2013

Probing the Structure of Caveolin-1

Jin Woo Lee
Lehigh University

Probing the Structure of Caveolin-1

by

Jin Woo Lee

A Dissertation

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Chemistry

Lehigh University

January 2013

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Jin Woo Lee
Probing the Structure of Caveolin-1

_____
Defense Date

_____
Dissertation Director
Kerney Jebrell Glover

_____
Approved Date

Committee Members:

_____
Tianbo Liu

_____
Kai Landskron

_____
Jennifer Swann

ACKNOWLEDGMENTS

Many people contributed to this dissertation in numerous ways, and I am grateful to all of them. First and foremost, I would like to thanks my advisor, Professor Kerney Jebrell Glover for his helpful advice and encouragement. I appreciate all his contributions and for supporting me on this cutting edge research project, being there to discuss projects together to help navigate through problems, and for teaching me how to think as a true scientist. Besides my advisor, I would like to thank the member of my thesis committee: Professor Kai Landskron, Professor Tianbo Liu, and Professor Jennifer Swann for their helpful scientific advice and many insightful suggestions. I would also like to thank Professor Fang Tian at Pennsylvania State University College of Medicine at Hershey for useful advice and discussion about NMR experiments. I would like to give a special thanks to my lab-mates: Monica, Kyle, and Sarah. I do not think I can say thank you enough for their thoughtful discussions, and advice. The work presented in this thesis would not have been accomplished without the help from members of the Glover laboratory. I would also like to thank the undergraduates who made many contributions to my research. Lastly, I would like to give thanks to my family for their love and infinite support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDIX

LIST OF ABBREVIATIONS

BME: 2-Mercaptoethanol

CD: Circular Dichroism Spectroscopy

CSI: Chemical Shift Indexing

DPC: dodecylphosphocholine

EDTA: Ethylenediaminetetraacetic acid

HCl: Hydrochloric Acid

HFIP: Hexafluoroisopropanol

HPLC: High Performance Liquid Chromatography

HSQC: Heteronuclear Single Quantum Coherence

IPTG: Isopropyl β-D-1-thiogalactopyranoside

LMPG: lyso-myristoylphosphatidylglycerol

MALDI-TOF: Matrix-Assisted Laser Desorption/ionization time-of-flight

MW: Molecular Weight

NMR: Nuclear Magnetic Resonance Spectroscopy

PFOA: Perfluorooctanoic Acid

SDS: Sodium Dodecyl Sulfate

TEV: Tobacco Etch Virus Protease

TFA: Trifluoroacetic Acid

TFE: Trifluoroethanol

TROSY: Transverse Relaxation Optimized Spectroscopy

WT: wild-type

Abstract

The topic of my doctoral dissertation is probing the structural features of caveolin-1. Caveolin-1 is responsible for the formation of caveolae which are 50-100 nm diameter invaginations in cell membrane. The unique feature of caveolin-1 is the transmembrane domain which has a horseshoe conformation where both the N- and C- termini are facing the cytoplasm and the polypeptide chain turns inside the lipid bilayer. In Chapter 1, expression and purification methods are discussed in order to obtain high yield of hydrophobic proteins. Trp Leader fusion system was employed to send toxic membrane proteins into the inclusion bodies to protect the cells. Also, optimization of purification methods are discussed to obtain high purity membrane proteins, especially caveolin-1 constructs. In Chapter 2, the horseshoe topology of caveolin-1 transmembrane domain (96-136) was investigated. Caveolin-1(96-136) displays a helix-break-helix motif and G108, I109, and P110 were responsible for the break region which is a putative turn. Various point mutations were performed to give insight into these three residues. In Chapter 3, the caveolin-1 P132L mutation was probed. The P132L mutation is a naturally occurring point mutation that is found in 16% of breast cancer patients. Comparison of the wild-type and P132L mutation shows an extension of the secondary helix by 4 amino acids. In Chapter 4, a new construct containing both the scaffolding domain and the transmembrane domain was studied to figure out whether the scaffolding domain is part of the transmembrane domain or a separate domain. CSI analysis of caveolin-1(82-136) shows that the scaffolding domain is an extension of the first helix in the transmembrane domain which suggests that the scaffolding domain is not a separate domain. In Chapter 5 various methods of incorporation of caveolin-1 constructs into bicelles and application of bicelles are investigated. Various methods were able to incorporate small amount of

caveolin-1 constructs into the bicelles but not at a concentration for NMR studies. Moreover, bicelles were used as templating medium to form periodic meso-porous material.

Introduction

Caveolae

The cell membrane is a semipermeable lipid bilayer found in all cells that contains a variety of biomolecules (Figure 1a). Caveolae are 50 to 100 nanometer invaginations located in the plasma membrane of cells (1, 2, 3) (Figure 1b). Caveolae are particularly abundant in endothelial cells, smooth muscle cells, macrophages, cardiac myocytes, and fibroblasts (2, 3, 4, 5). Caveolae act in a variety of cellular process including cell-cell signaling, membrane trafficking, calcium signaling, and lipid recycling (6). These long lived highly-curved structures serve as a platform for signaling processes and significantly increase the effective surface area of the membrane, allowing for greater receptor packing density. The major constituents of caveolae include a array of biomolecules such as cholesterol, sphingolipids, and most notably, the caveolin protein (6, 7, 8). Although, these components are distributed throughout the membrane, they are highly enriched in caveolae. The roles of these biomolecules are proposed to promote bilayer rigidity membrane curvature, allowing caveolae to exist as static structures. Although the exact role of cholesterol plays in caveolae has not been completely fleshed out, it is clear that these biomolecules are important because a loss in caveolae is detected when cholesterol is removed from plasma membrane (7, 8). The most abundant and important protein in caveolae is caveolin, which is classified as an integral membrane protein. Significantly, it has been shown that caveolae are obliterated when the caveolin gene is silenced (9). Moreover, misregulation and mutations to caveolin have been implicated in a variety of disease states including cancer, muscular dystrophy, Alzheimer's, and heart disease (10, 11,

12, 13, 14). Therefore, caveolin is necessary for the caveolae formation and proper cell function.

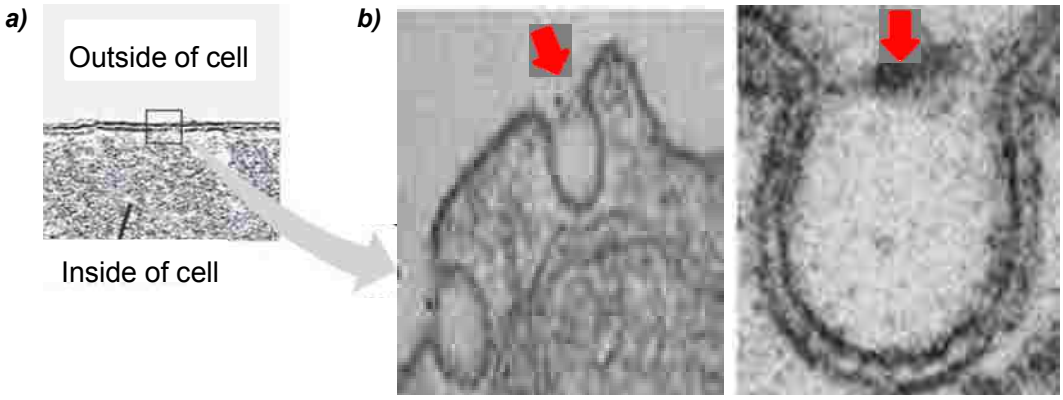


Figure 1. ***a)*** Cell membrane. ***b)*** Tunneling electron micrograph of caveolae, Stan, R. Biochim Biophys Acta (2005) 1746, 334-348. Red arrows denote caveolae.

Caveolin

Caveolin exists as three isoforms: caveolin-1, caveolin-2, and caveolin-3 which have very similar primary sequences (Figure 2). Each protein is approximately 24 kD in size, and caveolin-1 appears to be the most ubiquitous and is necessary for caveolae formation. Caveolin-2 is often co-expressed with caveolin-1 in most cells, while caveolin-3 is located primarily in skeletal and muscle cells. Importantly, caveolins are thought to form high-order homo- and hetero- oligomers *in vivo* that create a scaffold that supports caveolae formation (15, 16).

Previous investigations have postulated that the caveolin protein can be divided into four distinct domains: a soluble N-terminal domain, a scaffolding domain, a transmembrane domain, and a C-terminal domain. The N- terminal domain is soluble, and is exposed to the cytoplasm and not likely interacting with the cell membrane. Studies have revealed that the N-terminal domain does not appear to be structured (17). The scaffolding domain is predicted to form an amphipathic helix that rests on the surface of the membrane. The transmembrane domain of caveolin is buried deep inside of the membrane and adopts a putative horseshoe topology, in that the N- and C- termini reside within the cytoplasm. The C-terminal domain consists of an amphipathic helix that is proposed to rest on the membrane surface. Moreover, caveolin-1 has three palmitoylation sites at cysteine 133, cysteine 143, and cysteine 156 which help in anchoring the protein to the membrane (18, 19). A highly unusual feature of the caveolin protein is the length of the transmembrane domain (Figure 3). The transmembrane domain is proposed to be 33 amino acids long and therefore is too long to span the bilayer once and too short to span the bilayer twice since the average single spanning transmembrane domain requires 20-25 amino acids. In addition, there is no soluble region within the transmembrane region that would facilitate a complete

traverse into the extracellular space. This unusual length of the transmembrane domain gives rise to two plausible orientations of this protein. In one orientation, the transmembrane domain of caveolin penetrates the bilayer diagonally or as kinked helix to have N- and C- termini on different sides of membrane. The second possibility suggests that the transmembrane domain has a turn inside the lipid bilayer and therefore both the N- and C- termini are located on the same side of plasma membrane. Studies utilizing glycosylation mapping shows that both N- and C- termini of caveolin are facing same side of membrane (20). This unusual horseshoe topology is supported by a variety of *in vivo* studies which include immunofluoresence, immunoelectron microscopy, and antibody epitope mapping (15, 16, 20, 21, 22, 23, 24). The results of these studies have led to the postulation that the transmembrane domain forms a horseshoe structure in the membrane (Figure 4). Furthermore, the intra-membrane horseshoe conformation is thought to be a plausible structural mechanism that allows the protein mediated induction of membrane curvature. This can be visualized best by imagining caveolin as a wedge which intercalates in between lipids asymmetrically. This asymmetric insertion of caveolin into the lipid bilayer perturbs the inner leaflet more than outer leaflet. This unbalanced stretching curves the membrane to create caveolae (Figure 5).

Since the transmembrane domain might be critical for the mechanism of caveolae formation, the primary structure of caveolin-1 has been investigated for structural motifs exhaustively. One overarching theme is the level of conservation within the transmembrane domain across different species (E elegans to H. sapiens) and across isoforms. Specifically, two proline residues, proline 110 and proline 132, are conserved in all isoforms and different organisms (Figure 6). It seems as though these two proline residues are playing critical roles. The model of caveolin is based on

primary sequence arguments presented by Parton and co-workers suggest that proline at position 110 could be at pivot point of this unusual horseshoe topology (25). Moreover, Epand and co-workers showed in an *in vivo* study that the N-termini of caveolin are exposed to outside of the cell when proline 110 is mutated to alanine (26). It is overly presumptuous to identify proline 110 as the most critical amino acid for the horseshoe conformation because proline is found in many non-horseshoe turns and the study done by Epand and co-workers failed to identify the location of the C-terminus. However, proline 110 still seems to be a critical amino acid for the formation of the horseshoe topology. This is not surprising as some amino acids, particularly proline and glycine, are often found in the tight turns because of the rigidity of proline, which induces turns, and the high degree of freedom of glycine which is less constricted conformationally. Interestingly there is another conserved proline at residue 132 which is also in the transmembrane domain of caveolin. The mutation of this proline to leucine is a naturally occurring mutation often found in breast cancer patients (13, 27, 28, 29, 30, 31). A diminished number caveolae is an observed phenotype when this mutation occurs. The P132L mutation has been found to result in the retention of caveolin to the level of the Golgi apparatus. Very little or no caveolin reach the final destination which is the plasma membrane. It is still not clear how this mutation disrupts normal caveolin function but misfolded proteins are usually retained in the Golgi.

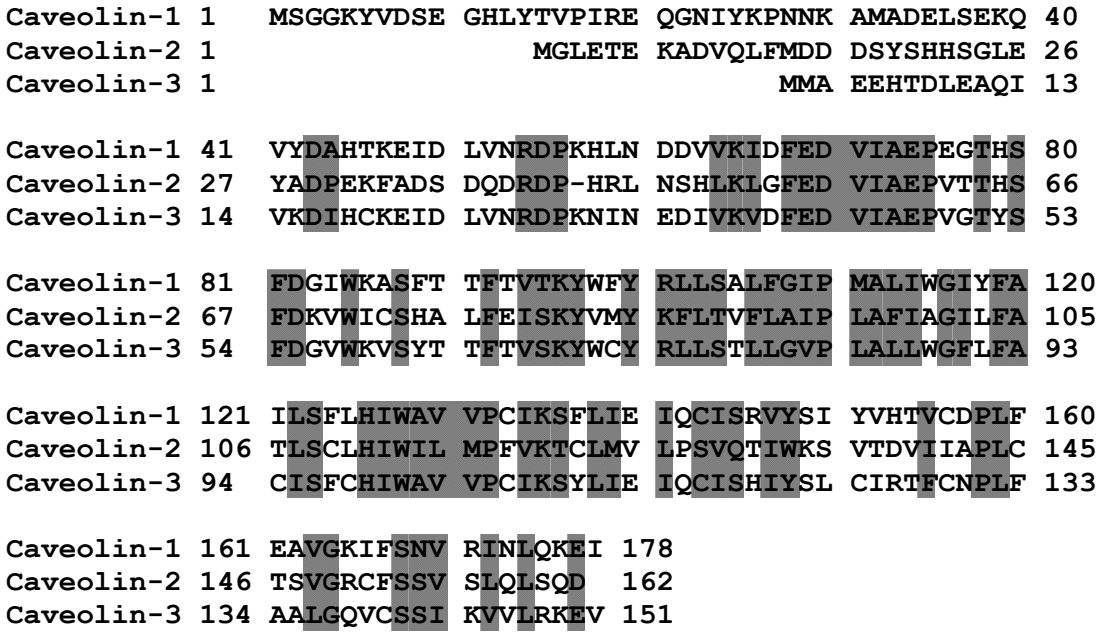


Figure 2. Sequence alignment of three caveolin isoforms (Homo Sapiens); caveolin-1, caveolin-2, and caveolin-3.

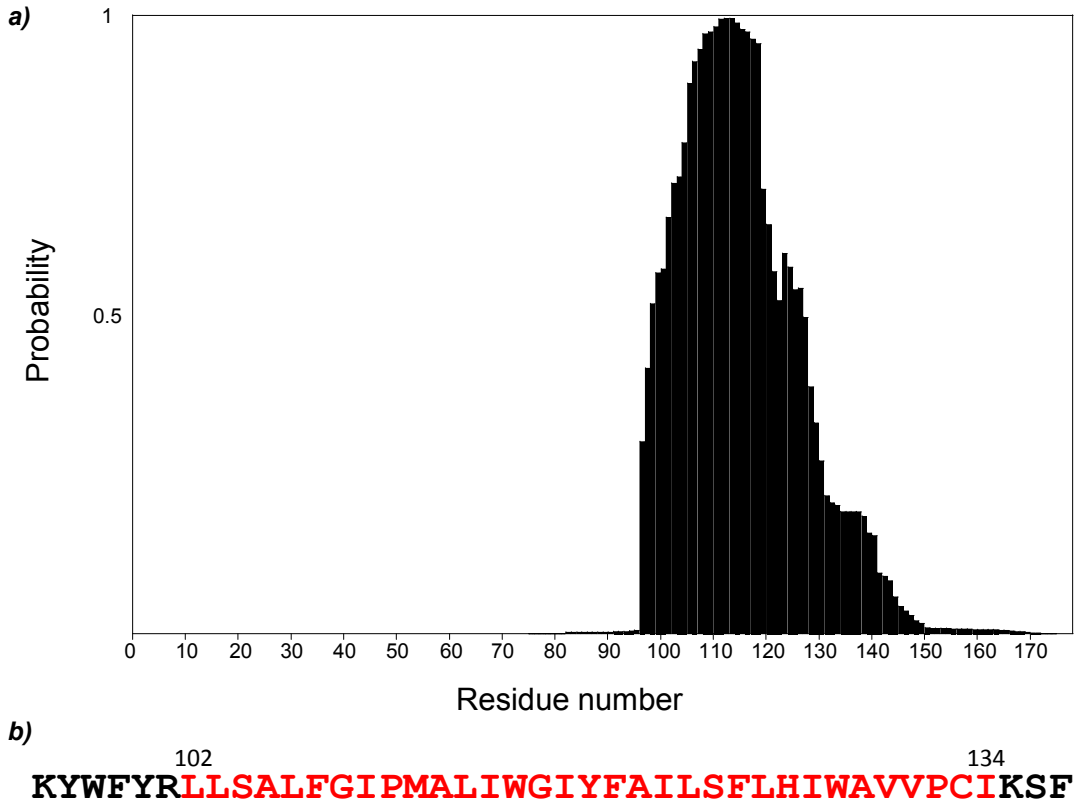
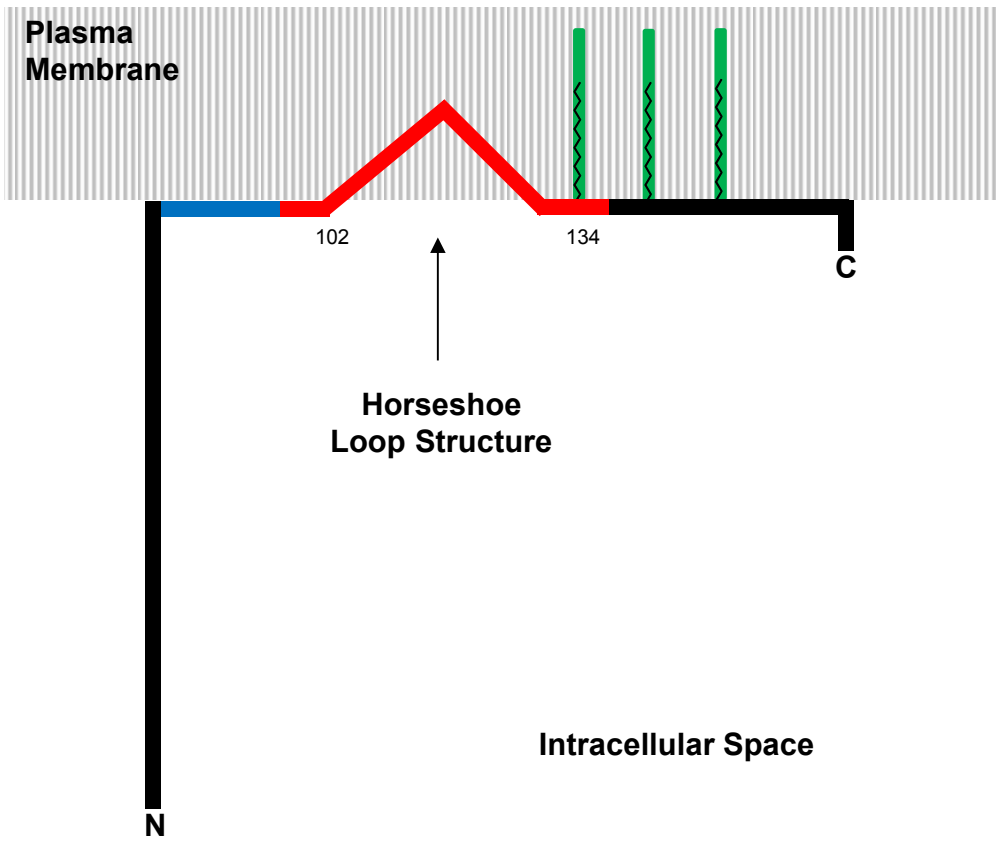


Figure 3. ***a)*** Hydropathy plot of caveolin-1 based on primary sequence. ***b)*** Primary sequece of caveolin-1 transmembrane domain. Hydrophobic stretch is approximately 33 amino acids (red).

a)



b)

MSGGKYVDSE GHLYTVPIRE QGNIYKPNNK AMADELSEKQ VYDAHTKEID
LVNRDPKHLN DDVVKIDFED VIAEPEGTHS FDGIWKASFT TFTVTKYWFY
RLLSALFGIP MALIWGIYFA ILSFLHIWAV VPCIKSFLIE IQCISRVYSI
YVHTVCDPLF EAVGKIFSNV RINLQKEI

Figure 4. ***a)*** Topology of caveolin-1. ***b)*** Sequence of caveolin-1. Blue denotes scaffolding domain, red denotes transmembrane domain, and green denotes palmitoylation.

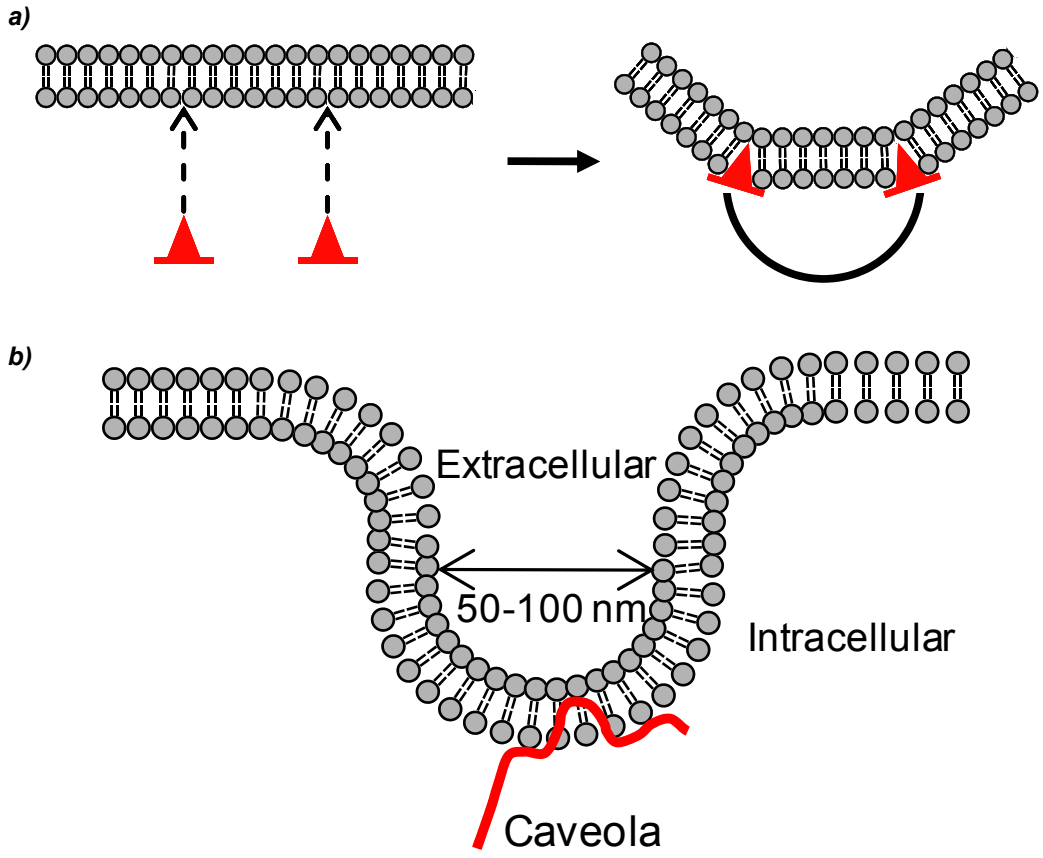


Figure 5. ***a)*** Proposed mechanism of caveolae fomation induced by caveolin.

b) Topology of caveolin in caveolae.

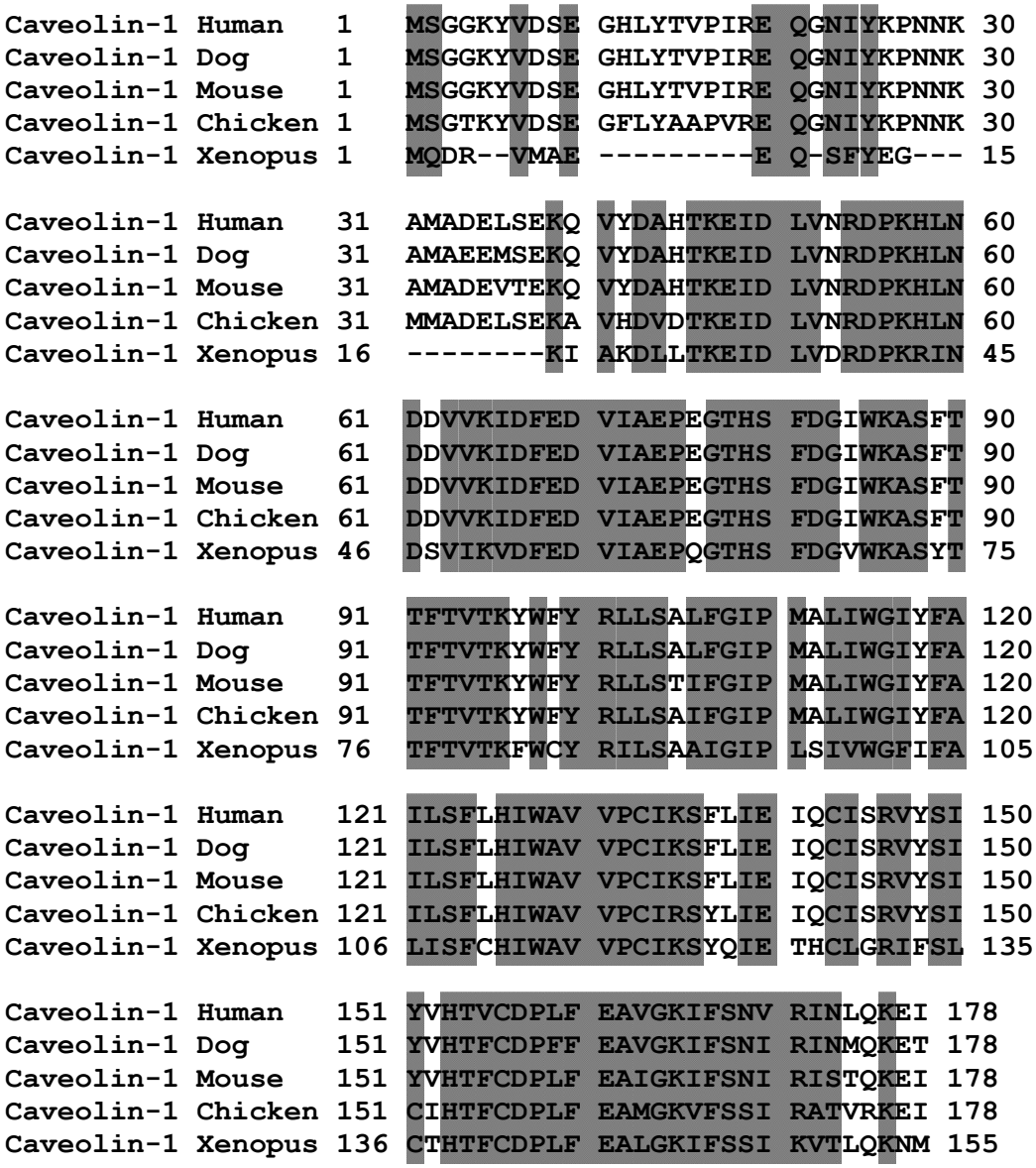


Figure 6. Sequence alignment of caveolin-1 in different organisms.

Membranes

As is the case for determining structural and functional aspects of any membrane protein in its most native like state, an accurate mimic of the cells plasma membrane is required to study the caveolin-1 protein. Membrane proteins including caveolin-1 are mostly composed of hydrophobic amino acids which cause them to be extremely insoluble in aqueous buffers. The insolubility of membrane proteins makes them more challenging to work with. Often times it is necessary to add strong detergents to aqueous buffers in order to solubilize membrane proteins. Unfortunately, the addition of detergents can influence the overall structure of the membrane protein introducing an unwanted bias. Therefore, it is critical to choose a detergent system that most closely mimics the cell membrane environment.

Lipids aggregates that are typically used to study membrane proteins are micelles and vesicles. Micelles have dynamic properties, but it is not the ideal mimic of the cell membrane because of its extreme curvature and lack of a planar bilayer region (Figure 7a). These deficiencies can lead to a biased conformation of membrane proteins that does not accurately depict the structure of these proteins in their native form. However, micelles are still mostly used in solution NMR studies because there is no other reliable mimic that can be adapted easily and the secondary structure of protein is usually conserved (32). Vesicles are perhaps the best mimic because they have less curvature than micelles as well as a distinct bilayer region, and also possess hollow centers much like an actual cell (Figure 7b). The biggest drawback to using vesicles, however, is that they lack the dynamic properties of micelles, and their relatively large size makes them difficult to adapt for NMR studies. Bicelles, on the other hand, combine the dynamic property of micelles with the bilayer properties of vesicles (33, 34, 35, 36, 37, 38, 39). A bicelle is a discoidal lipid aggregate composed of long-chain

phospholipids and detergent molecules (Figure 7c). The center, planar region of the bilayer is typically composed of a long-chain phospholipid, while the curved rim of the bicelle is composed of the short chain phospholipids or detergents that encapsulate and isolate the center bilayer from the aqueous surroundings. For example, a commonly used long-chain phospholipid is dimyristoyl-phosphatidylcholine (DMPC) and a short-chain is dihexanoyl-phosphatidylcholine (DHPC). The characteristics of the bicelle can be varied to suit many different applications by changing the length of the phospholipid chain or the head groups of the phospholipids thereby altering the charge characteristics of the bicelle. For example, the size of a bicelle is dependent on ***q***, which is the molar ratio of the long-chain lipid (DMPC) to the shorter chain lipid (DHPC). For most solution-phase studies, a ***q*** of 0.5 is used, however, larger bicelles can be used depending on the application simply by increasing the ***q***. Also at this ***q*** value, the bicelles have a diameter of 82 Å, which is an appreciable bilayer region sufficient to accommodate full-length proteins. Bicelles have unique properties based on the ***q*** value and identity of components making up the rim and planar regions.

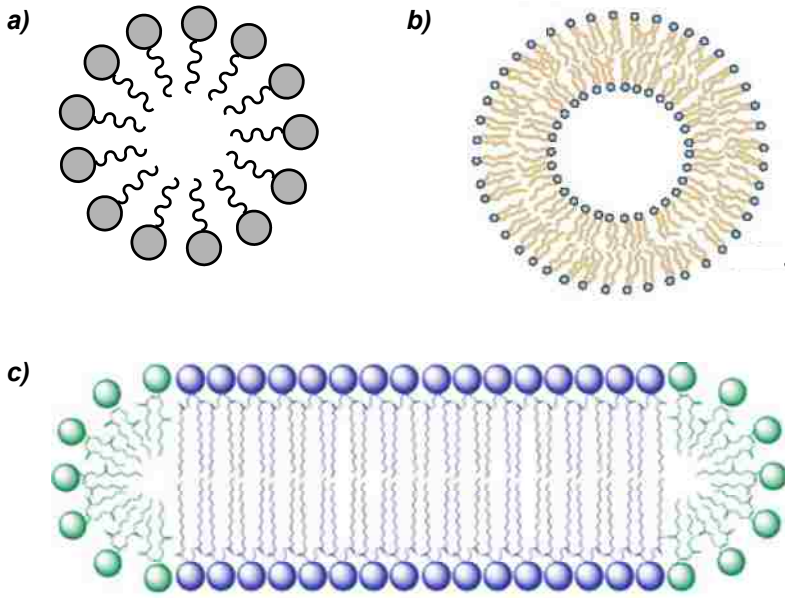


Figure 7. Mimic of cell membrane. ***a)*** Micelle. ***b)*** Vesicle. ***c)*** Bicelle.

Biophysical techniques

There are several methods to study the structure of proteins. One commonly used method to study secondary structure is circular dichroism spectroscopy (40, 41, 42). Circular dichroism spectroscopy is measures the differential absorption of left and right handed circularly polarized light. When circular polarized light passes through a sample that has chiral centers, the molecules will change polarization of the light. Based on changes of the circularly polarized light global secondary structural characteristics can be revealed. Advantageously, the common secondary structural motifs within proteins give rise to predictable patterns in circular dichroism spectra. For example, a protein that contains α-helical properties will give rise to a spectrum that has two minima at 208 nm and 222 nm. The β-sheet property shows minima at 217 nm and random coil shows a minimum at 204 nm. Based on this information, global secondary structure of protein can be revealed using computer program which will give the relative contribution of each structural motif to the entire spectrum (45, 46, 47).

To study the tertiary structure of proteins, two methods are primary used; nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography. Each method has advantages and disadvantages. In X-ray crystallography, samples have to be crystallized in order to carry out the analysis. This is especially difficult for membrane proteins as they often do not form usable crystals. Also, crystallized proteins can be forced to adopt a conformation that may not be representative of their native form in solution. For NMR, full-length proteins (>30kD) are difficult due to their long reorientation rates that results in acceptable spectral broadening. Therefore, separate experiments can be run to obtain information about small segments of a protein. Although, this is not ideal, NMR is advantageous compared to X-ray crystallography because the protein can be analyzed in solution, which more closely

mimics the natural dynamic environment of the protein. Moreover, it is also possible to investigate the influence of other solution properties that will undoubtedly be important to nuances in protein structure such as pH and temperature. For NMR analysis, using traditional one-dimensional NMR to solve protein structure is virtually impossible due to the presence of an overwhelmingly large number of resonances in a given protein sample. There would be a high degree of signal overlap which would preclude an accurate spectrum assignment. To circumvent this problem, two and three dimensional spectroscopic methods are utilized. Initially, a Heteronuclear Single Quantum Coherence (HSQC) spectrum will be obtained by incorporating uniformly isotopic labeled nitrogen into a protein and measuring the resonances of both ^{15}N and ^{1}H. The HSQC spectrum will provide critical information about the backbone of the protein because each peptide bond formed between amino acids contains an amide proton attached to a nitrogen atom which will give rise to a signal. Consequently, the number of peaks in the spectrum should correspond with the number of peptide bonds in the protein (40). The wide dispersion of the HSQC spectrum suggests that each amino acid is surrounded by unique environment. Therefore, a spectrum that has well dispersed peaks represents a well-structured protein. Therefore, based on quality of HSQC experiment, we can decide to move on to three dimensional experiments for backbone assignment.

Many *in vivo* studies have been carried out on caveolin, but there are few biophysical studies that are definitive in presenting a structural model for caveolin. This is mainly because of difficulty in handling this protein as it is extremely hydrophobic. Furthermore, there are no biophysical studies to our knowledge that have investigated the unusual structure of the transmembrane domain of caveolin. Therefore, we investigated the structure of caveolin-1 transmembrane domain. In

particular, we focused on two proline residues which reside in the transmembrane domain of caveolin-1. Caveolin-1 was chosen because it is the most ubiquitous and the transmembrane domains of three isoforms are highly conserved. To investigate the structure of the caveolin-1 transmembrane domain we utilized experiments as outline below. First we over-expressed caveolin-1 transmembrane domain. The Trp leader fusion system was used to bring caveolin into inclusion bodies to obtain high yield. Inclusion bodies are subjected to inclusion body purification methods to obtain extremely pure caveolin-1 protein. Cyanogen bromide cleavage is used to separate the target protein from the fusion partner and the protein is then isolated using C4 reverse phase HPLC and verified by MALDI-TOF mass spectroscopy (28, 43, 44). Second, the pure protein is incorporated into a membrane mimic and is then subjected to CD and NMR spectroscopy analysis to obtain structural information. Analyses of CD data were performed using a web-based computer program called Dichroweb and NMR data was analyzed using various computer programs such as NMRPipe and TALOS+ (45, 46, 47, 48, 49). After obtaining and analyzing wild-type caveolin-1 transmembrane domain, point mutations of proline residues were performed to investigate the role of these prolines. Moreover, a model of caveolin-1 was proposed based on direct structural experiments for the first time.

Chapter 1. Expression and Purification of Membrane Proteins

Abstract

With the growing interest in membrane protein structure and function, a reliable high-yielding method for the expression and purification of highly insoluble membrane proteins is desirable. For these studies, *E.coli* was used as the host for membrane protein expression. Over expression of membrane proteins is typically toxic to *E.coli*, therefore, the expression system that is described in this chapter uses a Trp leader protein fusion system where the membrane protein of interest is fused to the Trp leader protein sequence. During protein expression, the Trp Leader membrane protein fusion directed way into inclusion bodies in order to avoid premature cell death from toxicity of the membrane protein. The inclusion bodies are carried through a series of wash steps in order to obtain semi-pure proteins. For the purification of a small 2 kD membrane protein PFOA nickel-affinity chromatography was used to obtain highly pure fusion protein. Cyanogen bromide cleavage was performed to separate the target protein from the Trp Leader fusion partner, and C4 reverse phase high performance liquid chromatography was carried out to isolate the target membrane protein. The product was verified by matrix-assisted laser desorption and ionization-time of flight (MALDI-TOF) mass spectroscopy. This expression method was also adapted as a cost effective means to obtain isotopically labeled proteins.

Introduction

It is well known that approximately one third of the human genome encodes for membrane proteins. Membrane proteins play a critical role in cell signaling, signal transduction, and growth regulation (21). There are three classes of membrane proteins. First, integral membrane proteins (transmembrane protein) span membrane and are composed mostly of hydrophobic amino acids. Second, peripheral membrane proteins associated with lipid bilayer and only penetrate the surface of the lipid bilayer. Lastly, lipid anchor proteins attached to lipid bilayer anchoring to lipid head group. Integral membrane proteins have been shown to play an important role in the function of membrane proteins. For example, the pore structure of a transmembrane domain of an ion channel allows specific ions to pass through membranes (50, 51, 52). Therefore, reliable and efficient methods for the expression and purification of transmembrane domains are needed in order to drive progress in this area forward. In vitro research on integral membrane proteins is still in its infancy compared to soluble proteins due to the fact that membrane proteins are very hydrophobic and highly insoluble in traditional buffers which presents significant challenges when it comes to producing them. Membrane proteins can be synthesized using traditional molecular cloning methods whereby an organism is used as a vector for protein expression. Typically, *E.coli* is the organism of choice for over-expressing proteins because it is easy to culture and obtain protein yields are high (53, 54). However, in the case of membrane proteins, expression is difficult due to the fact that mammalian membrane proteins are highly toxic to the *E.coli* organism. The over-expression of these proteins leads to premature cell death before a sufficient protein yield can be achieved. To circumvent this problem, membrane proteins can be expressed directly into inclusion bodies. To achieve this a fusion system is utilized where one protein, usually a protein that *E.coli* expresses

natively , is attached to the target protein through the preparation of a gene that encodes the fusion protein. The fusion partner directs the entire fusion protein into inclusion bodies, which are comprised of misfolded, insoluble aggregated proteins. The inclusion bodies protect the host organism from premature cell death and premature degradation. This has allowed high expression levels of membrane proteins to be obtained using *E.coli*. Usually a series of buffered wash steps is employed to semi purify the inclusion bodies after they have been liberated from the *E.coli* cells (43).

Small proteins (approximately 2kD) are more difficult to express compared to larger membrane proteins (approximately 4-6kD) using the current established methods. The most common method of obtaining small proteins is solid-phase synthesis; however, there are limitations to this method as well. For example, nuclear magnetic resonance (NMR) studies require isotopic-labeled nitrogen and carbon in milligram quantities, but for solid-phase synthesis, this is not cost effective because reagents are very expensive and the yield is very low. Few approaches to expressing and purifying membrane proteins have been successful, yet there are no reliable or reproducible methods for preparing high quantities of small hydrophobic proteins. In this study, a reliable method for producing milligram quantities of small hydrophobic protein has been developed. This method can easily be adapted for the preparation of isotopically-labeled proteins for NMR characterization.

Material and Methods

Construct Design

Caveolin-1(96-136)

Caveolin-1(96-136) contains the full intact transmembrane domain of the caveolin-1 protein based on the hydrophobicity analysis of the protein's primary sequence. Since caveolin is known to create curvature in the membranes of cells, the working hypothesis is that the transmembrane domain of the protein is playing a key role. NMR is not conducive to large membrane protein structure, so the full-length caveolin protein structure was not initially investigated. The transmembrane region is predicted to form a horseshoe loop structure, which effectively curves the bilayer to form caveolae. The construct of the caveolin-1 protein was designed as a fusion to the Trp leader protein (Figure 1-1). The Trp leader and caveolin-1 transmembrane domain (96-136) fusion protein was used because the Trp leader protein promotes higher expression levels of caveolin-1(96-136). Trp leader promotes protein expression into inclusion bodies and brings the fused caveolin-1(96-136) into the inclusion bodies with it. Since caveolin-1(96-136) is extremely hydrophobic, expression under soluble conditions will quickly create a toxic environment that will kill the bacterial cells in which the protein is being expressed. This results in a lower protein yield. Therefore, the inclusion body expression method is ideal as it allows expression of caveolin-1(96-136) at much higher levels because the inclusion bodies protect the expression host from the cytotoxic caveolin protein. Additionally, a methionine residue is added between Trp leader and caveolin-1(96-136) which allows the caveolin-1(96-136) protein to be cleaved from the fusion protein using cyanogen bromide. Because cyanogen bromide cleavage occurs after all methionine residues, and caveolin-1(96-136) contains one

native methionine residue, this native methionine was mutated to leucine (M111L) based on sequence homology to caveolin-2 and caveolin-3 isoforms. Also one cysteine residue was mutated to serine (C133S) to avoid unwanted disulfide bonding. The mutation of cysteine to serine is considered conservative as this mutation was shown to not interrupt correct caveolin-1 traffinkng to the plasma membrane.

Caveolin-1(122-142)

The Caveolin-1(122-142) protein represents a region of the protein that has been implicated in a variety of disease states such as cancer. In particular, a P132L mutation in caveolin is associated with poorly functioning caveolin protein and serves as a marker for various disease states including breast cancer. In order to study the effects of this significant mutation, we have prepared a construct of the caveolin protein that encompasses residues 122-142. In this construct, the P132 residue resides at the center of the Caveolin-1(122-142) protein, and is flanked by 10 residues on each side. This construct was prepared in the same manner as the previous constructs using the Trp leader fusion system. However, because this protein is so small, a double fusion system was also employed. The ubiquitin protein was attached in tandem to the Trp leader protein followed by caveolin-1(122-142). To facilitate the purification of this construct, an octa-histidine tag was added to the front of the construct so that Ni-NTA affinity chromatography could be used for purification.

RfbP

The RfbP protein is approximately 25 kD in size and is native to the bacterium, Salmonella enteric. RfbP transfers a galactose moiety from uridine 5'-diphosphogalactose to undecaprenylphosphate to form undecaprenyl-pyrophosphate-galactose (55, 56). Undecaprenol is extremely hydrophobic and it is suspected that the

transmembrane domain of the RfbP protein may be required for proper function and activity. To decipher the mechanism of galactose transfer it is important to solve the structure of the transmembrane domain of this protein. To design the construct for these studies, the single spanning transmembrane domain of RfbP was used. This domain is approximately 20 amino acids in length and contains extremely hydrophobic residues. To enhance the solubility of this construct extra residues were added to both ends of the transmembrane domain. The final construct contained RfbP residues 18-51 which has an approximate molecular weight of 4.9 kD. The Trp leader fusion system was used to express this protein.

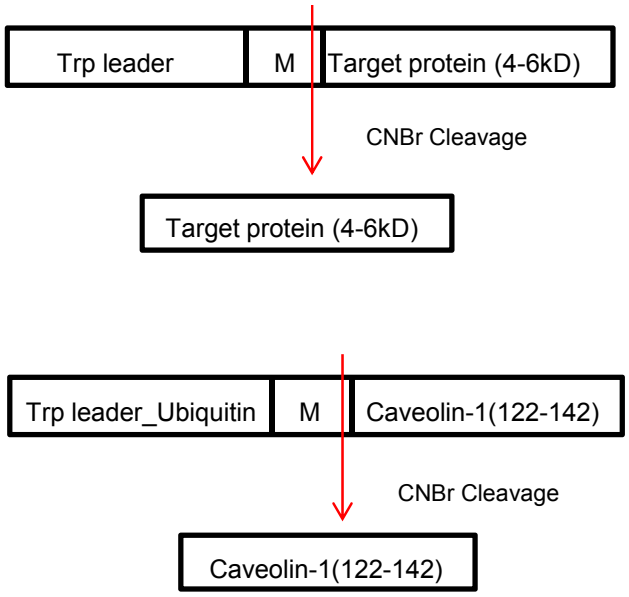


Figure 1-1. Design of fusion protein construct using Trp leader and Trp leader-Ubiquitin. Separation from the fusion protein by cyanogen bromide was accomplished by inserting a methionine between the fusion protein and the protein of interest. Target proteins investigated were various caveolin-1 constructs and RfbP.

Cloning and Expression

First, caveolin-1(96-136), caveolin-1(122-142) and RfbP(18-51) genes were prepared by obtaining two nucleotide oligomers with a 25 base pairs complementary overlap (Figure 1-2). The caveolin-1(96-136) and RfbP (18-51) genes were each cloned into the pET-24a vector containing the Trp leader fusion using the BamH1 and EcoR1 restriction sites. The Caveolin-1(122-142) gene was cloned into a pET-24a vector containing the N-terminal octa-histidine Trp leader-ubiquitin sequence using BamH1 and EcoR1 as well (Figure 1-3).

The forward and reverse primers (Appendix 1-1) were prepared at a 2.0 µg / µL concentration in sterile distilled water. A PCR reaction was prepared using 1 µL of each primer, 5.0 µL of 10X Thermapol buffer, 0.4µL of dNTPs (100 mM total), and 1µL of Taq polymerase on a 50-µL scale with volume adjustment using sterile water. The thermal cycler was set to the following cycling conditions: 1 cycle at 95 °C for two minutes, 5 cycles at 95°C for 15 seconds, 45 °C for 15 seconds, 72 °C for 1 minute, 1 cycle at 72 °C for 5 minutes, 4 °C for an indefinite period (Appendix 1-2). The PCR product was purified using the “PCR Quick cleanup procedure” outlined in the Qiagen Gel Extraction kit. The purified PCR product was digested with the restriction enzymes BamH1 and EcoR1 to create “sticky ends”, which allows the gene to be incorporated into the desired vector. The desired vector was digested with BamH1 and EcoR1 in a separate reaction. The digest reactions were carried out overnight at room temperature on a 50 µL scale consisting of 42.5 µL of PCR product or pET-24a vector, 5 µL of 10X EcoR1 buffer, 1 µL of BamH1, 1 µL of EcoR1, and 0.5 µL of 100X BSA. The pET-24a vectors containing Trp leader or Trp leader-Ubiq fusion were obtained from the Komives group at UCSD. The DNA encoding the desired protein and the vector were quantitated by 1% agarose gel electrophoresis. Insertion and ligation of

DNA into the vector were carried out at room temperature overnight. The reaction mixture containing a 1:5 ratio of vector to DNA with appropriate amount of 10X T4 DNA ligase buffer (based on concentration of vector and DNA) and 1 µL of T4 DNA ligase. After ligation, 1 µL of ligation reaction mixture was transformed into 50 µL XL1 blue subcloning grade cells (Appendix 1-3). The cells were grown overnight at 37 ºC on an LB agar plate containing kanamycin, which allows for the selective growth of bacteria containing the ligated vector. The pET-24a vector confers kanamycin resistance which allows for selective screening. Colonies were picked and grown at 37 °C overnight in 5 mL of LB broth containing kanamycin. The cells were harvested and the DNA was isolated using the Qiagen Miniprep kit (Valencia,CA) procedure. The purified DNA was sent to Genewiz (South Plainfield, NJ) to confirm that the correct sequence had been cloned in. The sequencing results confirmed successfully transformed cells. The sequenced DNA (20 ng) was transformed into *E.coli* BL21 (DE3) cells (100 µL) and was then plated kanamycin MDAG plate overnight at 37 ºC (Appendix 1-4). BL21 (DE3) cells are equipped with the necessary cellular machinery which will allow them to express the protein of interest. To test the transformed colonies ability to express the target protein, a small scale growth was performed with 1 mM IPTG induction. The expression level of target protein can be checked by running SDS-PAGE electrophoresis (Figure 1-4). Once the colony is chosen, a large scale growth is performed.

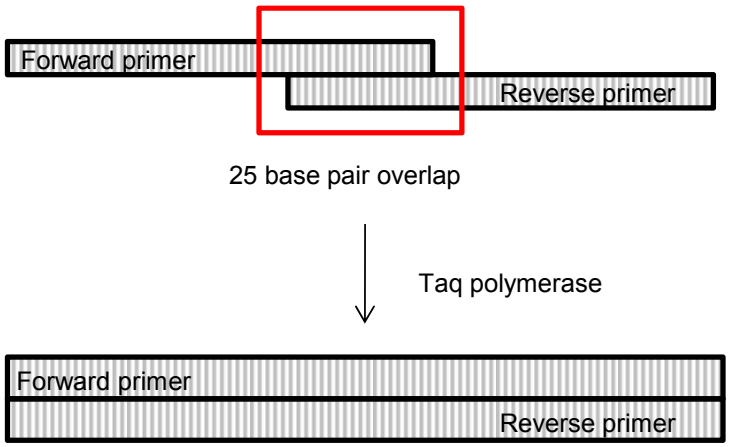


Figure 1-2. Synthetic preparation of target genes with 25 base pairs overlap.

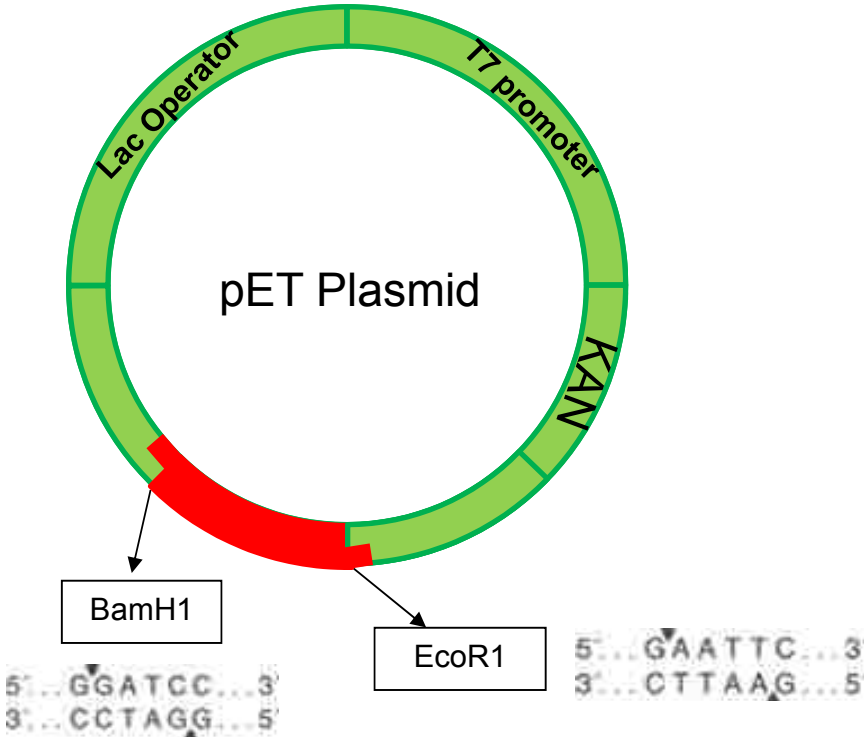


Figure 1-3. Map of the pET-24a vector expression system containing Trp leader or Trp leader-Ubiquitin. Synthesized gene (red) was inserted using BamH1 and EcoR1 restriction sites.

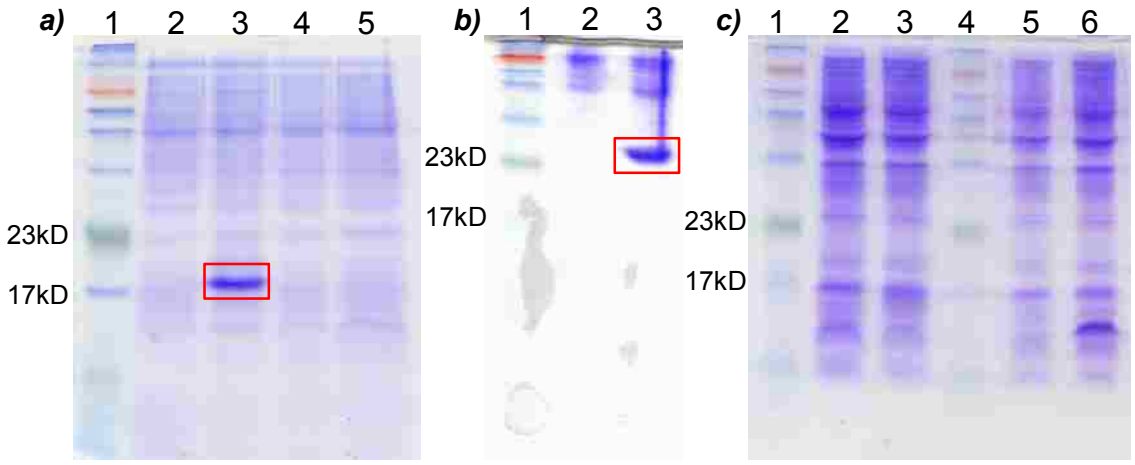


Figure 1-4. ***a)*** Test expression of caveolin-1(96-136). Lane 1. Molecular marker, lane2 and 3 Uninduced / Induced Trp leader fusion protein, lane 4 and 5. uninduce/induced Ubiquitin fusion protein. ***b)*** Test expression of caveolin-1(122-142). Lane 1. Molecular marker, lane2 and 3. Uninduced / Induced Trp leader-Ubiquitin fusion protein. ***c)*** Test expression of caveolin-1(122-142). Lane 1. Molecular marker, lane2 and 3 Uninduced / Induced Ubiquitin fusion protein, lane 4. Molecular marker, lane 5 and 6 Uninduce / Induced KSI fusion protein.

Cell growth and Isotope labeling

A large-scale cell growth is performed using the auto-induction method developed by Studier *et al* (57). For NMR studies the ^{15}N-labeled protein constructs have been successfully prepared using minimal auto-inducing media containing $^{15}NH_4Cl$ as the sole nitrogen source for the bacterial cells. Moreover, other labeling methods are optimized to obtain protein for NMR structural studies. The cells are grown to saturation in the auto-induction media and harvested (Figure 1-5). The cells were harvested by centrifugation at 5,000 x g for 30 minutes at 4 ˚C. The pelleted cells were washed and resuspended in ice cold saline solution (0.9% NaCl). In this section, various cell growth conditions were explored in order to optimize the yield of each of the protein constructs.

Rich media cell growth

A large-scale (6-liter flask, 1-liter of media) cell growth is performed using the auto-induction method developed by Studier *et al* (57). 1-liter ZYM-5052 media is prepared in 6-liter Erlenmeyer flask and incubated at 37ºC for at least an hour to reach thermal equilibrium (Appendix 1-5). One milliliter of starter (5mL growth in non-inducing MDAG media (Appendix 1-6) was added for every liter of media at a shaking speed of 225 rpm at 37 ºC to reach saturation point. The bacterial growth of the culture was monitored every hour at 600 nm until the optical density of the growth no longer changed. The cells are grown to saturation in the auto-induction media and are harvested as previously described.

Uniformly labeled ^{15}N cell growth in minimal media

A large scale growth for ^{15}N labeled target protein was performed using the auto-induction method. N5052 minimal media was chosen because ammonium chloride is the sole nitrogen source in the media (Appendix 1-5). Substitution of ^{14}N ammonium chloride to ^{15}N ammonium chloride can effectively label protein with ^{15}N nitrogen. Like the rich media cell growth method, 1 liter of sterile minimal media is prepared and incubated at 37 ºC for one hour to reach temperature equilibrium then 1 mL of MDAG starter is added to 1 L of media. The growth is incubated at 37ºC while a shaking speeds of 225 rpm for approximately 18 hours in order to reach the saturation point.

^{13}C, ^{15}N labeling with exchanging media method

To assign the protein backbone using solution NMR, nitrogen and carbon atoms must be labeled with ^{15}N and ^{13}C. In order to incorporate these isotopes into our protein, a cell growth was performed in M9 minimal media using a procedure that was optimized by Marley et al (58). This protocol uses ^{13}C glucose and $^{15}NH_4Cl$ as the only carbon and nitrogen sources, which forces the bacteria to incorporate these labels into the protein as it is synthesized. The cell growth was started by adding 1 mL of starter to 1-liter of LB broth. As soon as the optical density (600 nm) reaches 0.6 - 0.8, harvest the cells by centrifuging at 5000 x g for 30 minutes at 25 ºC. Next, the cells were washed with 250 mL of wash buffer which is M9 minimal media without any nitrogen ($^{15}NH_4Cl$) or carbon (^{13}C glucose) source. After resuspension in the wash buffer, the cells were centrifuged at 5000 x g for 30 minutes at 25ºC. After the cells have been pelleted a second time, they were resuspended in 250 mL of M9 minimal media containing ^{13}C glucose and ^{15}N nitrogen and was transferred to a sterile 6 Liter

flask. The culture was incubated at 37 °C for one hour, with shaking at 225 rpm. Protein expression was induced by the addition of 0.8 mM IPTG. The culture was incubated for another 6-8 hours and harvested. Incorporation of ^{2}H labeling of the protein for beta carbon information in the protein, the water is exchanged to deuterium in the M9 minimal media and is incubated twice as long in buffers before and after IPTG induction (approximately 2 hours before IPTG induction and 12hours after IPTG induction).

Specific labeling method

To facilitate protein backbone assignments, specific amino acid labeling is necessary. Isotopically labeled valine, alanine, phenylalanine, isoleucine, and leucine are introduced to the media during growth. Slight modifications were made to the procedure previously described and the optimization of these growths was based on the method of Truhlar et al (59). For example, to obtain only ^{15}N-alanine-labeled protein, M9 minimal media supplemented with ^{15}N alanine and a 5-fold excess of other ^{14}N amino acids is prepared. First, 1 Liter of M9 minimal media is prepared in a 6 L beaker. To the M9 media 500 mg of ^{14}N amino acids and 100 mg of ^{15}N amino acid were added. After all the amino acids have been added to the media, it was sterile filtered using a 0.2 µm filter and incubated at 37 ºC for one hour. 1 mL of MDAG starter was added to 1 Liter of media and the optical density of the growth was monitored at 600 nm to a final OD of 0.6. IPTG was added to a final concentration of 1 mM to induce protein expression and the culture was left to grow for 6 hours at 37 °C with shaking at 225 rpm.

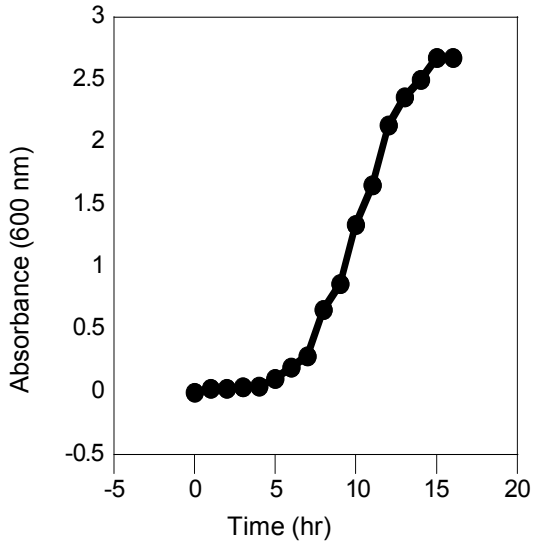


Figure 1-5. Cell growth profile of Trp leader caveolin-1 fusion protein.

Purification

Trp Leader fusion proteins

To isolate the protein from *E.coli* cells, a series of selective wash steps were performed (Figure 1-6) (43). In the first wash step, the cell pellet from a 1-liter growth was resuspended in 200 mL of a buffer containing 20% (w/v) sucrose, 20 mM Tris-HCl pH 8.0 with 1 mM βME and 1 mM EDTA. Beta-mercaptoethanol was added to a final concentration of 1 mM during the wash steps in order to prevent any undesired disulfide bond formation and to keep the methionine residue reduced, which is essential to the cyanogen bromide reaction. The resuspended material was sonicated using a Branson tip sonifier at a power output level of 10 and a 50% duty cycle with vigorous stirring for 15 minutes. The inclusion bodies and the soluble proteins are liberated from the bacterial cells. The insoluble inclusion bodies and cell debris are separated from the aqueous proteins by centrifugation at 27,500 x *g* for 2 hours at 10°C. The pellet containing the insoluble material is subjected to a second wash step using 200 mL of 20 mM Tris-HCl, pH 8.0 buffer containing 1% Triton X-100 (v/v) and 1 mM BME. The mixture is sonicated again to resuspend the insoluble material using the conditions above. The mixture is subjected to centrifugation at 27,500 x *g* for 1 hour at 10°C. The second wash step is designed to separate the broken cell membrane pieces along with the membrane proteins from the highly insoluble inclusion bodies. The desired protein is protected from this second wash step by the highly insoluble inclusion bodies, which do not readily dissolve in Triton X-100. Therefore, they are not removed with the supernatant in the second wash step and only membrane proteins that are soluble in Triton X-100 are removed. The last wash step uses 100 mL of 20 mM Tris-HCl pH 8.0 containing 60% (v/v) isopropanol (Mallinckrodt, Hazelwood, MO) which removes any traces of the Triton X-100 detergent.

Resuspended pellets are centrifuged at 50,000 x *g* for 1 hour at 10°C. Isopropanol is effective at increasing the critical micelle concentration of the Trition X-100 detergent, which makes it very effective at removing all traces of Triton X-100. After the isopropanol wash the semi-pure inclusion bodies are dried well and solubilized in 88% formic acid with a dounce homogenizer and centrifuged at 50,000 x *g* for 30 minutes at 25°C. The highly insoluble inclusion bodies are soluble in 88% formic acid (JT Baker, Phillipsburg, NJ), which is advantageous because the cyanogen bromide cleavage requires highly acidic conditions for optimal cleavage. Cyanogen bromide is added to the supernatant to cleave the fusion protein at the methionine position thereby liberating the caveolin protein from the Trp Leader fusion protein. The cyanogen bromide reaction is left to proceed at room temperature for approximately 18 hours, and it is subsequently dried down using a vacuum concentrator (Thermo Scientific) to remove unreacted cyanogen bromide and formic acid. The dry pelleted target protein was resuspended in a small volume of 88% formic acid; approximately 5 mL per 25 mg of dry protein. The dry protein was subjected to reverse-phase HPLC to separate the Trp Leader protein from target protein using a C4 bonded stationary phase. To separate the target protein from Trp Leader, a mobile phase with a 1% gradient was used beginning with 100% of solution containing 80% water / 20% acetic acid and changing to a solution of 80% 1-butanol / 20% acetic acid. All peaks were collected and the presence of the target protein was verified using MALDI-TOF mass spectrometry. The target protein was lyophilized from HFIP and stored at -20 ºC for later usage.

Trp Leader_Ubiquitin Fusion Proteins

A 2-liter cell pellet is processed the same as the Trp leader fusion protein mentioned above to remove the unwanted soluble and membrane-bound proteins.

Beta-mercaptoethanol can be added to a final concentration of 1 mM during the wash steps. The semi-pure pellet was dissolved in 8% (w/v) PFOA (pentadecafluorooctanoic acid) (Oakwood Products, Inc. West Columbia, SC), 25 mM phosphate pH 8 buffer. To facilitate the dissolution of the pellet, it was sonicated for 15 minutes to obtain a homogeneous solution. The solution was centrifuged at 50,000 x g for 30 minutes at 20 °C. The supernatant was retained and filtered through a 0.2 micron filter before loading onto a nickel affinity column. The bound material was washed with 3 column volumes of 1% (w/v) PFOA, 25 mM phosphate pH 8 buffer. The bound material is eluted from the column using 1% (w/v) PFOA, 25 mM phosphate pH 4 solution (Note: a 1% PFOA, 25 mM phosphate pH 8 buffer containing 250 mM imidazole can also be used to elute the nickel column). Pure full-length fusion protein was successfully recovered and dialyzed immediately against 15-liters of 10 mM ammonium sulfate pH 8 containing 1 mM βME (betamercaptanol) to remove PFOA. Several buffer exchanges were performed to completely remove PFOA. It is evident that most of the PFOA was removed because the fusion protein precipitates out of solution in the dialysis bag. After dialysis, the precipitated protein was collected by centrifugation at 15,000 x g for 30 minutes at 20°C. The pellet was dissolved in 70% formic acid and a 100-fold excess of cyanogen bromide was added to the cleavage reaction. The reaction was bubbled under a blanket of nitrogen for 5 minutes and left to proceed rotating in the dark for 16 hours at room temperature. This was sufficient for complete cleavage of the target protein. The reaction mixture was dried down in a SpeedVac® concentrator to remove excess formic acid and cyanogen bromide. The cleaved protein was dissolved in 88% formic acid (approximately 25 mg in 5 milliliters) and injected onto a C4 reverse-phase HPLC column. The column was run at a flow rate of 10 mL per minute at a 0.6 % mL per minute gradient using a 20% acetic acid and 80% water to 20% acetic acid and 80% butanol solution. The desired protein was collected and

verified by MALDI-TOF mass spectroscopy. The protein was dried out of the HPLC solvent and resuspended in a 20% HFIP solution, immediately frozen in liquid nitrogen and lyophilized to obtain a powder. The lyophilized powder is stable stored at -20 ºC for several months.

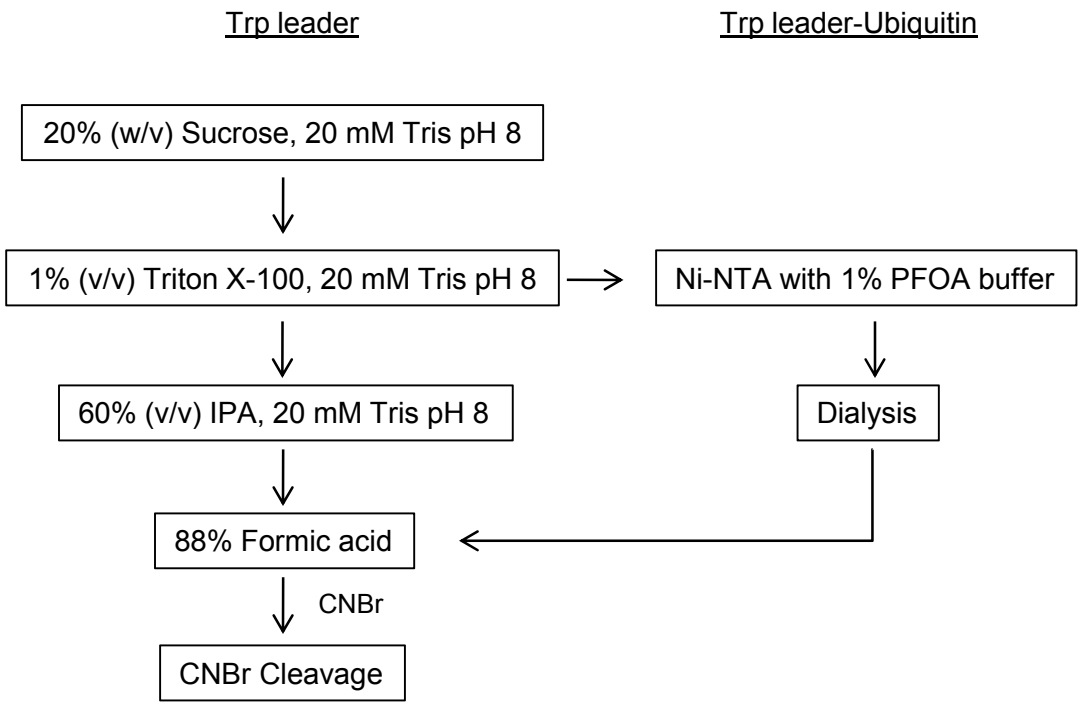


Figure 1-6. Flow chart of Trp leader and Trp leader-Ubiquitin fusion protein purification steps.

Results and Discussion

Expression

One of the biggest challenges in working with integral membrane proteins is that it is difficult to obtain purified material for in vitro studies. Extremely hydrophobic transmembrane proteins especially mammalian proteins ranging in size from 25 to 45 amino acids are poorly expressed in *E.coli*, in general, therefore it is hard to achieve sufficient amounts for structural studies. To circumvent this problem, the transmembrane domains of membrane proteins are typically expressed as fusion proteins. For our studies Trp leader was used because it is native to *E.coli*, is readily expressed and brings attached proteins into inclusion bodies. This fusion system dramatically improves expression level of the transmembrane proteins. To find a suitable fusion partner, several proteins were chosen and tested for their expression level of transmembrane domains of membrane proteins. Expression with the Trp leader fusion system directs toxic membrane proteins into inclusion bodies so that high expression levels can be achieved. This is likely due to the fact that mammalian membrane proteins are toxic to the cell during overexpression. The *E.coli* membrane has limited space to accommodate membrane protein insertion. Therefore, unless protein is precipitated into inclusion bodies the membrane protein will easily saturate the *E.coli* membrane and lead to premature cell death. Another well-known fusion partner, ubiquitin was also tested. The ubiquitin fusion system is known to bring expressed proteins into soluble regions of the cell which can make purification more facile by using conventional methods. Moreover, another fusion system, KSI (ketosteroid isomerase) was chosen and tested. KSI is a widely used fusion system that promotes protein expression into inclusion bodies. However, the desired studies, KSI fusion system were not able to over-express target proteins (Figure 1-4). Very low

protein yields were obtained using the ubiquitin fusion system and most of the protein expressed was truncated. The Trp leader proved to be the most successful at producing milligram quantities of full-length protein.

To optimize the Trp leader system several conditions of Trp leader were tested. The Trp leader fusion system was tested using a His-12 tag, but the protein expression level was low. Repeating codons of the His-12 tag were not favorable for *E.coli* and resulted in low expression level. The highest expression level was obtained when the codons were optimized and the His-tag was removed.

Although the caveolin-1 construct only expressed well with the Trp leader fusion partner, the RfbP(18-51) construct expressed well with the Ubiquitin fusion partner. This is likely because the RfbP(18-51) construct had more soluble amino acids compared to the caveolin-1 constructs. However, the Trp leader fusion system appeared to be most effective and reliable when expressing extreme hydrophobic proteins. In addition, expression of small proteins (approximately 2 kD) can be enhanced by adding ubiquitin which is a well-expressing protein in *E.coli* following Trp leader. The expression level of the Trp leader fusion protein was sufficient to obtain milligram quantities for NMR structural determination. Using these methods caveolin-1(96-136), caveolin-1(122-142), and RfbP(18-51) were successfully expressed.

Purification

Optimization of inclusion body purification

Once the protein is successfully expressed in the Trp leader fusion system, it is purified and isolated from the fusion protein. Several wash steps will lead to semi-pure protein and cyanogen bromide cleavage will isolate the target protein by cleaving after a single methionine residue, which is incorporated between Trp leader and the target protein. The first wash step with 20% sucrose was effective at removing most of the soluble proteins. This step was optimized using different concentrations of sucrose and verifying the presence of protein by SDS-PAGE electrophoresis (Figure 1-7). Based on the experimental data, adding sucrose to the first purification step was effective, but the concentration of sucrose had to be optimized to use as a reliable method. Therefore, different concentrations of sucrose were used and 20% sucrose made the buffer dense enough such that most of the unwanted proteins floated and the inclusion bodies pelleted. The second wash step with 1% Triton X-100 was effective at removing most of the membrane proteins that are soluble in 1% Triton X-100. The target protein was extremely hydrophobic and remained in the inclusion bodies. Several different detergents were tested such as Tween-20, Triton X-100, NP-40 and SDS. Tween-20 was too mild and could not remove the soluble membrane proteins effectively. The Triton X-100 was mild enough to preserve the inclusion bodies and wash away most of the other proteins. The NP-40 was slightly stronger compared to Triton X-100 and there is some loss of the target proteins. SDS was too strong to use because it dissolved the inclusion bodies. Based on the experiments, the most effective and reliable detergent to purify the inclusion bodies was Triton X-100. After the detergent wash step, a 60% isopropyl alcohol (IPA) wash was performed to remove Triton X-100 effectively by increasing the critical micelle concentration. The remaining

inclusion body pellet was dissolved in 88% formic acid for cyanogen bromide cleavage. Using this method caveolin-1(96-136), caveolin-1(82-136), and RfbP was successfully purified.

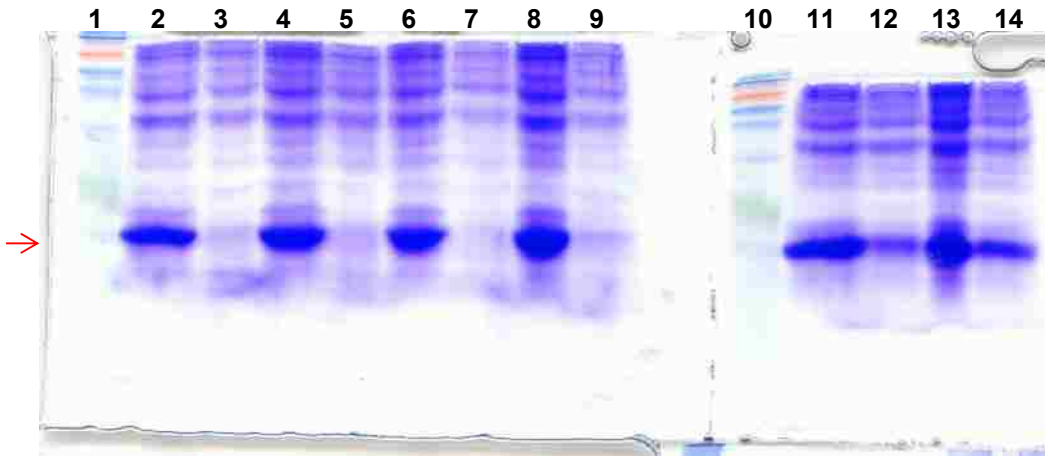


Figure 1-7. SDS-PAGE gel of sucrose concentration test. Lanes 1 and 10 is the MW ladder; lanes 2 and 3 are pre/post centrifugation in 20 mM Tris, 0.5 M sucrose, pH 8; lanes 4 and 5 are pre/post centrifugation in 20 mM Tris, 0.7 M sucrose, pH 8; lanes 6 and 7 are pre/post centrifugation in 20 mM Tris, 0.9M sucrose, pH 8; lanes 8 and 9 are pre/post centrifugation in 20 mM Tris, 1.1 M sucrose, pH 8; lanes 11 and 12 are pre/post centrifugation in 20 mM Tris, 1.3 M sucrose, pH 8; lanes 13 and 14 are pre/post centrifugation in 20 mM Tris, 1.5 M sucrose, pH 8.

Ubiquitin_RfbP

Purification of the RfbP(18-51) construct with the ubiquitin fusion partner was performed using Ni-NTA affinity chromatography. However, the protein was not soluble in 8 M Urea or 6 M GuHCl buffer alone but with aid of 0.1% SDS it was soluble in the buffer system. However, the problem with this method was that the Ni-NTA column has minimal compatibility with SDS detergent can solubilize the target protein. Therefore, other mild detergents were tested to find the optimal detergent that is soluble and compatible with Ni-NTA affinity column. Mild detergents such as Tween-20 or Triton X-100 were not powerful enough. Therefore, 0.1% SDS with 8 M urea was used to solubilize the protein and load the column. The column was washed using buffer containing 250 mM imidazole, 0.1% SDS and 8 M Urea. After purification SDS cannot be completely reomoved. Therefore, we were not able to find an optimal cleavage method to isolate the target protein.

Purification using PFOA

Small hydrophobic proteins such as caveolin-1(122-142) were not able to be purified using the established inclusion body purification method. The extinction coefficient of the small peptide is very low and difficult to detect compared to the impurities. Although it looks pure in the SDS-PAGE electrophoresis gel, there are still traces of unwanted proteins that strongly absorb in the reverse phase HPLC at a wavelength of 280 nm. Therefore, a more effective purification step is required to isolate the desired protein. After over-expressing the caveolin-1(122-142) hydrophobic peptide, it needed to be solubilized. Because integral membrane proteins are usually very hydrophobic, they do not dissolve in aqueous buffers. Therefore, detergents are usually used to promote solubilization of membrane proteins. In many cases, common

detergents such as Triton X-100, NP-40, and Tween-20 are successful in dissolving hydrophobic proteins. On the other hand, integral membrane proteins are among the most hydrophobic and traditional detergents are not strong enough to solubilize these proteins. For that reason, the addition of 0.1% SDS with a chaotropic agent such as 8 M urea can be used as an alternative to reconstitute highly insoluble proteins. However, there are disadvantages associated with the use of SDS. For example, SDS is difficult to remove after purification and it is minimally compatible with nickel affinity chromatography. Therefore, using nickel affinity chromatography was of limited utility in purifying highly insoluble proteins, since these proteins are not soluble in 8 M urea or 6 M Gu-HCl, and they are only slightly soluble in low concentrations of SDS. Higher concentrations of SDS are not compatible with the nickel resin. Membrane proteins also dissolve particularly well in strong organic acids or highly fluorinated solvents like trifluoroethanol (TFE) or hexafluoroisopropanol (HFIP), but these solvents are not compatible with the nickel resin. The fluorinated detergent, pentadecafluorooctanoic acid (PFOA), is very effective at solubilizing extremely hydrophobic proteins at neutral pH. Also, this fluorinated compound is completely compatible with nickel affinity chromatography according to Bear et al (60). Since PFOA is compatible with nickel affinity chromatography, only two wash steps are required to semi-purify the inclusion bodies. Thus, we designed our purification system for small proteins with two inclusion body washes: wash 1 to remove the bulk of soluble proteins and wash 2 to remove the semi-hydrophobic membrane proteins that are still soluble in Triton x-100 detergent buffer. After two wash steps and nickel affinity chromatography, the isolated full-length fusion protein was obtained. Based on the results of the inclusion body preparation, the SDS-PAGE gel shows relatively pure material that has a MW (molecular weight) of approximately 25 kD. Therefore, the washes appear to very effective in removing many undesirable proteins from the fusion protein. Also, after the nickel column, the

elution of the fusion protein looks relatively pure (Figure 1-8). Our optimized construct was designed with eight histidines in front of Trp leader protein for nickel affinity purification steps with PFOA. Furthermore, we substituted an elution buffer containing imidazole at a pH of 8 for a low pH buffer, which we found to improve peak sharpness dramatically. Lastly, we found that the position of the His-tag affects the binding of the protein to the nickel affinity column. The N-terminal octa-histidine tag appears to provide optimal binding of our fusion protein to the nickel resin. Once target protein was eluted, PFOA was removed by extensive dialysis with buffer exchange every day.

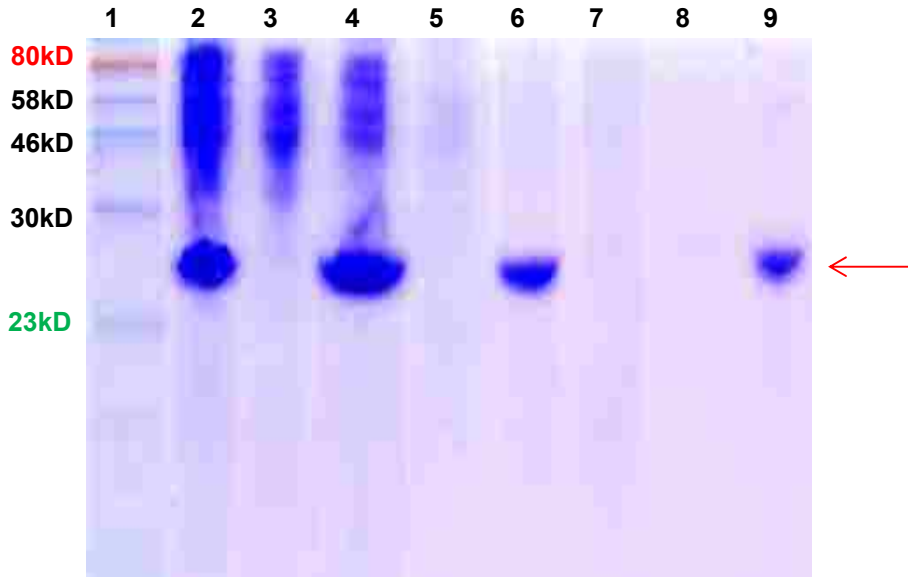


Figure 1-8. SDS-PAGE gel of Trp leader-ubiquitin-M-caveolin-1(122-142). Lane 1 is MW ladder; lane 2 and 3 is pre/post centrifugation in 20 mM Tris, 20% (w/v) sucrose, pH 8; lane 4 and 5 is pre/post centrifugation in 20 mM Tris, 1% Triton X-100, pH 8; lane 6 is post centrifugation in 25 mM phosphate, 8% PFOA, pH s8; lane 7 is flowthrough from Ni-NTA affinity column; lane 8 is wash with 25 mM phosphate, 1% PFOA, pH 8; and lane 9 is elution with 3 mM acetate, 1% PFOA, pH 4.

Cleavage method.

A variety of enzymatic and chemical cleavage methods were tested to figure out the best cleavage method. The enzymatic cleavage method such as TEV or Thrombin was not compatible with the detergent buffer or high concentration of urea buffer. The enzymes were denatured and were not able to the cleave target protein form the fusion protein. Thus we tried to exchange buffers that are optimized to enzyme use with an amicon ultra concentrator. However, the protein was precipitating out as soon as the buffers were exchanged. Therefore, the buffer exchange was performed while the protein was bound to Ni-NTA resin. But this method was not reliable, whenever we performed buffer exchange step lead to precipitation of protein. Although we can cleave the target protein, it precipitated out inside the resin and only the fusion partner was isolated. The target protein was precipitated inside the resin and we were never able to re-dissolve and isolate them. The membrane proteins are mostly hydrophobic residues which promotes aggregation without detergents. Therefore, we adopted an inclusion body purification method and tried other chemical cleavage methods. Hydroxylamine was the other choice instead of cyanogen bromide cleavage. The drawback of cyanogen bromide cleavage was it cleaves after methionine residues making the mutation of construct necessary if construct does not contain methionine. The hydroxyl amine cleavage method cleaves between Asn and Gly residues (61, 62). It is unlikely to have Asn-Gly sequence in the transmembrane domain since Asn is not very hydrophobic. The full length fusion protein was semi purified using an inclusion body purification method and then subjected to hydroxyl amine cleavage. The problem with this cleavage method was the buffer condition was basic and not able to solubilize proteins. Therefore we also adjusted buffer conditions by adding detergent into the buffer however the reaction did not proceed well with

detergents around. The yield of this cleavage method was less than 30% not enough for structural studies. Thus, we adopted the cyanogen bromide cleavage method which was able to get close to 100% completion of the cleavage products. Therefore, the cyanogen bromide cleavage method was performed to separate the target protein from the fusion partner. Because membrane proteins are particularly soluble in organic acids, 70%-88% formic acid was used as the reaction medium. Formic acid is well suited for the cyanogen bromide cleavage because it provides a highly acidic medium required for successful cleavage. Alternatively, trifluoroacetic acid (TFA) or hydrochloric acid (HCl) solution can also be used. However, formic acid proved to be the most effective at cleaving hydrophobic proteins from fusion proteins in large scale preparations. The drawback to using cyanogen bromide is that all native methionine residues in a membrane protein sequence must be mutated. Therefore, a hydrophobic protein that contains several methionine residues in its native sequence, may have to utilize an alternative cleavage method. However, other cleavage buffer conditions were not compatible with this hydrophobic protein. Moreover, the enzymetic cleavage method in PFOA buffer was inactive because PFOA denatures enzyme. The target proteins were effectively separated from the fusion partner using the cyanogen bromide method.

Separation and Verification

After cleaving the target protein from its fusion partner, the protein is separated using C4 reverse phase HPLC. Up to date, typically proteins are purified using a standard reverse-phase C18 column with a water/acetonitrile gradient with 0.1% trifluoroacetic acid. However, due to the hydrophobicity of our protein, we used a C4 column instead of longer carbon chain columns to prevent the hydrophobic protein from binding very tightly on the column. Moreover, the conventional buffer conditions

such as water / acetonitrile gradient to separate proteins was not able to be used in this case because those conditions were not strong enough to elute our target protein. Instead of the typical method, the gradient system from 80% water/20% acetic acid to 80% 1-butanol/20% acetic acid was used. The 20% acetic acid was critical to mix the two buffer and butanol buffer was strong enough to elute very hydrophobic proteins. Usually 1%/min gradient was used to isolate the target protein and elution is usually around 35% of buffer with butanol. To prevent aggregation of the target protein, it is critical to redissolve the protein in 88% formic acid immediately before loading it onto the column. The HPLC spectrum of the cyanogen bromide reaction shows three major peaks (Figure 1-9). The first peak is formic acid; the second large and somewhat broad peak represents the cleaved fusion partner portion of the fusion protein. The last peak in the spectrum represents the cleaved target protein. We expect the target protein to elute last due to the fact that it is more hydrophobic than the fusion partner. Moreover, nickel chromatography was critical to obtain pure protein for caveolin-1(122-142) based on HPLC trace comparison. Executing extra inclusion body washes was not helpful to the purity of the protein on the HPLC trace. Although SDS-PAGE gel looks clean after two inclusion body washes, it was unable to verify the protein peak in HPLC. Our results are confirmed by MALDI-TOF mass spectrometry (Figure 1-10). Most of the spectra were obtained using reflectron positive ion mode. For small proteins, such as caveolin-1(122-142), α-cyano-4-hydroxycinnamic acid was used as the matrix. A mono isotopic molecular weight of Caveolin-1(122-142) is small enough to obtain high resolution spectra. It was also possible to detect different isotopic peaks along with sodium and formate adducts of caveolin-1(122-142). The MALDI-TOF spectrum confirms that caveolin-1(122-142) was successfully separated from the fusion protein. For 4kD to 6kD proteins, sinapinic acid is used as the matrix and the

spectra were obtained with high resolution that was able to detect sodium and formate. This confirms that the target protein is successfully isolated. After verification by MALDI mass spectroscopy proteins are dried down with SpeedVac and redissolved into HFIP and lyophilized for further usage. HFIP is often used to deaggregate membrane proteins and is the only organic solvent that will dissolve extremely hydrophobic proteins and will lyophilize nicely.

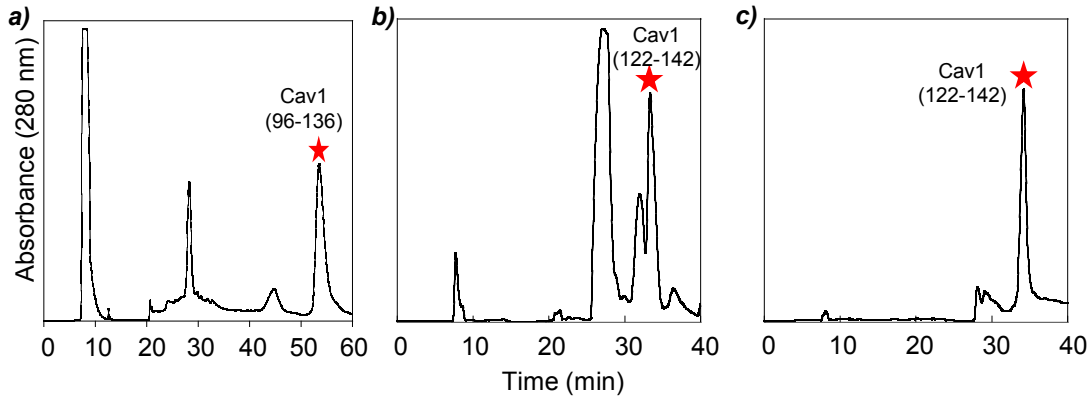


Figure 1-9. C4 reverse phase HPLC. Target protein is isolated from the fusion protein at flow rate 10 mL/min using buffer condition 20% acetic acid 80% water to 20% acetic acid 80% butanol. ***a)*** Purification of caveolin-1(96-136) using 1%/min gradient. ***b)*** First purification of caveolin-1(122-142) using 0.8%/min gradient. *c)* Second purification of caveolin-1(122-142) using 0.8%/min.

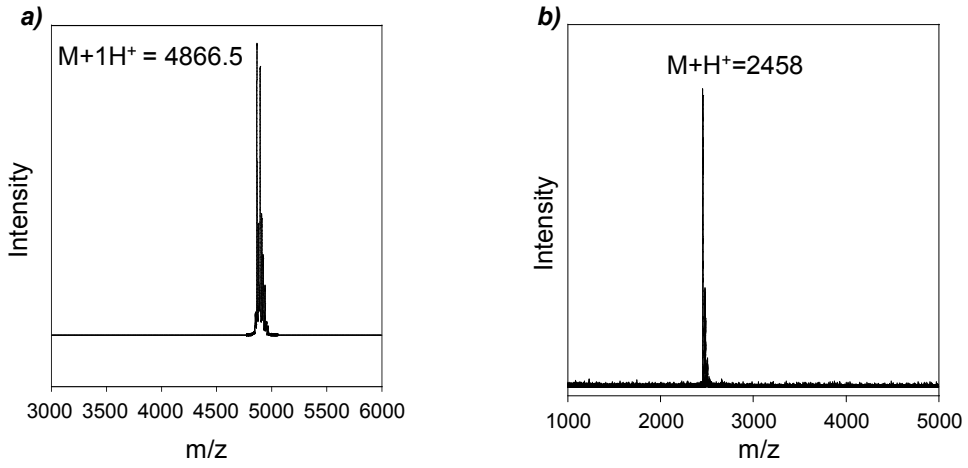


Figure 1-10. ***a)*** Mass spectrum of caveolin-1(96-136). ***b)*** Mass spectrum of caveolin-1(122-142).

Conclusion

Although a lot of challenges remain in the field of membrane proteins, our hope is that the method outlined here will serve as an aid to those who wish to obtain highly insoluble proteins for membrane protein studies. It is our belief that the Trp leader fusion expression system could be a universal method to produce high yields of small membrane proteins. Moreover, this method could potentially replace the highly expensive solid-phase protein synthesis method. Currently, the only drawback to the method at hand is that it relies on the cyanogen bromide method of cleavage, which requires that all methionine residues in the native protein sequence be mutated to avoid truncation of the desired protein. It may be desirable to investigate additional cleavage methods for proteins that contain native methionine residues in their primary structure.

Appendix 1-1. Design of oligomers

Caveolin-1(96-136)	
5’ oligomer	CGCGGATCCATGAAATACTGGTTCTACCGTCTGCTGTCTGC GCTGTTCGGTATCCCGCTGGCGCTGATCTGGGGTATCTAC TTCGC
3’ oligomer	CCGGAATTCTCAAGATTTGATAGACGGAACAACCGCCCAG ATGTGCAGGAAAGACAGGATCGCGAAGTAGATACCCCAGA TCAGCG
Caveolin-1(122-142)	
5’ oligomer	CGCGGATCCATGCTGTCTTTCCTGCACATCTG
3’ oligomer	CCGGAATTCTTACTGGATTTCGATCAGGAAAG
RfbP	
5’ oligomer	CGCGGATCCATGCGTTTCCTGAAACGTACCTTCGACATCGT TTCTTCTATCATCATCCTGATCATCGCGTCTCCGC
3’ oligomer	CCGGAATTCTCAACCGTCACGGGTAACTTTGTACCACAGGT AGATCAGCAGCGGAGACGCGATGATCAGGATGAT

Appendix 1-2. PCR cycle for the synthetic preparation.

1 cycle	95ºC	2 minutes
5 cycles	95ºC	15 seconds
	45ºC	15 seconds
	72ºC	1 minute
1 cycle	72ºC	5 minutes
	4ºC	∞

Appendix 1-3. Transformation Protocol (XL-1 blue)

1. Take 50 μL of cells from -80ºC freezer and thaw on ice
2. Put competent cells into pre-chilled 15 mL culture tubes
3. Add 1 μL of ligation reaction mixture
4. Keep in the ice for 30 minutes
5. Heat shock at 42ºC for 90 seconds
6. Keep in the ice for 2 minutes
7. Add 400 μL S.O.C.
8. Incubate at 37ºC for 1 hour at shaking speed 225 rpm.
9. Pour into LB plate with Kanamycin.
10. Incubate overnight at 37ºC

Appendix 1-4. Transformation Protocol (BL21(DE3))

1. Take 50 μL of cells from -80ºC freezer and thaw on ice
2. Put competent cells into pre-chilled 15 mL culture tubes
3. Add 20 ng of DNA
4. Keep in the ice for 30 minutes
5. Heat shock at 42ºC for 90 seconds
6. Keep in the ice for 2 minutes
7. Add 400 μL S.O.C.
8. Incubate at 37ºC for 1 hour at shaking speed 225 rpm.
9. Pour into LB plate with Kanamycin.
10. Incubate overnight at 37ºC

Appendix 1-5. Reagents for media preparation

ZYM-5052 (2 Liters)	N-5052 (2 Liters)
1916 mL H_2O 20 grams NZ-Amine AS 10 grams Yeast Extract 4mL $MgSO_4$ 400 µL 1000X Trace Metal 40 mL 50X5052 40 mL 50XM	14.196 grams Na_2HPO_4 13.609 grams KH_2PO_4 1.42 grams Na_2SO_4 1956 mL H_2O 5.45 grams $^{15}NH_4Cl$ 4mL $MgSO_4$ 400 µL 1000X Trace Metal 40 mL 50X 5052

Appendix 1-6. Reagents for starter preparation

MDAG (5 mL)
4.46 mL H_2O 10 µL $MgSO_4$ 1 µL 1000X Trace Metal 50 µL 25% Aspartate 100 µL 50XM 62.5 µL 40% Glucose 100 µL 17 Amino Acids 40 µL Methionine 5 µL Kanamycin (or appropriate antibiotic) A tip of stock freeze

Chapter 2. Caveolin-1 transmembrane domain

Abstract

Caveolin is an integral membrane protein that is found in high abundance in caveolae. Caveolin possesses an unusual horseshoe topology in the membrane which is thought to facilitate its many functions. In particular, a conserved proline residue (P110) in the caveolin transmembrane domain appears to be critical for proper orientation. To study the effects of this important residue, we have prepared a construct of the caveolin protein that encompasses the intact transmembrane domain (96-136). Caveolin (96-136) was over-expressed and isotopically labeled in *E.coli*, purified to homogeneity, and incorporated into lyso-myristoylphosphatidylglycerol micelles. Circular dichroism spectroscopy reveals that the secondary structure of caveolin (96-136) contains mostly α-helical character. Furthermore, NMR spectroscopy reveals that the transmembrane domain is composed of a helix-break-helix motif which allows caveolin to adopt it's unique conformation. Also, alanine scanning was carried out to probe the structural significance of residues 108-110 which create the helix-break in the transmembrane domain. Results indicate that position 108 requires small side chain amino acids, position 109 requires β-branched amino acids, and postion 110 requires proline. This research will ultimately lead to the elucidation of three-dimensional structure of caveolin on an atomic level.

Introduction

The most distinguishing feature of caveolin is that both the N- and C- termini are located on the same side of the plasma membrane (20, 23). This unusual topology is supported by a variety of *in vivo* studies which include immunofluoresence, immunoelectron microscopy, and epitope mapping (20, 23, 63, 64, 65). In addition, based on hydropathy prediction the transmembrane domain is approximately 33 residues (residues 102 to 135). The length of the transmembrane domain is unusual because it is too long to span the transmembrane domain once and too short to span twice. Combination of these studies have led to the postulation that the transmembrane domain forms a unique horseshoe structure in the membrane (turn inside of membrane).

Previous studies on caveolin-1 were carried out *in vivo*, but there are few biophysical studies on caveolin-1 particularly few on the transmembrane domain. Furthermore, there is no direct evidence to confirm the unique horseshoe conformation of caveolin-1 in the membrane at the molecular level. Without knowing the structure of the protein at the molecular level it is difficult to understand the role that caveolin-1 plays. Therefore, it is critical to determine the three-dimensional structure of caveolin-1 in lipid bilayers as this information will enhance future studies.

Interestingly, in the transmembrane domain of caveolin, two conserved proline residues (P110 and P132) appear to be critical for proper biological function. For example, a point mutation of proline at residue 132 to leucine (P132L) has been strongly correlated with breast cancer. In 16% of breast cancer cases, the P132L mutation was found in patients (27, 66, 67, 68, 69, 70, 71, 72, 73). Also, a proline to alanine mutation at position 110 (P110A) showed that the N-terminus of caveolin-1 was

exposed to the extracellular domain instead of facing the intracellular domain which could mean losing horseshoe conformation (26). Therefore, proline at position 110 (P110) in the transmembrane domain could be a candidate for the amino acid that is responsible for the turn in the horseshoe conformation. It is necessary to directly observe the structure of the caveolin transmembrane domain to investigate the role of P110.

The intact transmembrane domain of caveolin-1(96-136) was expressed and purified for biophysical studies. Pure protein was reconstituted into LMPG micellar solution to probe the structure by circular dichroism and NMR spectroscopy. Circular dichroism data demonstrate that caveolin-1(96-136) is highly helical and NMR data show that transmembrane domain exhibits helix-break-helix. These two conclusions imply that caveolini-1 transmembrane domain is likely responsible for the formation of the unique horseshoe conformation.

Mateial/Methods

Expression and purification

All mutant constructs were prepared using the modified QuikChange site-directed mutagenesis method (Appendix 2-1, Appendix 2-2). Reaction was performed with 50ng of vector which contains wild-type caveolin-1(96-136), 100ng of forward primer, 2.5uL of 9ºN ligase buffer, 2.5uL of 10X pfu buffer, 0.2uL of DMSO, 1.0uL of dNTPs, 1.0uL of pfu turbo, 1.0uL of 9ºN ligase, and bring the volume to 25uL with sterile water. The thermal cycler was set to the following cycling conditions: 1 cycle at 95 °C for 1 minute, 30 cycles at 95°C for 1 minute, 50 °C for 1 minute, 65 °C for 13 minutes, 4 °C for an indefinite period. Reaction products are digested with Dpn1 overnight at room temperature and transformed into XL-1 blue subcloning grade cells. After sequence analysis it is transformed into BL21(DE3) cells for protein expression.

The expression was performed with Trp leader fusion protein and purification was done using inclusion body purification method. The detail protocols are described in expression and purification chapter. For three-dimensional experiments, minimal ^{15}N, ^{13}C-labeling is required. To incorporate these isotopes into our protein, a cell growth was performed in M9 minimal media using a procedure that was optimized by Marley et al (58). This protocol uses ^{13}C glucose and $^{15}NH_4Cl$ as the only carbon and nitrogen source to label both of them. For specific amino acid labeling, the cell growth was done using M9 minimal media with 5 fold excess of ^{14}N 19 amino acids related to the desired ^{15}N amino acid (59). The growth was induced with 1 mM IPTG at O.D. 0.6 and grown for 6- 8 hours. After harvesting cell same purification was done to obtain NMR sample.

Sample preparation

To prepare a caveolin-1(96-136) NMR sample, various detergents were screened to determine which detergent was compatible with our sample and which detergent gave the best resolution in the NMR. Some detergents were not strong enough to solublize caveolin-1(96-136) while others, particularly ionic detergents, were too strong and can denature caveolin-1(96-136). For our studies we found that LMPG (*lyso*myristoylphosphatidylglycerol) was the only detergent that was both powerful enough to solubilize the hydrophobic caveolin-1(96-136) and also gave a well-dispersed HSQC spectrum. The LMPG micellar solution was prepared with 100 mM LMPG, 100 mM NaCl, 20 mM phosphate, pH 7. The NMR sample was prepared by dissolving 2.5mg ^{15}N labeled caveolin-1(96-136) in approximately 500 µL of the LMPG micellar solution. Vigorous vortexing and a brief heat shock were applied for effective sample preparation. The peptide concentration was verified by Biospec spectrometer at 280 nm and an extinction coefficient of ε=20970 $M^{-1}cm^{-1}$.

The sample used for circular dichroism studies were prepared in a similar manner to the NMR sample except that the protein concentration was approximately 0.05 mM. Therefore, the CD sample was obtained by diluting the NMR sample with the LMPG micellar solution after the NMR experiments were carried out.

CD and NMR

All circular dichroism experiments were performed using a JASCO CD Spectrophotometer (Easton, MD). The experiments were carried out at 25℃. Spectra were obtained using step mode in the wavelength range of 260 to 190 nm using 16 accumulations. A background spectrum was obtained using the LMPG micellar solution containing no caveolin protein and was subsequently subtracted from the spectrum of the protein sample. In all cases a 0.1 mm path length cuvette was used. The sample volume was approximately 250 µL. The data were processed using web-based computer program, Dichroweb (45, 46, 47).

The HSQC spectrum of caveolin-1(96-136) in LMPG micelles was acquired at 37°C using the TROSY-HSQC pulse sequence on a 600 MHz Bruker (Billerica, MA) Advance II spectrometer equipped with a cryoprobe. The instrument is located at Penn State Hershey Medical College in Hershey, PA. The spectrum was acquired with 256 (^{15}N dimension) × 2048 (^{1}H dimension) complex points using a total of 16 scans. The data were Fourier transformed using the NMRPipe program (48). After the HSQC spectrum was acquired, a series of other experiments were carried out including three-dimensional experiments such as HNCA and HN(CO)CA to obtain detailed information about the backbone structure of caveolin-1(96-136). To run the three-dimensional experiments ^{15}N and ^{13}C labeling is required. Specific amino acid ^{15}N labeling is also used to assist in assigning the resonances of the peptide backbone. Using specific amino acid labeling the HSQC spectrum shows only peaks that are labeled with ^{15}N amino acids. Overlaying the specific amino acid-labeled spectrum with the uniformly ^{15}N-labeled spectrum allows us to confirm the identity of the peaks in the HSQC spectrum. This information is used to analyze and check the backbone assignment made from the HNCA and HN(CO)CA experiments. The spectra were processed

using NMRPipe and Sparky (45, 46, 47, 74). Secondary structure information was obtained using Cα chemical shifts as described in Wishart et al (75). Dihedral angles for wild-type were obtained using the computer program TALOS+ and H, N, CO, Cα, and Cβ chemical shift data were used (49). After TALOS+ preliminary structures were built using the computer program CS-ROSETTA which build structure based on chemical shifts and side chain packing (76).

Results and Discussion

CD

The secondary structure of caveolin-1(96-136) was revealed using circular dichroism spectroscopy in LMPG detergent micelles. The LMPG micellar solution was chosen for consistency with other experiments because the LMPG offered the highest quality NMR spectrum. The data obtained from the circular dichroism experiments was processed using a web-based computer program called Dichroweb, which determined the percent α-helicity of caveolin-1(96-136) (45, 46, 47). The results of the Dichroweb analysis show that there are two minima at 208 and 222 nm which is indicative of α-helical properties. Analysis of the spectrum using the CDSSTR and K2D algorithms (Dichroweb) indicates that α-helicity of caveolin-1(96-136) is greater than 57%. Also the percentage of turns in the peptide is determined to comprise approximately 10% of the peptide structure, which means that 3-4 amino acid residues could be found in turn regions of the structure. However, mutation of proline 110 to alanine results in a circular dichroism spectrum that is remarkably similar to that of the wild-type. Therefore, mutation of proline 110 to alanine does not dramatically alter the secondary structure of Caveolin-1(96-136). Also other mutations G108A and I109A show similar secondary structure to that of the wild-type. (Figure 2-1).

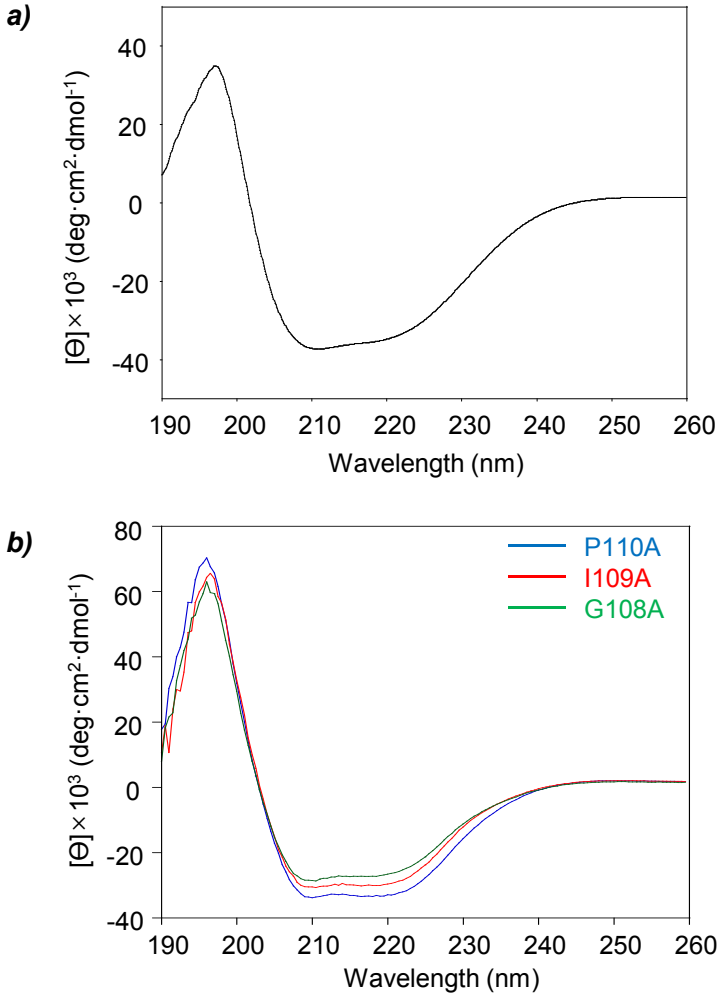


Figure 2-1. ***a)*** Circular dichroism spectroscopy of caveolin-1(96-136). Using programs in Dichroweb revealed mostly α-helical character in caveolin-1(96-136). ***b)*** Circular dichroism spectroscopy of caveolin-1(96-136) mutants; P110A (blue), I109A (red), and G108A (green).

NMR backbone assignment

The wide dispersion of the HSQC spectrum in LMPG micelles suggests that each amino acid is surrounded by a unique environment. The caveolin-1(96-136) HSQC peak dispersion is greater compared to a typical single-spanning transmembrane alpha helix, which is not surprising because the transmembrane domain of caveolin has such an unusual topology. After obtaining a high quality HSQC spectrum, backbone assignment is necessary to get structural information. In order to assign the backbone of caveolin-1(96-136) ^{13}C- and ^{15}N- labeling was required. For initial experiments, ^{13}C-, ^{15}N- labeled caveolin-1(96-136) was reconstituted into LMPG micelles, and an HSQC spectrum was obtained to check the signal-to-noise ratio and the intensity of peaks (Figure 2-2). This step is necessary to determine if signals were intense enough for a series of three-dimensional experiments. For the three-dimensional experiments, the HNCA experiment was executed first because it is one of the most sensitive the three-dimensional experiments. The HNCA experiments provide both intra- and inter- molecular Cα information from which the backbone of caveolin-1(96-136) can be identified. Intra-molecular Cα gives a strong signal while inter-molecular Cα gives a weak signal. By matching the strong peak of the starting peak to the weak peak of the other peak this approach allows us to assign the backbone of the protein (Figure 2-3). Unfortunately, chemical shift degeneracy can lead to ambiguous backbone assignments. Therefore, it is desirable to use additional experiments to aid backbone assignment. We chose to specifically label each amino acid using the ^{15}N isotope. In the caveolin-1(96-136) protein we chose to specifically label leucine, isoleucine, alanine, phenylalanine, and valine each of which was carried out in five separate preparations of the protein. These amino acids were chosen because leucine and isoleucine are the most abundant residues in the peptide which would give us the

most information. Also, alanine and phenylalanine are well-dispersed throughout the peptide based on primary sequence analysis. Labeling specific amino acid helps to provide a starting point when assigning the peaks in the HSQC spectrum (Figure 2-4, Figure 2-5, Figure 2-6). Initially, ^{15}N labeled glycine was used as a starting point in assigning chemical shifts because it gives a unique chemical shift in the HSQC spectrum compared to the other amino acids. Therefore, without specific labeling glycine would be the only known starting points which would make backbone assignment very difficult. Also knowing which peak represents which amino acid, it is very useful to verify that the direction of the backbone assignment is correct. Also HN(CO)CA experiment was performed to help backbone assignment. This experiment only provides inter Cα values which can be compared to HNCA experiment and help identifying inter and intra Cα peaks. Moreover, HNCO and HNCACB experiments are executed to obtain Cβ and CO information. The HNCACB information helps identifying threonine, serine and alanine because chemical shifts of Cβ are unique compare to other amino acids. The HNCO experiment is one of the most sensitive three dimensional experiments along with HNCA. These experiments are helpful to verify ambiguous peaks.

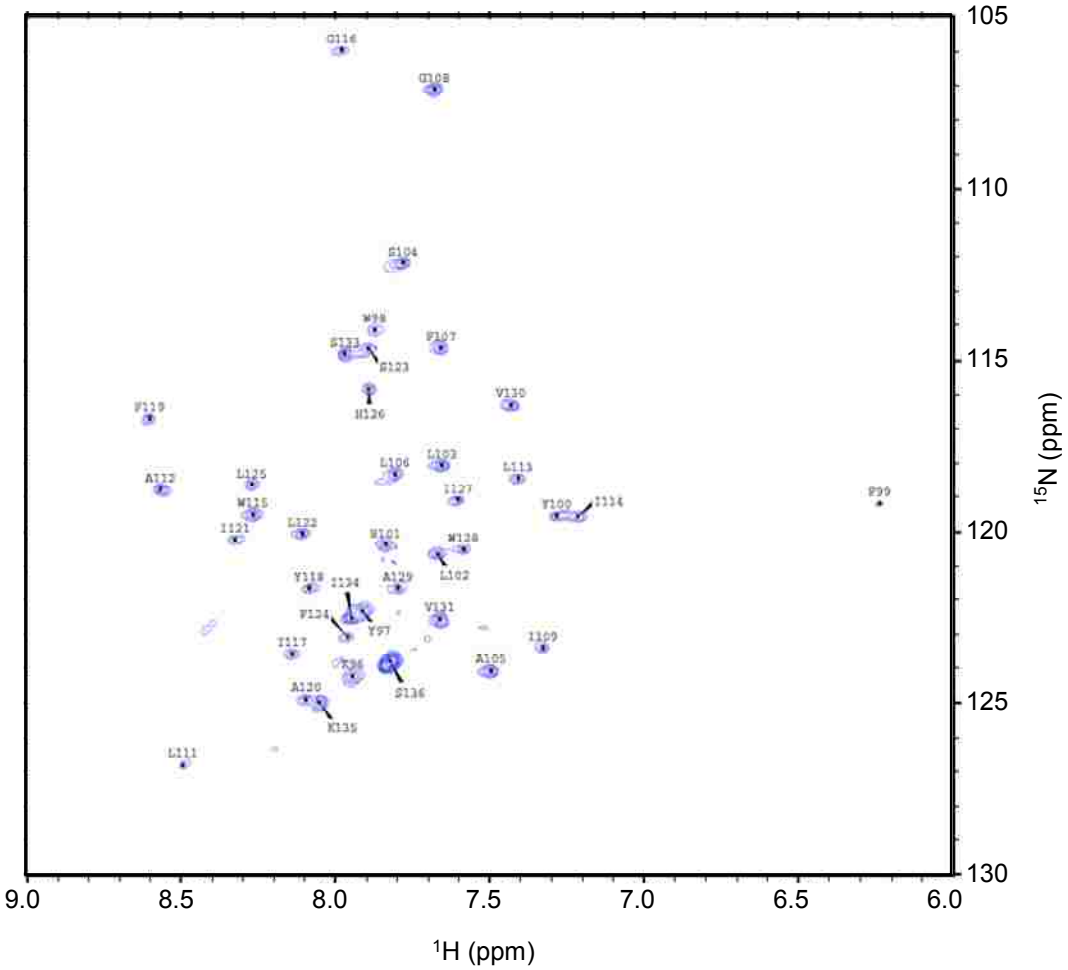


Figure 2-2. ^{1}H-^{15}N TROSY-HSQC spectrum of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136) in 100 mM LMPG, 100 mM NaCl, 20mM Phosphate pH 7. The spectrum was recorded at 37ºC on a 600 MHz spectrometer with 256 complex points in t1, 2048 complex points in t2.

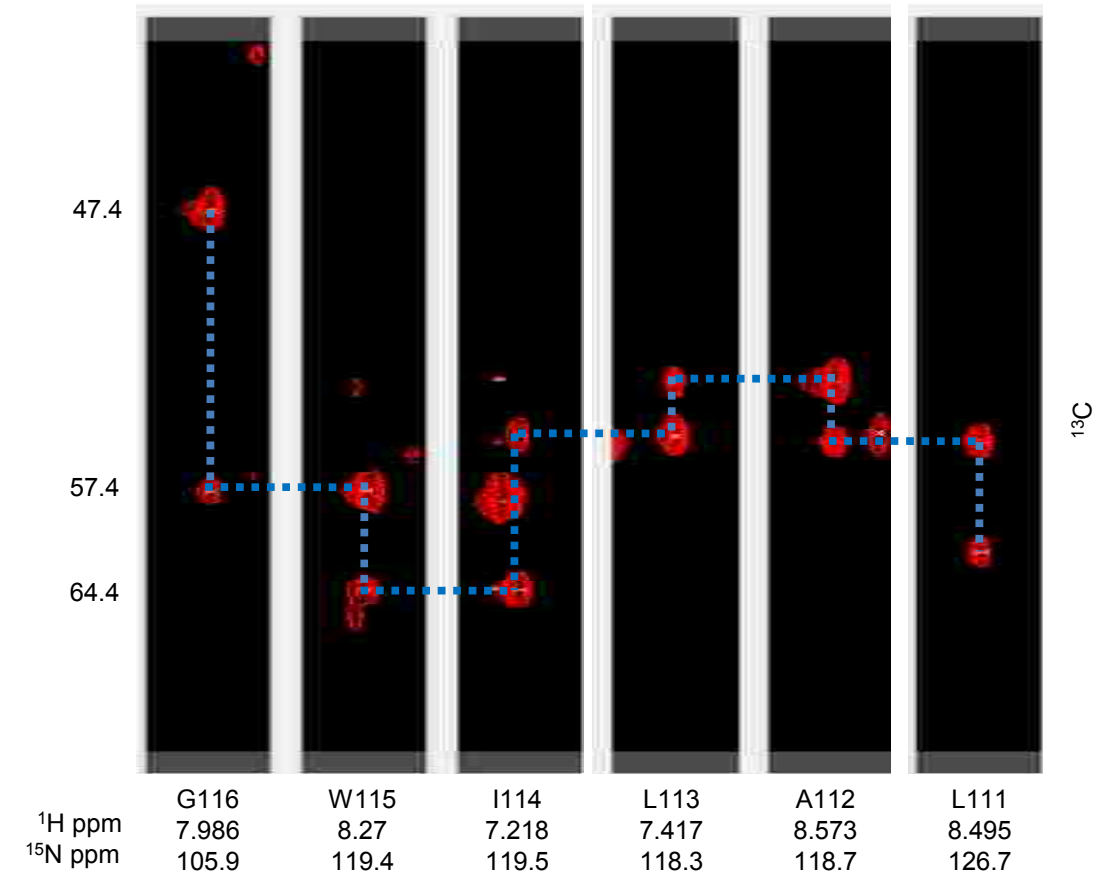


Figure 2-3. Walk through of backbone assignment of caveolin-1(96-136) using strips of HNCA experiment.

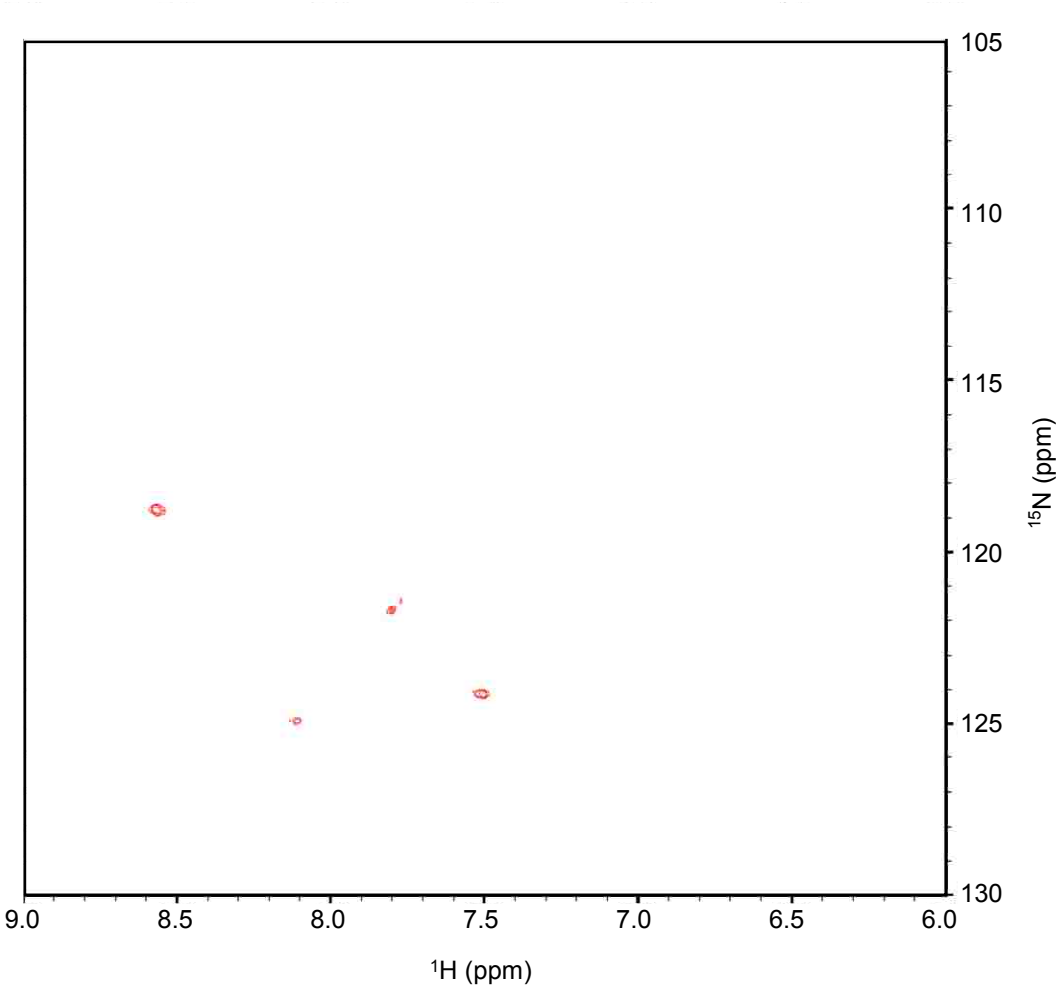


Figure 2-4. ^{1}H-^{15}N TROSY-HSQC spectrum of alaine specific ^{15}N-labeled caveolin-1(96-136) in 100 mM LMPG, 100 mM NaCl, 20mM Phosphate pH 7. The spectrum was recorded at 37ºC on a 600 MHz spectrometer with 256 complex points in t1, 2048 complex points in t2.

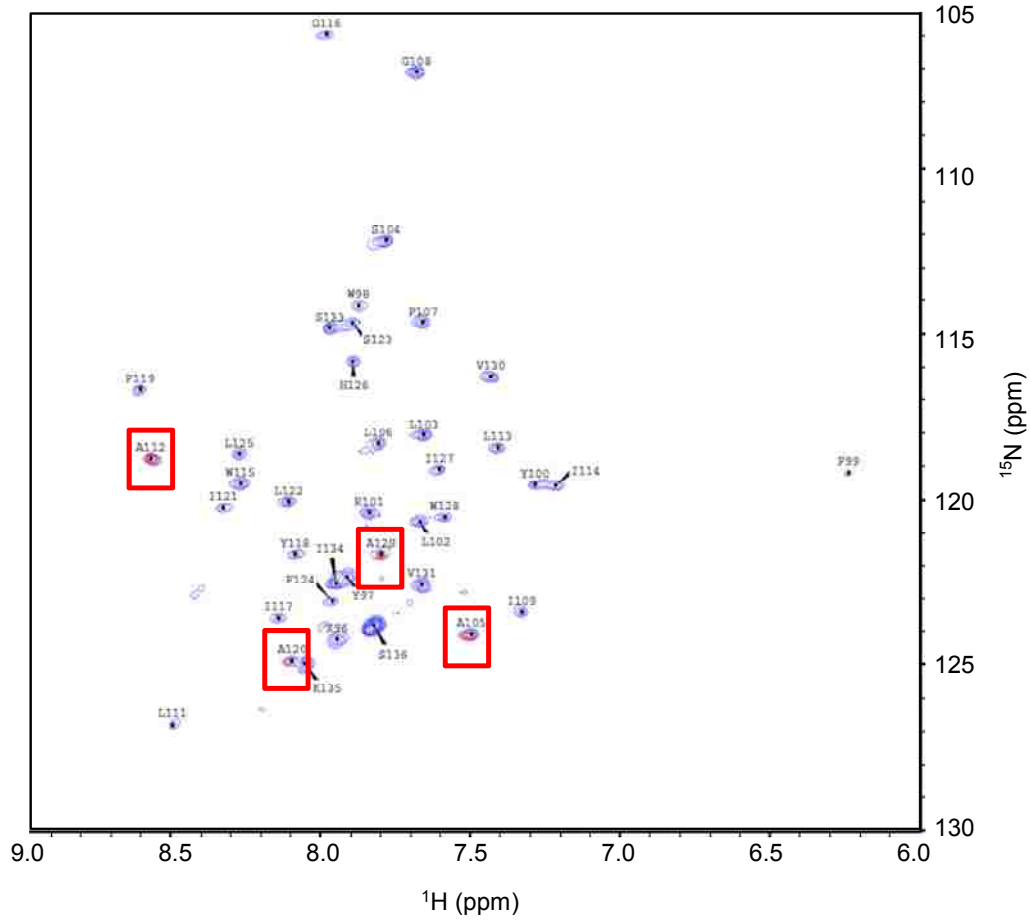


Figure 2-5. Overlay of ^{1}H-^{15}N TROSY-HSQC spectrum of alaine specific ^{15}N-labeled caveolin-1(96-136) (Red) and uniformly ^{15}N-labeled Caveolin-1(96-136) (Blue).

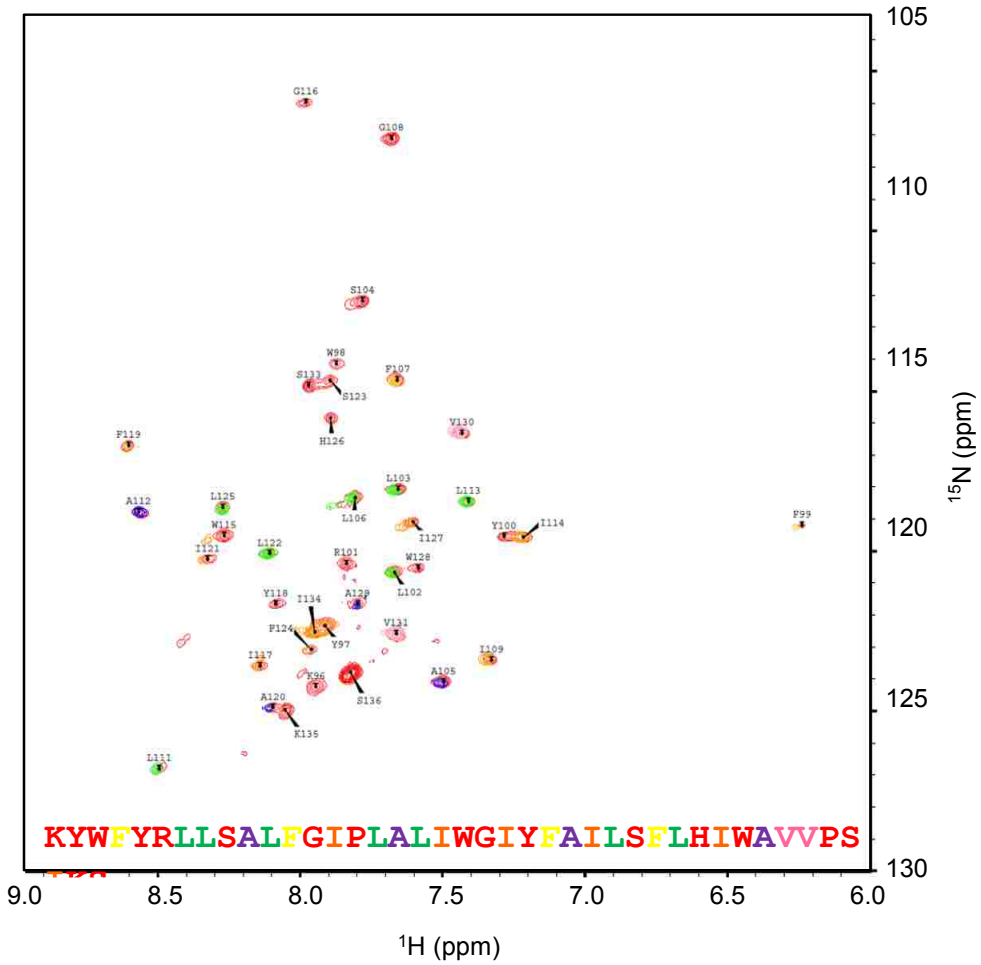


Figure 2-6. Overlay of ^{1}H-^{15}N TROSY-HSQC spectrum uniformly ^{15}N-labeled caveolin-1(96-136) with specific ^{15}N-labeled caveolin-1(96-136); Phe (yellow), Leu (green), Ala (purple), Ile (orange), Val (pink).

Analysis of CSI

After obtaining Cα information of all the peaks, their chemical shifts were indexed (CSI). The chemical shift index analyzes the differences between the chemical shifts of the Cα of each amino acid in the structure of the protein versus the chemical shift given by the same amino acid from a random coil structure. If the ΔCα is positive, it represents α-helica character and if ΔCα is negative it represents β-sheet character. But for both cases, a series of ΔCα has to be consecutive to represent secondary structure. In other words, if two or three amino acids ΔCα is positive or negative it does not represent secondary structure. After processing all of the residues, a helix-break-helix motif was found in the transmembrane domain of caveolin-1 (Figure 2-7). The first helix starts at Tyr97 and ends at Phe107 because the ΔCα's are all positive. The second helix starts at Leu111 and ends at Ala129 because the ΔCα's are positive. The breakage between these two helices spans Gly108 to Pro110 because ΔCα is negative but it is not enough to say there is a β-sheet. The remaining amino acid residues are situated toward each of the ends of the peptide and they are not well structured. The C-terminal region is a dynamic portion of this protein. Therefore, Gly108, Ile109, and Pro110 are potential turn point for the unusual horseshoe structure because these amino acids are located at a critical place in the transmembrane domain and create a break in the helical structure. This is the first experimental data presented of a helix-break-helix motif in the caveolin-1 transmembrane domain. Moreover, potential candidates for the critical amino acids responsible the for horseshoe conformation were revealed. This data is verified by the TALOS+ computer program which is a program that can estimate torsion angles based on chemical shifts (49). This program was used to obtain the torsion angle phi (φ) and psi (ψ) of caveolin-1(96-136), both of which are necessary for the three-dimensional structure of caveolin-

1(96-136). When the N, H, CO, Cα, and Cβ chemical shifts for the caveolin-1(96-136) were entered into the computer program TALOS+, residues Y97–V131 were confidently predicted to have dihedral angles and residues P132–K135 were predicted to be dynamic. This information can then be entered into a program called CS-Rosetta (76, 77). This program is a powerful tool that is used to build a protein structure using all of the detailed information provided by the experiments performed herein. CS-Rosetta, the program generates a series of possible structural models all of which satisfy the torsion angles specified, the proper bond length and van der Waals contact of the amino acid residues, and lowest energy states as it relates to the side chain packing of the residues. The data generated from CS-Rosetta can be subsequently visualized using XPLOR (NIH) where further refinements to the structure can also be made (Figure 2-8) (78, 79).

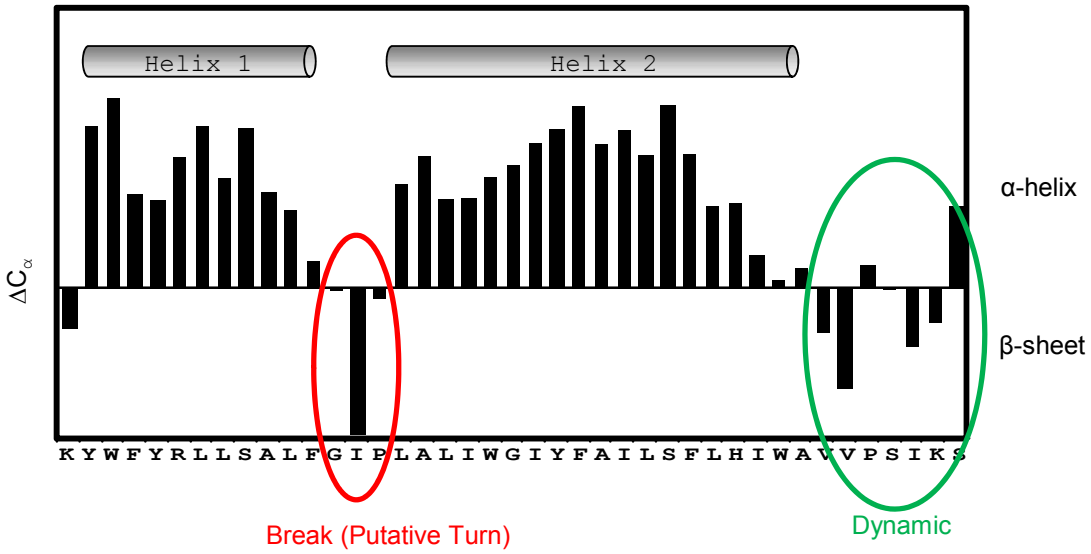


Figure 2-7. Chemical shift indexing of caveolin-1(96-136). Caveolin-1(96-136) displays a helix–break–helix structure. Red highlight denotes breakage region and green highlights denotes dynamic region.

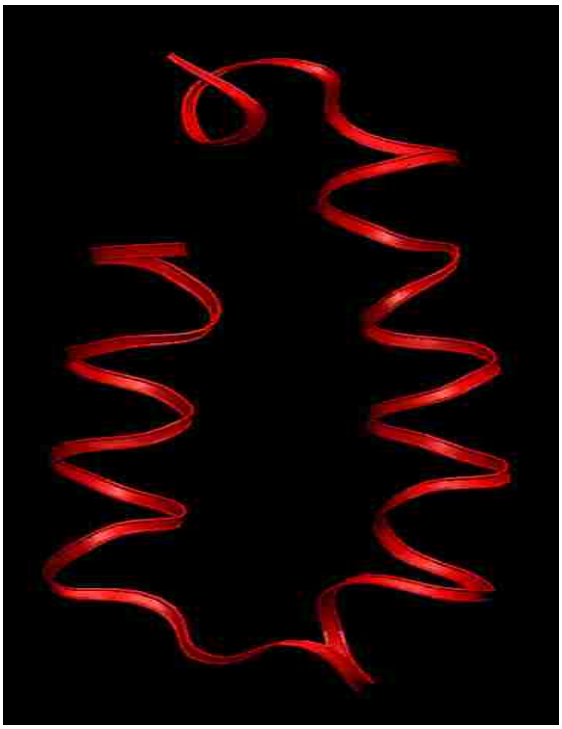


Figure 2-8. Model of Caveolin-1(96-136) created by CS-ROSETTA and visualized by X-PLOR.

Alanine Scanning

Alanine scanning is often carried out to determine the contribution of a specific amino acid to the protein structure and function. Alanine is usually selected because it has a relatively small and hence neutral side chain. Three residues are of particular interest; G108, I109, and P110, which we hypothesize to be responsible for the break in the structure of caveolin-1 transmembrane domain, which was revealed by CSI. Therefore, mutations to alanine were introduced at position 108, 109, and 110 in separate protein preparations and the effect on the NMR structure of these mutations was assessed. Moreover, F107 was also mutated to alanine and investigated since the dihedral angle created by TALOS+ is not exactly an α-helix dihedral angle. The alanine mutations were achieved by using a modified Quikchange site-directed mutagenesis protocol (Agilent Tech) and a large-scale growth was carried out following the procedure outlined in the previous chapter. Overlaying the F107A HSQC spectrum with the wild-type HSQC spectrum shows that most of peaks overlap (Figure 2-9). This suggests that the structure of the F107 mutation is similar to the wild-type. Also overlaying the G108A HSQC spectrum with the wild-type HSQC spectrum shows that most of the peaks overlap but the residues surrounding position 108 are slightly shifted (Figure 2-10, Figure 2-11). For example, residues surrounding G108 (F107 and I109) had slightly shifted chemical shifts but the V130 which is far away from G108 remained unchanged. The G108A mutation created small changes around residue 108 but overall the conformation was same. This result suggests that even though glycine has a small side chain (-H) this site can tolerate the bulkiness of the methyl group side chain introduced by the alanine. On the other hand, the I109A and P110A HSQC spectra looked drastically different compared to the wild-type spectrum. The I109A and P110A mutation appeared to change the overall conformation (Figure 2-12, Figure

2-13). Therefore, a β-branched bulky side chain like that of isoleucine appears to be critical at position 109. It may be that the alanine side chain is too small to maintain the structure because of loss of steric hindrance. Also, P110A mutation was not able to keep native conformation (Figure 2-14, Figure 2-15). Proline is unique due to its conformational rigidity. It is not surprising that by breaking this rigidity by inserting an alanine at the P110 position results in a loss of the native structure. Further analysis of the HSQC spectrum suggests two possible explanations, one is the loss of the native conformation and the other is that aggregation may be occurring. Aggregation is not an event that occurs randomly in a well-structured peptide. For aggregation to take place, a loss in native conformation must occur first. Therefore, it is more likely that the I109A and P110A mutations cause a conformational change that results in a loss of the horseshoe structure thereby promoting aggregation. To probe the details surrounding this phenomenon, specific point mutations were carried out at these three positions.

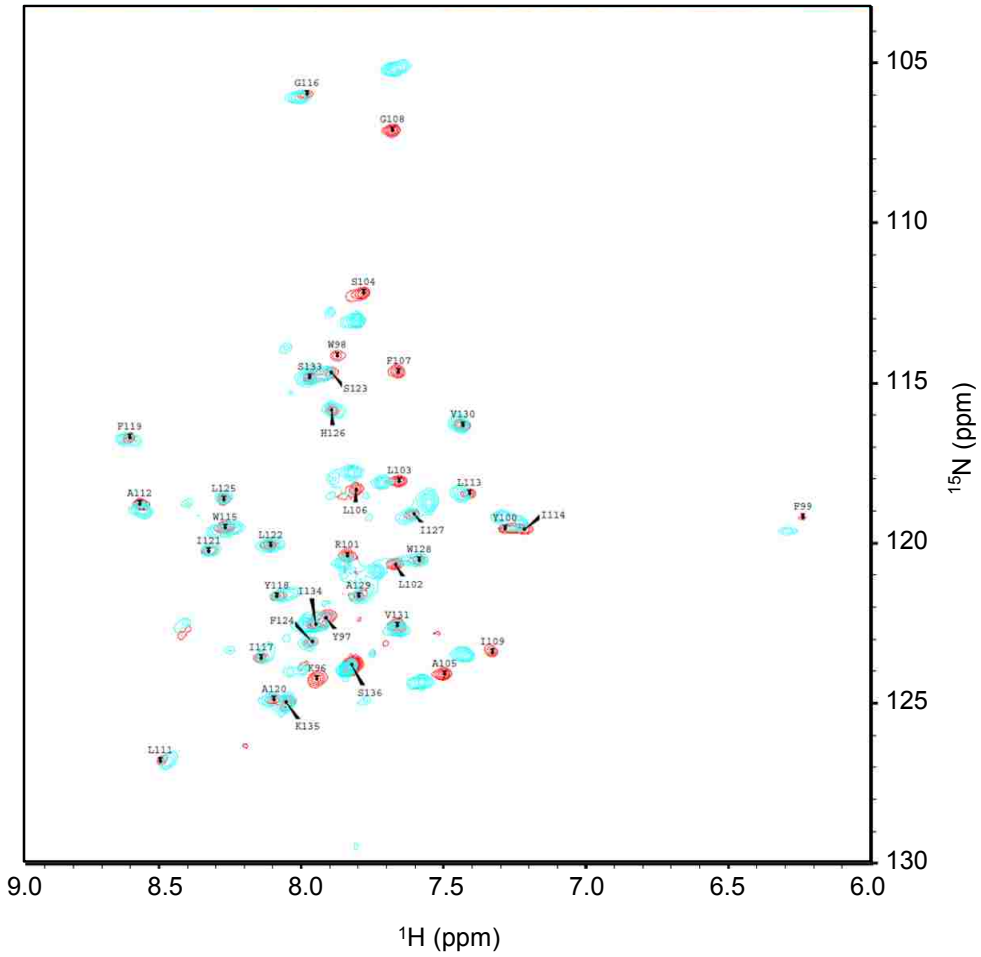


Figure 2-9. Overlay of ^{1}H-^{15}N TROSY-HSQC spectra of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136) wild-type (red) and caveolin-1(96-136)F107A mutant (blue).

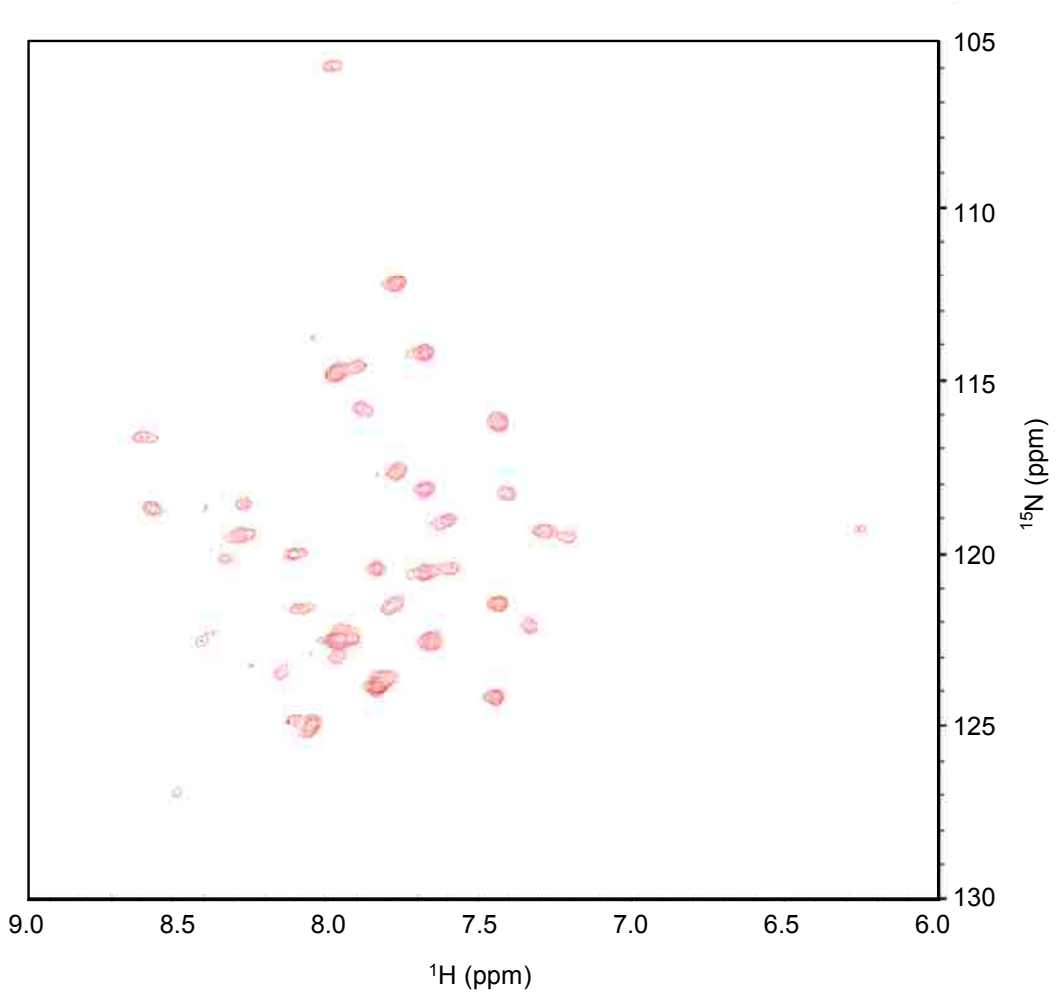


Figure 2-10. ^{1}H-^{15}N TROSY-HSQC spectrum of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136)G108A mutant.

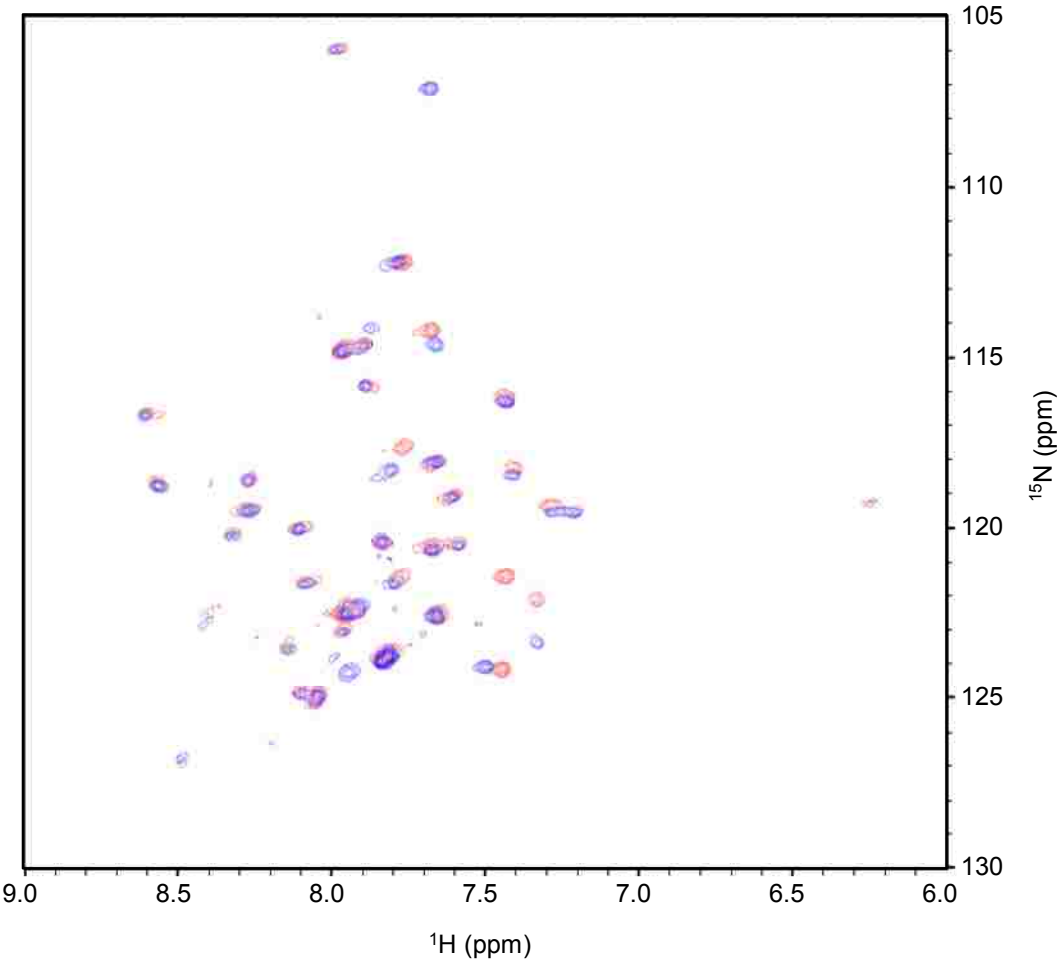


Figure 2-11. Overlay of ^{1}H-^{15}N TROSY-HSQC spectra of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136) wild-type (blue) and caveolin-1(96-136)G108A mutant (red).

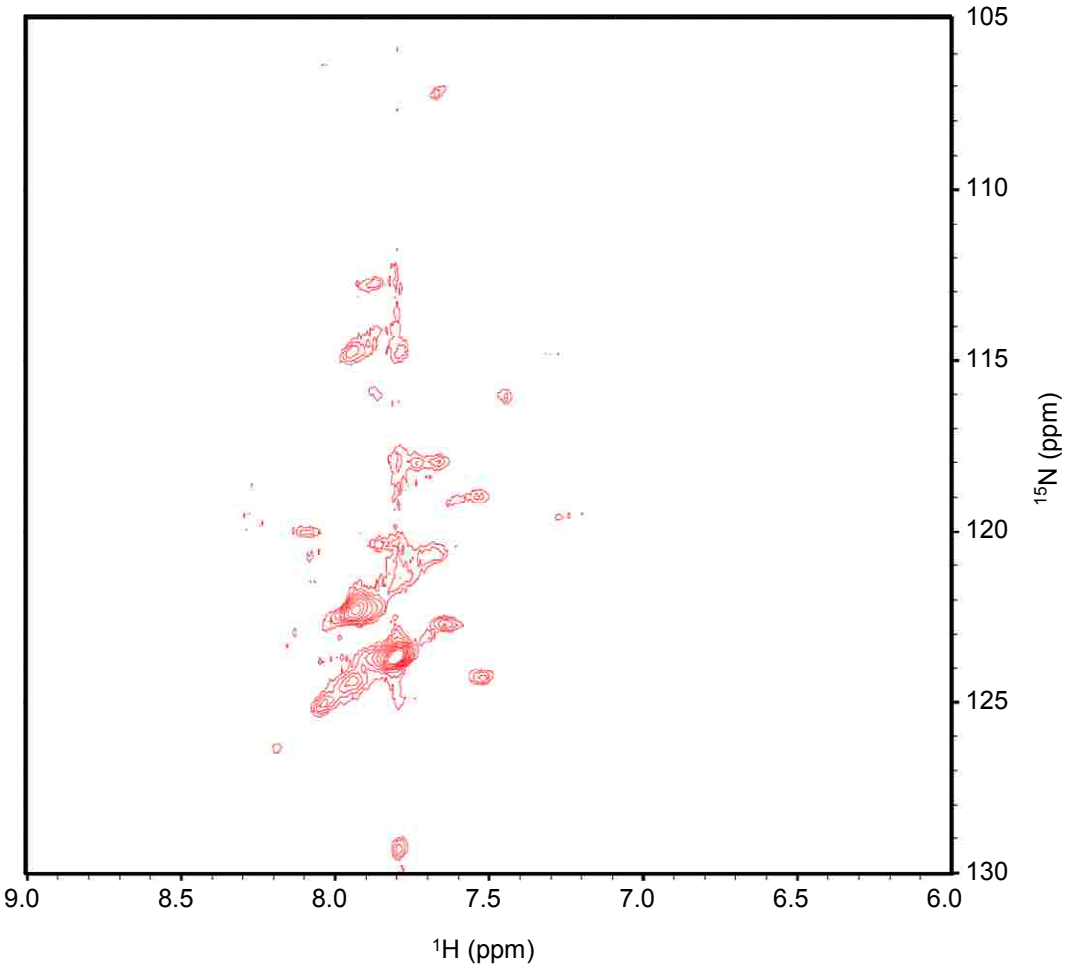


Figure 2-12. ^{1}H-^{15}N TROSY-HSQC spectrum of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136)I109A mutant.

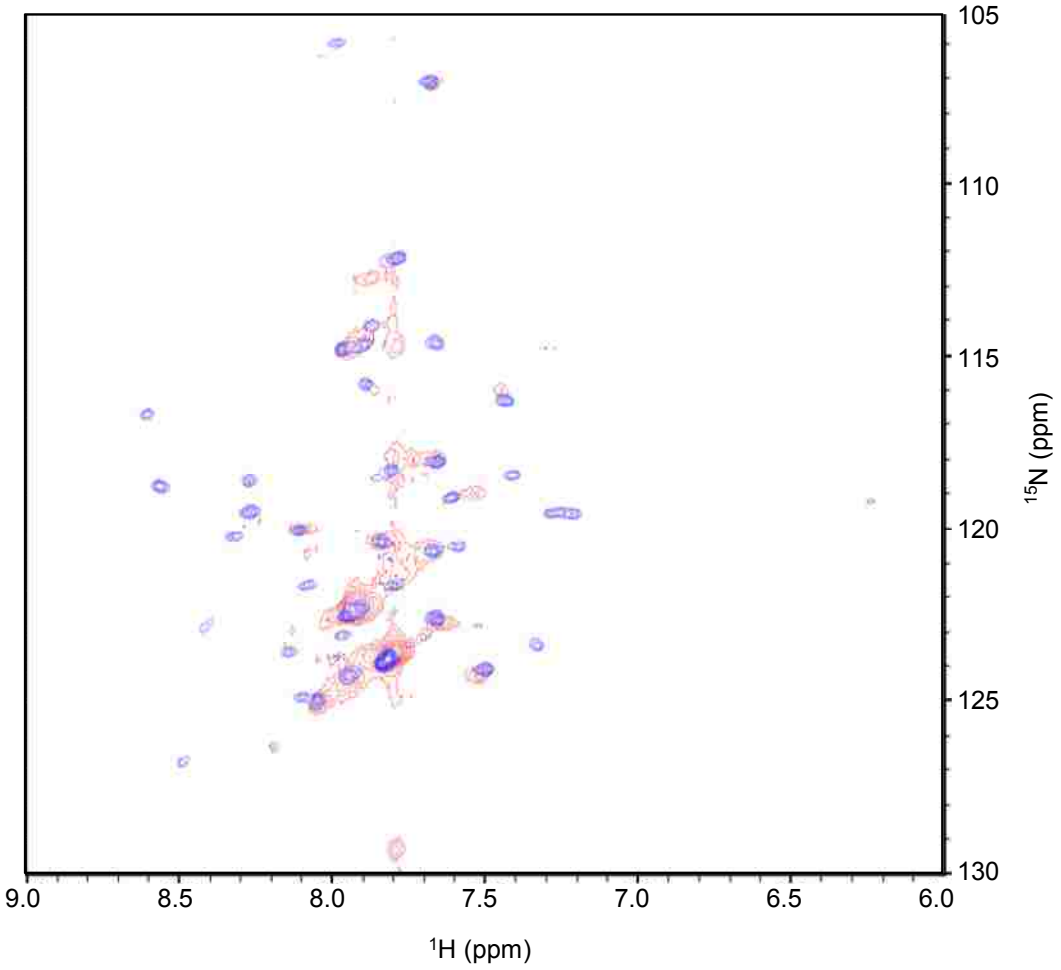


Figure 2-13. Overlay of ^{1}H-^{15}N TROSY-HSQC spectra of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136) wild-type (blue) and caveolin-1(96-136)I109A mutant (red).

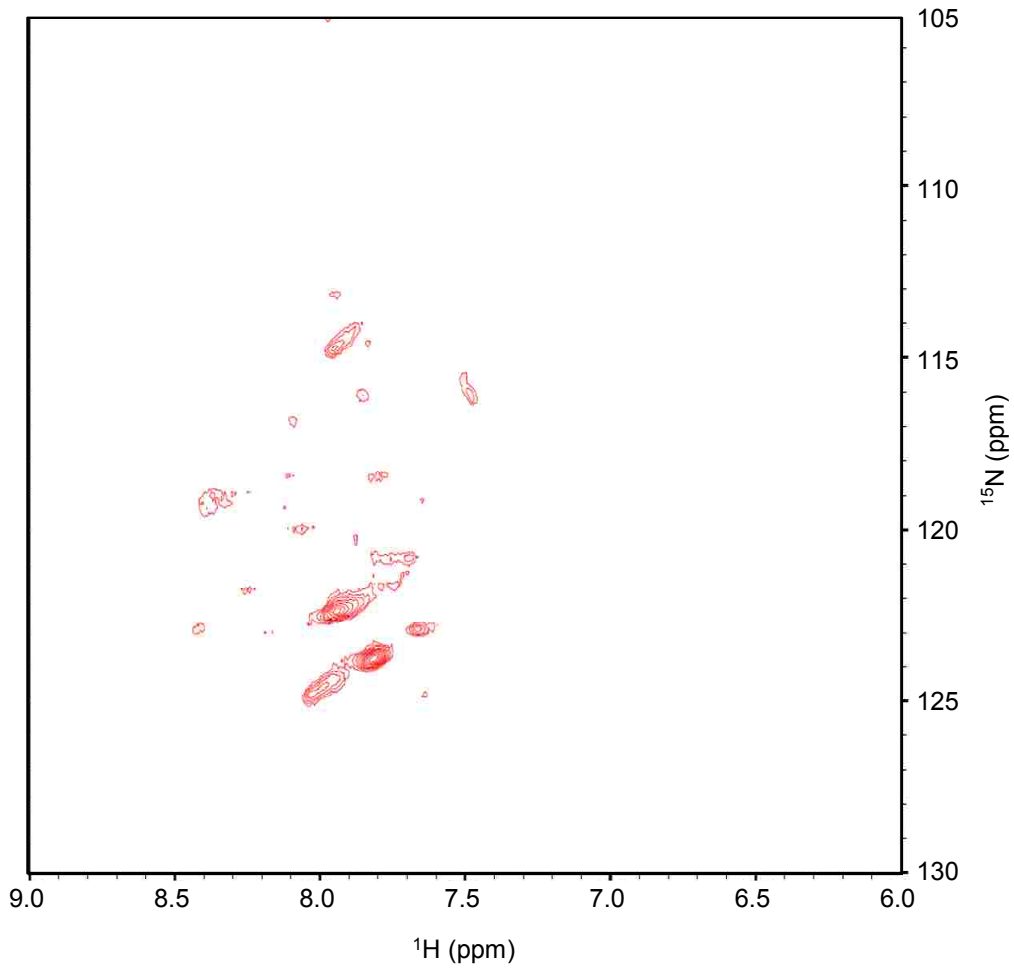


Figure 2-14. ^{1}H-^{15}N TROSY-HSQC spectrum of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136)P110A mutant.

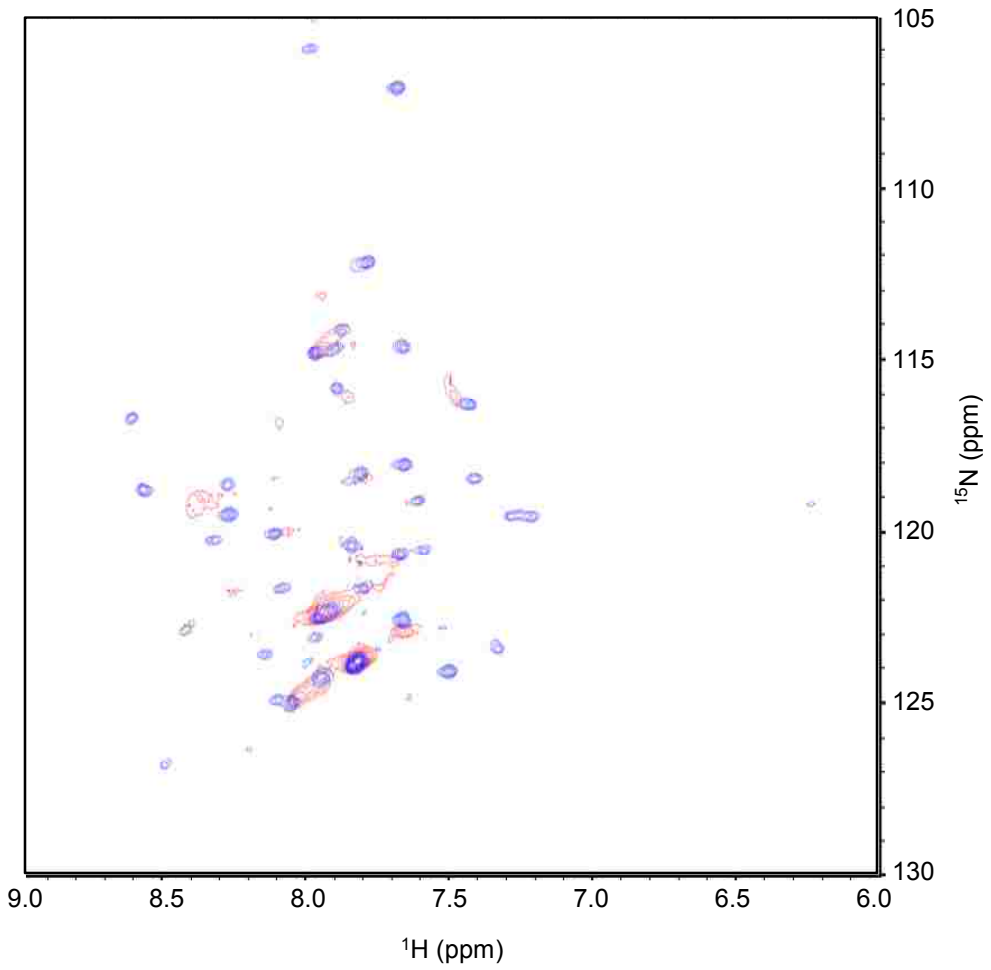


Figure 2-15. Overlay of ^{1}H-^{15}N TROSY-HSQC spectra of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136) wild-type (blue) and caveolin-1(96-136)P110A mutant (red).

Point Mutation

To probe the nature of the side chains at positions 108-110 specific point mutations were introduced similar to the alanine scanning approach. The glycine at position 108 (G108) was mutated to leucine and isoleucine to observe the limits of the side chain bulkiness and also to investigate the β-branched side chain effects. The isoleucine at position 109 (I109) was mutated to leucine and valine to determine if the β-branched side chain is critical at position 109 or if simply the bulky side chain is necessary. The proline at position 110 (P110) was mutated to glycine because glycine has the most conformational freedom. Glycine has the ability to adopt a similar angle to that of proline, which would suggest that the peptide would still be able to maintain its native conformation.

The overlay of the G108L and G108I to wild-type spectrum show that both result in a loss of conformation (Figure 2-16, Figure 2-17). Therefore, at position 108 leucine and isoleucine are not tolerated. The limits of the bulkiness of the amino acid side chain revealed that only a proton or a methyl group are tolerated. Therefore at position 108, glycine and alanine are the only two amino acids that are tolerated.

At position 109, leucine appeared to be meta-stable while valine was stable (Figure 2-18, Figure 2-19, Figure 2-20). This was evident by the fact that both I109L and I109V gave reasonable spectra but over time the leucine HSQC spectrum resolution was lost (Figure 2-19). The I109L spectrum was well overlapped with the wild-type spectrum but several new peaks were observed which was indicative of meta-stable state. The I109V spectrum was very similar to the G108A spectrum with a small shift around position 109, but the overall spectrum was unchanged. Therefore, it appears that it is essential that a β-branched side chain be present to preserve the

structure and it is not simply a matter of introducing steric bulk at this position to stabilize the structure. Beta-branched side chain amino acids are required at position 109 to sustain the structure because a beta-branch is closer to backbone of the protein which creates more rigidity that is required to maintain unique putative turn conformation.

At position 110, glycine was not able to maintain the structure (Figure 2-21). Specific peaks in the spectrum were not distinguishable. The P110G spectrum was similar to the P110A spectrum, which suggests a loss in conformation. Although glycine has the conformational freedom to adopt a similar angle to that of proline, it was not stable enough to maintain that angle. Therefore, glycine may be too flexible to maintain one conformation which means that proline is most likely the only amino acid that can be tolerated at position 110.

The conclusions of the point mutation study are that position 108 requires a small side chain amino acid residue such as glycine and alanine, β-branched amino acids like isoleucine and valine are required at position 109, and proline is likely the only amino acid that can be at position 110.

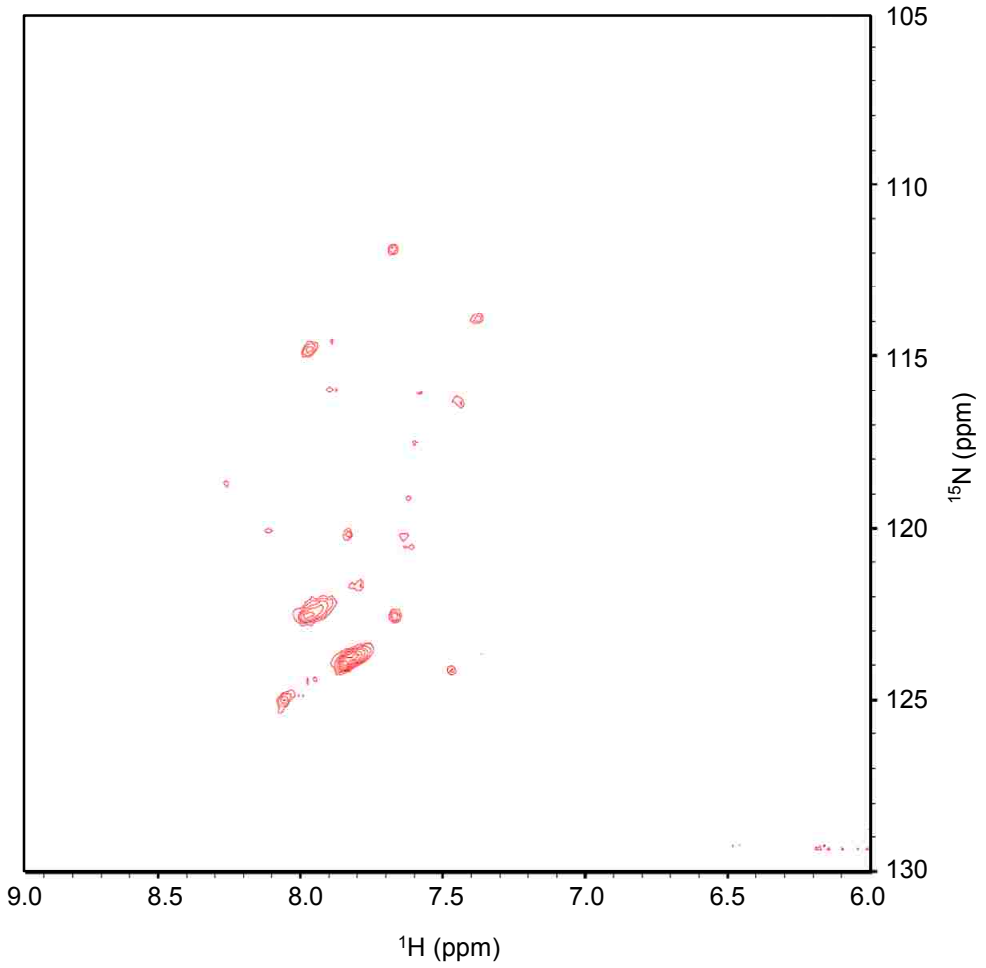


Figure 2-16. ^{1}H-^{15}N TROSY-HSQC spectrum of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136)G108I mutant.

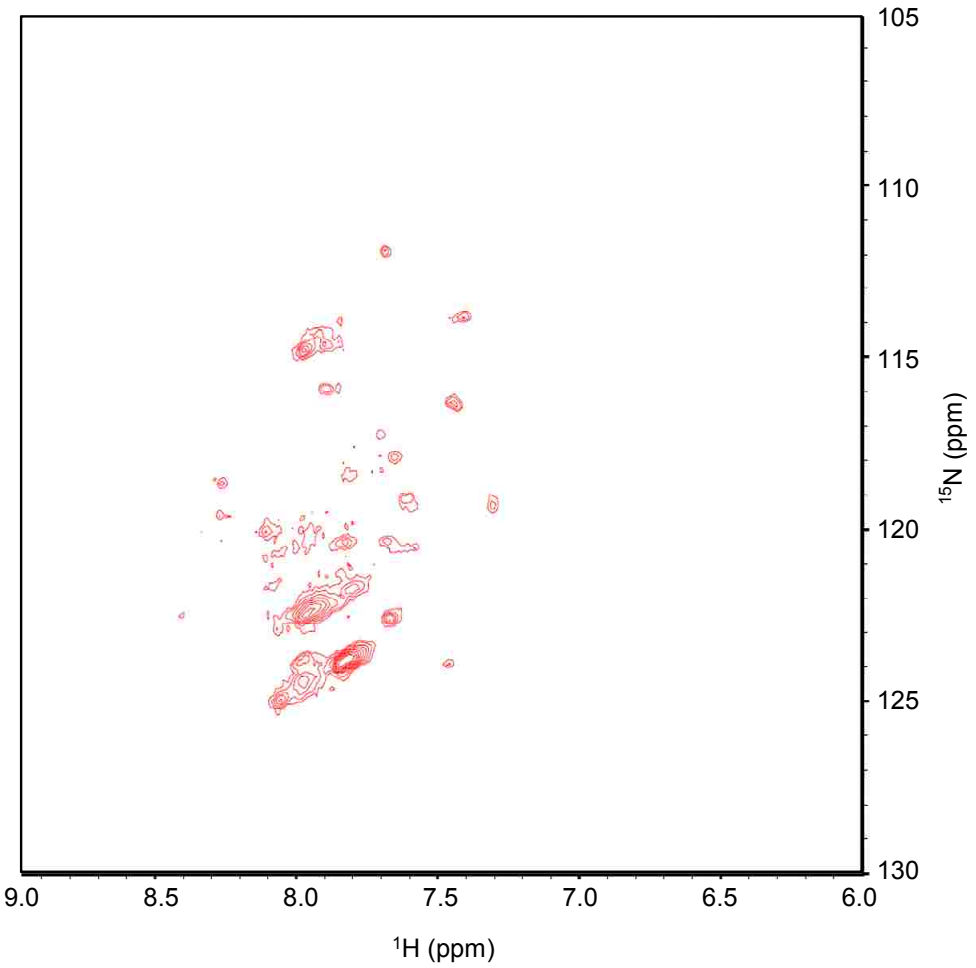


Figure 2-17. ^{1}H-^{15}N TROSY-HSQC spectrum of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136)G108L mutant.

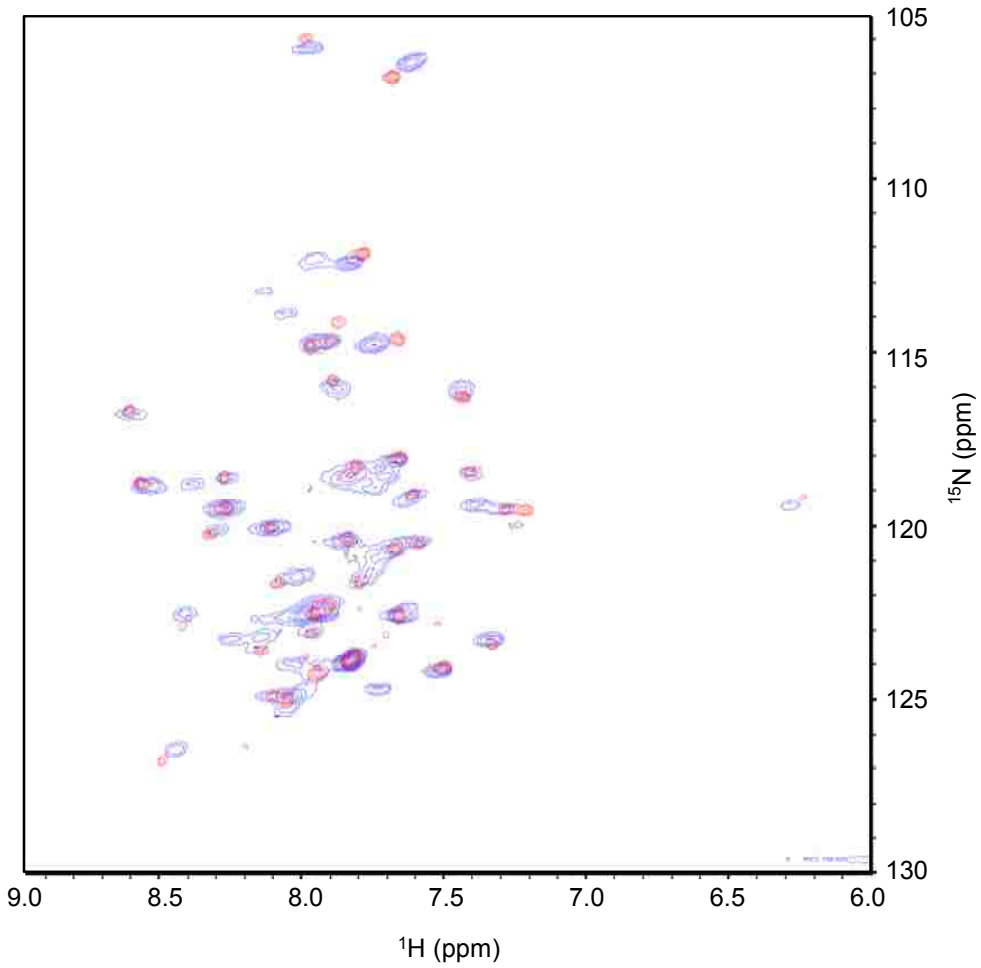


Figure 2-18. Overlay of ^{1}H-^{15}N TROSY-HSQC spectra of uniformly ^{1}H- and ^{15}N-labeled Caveolin-1(96-136) wild-type (red) and Caveolin-1(96-136)I109L mutant (blue).

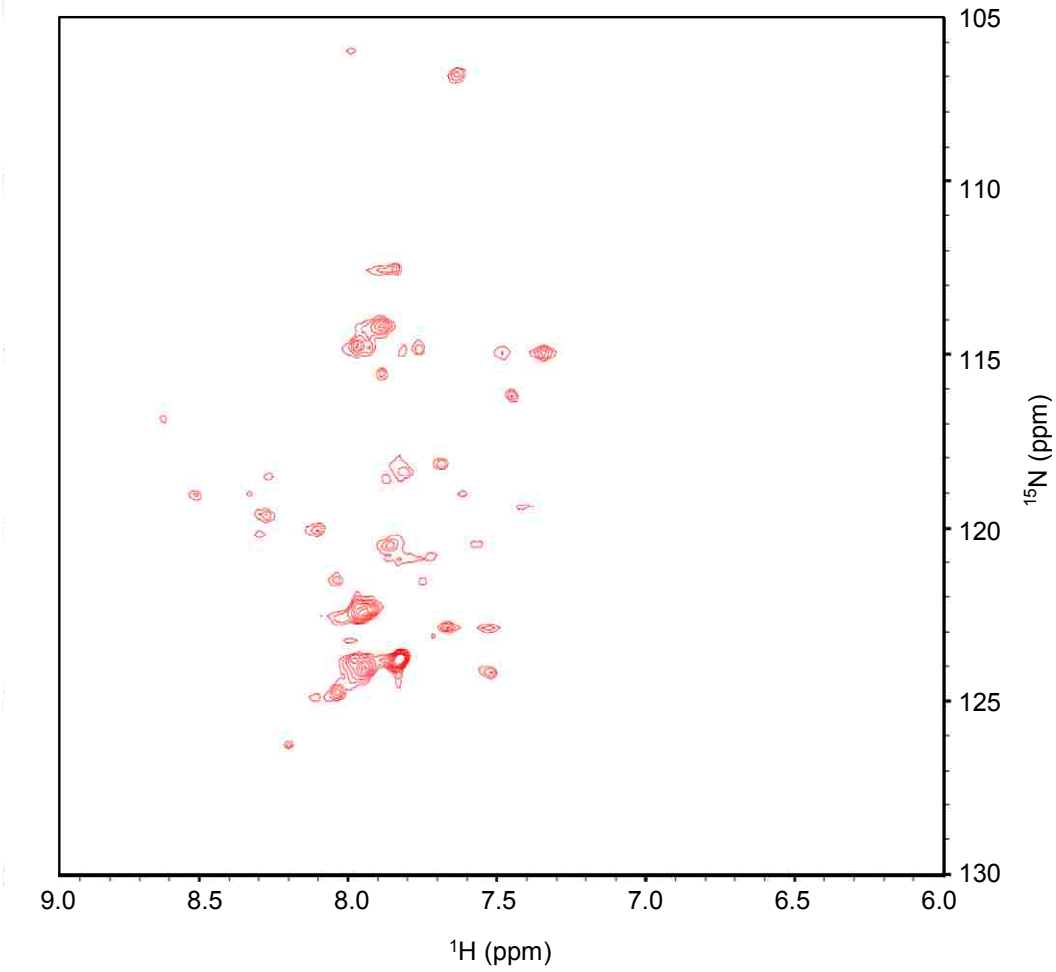


Figure 2-19. ^{1}H-^{15}N TROSY-HSQC spectrum of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136)I109L mutant after 1 week.

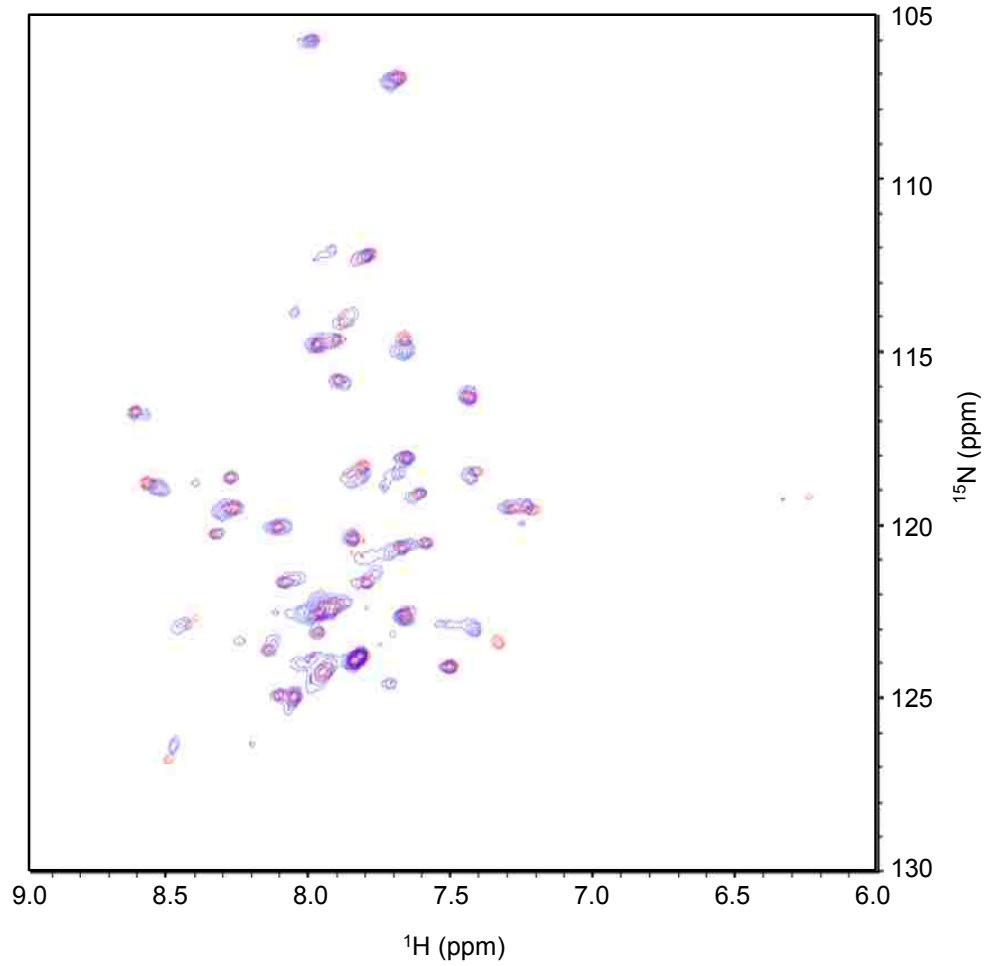


Figure 2-20. Overlay of ^{1}H-^{15}N TROSY-HSQC spectra of uniformly ^{1}H- and ^{15}N-labeled caveolin-1(96-136) wild-type (red) and caveolin-1(96-136)I109V mutant (blue).

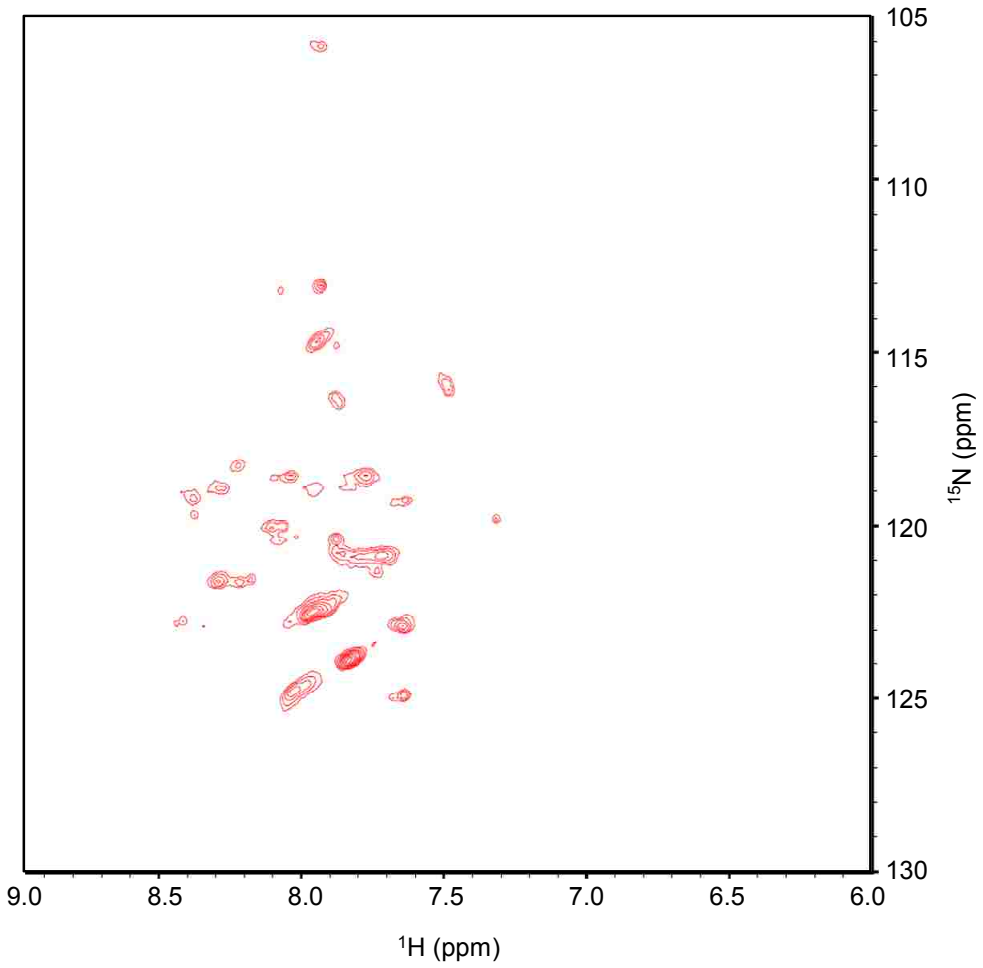


Figure 2-21. Overlay of ^{1}H-^{15}N TROSY-HSQC spectra of uniformly ^{1}H- and ^{15}N-labeled Caveolin-1(96-136) wild-type (red) and Caveolin-1(96-136)I109V mutant (blue).

Conclusion

Caveolin misregulation has been implicated in many different diseases including cancer, Alzheimer's disease, and muscular dystrophy. However, there are few biophysical studies that have probed the structure of this protein, which is not surprising given the extreme hydrophobicity of this integral membrane protein. Using LMPG detergent micelles we have utilized techniques such as circular dichroism and solution NMR to study the characteristics of the caveolin-1(96-136) structure which has never been done for caveolin before. Based on this data we have found that in addition to proline at position 110, two other residues are critical for the helix-break-helix motif observed in this protein. Glycine 108 and isoleucine 109 are also critical for the structure. Additional NMR experiments also reveal that glycine and alanine are tolerable at position 108, β-branched amino acids like valine and isoleucine are required at position 109 and proline is essential at position 110. This helix-break-helix motif is the first critical structural data of the caveolin-1 transmembrane domain that is based on experimental data. This finding will bring us close to elucidating the entire three-dimensional structure of caveolin-1 on an atomic level. In addition, this structural information will allow us to begin to answer questions about the role that this membrane protein plays in the cell, and how it creates the scaffold that causes membrane curvature. Moreover, we hope that these studies will provide a useful platform from which more structural studies on other membrane proteins can be accomplished. Ultimately, these studies will greatly advance the field of membrane protein science and will allow scientists to target these proteins for possible therapeutic intervention.

Appendix 2-1 Primers for point mutation

P110A	GCTGTTCGGTATCGCGCTGGCGCTGAT
I109A	CTGTCTGCGCTGTTCGGTGCGCCGCTGCGCCTGATCTG
G108A	CTGTCTGCGCTGTTCGCGATCCCGCTGGCGCTG
F107A	CTGTCTGCGCTGGCGGGTATCCCGCTGG
G108I	GTCTGCGCTGTTCATCATCCCGCTGGCGC
G108L	GTCTGCGCTGTTCCTGATCCCGCTGGCGC
I109V	CTGCGCTGTTCGGTGTTCCGCTGGCGCTGATC
I109L	CTGCGCTGTTCGGTCTGCCGCTGGCGCTGATC
P110G	CGCTGTTCGGTATCGGCCTGGCGCTGATCTG

Appendix 2-2. Protein constructs for mutation studies.

Caveolin-1(96-136) M111LC133S	KYWFYRLLSALFGIPLALIWGIYFAILSFLHIWAVVPSIKS
Caveolin-1(96-136) M111LC133SG108A	KYWFYRLLSALFAIPLALIWGIYFAILSFLHIWAVVPSIKS
Caveolin-1(96-136) M111LC133SI109A	KYWFYRLLSALFGAPLALIWGIYFAILSFLHIWAVVPSIKS
Caveolin-1(96-136) M111LC133SP110A	KYWFYRLLSALFGIALALIWGIYFAILSFLHIWAVVPSIKS
Caveolin-1(96-136) M111LC133SG108L	KYWFYRLLSALFLIPLALIWGIYFAILSFLHIWAVVPSIKS
Caveolin-1(96-136) M111LC133SG108I	KYWFYRLLSALFIIPLALIWGIYFAILSFLHIWAVVPSIKS
Caveolin-1(96-136) M111LC133SI109L	KYWFYRLLSALFGLPLALIWGIYFAILSFLHIWAVVPSIKS
Caveolin-1(96-136) M111LC133SI109V	KYWFYRLLSALFGVPLALIWGIYFAILSFLHIWAVVPSIKS
Caveolin-1(96-136) M111LC133SP110G	KYWFYRLLSALFGIGLALIWGIYFAILSFLHIWAVVPSIKS

Chapter 3. Probing Caveolin-1(96-136)P132L

Abstract

Caveolin-1 contains two highly conserved proline residues within the transmembrane domain at positions 110 and 132. It has been shown that the proline at position at 110 is critical for the formation of horseshoe conformation. A mutation at the position of the other highly conserved proline (132) to leucine has been shown to be involved with breast cancer pathogenesis. Previous studies show that P132L mutant retained in the Golgi, which is indicative of misfunction, but there are no studies that investigate this mutation in atomic detail. We utilized NMR spectroscopy methodologies to find differences between caveolin-1 WT and the P132L mutant on the atomic level. The Trp leader fusion over-expression system was adopted and the purification was carried out by isolating the inclusion bodies of the WT and P132L fusion to obtain pure caveolin-1 constructs; caveolin-1(96-136), caveolin-1(122-142), and caveolin-1(62-178). The pure protein was reconstituted into a detergent micellar solution (DPC, and LMPG) and was subjected to various NMR experiments such as HSQC, HNCA, HN(CO)CA. The HSQC data of caveolin-1(96-136)P132L suggests that there is a line broadening of the peaks compared to WT caveolin-1(96-136). Furthermore, chemical shift indexing suggested that 4 additional amino acids were included in helix-2 compared to the WT protein. Although there are 4 extra amino acid extension the helix-break-helix motif was still preserved in the mutant. These data suggest that the caveolin-1 P132L mutant is local conformational change and not global conformational change. The findings in this study will lead us to a better understanding of caveolin-1 P132L mutation and its role in the metastasis of breast cancer.

Introduction

Caveolin is an intergral membrane protein that is essential to form caveolae which are invaginations in cell membrane (23, 24). The loss of caveolin expression leads the to loss of caveolae formation and this leads to disease states as caveolae are a major hub of cell signaling activities. Thus mutations of caveolin and its misregulation have been closely linked to a number of disease states including heart disease, Alzheimer's, and cancer (80, 81). For example, lack of caveolin-1 leads to lesion formation in mouse mammary tissue that is a precursor to tumor development (9). This study shows that caveolin is necessary for proper cell function and when it is not regulated properly it leads to disease state. Disease states do not only occur when caveolin is completely silenced, another situation is where a caveolin-1 point mutation has occurred. In particular, the P132L mutation is a well-known and well-studied point mutation *in vivo*. This mutation is a naturally occurring and is found in approximately 16% of breast cancer patients (13, 29, 31). The P132L mutant is retained to the Golgi apparatus and caveolin does not reach its terminal destination in the plasma membrane, hence, there is a lack of caveolae. Therefore, the expression of the P132L mutant results in loss of the caveolae in the plasma membrane. Moreover, the P132L mutant acts in a dominant negative manner which means when both mutant and wild-type are co-expressed, both are retained in Golgi apparatus (13). Since it is a typical phenomenon that misfolded proteins are retained in Golgi it is clear that the P132L mutation leads to a misfolded protein, and therefore, leads to breast cancer. Therefore, a definite structural change in the caveolin protein has occurred. However, there are no biophysical studies that have probed the structural changes when the P132L mutation occurs. For example, is it global conformational changes or local small changes which contribute to the non-functionality of this misfolded protein. In this

study we utilized NMR spectroscopy and chemical shift indexing method to analyze conformational changes in the caveolin-1 transmembrane domain. Herein, we present a series of experiments which will pin down the nature of the caveolin-1 P132L mutation. Specifically, we will determine if the P132L mutation leads to local changes that will cause misfolding or global structural changes in the key transmembrane domain such as loss of the helix-break-helix motif. Using NMR and CSI will be able to investigate this question.

Material/Methods

Cloning/ Expression and Purification

Three constructs were prepared to address questions surrounding the nature of the P132L mutation. Caveolin-1(62-178), caveolin-1(96-136), and caveolin-1(122-142) was designed to probe P132L mutation and its effect on the global or local structural changes. Both wild-type and mutants are prepared for comparison. Preparation of wild-type construct of caveolin-1(96-136) and caveolin-1(122-142) is described in previous chapter and caveolin-1(62-178) construct was purchased from Genscript (Piscataway, NJ) and was prepared using the same method as caveolin-1(96-136). All caveolin-1 P132L mutant constructs were prepared using the modified QuikChange site-directed mutagenesis method. A primer was designed using web-based program primerX (Bioinformatics.org),

5′-CATCTGGGCGGTTGTTCTGTCTATCAAATCTTGAGA-3′.

The reaction was performed using the 9ºN ligase and pfu turbo method detailed in the previous chapter. Reaction products are digested and transformed into XL-1 blue cells. The mutation is confirmed *via* DNA sequencing and is then transformed into BL21(DE3) cells for protein expression. Expression was performed using the same method as the wild-type with Trp leader fusion protein and purification was accomplished using inclusion body purification method. Detailed information concerning expression and purification are outlined in previous chapters. For the backbone assignment of caveolin-1(96-136)P132L, uniformly ^{15}N, ^{13}C-labeled caveolin-1(96-136)P132L is required. The expression of uniformly labeled ^{15}N and ^{13}C caveolin-1(96-136)P132L was prepared as described in Marley et al (58). Moreover, to aid in backbone assignments specific amino acid labeling of caveolin-1(96-136)P132L was

employed. Labeled amino acids were incorporated by expressing caveolin-1(96-136)P132L according to the procedure outlined in Truhlar et al (59). The previously discussed methods of purification were carried out to obtain an NMR sample (see chapter 1 methods section)

Sample preparation

For circular dichroism studies, 3 mg of protein was dissolved in 2.84 mL of TFE with 38.50 mg of DMPC. This solution is injected into water and is immediately frozen and lyophilized to obtain a nice powder. The powder was reconstituted with 298 µL of water and 12 µL of 250mM phosphate. The sample appeared milky and vigorously vortexing was used to obtain homogeneity. 200 µL of 25% DHPC solution was added while vortexing and sample became clear. The sample was filtered with a 0.2 µm regenerated cellulose spin filter to remove dust. The concentration was adjusted using bicelle buffers.

For NMR samples, 500 µL of LMPG micellar solution (100 mM LMPG, 100 mM NaCl, 20 mM phosphate pH 7, 10% D2O) or DPC micellar solution (100 mM DPC, 100 mM NaCl, 20 mM phosphate pH 7, 10% D2O) was added to lyophilized caveolin constructs. The sample was vortexed vigorously and briefly heated in a water bath at 42 °C until clear. The final protein concentration was adjusted to approximately 1 mM using a Biospec nano drop spectrometer operating at the 280 nm wavelength. The clear solution was spin filtered through a 0.2 µm filter to remove any particulates. The sample was then loaded into an NMR tube.

CD/NMR

Circular dichroism experiments were performed using a JASCO CD Spectrophotometer (Easton, MD). The experiment was performed scanning

wavelength ranges of 260 to 190 nm at 25 °C. A background spectrum was obtained by running a sample without protein and was subtracted. Approximately 250 µL of sample was loaded into a 0.1 mm path length cuvette. Spectra were analyzed using the CDSSTR algorithm in Dichroweb (45, 46, 47).

The NMR spectra were acquired at 37 °C using a 600 MHz Bruker Advance II NMR spectrometer (Billerica, MA) equipped with a cryoprobe at Penn State Hershey Medical College (Hershey, PA). TROSY based HSQC was performed to compare wild-type and P132L constructs first. For backbone assignments of caveolin-1(96-136)P132L TROSY based HSQC, HNCA, and HN(CO)CA experiments were performed. Specific amino acid labeling (valine, serine, leucine) was used to clarify ambiguous peaks. The spectra were processed using NMRPipe and Sparky (48, 74). Secondary structure information was obtained using Cα chemical shifts as described in Wishart et al (75).

Results and Discussion

The naturally occurring mutant caveolin-1 P132L has been implicated in the pathogenesis of breast carcinomas. To date, molecular biology experiments have shown that it is likely due to misfolding of the protein, however, there is no definite answer due to lack of structural evidence (13, 28). Therefore, we wanted to examine the secondary structure of caveolin-1 P132L mutant by CD and NMR spectroscopy, which gives structural information. First caveolin-1(122-142) construct was prepared in DPC micellar solution since DPC is a frequently used membrane mimic in the NMR studies and is powerful enough to dissolve caveolin-1(122-142). The caveolin-1(122-142) construct was chosen because the P132 residue resides at the center of the caveolin-1(122-142) peptide, and it is flanked by 10 residues on each side. In addition, this construct will only show approximately 20 peaks in the HSQC spectrum which will be easy to assign the peaks. The HSQC data of caveolin-1(122-142) shows well dispersed peaks in the spectra which indicate a well-structured protein. The comparison of wild-type and the P132L mutant suggests there is a structural change. The chemical shifts of the P132L mutant are moved compared to the wild-type protein suggesting conformational differences but it is hard to distinguish between global or local conformational changes because the construct was too short (Figure 3-1). Therefore, we investigated a caveolin-1(62-178) construct that contains the entire membrane interacting portion. First the construct was reconstituted into DPC micellar solution and an HSQC experiment was performed. The number of peaks in the spectrum was much less than expected. Thus detergent screening was performed to find a better mimic for NMR studies. The LMPG micellar solution system was chosen because the spectrum was well dispersed and the number of peaks closely corresponds to the number of amino acids in the HSQC experiments (Figure 3-2).

Therefore, comparison of wild-type and the P132L mutant was performed in LMPG micellar solution. The spectra of wild-type and P132L comparison show that there is local conformational changes not global conformational change (Figure 3-3, Figure 3-4). Although it is clear from these experiments that the P132L is a local structural mutation, peak assignments are required to indicate the nature of the structural change. Because of the degeneracy apparent in the spectrum, it is not easy to tackle this problem. Therefore, caveolin-1(96-136) was chosen to investigate this mutation since caveolin-1(96-136) contains the core of caveolin-1.

CD experiments were carried out to examine the difference between wild-type and P132L caveolin-1 in their global secondary structure. Examination of these two spectra indicates very similar secondary structure (WT is 66% α-helicity and P132L is 62% helicity here) (Figure 3-5). This suggests that there is not a dramatic conformational change, and that the nature of the mutation results in local conformational changes which correspond to the results of other constructs. To give additional support to the results from the CD study, NMR structural studies were undertaken to observe the changes that occur in the transmembrane domain of the wild-type versus the P132L mutant. The NMR samples were prepared using LMPG detergent instead of bicelles which is used in CD studies because the concentration of protein in bicelles is not high enough to run NMR experiments. First the HSQC spectrum of caveolin-1(96-136)P132L mutant spectrum was acquired (Figure 3-6). Comparison of the wild-type and mutant spectra show that mutant spectrum line widths are broader and the peaks are more shifted for residues adjacent to the point mutation. Also, several three dimensional spectra were acquired to assign the backbone as was done for the wild-type protein. A specific amino acid labeling method was also undertaken to help backbone assignment. Once the peaks in the spectra were

assigned unambiguously, we analyzed the secondary structure by using the Cα chemical shift index developed by Wishart et al (75). We have confidence that this method is sufficient to draw conclusions as to the nature of the P132L mutant since the prediction of secondary structure for the wild-type protein using Cα chemical shift data agreed nicely with the data obtained from computer program TALOS+ which requires much more data input and is therefore much more restrictive (49). As previously described, chemical shift indexing examines the difference between obtained Cα chemical shifts of the protein and compares it to random coil Cα chemical shifts to reveal the secondary structure of caveolin-1(96-136)P132L. Chemical shift indexing of caveolin-1(96-136)P132L revealed two regions which have consecutively positive ΔCα values which indicates that the mutant chemical shift indexing was similar to the wild-type (Figure 3-7). Based on chemical shift indexing data, the caveolin-1(96-136)P132L mutant still retains the helix-break-helix motif, and therefore is similar in secondary structure to the wild-type protein. Specifically, the caveolin-1(96-136)P132L mutants first α-helix encompassed residues 97–107 followed by a break from residues 108–110 and then a second α-helix from residues 111–133. After residue 133 there is an unstructured region from residues 134–136. Interestingly, the secondary structure trend of the P132L mutant is also consistent with the horseshoe topology model of caveolin-1. The break at residues 108–110 is the putative turn point that allows helix-1 and helix-2 to be on the same side of membrane. Therefore, CD and NMR data suggest that caveolin-1 P132L mutant is not changing the structure dramatically because the mutant contains a similar amount of α-helicity and also retained the helix-break-helix motif which is likely to be a critical for the proper orientation of the caveolin protein. We are confident that the only structural changes that have occurred are localized to the region of the P132L mutation as helix-1 is not affected and the break is still occurs at the same amino acids; Glycine, Isoleucine, and Proline. Helix-2 started

at same residue at position 111 but terminated at a different residue than wild-type. The second helix of the wild-type protein encompasses residues 111 to 129. This is in contrast with the mutant protein where the second helix was extended to residue 133. This means that there are four extra amino acids (1 helical turn) that are more structured compare to wild-type. It is not that surprising that substituting proline to leucine created a more structured helix. Proline is known to induce breaks in helical regions by creating kinks due to its unusual rigidity in allowed phi and psi angles. Thus the mutation of the proline residue at position 132 to leucine changed the dynamic region to a more structured one. Moreover, studies based on gel filtration of wild-type and various mutants as well as analytical ultracentrifugation show that P132L mutant forms a dimer whereas the wild-type behaved as monomer (28). The extension of the protein by one extra helical turn is likely opening up a dimerization region around position 132. In the wild-type protein the proline at position 132 is critical to cause helix-2 to terminate at position 129 which prevents biologically irrelevant dimerization. However, when proline is mutated to leucine causing the elongation of helix-2 by 4 extra amino acids helix-2 contains an exposed dimerization region. This could be a possible new mechanism based on a structural point of view to explain why this mutation aids in the development of breast cancer state. As dimerization occurs, caveolin-1 is retained in Golgi and never reaches to the plasma membrane and leads to the loss of caveolae.

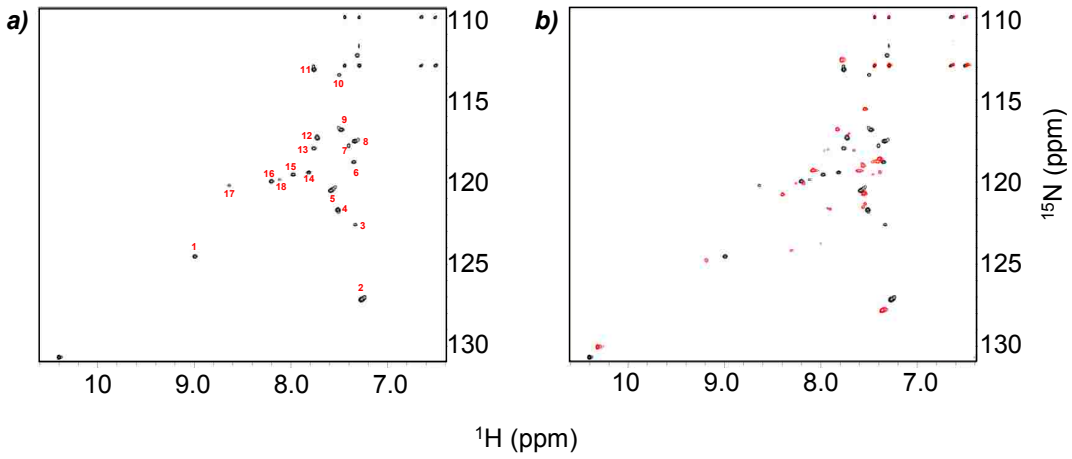


Figure 3-1. ***a)*** ^{1}H–^{15}N TROSY HSQC spectrum of caveolin-1(122-142) in 100 mM Dodecylphosphocholine micelles. ***b)*** Overlay of TROSY-HSQC spectra of wild-type (black) and P132L mutant (red).

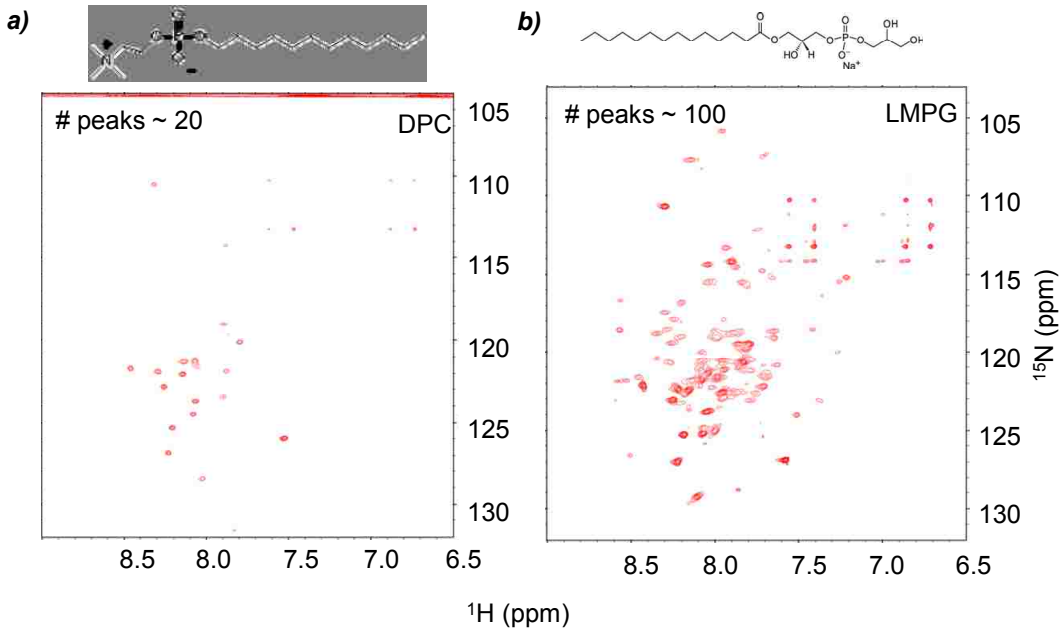


Figure 3-2. ^{1}H–^{15}N TROSY-HSQC spectra of caveolin-1(62-178). ***a)*** 100 mM Dodecylphosphocholine micelles. ***b)*** 100 mM lysomyristoylphosphatidylglycerol micelles.

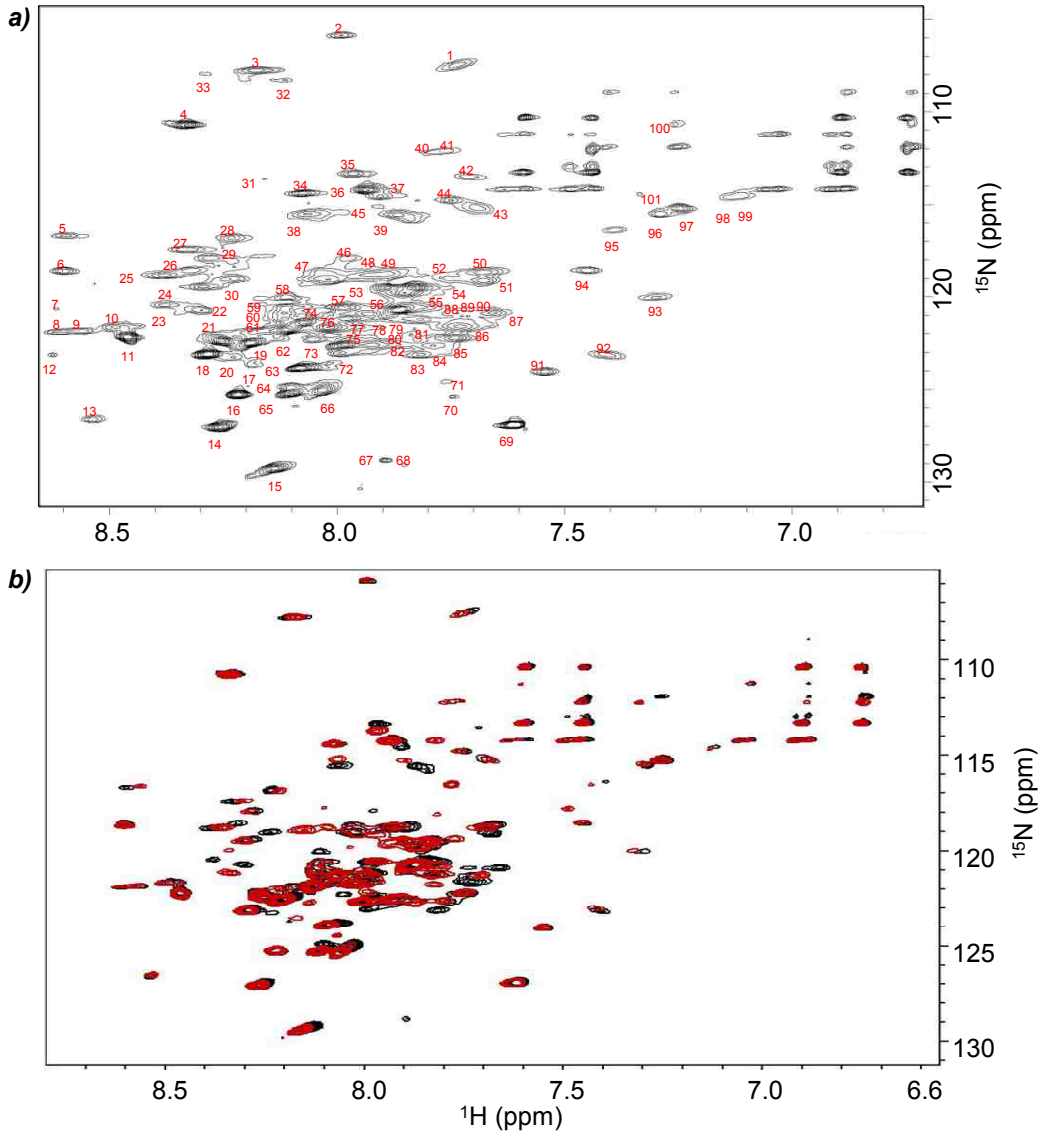


Figure 3-3. ***a)*** TROSY-HSQC of ^{15}N labeled caveolin-1(62-178) (0.5 mM); 20 mM phosphate buffer pH 7.0, 100 mM lysomyristoylphosphatidylglycerol (LMPG), 100 mM NaCl, 10% D_2O, 37°C. ***b)*** Overlay of TROSY-HSQC of ^{15}N labeled caveolin-1(62-178) wild-type (black) and P132L (red).

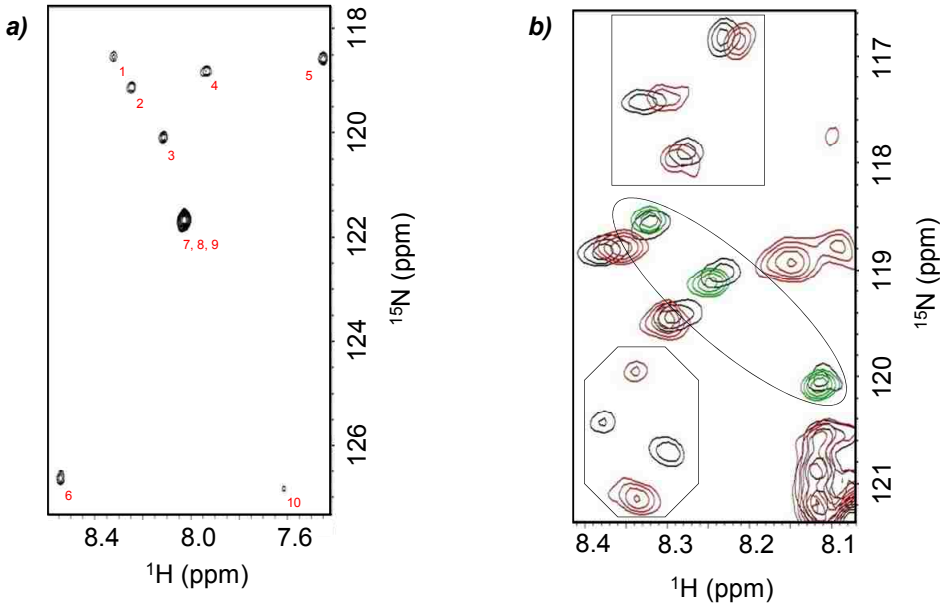


Figure 3-4. ***a)*** TROSY-HSQC spectrum of caveolin-1(62-178) with leucine residues ^{15}N labeled. ***b)*** Overlay of TROSY-HSQC spectra of ^{15}N labeled caveolin-1(62-178) (black), ^{15}N labeled caveolin-1(62-178)P132L (red), and ^{15}N leucine caveolin-1(62-178) (green).

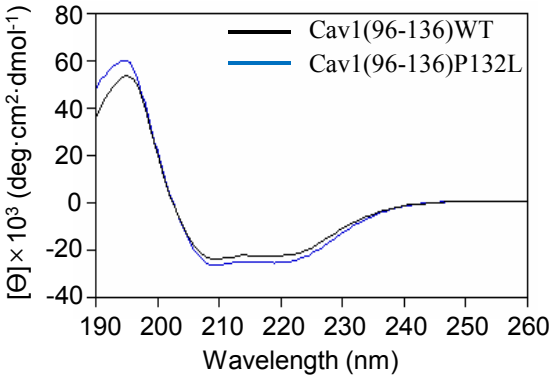


Figure 3-5. Overlay of Circular Dichroism spectra of caveolin-1(96-136) wild-type (black) and P132L mutant (blue).

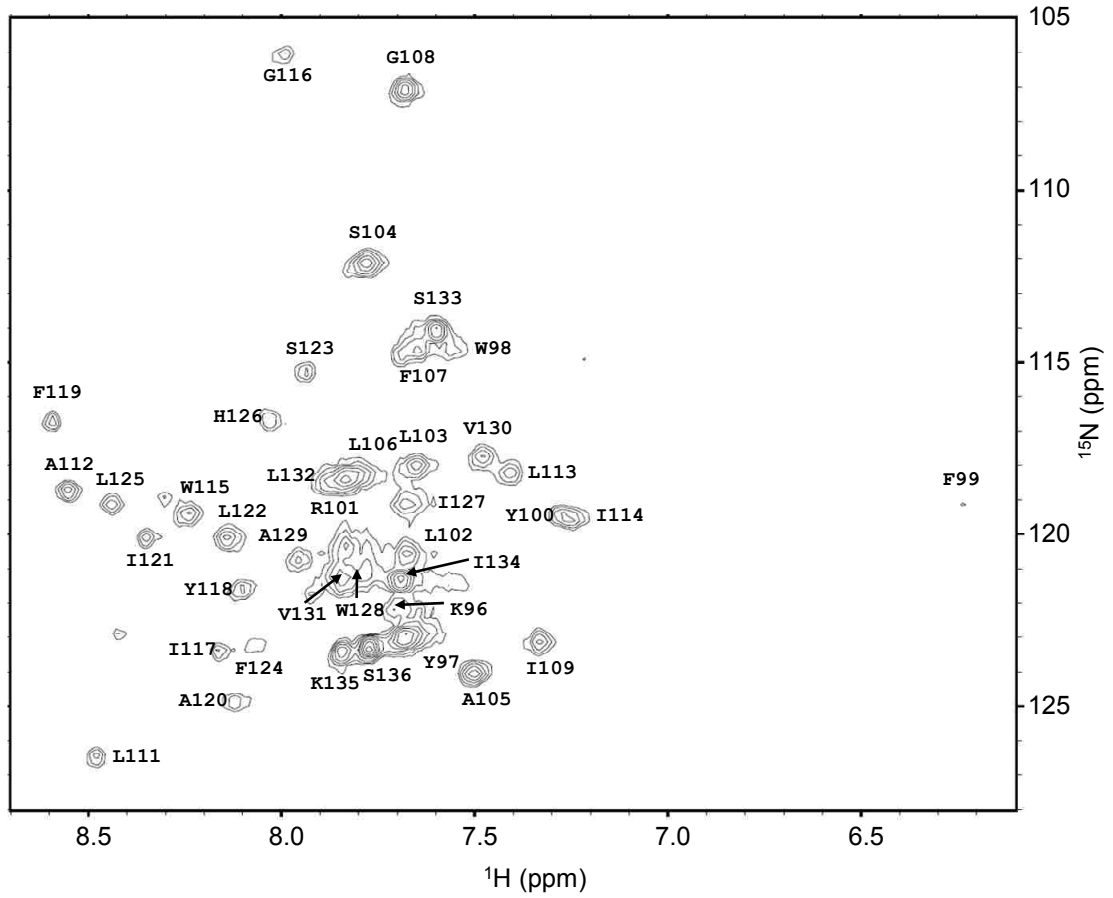


Figure 3-6. ^{1}H–^{15}N TROSY-HSQC spectrum of Caveolin-1(96-136)P132L. The spectrum was acquired with 256 complex points in t1 (^{15}N) and 2048 complex points in t2 (^{1}H).

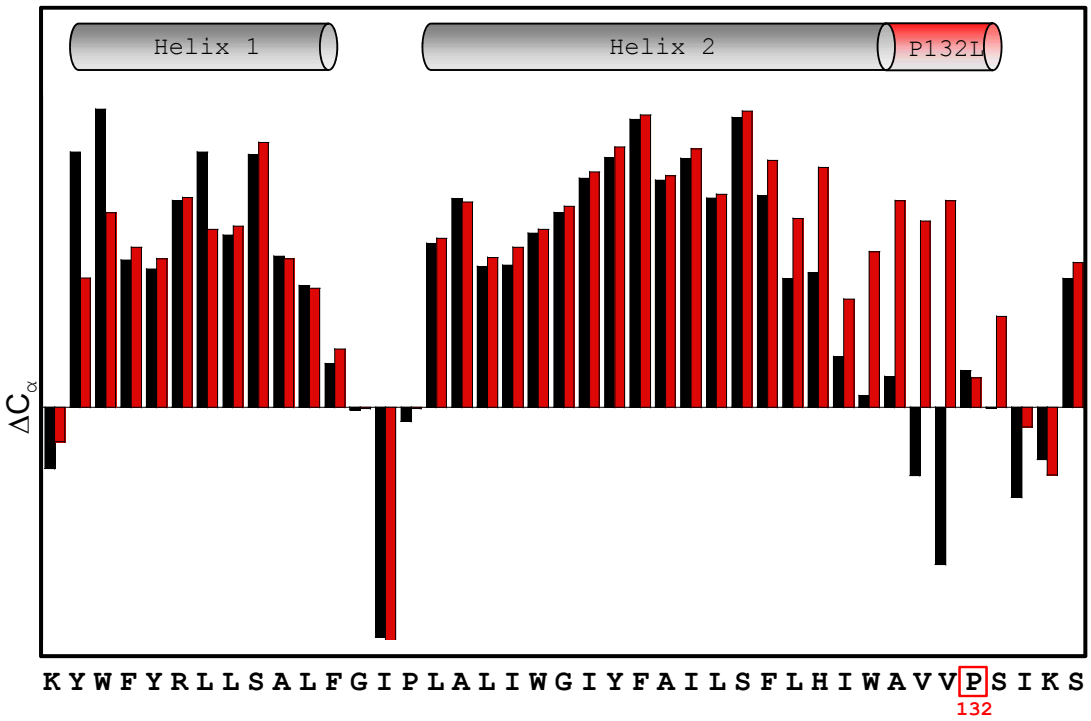


Figure 3-7. Chemical shift analysis of caveolin-1(96–136) and caveolin-1(96–136)P132L. Wild-type is depicted with black bars, and the mutant is depicted with red bars. Caveolin-1(96–136)P132L shows an extension of helix 2 by four residues.

Conclusion

Caveolin misregulation has been implicated in many different diseases including cancer, Alzheimer's disease, and muscular dystrophy. Interestingly, the caveolin-1 P132L mutation is found in 16% of breast cancer patients. Many *in vivo* studies have shown that the caveolin-1 P132L mutant was retained in the Glogi but there is no biophysical data to clarify the mechanism. Our over-expression system and reconstitution methods allow us to tackle the question of why P132L is so disruptive to normal caveolin-1 function. Using NMR to solve the secondary structure, residue by residue information revealed conformational differences between P132L and wild-type proteins. The NMR studies reveal that P132L was not changing its conformation globally because it still has the critical helix-break-helix motif. Local changes are shown to be occurring in the caveolin-1(96-136)P132L mutant by extending the second transmembrane α-helical region by 4 additional amino acids at the end of helix-2 using chemical shift indexing NMR analysis. This small extension might expose dimerization region. Dimerization of caveolin-1 P132L mutant suggest proline has an important role in stopping helix-2 at position 129 but when proline is mutated to leucine it expose dimerization sites. This dimerization could be a new mechanism of breast cancer. Dimerization sequesters caveolin-1 P132L to the Golgi and results in a loss of caveolae in the plasma membrane. This finding will lead us to attain a better understanding of the caveolin related breast cancer pathogenesis and will ultimately allow therapeutic interventions to be investigated.

Chapter 4. Extended Studies of Caveolin-1 Transmembrane Domain

Abstract

Caveolin-1 has a typical topology compared to most other membrane proteins. Both the N- and C- temini are facing towards the cytoplasm. Based on the unique topology and hydrophobic characteristics of caveolin-1 a model was proposed using primary sequence analysis. However, there are no detailed biophysical studies that support the topology model. Therefore, various caveolin-1 constructs were prepared and CD and NMR spectroscopy techniques were utilized to build a topology model of caveolin-1. The NMR spectrum of caveolin-1(82-136) was assigned and analyzed using Cα chemical shift indexing. The data suggests that there are two α-helices of similar lengths in the caveolin-1(82-136); helix-1 includes residues 87-107 and helix-2 includes residues 111-129. In addition, caveolin-1(82-136) still contains helix-break-helix motif like caveolin-1(96-136). This helix-break-helix motif has similar length helices that may insert into lipid bilayers asymmetrically in the bilayer (confined to one leaflet) and curve the membrane to form caveolae. Moreover, analysis of CD data and helical wheel projection show that there is an amphipathic helix at the C-terminal domain starting after position 143. From this information we proposed a model of the caveolin-1 membrane interacting domain. This will allow us to better understand the role of caveolin-1 in modulating membrane curvature.

Introduction

The caveolin-1 transmembrane domain plays a critical role in various cell types (29, 30, 82, 83). Although it is critical to study the transmembrane domain, the extreme hydrophobicity has limited progress in this area. Until now, most of the work on caveolin focused on the scaffolding domain (residues 82-101) of the protein. This region of the protein has been implicated as a cholesterol binding region and it is suspected to interact with various signaling molecules (84, 85, 86, 87, 88). The scaffolding domain is a short twenty amino acid segment that immediately precedes the transmembrane domain. Based on a helical wheel analysis, the scaffolding domain forms an amphipathic α-helical structure, which suggests that it is located at the surface of the plasma membrane. Also, recent studies using synthetic peptides modeled after the scaffolding domain suggest that the presence of the transmembrane domain may be critical for the structure of the scaffolding domain. Le Lan and colleagues found that the scaffolding domain was poorly structured unless five residues from the transmembrane domain were included in the peptide. This data suggests that the conformation of the scaffolding domain may be dependent on the transmembrane domain, highlighting the importance of the transmembrane domain. In addition, Parton and co-workers built a model of caveolin-1 based on primary sequence analysis. In this model, the caveolin-1 scaffolding domain is a separate domain from transmembrane domain suggesting there is a break between the helix of the scaffolding domain and helix-1 of the transmembrane domain. Moreover, previous data from caveolin-1(96-136) suggest that the first helix starts at residue Y97. Therefore, it is not clear that where the caveolin-1 transmembrane domain starts. Because the previous construct was only residues 96-136, we were not able to determine the starting position of the first helix. For this reason, we designed a

construct encompassing residues 82 to 136 to capture the full essence of the scaffolding domain and the transmembrane domain which contains the critical horseshoe conformation of the protein. Caveolin-1(82-136) was expressed as a Trp leader fusion protein to maximize the protein yield. A methionine residue was included after the Trp leader sequence so that the target protein could be isolated after purification. NMR spectroscopy was utilized to investigate the question of where helix-1 of the transmembrane domain starts. Moreover, we also designed longer constructs (caveolin-1(62-136) and caveolin-1(96-178)) and utilized circular dichroism spectroscopy in combination with CSI to elucidate other portion of caveolin-1.

Material and Methods

Expression and Purification

The genes of constructs were purchased from Genscript and inserted into pET-24a vector containing Trp leader fusion protein as previously described. Expression and purification was also performed by the same method as caveolin-1(96-136).

For NMR studies of caveolin-1(82-136), isotopic labeling was necessary. All samples were prepared as previously described except the specific amino acid labeling cell growth was scaled down to 500 mL and triple labeled (^{2}H, ^{13}C, ^{15}N) sample was optimized from ^{13}C and ^{15}N double labeling method. Deuterated water was used instead of water in M9 media and the incubation time before and after IPTG (isopropyl β-D-1-thiogalactopyranoside) was doubled. Detailed methods are described in expression and purification chapter.

Sample Preparation

For NMR studies lyophilized protein was reconstituted into 500 µL of LMPG buffer (100 mM lyso-myristoylphosphatidylglycerol, 100 mM NaCl, 20 mM phosphate pH 7, and 10% D_2O). Clear solutions were obtained and subjected to 0.2 µm regenerated cellulose spin filter. The final concentration was approximately 0.7 mM to 1.0 mM. Also circular dichorism samples were prepared in the same way as NMR samples using LMPG buffer without salt. The final concentration of samples for the CD study was 0.2 mM.

CD

Circular dichroism spectra were acquired at 25°C using a JASCO circular dichroism spectrophotometer (Easton, MD). A quartz cuvette with a path length of 0.1

mm was utilized. The spectra were obtained from 260 to 190 nm using step mode. For each experiment, 8 accumulations were averaged. A background spectrum using buffer was subtracted from the sample spectra. Spectra were analyzed using the CDSSTR algorithm in Dichroweb (45, 46, 47). This methodology was used for the following constructs, caveolin-1 (62-136), caveolin-1 (82-136), caveolin-1 (96-136), and caveolin-1 (96-178).

NMR

All NMR spectra were acquired at 37°C using a 600 MHz Bruker Avance II spectrometer (Billerica, MA) equipped with a cryoprobe. For analysis and backbone assignments the following TROSY-HSQC-based pulse sequences were employed: HSQC, HNCA, HNCACB, and HN(CO)CA. The spectra were processed using NMRPipe and Sparky (48, 74). For chemical shift indexing, $\Delta C\alpha$ was obtained by subtracting observed Cα chemical shifts from random coil chemical shifts as described by Wishart et al (75). To aid backbone assignments, specific amino acid labeling was employed (Gly, Phe, Tyr, Leu, Ile, and Val). All of the information was compiled and input into a computer program, TALOS+, to obtain dihedral angles. This data was then subjected to a computer program CS-ROSETTA to obtain preliminary structure (49, 76, 77).

Result and Discussion

Caveolin-1 transmembrane domain is thought to encompass residue L102 to residue I134 based on primary sequence analysis. However, previous NMR studies on caveolin-1(96-136) show that the transmembrane domain helix-1 encompasses residues Y97 to F107 and helix-2 encompasses residues L111 to A129. It reveals that helix-1 is composed of 11 amino acids and helix-2 is composed of 21 amino acids. Moreover, the unique feature of caveolin is that N-and C- termini are facing same side. To build model based on this information helix-1 and helix-2 have to insert into lipid bilayer at different angles (Helix-1 has to insert more perpendicular to the lipid bilayer). But chemical shift indexing of caveolin-1(96-136) shows that helix-1 starts without an unstructured region unlike helix-2 which has a dynamic region at the end of helix. Therefore, there are two possible explanations for this result. The first possible explanation is caveolin-1(96-136) construct was too short to contain the full transmembrane domain. So ideally helix-1 and helix-2 should be similar in length to insert into the lipid bilayer which the same tilt. Second is that there is a break portion before helix-1 and there is another helix (scaffolding domain) to help with the incorporation of the caveolin transmembrane domain into bilayers. In addition previous studies show that caveolin contains a scaffolding domain that occurs before the transmembrane domain. However it is not clear if this domain is separate from helix-1 or is a continuation of helix-1. Therefore we extended the transmembrane domain construct to investigate structural aspects of the enlongated construct, caveolin-1(82-136). The expression and purification methods were performed using a similar method to caveolin-1(96-136). Also pure caveolin-1(82-136) was reconstituted into LMPG micellar buffer and subjected series of NMR experiments.

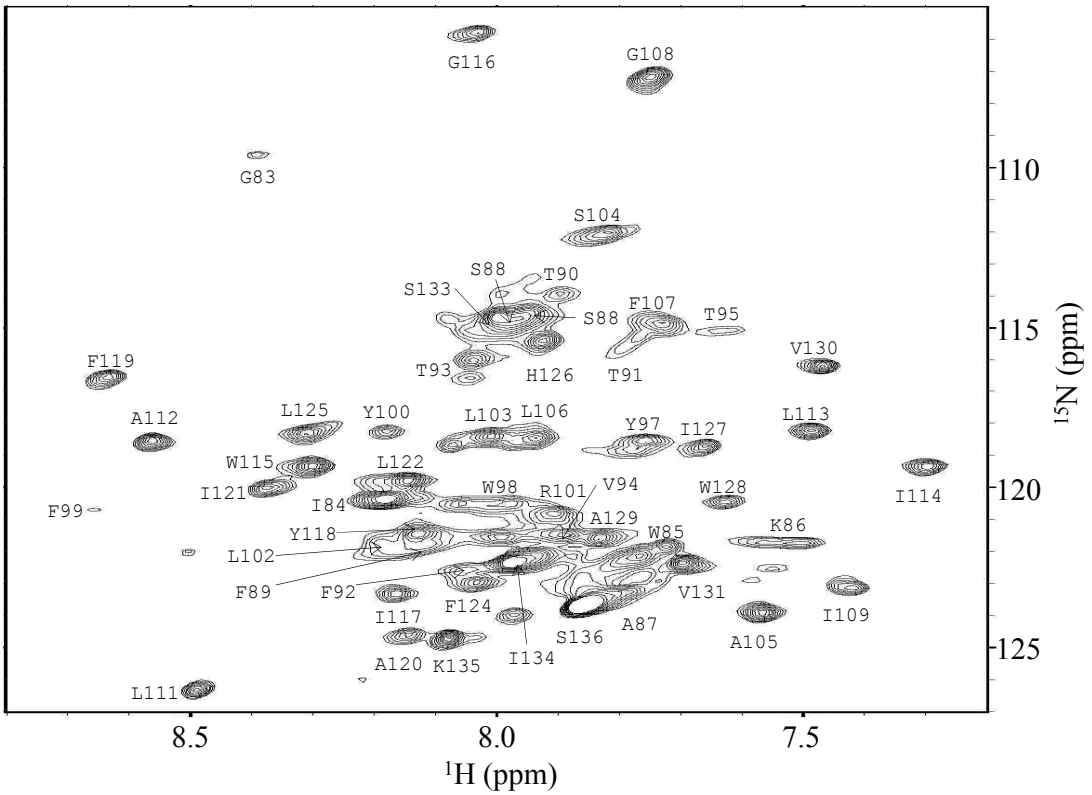


Figure 4-1. Assigned ^{1}H-^{15}N TROSY-HSQC spectrum of caveolin-1(82-136) in 100 mM 1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) micelles; 100 mM NaCl; 20 mM phosphate pH 7.0. The spectrum was acquired with 256 complex points in the t1 dimension (^{15}N) and 2048 complex points in the t2 dimension (^{1}H).

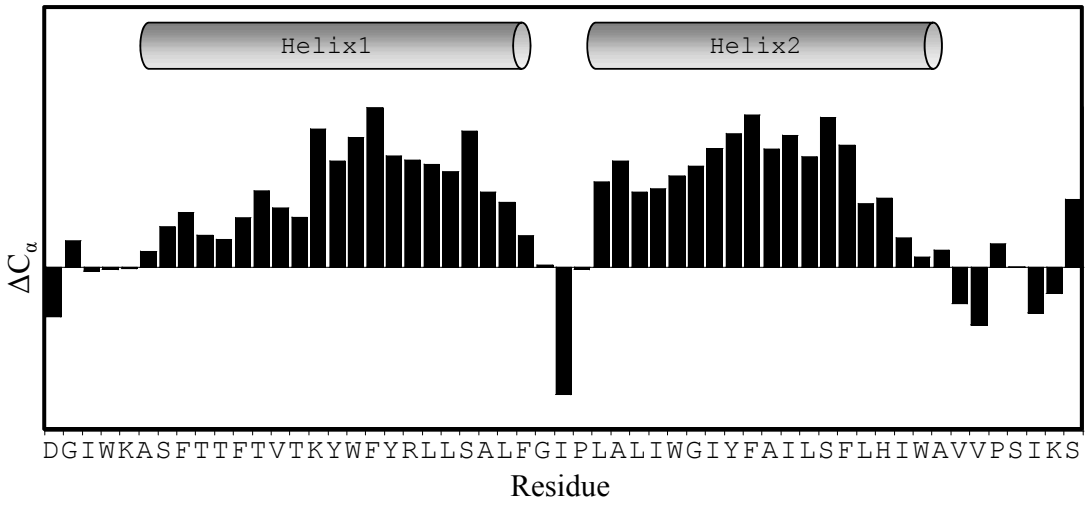


Figure 4-2. Chemical shift index plot of caveolin-1(82-136). Helix1 spans residues 87-107 and helix2 spans residues 111-129. The break between helix1 and helix2 spans residues 108-110.

To assign the backbone resonances employing NMR experiments, uniformly labeled caveolin-1(82-136) (^{2}H,^{13}C, ^{15}N) in LMPG micelles was performed using the following two and three dimensional TROSY-HSQC-based experiments: HSQC, HNCA, HNCACB and HN(CO)CA (Figure 4-1). Because of degeneracy, specific amino acid labeling was employed to aid in backbone assignments (Gly, Phe, Tyr, Leu, Ile, and Val). Based on Cα chemical shift indexing, residues 87-107 and residues 111-129 have consistently positive ΔCα's which is indicative of an α-helical structure and residues 108-110 which are in between do not have positive ΔCα's (Figure 4-2). Thus caveolin-1(82-136) still has helix-break-helix motif and helix-1 starts at position A87 instead of Y97. The transmembrane domain is comprised of two α -helices of near equal length. Helix 1, residues 87-107, is composed of 21 amino acids (a length of 31.5 Å), helix-2, residue 111-129, is composed of 19 amino acids (a length of 28.5 Å). No break was observed before residue Y97 which suggests that the scaffolding domain is a part of helix-1. Moreover, these similar lengths of two helices will insert caveolin into bilayer in similar angles. The preliminary model was created inputting compiled information into CS-ROSETTA computer program (Figure 4-3).

Next, we investigated two extended caveolin-1 transmembrane domain (residues 62-81 and residues 137-178). Previous studies show that caveolin-1 N-terminal domain is unstructured dynamic portion of the protein (17). To investigate this we extended construct from residue 62 to 136. Comparison of caveolin-1(62-136) and caveolin-1(82-136) could obtain information about residues 62-81. To obtain information of C-terminal domain we designed construct from reside 96 to 178. Comparison of caveolin-1(96-136) and caveolin-1(96-178) could obtain information about residues 137-178. The construct was prepared using the same methodology as previous described. The key difference of this study from other studies is that we

contain core of the caveolin protein (helix-break-helix motif). However, it is not feasible to examine the entire construct using NMR to obtain Cα information that can be used to analyze the secondary structure residue by residue using chemical shift indexing because of the degeneracy of the spectra. Therefore, we used circular dichroism spectroscopy to obtain global secondary structure information. Helical wheel analysis was also performed to investigate the C-terminal domain. Samples were prepared in a LMPG buffer similar to that used for analyzing the NMR samples but without salt in order to prevent light scattering. The previous chapter analyzing caveolin-1(96-136) show that 30 residues are α-helical (11 residues in helix-1 and 19 residues in helix-2) which is 73% α-helicity based on CSI. The circular dichorism data of caveolin-1(96-136) analyzed by CDSSTR method in Dichoweb show that 79% is α-helical which corresponds to 32 amino acid residues. In addition, caveolin-1(82-136) chemical shift indexing show that 40 residues are α-helical (helix-1 contains 21 residues and helix-2 contains 19 residues) which is 73% α-helicity. CD data analysis of caveolin-1(82-136) shows that 74% is α-helical which corresponds to 41 residues. From this comparison, the CD percentage values of α-helicity converted to the number of amino acids are well in the agreement with the number of amino acids analyzed by the chemical shift indexing method (approximately 1-2 amino acid error range). This indicates that CD analysis is accurate enough to obtain the number of amino acids that correspond to the α-helix. A series of the spectra obtained from different truncated constructs were analyzed using the CDSSTR algorithms (Dichroweb). The analyses indicate that approximately 79%of caveolin-1(96-136) is α-helical, caveolin-1(96-178) is 78% α-helical, caveolin-1(82-136) is 74% α-helical, and caveolin-1(62-136) is 72% α-helical (Figure 4-4). Therefore, caveolin-1 is mainly composed of α-helical secondary structure. To investigate 62 to 81 regions, CD data of caveolin-1(62-136) and CSI data of caveolin-1(82-136) were taken for comparison. From CSI analysis caveolin-1(82-

136) contains 40 amino acids that are α-helical. From CD data analysis caveolin-(62-136) has 72% α-helical which corresponds to 52 amino acids. The combination of these two results show that residues 62-81 contain 12 amino acids that have α-helical properties and 7 amino acids that have unstructured character. To figure out where this helix is, we made an assumption that proline at position 75 induces a break in the helix like the other prolines in the transmembrane domain. Therefore, a 12 amino acid long α-helix has to be before or after proline at position 75. But to place 12 amino acids α-helix after proline at 75 is not likely to occur. Because if the helix starts right after proline at position 75 the 12 amino acid long helix will end at residue 87. This means the caveolin-1 transmembrane domain helix-1 extend to residue 76 since transmembrane domain of helix-1 the starts at residue 87. Based on CSI, residues 82-86 are unstructured which suggests the 12 amino acid long helix cannot be after proline 75. Therefore, a 12 amino acid long helix will be before proline 75 which will create a helix starting at residue 62 or 63. Another possibility is a kinked helix (helix-break-helix motif) is present around proline at position 75 if the assumption is incorrect. However, because of limitations of CD spectroscopy it is difficult to address this question.

To investigate the C-terminal domain of this protein, caveolin-1(96-136) and caveolin-1(96-178) data were compared. The CSI analysis of caveolin-1(96-136) shows that 30 amino acids are α-helical. The CD data analysis of caveolin-1(96-178) shows that 78% are α-helical which mean 65 amino acids have α-helix properties. From this information we can conclude that approximately 35 amino acids are α-helical in the C-terminal domain of the protein (residue 137-178) and 5 amino acids are unstructured.

To pin point where this helix starts, analysis of primary sequence was performed using helical wheel analysis (Figure 4-5). Based on *in vivo* studies, there are three cysteine sites that are responsible for palmitoylation in the C-terminal domain at position 133, 143, and 156. To avoid the penalty of placing a palmitoyl group into the cytoplasm the three cysteine sites have to face membrane. However, the cysteine at position 143 is opposite to other cysteine sites that are not facing the membrane if the helix starts at residue 133 based on helical wheel analysis. Therefore, the helix has to start after residue 143 for all three cysteine sites because facing pamiltoylated cysteine sites towards hydrophilic environment is not energetically favorable. Moreover, caveolin-1(96-136) CSI shows that residues 130 to 136 are dynamic. Residues up to 143 must be unstructured to keep residue 143 facing membrane and this result is consistent with CD data, agreeing with number of helical residues. From this analysis, 7 amino acids are unstructured (5 unstructured amino acids ± 2 amino acids). Adding 7 amino acids after 136 will make unstructured region up to position 143. Therefore, C-terminal amphipathic helix will start after cysteine 143 and will go to the end of sequence. Using the combination of all this data, we can propose a model that contains the complete membrane interacting domain of caveolin-1 based on experimental data (Figure 4-6).

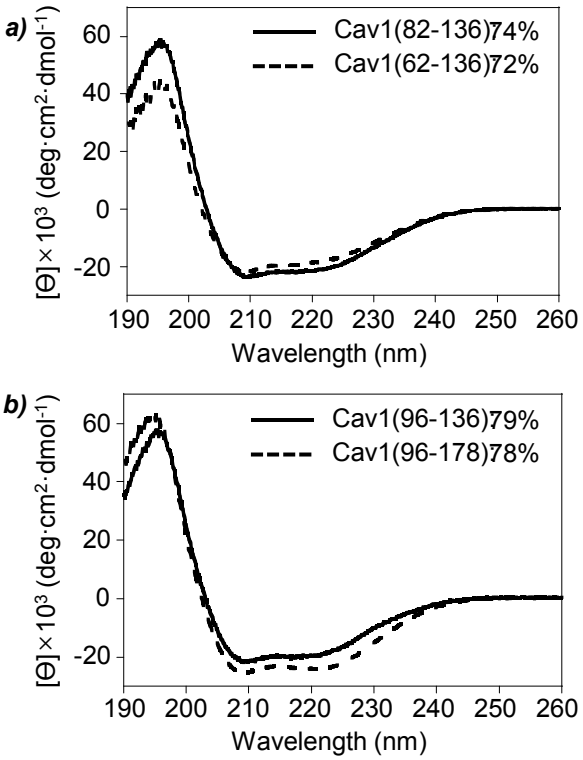


Figure 4-3. Circular dichroism spectra of various caveolin constructs in 100 mM LMPG; 20 mM phosphate pH 7.0. *a)* Caveolin-1(82-136) (solid line) and caveolin-1(62-136) (dashed line). *b)* Caveolin-1(96-136) (solid line) and caveolin-1(96-178) (dashed line).

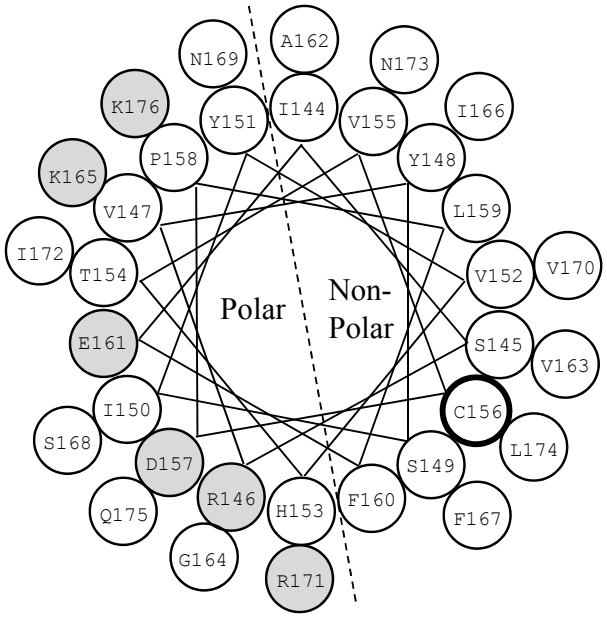


Figure 4-4. Helical wheel projection of caveolin-1(144-178). Charged residues are represented by shaded circles. Bolded circle represents cysteine 156 which is palmitoylated in vivo.

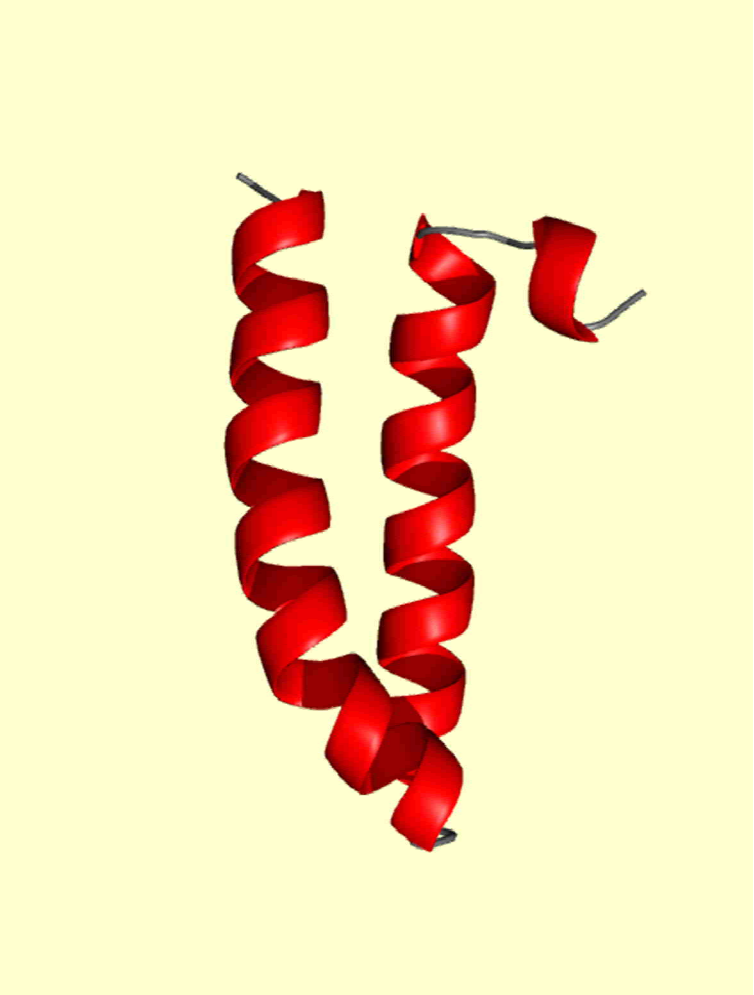


Figure 4-5. Preliminary model of caveolin-1(82-136) based on CS-ROSETTA.

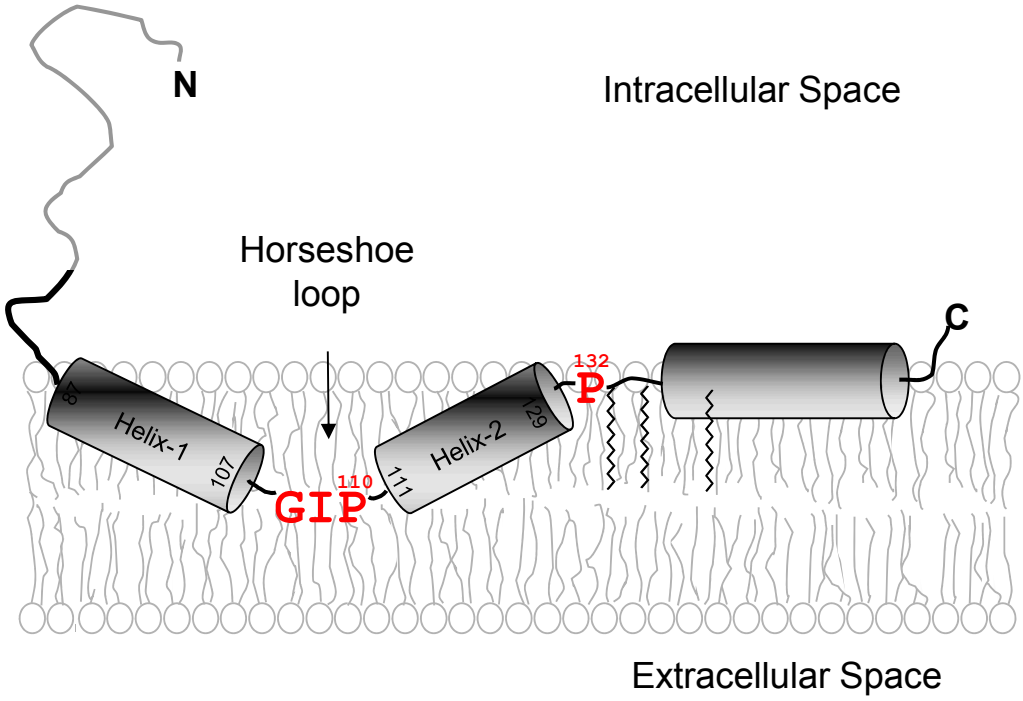


Figure 4-6. Topology model of caveolin-1 in a DMPC bilayer.

Conclusion

Approximately 50% of caveolin-1 is interacting or associated with the membrane. It is critical to elucidate the structure of the caveolin-1 membrane interacting domain. The structural analysis of the entire caveolin-1 membrane interacting domain is essential to understand role of caveolin-1. In this study we used CD and NMR spectroscopy to investigate the membrane interacting domain of caveolin-1. CSI indexing compiled from NMR data shows that the caveolin-1 transmembrane domain has a helix-break-helix motif and the first helix encompass residues 87-107 and the second helix encompass residues 111-129. Moreover, other putative membrane interacting helices are observed using the combination of CD and CSI. From this information we could propose a model based on experimental results. The transmembrane domain inserts into the lipid bilayer symmetrically and there is a long amphipathic helix in the c-terminal domain starting after residue 143. This is the first model caveolin-1 based on experiments and this model will provide better insights into the role of caveolin-1.

Chapter 5. Bicelles and Application

Abstract

Membrane mimics are critical to study membrane proteins in vitro. Until now, micelles and vesicles are the two most utilized systems to study membrane proteins. For solution NMR structural studies, micelles are the most common mimic to gather structural information since vesicles are huge and do not tumble fast enough. Although the micelle system is excellent to apply to solution NMR studies for membrane proteins, the drawback is that micelles do not have a planar bilayer region and the high curvature of micelle can create structural artifacts. Therefore, better systems are required to obtain more accurate structural information. The bicelles are discodial lipid aggregates composed of long and short chain lipids. The long chain lipid of bicelles creates planar bilayer and short chain lipids encapsulate the rim. The different ratio of long and short chain lipids (***q***) enables the use of bicelles to both solid and solution state NMR studies. Incorporation of various caveolin-1 constructs (especially 96-136) is performed to find out the best way to reconstitute membrane protein into bicelles such as preforming bicelles, modified ethanol injection, and PFOA vesicle methods to apply to solution NMR studies. In addition, other applications of bicelles are studied instead of their use as a membrane mimic for solution NMR studies. Periodic meso-porous material was developed using bicelles and micelles. Since bicelles have unique properties that micelles do not have like size tuning and magnetic field alignment they may be desirable in the generation of novel porous materials. The different lipid concentration bicelles are tested with ***q***=0.5 value. Moreover, bicelle and tetramethoxysilane (TMOS) concentration was varied to obtain porous materials. The

material obtained using bicelles had novel pillared lamellar structure with high surface areas.

Introduction

A major hindrance of studying membrane proteins in vitro has been a lack of "native-like" medium which can support structure, function, and activity. Most often and most widely, detergents are used to solubilize membrane proteins. However, detergents are not the best mimic of the membrane for biophysical studies for several reasons. First, detergents usually lack the characteristics of natural phospholipids which can bias structural determination. Second, the high curvature of micelles could create artifact in the structures. Lastly, micelles do not have true bilayer to support membrane proteins. Another heavily used membrane mimic is vesicles. Vesicles are used instead of micelles for membrane protein studies since they have a true bilayer and are therefore most similar to the cell membrane system. It is a good mimic to study functions but for structural studies such as NMR it is not adoptable because of the large size and lack of dynamic properties of vesicles. Therefore, a better mimic of the membrane is desirable to obtain more accurate structural information of membrane proteins. In particular, a bilayer with phospholipids would serve as a better environment for membrane proteins such as vesicles and sufficient fast reorientation rates like micelles to able to obtain high resolution solution NMR spectra for structural determination.

The bicelles have become promising model for structural studies using NMR and other functional studies. A bicelle is a discoidal lipid aggregate composed of long chain phospholipids and detergent or short chain phospholipids (Figure 5-1). The features of bicelles have a combination of advantages from both vesicles and micelle system. Bicelles have a true bilayer like vesicles and dynamic properties like micelles. The structure of bicelles have a central bilayer formed by the long-chain phospholipid and are encapsulated by a ring of detergent or short chain lipids that shields the long-

chain lipid tails from water. The central bilayer region is most commonly composed of DMPC and the rim region is composed of either a bile-salt derivative such as CHAPSO or a short-chain phospholipid such as DHPC (33, 34, 35, 36, 37, 38, 39). There are several advantages in bicelles. First, the thickness can be changed by using other chain length phospholipids instead of DMPC. Second, the surface charge can be changed. The bilayer region can be doped with phospholipids that have identical chain lengths but different head groups in order to alter the charge characteristics of the membrane. Lastly, important advantage of bicelles is that their size and properties can be varied by altering both the total lipid concentration and the molar ration (q) of DMPC to that of DHPC. For most solution phase NMR experiments, bicelles with a q value of 0.5 are utilized. At this q the bicelles still have an appreciable bilayer region (82 A diameters) to accommodate full-length proteins and they tumble fast to have sufficient reorientation times.

The attention to periodic meso-porous materials is growing fast because of variety applications such as catalysis, separation, chemical storage, and delivery (89, 90). The most widely employed method is templation of micelles especially using cationic or non-ionic detergents (91). The new micelle system was a driving force of the fast growth of this area in material science because different micelle systems produce different meso-porous materials. However, there is a limitation using micelle systems like similar pore size because pore size is dependent on the size of micelles. To circumvent this problem bicelles systems are applied to obtain new periodic meso-porous materials. The ability of tuning size of bicelles and the altering charges of head group will open up new perspective for templating.

In this chapter, various method of incorporation of membrane protein especially caveolin-1(96-136) was performed to find out best way to incorporate membrane

proteins into bicelles. Moreover, investigation of application of bicelles for templation of periodic meso-porous materials will be introduced.

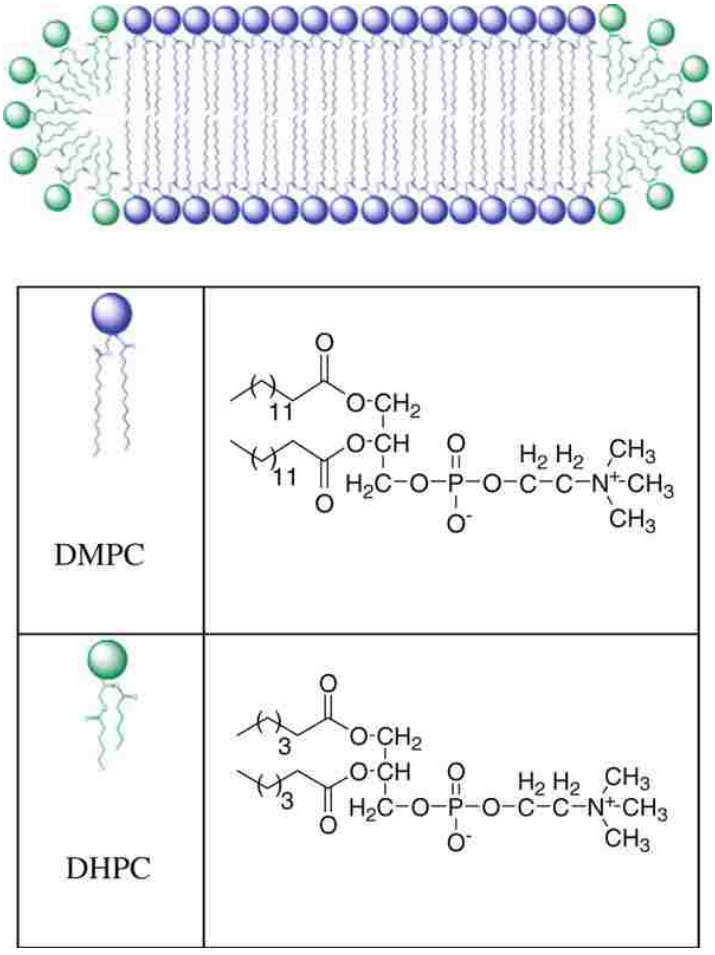


Figure 5-1. Schemetic representation of DMPC bicelles with DHPC rim.

Material/ Methods.

Preparation and incorporation of membrane protein into bicelles

Modified Ethanol injection method.

38.5 mg of DMPC and 3mg of Caveolin-1(96-136) were dissolved into 2.84 mL TFE. Inject into 50 mL of water quickly and freeze immediately. The sample was lyophilized to achieve a nice fluffy powder. The sample was then vortexed vigorously to obtain milky homogeneous solution in 310 µL of 20 mM phosphate buffer. 200 µL of 25% (w/w) DHPC was added while mixing to form clear bicelles solution.

Co-lyophilization method

38.5mg of DMPC and 3mg of Caveolin-1(96-136) were dissolved in 5mL of 50% HFIP and lyophilized immediately. The sample was vortexed until milk-like suspension is obtained in 310 µL of 20 mM phosphate buffer. 200 µL of 25% (w/w) DHPC was slowly added into resuspended DMPC while mixing. After an extra one minute of voltexing, clear bicelle solution is obtained.

Preforming bicelle method

38.5 mg of DMPC was resuspended with 310 µL of 20 mM phosphate buffer water and voltexed until homogeneous milk-like suspension is obtained. 200 µL of 25% (w/w) DHPC was slowly added into resuspended DMPC while mixing. After an extra one minute of voltexing, clear bicelle solution is obtained. The solution was spun for one hour at 20,000xg to remove bubbles from the solution. 3 mg of lyophilized caveolin-1(96-136) was added to preformed bicelles.

PFOA dialysis method

38.5 mg of DMPC and 3 mg of caveolin-1(96-136) were dissolved in 2.84 mL of PFOA buffer (300 mM PFOA and 10 mM Tris, pH 8). Place dissolved sample in hot water until solution turns clear. Load into 10,000 molecular weight cut off dialysis cassette and dialyze against 2-liter of 10 mM ammonium sulfate, pH 8. Dialysis was performed for 3 days with every 12 hour buffer exchange and final exchange was performed with NMR buffer condition (100 mM NaCl and 20 mM phosphate, pH 7). 200 µL of 25% DHPC was added and concentrated to obtain proper volume.

Templating for meso-porous materials

1 mg of ammonium fluoride was dissolved in bicelle solution and added 218 mg of the tetramethoxysilane (TMOS). During the addition of the tetramethoxysilane, the bicelle solution was continuously sonicated at room temperature. After the addition of the TMOS a white precipitate formed immediately. The mixture was sonicated for another 2 min and the precipitate was filtered off, washed with acetone, and vacuum-dried at room temperature (92).

Result and Discussion

Bicelles are a mimic of the membrane that has both advantages of micelles and vesicles. It has the dynamic properties of micelles and a planar bilayer region like vesicles. Bicelles are a great medium to study protein lipid interactions in the solution NMR state. In addition, bicelles can adopt cholesterol up to 20% which will be useful to study membrane proteins that have interactions with cholesterol. However, it is not as commonly used to perform membrane protein structural studies compared to micelles because it is not easy to incorporated membrane proteins into bicelles especially at the concentration for NMR experiments. In addition most research has found that there is little difference in secondary structure of protein when it is observed in micelles and bicelles. However, there is no doubt that bicelles are a better mimic of the membrane and further studies can be only achieved using bicelles such as protein-lipid NOEs. Therefore, incorporation of caveolin-1(96-136) constructs was performed to optimize incorporation methods. First, a preforming bicelle method was performed. Bicelles are prepared by mixing DMPC and DHPC with buffer and loaded into micro centrifuge tubes that contained lyophilized caveolin-1(96-136). As caveolin-1(96-136) was incorporated high concentration to achieve NMR concentration precipitation usually occurred. This may result of caveolin-1(96-136)'s extremely hydrophobic nature. It is possible that some of caveolin-1(96-136) interacts with each other before protein have chance to interact with lipids. Thus other methods are employed to address this problem. The modified ethanol injection method was performed using TFE instead of ethanol which is a well-established method to obtain unilamellar vesicles. To make sure caveolin-1(96-136) was well mixed with lipid to ensure chances of interacting with lipids are higher than aggregation process. A homogeneous milky suspension was obtained after resuspension with phosphate buffer but still some precipitated out when

DHPC was added. This may be because TFE is not strong enough to break down aggregation in the first place. Therefore, co-lyophilizing caveolin-1(96-136) with DMPC lipid using the HFIP method was utilized. However, precipitation was still observed as increment of concentration. This result lead us to believe it may be that lyophilization was too harsh for the vesicle system with caveolin-1(96-136). Vesicles can burst when lyophilization is taking a place to remove water inside and the protein is not situated properly in the vesicles and when hydrated with buffer and some protein which is not in the right conformation will precipitate out. Therefore, we adapted the method forming vesicles with PFOA. PFOA is perfluorinated detergent which is powerful enough to solubilize inclusion bodies. The caveolin-1(96-136) and DMPC was dissolved in PFOA buffer and dialyzed a couple of days with extensive buffer exchanges to form vesicles in the most gentle way. Most of the PFOA was removed in the first 12 hours and vesicles were formed within 3 days. After obtaining vesicles DHPC was added to form bicelles and the precipitation was not observed immediately. But the problem was the volume of sample was higher than we needed. Therefore, a method to keep volume suitable for NMR experiments was required. A concentration method was performed to obtain required volume size but tremendous protein loss was observed after concentration. Therefore, the drying down method was performed but the protein started precipitating out as soon as the required volume is reached. From these experiments, the incorporation of caveolin-1(96-136) construct into bicelles for CD studies was achieved but not for NMR studies. As caveolin-1(96-136) was titrated into bicelles system using various methods it was clear that bicelles can handle up to a certain concentration but it is not high enough for NMR experiments. Tuning the bicelle size was also performed to check preference of size of bicelles for caveolin-1(96-136). The sizes were q=0.2 to 1.0 with 15% lipid concentration. Also the concentrations of bicelles are brought up to 20% to check concentration of lipid will help incorporation.

However, the incorporation of caveolin-1(96-136) in bicelles suitable for NMR studies was not achieved. More optimization is required to obtain high concentration of membrane protein into bicelles such as changing the length of long chain lipids or charge of lipid head group.

Instead of using bicelles as mimic of membrane, bicelles were adopted for the use as a template of periodic mesoporous materials. Until now, micelle systems were the most commonly used template for periodic mesoporous materials. However, using bicelles as template has several advantages because bicelles can tune the size of pores. For these studies, different concentrations of bicelles are utilized, 5%, 12%, and 20%. Also the ratio of TMOS and bicelles was varied to investigate effect of formation of meso-porous materials. From data the ratio between TMOS and bicelles is critical to form novel hexagonal pillared lamellar structure or non-pillared lamellar structure (92). Higher concentration of TMOS lead to novel hexagonal pillared lamellar structure (Figure 5-2). The meso-porous material that is created by using 5% bicelles concentration with higher concentration of TMOS for approximately 2 minutes of reaction time showed that the material has a very high surface area. These data suggest that bicelles can be adapted to other applications and not only to membrane protein research.

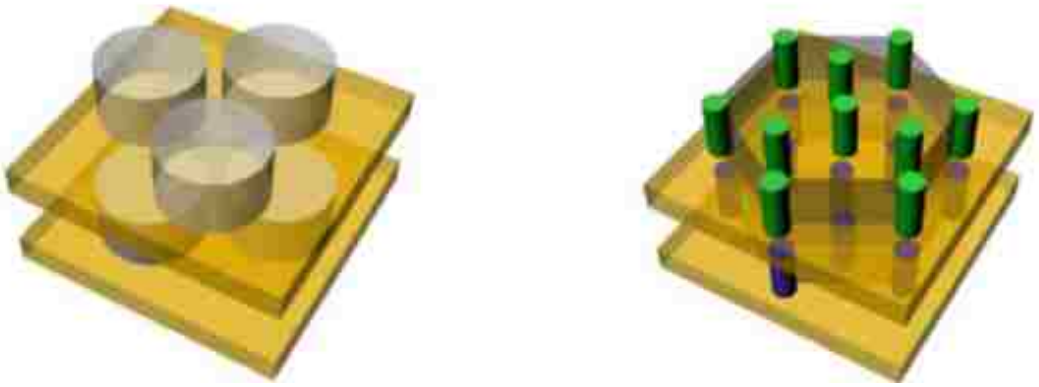


Figure 5-2. Schemetic representation of bicelle templated silica structure.

Grey: bicelles, Orange: silica layers, Green and blue: silica pillars.

Conclusion

Bicelles have been characterized and applied as a membrane mimic for membrane protein studies. In particular, the advantages of bicelles are desirable for membrane protein NMR. The magnetic alignment property of high ***q*** bicelles facilitates the application of bicelles to solid state NMR structural determination. The fast tumbling property of low ***q*** bicelles enables bicelles to be adapted in solution state NMR structural determinations. Therefore, various methods of reconstituting caveolin-1 constructs into bicelles were performed but none of the methods are able to reach solution NMR concentration (approximately 0.5-1.0 mM concentration). Because the function of caveolin-1 is proposed to curve cell membrane, it may destabilize bicelles at high concentration. More optimization is needed to incorporate caveolin-1 constructs into bicelles such as changing the head group of lipids to change surface charges or change the lipid tail length to find the optimum length for proteins. Although incorporation of caveolin-1 constructs into bicelles is not easy they still have the potential to be used in place of micelle systems for solution NMR studies. Bicelles can also be applied to other studies that use micelle systems as a membrane mimic. The formation of periodic meso-porous materal was performed and displayed novel hexagonal pillared lamellar structure. The advantage of tuning the size of the bicelles could open more possibilities of different size porous materials. Moreover, the magnetic alignment property of bicelles is alluring to add new perspective to templating such as magnetic field directed templating. Perspective porous material applying bicelles for templating can be applied in the biophysical studies, and may serve as new medium for membrane protein alignment for solution NMR studies to obtain long range restraints.

References

1. Stan, R. V. (2005) Structure of Caveolae. *Biochimica et Biophysica Acta. 1746*, 334-348.

2. Lisanti, M. P., and Williams, T. M. (2004) The Caveolin Genes: From Biology to Medicine. *Annual Medicine. 36*, 584-595.

3. Krajewska, W. M., and Maslowska, I. (2004) Caveolins: Structure and Function in Signal Transduction. *Cellular & Molecular Biology Letters. 9*, 195-220.

4. Williams, T. M., and Lisanti, M. P. (2004) The Caveolin Proteins. *Genome Biology. 5*.

5. Williams, T. M., and Lisanti, M. P. (2004) The Caveolin Genes: From Cell Biology to Medicine. *Annals of Medicine. 36*, 584-595.

6. Spisni, E., Tomasi, V., Cestaro, A., and Tosatto, S. C. (2005) Structural Insights into the Function of Human Caveolin 1. *Biochem. Biophys. Res. Commun. 338*, 1383-1390.

7. Chao, W. T., Fan, S. S., Chen, J. K., and Yang, V. C. (2003) Visualizing Caveolin-1 and HDL in Cholesterol-Loaded Aortic Endothelial Cells. *Journal of lipid research. 44*, 1094-9.

8. Frank, P. G., Cheung, M. W. C., Pavlides, S., Llaverias, G., Park, D. S., and Lisanti, M. P. (2006) Caveolin-1 and Regulation of Cellular Cholesterol Homeostasis. *American Journal of Physiology-Heart and Circulatory Physiology. 291*, H677-H686.

9. Drab, M., Verkade, P., Elger, M., Kasper, M., Lohn, M., Lauterbach, B., Menne, J., Lindschau, C., Mende, F., Luft, F. C., Schedl, A., Haller, H., and Kurzchalia, T. V. (2001) Loss of Caveolae, Vascular Dysfunction, and Pulmonary Defects in Caveolin-1 Gene-Disrupted Mice. *Science. 293*, 2449-2452.

10. Karam, J. A., Lotan, Y., Roehrborn, C. G., Ashfaq, R., Karakiewicz, P. I., and Shariat, S. F. (2007) Caveolin-1 Overexpression is Associated with Aggressive Prostate Cancer Recurrence. *The Prostate. 67*, 614-622.

11. Xia, H., Khalil, W., Kahm, J., Jessurun, J., Kleidon, J., and Henke, C. A. (2010) Pathologic Caveolin-1 Regulation of PTEN in Idiopathic Pulmonary Fibrosis. *Am. J. Pathol. 176*, 2626-2637.

12. Ohsawa, Y., Toko, H., Katsura, M., Morimoto, K., Yamada, H., Ichikawa, Y., Murakami, T., Ohkuma, S., Komuro, I., and Sunada, Y. (2004) Overexpression of P104L Mutant Caveolin-3 in Mice Develops Hypertrophic Cardiomyopathy with Enhanced Contractility in Association with Increased Endothelial Nitric Oxide Synthase Activity. *Human Molecular Genetics. 13*, 151-157.

13. Lee, H., Park, D. S., Razani, B., Russell, R. G., Pestell, R. G., and Lisanti, M. P. (2002) Caveolin-1 Mutations (P132L and Null) and the Pathogenesis of Breast Cancer - Caveolin-1 (P132L) Behaves in a Dominant-Negative Manner and Caveolin-1 (-/-) Null Mice show Mammary Epithelial Cell Hyperplasia. *American Journal of Pathology. 161*, 1357-1369.

14. Smythe, G. M., Eby, J. C., Disatnik, M. H., and Rando, T. A. (2003) A Caveolin-3 Mutant that Causes Limb Girdle Muscular Dystrophy Type 1C Disrupts Src Localization and Activity and Induces Apoptosis in Skeletal Myotubes. *J Cell Sci. 116*, 4739-49.

15. Schlegel, A., and Lisanti, M. P. (2000) A Molecular Dissection of Caveolin-1 Membrane Attachment and Oligomerization - Two Separate Regions of the Caveolin-1 C-Terminal Domain Mediate Membrane Binding and oligomer/oligomer Interactions *in vivo*. *Journal of Biological Chemistry. 275*, 21605-21617.

16. Das, K., Lewis, R. Y., Scherer, P. E., and Lisanti, M. P. (1999) The Membrane-Spanning Domains of Caveolins-1 and -2 Mediate the Formation of Caveolin Hetero-Oligomers. *Journal of Biological Chemistry. 274*, 18721-18728.

17. Fernandez, I., Ying, Y., Albanesi, J., and Anderson, R. G. (2002) Mechanism of Caveolin Filament Assembly. *Proc. Natl. Acad. Sci. U. S. A. 99*, 11193-11198.

18. Dietzen, D. J., Hastings, W. R., and Lublin, D. M. (1995) Caveolin is Palmitoylated on Multiple Cysteine Residues. Palmitoylation is Not Necessary for Localization of Caveolin to Caveolae. *J. Biol. Chem. 270*, 6838-6842.

19. Uittenbogaard, A., and Smart, E. J. (2000) Palmitoylation of Caveolin-1 is Required for Cholesterol Binding, Chaperone Complex Formation, and Rapid Transport of Cholesterol to Caveolae. *J. Biol. Chem. 275*, 25595-25599.

20. Monier, S., Parton, R. G., Vogel, F., Behlke, J., Henske, A., and Kurzchalia, T. V. (1995) VIP21-Caveolin, a Membrane Protein Constituent of the Caveolar Coat, Oligomerizes *in vivo* and in Vitro. *Mol. Biol. Cell. 6*, 911-927.

21. Spisni, E., Tomasi, V., Cestaro, A., and Tosatto, S. C. E. (2005) Structural Insites into the Function of Human Caveolin 1. *Biochemical and Biophysical Research Communications. 338*, 1383-1390.

22. Scherer, P. E., Tang, Z. L., Chun, M. Y., Sargiacomo, M., Lodish, H. F., and Lisanti, M. P. (1995) Caveolin Isoforms Differ in their N-Terminal Protein-Sequence and Subcellular-Distribution - Identification and Epitope Mapping of an Isoform-Specific Monoclonal-Antibody Probe. *Journal of Biological Chemistry. 270*, 16395-16401.

23. Kurzchalia, T. V., Dupree, P., Parton, R. G., Kellner, R., Virta, H., Lehnert, M., and Simons, K. (1992) VIP21, a 21-kD Membrane Protein is an Integral Component of Trans-Golgi-Network-Derived Transport Vesicles. *J. Cell Biol. 118*, 1003-1014.

24. Glenney, J. R., and Soppet, D. (1992) Sequence and Expression of Caveolin, a Protein Component of Caveolae Plasma Membrane Domains Phosphorylated on Tyrosine in Rous Sarcoma Virus-Transformed Fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America. 89*, 10517-10521.

25. Parton, R. G., Hanzal-Bayer, M., and Hancock, J. F. (2006) Biogenesis of Caveolae: A Structural Model for Caveolin-Induced Domain Formation. *J. Cell. Sci. 119*, 787-796.

26. Aoki, S., Thomas, A., Decaffmeyer, M., Brasseur, R., and Epand, R. M. (2010) The Role of Proline in the Membrane Re-Entrant Helix of Caveolin-1. *J. Biol. Chem.*

27. Koike, S., Kodera, Y., Nakao, A., Iwata, H., and Yatabe, Y. (2010) Absence of the Caveolin-1 P132L Mutation in Cancers of the Breast and Other Organs. *J. Mol. Diagn. 12*, 712-717.

28. Rieth, M. D., Lee, J., and Glover, K. J. (2012) Probing the Caveolin-1 P132L Mutant: Critical Insights into its Oligomeric Behavior and Structure. *Biochemistry.*

29. Bonuccelli, G., Casimiro, M. C., Sotgia, F., Wang, C., Liu, M., Katiyar, S., Zhou, J., Dew, E., Capozza, F., Daumer, K. M., Minetti, C., Milliman, J. N., Alpy, F., Rio, M. C., Tomasetto, C., Mercier, I., Flomenberg, N., Frank, P. G., Pestell, R. G., and Lisanti, M. P. (2009) Caveolin-1 (P132L), a Common Breast Cancer Mutation, Confers Mammary Cell Invasiveness and Defines a Novel Stem cell/metastasis-Associated Gene Signature. *Am. J. Pathol. 174*, 1650-1662.

30. Shatz, M., Lustig, G., Reich, R., and Liscovitch, M. (2010) Caveolin-1 Mutants P132L and Y14F are Dominant Negative Regulators of Invasion, Migration and Aggregation in H1299 Lung Cancer Cells. *Exp. Cell Res. 316*, 1748-1762.

31. Mercier, I., Bryant, K. G., Sotgia, F., Bonuccelli, G., Witkiewicz, A. K., Dasgupta, A., Jasmin, J. F., Pestell, R. G., and Lisanti, M. P. (2009) Using Caveolin-1 Epithelial Immunostaining Patterns to Stratify Human Breast Cancer Patients and Predict the Caveolin-1 (P132L) Mutation. *Cell. Cycle. 8*, 1396-1401.

32. Van Horn, W. D., Ogilvie, M. E., and Flynn, P. F. (2008) Use of Reverse Micelles in Membrane Protein Structural Biology. *Journal of Biological NMR. 40*, 203-211.

33. Houdai, T., Matsumori, N., and Murata, M. (2008) Structure of Membrane-Bound Amphidinol 3 in Isotropic Small Bicelles. *Org Lett. 10*, 4191-4.

34. Lau, T. L., Partridge, A. W., Ginsberg, M. H., and Ulmer, T. S. (2008) Structure of the Integrin beta3 Transmembrane Segment in Phospholipid Bicelles and Detergent Micelles. *Biochemistry. 47*, 4008-16.

35. Glover, K. J., Whiles, J. A., Vold, R. R., and Melacini, G. (2002) Position of Residues in Transmembrane Peptides with Respect to the Lipid Bilayer: A Combined Lipid NOEs and Water Chemical Exchange Approach in Phospholipid Bicelles. *Journal of Biomolecular Nmr. 22*, 57-64.

36. Whiles, J. A., Glover, K. J., Vold, R. R., and Komives, E. A. (2002) Methods for Studying Transmembrane Peptides in Bicelles: Consequences of Hydrophobic Mismatch and Peptide Sequence. *Journal of Magnetic Resonance. 158*, 149-156.

37. Glover, K. J., Whiles, J. A., Wu, G., Yu, N., Deems, R., Struppe, J. O., Stark, R. E., Komives, E. A., and Vold, R. R. (2001) Structural Evaluation of Phospholipid Bicelles

for Solution-State Studies of Membrane-Associated Biomolecules. *Biophys. J. 81*, 2163-71.

38. Whiles, J. A., Brasseur, R., Glover, K. J., Melacini, G., Komives, E. A., and Vold, R. R. (2001) Orientation and Effects of Mastoparan X on Phospholipid Bicelles. *Biophysical Journal. 80*, 280-293.

39. Glover, K. J., Wood, M. J., Vold, R. R., and Komives, E. A. (2000) Analysis of a Transmembrane Peptide: Prion(110-137) in Bicelles. *Biophysical Journal. 78*, 158A-158A.

40. Le Lan, C., Neumann, J. M., and Jamin, N. (2006) Role of the Membrane Interface on the Conformation of the Caveolin Scaffolding Domain: A CD and NMR Study. *FEBS Lett. 580*, 5301-5305.

41. Loudet, U., Khemtemourian, L., Aussenac, F., Gineste, S., Achard, M. F., and Dufourc, E. J. (2005) Bicelle Membranes and their use for Hydrophobic Peptide Studies by Circular Dichroism and Solid State NMR. *Biochimica Et Biophysica Acta-General Subjects. 1724*, 315-323.

42. Goetz, M., Carlotti, C., Bontems, F., and Dufourc, E. J. (2001) Evidence for an Alpha-Helix -> Pi-Bulge Helicity Modulation for the neu/erbB-2 Membrane-Spanning Segment. A H-1 NMR and Circular Dichroism Study. *Biochemistry. 40*, 6534-6540.

43. Diefenderfer, C., Lee, J., Mlyanarski, S., Guo, Y., and Glover, K. J. (2009) Reliable Expression and Purification of Highly Insoluble Transmembrane Domains. *Anal Biochem. 384*, 274-8.

44. Lee, J., and Glover, K. J. (2012) The Transmembrane Domain of Caveolin-1 Exhibits a Helix-Break-Helix Structure. *Biochim. Biophys. Acta.*

45. Whitmore, L., and Wallace, B. (2004) DICHROWEB, an Online Server for Protein Secondary Structure Analyses from Circular Dichroism Spectroscopic Data RID C-5651-2008 RID C-3753-2008. *Nucleic Acids Res. 32*, W668-W673.

46. Lobley, A., Whitmore, L., and Wallace, B. (2002) DICHROWEB: An Interactive Website for the Analysis of Protein Secondary Structure from Circular Dichroism Spectra RID C-5651-2008 RID C-3753-2008. *Bioinformatics. 18*, 211-212.

47. Lobley, A., and Wallace, B. (2001) Dichroweb: A Website for the Analysis of Protein Secondary Structure from Circular Dichroism Spectra. RID C-3753-2008. *Biophys. J. 80*, 373A-373A.

48. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR. 6*, 277-293.

49. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: A Hybrid Method for Predicting Protein Backbone Torsion Angles from NMR Chemical Shifts. *J. Biomol. NMR. 44*, 213-223.

50. Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., and MacKinnon, R. (1998) The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science. 280*, 69-77.

51. Noskov, S. Y., and Roux, B. (2006) Ion Selectivity in Potassium Channels. *Biophys. Chem. 124*, 279-291.

52. Tao, X., Avalos, J. L., Chen, J., and MacKinnon, R. (2009) Crystal Structure of the Eukaryotic Strong Inward-Rectifier K+ Channel Kir2.2 at 3.1 A Resolution. *Science. 326*, 1668-1674.

53. Molina, D. M., Cornvik, T., Eshaghi, S., Haeggstrom, J. Z., Norland, P., and Sabet, M. I. (2008) Engineering Membrane Protein Overproduction in Escherichia Coli. *Protein Science. 17*, 673-680.

54. Studier, F. W. (2005) Protein Production by Auto-Induction in High Density Shaking Cultures. *Protein Expr Purif. 41*, 207-34.

55. Wang, L., Liu, D., and Reeves, P. R. (1996) C-Terminal Half of Salmonella Enterica WbaP (RfbP) is the Galactosyl-1-Phosphate Transferase Domain Catalyzing the First Step of O-Antigen Synthesis. *J. Bacteriol. 178*, 2598-604.

56. Wang, L., and Reeves, P. R. (1994) Involvement of the Galactosyl-1-Phosphate Transferase Encoded by the Salmonella Enterica rfbP Gene in O-Antigen Subunit Processing. *Journal of bacteriology. 176*, 4348-56.

57. Studier, F. W. (2005) Protein Production by Auto-Induction in High Density Shaking Cultures. *Protein expression and purification. 41*, 207-34.

58. Marley, J., Lu, M., and Bracken, C. (2001) A Method for Efficient Isotopic Labeling of Recombinant Proteins RID A-7811-2009. *J. Biomol. NMR. 20*, 71-75.

59. Truhlar, S. M., Cervantes, C. F., Torpey, J. W., Kjaergaard, M., and Komives, E. A. (2008) Rapid Mass Spectrometric Analysis of 15N-Leu Incorporation Fidelity during Preparation of Specifically Labeled NMR Samples. *Protein Sci. 17*, 1636-1639.

60. Ramjeesingh, M., Huan, L. J., Garami, E., and Bear, C. E. (1999) Novel Method for Evaluation of the Oligomeric Structure of Membrane Proteins. *Biochem J. 342 (Pt 1)*, 119-23.

61. Zhou, L., Lai, Z. T., Lu, M. K., Gong, X. G., and Xie, Y. (2008) Expression and Hydroxylamine Cleavage of Thymosin Alpha 1 Concatemer. *J Biomed Biotechnol. 2008*, 736060.

62. Heung-Bok Park, Sang-hyun Pyo, Seung-Suh Hong,Jin-Hyun Kim. (2001) Optimization of the Hydroxylamine Cleavage of an Expressed Fusion Protein to Produce a Recombinant Antimicrobial Peptide. *Biotechnology letters. 23*, 637-641.

63. Das, K., Lewis, R. Y., Scherer, P. E., and Lisanti, M. P. (1999) The Membrane-Spanning Domains of Caveolins-1 and -2 Mediate the Formation of Caveolin Hetero-Oligomers. Implications for the Assembly of Caveolae Membranes *in vivo*. *J. Biol. Chem. 274*, 18721-8.

64. Schlegel, A., Schwab, R. B., Scherer, P. E., and Lisanti, M. P. (1999) A Role for the Caveolin Scaffolding Domain in Mediating the Membrane Attachment of Caveolin-1 - the Caveolin Scaffolding Domain is both Necessary and Sufficient for Membrane Binding in Vitro. *Journal of Biological Chemistry. 274*, 22660-22667.

65. Mori, M., Ishikawa, G., Takeshita, T., Goto, T., Robinson, J. M., and Takizawa, T. (2006) Ultrahigh-Resolution Immunofluorescence Microscopy using Ultrathin Cryosections: Subeellular Distribution of Caveolin-1 Alpha and CD31 in Human Placental Endothelial Cells. *Journal of Electron Microscopy. 55*, 107-112.

66. Li, T., Sotgia, F., Vuolo, M. A., Li, M., Yang, W. C., Pestell, R. G., Sparano, J. A., and Lisanti, M. P. (2006) Caveolin-1 Mutations in Human Breast Cancer: Functional Association with Estrogen Receptor Alpha-Positive Status. *Am. J. Pathol. 168*, 1998-2013.

67. Li, T., Sparano, J. A., Sotgia, F., Vuolo, M. A., Li, M., Yang, W., Negassa, A., and Lisanti, M. P. (2006) Caveolin-1 Mutations are Associated with Relapse in Estrogen Receptor-Positive Breast Cancer and Sensitivity to Estrogen. *Breast Cancer Research and Treatment. 100*, S36-S36.

68. Li, T. H., Sotgia, F., Vuolo, M. A., Lo, M. M., Yang, W. C., Pestell, R. G., Sparano, J. A., and Lisanti, M. P. (2006) Caveolin-1 Mutations in Human Breast Cancer - Functional Association with Estrogen Receptor Alpha-Positive Status. *American Journal of Pathology. 168*, 1998-2013.

69. Pinilla, S. M. R., Honrado, E., Hardisson, D., Benitez, J., and Palacios, J. (2006) Caveolin-1 Expression is Associated with a Basal-Like Phenotype in Sporadic and Hereditary Breast Cancer. *Breast Cancer Research and Treatment. 99*, 85-90.

70. Rodriguez, O., Fricke, S., Chien, C., Dettin, L., VanMeter, J., Shapiro, E., Dai, H. N., Casimiro, M., Ileva, L., Dagata, J., Johnson, M. D., Lisanti, M. P., Koretsky, A., and Albanese, C. (2006) Contrast-Enhanced *in vivo* Imaging of Breast and Prostate Cancer Cells by MRI. *Cell Cycle. 5*, 113-119.

71. Sotgia, F., Rui, H., Bonuccelli, G., Mercier, I., Pestell, R. G., and Lisanti, M. P. (2006) Caveolin-1, Mammary Stem Cells, and Estrogen-Dependent Breast Cancers. *Cancer Research. 66*, 10647-10651.

72. Bouras, T., Lisanti, M. P., and Pestell, R. G. (2004) Caveolin-1 in Breast Cancer. *Cancer Biology & Therapy. 3*, 931-941.

73. Sloan, E. K., Stanley, K. L., and Anderson, R. L. (2004) Caveolin-1 Inhibits Breast Cancer Growth and Metastasis. *Oncogene. 23*, 7893-7897.

74. Goddard, T. D., and Kneller, D. G. *SPARKY 3.* University of California, San Francisco.

75. WISHART, D., and SYKES, B. (1994) The C-13 Chemical-Shift Index - a Simple Method for the Identification of Protein Secondary Structure using C-13 Chemical-Shift Data. *J. Biomol. NMR. 4*, 171-180.

76. Shen, Y., Vernon, R., Baker, D., and Bax, A. (2009) De Novo Protein Structure Generation from Incomplete Chemical Shift Assignments. *J. Biomol. NMR. 43*, 63-78.

77. Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D., and Bax, A. (2008) Consistent Blind Protein Structure Generation from NMR Chemical Shift Data. *Proc. Natl. Acad. Sci. U. S. A. 105*, 4685-4690.

78. Schwieters, C. D., Kuszewski, J. J., and Clore, G. M. (2006) Using Xplor-NIH for NMR Molecular Structure Determination. *Prog Nucl Magn Reson Spectrosc. 48*, 47-62.

79. Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR Molecular Structure Determination Package. *J. Magn. Reson. 160*, 65-73.

80. Engelman, J. A., Zhang, X. L., Galbiati, F., Volonte, D., Sotgia, F., Pestell, R. G., Minetti, C., Scherer, P. E., Okamoto, T., and Lisanti, M. P. (1998) Molecular Genetics of the Caveolin Gene Family: Implications for Human Cancers, Diabetes, Alzheimer Disease, and Muscular Dystrophy. *American Journal of Human Genetics. 63*, 1578-1587.

81. Nishiyama, K., Trapp, B. D., Ikezu, T., Ransohoff, R. M., Tomita, T., Iwatsubo, T., Kanazawa, I., Hsiao, K. K., Lisanti, M. P., and Okamoto, T. (1999) Caveolin-3

Upregulation Activates Beta-Secretase-Mediated Cleavage of the Amyloid Precursor Protein in Alzheimer's Disease. *Journal of Neuroscience. 19*, 6538-6548.

82. Couchoux, H., Allard, B., Legrand, C., Jacquemond, V., and Berthier, C. (2007) Loss of Caveolin-3 Induced by the Dystrophy-Associated P104L Mutation Impairs L-Type Calcium Channel Function in Mouse Skeletal Muscle Cells. *Journal of Physiology-London. 580*, 745-754.

83. Hue-Beauvais, C., Pechoux, C., Bouguyon, E., Chat, S., Truchet, S., Pauloin, A., Le Gouar, Y., and Ollivier-Bousquet, M. (2007) Localisation of Caveolin in Mammary Tissue Depends on Cell Type. *Cell and Tissue Research. 328*, 521-536.

84. Pani, B., and Singh, B. B. (2009) Lipid rafts/caveolae as Microdomains of Calcium Signaling. *Cell Calcium. 45*, 625-633.

85. Calaghan, S., Kozera, L., and White, E. (2008) Compartmentalisation of cAMP-Dependent Signalling by Caveolae in the Adult Cardiac Myocyte. *Journal of Molecular and Cellular Cardiology. 45*, 88-92.

86. Luoma, J. I., Boulware, M. I., and Mermelstein, P. G. (2008) Caveolin Proteins and Estrogen Signaling in the Brain. *Molecular and Cellular Endocrinology.*

87. Patel, H. H., Murray, F., and Insel, P. A. (2008) Caveolae as Organizers of Pharmacologically Relevant Signal Transduction Molecules. *Annu. Rev. Pharmacol. Toxicol. 48*, 351-391.

88. Quest, A. F. G., Gutierrez-Pajares, J. L., and Torres, V. A. (2008) Caveolin-1: An Ambiguous Partner in Cell Signalling and Cancer. *Journal of Cellular Molecular Medicine. 12*, 1130-1150.

89. Ma, X., Wang, X., and Song, C. *"Molecular Basket" Sorbents for Separation of CO2 and H2S from various Gas Streams.* - American Chemical Society, .

90. Dubey, A., Mishra, B. G., and Sachdev, D. (2008) RETRACTED: Catalytic Applications of Ordered Mesoporous Magnesium Oxide Synthesized by Mesoporous Carbon. *Applied Catalysis A: General. 338*, 20-26.

91. Cao, L., Man, T., and Kruk, M. *Synthesis of Ultra-Large-Pore SBA-15 Silica with Two-Dimensional Hexagonal Structure using Triisopropylbenzene as Micelle Expander.* - American Chemical Society, .

92. Mohanty, P., Lee, J., Glover, K., and Landskron, K. (2011) Discoid Bicelles as Efficient Templates for Pillared Lamellar Periodic Mesoporous Silicas at pH 7 and Ultrafast Reaction Times. *Nanoscale Research Letters. 6*, 1-5.

Jinwoo Lee

13 Duh Drive 112

Bethlehem, PA 18015

(610)-573-9530

EDUCATION

Doctor of Philosophy in Chemistry, January 2013

Lehigh University, Bethlehem, Pennsylvania

Concentration: Biochemistry

Research: Probing the Caveolin-1 Structure.

Advisor: K. Jebrell Glover, Ph.D.

Master of Science in Chemistry, September 2008

Lehigh University, Bethlehem, Pennsylvania

Concentration: Biochemistry

Research: Structural Investigation of Caveolin-1.

Advisor: K. Jebrell Glover, Ph.D.

Bachelor of Science in Chemistry, February 2006

Hanyang University, Seoul, Korea

Concentration: Organic Chemistry

Research: Gold(I)-Catalyzed Intramolecular Hydroamination of Alkyne with Trichloroacetimidates.

Advisor: Seunghoon Shin, Ph.D.

RESEARCH EXPERIENCE

Graduate Student, Fall 2006 - Present

Professor K. Jebrell Glover, Ph.D.: Lehigh University, Bethlehem, Pennsylvania;

Thesis: Investigation of the Three-dimensional Structure of Caveolin-1.

Expertise in the Cloning, Expression, Purification, Reconstitution, and NMR analysis of membrane proteins.

Undergraduate Research Assistant, Spring 2005 - Summer 2006

Professor Seunghoon Shin, Ph.D.: Hanyang University, Seoul, Korea

A study of the gold(I)-catalyzed intramolecular hydroamination of trichloroacetimidates derived from propargyl and homopropargyl alcohols.

TEACHING EXPERIENCE

Teaching Assistant, 2006 – 2011

Lehigh University, Bethlehem, Pennsylvania

Courses: General Chemistry Lab (Chem 30)

Fall 2006, Spring 2007, Fall 2007, Spring 2008, Summer 2008, Fall2008,

Spring 2009, Fall 2009, Spring 2010, Fall 2010

Organic Chemistry Lab (Chem 51)

Summer 2007, Summer 2010

PUBLICATIONS

Root, K.T.*; **Lee, J.***; Glover, K.J., “Caveolin Revealed: An Experimental determined Topological Model.” Manuscript in preparation.

Rieth, M.D.*; **Lee, J.***; Glover, K.J., “Probing the caveolin-1 P132L mutant: Critical insights into its oligomeric behavior and structure.” Biochemistry 2012, 51 (18), 3911-3918.

Lee, J.; Glover, K.J., “The transmembrane domain of caveolin-1 exhibits a helix-break-helix structure.” Biochim. Biophys. Acta 2012, 1818, 1158-1164.

Mohanty, P.; **Lee, J.**; Glover, K.J.; Landskron, K., “Discoid bicelles as efficient templates for pillared lamellar periodic mesoporous silicas at pH 7 and ultrafast reaction times.” Nanoscale Res. Lett. 2011, 6, 61-65.

Diefenderfer, C.; **Lee, J.**; Mlyanarski, S.; Guo, Y.; Glover, K.J., “Expression and Purification of Transmembrane Domains for Biophysical Studies.” Analytical Biochemistry 2009, 384, 274-278.

Kang J.E.; Kim H.B.; **Lee J.W.**; Shin S., "Gold(I)-Catalyzed Intramolecular Hydroamination of Alkyne with Trichloroacetimidates" Org. Lett. 2006, 8, 3537–3540.

* denotes co-first authors.

PRESENTATIONS

Biophysical Society Pennsylvania Network Meeting, September 2012

-Poster presentation: Revealing the structure of the transmembrane domain of caveolin-1.

FASEB conference (Molecular Biophysics of Membrane), June 2012

-Talk and Poster: Revealing the structure of the transmembrane domain of caveolin-1

Delaware Membrane Protein Symposium, May 2012

-Poster presentation: Revealing the structure of the transmembrane domain of caveolin-1.

Biophysical Society, March 2011

-Poster presentation: Investigation of the three-dimensional structure of the caveolin-1 membrane interacting domain.

Delaware Membrane Protein Symposium, May 2011

-Poster presentation: Investigation of the three-dimensional structure of the caveolin-1 membrane interacting domain.

Delaware Membrane Protein Symposium, November 2009

-Poster presentation: Insights into the Secondary Structure of Caveolin.

AFFILIATIONS

Sigma Xi, 2010 – present

American Chemical Society, 2010 – present

Biophysical Society, 2011 – present

AWARDS

Student Chemistry Foundation Fellowship, 2011-2012