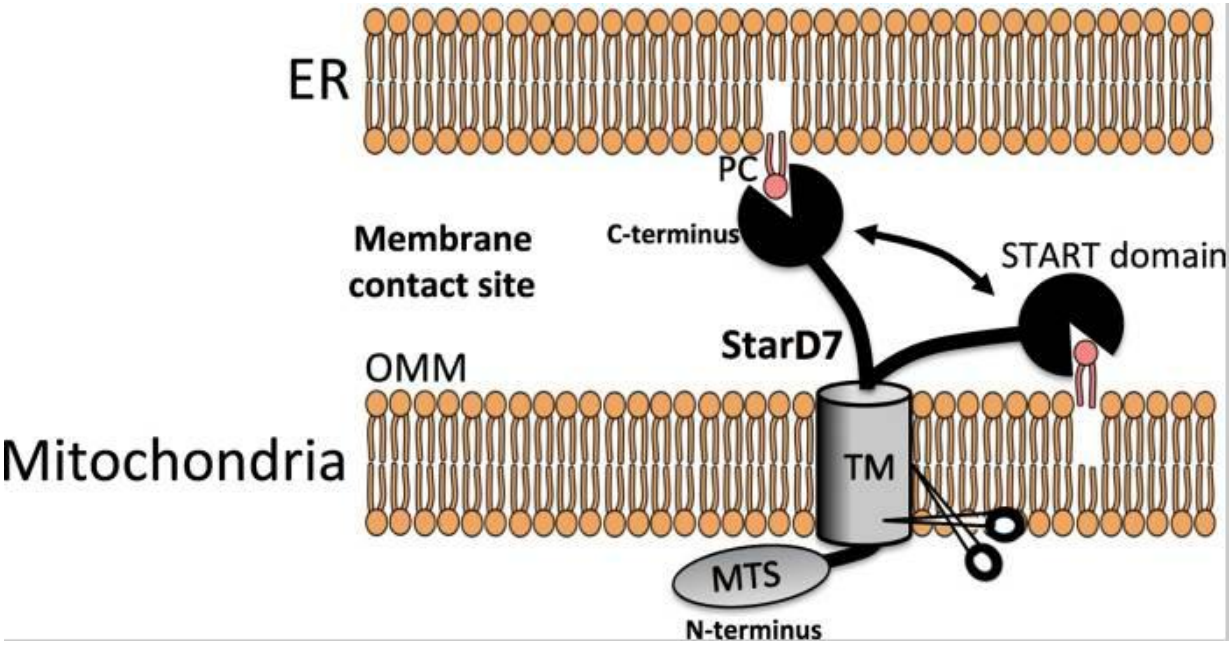

Figure 3. Diagram of one possible StarD7 mechanism. StarD7 is embedded in the outer mitochondrial matrix via its transmembrane domain, and the START domain extends into the cytoplasm and shuttles PCs [9].

The second suggests that StarD7 localizes to both the intermembrane space of the ER as well as the cytosol. StarD7 here serves as an intramitochondrial lipid transfer protein for PC, allowing PC to build up in the inner mitochondrial membrane [10]. This is consistent with the composition of the

mitochondrial membranes, in which PC is found in significantly higher concentrations in the inner mitochondrial membrane versus the outer mitochondrial membrane [11].

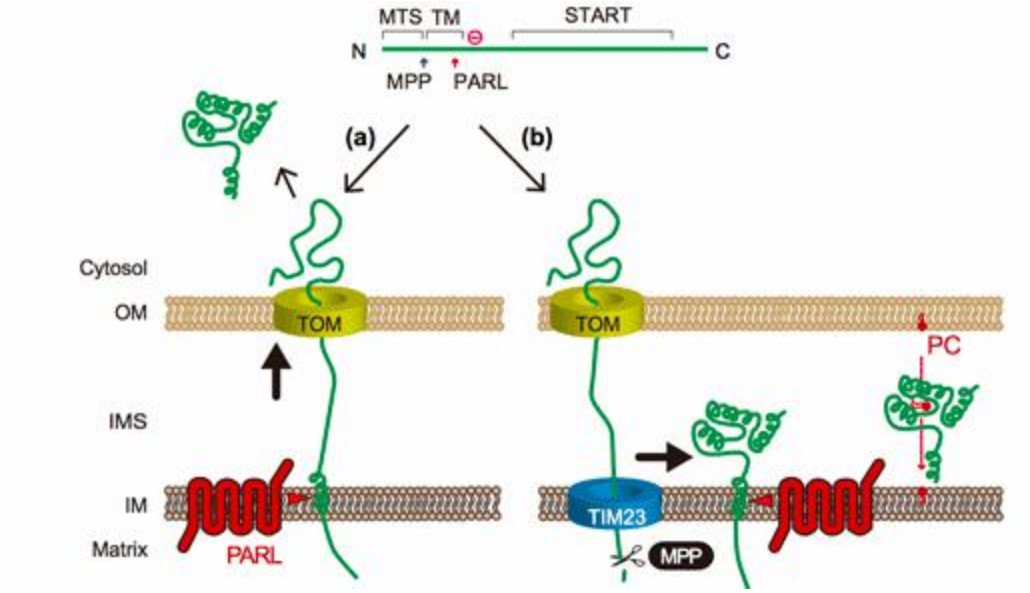

Figure 4. Diagram of another possible StarD7 mechanism in which StarD7 serves as an intramitochondrial lipid shuttle protein, exchanging PC from the outer mitochondrial membrane to the inner mitochondrial membrane [11].

StarD7 is known to be expressed in nasal epithelial cells, and is noticeably down-regulated in nasal epithelial cells in those experiencing asthma attacks versus healthy patients and those with stable asthma [1]. The phenotypic purpose of StarD7 potentially lies in its role in preventing an exaggerated allergic response by protecting mucosal tissues. Mutations in StarD7 have been shown to disrupt mitochondrial PC concentrations and lead to adverse effects. For instance, StarD7 is upregulated in cancer cells, leading to a higher frequency of metastasis, suggesting that overexpression of StarD7 may lead to cancer phenotypes [8]. StarD7 knockout mice are nonviable, while the heterozygous variant experienced significant increases in pulmonary inflammation, mucous cell metaplasia, and airway hyperresponsiveness [2]. Additionally, knockdown of StarD7 in HEPA-1 cells significantly reduced PC concentrations in the mitochondria, oxygen consumption rates and lowered activity of respiratory enzymes (Fig. 5). StarD7 knockout mice cells were also observed to have an increase lung epithelial

permeability. These data emphasizes the importance of StarD7's role in maintaining the mitochondrial phospholipid membrane composition and cellular homeostasis. Elucidating the structure of StarD7 using X-ray crystallography would allow us to better understand these biological mechanisms.

We set out to determine the first structure of StarD7 using X-ray crystallography to gain insights into the protein necessary for cellular metabolism and homeostasis. StarD7 potentially binds specific PCs in order to regulate mitochondrial PC composition [12]. Elucidating the protein's structure will help us understand the mechanism of PC's binding to StarD7 and allow us to create specific mutations to better understand the StarD7 phenotypes.

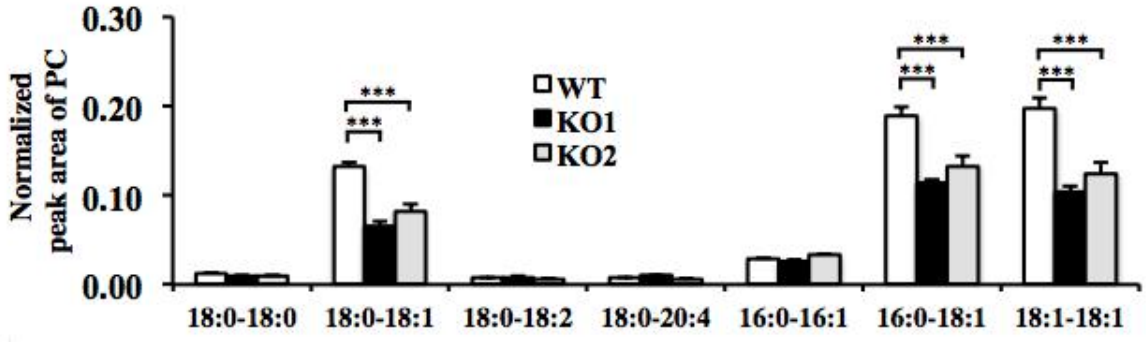

Figure 5. Mitochondrial PC content was reduced in knockout HEPA-1 cell models. Graph B shows the effect of two different knockdowns, two clones with deletions in both StarD7 alleles. It is important to note that only certain PCs were significantly reduced, while others remained relatively unchanged [12].

Chapter 1

Methods

Plasmids:

We used two different cell systems, an *E.coli* bacterial system using BL21(DE3) cells and a HEK293T mammalian cell system. We used a truncated transcript that contained residues 141 - 326 (short transcript) as well as a longer version that contained residues 112 – 370 (long transcript). These transcripts were put into either a SUMO vector, pET15b vector, or a pMCSG9 vector. For the mammalian expression, we used the pLVX ZsGreen plasmid. By using a variety of different constructs with a variety of different tags, we were able to determine which method best solubilizes and purifies the StarD7 protein.

Recombinant Protein Expression:

StarD7 was expressed in BL21(DE3) *E. coli* cells and grown in 6 liters of terrific broth media to an OD600 0.6. Protein expression was induced upon the addition of 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown at 18°C for 18 – 20 hours. Cells were harvested by spinning at 3,200 x rpm for 15 minutes. Cells were lysed in lysis buffer containing 20 mM Tris HCl pH 7.4, 150 mM NaCl, 5% Glycerol, 25 mM imidazole, lysozyme, DNase A, lecithin (for binding pocket stabilization) and phenylmethylsulfonyl

fluoride (PMSF). The mixture was agitated for 10 minutes at 4°C followed by sonication at a pulse timing of one second on and one second off for a total of eight minutes. Cell lysates were then spun at 16,000 x G for 30 minutes, and the supernatant was collected.

Test Expressions

Multiple growth conditions were screened using 5 mL test expressions of BL21(DE3) *E. coli* transiently transformed with the StarD7 vector in lysogeny broth (LB) grown at 37°C. Protein expression was induced with 0.5 mM IPTG once cultures reach OD600 0.6. Following the addition of IPTG, the temperature was either reduced to 30°C and grown for 4 hours, or reduced to 18°C and grown for 18 – 20 hours. Additionally, 0.1 % lecithin was added to the LB media for some samples before inoculation. A small sample of cells were harvested prior to induction and following induction, and were analyzed using SDS- PAGE.

Protein Purification:

StarD7 (grown at 18°C overnight) was purified from soluble lysates using nickel affinity chromatography, a protein purification technique that takes advantage of the specific coordinate covalent bond of amino acids, particularly histidine, to metals. The eluted StarD7 was collected and then 200 uL of TEV protease was added to the protein solution to cleave the SUMO tag. The solution was placed in a 10,000 Dalton pore membrane and put into dialysis buffer containing 20 mM Tris HCl pH 7.4, 150 mM NaCl, 5% Glycerol overnight to remove imidazole. The protein solution was then run over a post-TEV/thrombin Ni-column, and the flow through, which contained cleaved StarD7, was collected. Finally, the protein was purified using an anion exchanger Q column, and the StarD7 elution was once more collected and the purified protein was stored at -80°C.

Alternatively, StarD7 (grown at 18°C overnight) eluted from the nickel affinity column was then collected and run over an anion exchanger Q column. StarD7 was then cleaved with either TEV or thrombin using of 200 uL of protease, placed into a membrane with 10,000 Dalton pores, and then put in dialysis buffer overnight. The protein solution was then run over a post-TEV/thrombin Ni- column,

and the flow through containing StarD7 was collected and stored at -80°C, or further purified using a size exchange chromatography column.

Thermal Shift Assay:

An initial screen of differing protein and SYPRO Orange concentrations was used to determine optimal conditions to stabilize the StarD7 protein using a thermal shift assay. Thermal shift assays quantify the changes in protein when they are heat denatured under set conditions. The differing conditions we looked at were varying buffers, pHs, and protein/dye conditions.

Viral Transduction:

A pLVX plasmid containing long terminal repeats, the StarD7 short transcript, an internal ribosome entry site (IRES), and a GFP 2S-Green tag was transfected into adherent HEK 293T cells along with a viral plasmid to make and harvest the virus. Then we determined multiplicity of infection (MOI) via serial dilutions, and transduced the virus into two batches of suspension HEK 293F cells, one with an MOI of 9 and another with an MOI of 0.9. Then, the amount of StarD7 produced, or in other words, how well the cells were transduced, was determined using the fluorescence and a cell counter.

Chapter 2

Isolating StarD7 – pET SUMO Vector

We used a "long" StarD7 transcript in which StarD7 residues 112 – 370 were cloned into a pET SUMO vector containing a 6x His tag and a small ubiquitin-related modifier (SUMO) fusion tag, followed by a TEV cleavage site (Fig. 6). The molecular weight of the uncut sequence is 43.8 kDa and its isoelectric point is 6.03. Once the fusion tag is removed, the molecular weight of StarD7 is 30.6 kDa and the isoelectric point is 6.55.

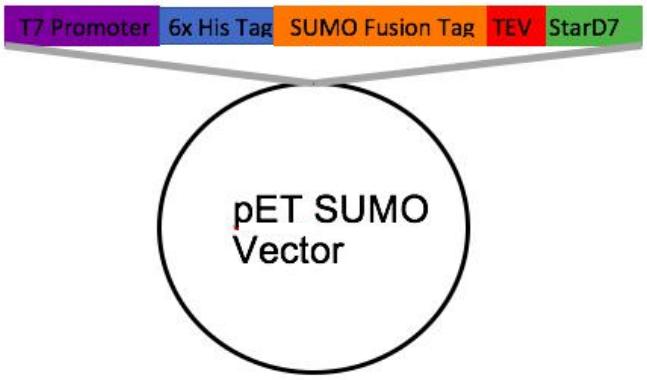

Figure 6. StarD7 SUMO Sequence:

MGHHHHHHGSLQDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAK
RQGKEMDSLRFLYDGIRIQADQAPEDLDMEDNDIIEAHREQIGGENLYFQSNAFQSSGVQHHPP
EPKAQTEGNEDSEGKEQRWEMVMDKKHFKLWRRPITGTHLYQYRVFGTYTDVTPRQFFNVQ
LDTEYRKKWDALVIKLEVIERDVVSGSEVLHWVTHFPYPMYSRDYVYVRRYSVDQENNMMV
LVSRAVEHPSVPESPEFVRVRSYESQMVIRPHKSFDENGFDYLLTYSDNPQTVFPRYCVSWMV
SSGMPDFLEKLHMATLKAKNMEIKVKDYISAKPLEMSSEAKATSQSSERKNEGSCGPARIEYA

6x His + SUMO Tag
TEV Cleavage site
MW: 43.8 kD
pI: 6.03
Cut StarD7:
MW: 30.6 kD
pI: 6.55

The goal of these experiments was to purify StarD7 from a bacterial system for the purposes of crystallography. StarD7 was recombinantly expressed using BL21(DE3*) E. coli* as an expression system. Large 6 L growths of StarD7 were grown and induced. The first method, or Method 1, starts off

with nickel affinity chromatography. Protein was eluted with increasing concentrations of imidazole in a step gradient.

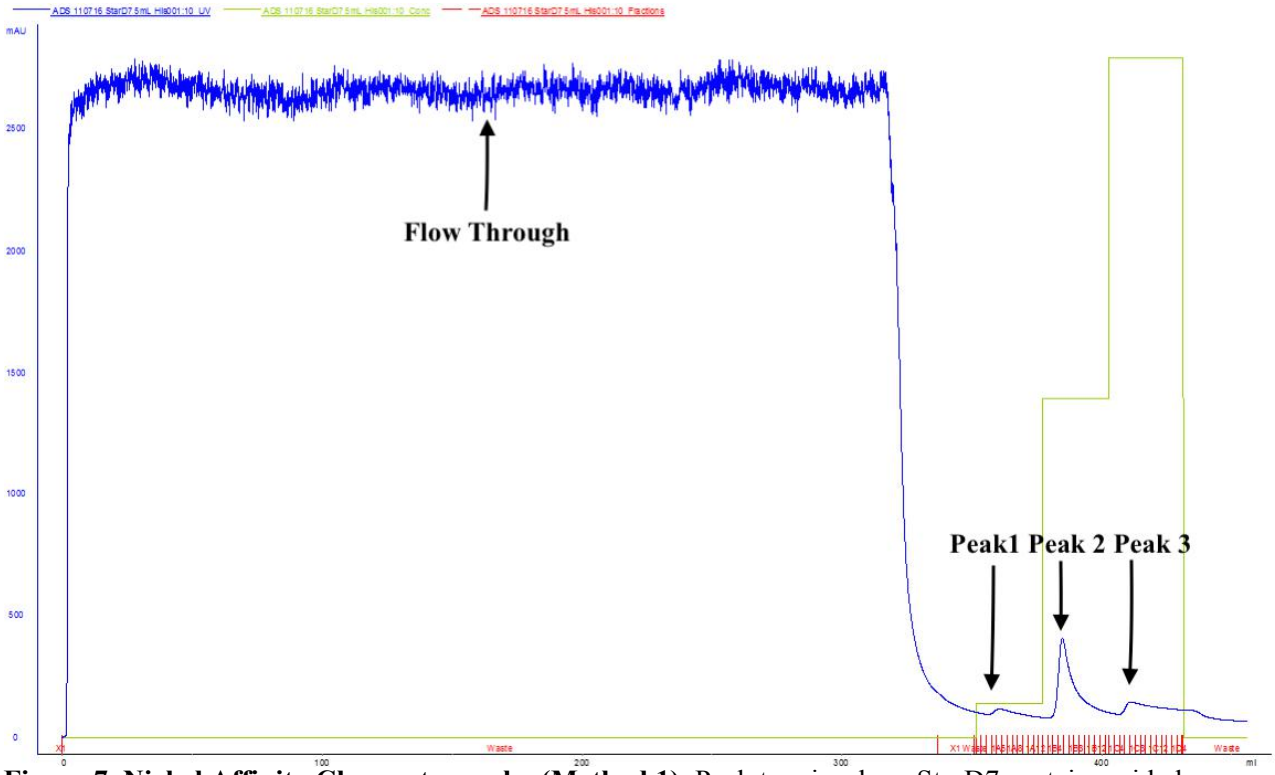

Figure 7. Nickel Affinity Chromatography (Method 1). Peak two is where StarD7 protein resided, and was ultimately collected and TEV-cleaved. The blue line corresponds to the UV 280. The green line corresponds to the concentration of imidazole within the elution buffer. The elution buffer was stepwise increased from 5% to 50% to 100% imidazole (250 mM).

A gel was run with the flow through from the nickel affinity chromatography column as a control and took samples from Peaks 1, 2, and 3, as indicated by Figure 1. These samples were run through SDS-PAGE showing that the majority of our protein was in Peak 2, which concurs with the chromatograph, as Peak 2 is the highest peak, suggesting that this is where the majority of StarD7 protein was eluted.

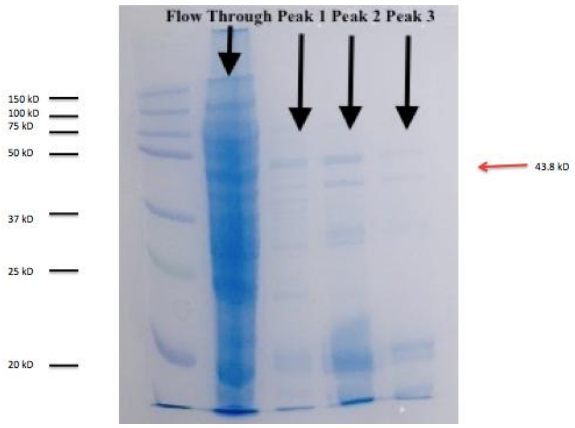

Figure 8. SDS-PAGE of nickel affinity chromatography results. Peak two (labeled in Figure 1) has the highest concentration of our protein.

The StarD7 protein solution was run over a Post-TEV Nickel Affinity column in order to remove all the TEV protease and cleaved SUMO tags. The flow through was collected, as cut StarD7 did not bind to the Nickel resin.

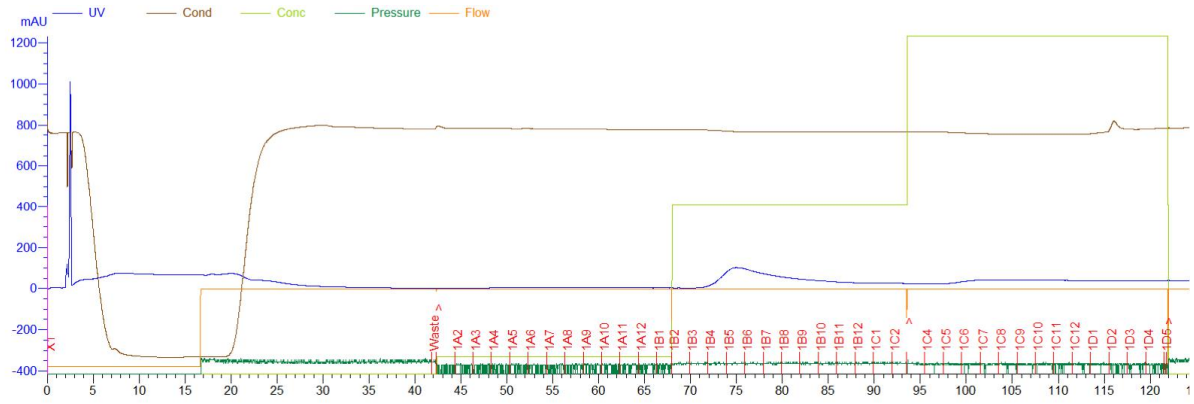

Figure 9. Post-TEV column (Method 1). The flow through is were StarD7 is located. The brown line corresponds to the measure of conduction. The orange line corresponds to the flow rate. The dark green line corresponds to pressure. The blue line corresponds to the UV 280. The green line corresponds to the concentration of imidazole within the elution buffer. The elution buffer was stepwise increased from 5% to 50% to 100% imidazole 250 mM.

Method 1 finished with running the StarD7 over an anion exchanger Q column. The Q column is positively charged, and our protein with a predicted pI of 6.55 will be negatively charged in a basic buffer (pH 8.4) , thus will bind to the Q column and elute in higher concentrations of salt. The flow through containing StarD7 is

collected at the end, and the peaks in Figure 4 indicate where protein could potentially be. This method did not yield enough viable protein, so we tried to optimize the protocol by changing the order of the purification methods.

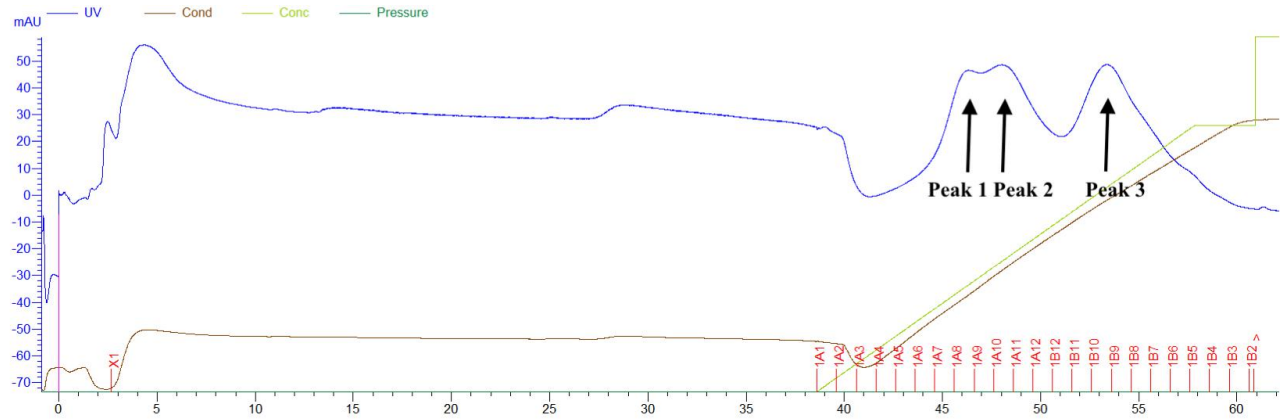

Figure 10. Anion exchanger Q column results (Method 1). The protein was ultimately eluted in Peak 3 (labeled here) and collected. The concentration was too low to be usable. The blue line corresponds to the UV 280. The green line corresponds to the increase in salt concentration (20 mM to 1 M NaCl on a gradient). The elution buffer was stepwise increased from 5% to 50% to 100% imidazole (250 mM).

The second method, or Method 2, also began with a nickel affinity chromatography column. The high imidazole content allowed for the elution of the protein off the histidine binding column. The eluted StarD7 was isolated in the peak indicated in Figure 5.

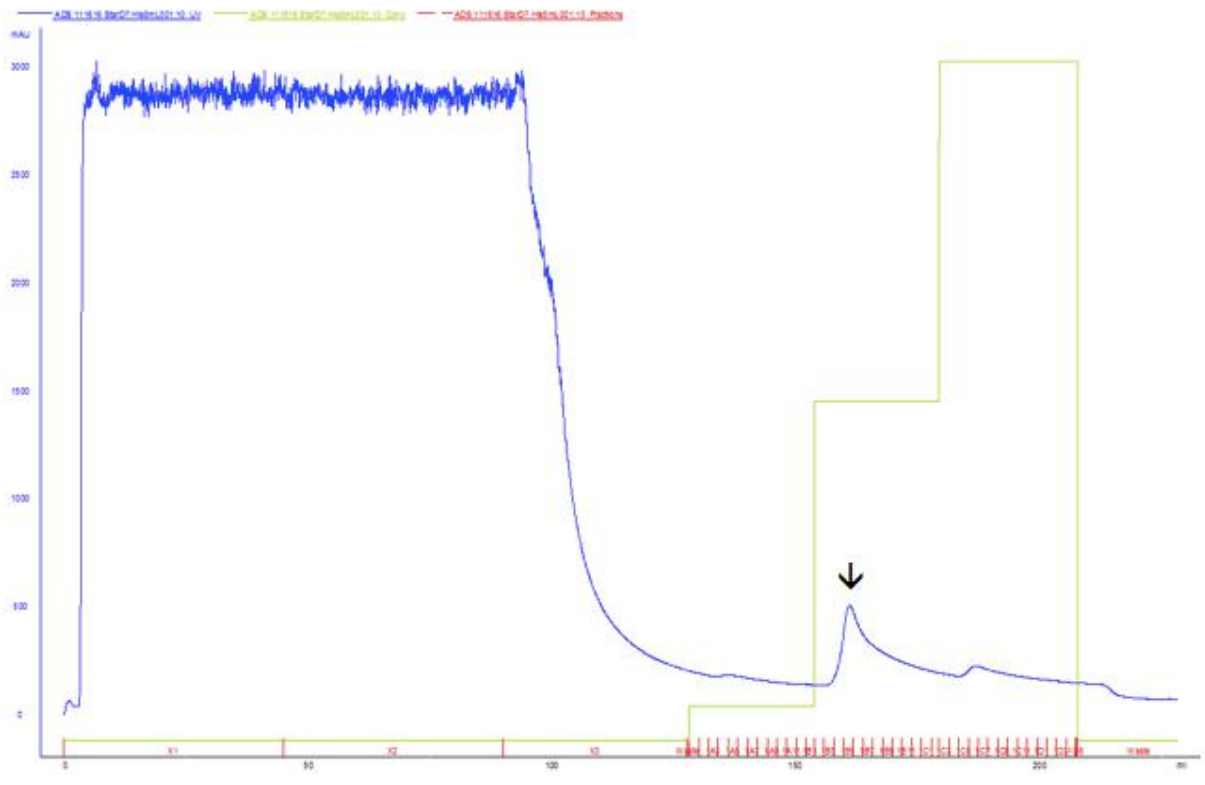

Figure 11. Nickel Affinity Chromatography (Method 2). The indicated peak is where StarD7 protein was eluted and ultimately run over a Q column. The blue line corresponds to the UV 280. The green line corresponds to the concentration of imidazole within the elution buffer. The elution buffer was stepwise increased from 5% to 50% to 100% imidazole (250 mM).

The StarD7 was then run over a Q column instead of TEV cleaving the protein first in order to purify the protein solution further. The flow through containing StarD7 is collected at the end, and the peaks in Figure 6 indicate where protein could potentially be.

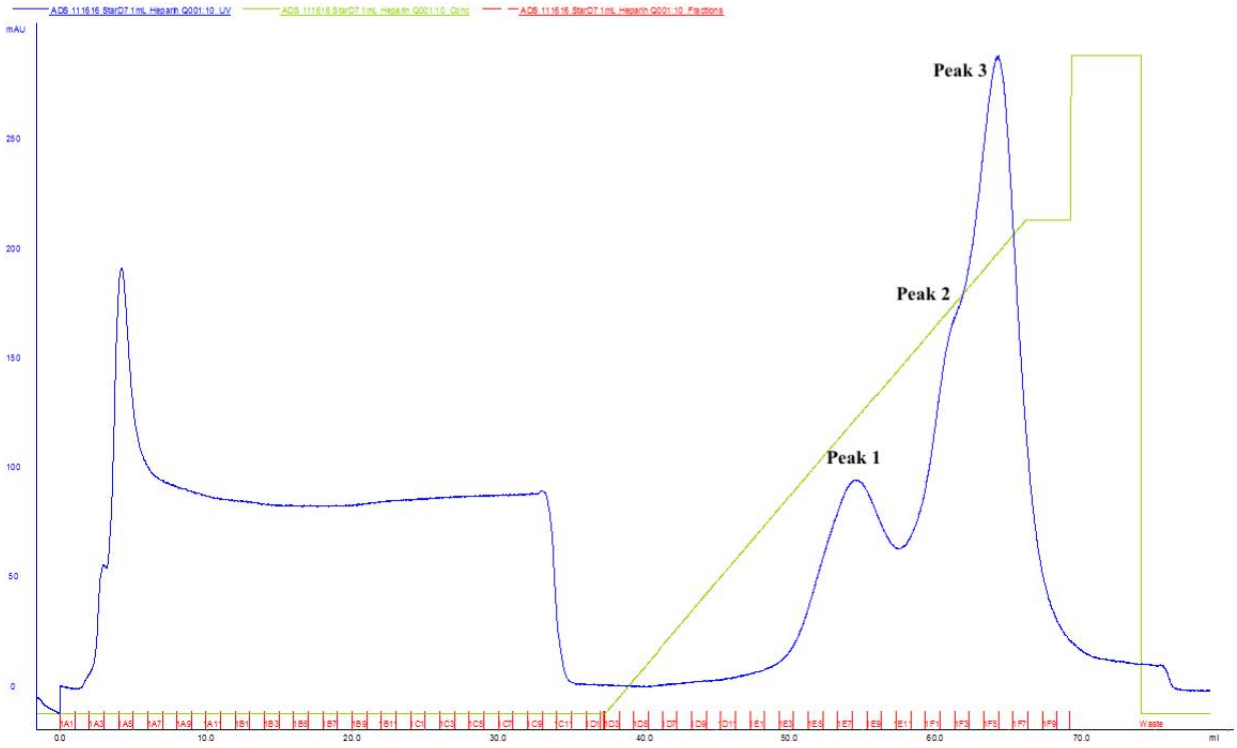

Figure 12. Anion exchanger Q column results (Method 2). There were multiple peaks, but gel electrophoresis (Figure 8) revealed that StarD7 was in Peak 2. No TEV-cleavage has occurred. The blue line corresponds to the UV 280. The green line corresponds to the concentration of imidazole within the elution buffer. The elution buffer was stepwise increased from 5% to 50% to 100% imidazole (250 mM).

Samples from Peaks 1, 2, and 3, as indicated by Figure 6, were then run through SDS-PAGE and found that the majority of our protein was in Peak 2, which was confusing as peak 3 contained the most protein based of the chromatogram. Therefore, peaks 1, 2, and 3 were all collected for use in the next step. There were a few bands that did not appear to be StarD7, indicating that the protein was not pure.

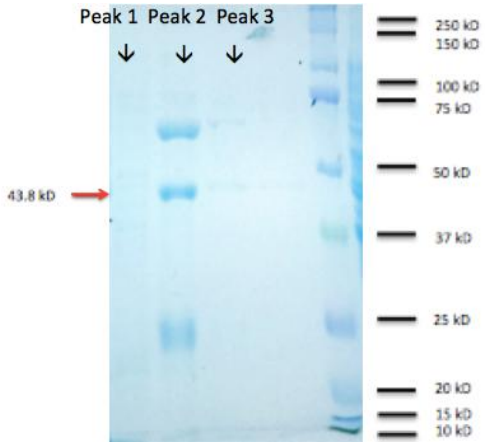

Figure 13. Gel electrophoresis of Q column results. Most of StarD7 is in Peak 2 at the uncut size 43.8 kD; however, there are still a few nonspecific bands, suggesting that the protein is not pure. The protein is TEV- cleaved and dialyzed after this step.

The StarD7 protein solution was then TEV cleaved and run over a Post-TEV Nickel Affinity column in order to remove all the TEV protease and cleaved SUMO tags. The StarD7 protein solution was run over a Post-TEV column, as indicated by Figure 8, in order to remove all the TEV protease and cleaved SUMO tags. The flow through was collected, as cut StarD7 did not bind to the Ni resin. Unfortunately, the protein concentration was too low and no protein was present after running on a SDS-PAGE gel.

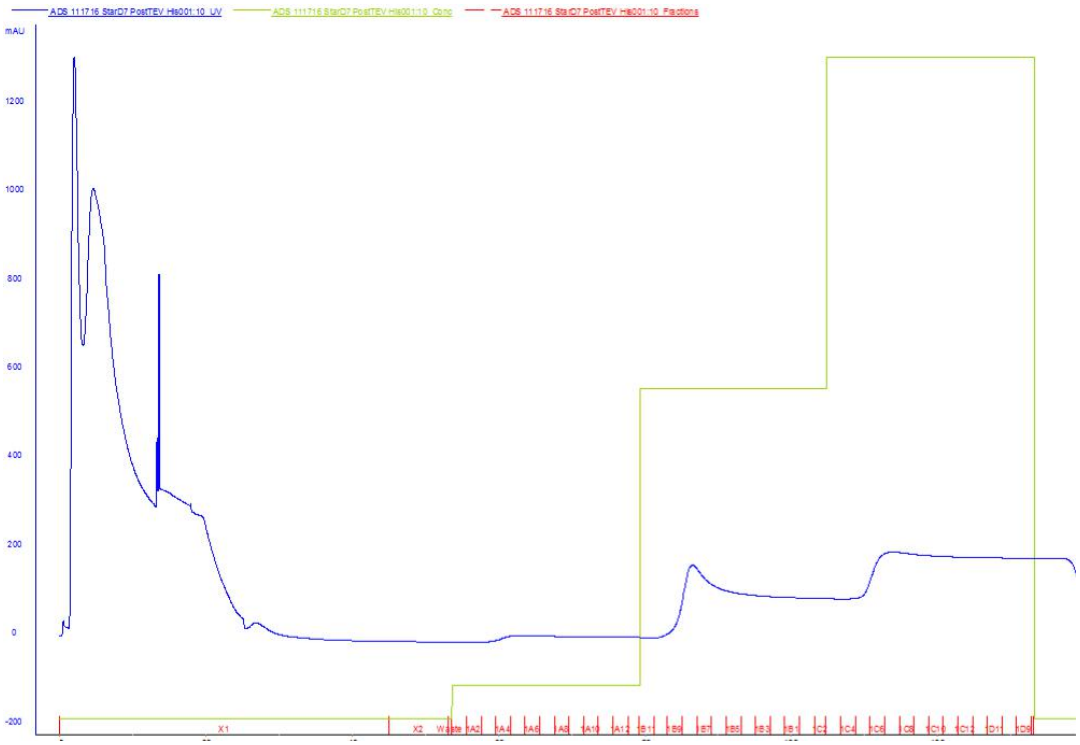

Figure 14. Post-TEV column (Method 2). The indicated peak is where the StarD7 protein was eluted. This protein was collected and frozen down to -80oC. The concentration was too low (1.93 mg/mL) to be crystallized. The blue line corresponds to the UV 280. The green line corresponds to the concentration of imidazole within the elution buffer. The elution buffer was stepwise increased from 5% to 50% to 100% imidazole (250 mM).

Test expressions were conducted in order to optimize growing conditions for StarD7 protein so that even if some protein was lost in the process, the concentration would still be high enough to be used for crystallography purposes. The initial test expressions revealed that StarD7 was most specifically induced during a day growth for 4 hours at 30°C opposed to a night growth, as indicated by Figure 9. The induction of the protein appeared unaffected by the addition of lecithin. This was to provide ligand for StarD7, but lecithin didn't appear to help induction. We determined that the difference in day growth conditions were not significant enough to be incorporated into our protocol, and ultimately we switched over to overnight growths in future experiments.

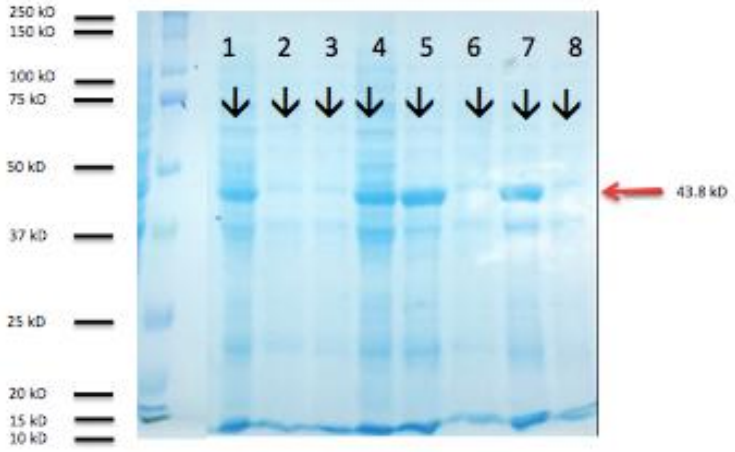

Figure 15. Test expressions for StarD7 induction. 1 – Post-induction night growth (+) lecithin. 2 – Pre- induction night (18°C) growth (+) lecithin. 3 – Pre-induction night growth (-) lecithin. 4 – Post-induction night growth (-) lecithin. 5 – Post-induction day growth (30°C) (+) lecithin. 6 – Pre-induction day growth (+) lecithin. 7 – Post-induction day growth (-) lecithin. 8 – Pre-induction day growth (-) lecithin. As indicated by the red line, the bands fall within the range of uncut StarD7 with SUMO tag. There were no significant differences in induction seen.

Chapter 3

Isolating StarD7 – pET15b vector

Next, we decided to clone the same long StarD7 sequence into a pET15b vector instead of a SUMO pET vector with a thrombin cleavage site instead of a TEV cleavage site to see if this would improve our protein yield and purity. PCTP, a similar protein, was purified with just a His tag, suggesting that we may have success with StarD7 purification using a His tag.

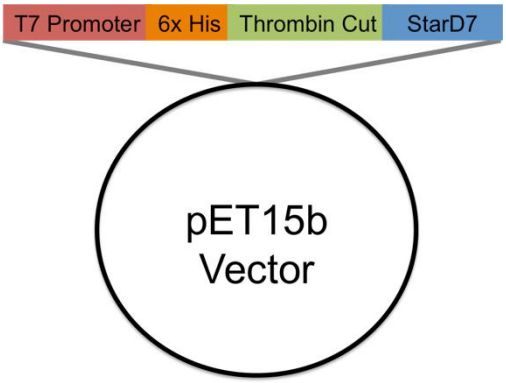

Figure 16. StarD7 pET15b Sequence:

MGSSHHHHHHSSGLVPRGSHIQEEELQRSINEMKRLEEMSNMFQSSGVQHHPPEPKAQTEGNE
DSEGKEQRWEMVMDKKHFKLWRRPITGTHLYQYRVFGTYTDVTPRQFFNVQLDTEYRKKWD
ALVIKLEVIERDVVSGSEVLHWVTHFPYPMYSRDYVYVRRYSVDQENNMMVLVSRAVEHPSV
PESPEFVRVRSYESQMVIRPHKSFDENGFDYLLTYSDNPQTVFPRYCVSWMVSSGMPDFLEKL
HMATLKAKNMEIKVKDYISAKPLEMSSEAKATSQSSERKNEGSCGPARIEYA

MW: 35.2 kD
pI: 6.47
Cut StarD7:
MW: 33.4
pI: 6.10
His Tag
Thrombin cut site

We used BL21(DE3)/pG-Tf2 cells that co-express chaperone proteins to grow the *E. coli*. The chaperone proteins were first induced with 0.5 mM tetracycline HCl. StarD7 was induced with 0.5 mM IPTG once cultures reached OD600 0.6. Cultures grew overnight at 16°C.

Cells were resuspended in buffer for 10 minutes at 4^{o}C and sonicated. Cell lysates were then spun at 16,000 x G for 30 minutes, and the supernatant was collected.

StarD7 was purified from soluble lysates using nickel affinity chromatography.

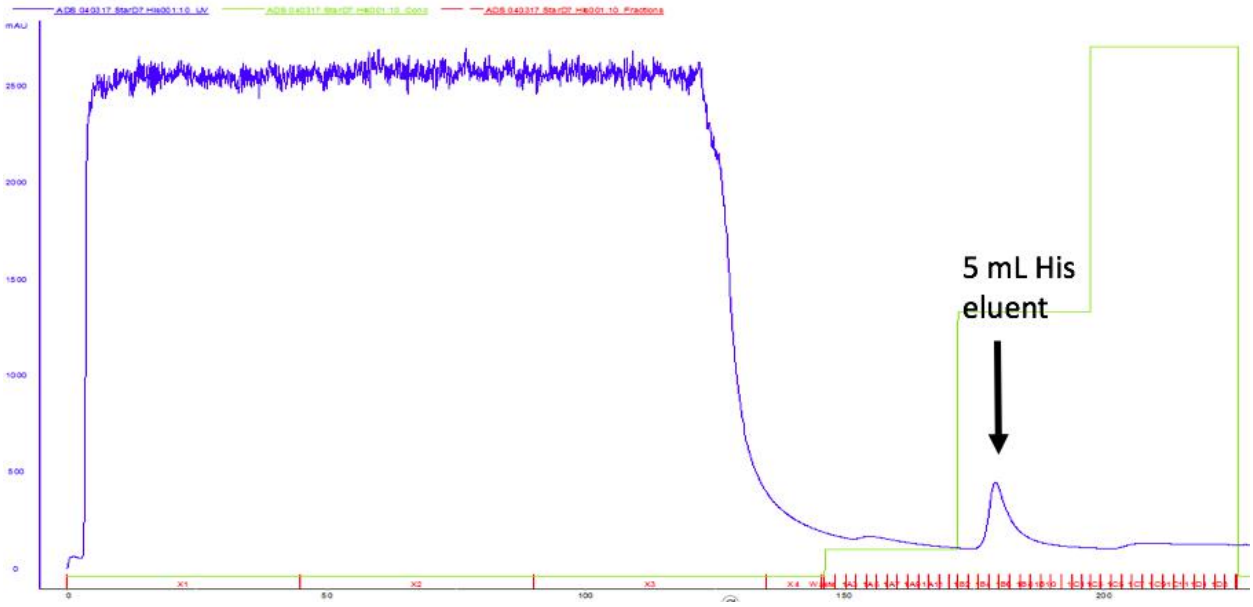

Figure 17: First-pass purification of His-tagged StarD7. We used Nickel Affinity Chromatography to purify the crude protein. This run was using protein isolated from BL239DE3)/pG-Tf2 cells grown in LB media. Peak two is where StarD7 protein resided, and was ultimately collected and thrombin-cleaved. The blue line corresponds to the UV 280. The green line corresponds to the concentration of imidazole within the elution buffer. The elution buffer was stepwise increased from 5% to 50% to 100% imidazole (250 mM).

The eluted StarD7 was collected and then 25 units of thrombin was added to the protein solution. The solution was placed into a membrane with 10,000 Dalton pores and put into dialysis buffer containing 20 mM Tris HCL pH 7.4, 150 mM NaCl, 5% glycerol overnight at room temperature.

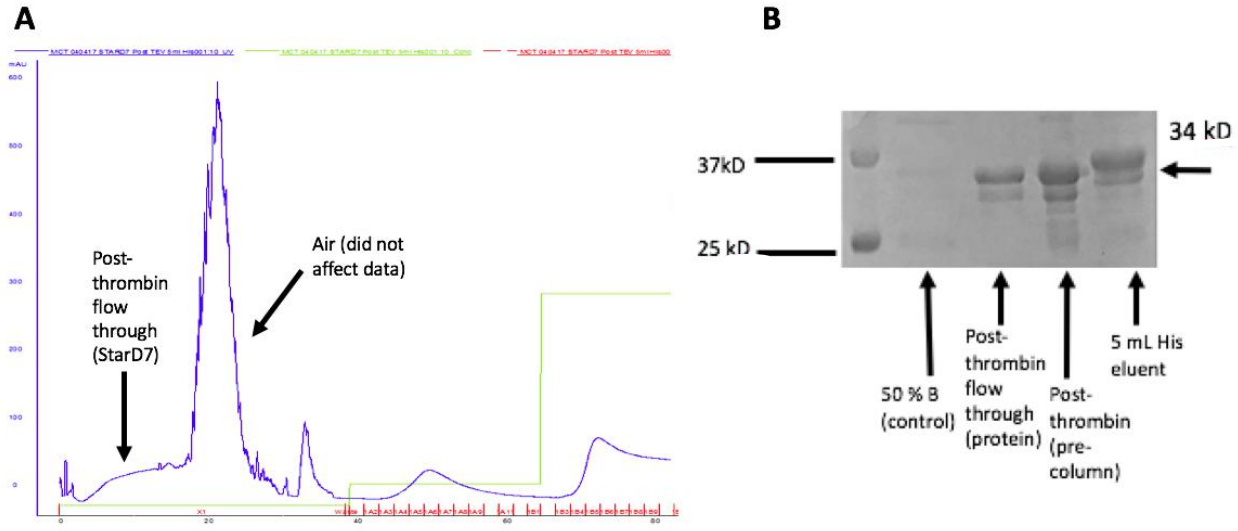

Figure 18. Purification of STARD7 following tag cleavage. *A.* We used nickel affinity chromatography to purify the protein whose His tag was cut using thrombin. The indicated peak is ultimately where the StarD7 protein resided. *B.* A gel was run with StarD7 samples and 50% B peak to ensure complete thrombin cleavage. Two bands at 34 kD are at the expected molecular weight for StarD7.

Thermofluor was used to determine optimal buffer conditions for the purified protein.

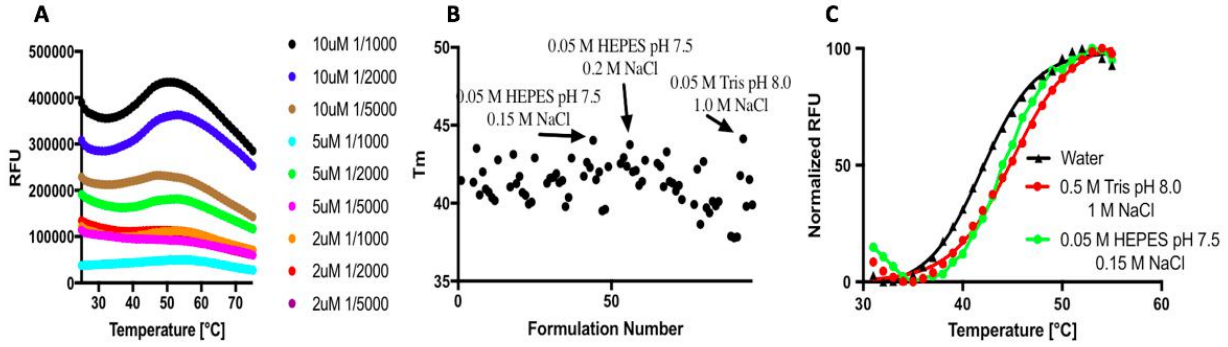

Figure 19. Thermofluor analysis of StarD7 for stabilizing conditions. Thermofluor was used to test for optimal conditions in which StarD7 was stabilized. *A*. An initial screen of differing protein and SYPRO Orange concentrations suggests that 10 µM protein and 1/2000 dye (blue line) yields the strongest signal. *B and C.* We tested the stabilizing effects of 96 different buffer conditions by examining the Tm of StarD7 after incubation with buffers. We found that 0.05M HEPES pH 7.5 and 150 mM NaCl increased StarD7's T_m. We used these buffer conditions moving forward.

The protein was then purified using size exclusion chromatography, using the 0.05 HEPES pH 7.5, 150 mM NaCl buffer that most stabilized the StarD7 protein.

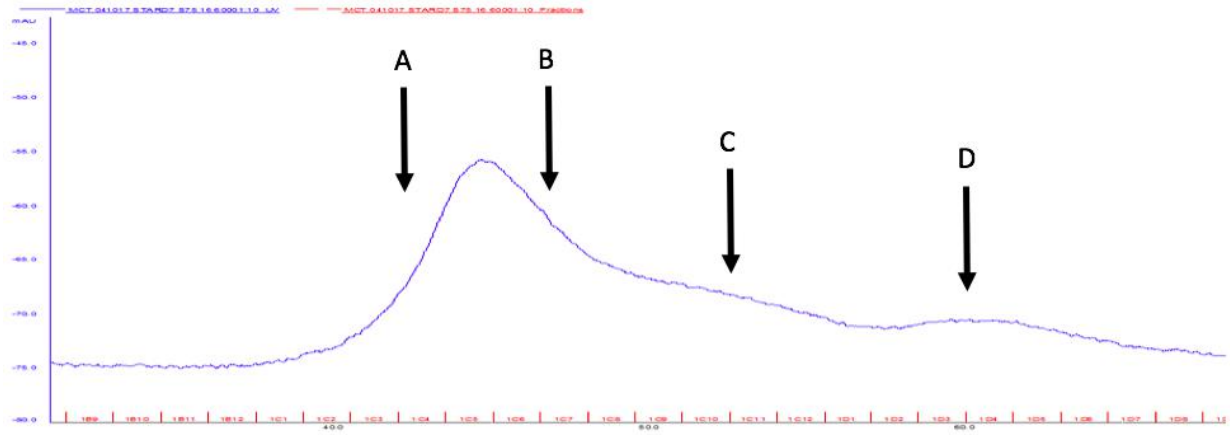

Figure 20: Size-Exclusion Chromatography (SEC) of StarD7. The protein was run over a size-exclusion chromatography column using a buffer containing 0.05M HEPES pH 7.5, 150mM NaCl, and 5% glycerol in order to further purify the protein and remove any nonspecific proteins. Samples were collected from Peaks A, B, C, and D.

We used SDS-PAGE to confirm the presence of pure StarD7 protein.

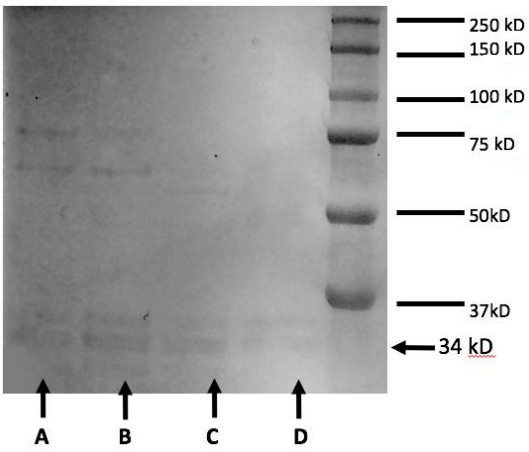

Figure 21. SDS-PAGE of StarD7 from SEC. A gel was run with samples A, B, C, and D from SEC. These samples were run on SDS-PAGE and we found that the majority of our purest protein was in Peak D (Figure 5). Although we had low amounts of protein present in this peak, the sample in peak D did not have any other nonspecific proteins.

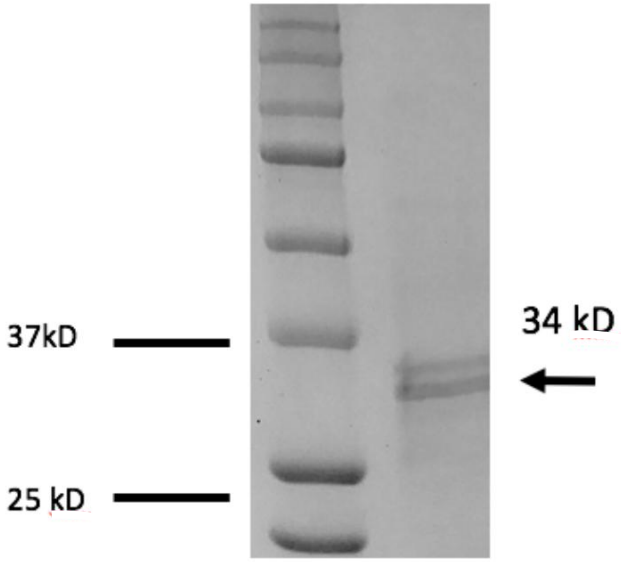

Figure 22: Final StarD7 purified protein concentrated. The protein was collected from peak D in Figure 5 and concentrated the protein down, and ultimately we collected 56 µM of protein and analyzed it using SDS-PAGE. The two pure bands are likely both StarD7.

We attempted to set crystal trays with a 24-well crystal screen , but did not have enough isolated protein to form viable crystals. Any crystal screens that we set using the hang drop crystal method ended up crashing out and we did not have enough protein left to set up additional screens to determine successful stabilizing conditions.

Chapter 4

Isolating StarD7 – Truncated Transcript

The StarD7 transcript was truncated, using primers and restriction enzymes in a PCR, from residues 112 – 370 to residues 137 – 321, and this adjustment significantly improved the solubility and crystallization potential of StarD7. We removed the N-terminus disordered regions, as this adjustment would significantly help the StarD7 protein fold into its proper structure (Fig. 17). This difference in crystallization potential can be seen in Figure 18, where the crystallization potential group drastically improves once the disordered regions of the StarD7 long transcript are removed.

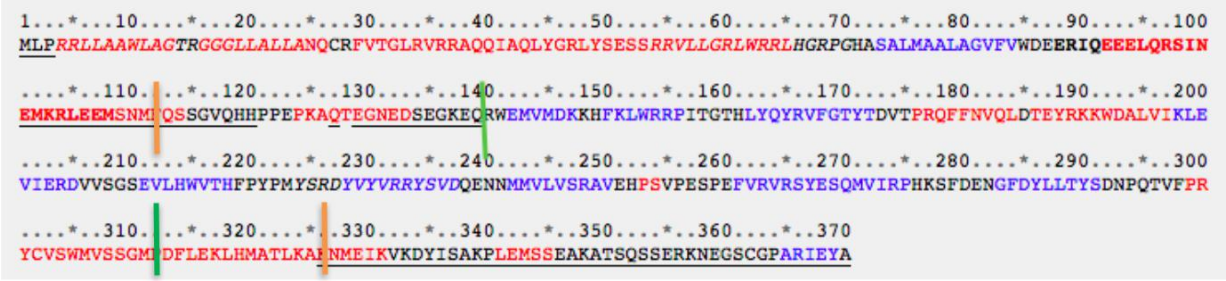

Figure 23. The orange indicates the N-terminus and C-terminus of the long transcript, while the orange indicates the N-terminus and C-terminus of the short transcript. It is important to note that the underlined regions are disordered and would impair the crystallization of the StarD7 protein, so the truncation removes all the predicted disordered regions. Red regions are predicted alpha helices, while blue regions are predicted beta sheets [13].

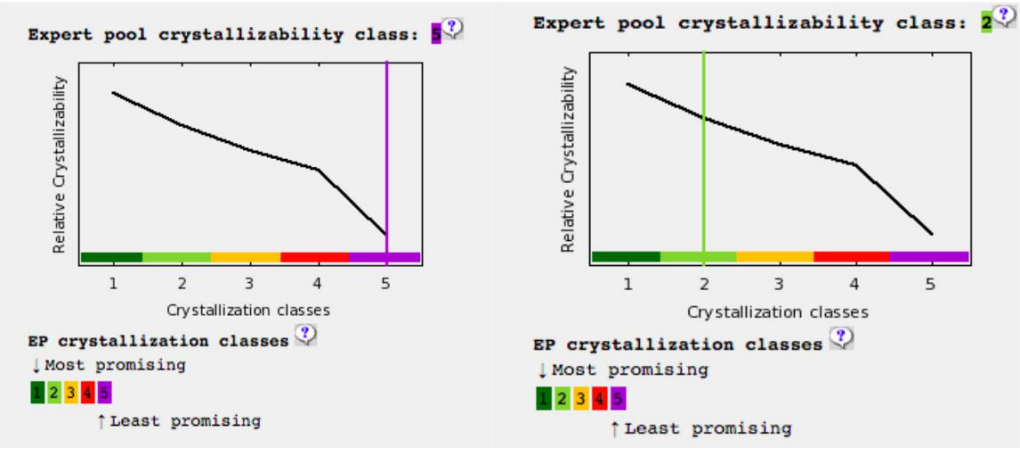

Figure 24. The graph on the right shows the crystallization potential of the truncated StarD7 transcript, versus the crystallization potential of the long version of our transcript on the left, after removing the disordered regions of the StarD7 gene [13].

Chapter 5

Crystallizing Truncated StarD7 – Two Tags

The StarD7 truncation (residues 141 – 326) was cloned into a pMCSG7 and pMCSG9 vector containing either a Lic His tag or a Maltose Bindin Protein (MPB)-His tag, followed by a TEV cleavage site. It was important to

determine which tag further improved the solubility of the StarD7 protein in the *E.coli* system.

MHHHHHHSSGVDLGTENLYFQSNAWEMVMDKKHFKLWRRPITGTHLYQYRVFGTYTDVTPR
QFFNVQLDTEYRKKWDALVIKLEVIERDVVSGSEVLHWVTHFPYPMYSRDYVYVRRYSVDQE
NNMMVLVSRAVEHPSVPESPEFVRVRSYESQMVIRPHKSFDENGFDYLLTYSDNPQTVFPRYC
VSWMVSSGMPDFLEKLHMATLKA

Lic His Tag
TEV Cleavage site
MW: 25 kD
pI: 6.67
Cut StarD7:
MW: 22 kD
pI: 7.23

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFW
AHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPP
KTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGA
KAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTF
KGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAK
DPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNN
NNNNNLGIEGRENLYFQSNAWEMVMDKKHFKLWRRPITGTHLYQYRVFGTYTDVTPRQFFN
VQLDTEYRKKWDALVIKLEVIERDVVSGSEVLHWVTHFPYPMYSRDYVYVRRYSVDQENNM
MVLVSRAVEHPSVPESPEFVRVRSYESQMVIRPHKSFDENGFDYLLTYSDNPQTVFPRYCVSW
MVSSGMPDFLEKLHMATLKA

MBP-His Tag
TEV Cleavage site
MW: 25 kD
pI: 6.67
Cut StarD7:
MW: 22 kD
pI: 7.23

StarD7 was expressed in BL21(DE3) *E. coli* cells and grown in 6 liters of terrific broth media to an OD600 0.6. Protein expression was induced upon the addition of 0.5 mM IPTG and grown at 18°C for 18 - 20 hours. Cells were harvested by spinning at 3,200 x g for 15 minutes. Cells were lysed in lysis buffer containing 20 mM Tris HCL pH 7.4, 150 mM NaCl, 5% Glycerol, 25 mM imidazole, lysozyme, DNase, and phenylmethylsulfonyl fluoride (PMSF). The mixture was agitated for 10 minutes

at 4°C and was sonicated. Cell lysates were then spun at 16,000 x rpm for 30 minutes, and the supernatant was collected.

StarD7 (grown at 18°C overnight) was purified from soluble lysates using Nickel affinity chromatography. The eluted StarD7 was collected and then 200 ug of TEV was added to the protein solution. The solution was placed into a membrane with 10,000 Dalton pores and put into dialysis buffer containing 20 mM Tris HCL pH 7.4, 150 mM NaCl, 5% Glycerol overnight in order to cleave the TEV tag. The protein solution was then run over a post-TEV Nickel affinity column, and the flow through, which contained StarD7, was collected. Finally, the protein was purified using an anion exchanger Q column, and the StarD7 elution was once more collected and the purified protein was stored at -80°C.

StarD7-MBP-His:

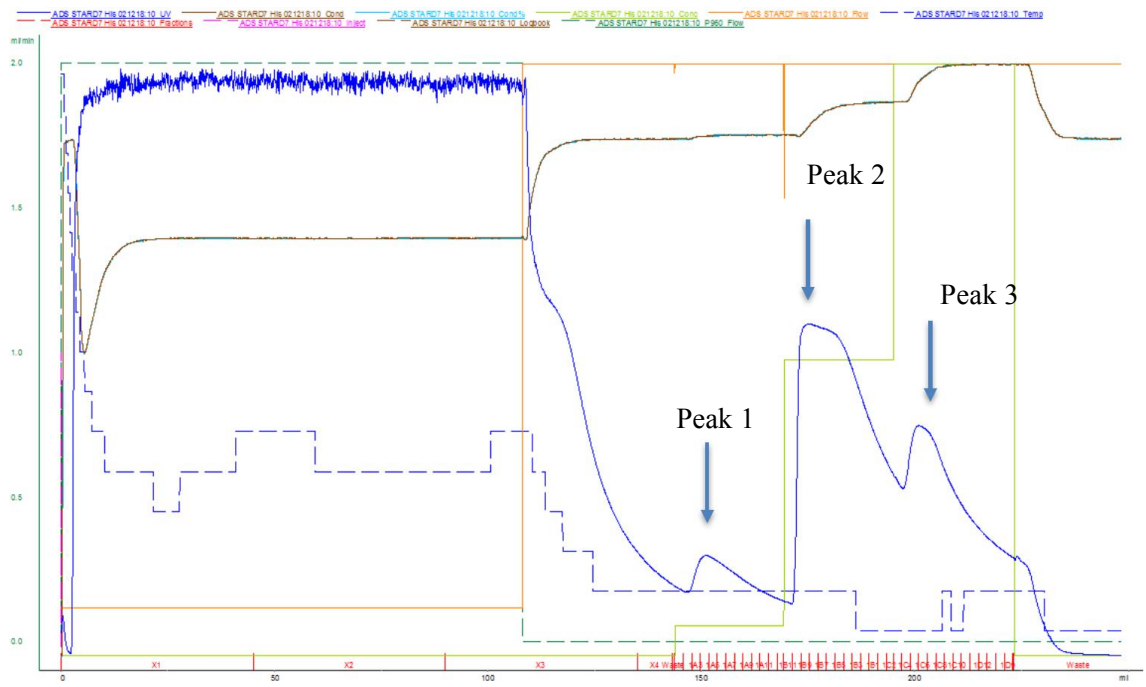

Figure 25: First-pass purification of MBP-tagged StarD7. We used Nickel Affinity Chromatography to purify the crude protein. This run was using protein isolated from BL239DE3)/pG-Tf2 cells grown in LB media. Peak two is where StarD7 protein resided, and was ultimately collected and TEV-cleaved. The blue line corresponds to the UV 280. The green line corresponds to the concentration of imidazole

within the elution buffer. The elution buffer was stepwise increased from 5% to 50% to 100% imidazole (250 mM).

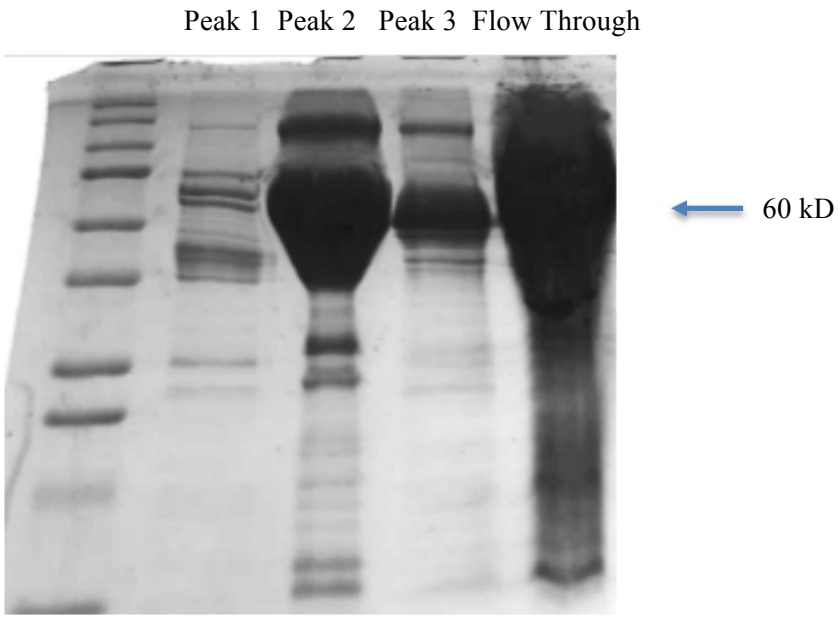

Figure 26. SDS-PAGE confirmed the presence of StarD7, particularly concentrated in the highest peak (band 2). It is important to note that because StarD7 is MBP-tagged, it resides at 60 kD.

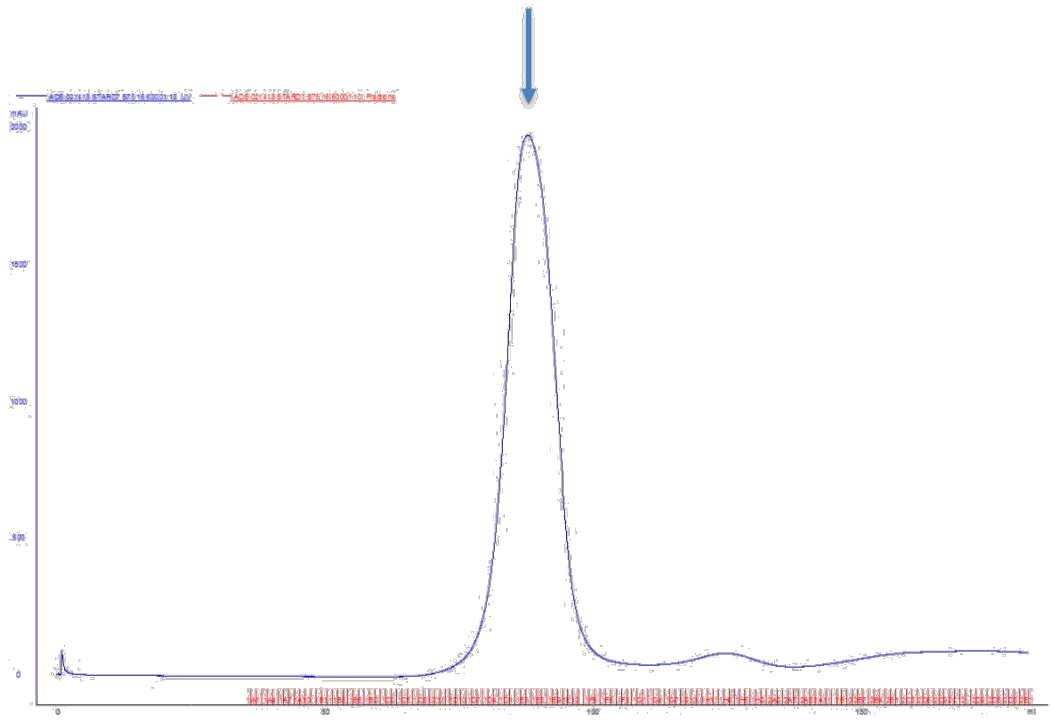

Figure 27: Size-Exclusion Chromatography (SEC). The protein was run over a size-exclusion chromatography column in order to remove the tag after a TEV cleavage.

Size-exclusion chromatography did not remove the HSP-60's, as they still strongly associated with the protein of interest.

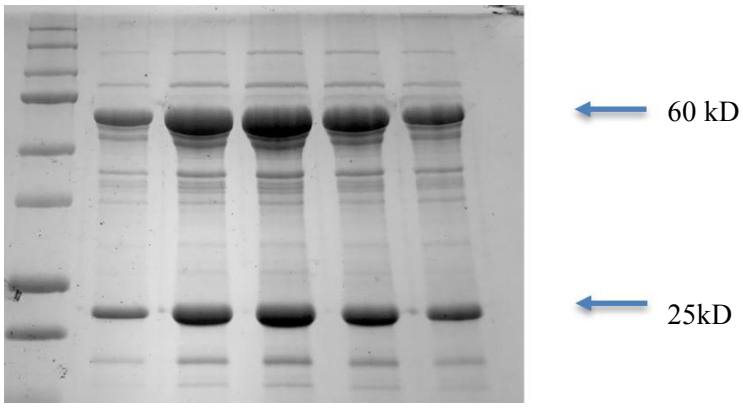

Figure 28. SDS-PAGE was conducted using multiple sample of the main peak. The gel confirmed the presence of StarD7 post-SEC, as cut StarD7 should reside at 25 kD. It is important to note the presence of heat shock proteins (HSP-60s) at 60 kD throughout all samples, indicating the size exclusion was unsuccessful, and the HSP-60s could not be activated by ATP and $MgCl_2$ to the point of releasing StarD7.

At 60 kD, we noticed a specific protein that seemed to tether to StarD7 and did not dissociate during purification efforts. We determined through SDS-PAGE that this could not be uncut protein, and we ultimately identified the ptotein as heat shock protein 60s (HSP-60s). The presence of HSP-60s is likely the result of the truncated StarD7 transcript's exposed hydrophobic face, which associated and was stabilized by the heat shock proteins. In an effort to disassociate the isolated StarD7 proteins and HSP-60, we added ATP and magnesium chloride, $MgCl_2$. Many chaperone proteins use ATP as an energy source while aiding in protein folding and magnesium stabilizes triphosphate groups.

StarD7-Lic-His:

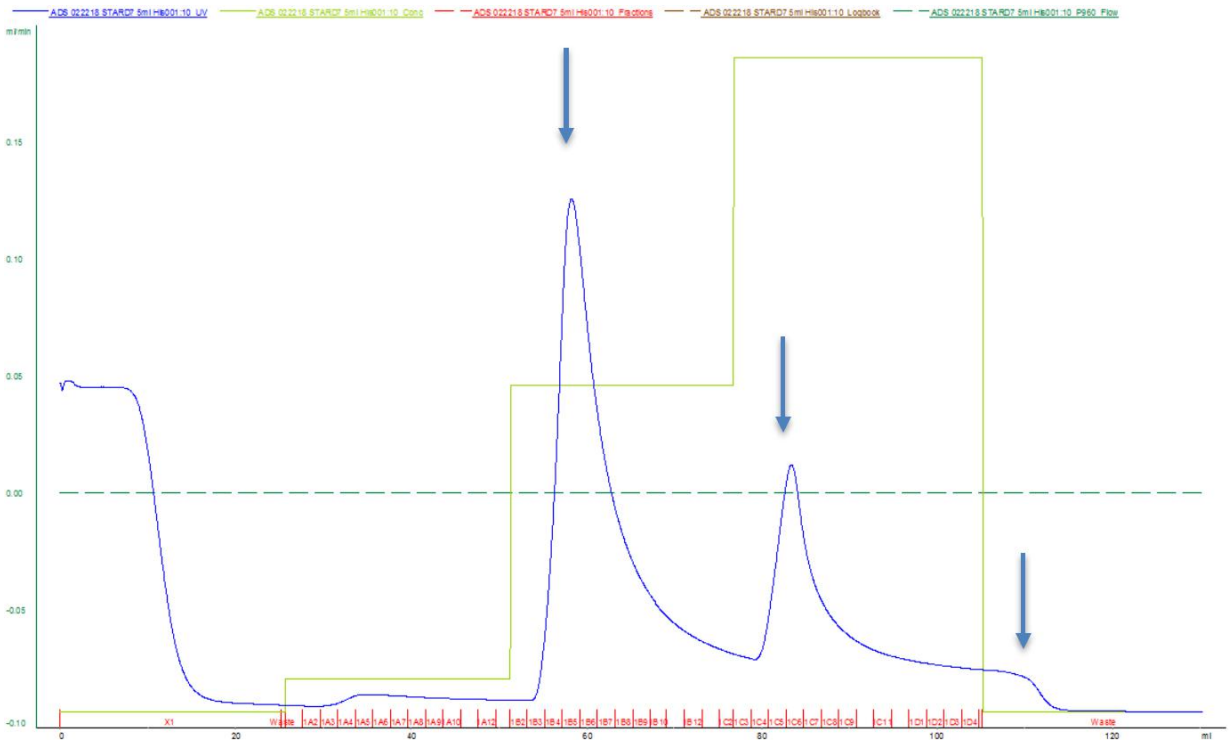

Figure 29. First-pass purification of Lic-tagged StarD7. We used Nickel Affinity Chromatography to purify the crude protein. This run was using protein isolated from BL239DE3)/pG-Tf2 cells grown in LB media. We ran all three peaks on a gel.

Unfortunately, there was no StarD7 protein isolated using the Lic-His tag in any of the peaks tested from the nickel affinity chromatography experiment. SDS-PAGE confirmed that there was no detectable protein in any of the identified peaks.

Chapter 6

StarD7 in a Mammalian Cell Line System

With the bacterial systems proving ineffective, the next experiment involved expressing protein in a mammalian system so that more relevant molecules will bind and be later analyzed through mass spectrometry, and protein will possibly be purified from this system.

To do this we cloned in the following vector, which developed a stable mammalian cell line:

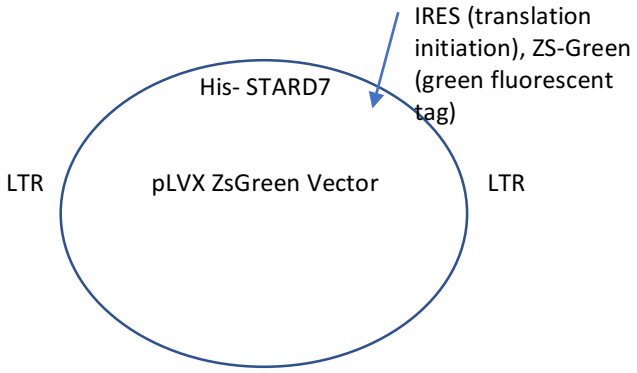

Figure 30. PLVX vector containing truncated, His-tagged StarD7, LTR regions, an IRES site, and a ZS-Green tag for fluorescent detection.

First, the viral plasmid and StarD7 + Vector were transfected into adherent HEK293T cells. Media was collected and virus was harvested using a PEG solution. Then, HEK 293F cells were transduced with several dilutions of virus. The virus titer and multiplicity of infection (MOI) was determined through re-suspending transduced cells in PBS and measuring GFP fluorescence using a cell counter. Two lines stably expressing StarD7 with an MOI 9 and MOI 0.9 were used. We tested both cell lines for expression of StarD7; however, both yield no pure protein.

Chapter 7

Fixing the Truncated StarD7 with a Chimera

Looking back on experiments, the StarD7 transcript and system with the most success was the StarD7 truncation with a MBP-His tag in the *E.coli* bacterial system. Despite the high yield of protein, we were unable to purify it completely, as HSP60 still associated with STARD7. We then examined a SWISS model of our protein, which made predictions about the secondary structure of StarD7. One of the issues posed by this truncation was the exposure of a hydrophobic **β** sheet face, which left the StarD7 protein vulnerable to association with other nonspecific proteins.

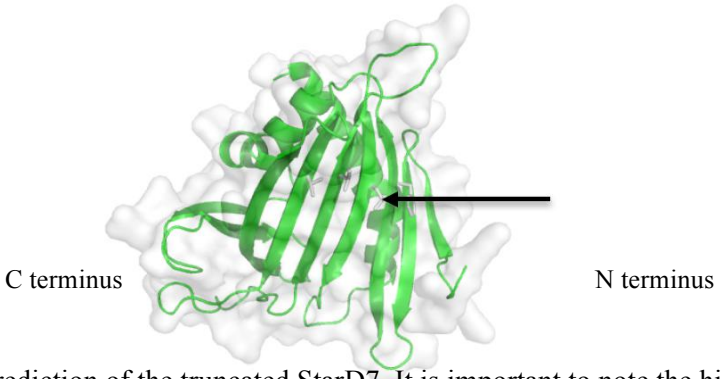

Figure 31. A structural prediction of the truncated StarD7. It is important to note the highly nonpolar **β** sheet that is exposed, and could have potentially been associating with the HSP-60s during our purification experiments.

Figure 32. Swiss model makes a structural model of your protein based of the sequence and known crystal structures of similar proteins. Swiss model made a model based off the PCTP structure, which is in the same group of START domain proteins. The red regions are low quality predictions, while the blue regions are high quality predictions.

Looking back at Figure 17, we see that by truncating the StarD7 transcript, we have removed a predicted alpha helix, which could have potentially protected the exposed nonpolar beta sheet; however, it is likely that the disorder made crystallization difficult with our long transcript.

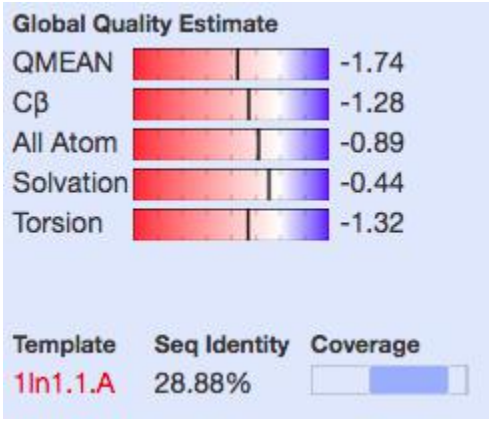

Figure 33. SWISS, a homology modeling server model predictions show a 28.8% similarity between PCTP and StarD7, and much of the conserved region lies in the binding pocket. This is likely due to the similar functionalities of PCTP and StarD7: to bind and shuttle PC lipids.

Phosphatidylcholine Transfer Protein, or PCTP, is homologous to StarD7 and serves a similar purpose as that of StarD7: to shuttle PCs to mitochondria. By cloning a PCTP transcript into the StarD7 transcript and vector, a chimera with a helix covering the hydrophobic sheet face could be created and aid in folding (Fig. 28). However, multiple attempts at cloning and PCR protocols failed.

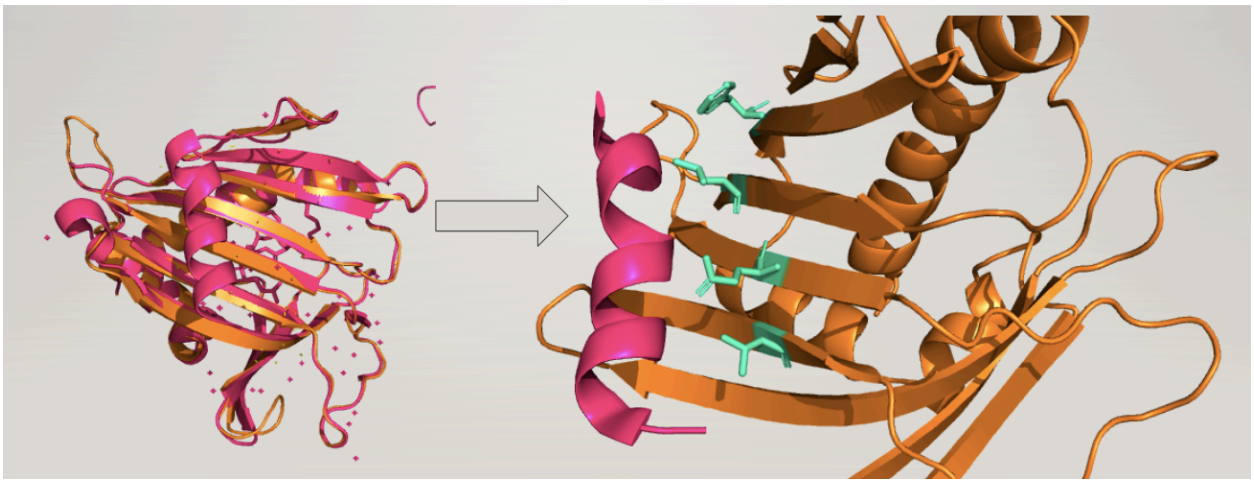

Figure 34. This structural prediction shows StarD7 and PCTP overlaid on the left. Taking a closer look on the right, the pink PCTP helix could protect the hydrophobic face on the StarD7 truncation, improving crystallization potential.

Chapter 8

Future Directions and Conclusions

The current issue is that StarD7 protein is lost every time the protein is run through a column, and the protein is never pure enough to isolate and crystallize. The goal is to find the optimal conditions to isolate an abundance of StarD7 protein at a high purity. Once the endogenous lipid that binds to StarD7 is identified, the lipid will be used to crystallize the isolated StarD7. Crystallization conditions will be screened to identify optimal conditions to grow StarD7 crystals that will ultimately diffract X-rays and yield a crystal structure.

Determining a successful PCR protocol for the cloning of PCTP into the StarD7 MBP-His tagged transcript and may prove to be successful in the *E.coli* system, as this would likely solve the problem of associated HSP-60s. Alternatively, if the cloning protocol is unsuccessful, we could synthesize the chimeric gene from scratch and use the same bacterial cell protocol to isolate the StarD7 protein. To do this, we would insert the sequence of the α helix of the PCTP into the vector with the StarD7 short transcript in an effort to shield the hydrophobic β sheet. Additionally, using alternate orthologs, such as the mouse StarD7 transcript, could improve the solubility of the StarD7 transcript and the resulting in better purification of the StarD7 protein. The zebrafish ortholog could also have an N-terminus helix that may aid in the protection of the exposed hydrophobic face on the StarD7 protein.

BIBLIOGRAPHY

[1] Vance, J. E. (2015). Phospholipid synthesis and transport in mammalian cells. *Traffic, 16*(1), 1-18. doi:10.1111/tra.12230

[2] Flis, V. V., & Daum, G. (2013). Lipid Transport between the Endoplasmic Reticulum and Mitochondria. *Cold Spring Harbor Perspectives in Biology, 5*(6). doi:ARTN a013235 10.1101/cshperspect.a013235

[3] Vance, J. E. (2014). MAM (mitochondria-associated membranes) in mammalian cells: Lipids and beyond. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids, 1841*(4), 595-609. doi:10.1016/j.bbalip.2013.11.014

[4] Yang, L., Na, C. L., Luo, S. Y., Wu, D., Hogan, S. M., Huang, T. S., & Weaver, T. E. (2017). The Phosphatidylcholine Transfer Protein Stard7 is Required for Mitochondrial and Epithelial Cell Homeostasis. *Scientific Reports, 7*. doi:ARTN ł6416 10.1038/srep46416

[5] Tsujishita, Y., & Hurley, J. H. (2000). Structure and lipid transport mechanism of a StAR-related domain. *Nature Structural Biology, 7*(5), 408-414.

[6] Alpy, F., & Tomasetto, C. (2005). Give lipids a START: The StAR-related lipid transfer (START) domain in mammals. *Journal of Cell Science, 118*(13), 2791-2801. doi:10.1242/jcs.02485

[7] Hu, J., Zhang, Z., Shen, W. J., & Azhar, S. (2010). Cellular cholesterol delivery, intracellular processing and utilization for biosynthesis of steroid hormones. *Nutr Metab (Lond), 7*, 47. doi:10.1186/1743-7075-7-47

[8] Horibata, Y., & Sugimoto, H. (2010). StarD7 mediates the intracellular trafficking of phosphatidylcholine to mitochondria. *J Biol Chem, 285*(10), 7358-7365. doi:10.1074/jbc.M109.056960

[9] Horibata, Y., Ando, H., Satou, M., Shimizu, H., Mitsuhashi, S., Shimizu, Y., . . . Sugimoto, H. (2017). Identification of the N-terminal transmembrane domain of StarD7 and its importance for mitochondrial outer membrane localization and phosphatidylcholine transfer. *Scientific Reports, 7*. doi:ARTN 8793 10.1038/s41598-017-09205-1

[10] Saita, S., Tatsuta, T., Lampe, P. A., Konig, T., Ohba, Y., & Langer, T. (2018). PARL partitions the lipid transfer protein STARD7 between the cytosol and mitochondria. *Embo Journal, 37*(4). doi:UNSP e9790910.15252/embj.201797909

[11] Vance, J. E. (2015). Phospholipid Synthesis and Transport in Mammalian Cells. *Traffic, 16*(1), 1-18. doi:10.1111/tra.12230

[12] Horibata, Y., Ando, H., Zhang, P., Vergnes, L., Aoyama, C., Itoh, M., . . . Sugimoto, H. (2016). StarD7 Protein Deficiency Adversely Affects the Phosphatidylcholine Composition, Respiratory Activity, and Cristae Structure of Mitochondria. *J Biol Chem, 291*(48), 24880-24891. doi:10.1074/jbc.M116.736793

[13] Slabinski, L., Jaroszewski, L., & Godzik, A. (n.d.). Retrieved April 09, 2018, from http://ffas.burnham.org/XtalPred-cgi/xtal.pl