4-21-2009

Expression and Purification of Engineered Calcium Binding Proteins

Adriana P. Castiblanco
adripaka@yahoo.com

EXPRESSION AND PURIFICATION OF ENGINEERED CALCIUM BINDING PROTEINS

by

ADRIANA PATRICIA CASTIBLANCO ALFONSO

Under the Direction of Dr. Jenny J. Yang

ABSTRACT

Previous studies in Dr. Yang's laboratory have established a grafting, design, and subdomain approach in order to investigate the properties behind Ca^{2+}-binding sites located in Ca^{2+}-binding proteins by employing engineered proteins. These approaches have not only enabled us to isolate Ca^{2+}-binding sites and obtain their Ca^{2+}-binding affinities, but also to investigate conformational changes and cooperativity effects upon Ca^{2+} binding.

The focus of my thesis pertains to optimizing the expression and purification of engineered proteins with tailored functions. Proteins were expressed in *E. coli* using different cell strains, vectors, temperatures, and inducer concentrations. After rigorous expression optimization procedures, proteins were further purified using chromatographic and/or refolding techniques. Expression and purification optimization of proteins is essential for further analyses, since the

techniques used for these studies require high protein concentrations and purity. Evaluated proteins had yields between 5-70 mg/L and purities of 80-90% as confirmed by SDS-PAGE electrophoresis.

INDEX WORDS: Calcium, Engineered proteins, Protein expression, Protein purification, Chromatographic techniques.

EXPRESSION AND PURIFICATION OF ENGINEERED CALCIUM BINDING PROTEINS

by

ADRIANA PATRICIA CASTIBLANCO ALFONSO

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2009

EXPRESSION AND PURIFICATION OF ENGINEERED

CALCIUM BINDING PROTEINS

by

ADRIANA PATRICIA CASTIBLANCO ALFONSO

Committee Chair: Dr. Jenny J. Yang

Committee: Dr. Giovanni Gadda
Dr. Gangli Wang
Dr. Zhi-Ren Liu

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
May 2009

DEDICATION

I want to especially dedicate this thesis to my beautiful Conejin family: Hector, Stella, Diego, and Genesis. To my parents Hector and Stella, for all their exceptional and unconditional advice and support. I could not be what I am now without the values, beliefs, and principles that you have inculcated in me throughout my entire life. To my brother Diego, for always being supportive of my goals and for the many conversations that we have had about life, school, and our family's future. I want to also dedicate this thesis to my boyfriend Larry and to my true friends and close relatives. Thank you all for your support.

Adriana Castiblanco (May 2009)

DEDICACION

Quiero especialmente dedicar esta tesis a mi hermosa familia Conejin: Hector, Stella, Diego y Genesis. A mis padres Hector y Stella, por todo sus consejos y apoyo excepcional e incondicional. Yo no seria la persona que soy ahora de no ser por los valores, creencias, y principios que me han inculcado durante toda mi vida. A mi hermano Diego, por siempre estar cerca mio apoyandome y por todas las conversaciones que hemos tenido acerca de la vida, la universidad y el futuro de nuestra familia. Tambien quiero dedicar esta tesis a mi novio Larry y a mis verdaderos amigos y familiares cercanos. Muchas gracias a todos por su apoyo.

Adriana Castiblanco (Mayo 2009)

ACKNOWLEDGEMENTS

I want to express my gratitude and appreciation to Dr. Jenny J. Yang for her unconditional advice and guidance during my undergraduate and graduate career and for the opportunity to work in such a unique and interesting environment. I would like to especially thank my colleague and friend Dr. Yun Huang for her unrestricted support and for all the knowledge and good times that we have shared over the past years. Additionally, I will like to thank Dr. Shunyi Li for his guidance in all the purification techniques and especially for the tag-less purification method. I also have to express my gratitude to Michael Kirberger for his editorial assistance with various documents. And last but not least, I will like to recognize the rigorous and intense work of current and past members in Dr. Yang's research group and to the National Institutes of Health (NIH) supplement grant for promoting diversity and monetary support in the research field.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	After induction
AC	Affinity chromatography
BI	Before induction
BAB	Beads after binding
BAC	Beads after cleavage
BAE	Beads after elution
CaM	Calmodulin
CaSR	Calcium-sensing receptor
CD	Circular dichroism
CD2.D1	Domain 1 of cluster of differentiation 2
C/P	Cell pellet
ECD	Extracellular domain
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
Elu	Elution
EGFP	Enhanced green fluorescent protein
FHH	Familial hypocalciuric hypercalcemia
GPCR	G protein-coupled receptor
GST	Glutathione S-transferase
6x-his	Histidine tag
HPTH	Primary hyperparathyroidism
IEC	Ion exchange chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
LB	Luria-Bertani
NSHPT	Neonatal severe hyperparathyroidism
NMR	Nuclear magnetic resonance
MRI	Magnetic resonance imaging
MS	Mass spectroscopy
MW	Molecular weight
OD	Optical density
O/N	Overnight
P	Pellet
PBS	Phosphate-buffered saline
pI	Isoelectric point
PTH	Parathyroid hormone
RUB	Rubella virus
7TMD	Seven transmembrane domain
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
STIM1	Stromal interaction molecule 1
Sup	Supernatant
Thx	Thioredoxin tag
UV	Ultraviolet
Wst	Waste

1. INTRODUCTION

1.1 Ca^{2+}-binding proteins and their properties

Calcium (Ca^{2+}) is classified as an alkaline-earth metal and is one of the most abundant metallic elements on earth. It is involved in the regulation of numerous important cellular signaling pathways and biological processes in living organisms [1, 2]. In the Ca^{2+}-dependent signal pathways, Ca^{2+} binding usually results in large conformational changes which allow proteins to interact with downstream partners or other target molecules.

Ca^{2+} concentrations vary more than 1000-fold in intracellular and extracellular environments and that is why Ca^{2+}-binding proteins usually have different Ca^{2+}-binding affinities depending upon their functional locations [2-4]. These proteins function by experiencing Ca^{2+} induced conformational changes whereby Ca^{2+} influxes are translated into metabolic responses [5]. For example, eukaryotic cells in a resting state can present a cytoplasmic Ca^{2+} ion level between 10^{-7}-10^{-6} M, which can be considered quite low compared to an extracellular concentration of 10^{-3} M (Figure 1) [4, 6]. It is important to mention that external stimuli such as hormonal signals can change the cytoplasmic Ca^{2+} ion level to an activated state of ~10^{-6} M [4, 6]. Thus, a significant number of intracellular and extracellular Ca^{2+}-binding proteins play essential roles in maintaining a proper distribution of Ca^{2+} concentrations [3]. These Ca^{2+}-binding proteins have important intracellular and extracellular functions. For example, intracellular Ca^{2+}-binding proteins are important in Ca^{2+} buffer and transport systems, activation of enzymes, metabolism, gene expression, apoptosis, and Ca^{2+} storage just to mention a few intracellular functions (Figure 1) [2-4]. Additionally, extracellular Ca^{2+}-binding proteins are important in neurite extension, chemotactic and hormone activity and they are important extracellular matrix components (Figure 1) [2-4].

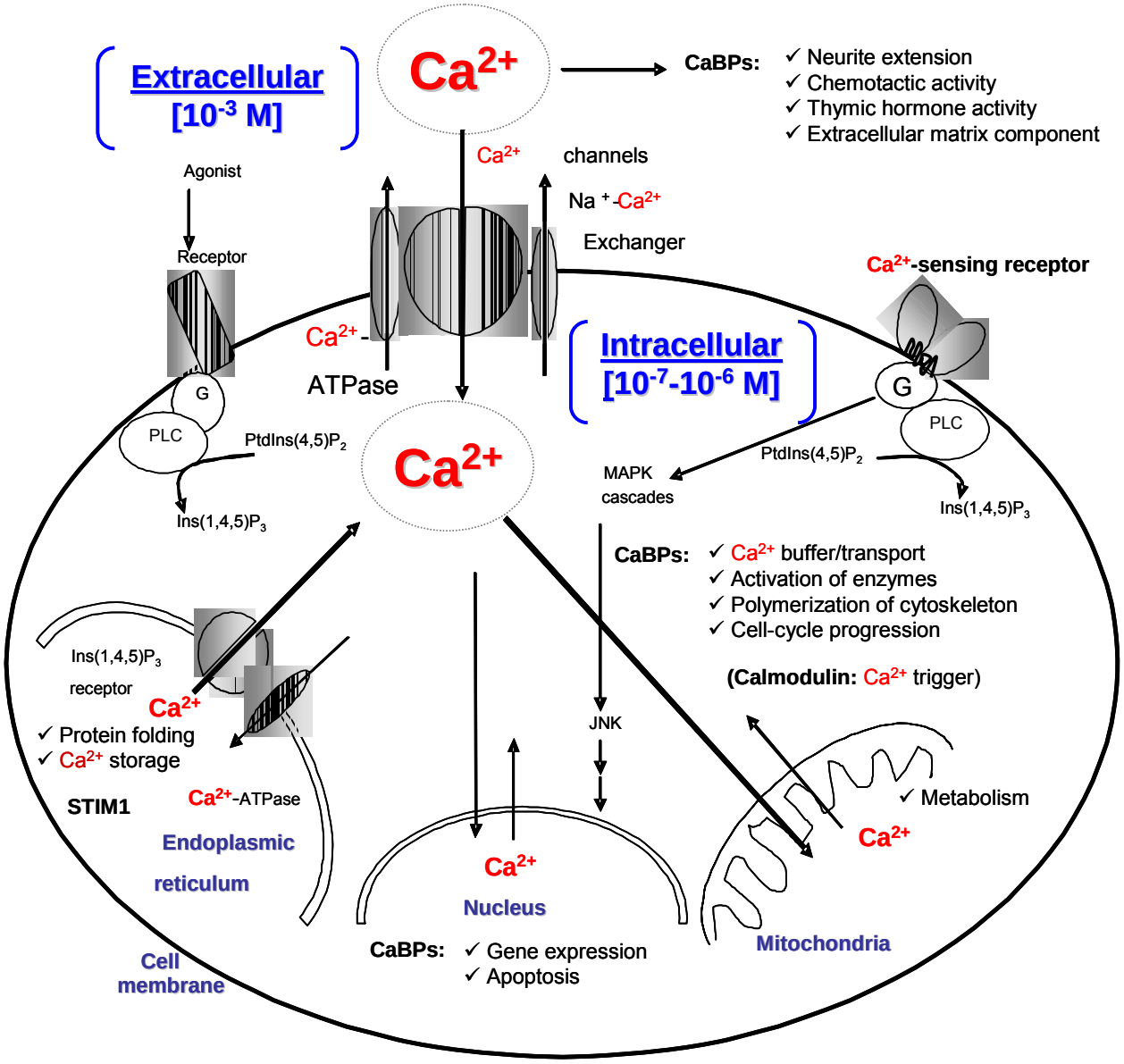


Figure 1. Intracellular and extracellular Ca^{2+} concentrations and important functions of Ca^{2+}-binding proteins (CaBPs) [2-4].

For example, calmodulin (CaM) is an extremely important and well studied intracellular Ca^{2+}-binding protein that belongs to the large EF-hand family. A typical EF-hand protein consists of two or more α–helices oriented ~90^{o} with respect to each other, and a loop around the Ca^{2+} ion in a helix-loop-helix formation [5, 7-9]. CaM contains four EF-hand motifs, and other EF-hand proteins may include two to eight EF-hand motifs [5, 7]. EF-hand proteins usually bind

to Ca^{2+} in a pentagonal bipyramidal arrangement with seven oxygen ligands (six of the ligands from the protein's side chains of Asp, Asn, Glu, and/or the main chain, and one from a water molecule) [5, 8]. EF-hand proteins have similar primary sequences and secondary structures but their cellular functions and the way they respond to Ca^{2+} are enormously diverse.

Non-EF-hand proteins, such as cell adhesion molecules, transmembrane proteins, and receptors like the Ca^{2+}-sensing receptor (CaSR) typically do not possess conserved Ca^{2+}-binding loops and helices, but they do possess Ca^{2+}-binding sites located at linker regions between domains or on extracellular domains [9]. For example, the CaSR senses extracellular Ca^{2+} concentrations and responds to those changes by modulating multiple intracellular signaling pathways through a series of Ca^{2+}-binding sites located in the extracellular domain (ECD) of this receptor [10]. This important receptor is expressed in a great number of tissues and it is associated with a large number of diseases including hypercalcaemia, hypocalcaemia, hypoparathyroidism, and hyperparathyroidism [11].

It is possible to find a great number of solved structures of Ca^{2+}-binding proteins, but many details related to Ca^{2+}-protein interactions and processes in which these special proteins achieve their Ca^{2+}-binding affinities are still unknown [8, 9, 12]. Ca^{2+}-binding proteins are characterized by having different Ca^{2+}-binding sites that interact with each other as they sense Ca^{2+}. Commonly-used experiments and investigations have been performed in order to study the total Ca^{2+} interaction with different proteins. Due to issues with cooperativity, efforts have been made to investigate isolated Ca^{2+}-binding sites either by grafting the sites onto scaffold proteins, separating domains, or inactivating binding sites, but these approaches, may produce results in the tested system that differ significantly from the intact protein, particularly with respect to Ca^{2+}-binding affinities [8, 9, 12].

1.2 Creating engineered Ca^{2+}-binding proteins

The first step in creating an engineered protein is to study in detail the chemical, biological, physical, and structural properties of the potential host protein. Stability of the host protein is an important characteristic because the protein must be able to tolerate mutations and/or addition of amino acid sequences while retaining its natural conformation and biochemical properties.

The development of engineered Ca^{2+}-binding proteins by predicting Ca^{2+}-binding sites and further designing or grafting them into a host protein can be an excellent way to understand and study Ca^{2+}-protein interaction in more detail. However it is important to emphasize that the host protein must retain its native structure after the Ca^{2+}-binding site has been designed or grafted and that the protein's environment should not affect the Ca^{2+}-binding properties [12].

Domain one of cluster of differentiation (CD2.D1) is a 99 residue long non-Ca^{2+}-binding protein that was chosen as the host protein for protein engineering [13, 14]. CD2.D1 exhibits several characteristics that make it an attractive scaffold for engineering, including: easy expression and purification; maintenance of its native folding state over a wide range of both pH (pH 1.0-10.0) and salt concentrations (0-4 M); and previous data for over 40 mutations shown not to produce significant deviations from its native fold [13, 14]. Also, well established X-ray and NMR structures for human (Figure 2A) and rat CD2 (Figure 2B) have been reported allowing biophysical analysis based on those results [15-18]. Therefore, the design of engineered Ca^{2+}-binding proteins using this protein is an excellent way to study Ca^{2+}-binding affinity and metal and Ca^{2+} selectivity by integrating Ca^{2+}-binding sites into CD2.D1.

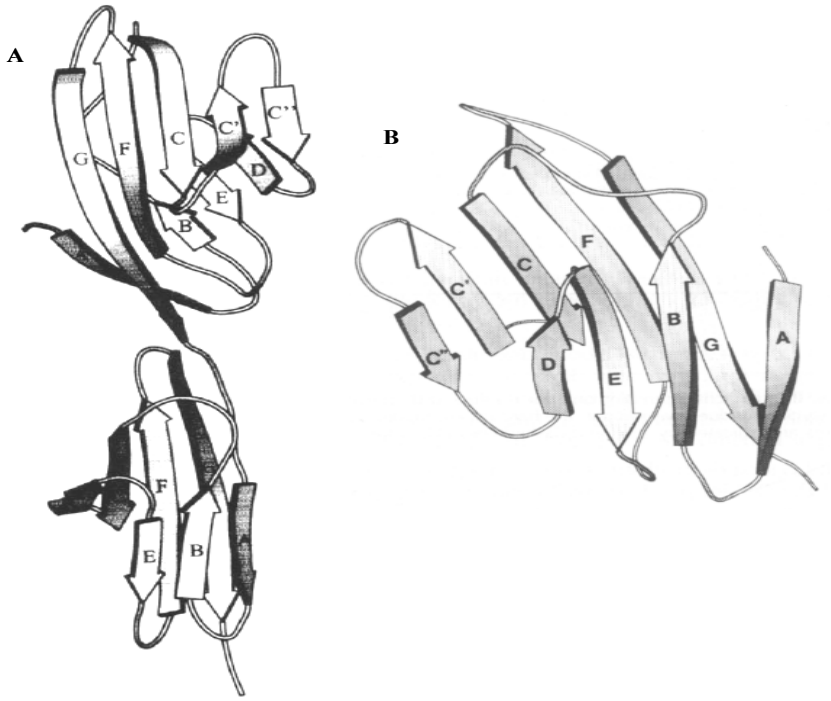


Figure 2. Structure of human and rat CD2. (A) Ribbon drawing for human CD2 domains I and II [17]. (B) Ribbon drawing for rat CD2 domain 1 [18].

1.3 Protein expression: *E. coli* as a protein producer

The bacterium *Escherichia coli* (*E. coli*) is the most widely studied prokaryotic organism and most commonly used bacterium for protein production in laboratory and industrial scenarios. It is a rod-shaped gram-negative bacterium containing a distinct surface layer composed of inner and outer phospholipid membranes and a rigid peptidoglycan cell wall between these two layers [19, 20]. This biochemical structure allows these gram-negative bacteria to adjust and adapt themselves to different environments and especially to temperature changes, since they also have the ability to physically change the diameter of their porins located on the cell membrane when they sense environmental changes [20]. *E. coli* has been selected for different studies because it is quite easy to grow in the laboratory and because its genetic mechanisms co-evolved along with those of mammals. These microorganisms can be grown in the laboratory in liquid rich or minimal media doubling in number about every 20-30 minutes depending upon the media

nutrients, usually at a temperature of 37 ^{o}C, and they can also be grown as colonies on agar plates containing nutrient medium and antibiotics to avoid contamination [19, 21].

Most bacteria reproduce via an asexual process called binary fission in which each cell increases in size and divides into two cells [22]. During this process there is an orderly increase in cellular structures and components, replication and segregation of the bacterial DNA, and formation of a septum or cross wall which divides the cell into two progeny cells [22]. A growing bacterial population doubles at regular intervals in an exponential way; but the exponential growth is only part of the bacterial life cycle, and it does not represent the natural pattern of growth of bacterial cultures [22].

Under optimum conditions, bacteria will increase in number until they either deplete the nutrients in the medium or accumulate sufficient waste products to inhibit growth. Under these conditions of saturation, the cells undergo a physiological change that allows for survival in the absence of growth. This state is called stationary phase, where the number of cells in a culture increases slowly, if at all. When stationary phase cells are diluted into fresh medium, they undergo another physiological change and begin to divide at an exponential rate known as the logarithmic or log phase. Most physiological experiments are conducted using rapidly dividing log phase cultures. Cell growth can be measured by following the increase in absorption of a culture at 600 nm using a spectrophotometer and checking the optical density (OD) at different time periods. A culture with an absorbance of 1.0 at 600 nm contains approximately 2 x 10^{8} cells/mL [22].

1.3.1 Factors that affect protein expression

Several factors may affect protein expression. For instance, selection of a proper cell strain is an essential factor that needs to be considered during bacterial transformation and

expression. This is a critical factor because protein expression can be strongly influenced by the way these cell strains respond at different expression temperatures and inducer amounts.

Protein expression in *E. coli* can follow two paths, where the protein is either expressed in a soluble, well-folded form or as inclusion bodies or similar insoluble aggregates (Figure 3). These inclusion bodies are usually inactive misfolded proteins that were formed due to over-expression. [23]. It is important to mention that sometimes certain proteins can be purified from inclusion bodies and therefore, inclusion body formation is desired and sometimes the best way to obtain pure and functional proteins [24].

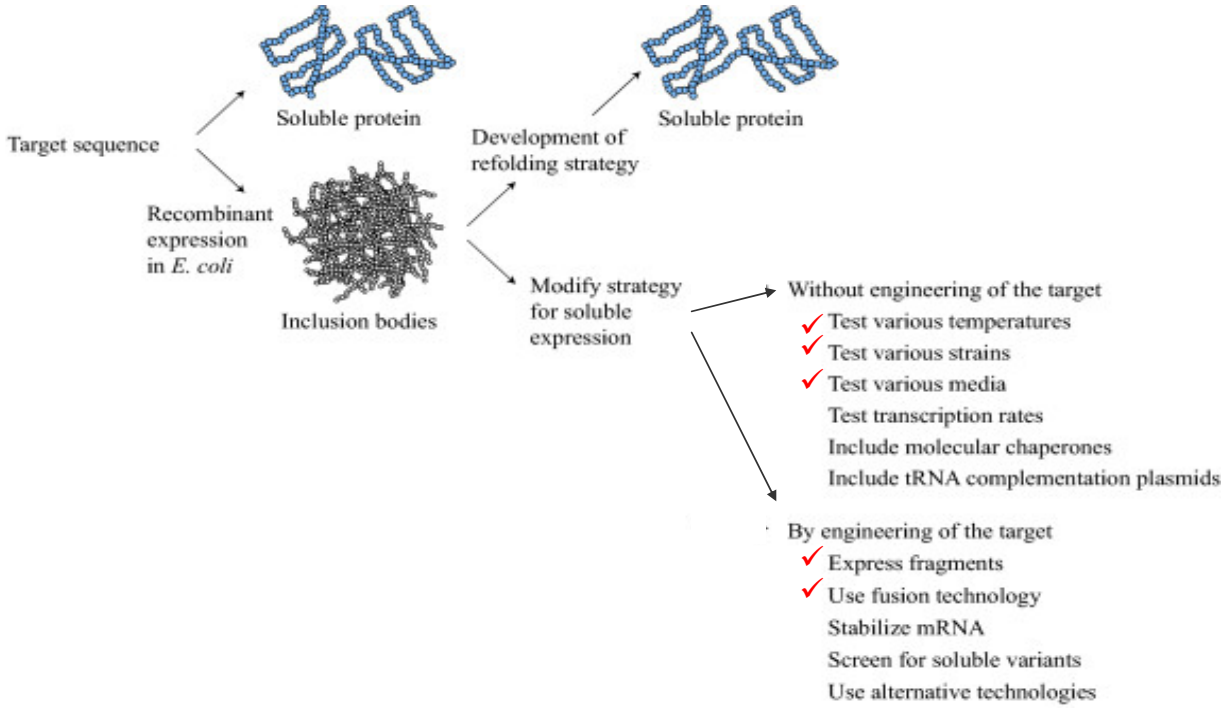


Figure 3. *E. coli* and the production of soluble or insoluble proteins and two strategies to treat and/or prevent inclusion body formation [25].

Experimental methods can be designed and modified to avoid inclusion bodies formation. First, a refolding strategy can be developed where denaturants are used to initially solubilize and unfold the protein, followed by removal of the denaturant to allow Domain 1 of cluster of

differentiation 2 the protein to refold into the native and active conformation (Figure 3) [25]. Another option is to modify the expression conditions in order to obtain soluble proteins. We can therefore test various temperatures, cell strains, and/or media conditions without changing the target sequence or we can also engineer the target by expressing fragments or making changes to the target sequence using fusion technology (Figure 3) [25]. Low expression temperature and low induction concentrations usually increases protein's solubility [25]. For example, temperatures around 30 ^{o}C activate a great number of *E. coli* chaperones stabilizing and favoring correct folding [23].

1.3.1.1 Protein expression using different cell strains

There are a great variety of cell strains that can be used in order to express proteins in *E. coli*. The most widely used cell strain is BL21 which does not encode *Ion*, *ompT*, and *ompP* proteases minimizing protein degradation during purification [19, 26]. There are different strains within BL21. For example, BL21(DE3) contains a T7 RNA polymerase gene that is integrated into the genome and that is under control of the *lacUV5* promoter, which is a lactose analog IPTG inducible system. The difference between BL21(DE3) and BL21(DE3)pLysS is that the pLysS strain has a plasmid encoding a T7 lysozyme gene that allows the expression of more toxic proteins in greater amounts since that gene is a natural inhibitor of T7 RNA polymerase and therefore, protein expression is tightly regulated [27]. Moreover, the cell wall lysis in the purification step is greatly enhanced in pLysS strain compared to DE3 because the lysozyme comes out and lyses the cells in addition to the method of cell disruption (French press or sonication).

Tuner is another strain that can be used for protein expression. Tuner strains are *lacZY* deletion mutants of BL21 which enable adjustable levels of protein expression throughout all

cells in a culture by allowing precise IPTG control [27]. Therefore, IPTG induction can occur in a true concentration-dependent fashion that is uniform throughout the culture and the expression can be regulated from very low expression levels up to highly induced expression levels [27].

Rosetta strains are derived from BL21 *lacZY* and tuner strains but they carry a pRARE derived plasmid that encodes several "rare" *E. coli* tRNA genes that can be encoded by a chloramphenicol resistant plasmid [27]. Some of the rare codon genes are AGG/AGA (arginine), AUA (isolucine), and CCC (proline) [27]. Therefore, these strains are designed to increase the expression of heterologous proteins whose genes encode numerous rare *E. coli* codons. Expression of these proteins can be dramatically increased when the level of rare tRNA is increased within the host.

Origami strains have mutations in both the thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes, which enhance disulfide bond formation in the cytoplasm. These strains are selectable on kanamycin, tetracycline, ampicillin, and chloramphenicol. OrigamiB is a derivative of origami and tuner strains and it contains *trxB* and *gor* mutations and it facilitates a precise IPTG control during protein expression [28].

It is possible that problems associated with low expression, toxicity or protein solubility may be mitigated by protein expression optimization using different cell strains. For general expression it might be useful to use BL21(DE3) but if we need to work with a toxic protein it might be better to utilize BL21(DE3)pLysS. On the other hand, if we are working with an insoluble protein tuner(DE3)pLacI might be the best way to express the target protein. Protein's low expression levels might be enhanced by using rosetta(DE3)pLysS or origamiB(DE3)pLysS. A brief description and applications of five cell strains are summarized in Table 1 and a more

detailed table that was modified from Novagen with the most commonly used cell strains in our lab can be observed in Table 2 [28].

Table 1. Common applications of five *E. coli* cell strains

	Cell Strain	Comments
General Expression	BL21(DE3)	T7-based expression upon IPTG induction.
Toxic Protein	BL21(DE3) pLysS	The pLysS plasmid produces T7 lysozyme to reduce basal level expression of the gene of interest.
Insoluble Protein	Tuner(DE3) pLacI	Contains a mutation in the lac permease (*lacZY*) gene. This enables adjustment of protein expression by using different IPTG concentrations.
Low Expression	Rosetta(DE3) pLysS	They supply tRNAs rarely used in *E. coli* providing a "universal" translation.
	OrigamiB(DE3) pLysS	Contains mutations in thioredoxin reductase and glutathione reductase genes, which enhances disulfide bond formation in the cytoplasm

Table 2. Description of cell strains employed in this thesis

Description	Cell strain	BL21(DE3)	BL21(DE3)pLysS	Rosetta(DE3)pLysS	Rosetta-gami(DE3)pLysS	OrigamiB(DE3)pLysS	Tuner(DE3)pLacI
BL21 - lacks of *ompT* (outer membrane protein T) membrane and cytoplasmic *Lon* proteases		X	X	X	X	X	X
DE3 - host carries a chromosomal copy of the T7 RNA polymerase gene under control of the *lacUV5* promoter		X	X	X	X	X	X
pLysS - provides T7 lysozyme (reduce basal expression of target genes)			X	X	X	X	
Carries *lacI*^{q} for overexpression of lac represor protein (suppresses basal expression from promoters)					X		
Lacks *trxB*, thioredoxin reductase (facilitates disulfide bond formation in cytoplasm)					X	X	
Lacks *gor*, glutathione reductase (facilitates disulfide bond formation in cytoplasm)					X	X	
Lacks *lacY* (providing a homogeneous uptake of IPTG into all cells in the population)				X			X
Overproduces *lacI* (suppress basal transcription of target genes controlled by lac operators)							X
Provides tRNAs for mammalian codons that rarely occur in *E. coli*				X	X	X	
Chloramphenicol resistance			X	X	X	X	X
Kanamycin resistance					X	X	
Tetracycline resistance					X	X	

Temperature and inducer concentrations vary between cell strains and protein expression can be highly affected by these factors. Protein solubility is actually enhanced when proteins are expressed at lower temperatures (30 ^{o}C, 25 ^{o}C, or lower), because these low temperatures slow down cells growth rate during bacterial expression, help to avoid formation of hydrophobic interactions, and proteases activity is highly reduced [20, 27]. Additionally, chromophore formation of certain proteins is often achieved at low temperatures.

1.4 Protein purification and its aims

Purification may be preparative or analytical. The aims in a preparative purification are to produce a relatively large quantity of purified proteins for subsequent use. Examples include the preparation of commercial products such as enzymes (e.g. lactase), nutritional proteins (e.g. soy protein isolate), and certain biopharmaceuticals (e.g. insulin) [29, 30]. The aim in an analytical purification is to produce a relatively small amount of protein for a variety of research or analytical purposes, including identification, quantification, and studies of the protein's structure, post-translational modifications and function [29]. If the protein is present in low amounts, a recombinant DNA technology can be used in order to develop cells that will produce large quantities of the desired protein, and in that way, create an expression system.

Additionally, an analytical purification generally utilizes certain protein properties in order to separate proteins, and depending on the source, purified proteins used for analytical purposes have to be brought into solution by breaking the tissue or cells containing them. There are several methods for achieving this including, repeated freezing and thawing, sonication, homogenization by high pressure, or permeabilization by organic solvents. The method of choice depends on the relative fragility of the protein and the robustness of the cells. After the extraction process, soluble proteins can be separated from cell membranes, DNA, and other

undesired components by centrifugation. The extraction process removes proteases, which will start digesting proteins in solution. If the protein is sensitive to proteolysis, it is usually desirable to proceed quickly, and keep the extract cooled to slow down proteolysis. Usually a protein purification protocol contains one or more chromatographic steps. The basic procedure in chromatography is to move the solution containing the protein through a column packed with various materials. Different proteins interact differently with the column material, and can thus be separated by the time required to pass the column or the conditions required to elute the protein from the column.

1.4.1 Commonly used purification techniques: separation based on chemistry

A great number of chromatographic methods for molecules separation involve a chemical interaction between the molecules of interest and the separation medium. These chemical interactions include dispersion forces, electrostatic dipole interactions, electron-donor acceptor forces, and covalent bonds formations. Dispersion force interactions are caused by induced dipole-induced dipole interactions and are therefore classified as non-specific interactions. Electrostatic dipole interactions are related to orientation of permanent dipoles and/or the interaction between a permanent dipole and an induced dipole. Electron donor acceptor interactions include overlap of electron clouds, sharing of lone electrons, and the overlap of orbitals to form bonds. Covalent bond formation involves a transition from intermolecular to intramolecular forces and it is somewhat related to an electron donor-acceptor interaction [29-31].

1.4.1.1 Ion exchange chromatography

Ion exchange chromatography separates compounds according to the nature and degree of their ionic charge. The column to be used is selected according to its type and strength of

charge. There are two types of ion exchange chromatography. One is cation exchange in which the resins have a negative charge and are used to separate positively charged molecules. The second one is anion exchange in which the resins have a positive charge and are used to retain and separate negatively charged compounds (Figure 4) [29, 32]. Before the separation begins a buffer is pumped through the column to equilibrate the opposing charged ions. Upon injection of the sample, solute molecules will exchange with the buffer ions as each competes for the binding sites on the resin. The length of retention for each solute depends upon the strength of its charge. The most weakly charged compounds will elute first, followed by those with successively stronger charges. Because of the nature of the separating mechanism, pH, buffer type, buffer concentration, and temperature all play important roles in controlling the separation [29, 32].

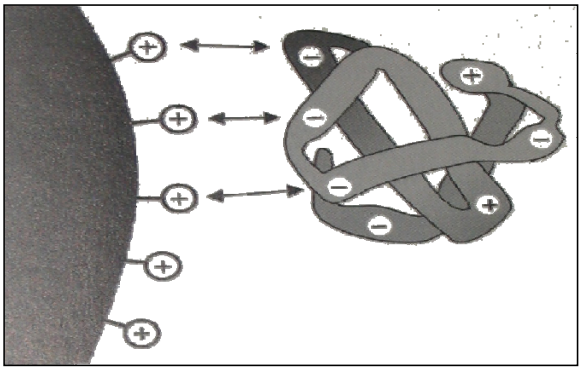


Figure 4. The principle of ion exchange chromatography (anion exchange) [32].

1.4.1.2 Affinity chromatography

Affinity chromatography is a method of separating biochemical mixtures, based on a highly specific biologic interaction, such as that between antigen and antibody, enzyme and substrate, or receptor and ligand [31, 33]. This chromatographic technique separates proteins on the basis of a reversible interaction between a protein and a specific ligand attached to a chromatographic matrix [33]. Therefore, the target protein or molecule is specifically and

reversibly bound to a complementary binding substance (affinity ligand) to then be desorbed from the stationary phase (Figure 5) [32].

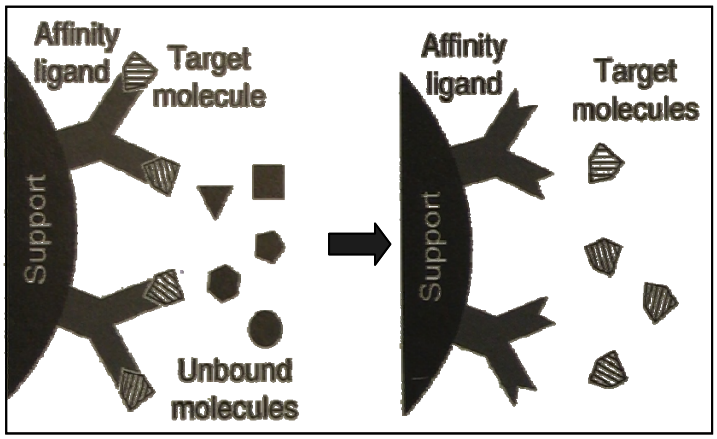


Figure 5. The principle of affinity chromatography [32].

This chromatography technique allows purification of proteins that have fusion tags, such as glutathione S-transferase (GST), polyhistidine (6x-his), and/or thioredoxin (thx) tags. The sample is applied under conditions that favor specific binding. Unbound material is washed away, and the bound target protein is recovered by changing conditions to those favoring desorption [32, 33]. Desorption is performed specifically using a competitive ligand or by using chemicals or conditions that can break bonds between the fusion tags and the target protein. Proteins are then collected, dialyzed in the desired buffer, and if necessary further purified using other chromatographic techniques.

1.4.1.3 Hydrophobic interaction chromatography

This type of technique separates biological macromolecules like proteins based on their hydrophobic surface (Figure 6) [32]. Its separation principle is still not well understood since most of the protein's tertiary structure hides the hydrophobic core and exposes the polar residues, but some proteins still present hydrophobic areas or patches on the surface [30-33].

This is a very powerful technique that allows the separation of closely related proteins and it can accomplish separations that ion-exchange or affinity chromatography can not accomplish.

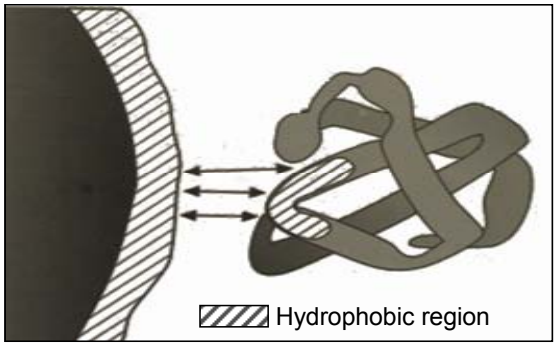
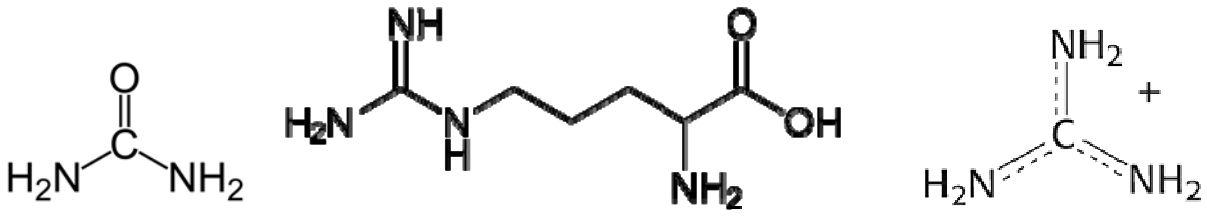


Figure 6. The principle of hydrophobic interaction chromatography [32].

1.4.1.4 Protein refolding and denaturant agents

Proteins can form inclusion bodies or inactive aggregates during the expression procedure and in some circumstances these inactive aggregates can be useful and advantageous in protein purification. The first step in protein purification using a refolding method is to isolate inclusion bodies and then use detergents and denaturants, such as urea, arginine, and/or guanidinium ions in order to completely unfold the protein (Figure 7). The following steps involve the removal of detergents and denaturants using dialysis or dilution strategies in order to slowly refold the protein into its native and active conformation. Sometimes, it is also necessary to further utilize other chromatographic techniques, such as ion exchange or affinity chromatography in order to further purify the protein.



Urea **Arginine** **Guanidinium ion**

Figure 7. Chemical structures of three commonly used denaturants.

This purification technique has its advantages and disadvantages. Some disadvantages are that it is time-consuming and there is no guarantee that the proteins will refold to its native conformation without losing activity and stability [26]. However, this purification strategy has been successfully demonstrated with small proteins, resulting in high yields at a low cost [26]. Additionally, insoluble aggregates are often protected against degradation by reducing the ability of proteases to damage the proteins. Therefore, protein refolding can be an excellent technique to purify proteins that were expressed as insoluble aggregates.

One of the most commonly used denaturants in protein refolding is urea or carbamide. This water-soluble organic compound, which can form up to eight hydrogen bonds (Figure 7) [34], has been widely used to solubilize proteins and especially inclusion bodies since it hydrogen-bonds to water [35]. The molecular basis for urea's denaturation abilities is still unknown, but it has been suggested that this compound can have a direct or an indirect effect in protein's stability. Thus, urea might be able to weaken protein's structure and stability by disrupting intramolecular hydrogen-bonds upon direct interaction with the protein [36, 37]. Or it might also indirectly affect protein's stability by altering the solvent environment and in that way facilitate exposure of hydrophobic residues [37]. An additional possibility that might explain urea's effect in protein's stability is that this effect might be a combination of protein-urea and solvent-urea interaction.

1.4.2 Detection of impurities using SDS-PAGE electrophoresis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is an excellent and relatively rapid way to determine a protein's molecular mass and at the same time analyze its purity and expression level since this technique separates proteins according to size [29]. This technique has been widely used to optimize purification steps and strategies in order to increase

protein's purity, expression, and stability [29]. The sample to be analyzed is mixed with a detergent (SDS) in order to denature the protein's secondary and tertiary structures, and at the same time apply a negative charge to the proteins since SDS is an anionic detergent containing a sulfate group. Therefore, proteins will be in a primary structure and they will be negatively charged allowing them to migrate toward the positive pole when an electric field is employed. Denatured samples will pass through a large pore 4% polyacrylamide gel (stacking gel) in order to have a starting point where all the proteins are concentrated to then pass through a small pore 3-30% polyacrylamide gel (resolving gel) containing different pore sizes and where smaller proteins will be able to go further down the gel whereas the larger proteins will stay on the top.

1.5 Objectives of this study

The main objectives of this study are to optimize the expression and purification of engineered Ca^{2+}-binding proteins. Protein's expression and purification are essential steps necessary to obtain proteins with high purity and yields. The successful optimization of these two important steps will allow us to obtain high quality and quantity proteins with the least possible amount of time and effort. These proteins will be used to perform future experiments (such as, metal binding studies, nuclear magnetic resonance (NMR), mass spectroscopy (MS), circular dichroism (CD), and titrations) in order to test their conformation and metal binding affinity and selectivity.

1.6 Overview

The research in this thesis is organized in the following way: In Chapter 1, I have introduced Ca^{2+}-binding proteins and some of their most important properties and functions in extracellular and intracellular environments. I have also introduced important concepts related to the

expression and purification, such as factors that affect protein expression in *E. coli* and important purification techniques that are commonly used to purify proteins.

In Chapter 2, materials and methods used to express and purify engineered Ca^{2+}-binding proteins are presented. The expression of some of these proteins was performed in two different media and four chromatographic techniques were used in order to purify engineered proteins depending upon the protein's physiochemical and conformational characteristics. The remaining chapters will focus on the expression and purification of engineered Ca^{2+}-binding proteins that were developed using either a grafting, subdomain, and/or design approach.

Chapter 3 introduces the grafting approach and how two host proteins, CD2 and EGFP, were used to insert Ca^{2+}-binding motifs to study metal-binding affinity, selectivity, and the creation of new EGFP-based Ca^{2+} sensors. This chapter presents the expression of CD2-WT and EGFP-WT and their variants, and it also includes protein purification using affinity (GST-tag or His-tag) and ion exchange chromatography.

Chapter 4 presents a brief relationship between CaM and the grafting approach. This chapter will mainly focus on the expression and purification of this important intracellular Ca^{2+}-binding protein.

Chapter 5 relates the grafting approach with the subdomain approach and how these approaches were used in order to study the CaSR and its Ca^{2+}-binding affinity and selectivity. This chapter presents the expression and purification of two engineered Ca^{2+}-binding sites that were grafted into CD2.D1 and purified using affinity (GST-tag) and ion exchange chromatography. It also includes the expression optimization of subdomain 1, 2, and 3 of the ECD of the CaSR and their purification employing affinity (6x-his-tag) chromatography for

subdomain 1 and 2, and a refolding method for subdomain 3. It will also discuss recent studies and accomplishments in the expression and purification of the ECD of the CaSR.

Chapter 6 introduces the design approach, the expression and purification of tag-less proteins CD2.7E15 and a variant using a refolding method and their applications as MRI contrast agents.

Chapter 7 is a concise summary of major conclusions and the significance of this thesis.

A summary of all studied engineered proteins by chapter and information related to their expression and purification can be observed in Table 3.

Table 3. Summary of engineered proteins studied by chapter

Chapter	Engineered Protein	Approach	Fusion Tag	Vector	Cell Strain	Purification Technique
3	CD2.STIM1	Grafting	GST	pGEX-2T	BL21(DE3)	AC and IEC
	CD2.STIM1.Mut1	Grafting	GST	pGEX-2T	BL21(DE3)	AC and IEC
	CD2.Rub.Ca	Grafting	GST	pGEX-2T	BL21(DE3)	n/a
	EGFP.WT	n/a	6x-His	pET28a	BL21(DE3)	AC
	EGFP.EF.172	Grafting	6x-His	pET28a	BL21(DE3)	AC
4	CaM.WT	n/a	n/a	pET20b	BL21(DE3)	HIC
5	CD2.CaSR.Site3	Grafting	GST	pGEX-2T	BL21(DE3)	AC and IEC
	CD2.CaSR.Site5	Grafting	GST	pGEX-2T	BL21(DE3)	AC and IEC
	CD2.CaSR.Mutants	Grafting	GST	pGEX-2T	BL21(DE3)	AC and IEC
	Subdomain 1	Subdomain	6x-His	pRSETa	Rosetta(DE3)pLysS	AC
	Subdomain 2	Subdomain	6x-His	pRSETa	Rosetta(DE3)pLysS	AC
	Subdomain 3	Subdomain	6x-His	pRSETa	Tuner(DE3)pLacI	Refolding and AC
	ECD	n/a	6x-His	pRSETa	Rosetta-gami(DE3)pLysS OrigamiB(DE3)pLysS	AC
6	CD2.7E15	Design	n/a	pET20b	Tuner(DE3)pLacI	Refolding and IEC
	CD2.7E15.52I	Design	n/a	pET20b	Tuner(DE3)pLacI	Refolding and IEC

2. MATERIAL AND METHODS

2.1 Cloning and transformation

Engineered proteins were cloned by standard PCR reactions by Dr. Yun Huang [38] and Dr. Yubin Zhou [39] and details can be found in their PhD dissertations. The competent cells used for the transformation procedures were BL21(DE3), BL21(DE3)pLysS, Tuner(DE3)pLacI, Rosetta-gami(DE3)pLysS, OrigamiB(DE3)pLysS and/or Rosetta(DE3)pLysS. The antibiotics used were ampicillin, kanamycin chloramphenacol, and/or tetracycline. To begin, 50 μL of the competent cell were added into labeled tubes, followed by 0.5-1.0 μL of DNA. The tubes were then placed on ice for 30 minutes, then in 42 ^{o}C water bath for exactly 90 seconds to allow the DNA transfer into cell. Following this heat shock, samples were returned to the ice bucket for 2 minutes. Next, 50 μL of Luria-Bertani (LB) media was added to the tubes and they were then placed in an incubator for 30 minutes at 37 ^{o}C. After 30 minutes, 50 μL of the cell culture were added drop wise to a labeled agar plate, and streaked across the agar surface using a sterile triangle. One additional plate was streaked (without DNA) as a control and both were placed at 37 ^{o}C overnight with the agar side of the plate facing up.

2.2 Protein expression

2.2.1 Luria-Bertani (LB) media

For overnight inoculation, a single and healthy clone was removed from a previously transformed agar plate using an inoculation loop. The loop was inserted and stirred to release the clone into a 1 L flask containing 500 mL LB media (10 g/L peptone from casein, 5 g/L yeast extract, and 10 g/L NaCl) and 100 mg/L ampicillin or 30 mg/L kanamycin. The flask was then placed in the shaker to incubate overnight at 37 ^{o}C at a speed of 220-230 rpm.

For the expression of engineered proteins, 50 mL of observable cell growth grown overnight was placed into a 4 L flask containing 1 L LB media. Next, either 1000 µL of ampicillin (100 mg/L) or 600 µL of 50 mg/mL kanamycin were added for a final concentration of 0.03 mg/mL and the flask was placed in the shaker at 37 °C. The optical density (OD) was checked at regular intervals to evaluate cell growth using a Shimadzu UV-1601 PharmaSpec UV-Vis spectrophotometer with UV Probe software (Shimadzu North America, Columbia, MD). When the OD reached approximately 0.6-0.8, cells were induced with 100-500 µL isopropyl ß–D-1-thiogalactopyranoside (IPTG) which is an effective inducer in the concentration range of 0.1-1.5 mM. Cells were collected the same day they were expressed at 37 °C or if they were expressed at lower temperatures (25 °C or 30 °C) cells were allowed to grow overnight with agitation and collected the following day. Cells were collected using high speed centrifugation (7,000 rpm) for 20 minutes and stored at -20 °C (Figure 8).

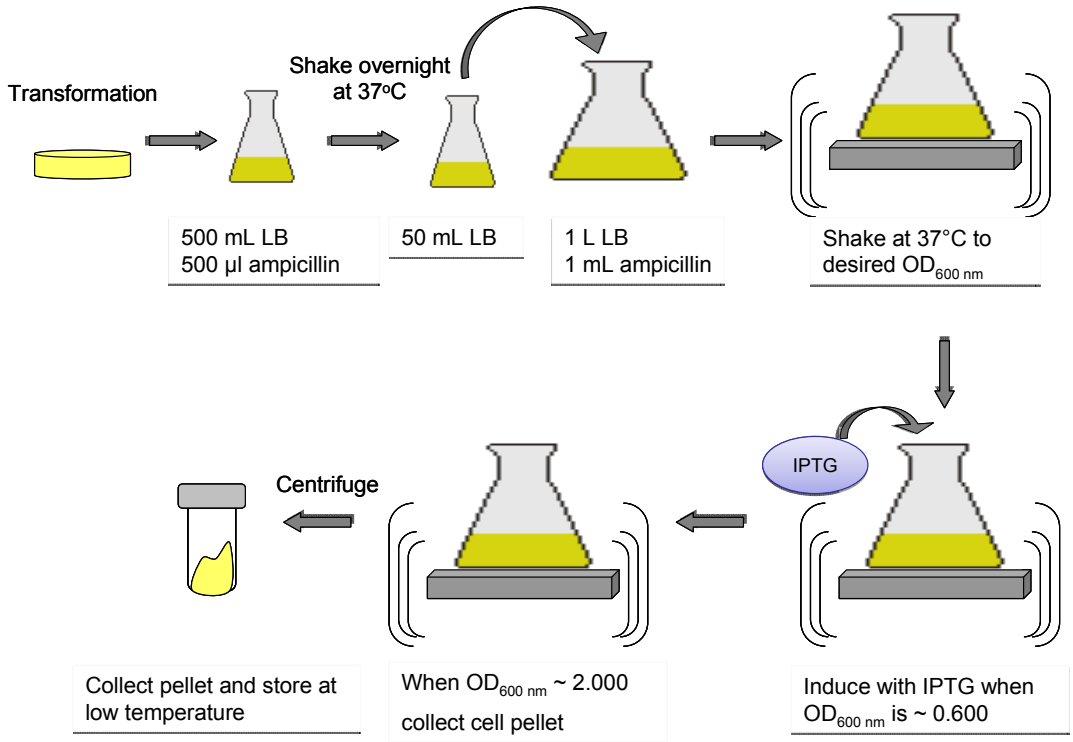


Figure 8. Diagram illustrating protein expression in *E. coli.*

2.2.2 Labeled ^{15}N minimal media

Labeled ^{15}N minimal media was prepared using the following chemicals: $K_2HPO_4 \cdot 3H_2O$ (10.4 g/L), KH_2PO_4 (4.4 g/L) $MgSO_47H_2O$ (0.05 g/L), $(NH_4)Fe(SO_4)_2 \cdot 6H_2O$ (0.007 g/L), 20% glucose, and 20% $^{15}NH_4Cl$. The expression procedure using minimal media is similar to that explained earlier for LB media (Figure 8). The only difference is the procedure used in the inoculation step. It was necessary to inoculate 10 mL of LB media, 10 μL ampicillin, and a healthy individual clone from an agar plate that contained 3% glucose in three to four different 50 ml plastic tubes for at least 8 hours. The best growing cultures were then transferred to a 1 L minimal media flask until the OD reached a 0.6 value. Cells were then induced with 150 μl IPTG and temperature was reduced to 25 ^oC overnight. The expressed and further purified proteins included ^{15}N labeled atoms that were used in NMR studies.

2.3 Protein purification

Engineered proteins were originated using a grafting, a subdomain, or a design approach. Some of these proteins had fusion tags which helped for a successful expression and purification and other proteins were expressed and purified as wild type proteins for use as controls in future experiments.

All of the engineered proteins that were grafted into CD2.D1 contained a GST-tag (pGEX-2T) and they were purified using affinity and ion exchange chromatography, whereas, those proteins that were grafted into EGFP contained a 6x-his-tag (pET28a) and were purified using nickel-sepharose affinity chromatography. Proteins generated from the subdomain approach also contained a 6x-his-tag (pRSETa). Subdomain 1 and 2 were therefore purified using nickel-sepharose affinity chromatography. However, an additional refolding technique was utilized to purify subdomain 3. Proteins that were generated using a design approach were tag-

less proteins (pET20b) and they were therefore purified using a refolding technique followed by ion exchange chromatography. CaM was another tag-less (pET20b) protein that was purified using hydrophobic interaction chromatography. Refer to Table 3 in Chapter 1 for a complete table that summarizes all the engineered proteins, fusion tags, vectors, and purification technique(s) used.

2.3.1 Affinity and ion exchange chromatography (GST-tag)

A 2 L expression cell pellet was re-suspended in 35-50 mL lysis buffer (1% sacosine and 1 mM EDTA in PBS, pH 7.4) and 1 M DTT was added to the solution to a final concentration of 5 mM. The resulting solution was subjected to sonication (Branson 450 Sonifier) 7-8 times for 30 seconds with 1-2 minute intervals between sonications in order to avoid protein degradation due to heat produced by the sonicator's sound waves.

The solution was then centrifuged at 17,000 rpm for 40 minutes (S34 rotor). The supernatant was subsequently filtered with a 0.45 μm filter (Whatman, Florham Park, NJ) and the filtered supernatant was then poured in glutathione (GST) agarose beads (GE healthcare) affinity columns in order to start protein binding. The flow through was repetitively passed over the column 3-5 times for further binding. The affinity columns were washed with 50 mL 1X PBS buffer. A 2:5 BioRad protein assay is used to determine if the protein was bound to the GST beads. If the binding was successful a color change was observed on a micro-plate reader from a copper to a blue color. The protein was cleaved adding 20-40 μL per column of thrombin. The cleavage was done overnight at 4 ^{o}C with agitation and for an additional 2-3 hours at room temperature the following day. The columns were then washed with 20-50 mL 1X PBS buffer per column and allowed to elute the protein. The BioRad protein assay was used to determine if all the protein was eluted (Figure 9).

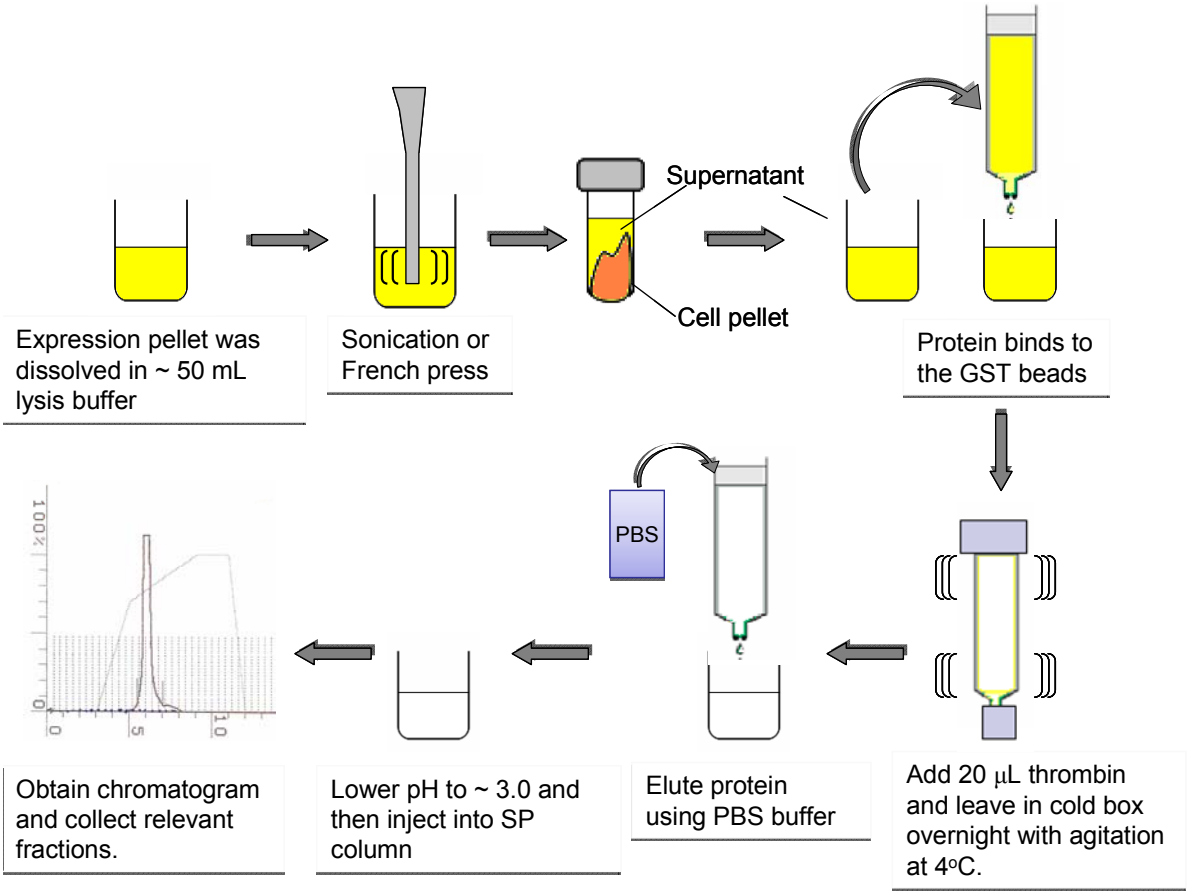


Figure 9. Diagram illustrating protein purification by GST-tag affinity column and ion exchange chromatography.

Proteins were further purified using ion exchange chromatography. The eluted protein's pH was lowered to a 3.0-4.0 range using acetic acid and it was subsequently filtered using 0.45 μm filter. Filtered solution was injected into SP column using a fast liquid chromatography (FPLC) system. Proteins are bound to the column in buffer A (50 mM sodium acetate-acetic acid buffer solution, pH 4.5) and they were eluted with buffer B (50 mM Tris-HCl, pH 8.0). The FPLC system displayed a chromatogram with a sharp peak when the protein was eluted allowing us to collect the relative fractions (Figure 9). Collected fractions were concentrated to 10 ml using an Amicon brand concentrating apparatus with a 3-10 kDa molecular weight cutoff

membrane. Concentrated solution was extensively dialyzed in 10 mM Tris-HCl buffer pH 7.4 to perform further studies.

2.3.2 Nickel-sepharose affinity chromatography (6x-his-tag)

A 2 L expression cell pellet was re-suspended in 20 mL extraction buffer (20 mM Tris, 100 mM NaCl, 0.1% Triton X-100) or 30-50 mL lysis buffer (1% sacosine and 1 mM EDTA in PBS, pH 7.4). Cell membranes were broken using sonication or French press. The resulting solution was centrifuged for 40 minutes at 17,000 rpm and the extracted supernatant was filtered with 0.45 μm pore size filter into a plastic tube. In order to improve protein purity, 1 M imidazole was optionally added to the solution so that the final concentration was 20-50 mM.

Purification of His-tag proteins was completed using a Hitrap 20 mL or 5 mL sepharose column and an ÄKTA prime FPLC instrument (Amersham Biosciences, Piscataway, NJ). Buffer A (referred to as binding buffer) was comprised of 1 M K_2HPO_4, 1 M KH_2PO_4, 250 mM NaCl, pH 7.4. Buffer B (referred to as elution buffer) was comprised of buffer A and 0.5 M imidazole. The column was first rinsed with a 100 mM EDTA and 1 M NaCl solution in order to remove any metals and undesirable impurities. It was then rinsed with ddH_2O twice. Following the cleaning step, the column was washed with 0.1 M $NiSO_4$ in order to bind Ni^{2+} onto the column, and it was then rinsed again with ddH_2O in order to remove any unbound $NiSO_4$. Filtered protein was then injected into the column and further eluted out in 8 mL fractions. Collected fractions were further purified by dialysis in 2 L of 10 mM Tris-HCl buffer pH 7.4 and using a dialysis bag with a molecular weight cutoff of 3500 Da. The dialysis solution was changed every 3 hours in order to remove imidazole (Figure 10).

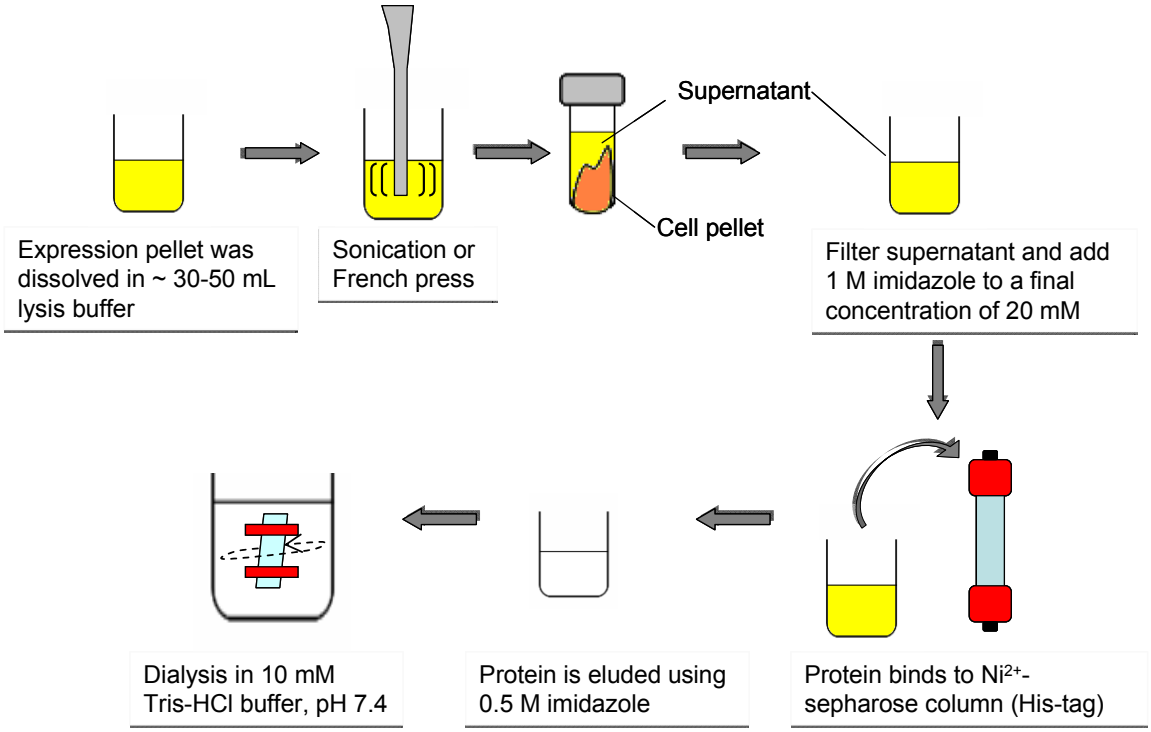


Figure 10. Diagram illustrating protein purification by 6x-his-tag affinity column step by step.

2.3.3 Hydrophobic interaction chromatography

A 2 L expression cell pellet was re-suspended in 30 mL homogenization buffer (2 mM EGTA, 1 mM DTT, 1 mM PMSF in 50 mM Tris-HCl, pH 7.5) and cell membranes were broken using French press. The resulting solution was subsequently heated at 80-85 ^{o}C in a water bath for about 5 minutes and this solution was centrifuged for 40 minutes at 17,000 rpm. The extracted supernatant was filtered with 0.45 µm pore size filter into a plastic tube and 1 M $CaCl_2$ was added to a final concentration of 5 mM. The supernatant was then loaded onto a phenylsepharose column that was previously equilibrated with wash 1 buffer (0.5 $CaCl_2$ in 50 mM Tris-HCl, pH 7.5). The binding process was initially done at room temperature with a 1 mL/min binding flow and the sample was recycled overnight at 4 ^{o}C to ensure complete protein binding to the column.

Next day, the column was washed with 150 mL wash 1 followed by 150 mL wash 2 buffer (0.5 mM $CaCl_2$, 500 mM NaCl in 50 mM Tris-HCl, pH7.5). Protein was eluted using elution buffer (5 mM Na_2EGTA in 50 mM Tris-HCl, pH 7.5) and the different fractions were checked using UV-visible spectrophotometry. Relative fractions were collected, concentrated, and dialyzed in 20 mM Tris-HCl buffer, pH 7.4 at 4 ^oC where the dialyzed buffer was exchanged every 2-3 hrs (Figure 11). The final protein concentration was calculated using UV-visible spectrophotometry, the protein's extinction coefficient, and the measured absorption at 277 nm.

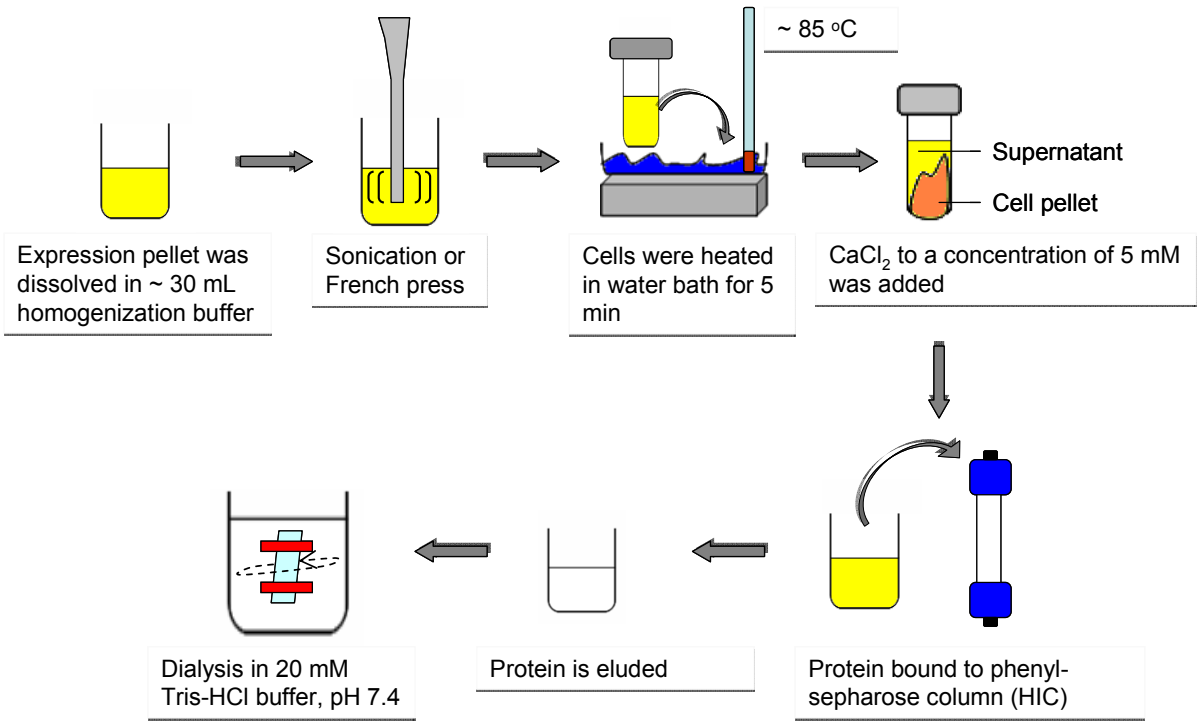


Figure 11. Diagram illustrating protein purification by hydrophobic interaction chromatography step by step (phenylsepharose column).

2.3.4 Urea refolding

A 1 L expression cell pellet was re-suspended in 30-35 mL PBS buffer, 60 µL benzonase nuclease, and 200 µL protease inhibitor. The resulting solution was subjected to French press three times. The solution was centrifuged for 30 minutes at 17,000 rpm. The supernatant was

discarded and the pellet was re-suspended in 20 mL 2% Titron X-100 in PBS solution and then centrifuged for 20 minutes at 17,000 rpm. This procedure was repeated 3-4 times. The washed pellet containing the protein was then transferred to two 50 mL plastic tubes and 25 mL of an 8 M urea (pH 7.3) and 5 mM EDTA (pH 11.3) solution was added to each tube and left at 4 ^{o}C overnight on a magnetic stirring plate. Next day, the solution was centrifuged for 20 minutes at 17,000 rpm. The supernatant was collected which contained the solubilized protein that was completely unfolded by the 8 M urea solution. It was then necessary to gradually reduce the urea concentration in order to refold the protein. The 8 M urea sample was then diluted to 4 M urea by adding 25 mL PBS buffer drop wise and by constantly stirring the sample at 4 ^{o}C for 2 hours. The 4 M urea solution was then transferred to two dialysis bags in order to dialyze them in 2 M urea for 4 hours. Dialysis bags were then transferred to a 2 L dialysis buffer (10 mM Tris-HCl buffer pH 7.4 and 200 µL protease inhibitor) overnight. The dialysis buffer was subsequently changed three times at three hour intervals in order to completely remove urea from solution. The dialyzed solution was further purified using ion exchange chromatography. Figure 12 illustrates a simplified diagram describing the refolding method process.

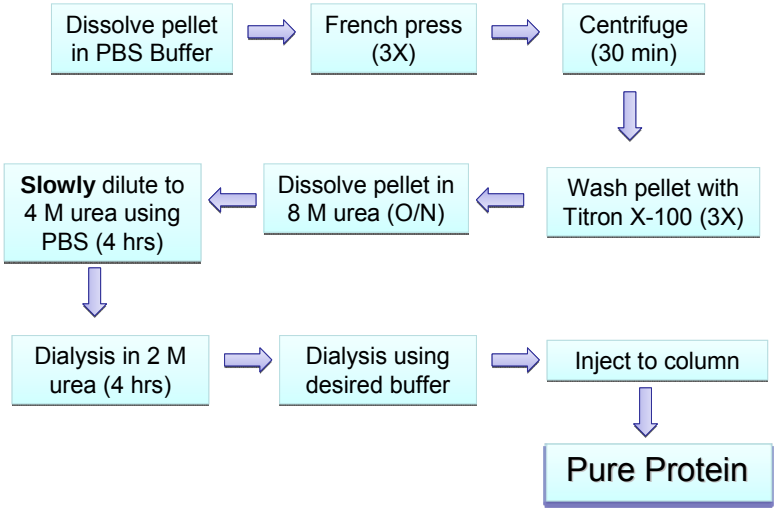


Figure 12. Diagram illustrating protein purification using 8 M urea refolding method.

2.4 SDS-PAGE electrophoresis

Samples to be analyzed were mixed with SDS-PAGE sample buffer (glycerol, 1 M Tris-HCl buffer pH 6.8, SDS, bromophenol blue, and dH_2O) and 5% β-mercaptoethanol to avoid S-S bridges formation. Mixed and denatured samples were heated in a boiling water-bath for 10 min. They were then loaded into a SDS casting gel containing a stacking gel on top (4% polyacrylamide) and a resolving gel on bottom (15% or 8% polyacrylamide). A constant current of ~ 110-120 V was applied and the samples were allowed to run until they migrated to the bottom of the resolving gel. The gel was stained using Coomassie brilliant blue staining buffer and 0.1% Coomassie blue R-250 methanol, acetic acid, dH_2O 40:10:50 for 20 min and then destained using destaining buffer solution (25% methanol, 7% acetic acid, and dH_2O).

2.5 Protein concentration calculation

Protein concentration was determined using UV-visible spectrophotometry, the UV absorbance at 280 nm or 277 nm (phenylalanine absorbance), the protein's molar extinction coefficient, and the Beer-Lambert Law,

$$A = \varepsilon\, b\, c$$

where A is the measured absorbance at 280 nm or 277 nm, ε is the molar extinction coefficient, b is the path length (1 cm), and c is the protein's concentration. The absorbance at 280 nm corresponds to tryptophan residues present in CD2 and its variants, while 277 nm corresponds to tyrosine residues present in CaM.

3. THE GRAFTING APPROACH: PROTEIN PURIFICATION BY AFFINITY AND ION EXCHANGE CHROMATOGRAPHY

3.1 Introduction

Most Ca^{2+}-binding proteins contain multiple binding sites and they are therefore able to bind Ca^{2+} cooperatively. The grafting approach is an excellent way to examine site-specific Ca^{2+}-binding properties. In this approach, a single EF-hand Ca^{2+}-binding loop or predicted Ca^{2+}-binding sites are inserted into a scaffold protein [12, 15]. Chosen scaffold or host proteins must adhere to a series of conformational and stability characteristics, and loop insertions into these host proteins must follow certain criteria that are essential for successful and accurate metal-binding and conformational studies. Domain 1 of CD2 (CD2.D1) and enhanced green fluorescence protein (EGFP) were the two chosen host proteins. Therefore, the usage of these proteins in combination with a grafting approach will allow us to create an environment to evaluate the intrinsic Ca^{2+}-binding affinities of isolated binding sites and it will also allow us to develop Ca^{2+} sensors by exploiting the intrinsic fluorescence properties of EGFP variants. In this chapter discussion will focus on the expression and purification of engineered Ca^{2+}-binding proteins in which Ca^{2+}-binding loops were inserted into CD2 and EGFP.

3.1.1 CD2.D1 as a host protein

CD2.D1 was chosen as a scaffold protein in order to create a model system and in that way study Ca^{2+}-binding by protein grafting. CD2 enhances T cell and NK cell activation by interacting with CD48 or CD58 on the antigen-presenting cells and is one of the most extensively studied non-Ca^{2+}-binding cell adhesion proteins [18]. This protein has high structural conformation stability and it is able to maintain its native structure when electrostatic changes

have been made to its structure [12]. Therefore, CD2 was converted into a specific receptor for Ca^{2+} by inserting Ca^{2+}-binding loops into this host stable protein (Figure 13).

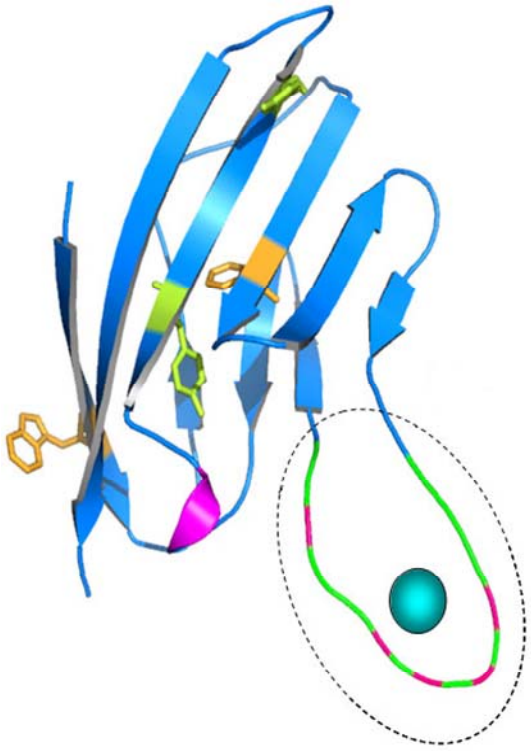
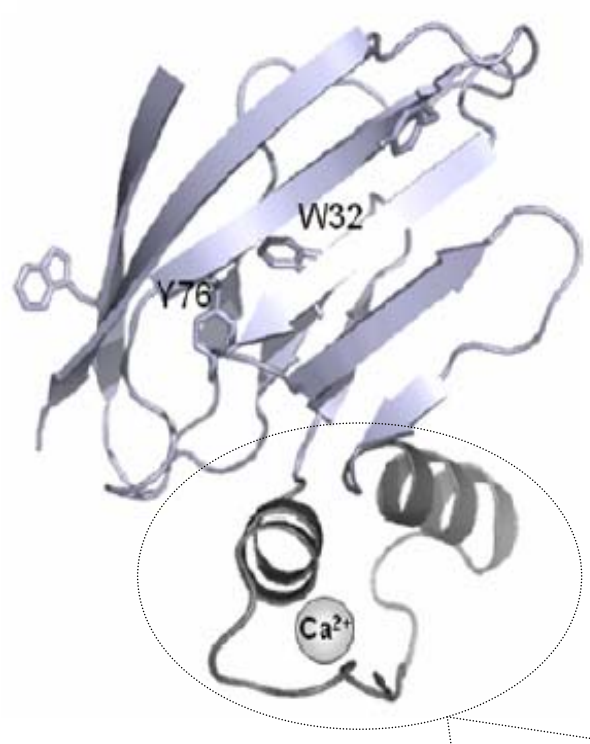


Figure 13. Model structure of CD2.D1 grafted with a putative Ca^{2+}-binding EF loop.

STIM1 is a 685 amino acid long transmembrane protein that is located on plasma and endoplasmic reticulum (ER) membranes [40]. The ER is essential in store-operated Ca^{2+} entry and actually STIM1 can be considered as a Ca^{2+} sensor in the ER. The N-terminal region of STIM1 facing the ER lumen contains a sterile α–motif (SAM) domain and an EF-hand motif [40, 41]. This EF-hand motif was inserted into CD2.D1 using the grafting approach in order to study its metal-binding activity and dimerization. Also, a STIM1 mutant (STIM1.Mut) was designed where two asparagines (D) in position 1 and 3 were mutated to alanines (A) in order to disrupt its Ca^{2+}-binding on the EF-loop. The model structure of CD2.STIM1 and CD2.STIM.mut with their inserted sequences can be observed in Figure 14.



STIM1 A V R N I H K L M D D D A N G D V D V E E S D E F L R E D
STIM1.MUT A V R N I H K L M A D A A N G D V D V E E S D E F L R E D

Figure 14. Model structure of CD2.STIM1 and CD2.STIM1.Mut. EF-hand motifs were inserted at position 52 connected by three Gly residues at both ends of the sequence.

The rubella virus is a member of the family *Togaviridae* (genus *Rubivirus*) and it causes a disease commonly known as rubella [42-44]. This single-stranded enveloped virus has a genome size of 9,762 nucleotides [42-44]. This genome consist of two long open reading frames (ORFs): a 5'-proximal ORF encoding nonstructural proteins (NSP) and a 3'-proximal ORF encoding structural proteins, one capsid protein, and two envelope glycoproteins [44, 45]. A Ca^{2+}- binding motif (12 residues long) located within the NS protease domain of the 5-proximal ORF of RUB was predicted and grafted between Gly53 and Ala54 by PCR into CD2 [44]. The prediction was done based on analysis of current data related to EF-hand proteins (Figure 15).

		1 3 5 7 9 12		
EF-hand		-n--nn--n	X-Y-ZG#-y--z	n--nn--n
CAM_EF1	11	EFKEAFSLF	DKDGDGTITTKE	LGTVMRSL 39
CAM_EF2	47	ELQDMINEV	DADGNGTIDFPE	FLTMMARK 75
CAYP_hum	97	VIAAAFAKL	DRSGDGVVTVDD	LRGVYSGR 115
CAYP_rab	97	VIAAAFAKL	DRSGDGVVTLGD	LRGVYSGR 115
RV1a_Fth	1197	SQRWSASHA	DASPDGTGDPLD	PLMETVGC 1225
RV1B_FAN	1197	SQRWSASHA	DASPDGTGDPLD	PLMETVGC 1225
RV1C_VEL	1197	SQRWSASHA	DASPDGTGDPLD	PLMETVGC 1225
RV1D_SAL	1197	SQRWSASHA	DASPDGTGDPLD	PLMETVGC 1225
RV1E_RAY	1197	SQRWSASHA	DASPDGTGDPLD	PLMETVGC 1225
RV2A_BRI	1197	SQRWIASHA	DASLDGTGDPLD	PLMETVGC 1225
RV2B_AN5	1197	SQRWSASHA	DASPDGTGDPLD	PLMETVGC 1225
RV2c_C74	1197	SQRWSASHA	DASPDGTGDPLD	PLMETVGC 1225

Figure 15. Sequence alignment of EF-hand Ca^{2+}-binding motif RUBCa (RV) and other EF-hand motifs [44]**.**

3.1.2 EGFP as a host protein

Green fluorescence protein (GFP) was first discovered and isolated by Shimomura and his group in 1962 from a jellyfish, *Aequorea victoria* [46]. Since then, a series of GPF variants has been created in order to enhance its biochemical characteristics. For example, enhanced GFP (EGFP) is a GFP variant in which point mutations (F64L and S65T) were done in order to enhance this protein's folding efficiency [46, 47]. EGFP has been used as a host protein to develop sensors by grafting Ca^{2+}-binding motifs into sensitive locations of this protein by choosing three different locations that were solvent accessible [48]. One of these sensors was EGFP-172-EF-III in which the Ca^{2+}-binding motif of CaM, loop III (DKDGNYISAAE) and the EF-hand motif (EEEIREAFRVFDKDGNGYISAAELRHVMTNL) were grafted into EGFP [48]. This sensor was named Ca-G1 and its motifs were inserted in position 1 of EGFP that corresponded to Glu172-Asp 173 within loop 9 [48] Figure 16.

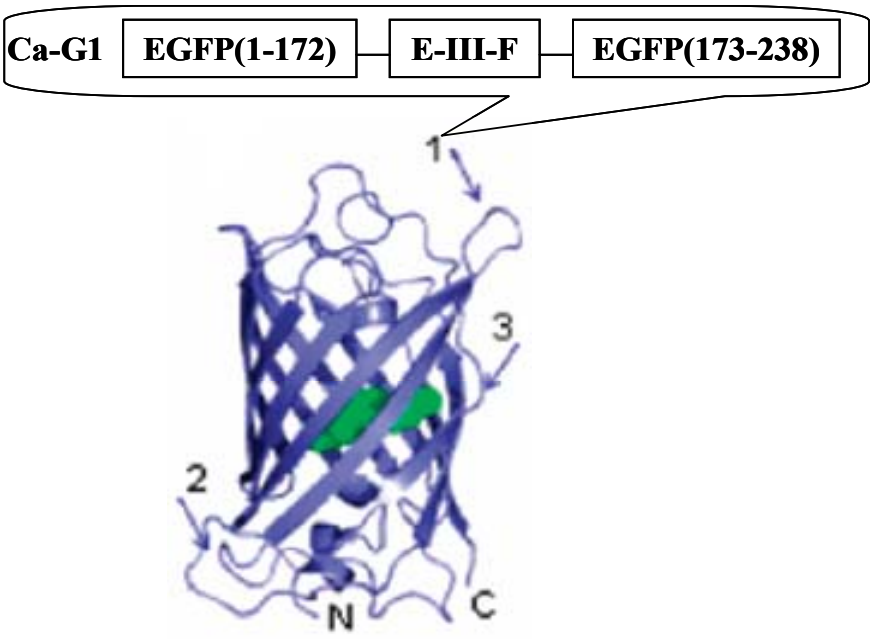


Figure 16. Model structure of a EGFP-based Ca^{2+} sensor (Ca-G1) located in position 1.

3.2 Results and discussion

Engineered proteins were generated using a grafting approach where CD2.D1 and EGFP were used as host proteins. The following section is divided into two parts. The first part includes the expression and purification of CD2-WT and its variants while the second part discusses the expression and purification of EGFP-WT and one of its variant.

3.2.1 Expression and purification of CD2-WT and variants by GST-fusion affinity and ion exchange chromatography

3.2.1.1 CD2-WT expression

CD2-WT was expressed in *E. coli* BL21(DE3) at 37 ^{o}C as a fusion engineered protein with a GST-tag in a pGEX-2T vector. The exponential expression growth curve for CD2-WT clearly shows that the proteins reached an OD at 600 nm of approximately 2.0, 3-4 hrs after induction (Figure 17A). The SDS gel confirmed that the protein was successfully expressed since no protein was observed BI, but after IPTG induction (AI) protein expression increased every hour, based on increased band thickness observed as a function of time (Figure 17B).

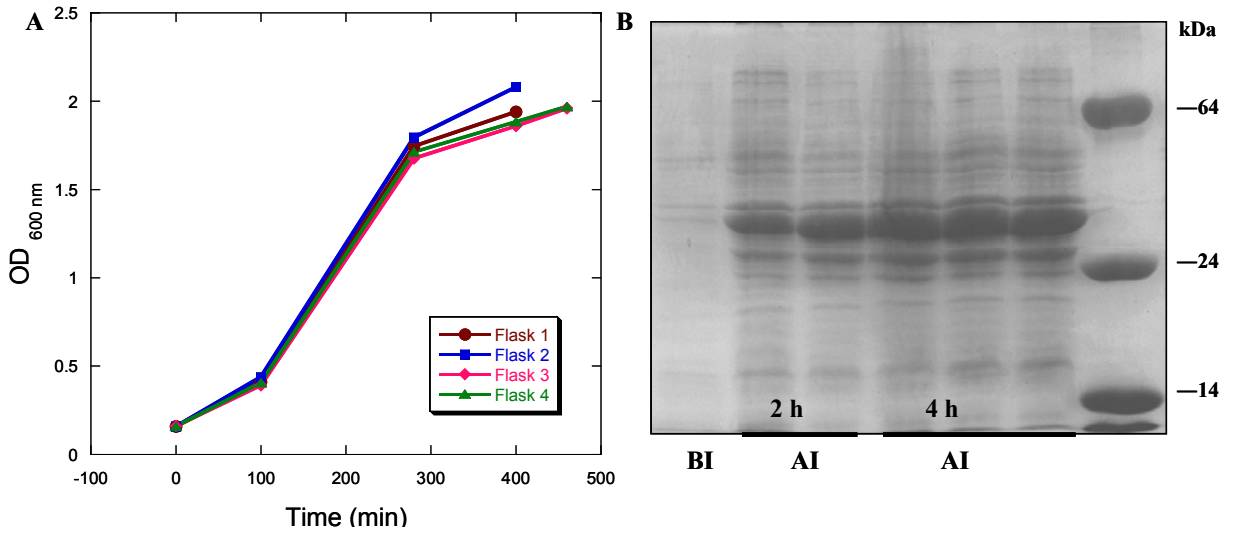


Figure 17. (A) CD2-WT expression growth curve (Vector: pGEX-2T, LB media, BL21(DE3), 37 ^{o}C expression). (B) CD2-WT expression in *E. coli*.

3.2.1.2 CD2-WT purification

CD2-WT was purified using affinity (GST-tag) and ion exchange chromatography. The purification SDS gel shows that most of the protein was in the soluble form (Sup) and that the GST-tag located on the protein bound to the column (BAB), even though some protein was still left in the waste (Wst) which is due to incomplete binding of the protein with the sepharose column. The GST-tag was effectively cleaved using thrombin (BAC), and protein was eluted out (Elu) using 1X PBS buffer (Figure 18A). The pH of the eluted protein was then reduced to ~ 3.0 using acetic acid for subsequent injection into an ÄKTA FPLC system containing an ion-exchange column. The chromatogram showed that the relevant fractions were 18 to 21 (Figure 18B) and the SDS gel confirmed that these fraction showed >95% purity (Figure 18A). These fractions were collected and subjected to dialysis in 10 mM Tris-HCl buffer, pH 7.4. The final concentration was 1.8 mM in ~ 2 mL and the final yield was 23.4 mg/L.

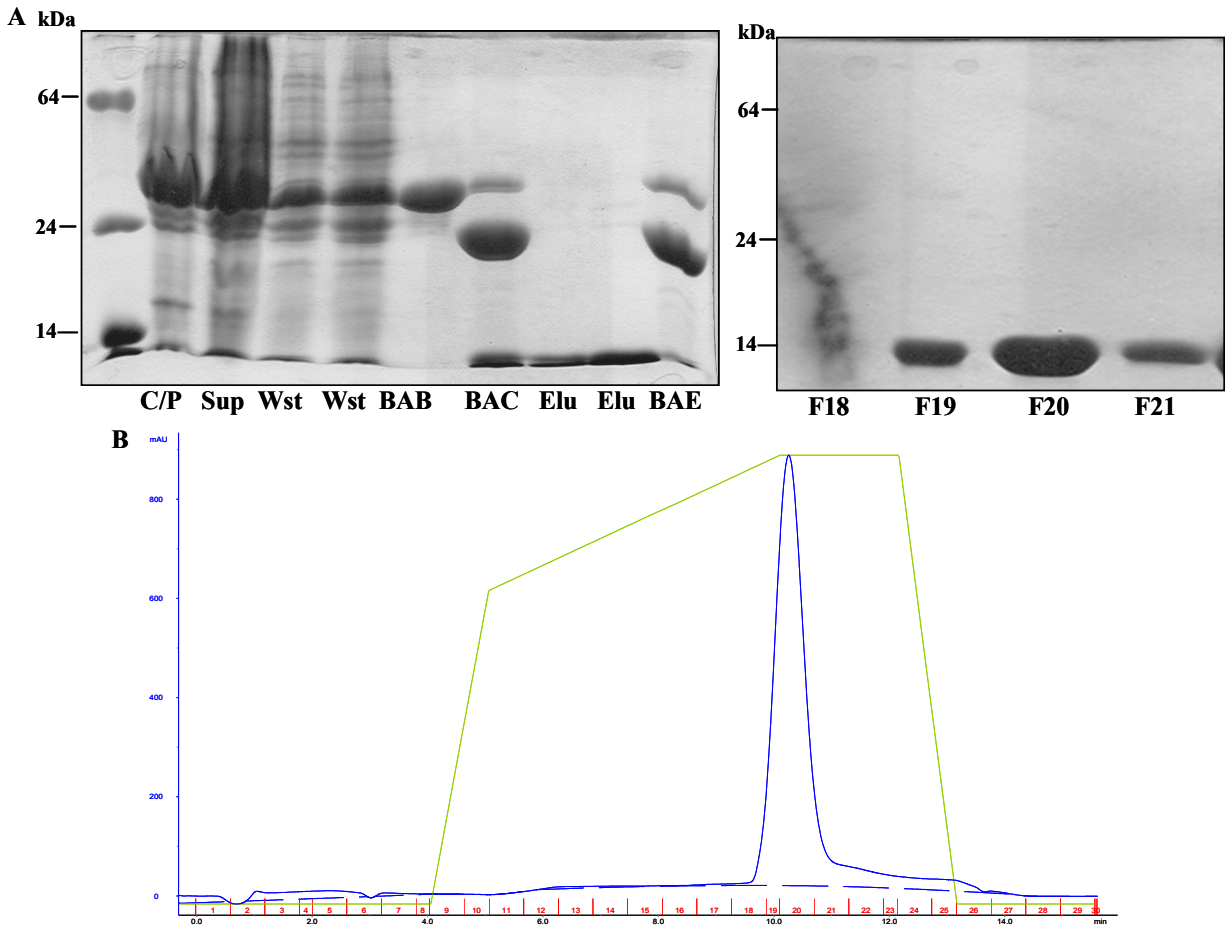


Figure 18. CD2-WT purification using affinity (GST-tag) and ion exchange chromatography. (A) SDS gels: GST-tag affinity column chromatography (left) eluted fractions after ion exchange chromatography (right) (B) ion exchange chromatogram (collected fractions 18-21).

3.2.1.3 CD2.STIM1 and CD2.STIM1.Mut expression

CD2.STIM1 and CD2.STIM.Mut were expressed in *E. coli* BL21(DE3) at 37 ^{o}C as a fusion engineered protein with a GST-tag in a pGEX-2T vector. The exponential expression growth curve for CD2.STIM1 clearly shows that the protein reached an OD at 600 nm of approximately 1.4, 4-5 hrs after induction (Figure 19A). The SDS gel confirmed that the protein was successfully expressed since no protein was observed BI but after IPTG induction (AI) protein expression increased every hour, based on increased band thickness observed as a function of time (Figure 19B). CD2.STIM1.Mut was expressed using the same conditions as

those used for CD2.STIM1 and a similar expression trend was observed. SDS gel electrophoresis confirmed the results.

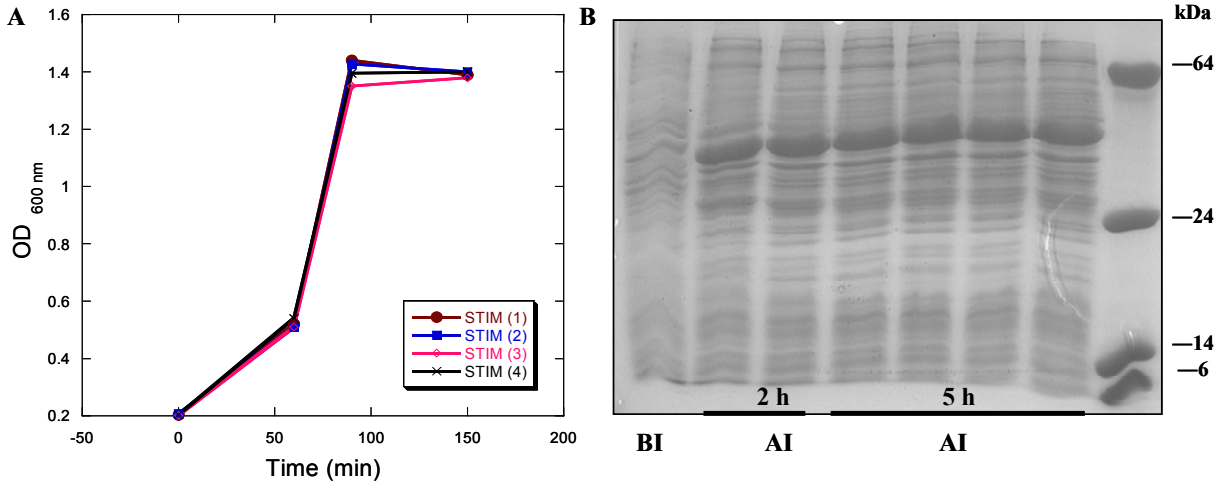


Figure 19. (A) CD2.STIM expression growth curve (Vector: pGEX-2T, LB media, BL21(DE3), 37 ^{o}C expression). (B) CD2.STIM expression in *E. coli*.

3.2.1.4 CD2.STIM1 and CD2.STIM1.Mut purification

These two engineered proteins were purified using affinity (GST-tag) and ion exchange chromatography. The purification SDS gel shows that most of the protein was in the soluble form (Sup) and that the GST-tag located on the protein bound to the column (BAB), even though some protein was still left in the waste (Wst) due to incomplete binding of the protein with the sepharose column. The GST-tag was effectively cleaved using thrombin (BAC), and protein was eluted out (Elu) using 1X PBS buffer (Figure 20). The pH of the eluted protein was then reduced to ~ 3.0 using acetic acid in order to be consequently injected into an ÄKTA FPLC system containing an ion-exchange column. The chromatogram showed that the relevant fractions were 17 to 19. These fractions were collected and subjected to dialysis in 10 mM Tris-HCl buffer, pH 7.4. The final concentration for CD2.STIM1 was 1.1 mM in ~2 mL and the final yield was 14.4

mg/L. A similar purification trend was observed for CD2.STIM1.Mut. Its final concentration was 770.9 µM in ~2mL and the yield was 10.0 mg/L.

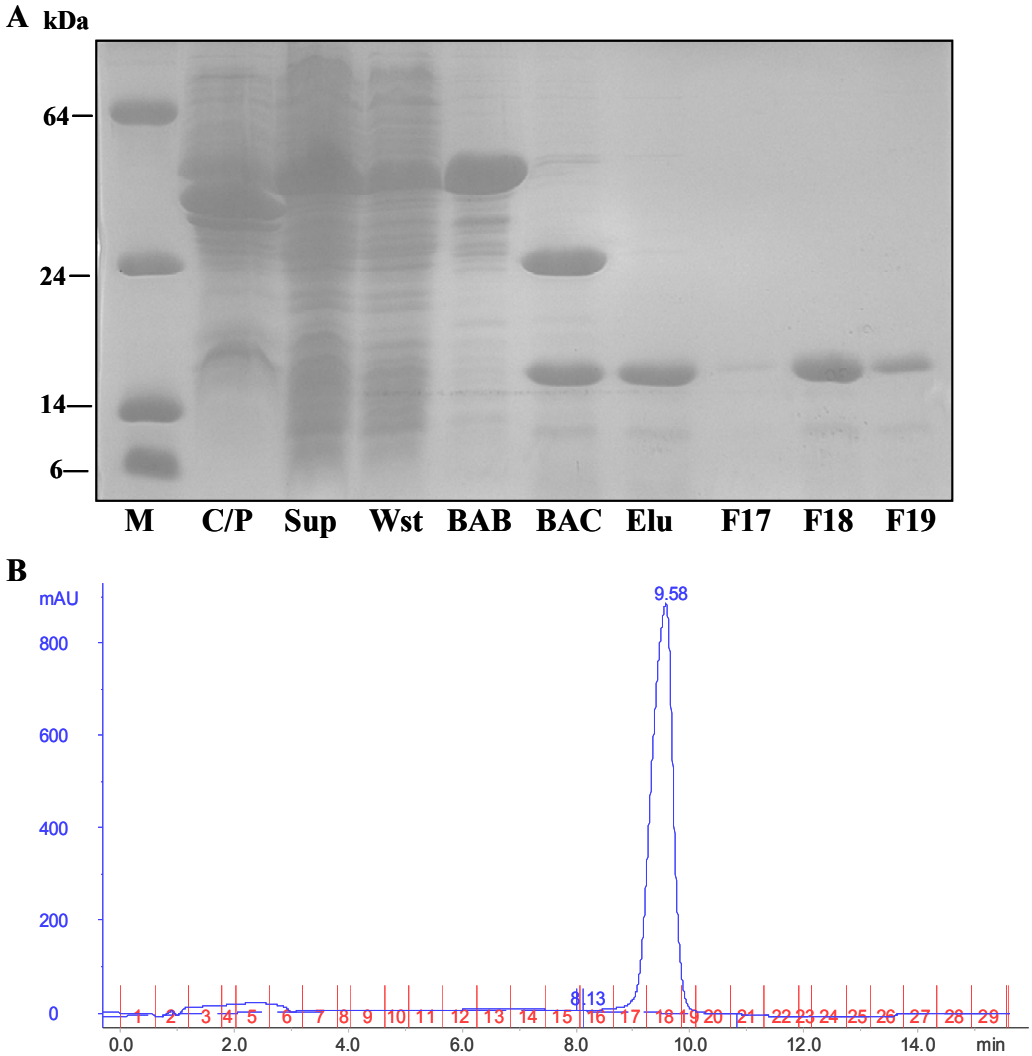


Figure 20. CD2.STIM purification using affinity (GST-tag) and ion exchange chromatography. (A) SDS gels: GST-tag affinity column chromatography and eluted fractions after ion exchange chromatography (B) ion exchange chromatogram (collected fractions 17-19).

3.2.1.5 CD2.RUB.Ca expression

This protein was also expressed in *E. coli* BL21(DE3) at 37 ^{o}C as a fusion engineered protein with a GST-tag in a pGEX-2T vector. The exponential expression growth curve for CD2.RUB.Ca clearly shows that the protein reached an OD at 600 nm of approximately 1.8, 3-4

hrs after induction (Figure 21A). The SDS gel confirmed that the protein was successfully expressed since no protein was observed BI but after IPTG induction (AI) the protein expression increased every hour, based on increased band thickness observed as a function of time (Figure 21B).

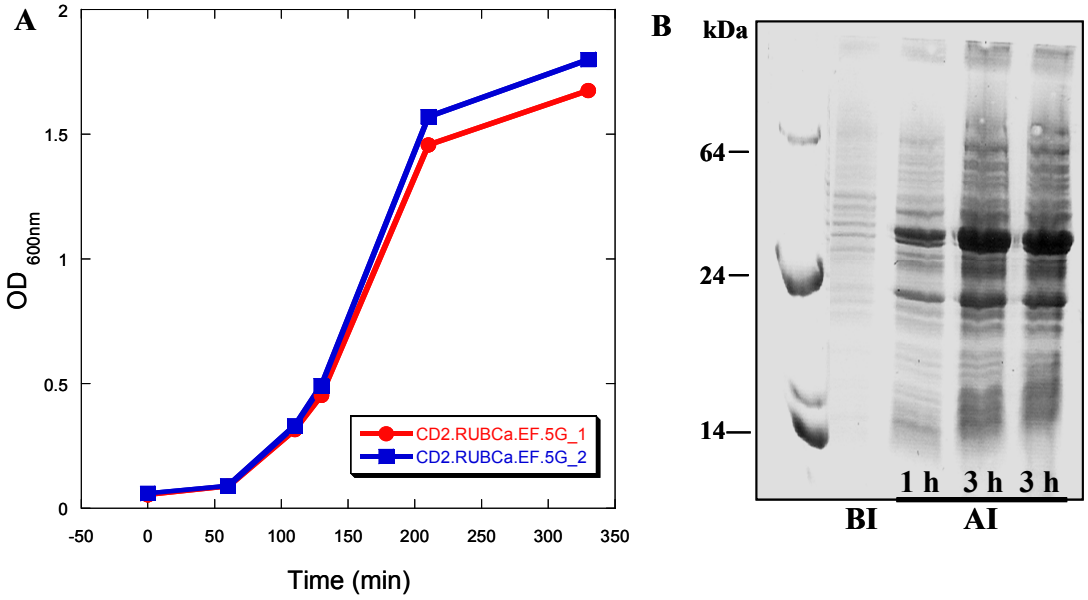


Figure 21. (A) Rub.Ca.EF expression growth curve (Vector: pGEX-2T, LB media, BL21(DE3), 37 ^{o}C expression). (B) RubCa.EF expression in *E. coli.*

3.2.2 Expression and purification of EGFP-WT and a variant using 6x-his-tag affinity chromatography

3.2.2.1 EGFP-WT and EGFP-EF-172 expression and purification

Green fluorescence proteins are different form CD2 since they contain a chromophore located at the center of a beta barrel structure. Chromophore formation is usually achieved at low temperatures since these temperatures allow proper chromophore maturation. The chromophore is responsible for this protein's characteristic green color after its expression in *E. coli* and further purification.

EGFP-WT and EGFP-EF-172 proteins were expressed in *E. coli* BL21(DE3) as fusion engineered proteins containing a 6x-his-tag in a pET28a vector. Cells were grown in LB media containing 30 μg/mL kanamycin at an initial temperature of 37 ^{o}C. Temperature was subsequently reduced after IPTG induction (OD 0.6) to 30 ^{o}C overnight in order to increase chromophore maturation. Both of these proteins were successfully expressed (Figure 22A and Figure 23A) and were further purified using 6x-his-tag affinity chromatography (Figure 22B and Figure 23B). Proteins were soluble after cells were disrupted using sound waves (sonication) and the collected fractions were dialyzed in 10 mM Tris-HCl buffer, pH 7.4 in order to remove the high imidazole concentration used to elute the protein. Fractions 12 and 13 for EGFP-WT and 13 and 14 for EGFP-EF-172 showed higher purity in the SDS gels (Figure 22B and Figure 23B) and they were used for crystallization procedures. The final yields for the collected fractions were 73.6 mg/L and 10.2 mg/L for EGFP-WT and EGFP-EF-172 respectively.

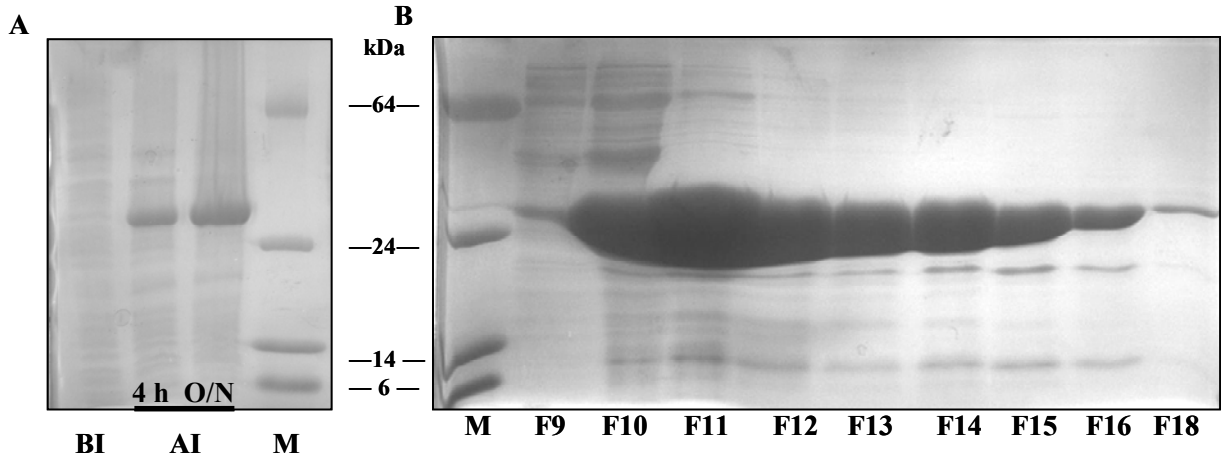


Figure 22. (A) EGFP-WT expression in *E. coli* (Vector: PET20a, LB media, BL21(DE3). (B) EGFP-WT purification SDS gel using 6x-his-tag affinity chromatography.

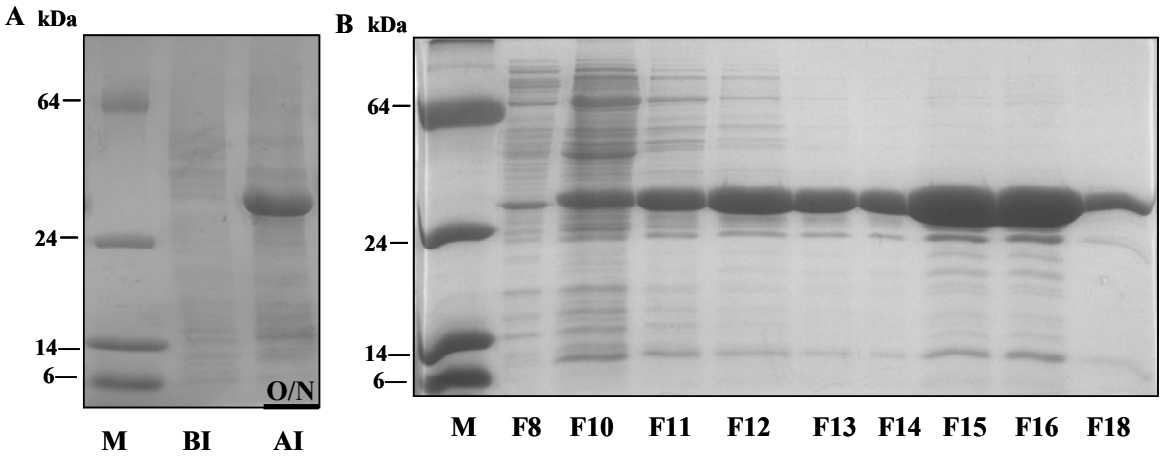


Figure 23. (A) EGFP-EF-III-172 expression in *E. coli* (Vector: PET20a, LB media, BL21(DE3). (B) EGFP-EF-III-172 purification SDS gel using 6x-his-tag affinity chromatography.

3.2.2.2 EGFP-WT and EGFP-EF-172 crystallization

Proteins are able to form crystals by aligning themselves in repeating series of "unit cells" forming a crystalline lattice, these crystalline lattices might be formed utilizing one of the most common crystallization techniques, the hanging-drop method [49]. In this method, a drop containing the purified protein, buffer, and precipitant is placed on a glass cover slip and it is allowed to equilibrate with larger reservoirs containing similar buffers and precipitants in higher concentrations and at different conditions.

To date, efforts to crystallize EGFP-WT in our laboratory have been unsuccessful. Samples were prepared for both EGFP-WT and EGFP-EF-172 at room temperature using the following precipitant solution: 10-15% PEG 4000 in 50 mM HEPES pH 8.0-8.5, 50 mM $MgCl_2$, 10 mM 2-mercaptopethanol. Phase separations and oil shape profiles were observed in the samples due either to low monovalent salt concentration or because the precipitant concentration was too high (Figure 24). The observed results indicate that salt and PEG concentrations need to be further optimized in order to obtain crystals for this protein.

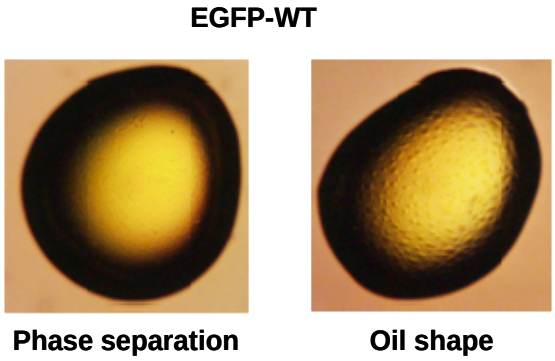


Figure 24. EGFP-WT crystallization. Phase separation and oil shape profiles were observed.

Similar conditions were used to grow crystals for EGFP-EF-172. Different PEG concentrations (10-15%) and different pH (8.0-8.5) were also tested. The best conditions that showed microcrystals were pH 8.0 and 8.4 and 10% and 13% PEG 4000 concentration (Figure 25). These conditions are still under investigation because these microcrystals appear to be overlapping each other, forming crystal clusters. These clusters can be avoided by reducing crystallization temperature in order to slow down the nucleation process.

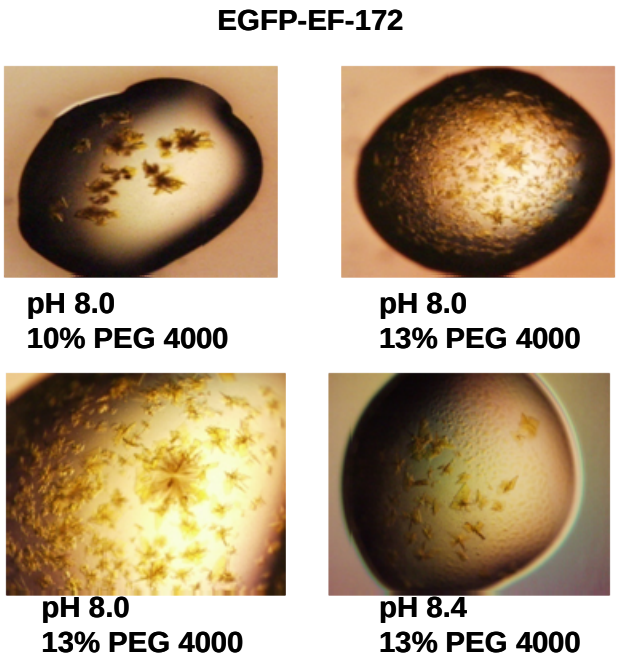


Figure 25. EGFP-EF-172 crystallization. Microcrystal formation and crystal clusters profiles were observed.

3.3 Summary

A grafting approach was used to insert Ca^{2+}-binding motifs into host proteins, CD2.D1 and EGFP. These host proteins are easily expressed and purified, tolerate mutations, and their native structures are not disrupted by strategic insertions of foreign sequences. Expression and purification of these engineered proteins are initial and essential steps to perform future physicochemical experiments (e.g. CD2, MS, NMR, titrations etc.). All engineered proteins were expressed at 37 ^{o}C in *E. coli* BL21(DE3) LB media conditions using IPTG as a protein inducer. CD2-WT, CD2.STIM1, and CD2.STIM1.Mut were purified using affinity (GST-tag) and ion exchange chromatography. Conversely, EGFP-WT and EGFP-172-III-172 were purified using 6x-his-tag affinity chromatography and the collected fractions were used to grow crystals using the hanging-drop method. Optimum salt and precipitant concentrations are still under investigation since we were unable to obtain EGFP-WT crystals, and crystal clusters were obtained for EGFP-EF-172. The purities of our engineered proteins were tested by SDS gel electrophoresis and their yields and concentrations were between 10-23 mg/L and 366-1800 μM in ~ 2 mL (Table 4).

Table 4. Concentrations and yields obtained for engineered proteins that were created using a grafting approach

Protein	Yield (mg/L)	Concentration (μM) in 2 mL
CD2.WT	23.4	1800
CD2.STIM1	14.4	1100
CD2.STIM1.Mut	10.0	771
EGFP-WT	73.6	657
EGFP-EF-172	10.2	366

4. CALMODULIN: AN INTRACELLULAR PROTEIN PURIFIED BY HYDROPHOBIC INTERACTION CHROMATOGRAPHY

4.1 Introduction

CaM is a small (~17 kDa) intracellular Ca^{2+}-binding protein with four EF-hand motifs or loops. This important and well-studied protein is able to regulate a series of important cellular functions including secretion, metabolism, motility, signal transduction, and cell division and growth [7]. Its primary structure is highly conserved in all cell types and it has been also identified in prokaryotic systems where it changes its conformation upon Ca^{2+} addition [50].

CaM has four Ca^{2+}-binding sites located in four different EF-loops since this protein has a helix-loop-helix conformation (Figure 26A). These loops have different physiochemical properties and amino acid residues that will influence their metal-binding affinities, especially its interaction with Ca^{2+} (shown as dark grey spheres). For example, EF-loop I has two positively changed residues at positions 2 and 11 while EF-loop IV contains an additional negatively charged residue at position 11 (Figure 26B). Previous researchers in Dr. Yang's laboratory have performed studies in which EF-loops of CaM were inserted into CD2 domain 1 (CD2.D1) in order to study their Ca^{2+}-binding properties and affinities [50]. Figure 26C illustrates the model structure of one of the grafted loops. CaM EF-loop III was grafted into CD2.D1 at position 52 connected by five Gly residues (3 at one end and 2 at the other end) in order to increase the loop's flexibility [50]. The insertion of the EF-loops did not change the host protein's conformation and it was concluded that EF-loop IV exhibited the weakest metal-binding affinity while EF-loop I exhibited the strongest binding affinity [50].

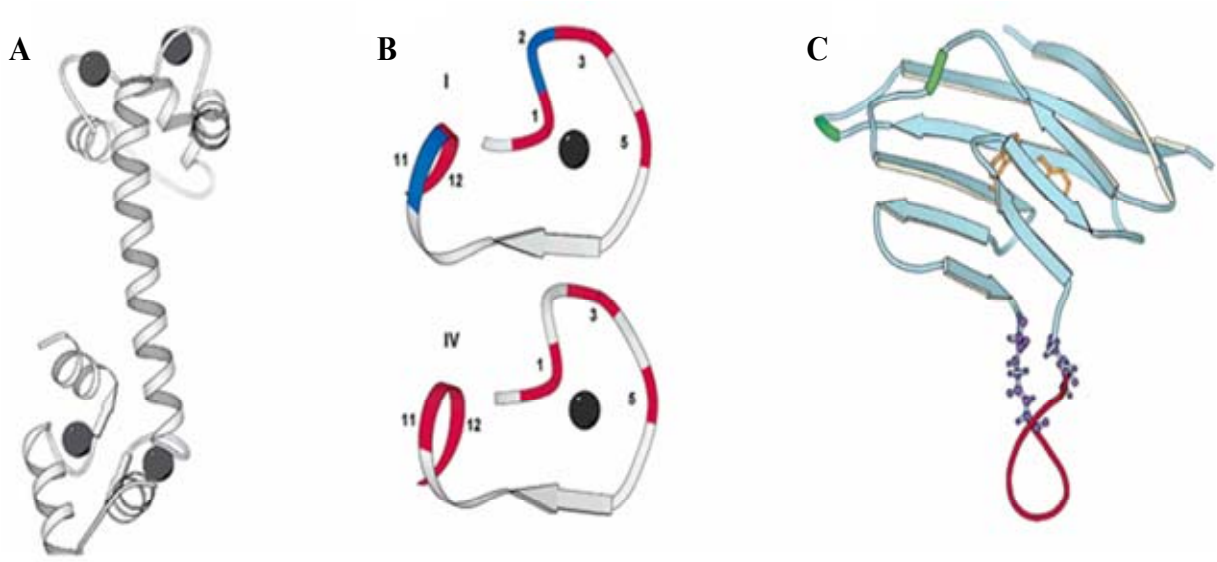


Figure 26. Three-dimensional structures of calmodulin using MOLSCRIPT.67. (A) Ribbon diagram of CaM-calcium loaded form. (B) Local structures of loop I (top) containing two positively charged residues and loop IV (bottom) containing an additional negatively charged residue at loop position 11. (C) Model structure of domain 1 of CD2 (CaM-CD2-III-5G-52) with EF-loop III grafted at position 52 (Trp-32 and Tyr-76 are highlighted in yellow) [50].

4.2 Results and discussion

4.2.1 CaM-WT expression using two growth media

CaM-WT was expressed in *E. coli* in a pET20b vector and using LB and ^{15}N labeled minimal media. The cell strain used was BL21(DE3) and the IPTG concentration used to induce protein expression was 150 mM in both media. The temperature for expression was different for both media since ^{15}N labeled minimal media provides fewer nutrients for bacterial growth compared to those provided by LB media. CaM-WT was successfully expressed at 37 ^{o}C in LB media and bacteria were collected after 6 h induction (Figure 27). In contrast, this protein was expressed at 25 ^{o}C in the minimal media and it was left in the shaker with agitation for an overnight period (Figure 27). A clear band was therefore observed around the 14 kDa marker band because the protein's molecular weight is ~17 kDa.

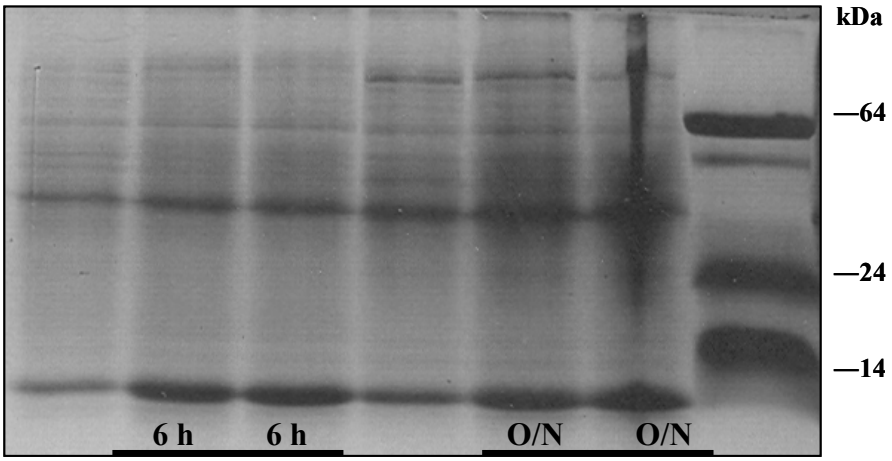


Figure 27. CaM-WT expression in two different growth media.

4.2.2 CaM-WT purification

CaM-WT was purified using hydrophobic chromatography and a phenyl-sepharose column. This well studied protein exposes a hydrophobic surface upon Ca^{2+} binding and this hydrophobic region allows for its purification since it can bind to the phenyl-sepharose column and it can be eluted out by using chelating agents like EGTA. The purification SDS gels for two proteins expressed in different media can be observed in Figure 28A (LB media) and Figure 28B (minimal media). The SDS gels clearly show that most of the protein was in the soluble form (Sup) after cells were disrupted with high pressure. It is also possible to observe that the protein bound completely to the phenyl-sepharose column because no protein was left in the waste (Wst) after intensive binding, and that the 8 mL eluted fractions contained high quantity of pure protein. Relevant collected fractions were concentrated and dialyzed in 20 mM Tris-HCL buffer, pH 7.4 and the dialyzed buffer was exchanged three times with time intervals between 2-3 h and sometimes for an overnight period.

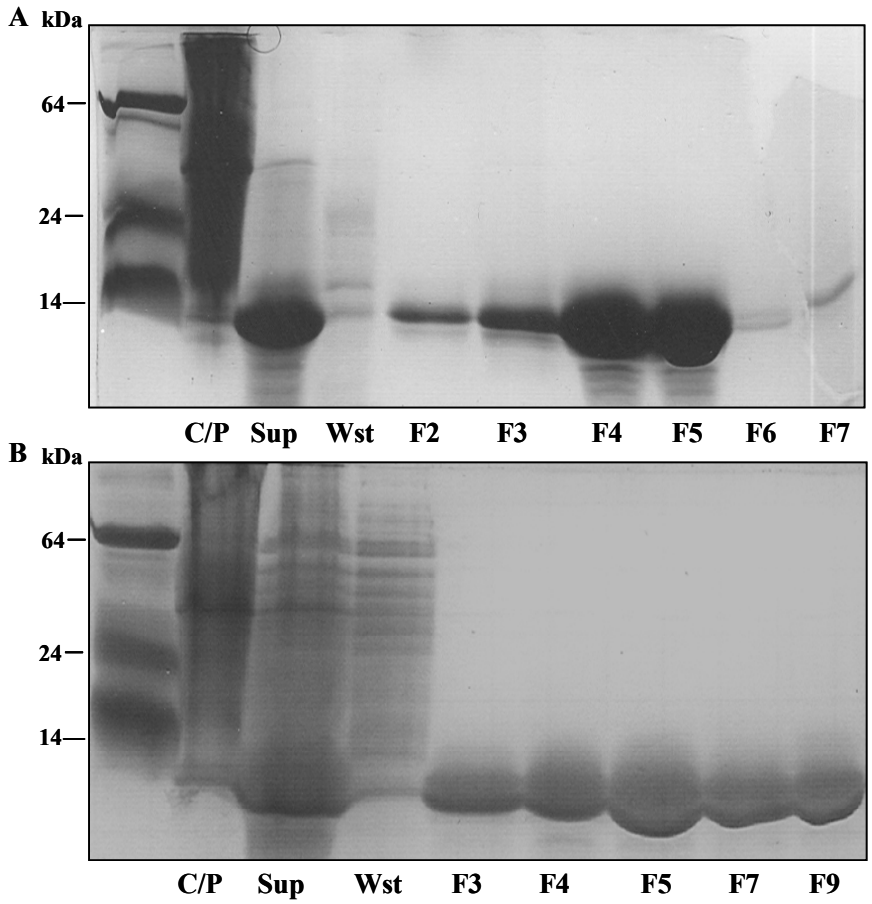


Figure 28. CaM-WT purification using hydrophobic chromatography. (A) CaM-WT purification and eluted fractions when protein was expressed in 1 L LB media. (B) CaM-WT purification and eluted fractions when protein was purified in 2 L ^{15}N labeled media.

Protein concentration was calculated using UV-visible spectroscopy, the extinction coefficient (3030 M^{-1} cm^{-1}), and the protein's absorbance at 277 nm. A typical UV-visible spectrum of CaM-WT can be observed in Figure 29 where it is also possible to observe that the UV absorbance at 277 nm for tyrosine and phenylalanine at 257 nm. For the pure protein expressed in LB media the $A_{277\ nm}$ was 0.2676, whereas for the pure protein expressed in minimal media the $A_{277\ nm}$ was 0.2074. The obtained absorbances at 277 nm were divided by the extinction coefficient and multiplied by a dilution factor of 50. The calculated final concentrations and yields of purified proteins were 4.4 mM and 74.8 mg/L (expressed in LB

media) and 3.4 mM and 57.8 mg/L (expressed in ^{15}N labeled minimal media) in ~2 mL. It is normal to obtain a higher amount of protein when the protein is expressed in LB media since the nutrients provided by these media are higher than those provided by minimal media.

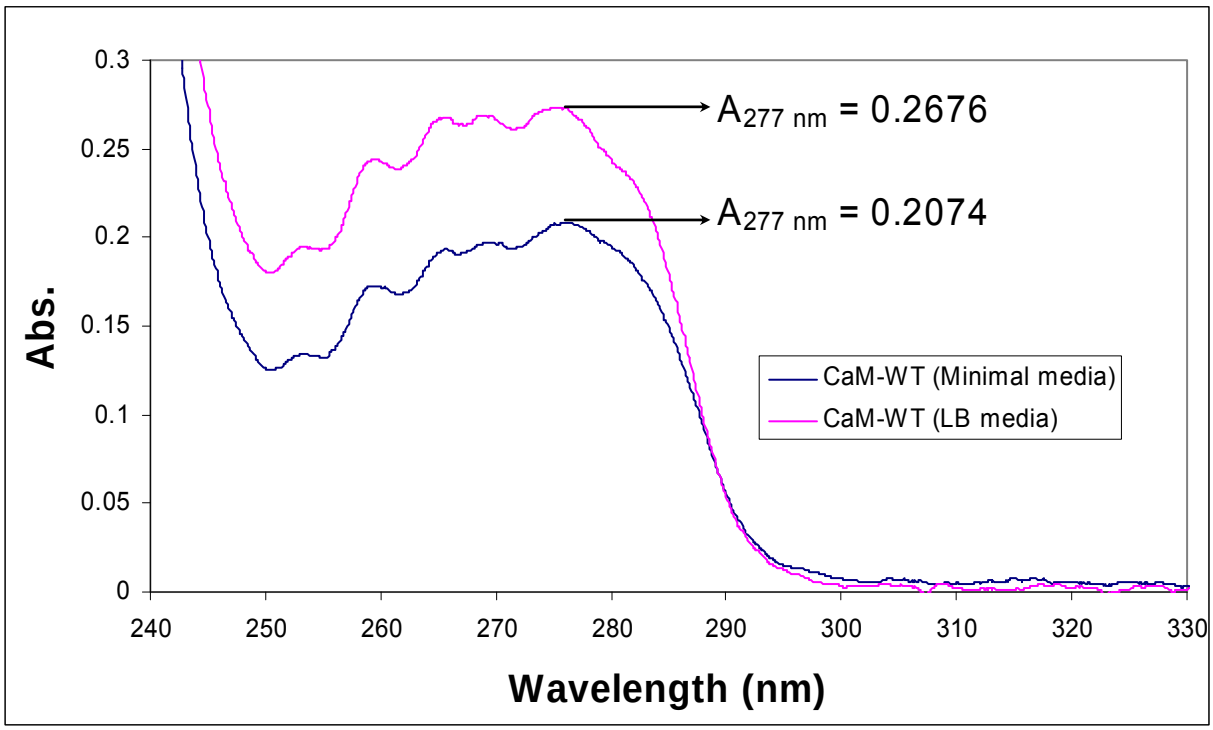


Figure 29. CaM-WT UV-visible spectrum after protein was expressed in two growth media and purified by hydrophobic chromatography. Protein concentrations were calculated using the protein's absorbance at 277 nm, its extinction coefficient, and Beer-Lambert Law.

4.2.3 Summary

CaM is an important intracellular Ca^{2+}-binding protein with four EF-hand motifs that undergoes a conformation change upon Ca^{2+} binding. CaM-WT was successfully expressed in *E. coli* BL21(DE3) using two different media and temperatures. Expression in LB media was performed at 37 ^{o}C whereas expression in ^{15}N minimal media was performed at 37 ^{o}C overnight because this media provided fewer nutrients to the cells and it required a longer incubation time. CaM-WT was subsequently purified using hydrophobic interaction chromatography (phenyl-

sepharose column) since this protein exposes a hydrophobic region when it binds to Ca^{2+} allowing its purification. The final yield for CaM-WT expressed in LB media was 74.8 mg/L which was higher to that obtained when it was expressed in ^{15}N minimal media (57.8 mg/L) due to the media nutrients and the expression time and temperature. Future experiments were performed with the purified protein. For example, CaM-WT expressed in ^{15}N minimal media was used to perform NMR experiments to further study the effect that Ca^{2+} might have on this protein's molecular structure.

5. THE ROLE OF CALCIUM AND THE EXTRACELLULAR CALCIUM-SENSING RECEPTOR

5.1 Introduction

The CaSR was cloned from bovine parathyroid tissue in 1993 by Dr. Brown and his colleagues and since then researchers have been rigorously working in order to study this important receptor [51-53]. The receptor is most highly expressed in organs like the parathyroid, thyroid, and kidney and it is a key and crucial player in Ca^{2+} homeostasis maintenance since it regulates the secretion of parathyroid hormone (PTH) [11, 52, 54, 55]. For example, primary hyperparathyroidism (HPTH) involves an over activity of the parathyroid glands causing PTH overproduction and elevated blood Ca^{2+} levels or a hypercalcemia condition [56]. Overproduction of PTH can lead to osteoporosis, kidney stones, peptic ulcers, and high blood pressure [55]. Several diseases have been associated with mutations in the CaSR that are intrinsically linked with Ca^{2+} homeostasis disorders. For instance, familial hypocalciuric hypercalcemia (FHH) and neonatal severe hyperparathyroidism (NSHPT) were associated with inactivating mutations affecting one or both alleles [52].

The CaSR belongs to family C of G protein-coupled receptors and it recognizes Ca^{2+} as its physiological ligand. It is comprised of a huge extracellular domain (ECD) of approximately 600 amino acid residues, a seven transmembrane domain (7 TMD), and an intracellular domain of approximately 200 amino acids [11, 55, 57]. The CaSR is a dimer and its ECD is connected to the 7 TMD by a cysteine-rich region. The crystal structure for this receptor has not yet been elucidated yet but a model structure based on the X-ray structure of mGLuR1 has been developed in order to illustrate the three important domains that comprise the CaSR (Figure 30).

This model structure illustrates the dimeric form of the CaSR and it also shows the location of the five predicted Ca^{2+}-binding sites.

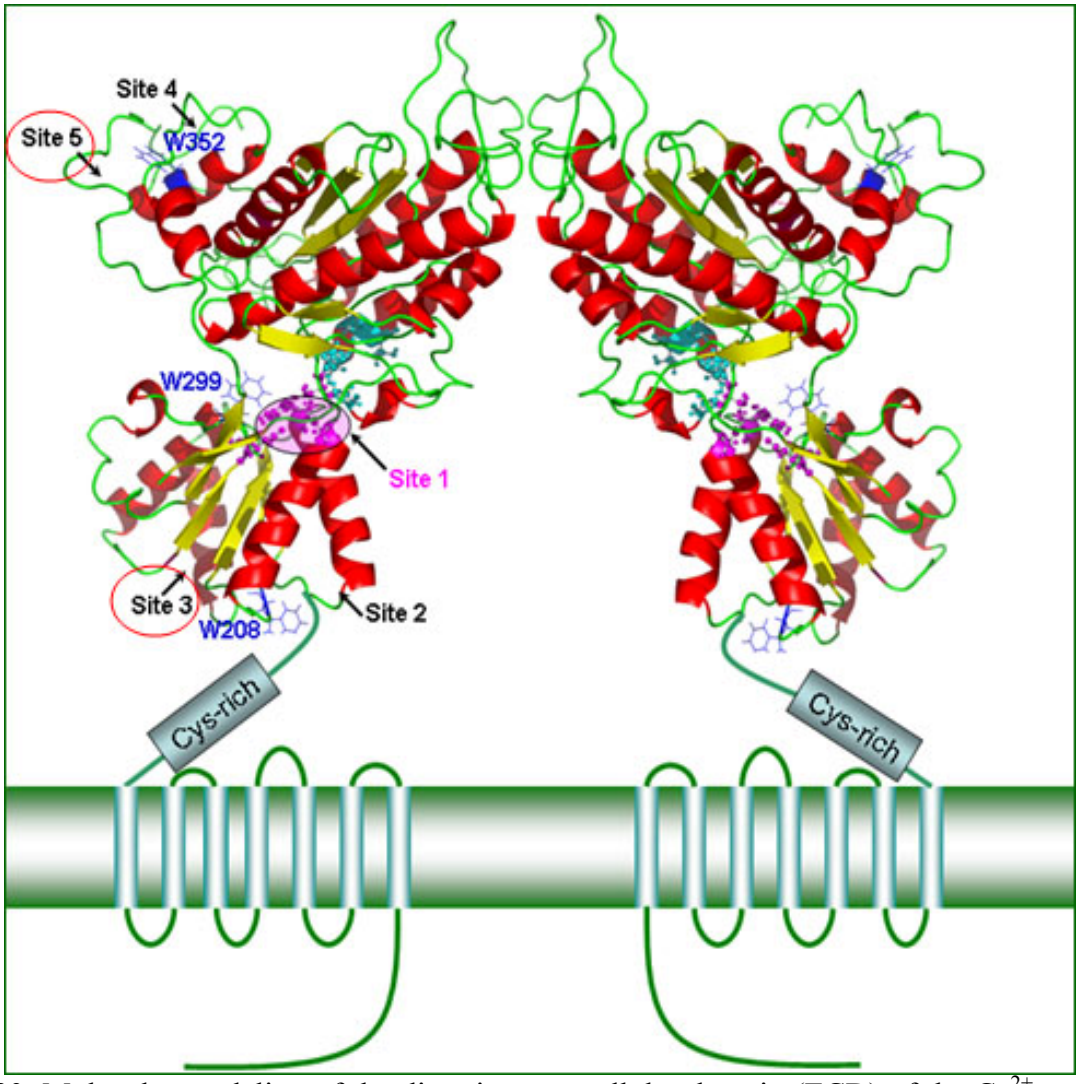


Figure 30. Molecular modeling of the dimeric extra cellular domain (ECD) of the Ca^{2+}-sensing receptor (CaSR) based on the ECD X-ray structure of mGluR1 [58].

5.2 Two approaches to study the metal-binding properties of the CaSR

5.2.1 Grafting approach

Five putative Ca^{2+}-binding sites were predicted and two of them (site 3 and 5) were grafted into CD2.D1 using a well-studied grafting approach developed in Dr. Yang's lab. Therefore, we have converted CD2 into a specific receptor for Ca^{2+} by inserting a Ca^{2+}-binding

site into CD2.D1 between glycine 53 (G53) and serine 52 (S52) (Figure 31). The location of the Ca^{2+}-binding sites on the ECD (Figure 32A) and the proposed model structures for all five sites is illustrated in (Figure 32B) where sites 3 and 5 are circled in red. Lobe 1 of the ECD of the CaSR contains sites 4 and 5, whereas lobe 2 contains sites 1, 2, and 3. Site 1 is expected to be important in Ca^{2+}-induced conformational changes since this site is located at a hinge region connecting both lobes. Additionally, the physiochemical properties of CD2.CaSR site 3 and 5 and their mutants are summarized in Table 5 where it is possible to observe the individual grafted sequences, isoelectric points (pI), and molecular weights (kDa) for sites 3 and 5 and their mutants. Glycine linkers were added at each insertion site in order to study the loop's flexibility upon linker addition.

Table 5. Grafted sequences, isoelectic points (pI), and molecular weights for CD2.CaSR site 3, site 5, and mutants

Protein	Grafted Sequence	pI	MW (kDa)
CD2.CaSR5	GHEESGDRFSNSSTAFRPLCTGDENI	5.97	14.1
CD2.CaSR5.Mut1	GHAASGDRFSNSSTAFRPLCTGDENI	7.38	14.0
CD2.CaSR5.Mut2	GHEESGDRFSNSSTAFRPLCTGAANI	7.38	14.0
CD2.CasR5 linker	KSGGGHEESGDRFSNSSTAFRPLCTGDENIGGGSA	7.38	12.9
CD2.CaSR3	GIEKFREEAEERDIG	5.59	13.0
CD2.CaSR3.Mut1	GIAKFREEAEERDIG	7.19	12.9
CD2.CaSR3.Mut2	GIEKFRAAAEERDIG	7.19	12.9
CD2.CaSR3.Mut3	GIEKFREEAAARDIG	7.19	12.9
CD2.CaSR3 linker	KSGIEKFREEAEERDIGA	5.59	13.3

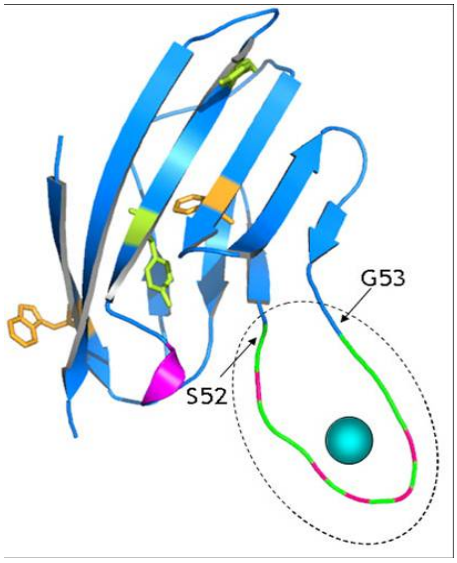
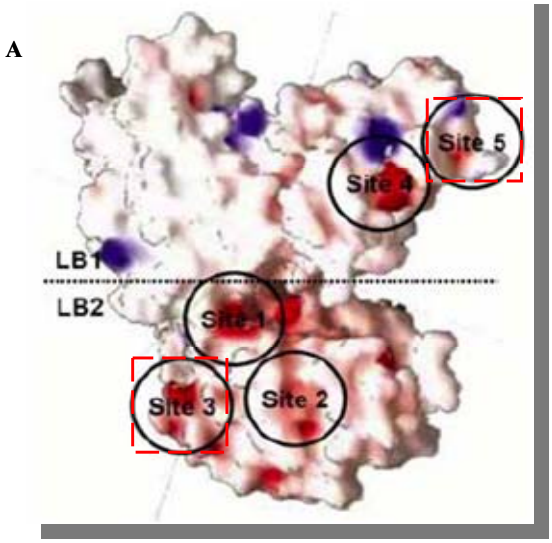


Figure 31. Model structure of CD2.D1 grafted with a putative Ca^{2+}-binding site inserted between serine 52 and glycine 53.



Predicted sites	Residues
○ Site1	S147, S170, D190, Y218, E297
△ Site2	D215, L242, S244, D248, Q253
◇ Site3	E224, E228, E229, E231, E232
☆ Site4	E350, E353, E354, N386, S388
▯ Site5	E378, E379, T396, D398, E399

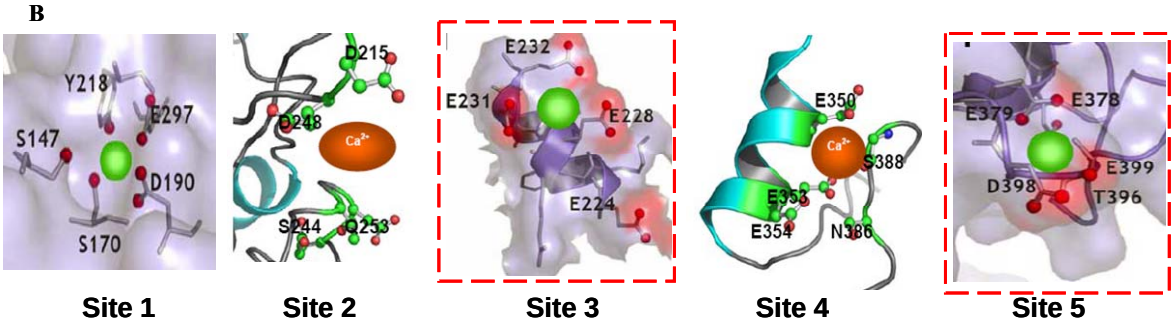


Figure 32. (A) ECD of the CaSR illustrating possible locations of five potential Ca^{2+}-binding sites in lobes 1 and 2 and the predicted residues at each site. (B) Proposed structures for potential Ca^{2+}-binding sites in the ECD of the CaSR. Site 3 and site 5 shown in red.

5.2.2 Subdomain approach

The strategy of the subdomain approach is to chop intact CaSR into three subdomains containing the identified Ca^{2+}-binding sites. The objective of this strategy is to understand the cooperativity of coupled Ca^{2+}-binding sites and to study Ca^{2+}-induced conformational changes [10]. Three constructs were designed encompassing subdomains that include different numbers of predicted Ca^{2+}-binding sites of the CaSR. Subdomain 1 contains CaSR sites 1, 2 and 3, subdomain 2 contains sites 2 and 3 in lobe 2, and subdomain 3 contains sites 4 and 5 in lobe 1 (Figure 33).

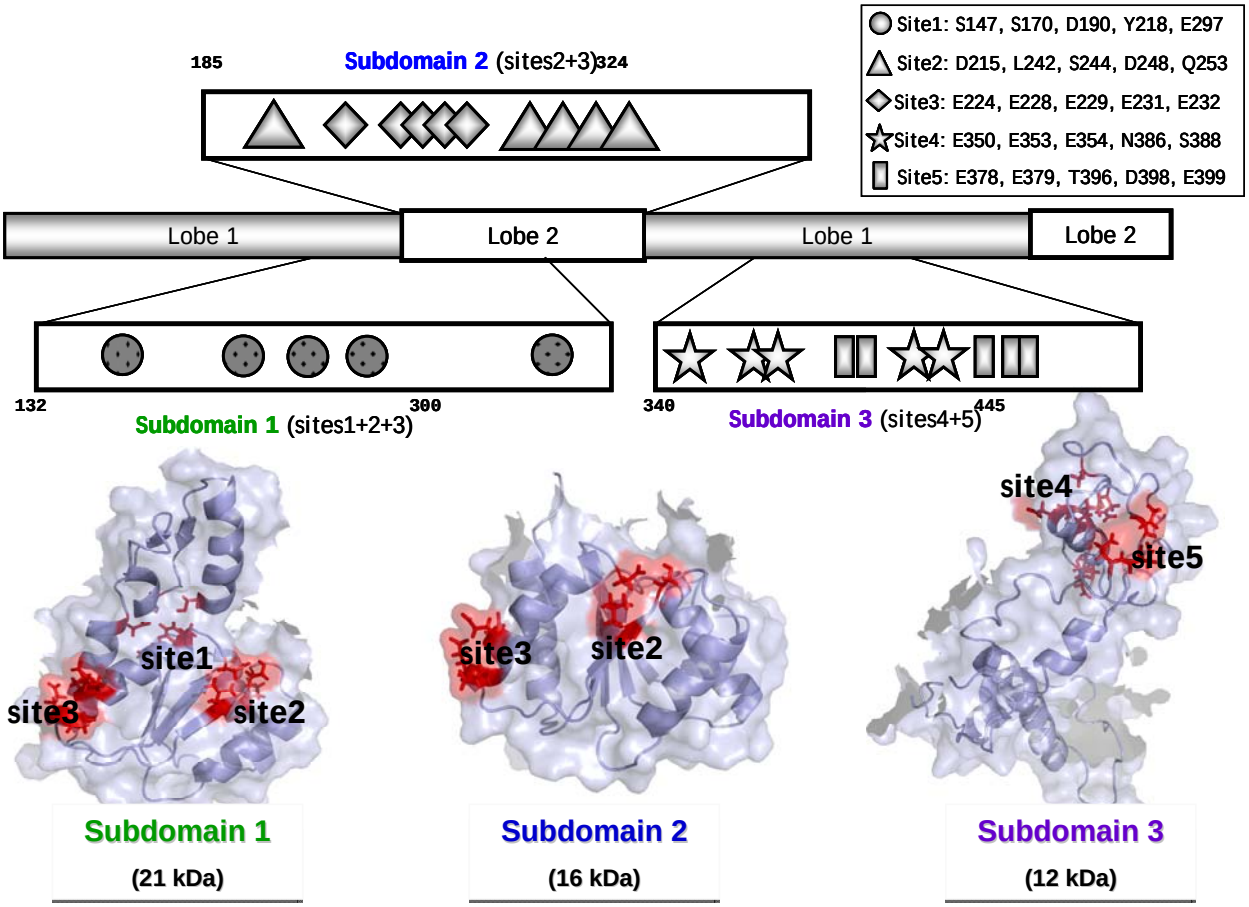


Figure 33. Diagram illustrating the subdomain approach: ECD of the CaSR is chopped into 3 subdomains. Subdomain 1 contains sites 1, 2, and 3. Subdomain contains sites 2 and 3. subdomain 3 contains sites 4 and 5.

The physiochemical properties of the three subdomains including the extinction coefficient, the isoelectric point, the number of cysteine residues, and the molecular weight (kDa) can be observed in Table 6.

Table 6. Physiochemical properties of three subdomains located in the ECD of the CaSR

Protein	Amino acid from CaSR	Extinction Coefficient (M^{-1} cm^{-1})	Cys No.	pI	MW (kDa)
Subdomain 1	S132-A300	36,130	1	4.59	22.9
Subdomain 2	R185-A324	34,850	1	4.62	20.3
Subdomain 3	A323-G494	27,550	5	5.05	23.8

5.3 Results and discussion

5.3.1 CD2.CaSR site 3 and 5 expression and purification using GST-fusion technology

5.3.1.1 CD2.CaSR site 3 and 5 expression in LB media

Site 3 and site 5 from the CaSR were inserted into CD2.D1 and they were successfully expressed as fusion proteins with a GST-tag in a pGEX-2T vector transformed into *E. coli* cell strain BL21(DE3) and IPTG was used as the inducer. The exponential expression growth curve clearly shows that the proteins reached an OD at 600 nm of approximately 2.0 300-350 min (3-4 hrs) after induction (Figure 34A). SDS gels confirmed that the proteins were successfully expressed since protein expression was not observed before induction (BI) but after IPTG induction (AI) protein's expression increased every hour as the bands were getting thicker and thicker as the hours were elapsing (Figure 34B).

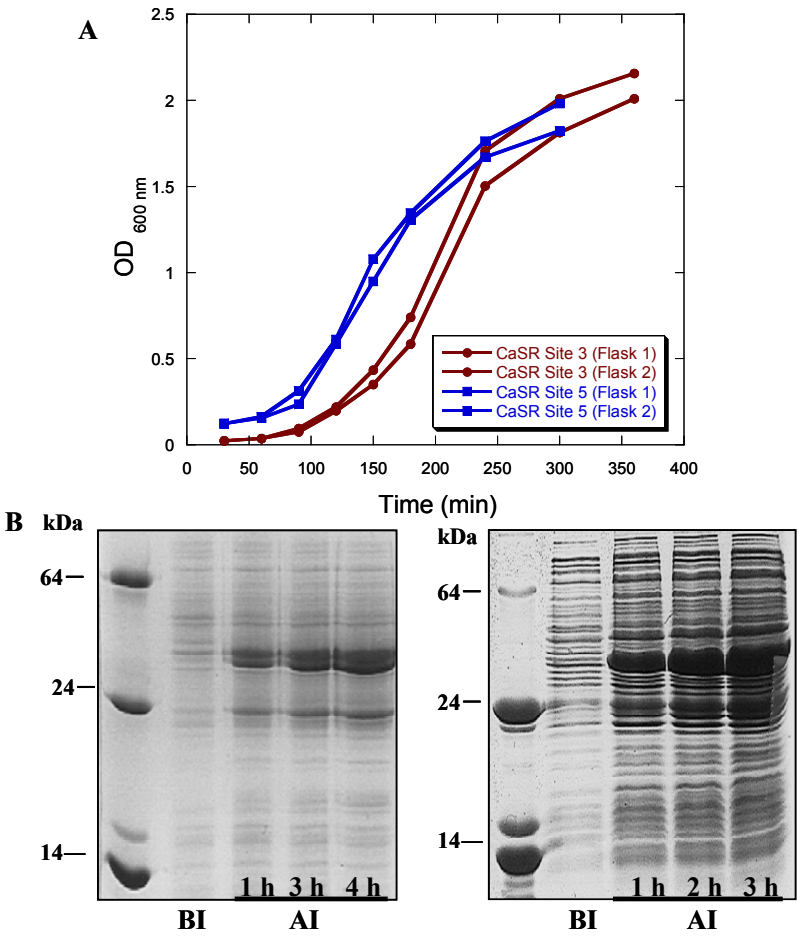


Figure 34. (A) CD2.CaSR site 3 and site 5 expression growth curve (Vector: pGEX-2T, LB media, cell strain: BL21(DE3), 37 ^{o}C expression). (B) SDS gels showing CD2.CaSR site 3 (left) and site 5 (right) expression.

These two proteins were purified using affinity (GST-tag) and ion-exchange chromatography. Engineered proteins are frequently subject to solubility problems but in this case the amount of soluble protein (in the Sup form) that was obtained after cell disruption was enough to perform future purification steps. The supernatant was poured into sepharose-columns where the GST was initially bound to the column (BAB), subsequently cleaved using thrombin (BAC), and eluted out (Elu) using 1X PBS buffer (Figure 35). The pH of the eluted protein was then reduced to ~3.0 using acetic acid for subsequent injection into an FPLC system containing an ion-exchange column. The chromatogram revealed relevant fractions that were collected and

subjected to dialysis in 10 mM Tris buffer pH 7.4. Both proteins showed a clear and pure band around ~14 kDa (site 5) and ~13 kDa (site 3) on the SDS gels Figure 35A and Figure 35B respectively.

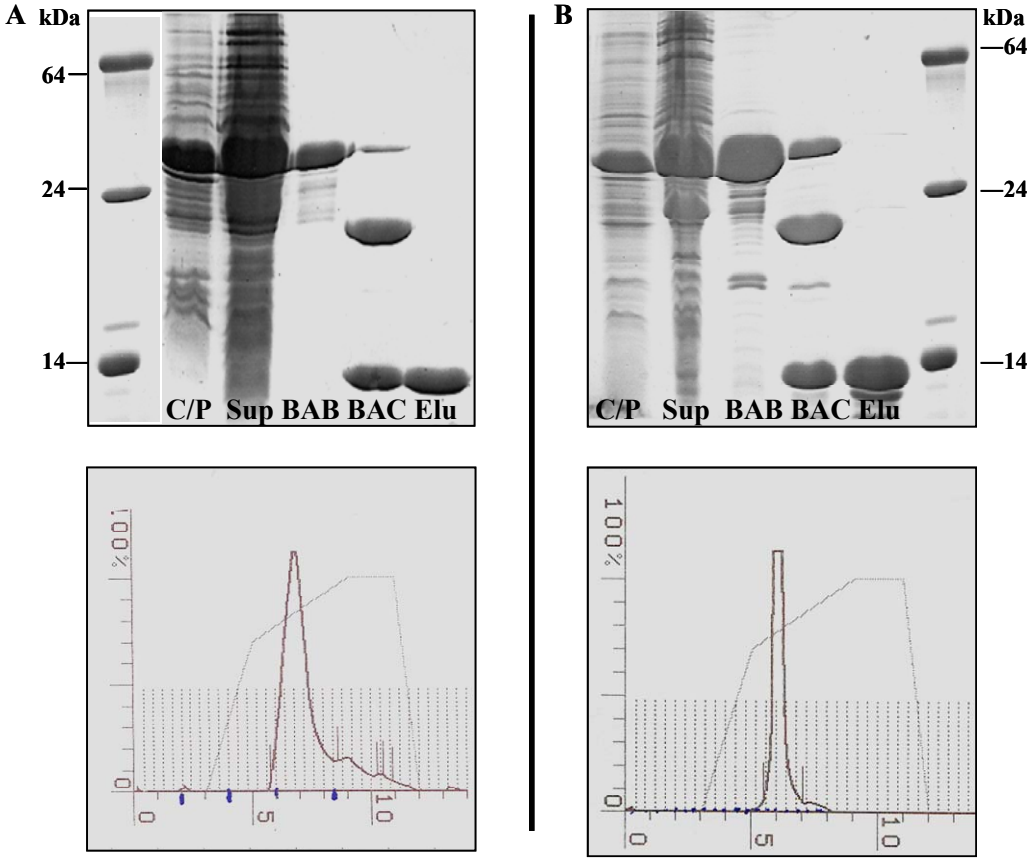


Figure 35. CaSR site 3 and site 5 purification using affinity (GST-tag) and ion exchange chromatography. SDS gels and ion exchange chromatograms for (A) CaSR site 5 and (B) CaSR site 3.

5.3.1.2 CD2.CaSR site 3 and 5 expression in minimal media

Additionally, efforts were made to express and purify CD2.CaSR site 3 and 5 in ^{15}N labeled minimal media for NMR analysis to obtain structural information of the predicted sites. Both proteins were successfully expressed and purified using similar approaches as those employed for regular LB expression. The expression level was reduced to that observed in LB media expression because minimal media provided fewer nutrients for bacterial growth (Figure

36A). Since the expression level was reduced, the final eluted protein amount was also reduced when it was purified by affinity and ion exchange chromatography as confirmed by the SDS gels of both proteins (Figure 36B).

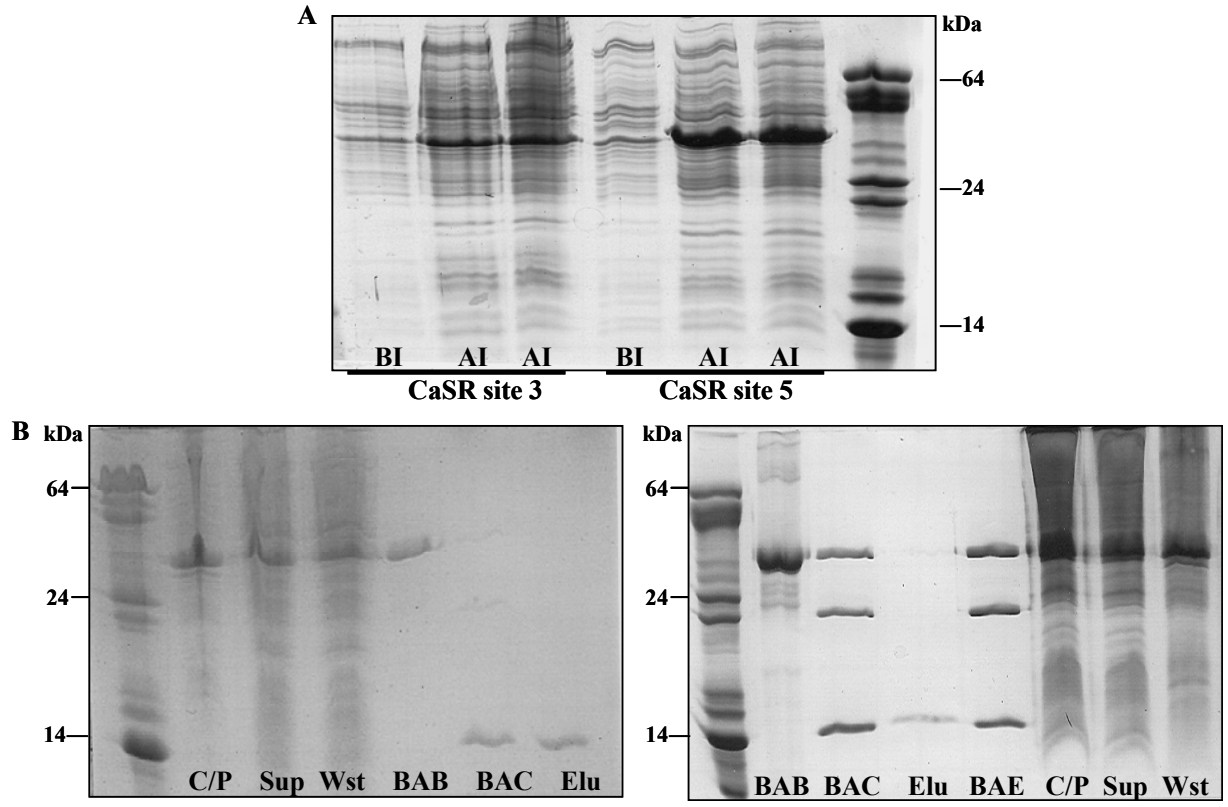


Figure 36. (A) Expression of CaSR site 3 and site 5 in ^{15}N labeled minimal media (BL21(DE3), 37 ^{o}C expression). (B) Purification of CaSR site 3 (left) and site 5 (right) by GST-tag affinity column.

5.3.2 Subdomain 1 expression and purification optimization

5.3.2.1 Expression and cell disruption using different vectors

Subdomain 1 expression was performed using fusion technology in order to compare the protein's expression and solubility after cell disruption. Three different vectors were evaluated: pGEX-2T (GST-tag), pET32a (6x-his and thx-tag), and pRSETa (6x-his tag). The expression

conditions were the same (*E. coli* BL21(DE3) at 37 ^{o}C) for all three vectors. Subdomain 1 showed a higher expression level when it was expressed using pGEX-2T and pET32a, but it was also expressed to a lower extent in pRSETa (Figure 37A). Protein solubility was investigated by breaking the cell using sonication. The protein formed inclusion bodies in the C/P in all three cases as observed in Figure 37B. Most of the protein was therefore identified within the C/P and not in the Sup form. The only vector that showed a significant amount of protein in the soluble form was pET32a but the thx-tag could not be removed from the target protein.

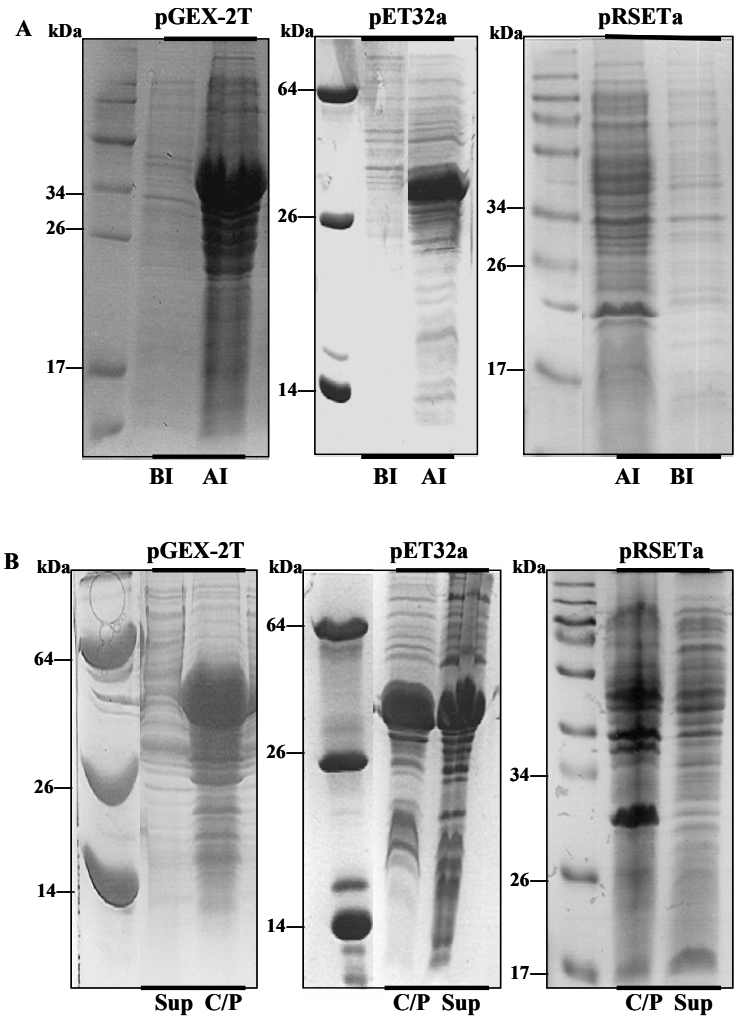


Figure 37. Subdomain 1 expression and cell disruption using three vectors. (A) Protein was expressed at 37 ^{o}C in *E. coli* BL21(DE3) using pGEX-2T (GST-tag), pET32a (6x-his-thx-tag), and pRSETa (6x-his-tag). (B) Cell disruption in order to investigate subdomain 1 solubility.

5.3.2.2 Subdomain 1 expression optimization using pRSETa

Subdomain 1 was cloned in pRSETa vector between *Bam*HI and *Eco*RI restriction sites by Dr. Yun Huang. This vector contained a 6x-his-tag region that was used for further purification. The 6x-his-tag not only facilitated the subdomain 1 purification, but the protein's expression was further optimized by utilizing different cell strains and temperatures.

Protein was transformed and expressed in four *E. coli* cell strains: BL21(DE3), BL21(DE3)pLysS, tuner(DE3)pLacI, and rosetta(DE3)pLysS. These cells strains have different properties and these properties had diverse effects in subdomain 1 expression and further purification. All the chosen cell strains were T7-based expression systems that control proteases at the cellular level. The pLysS plasmid provided a reduction of basal expression and therefore a tight expression was expected. Rosetta cell strain offered a universal translation because it provided tRNAs that rarely occur in *E. coli*. Additionally, both rosetta and tuner had deletions in the *lacY* gene (lac permease) that allowed these cell strains to control expression levels upon different IPTG concentrations.

Cells were induced when the OD reached ~0.6 with IPTG. The expression growth curves show that the maximum OD or the maximum amount of bacteria that grew in the media fell between 1.0-1.2 when the protein was expressed at 37 ^{o}C (Figure 38A), and that the OD was higher between 1.2-1.8 when the protein was expressed at 30 ^{o}C overnight (Figure 38B). A higher OD reading was observed for BL21(DE3) at both temperatures and rosetta(DE3)pLysS showed the lowest OD reading at 30 ^{o}C. These graphs illustrate typical exponential bacterial cell growth curves where bacteria exhibit a slow initial growth rate (~0 to 200 min) which increases to a faster exponential growth rate (~200 to 400 min) until it reaches a stationary phase (~400 to 1000 min).

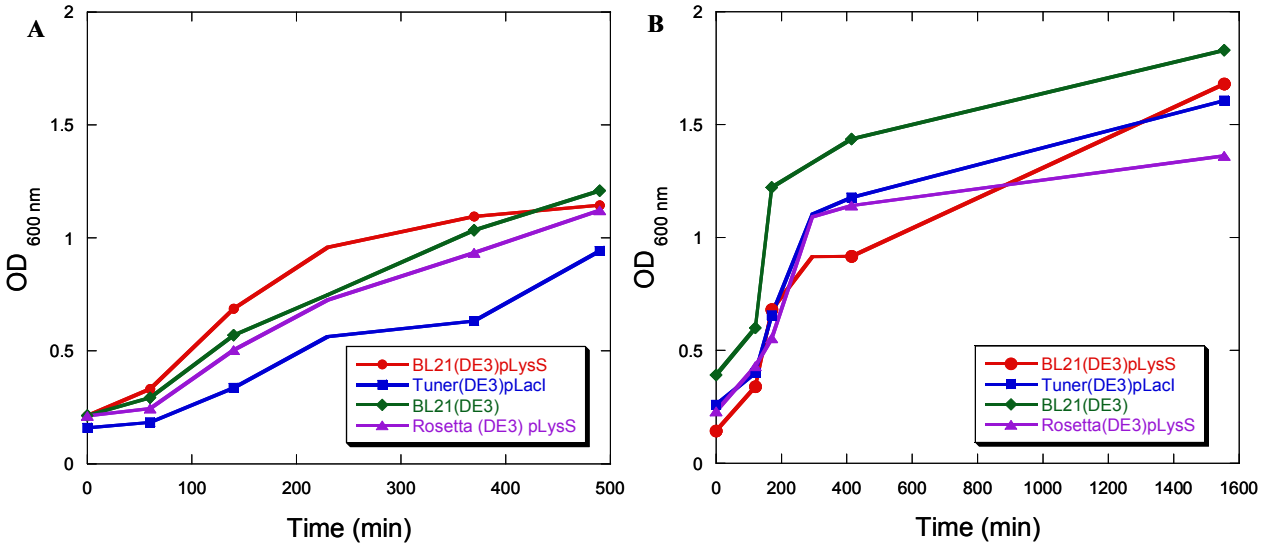


Figure 38. Subdomain 1 expression growth curve using four cell strains and two temperatures. (A) Protein expressed at 37 ^{o}C (B) Protein expressed at 30 ^{o}C overnight.

Additionally, the SDS gel shows that subdomain 1 was successfully expressed in all cell strains at 37 ^{o}C and that the expression band was thicker when this protein was expressed using rosetta(DE3)pLysS (Figure 39A). Additionally, this protein was not expressed when BL21(DE3) was used as the cell strain at 30 ^{o}C overnight even though its OD was the highest among the other three cell strains (Figure 38B and Figure 39B) . This means that having a high OD does not guarantee that the protein is being expressed correctly. In contrast, subdomain 1 was successfully expressed in the other three cell strains at 30 ^{o}C expression. The best cell strains that expressed subdomain 1 were rosetta(DE)pLysS and tuner(DE3)pLacI at 30 ^{o}C overnight (Figure 39B). The intensity change was calculated using ImageJ software (National Institutes of Health, Bethesda, MD). The cell strain that showed the higher intensity change BI and AI was rosetta(DE3)pLysS at both temperatures and the intensity change was more evident when cells were expressed at low temperatures (Figure 39C). This means that lowering the

temperature increased protein expression since proteins had more time to properly fold and the bacteria was able to grow longer since expression at lower temperatures is usually performed overnight.

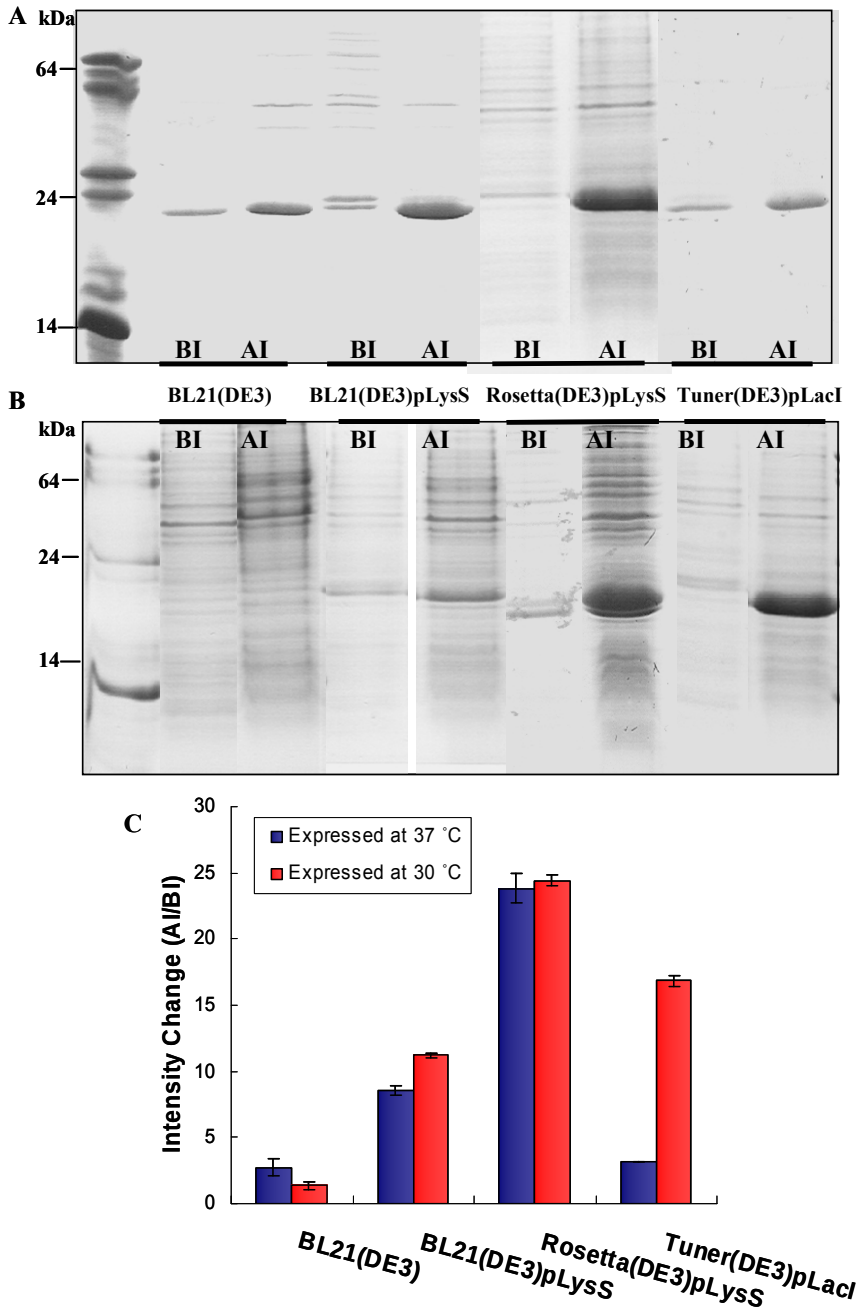


Figure 39. Subdomain 1 expression optimization using four cell strains and two temperatures. (A) Protein was expressed at 37 °C (B) Protein expressed at 30 °C overnight. (C) Intensity change AI/BI for subdomain 1 expression using different cell strains and temperatures.

5.3.2.3 Subdomain 1 purification using affinity chromatography

Expressed proteins were further purified using affinity chromatography and a Ni^{2+}-sepharose column using the 6x-his rich region provided by the pRSETa vector. Cells were disrupted using French press and it was found that a larger proportion of insoluble aggregates was formed when the protein was expressed at 37 ^{o}C. The cell strain that formed the highest amount of insoluble aggregates present in the C/P was BL21(DE3) and as a consequence this strain presented the lowest yield (1.2 mg/L) and concentration (26 μM) among all other cell strains (Table 7 and (Figure 40A). In comparison, BL21(DE3)pLysS and rosetta(DE3)pLysS exhibited better yields (4.0 mg/L and 6.2 mg/L) and concentrations (88 μM and 135 μM respectively) and, although the protein still formed inclusion bodies, a significant amount of protein was soluble and could be further injected into the FPLC system and eluted out with high (50 mM) imidazole concentration (Table 7, Figure 40B, and Figure 40C). Relevant eluted fractions from the FPLC system were collected, concentrated down, and dialyzed in 10 mM Tris-HCl buffer, pH 7.4 in order to remove imidazole from the protein.

Table 7. Subdomain 1 concentrations and yields after 6x-his-tag affinity chromatography

Cell strain	Concentration (μM)* (37 ^{o}C)	Yield (mg/L) (37 ^{o}C)	Concentration (μM)* (30 ^{o}C)	Yield (mg/L) (30 ^{o}C)
BL21(DE3)	26	1.2	n/a	n/a
BL21(DE3)pLysS	88	4.0	306	14.0
Tuner(DE3)pLacI	n/a	n/a	503	23.0
Rosetta(DE3)pLysS	135	6.2	788	36.1

* Protein volume ~ 2 mL

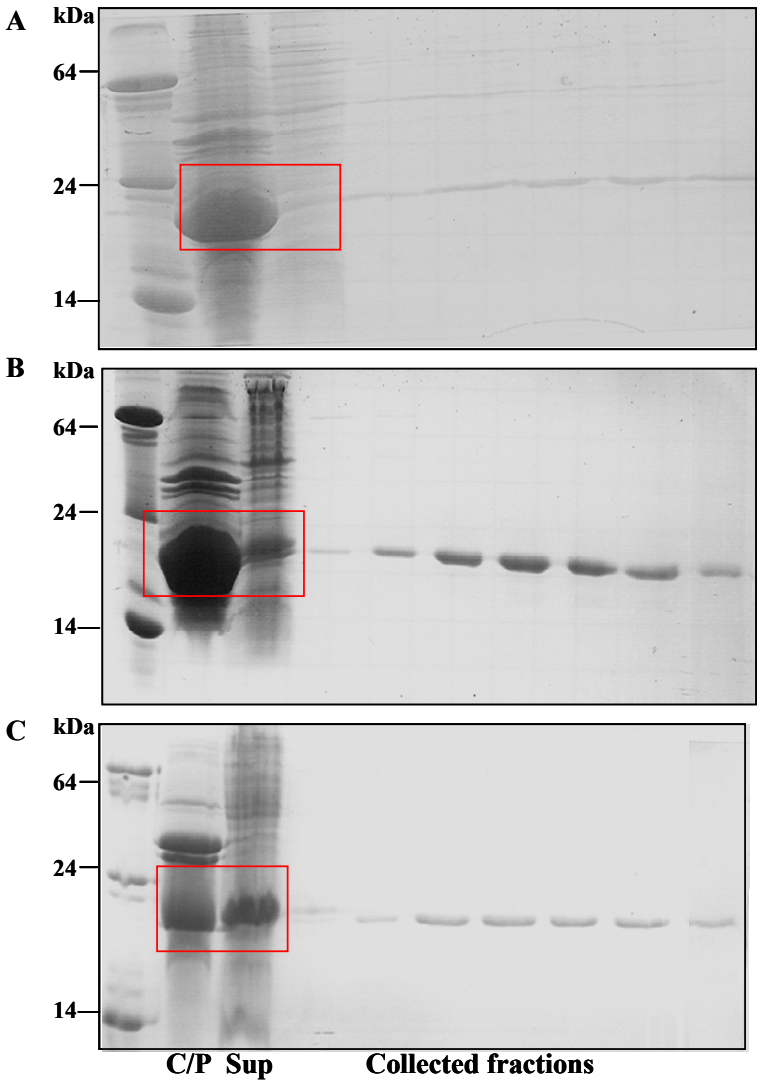


Figure 40. Subdomain 1 purification using 6x-his-tag affinity chromatography. Protein was expressed in LB media at 37 ^{o}C using three different cell strains: (A) BL21(DE3) (B) BL21(DE3)pLysS and (C) Rosetta(DE3)pLysS. Red boxes: most of the protein formed inclusion bodies or insoluble aggregates in the cell pellet and a low amount of soluble protein was present in the supernatant.

Subdomain 1 was more soluble when it was expressed at 30 ^{o}C and these results can be corroborated by the purification SDS gels when they were expressed in three different cell strains (Figure 41). Inclusion bodies were still present in C/P but to a lesser extent compared to expressed protein at 37 ^{o}C. We were able to obtain higher yields and concentrations when the protein was expressed at a lower temperature (Table 7 and Figure 42). Lower temperature

expression improved the protein's solubility because bacteria grew at a slower rate and the protein had more time to fold properly (Figure 41A-C). The cell strain that had the highest yield and concentration (in ~2mL) was rosetta(DE3)pLysS (36.1 mg/L and 788 µM), followed by tuner(DE3)pLacI (23.0 mg/L and 503 µM), and BL21(DE3)pLysS (14.0 mg/L and 306 µM) (Table 7).

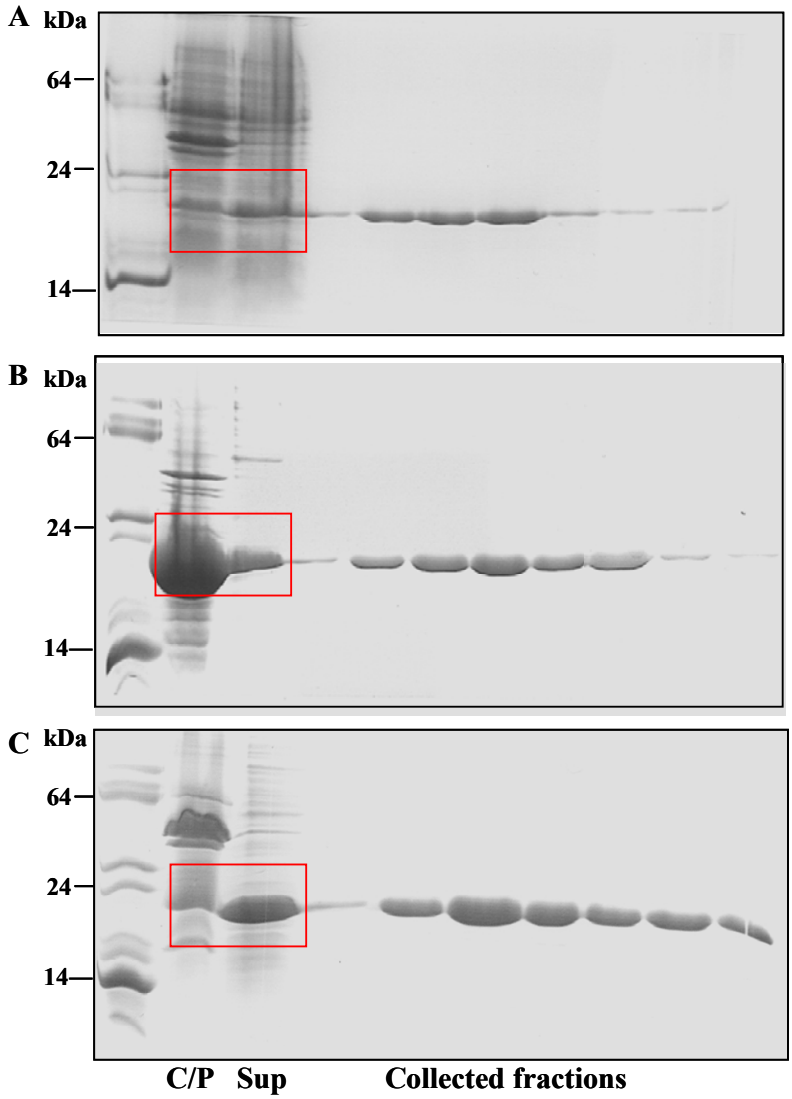


Figure 41. Subdomain 1 purification using 6x-his-tag affinity chromatography. Protein was expressed in LB media at 30 °C overnight using three different cell strains. (A) BL21(DE3)pLysS, (B) Tuner(DE3)pLacI, and (C) Rosetta(DE3)pLysS. Red boxes: protein formed a lower amount of insoluble aggregates in the cell pellet compared to protein expression at 37 °C.

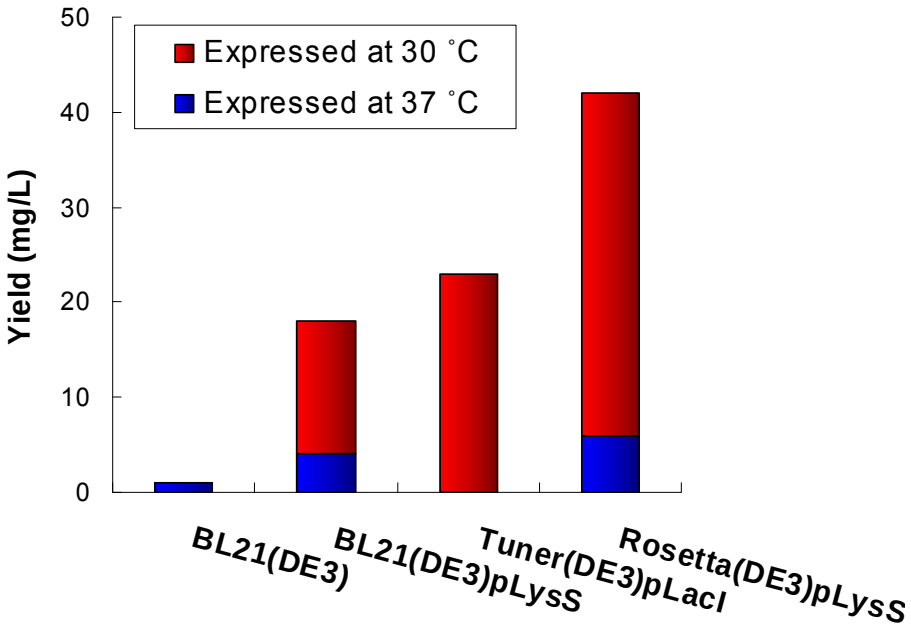


Figure 42. Subdomain 1 final yields were affected by expression temperatures and cell strains.

Soluble and insoluble fractions in the supernatant and in the cell pellet were compared using purification SDS gels and the respective fractions were calculated using ImageJ software (Figure 43). It was observed that soluble fractions were only 16% and 41% for BL21(DE3)pLysS and rosetta(DE3)pLysS respectively when the protein was expressed at 37 ^{o}C (Figure 43A). However, when the protein was expressed at 30 ^{o}C, soluble fractions for BL21(DE3)pLysS and rosetta(DE3)pLysS remarkably increased to 57% and 83% (Figure 43B). The cell strains that produced the higher percentages of insoluble aggregates were BL21(DE3) expressed at 37 ^{o}C with 92% in the insoluble form and tuner(DE3)pLacI at 30 ^{o}C with 85%. Tuner(DE3)pLacI was insoluble to a large extent due to overexpression of the protein resulting in production of inclusion bodies. Since this cell strain is IPTG dependent, protein expression was not regulated. On the other hand, better results were obtained with cell strains that had a pLysS plasmid because the cell strains that contained this plasmid allowed a tight and selective

expression. The results shown in Figure 43 confirmed that lower expression temperature enhanced protein solubility.

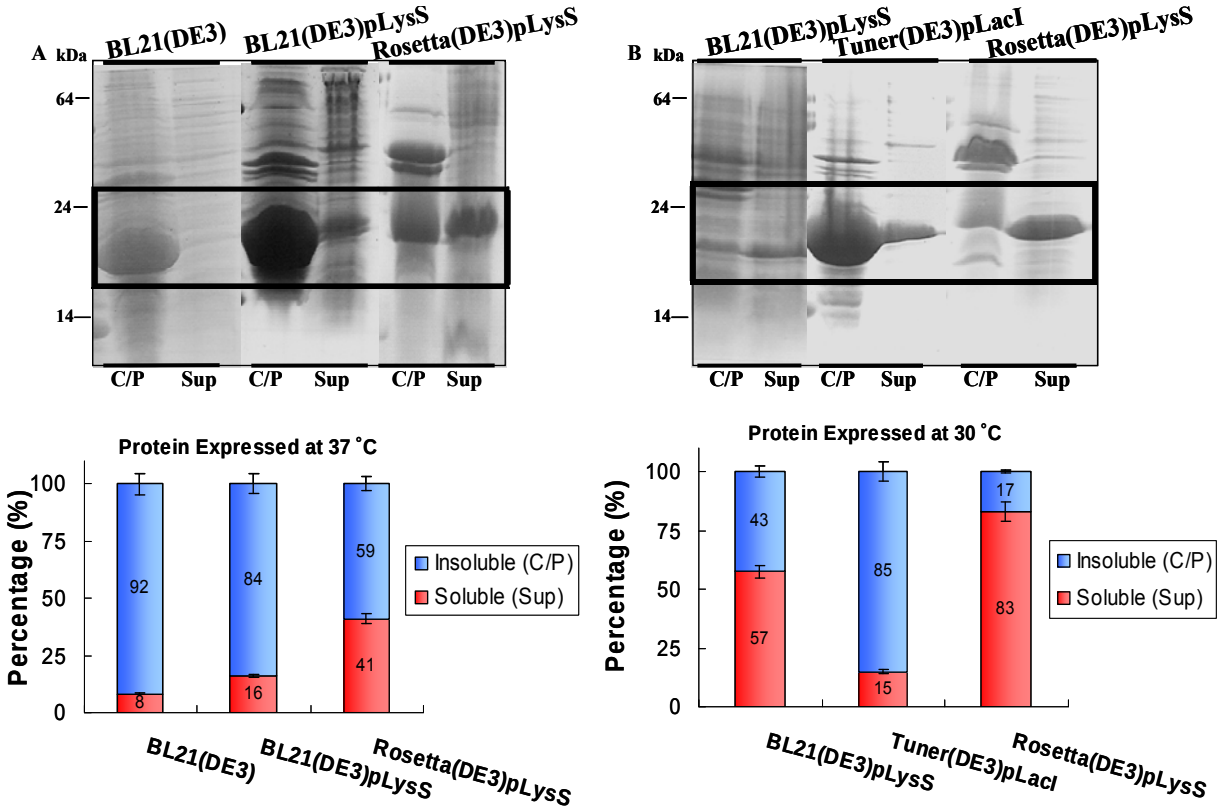


Figure 43. Expression temperature and cell strains affected subdomain 1 solubility. (A) Insoluble and soluble percentages when protein was expressed at 37 ^{o}C. (B) Insoluble and soluble percentages when protein was expressed at 30 ^{o}C overnight. Percentages were calculated using ImageJ software.

Protein concentration was calculated using the extinction coefficient of subdomain 1 (36,130 M^{-1} cm^{-1}) and its UV-vis absorbance at 280 nm.

5.3.3 Subdomain 2 expression and purification

The best expression and purification conditions used for subdomain 1 were used to express subdomain 2. This protein was therefore successfully expressed in *E. coli* rosetta(DE3)pLysS at a temperature of 30 ^{o}C overnight (Figure 44). Subdomain 2 molecular weight is ~20 kDa and a clear band can be observed close to the 24 kDa mark. Expressed protein

was further purified using 6x-his-tag affinity chromatography and the final yield was 7 mg/L and a concentration of 363 μM in ~2 mL (Figure 44).

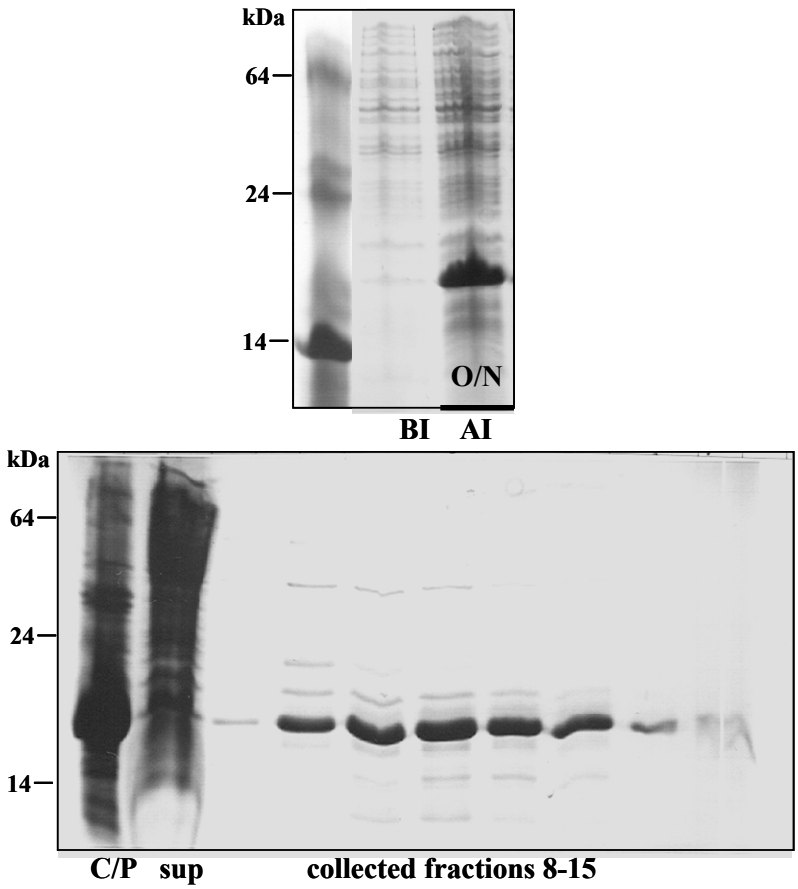


Figure 44. Subdomain 2 expression and purification using 6x-his-tag affinity chromatography. (LB media, rosetta(DE3)pLysS, 30 ^{o}C overnight expression).

5.3.4 Subdomain 1 and 2 expression in ^{15}N labeled minimal media

Additional efforts were made to express and purify subdomain 1 and 2 in ^{15}N labeled minimal media. These two engineered proteins were expressed in *E. coli* BL21(DE3)pLysS, rosetta(DE3)pLysS, and tuner(DE3)pLacI. The best cell strain that expressed these proteins with the least amount of insoluble aggregates in C/P was tuner(DE3)pLacI for both of these proteins (see red boxes in Figure 45A). The SDS gel clearly shows that the proteins were expressed and that even though a large amount of protein was present in the C/P a significant amount of protein

was in the Sup or soluble form. Additionally, these proteins were purified using 6x-his-tag affinity chromatography and the SDS purification gel for subdomain 1 can be observed in Figure 45B.

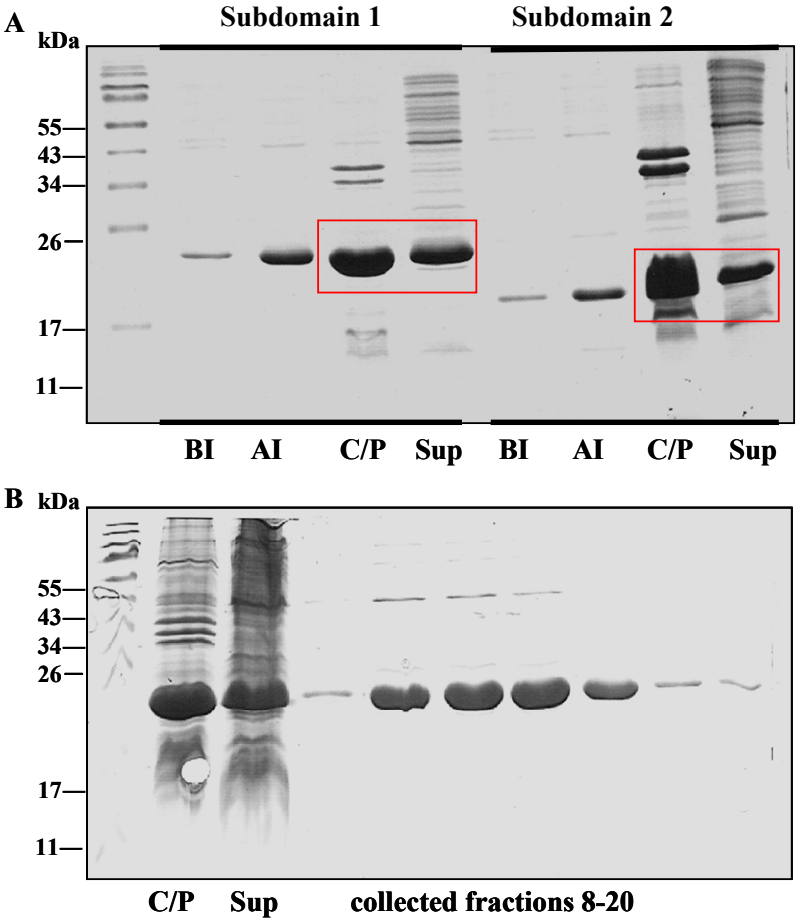


Figure 45. (A) Subdomain 1 and subdomain 2 expression using ^{15}N labeled minimal media (tuner(DE3)pLacI) and cell disruption using French press. (B) Subdomain 1 purification using 6x-his-tag affinity chromatography.

5.3.5 Subdomain 3 expression optimization

Four cell strains at different temperatures were tested: rosetta(DE3)pLysS at 30 ^{o}C overnight, BL21(DE3) at 37 ^{o}C, BL21(DE3)pLysS at 30 ^{o}C overnight, and tuner(DE3)pLacI at 37 ^{o}C in order to express this protein. The optimum expression conditions used for subdomain 1 and 2 did not express subdomain 3. One possible reason is that this protein has four more

cysteines and these residues could cause folding problems. This protein was only successfully expressed in tuner(DE3)pLacI at an expression temperature of 37 ^{o}C and 1 mL/L IPTG was used for induction (Figure 46). This protein was not expressed using any other of the tested cells strains (Figure 46).

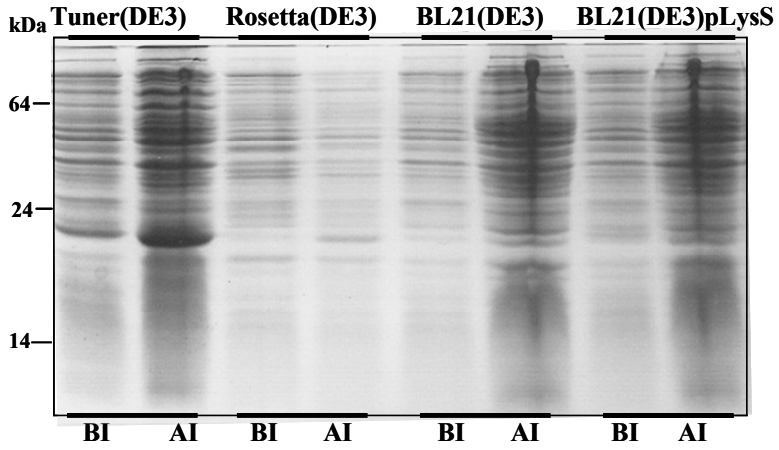


Figure 46. Subdomain 3 expression optimization using four cell strains and two temperatures. Protein was expressed at 37 ^{o}C in LB media.

5.3.6 Subdomain 3 purification optimization

Subdomain 3 was purified at first using His-tag affinity chromatography but the collected fractions contained a low amount of pure protein because most of the protein was in the pellet forming insoluble aggregates after the cell was disrupted using French press (Figure 47A). The pellet containing the protein was subsequently solubilized and completely unfolded using 8 M urea then refolded by gradually removing urea. Therefore, the 8 M urea solution was diluted to 4 M urea, dialyzed in 2 M urea, and then dialyzed in 10 mM Tris-HCl buffer, pH 7.4. The protein was then identified in the Sup after it was treated with urea (Figure 47A). The Sup containing the soluble and refolded protein was further purified using 6x-his-tag affinity chromatography (Figure 47B). Relevant fractions from the FPLC were collected (fractions 9 to 19), dialyzed in 10 mM Tris-HCL buffer pH 7.4, and concentrated down to ~2 mL. The calculated final

concentration using UV-visible spectroscopy was 300 μM in a 2 mL volume and the final yield was 7 mg/L.

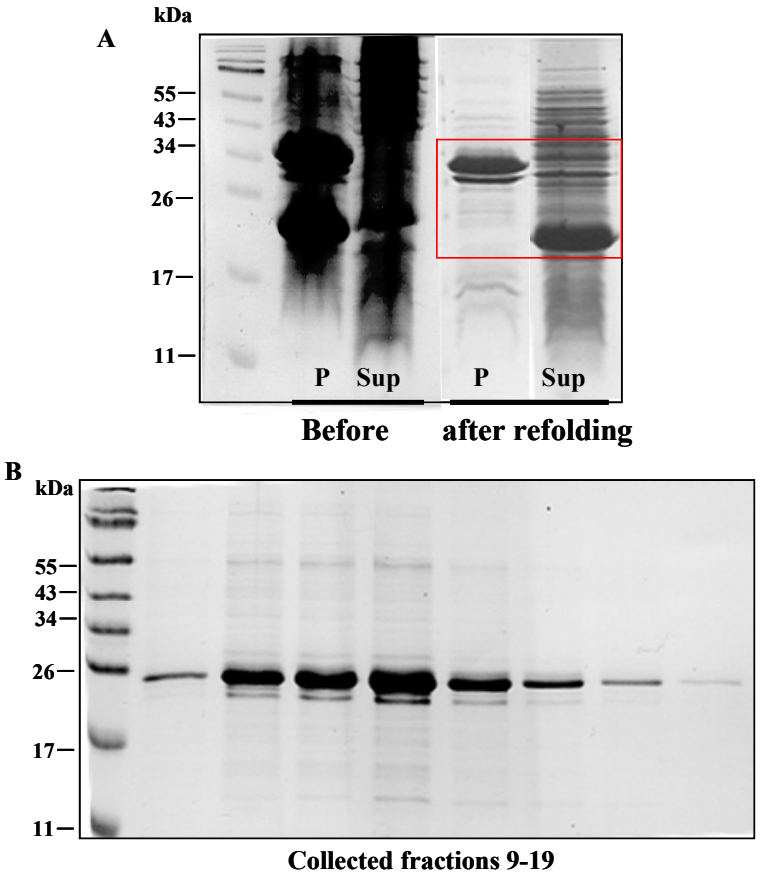


Figure 47. Subdomain 3 purification. SDS gels, (A) Subdomain 3 purification from insoluble aggregates. Pellet (P) and supernatant (sup) before and after refolding. Red box: protein is now in soluble supernatant form. (B) Subdomain 3 collected fractions after affinity chromatography.

5.3.7 Structural and conformational analyses for subdomains of the CaSR

Secondary and tertiary structures of the three subdomains and their conformational changes upon Ca^{2+} addition were investigated by Dr. Yun Huang using CD spectroscopy and intrinsic tryptophan (Trp) fluorescence (details can be observed in her thesis dissertation) [38]. The three subdomains exhibited typical helical secondary structure with negative signals at 208 nm and 222 nm (Figure 48A) as it was predicted by their model structures (Figure 48B). The

model structures of subdomain 1 and 2 clearly show a high globular helical content, however a lower number of helices was observed in subdomain 3 and that explains why subdomain 3 exhibited a slightly different CD spectrum compared to the other subdomains (Figure 48B).

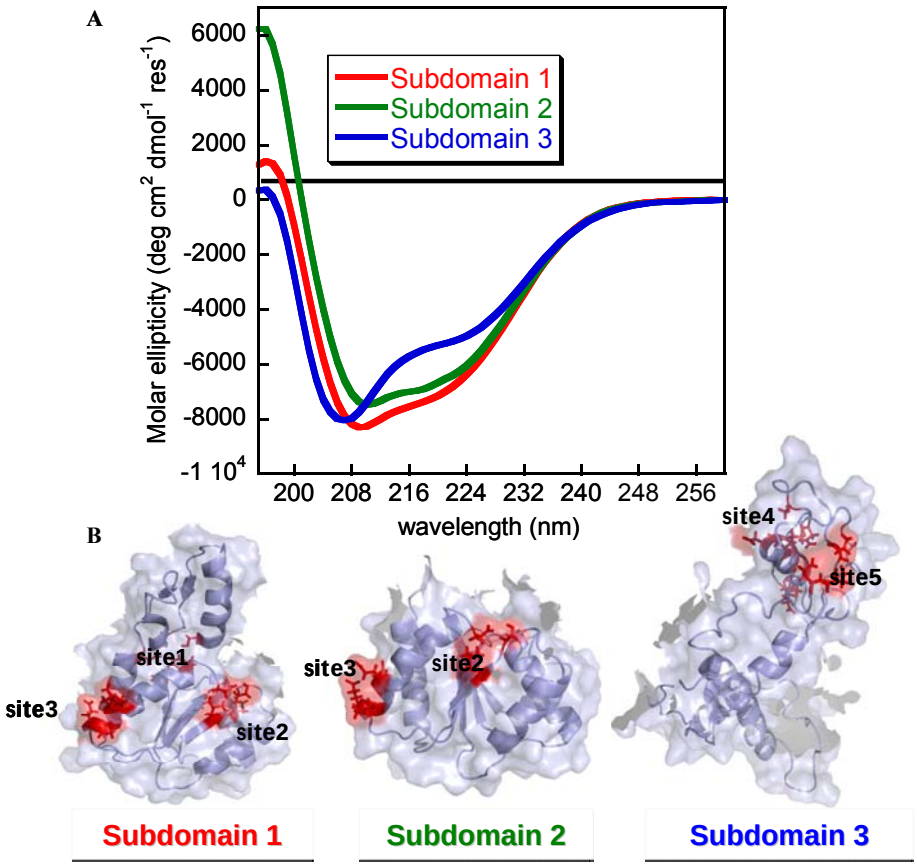


Figure 48. Structural analysis of subdomains using CD spectroscopy (A) Subdomains exhibited typical helical secondary structures. (B) Predicted model structures for each subdomain exhibiting high helical content.

Subdomains' conformational changes upon Ca^{2+} addition were investigated using CD spectroscopy and Trp fluorescence. CD spectra changes were observed for subdomain 1 and 3 after addition of saturating amounts of Ca^{2+} indicating that the secondary structure for these two subdomains is changing to a more helical form. Additional changes were observed for Trp fluorescence spectra indicating a change in the tertiary structure; this change was particularly

evident for subdomain 1 since an increase in the Trp fluorescence intensity was observed (Figure 49).

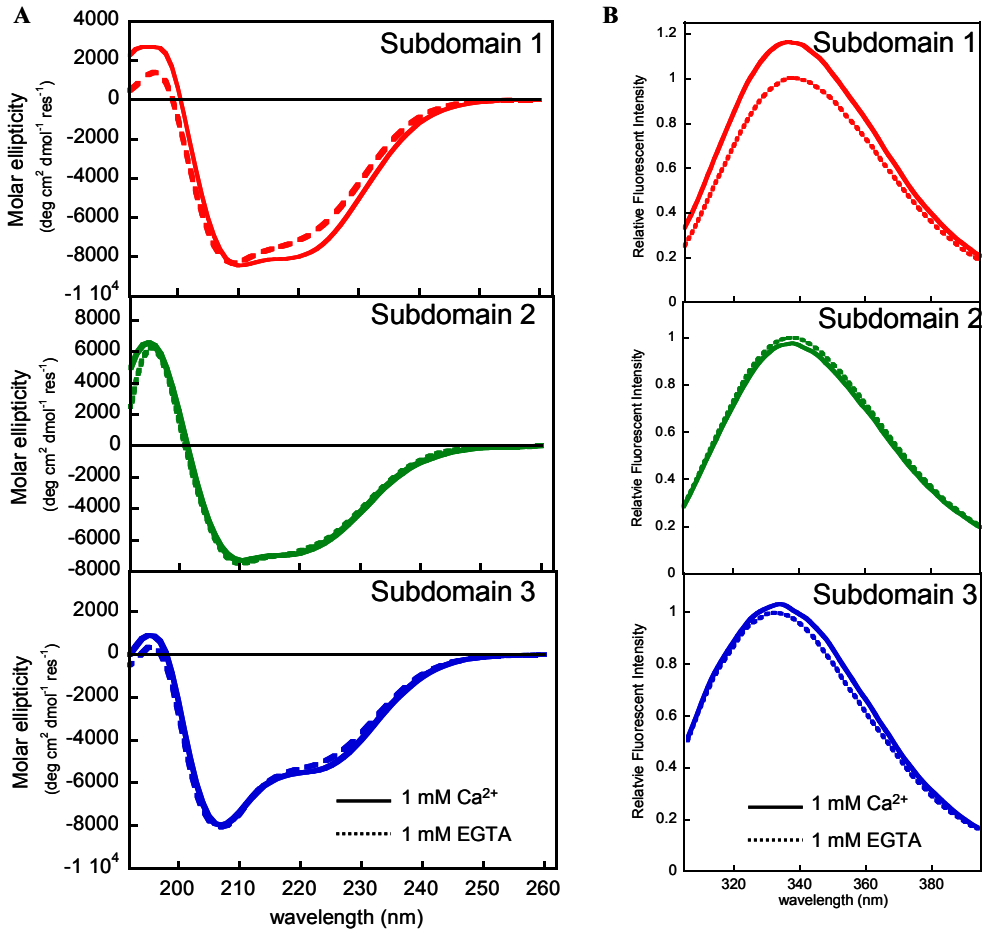


Figure 49. Conformational changes of the subdomains upon Ca^{2+} binding. (A) Far UV-CD spectra. (B) Trp fluorescence spectra [10].

5.3.8 Subdomain 1 metal-binding process

An additional metal-binding study that includes a Tb^{3+}-luminescence resonance energy transfer was performed by Dr. Yun Huang (details can be observed in her thesis dissertation) [38]. This metal-binding process confirmed that subdomain 1 exhibited a biphasic binding process as can be observed in Tb^{3+}-titration curve (Figure 50). This curve clearly shows that at least 2 metal-binding processes or components are occurring in subdomain 1. The first

component was attributed to site 1 and the second component corresponds to site 3 as it was concluded by Dr. Yun Huang [10].

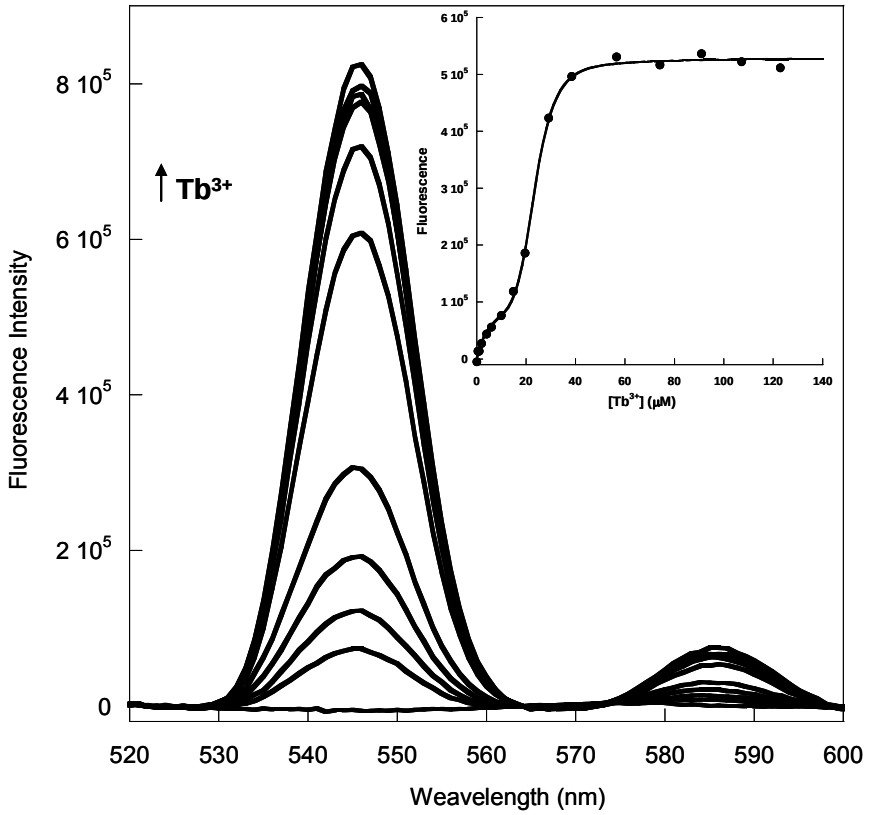


Figure 50. Tb^{3+} titration curve for subdomain 1. Subdomain 1 exhibited at least two metal-binding processes [38].

5.3.9 Extracellular domain (ECD) of the CaSR

The ECD of the CaSR is a huge domain with a molecular weight of ~73 kDa and this protein has not been expressed before in *E. coli*. This protein's expression and purification has been challenging due to its size, its extracellular environment, and the presence of ~20 cysteine residues in its original sequence and a rich cysteine region that connects the ECD with the 7TMD of the CaSR [51] (Figure 51). We have predicted five Ca^{2+}-binding sites located on the ECD and we have divided the ECD into three subdomains. The expression and purification of

the ECD is essential in order to further investigate how the three subdomains and their respective Ca^{2+}-binding sites interact with each other.

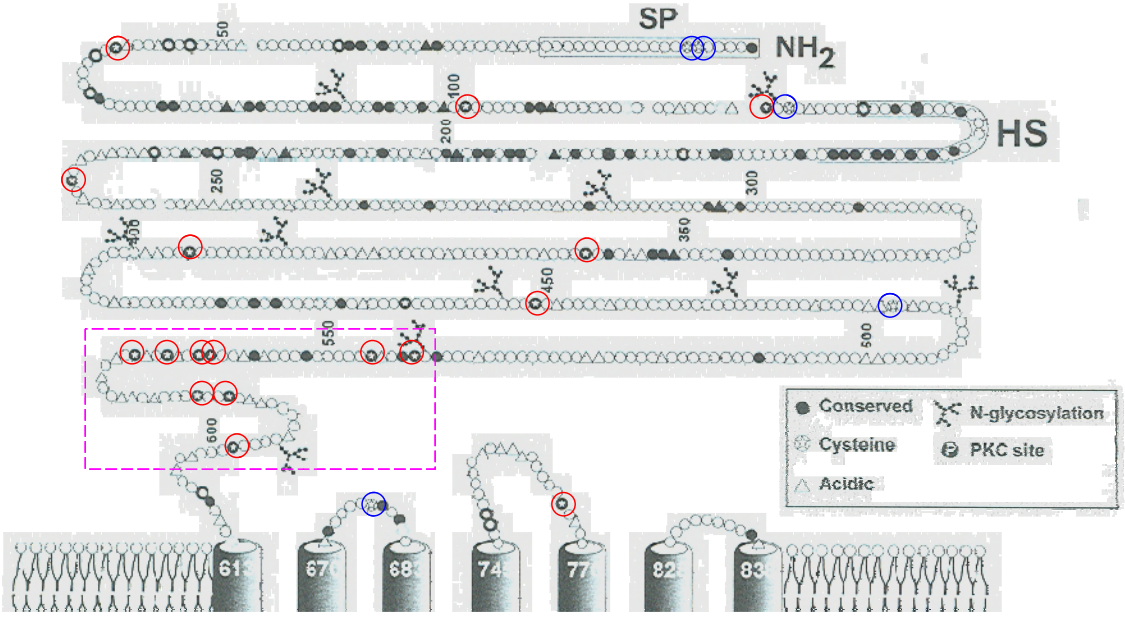


Figure 51. ECD of the CaSR contains 20 cysteine residues and a cysteine-rich region of 9 cysteine residues (pink box) that connects the ECD with a 7 TMD [51].

The following data show recent efforts made in order to express and purify this protein. It is important to mention that these are the best conditions that have been achieved to date, but further optimization procedures are being evaluated. The most recent purified protein using the best conditions up to this point showed low stability and degradation.

5.3.9.1 ECD expression optimization

ECD was cloned in a pRSETa vector containing a 6x-his-tag region in order to be purified using affinity chromatography in the same manner as its subdomains. Since the molecular weight of this protein is close to 73 kDa it was necessary to use 8% SDS gels in order to observe its purity and molecular weight band using gel electrophoresis.

This protein was first expressed in *E. coli* rosetta-gami(DE3)pLysS at 37 ^{o}C. Cells were growing at a very slow rate but they were still collected the same day after induction (3 h AI)

(Figure 52). Additional efforts were made to express this protein at 3 different temperatures 25 ^{o}C, 30 ^{o}C, and 37 ^{o}C overnight in order to determine if the temperature and expression time could affect the ECD expression. Temperature and expression time indeed affected this protein's expression because ECD was not expressed at low temperatures (25 ^{o}C and 30 ^{o}C) but it was expressed at 37 ^{o}C both during a single day and overnight expression (Figure 52). The SDS gel clearly shows that the protein was better expressed at 37 ^{o}C overnight. The same conditions were utilized to express the ECD in tuner(DE3)pLacI at 37 ^{o}C overnight but this cell strain expressed other proteins besides ECD. OrigamiB(DE3)pLysS at 30 ^{o}C overnight was also tested and actually a more pronounced expression band was observed in origamiB(DE3)pLysS when compared with the other two cell strains (Figure 53A). Additionally, origamiB(DE3)pLysS obtained the higher intensity change AI/BI followed by rosetta-gami(DE3)pLysS and tuner(DE3)pLacI (Figure 53B). Therefore, a higher expression level was observed with origami-B(DE3)pLysS and rosetta-gami(DE3)pLysS because these two cell strains contain mutations in the thioredoxin and glutathione reductase facilitating disulfide bond formation in the cytoplasm.

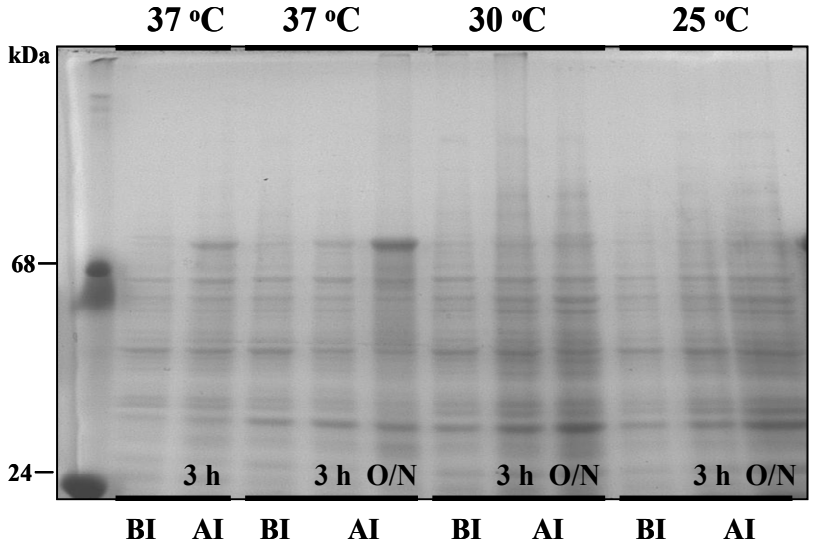


Figure 52. ECD of the CaSR expression at three temperatures (Vector: pRSETa, LB media).

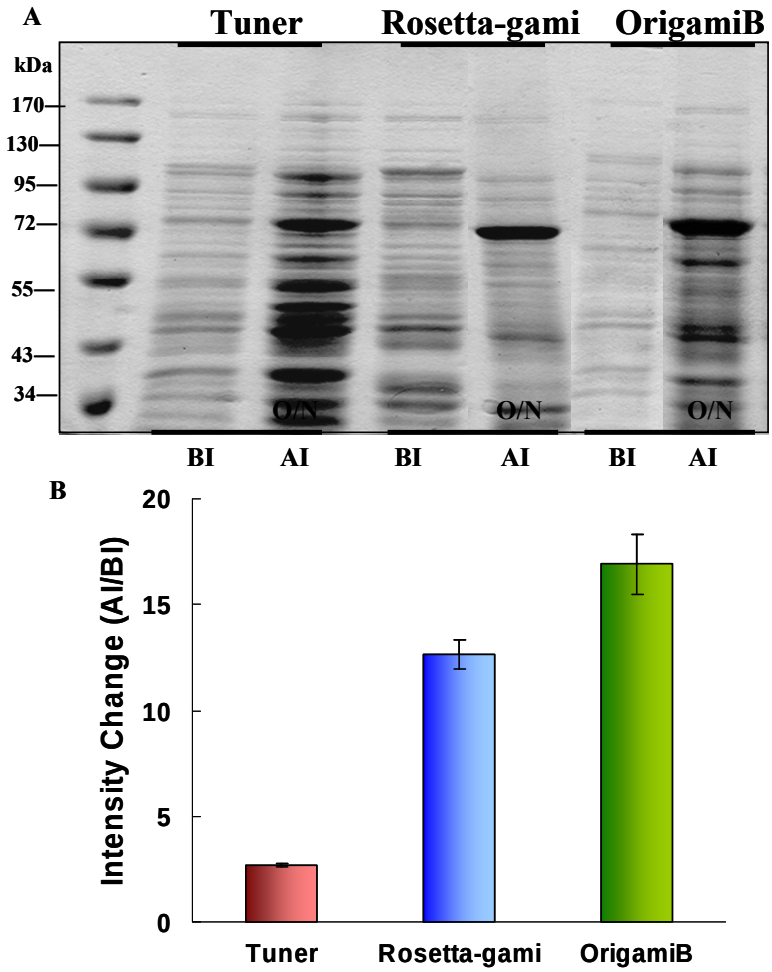


Figure 53. ECD of the CaSR expression. (A) Expression using different cell strains: tuner(DE3)pLacI, rosetta-gami(DE3)pLysS, and origamiB(DE3)pLysS at 37 ^{o}C and 30 ^{o}C overnight (Vector: pRSETa, LB media). (B) Intensity change AI/BI for ECD expression using different cell strains.

5.3.9.2 ECD purification optimization

Expressed protein was further purified using 6x-his-tag affinity chromatography. A few changes were made to the protocol used for the subdomains purification because when the same protocol was used for ECD purification a significant amount of protein degraded. Two major changes were therefore made to the usual 6x-his-tag affinity chromatography purification. The first change was dissolving the expression pellet in lysis buffer with two tablets of protease inhibitor. This step was done in order to prevent degradation of the protein by proteases. The

second change was that the initial imidazole concentration of the solution before it was injected into the FPLC was increased from 20 mM to 50 mM. This step was done in order to increase the binding competition between imidazole and protein so that the purity could be increased.

Protein was purified using affinity chromatography and the best results were obtained when the protein was expressed in rosetta-gami(DE3)pLysS at 37 ^{o}C overnight (Figure 54A). However, degradation was observed and the presence of insoluble aggregates in the C/P was also evident in all three cases (Figure 54A-C). Collected fractions from the His-tag column were dialyzed in 10 mM Tris-HCl buffer, pH 7.4 and consequently concentrated down to ~2 mL. The final concentration was calculated using UV-vis spectroscopy, protein's absorbance at 280 nm, and its extinction coefficient which is 110,300 M^{-1} cm^{-1}. The final yields were 12.7 mg/L, 3.9 mg/L, and 6.7 mg/L for rosetta-gami(DE3)pLysS, tuner(DE3)pLacI, and origamiB(DE3)pLysS respectively (Table 8). Optimum purification conditions for this protein are still under investigation but the best conditions to date were observed with rosetta-gami and origamiB since they facilitate disulfide-bond formation in the cytoplasm.

Table 8. ECD final yields when protein was expressed in different cell strains

Cell strain	Yield (mg/L)
Rosetta(DE3)pLysS	12.7
Tuner(DE3)pLacI	3.9
Origami-B(DE3)pLysS	6.7

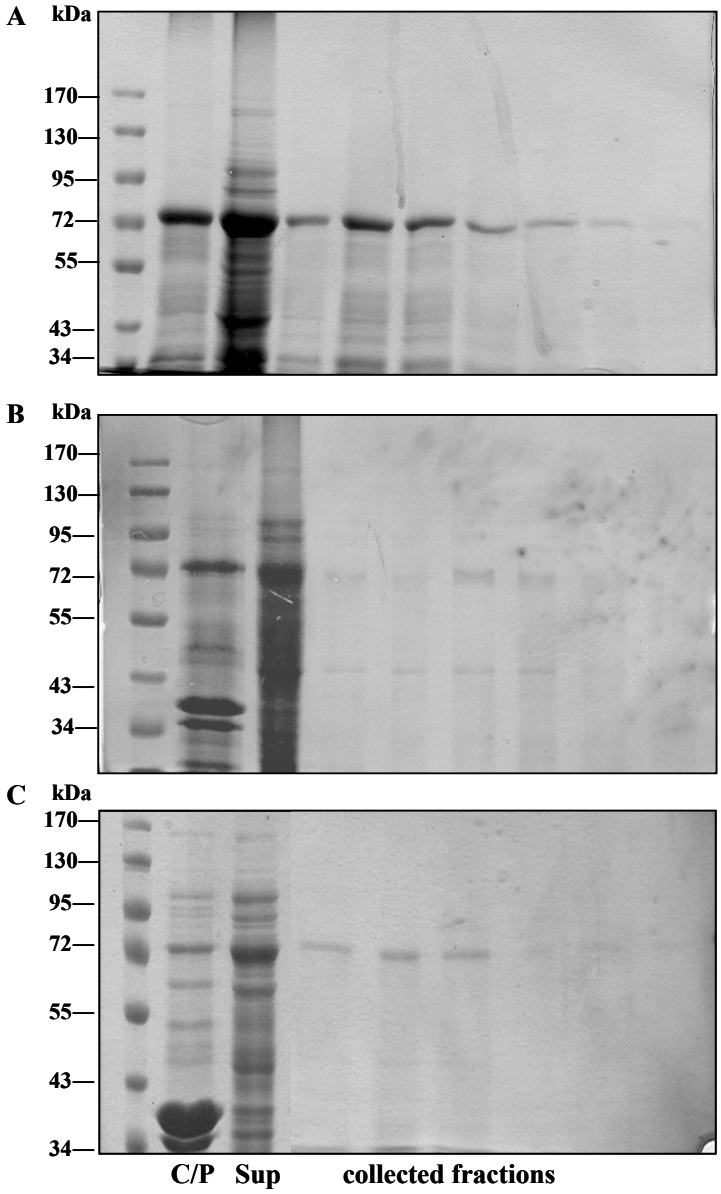


Figure 54. ECD of the CaSR purification using 6x-his-tag affinity chromatography. (A) Protein was expressed in *E. coli* rosetta-gami(DE3)pLysS, (B) tuner(DE3)pLacI, and (C) origamiB(DE3)pLysS.

5.4 Summary

Five predicted Ca^{2+}-binding sites of the CaSR (sites 3 and 5) were grafted into CD2 and two out of those five were expressed in *E. coli* BL21(DE3) at 37 ^{o}C. Since these sites were

previously cloned in a pGEX-2T vector containing a GST-tag, further purification was performed using affinity and ion-exchange chromatography. The predicted Ca^{2+}-binding sites were subjected to a series of experiments that confirmed that they are indeed Ca^{2+}-binding sites of the CaSR.

The ECD of the CaSR was cleaved into three subdomains that contained the predicted Ca^{2+}-binding sites. Subdomain 1 was first expressed in *E. coli* BL21(DE3) using three different vectors: pGEX-2T, pET32a, and pRSETa. This protein formed insoluble aggregates in all three cases and for that reason additional efforts were made to optimize subdomain 1 expression and purification utilizing pRSETa vector. This protein was therefore expressed in four different cell strains and two temperatures. Expressed proteins were then purified using 6x-his-tag affinity chromatography in order to determine whether the protein formed insoluble aggregates after cell disruption. The best results were obtained when subdomain 1 was expressed in rosetta(DE3)pLyS at 30 ^{o}C overnight. Therefore, lower temperatures slowed down the protein expression allowing proper protein folding and increasing its solubility.

The same expression and purification conditions were used to successfully express and purify subdomain 2. Therefore, subdomain 2 was also expressed in rosetta(DE3)pLysS at 30 ^{o}C overnight and purified using affinity chromatography.

On the other hand, subdomain 3 was expressed using a different cell strain tuner(DE3)pLacI since it was not expressed in rosetta(DE3)pLysS, BL21(DE3), or BL21(DE3)pLysS. This protein was purified from insoluble aggregates by refolding the protein followed by affinity chromatography.

All three subdomains were successfully purified with concentrations between 300 to 800 μM in a volume of ~ 2 mL and yields between 7 to 36 mg/L (Table 7). Additional efforts have

been recently made to express and purify the ECD of the CaSR. The best results show that this protein was successfully expressed in *E. coli* rosetta-gami(DE3)pLysS at 37 ^{o}C overnight and origamiB(DE3)pLysS at 30 ^{o}C overnight. This protein was purified using affinity 6x-his-tag chromatography. Additional efforts have been made in order to optimize the expression and purification of the ECD of the CaSR since this protein has shown low stability and concentrations due to its size and presence of a high number of cysteine residues.

6. THE DESIGN APPROACH: PROTEIN PURIFICATION FROM INCLUSION BODIES

6.1 Introduction

The purpose of the design approach is to use a computer algorithm to generate potential Ca^{2+}-binding sites and then incorporate those sites into a chosen host protein (CD2) using site-directed mutagenesis [15]. The advantage of this approach is that it will allow us to understand the factors that contribute to Ca^{2+} binding by investigating the relationship between net charges in the coordination sphere and the Ca^{2+}-binding affinities [13, 15]. It is important to mention that the electrostatic nature of Ca^{2+} binding is highly affected by charged residues, binding geometry, and protein environment [59].

Two proteins were designed in order to study the effect of local charges on metal-binding affinity, CD2.6D15 and CD2.7E15. The first one (CD2.6D15) has three negatively charged ligands and contains two mutations (N15D and N17D) and two natural residues as ligands (N60 and D62) as it can be observed in the NMR structure of this engineered protein (Figure 55A) [59]. The second one (CD27E15) has five negatively-charged ligands, three mutations (N15E, L58D, and K64D) and two natural residues as ligands (E56 and D62) as observed in the model structure shown in Figure 55B [59]. Previous studies have concluded that the incorporation of these Ca^{2+} ligands into the host protein did not reveal global conformational changes in the designed proteins [59]. It was also concluded that CD2.7E15 had a 12 to 14-fold increase in the binding affinity for Ca^{2+}, Tb^{3+}, and La^{3+} compared to CD2.6D15 [59]. However, the localized charges in CD2.7E15 decreased the host protein's thermal stability (T_m) from 61 ^oC to 41 ^oC

[59]. The designed approach can be therefore used to further study Ca^{2+} signaling and binding affinities properties.

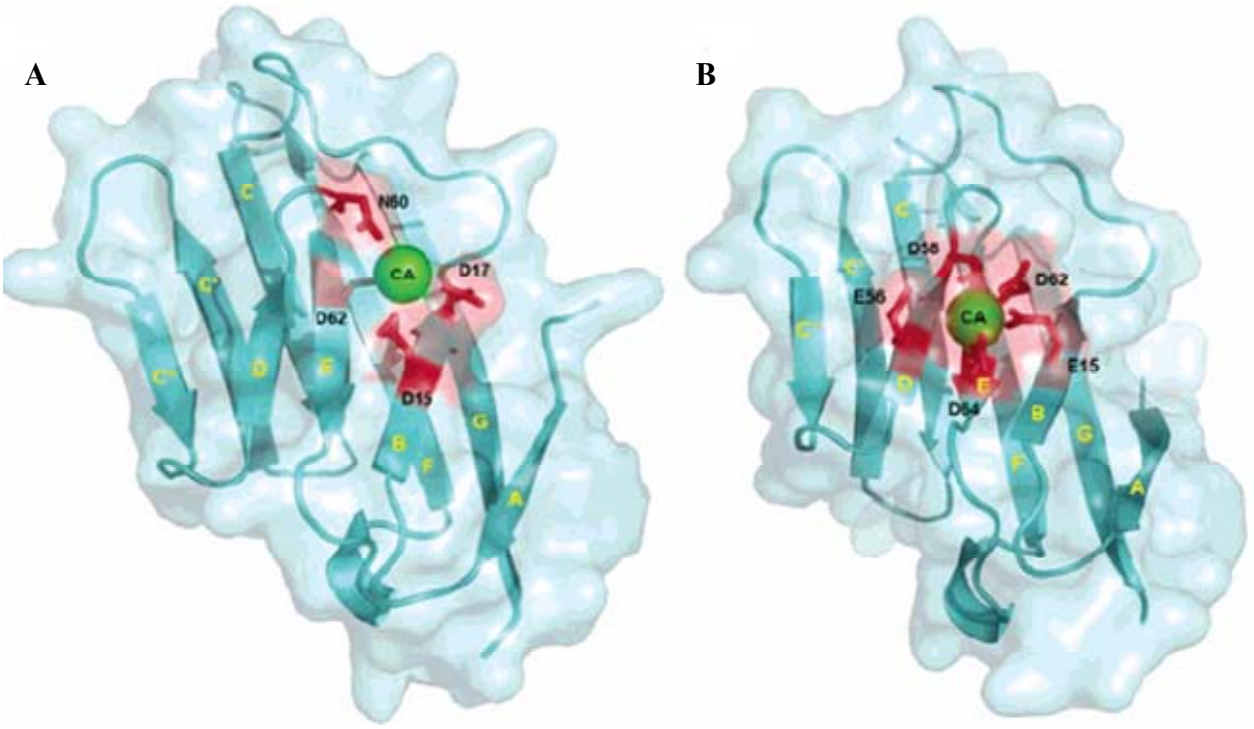


Figure 55. (A) NMR structure of CD2.6D15. (B) Model structure of CD2.7E15. Pictures generated using PyMol (DeLano Scientific LLC) [59].

CD2.7E15 was initially expressed in *E. coli* as a fusion GST-tag protein but its expression and purification levels were very low. Therefore, in order to improve this protein's expression and purification this protein has been expressed and purified as a tag-less protein. Also, a series of newly-designed proteins have been created using the design approach for the purpose of creating contrast agents in cancer cell targeting. One of these designed proteins is CD2.7E15.52I containing a bombesin insertion (Glu-Gln-Arg-Leu-Gly-Asn-Gln-Trp-Ala-Val-Gly-His-Leu-Met-NH_2). Bombesin possesses a high affinity for the gastrin releasing peptide which is over expressed on different types of cancer cells [60]. Bombesin addition can therefore enrich the environment in which the designed protein contrast agent moves and in that way increase the

intensity and targeting in tumor detection. This protein did not contain a tag and for that reason it was necessary to develop new and optimized conditions to express and purify tag-less protein.

6.2 Designed proteins as MRI contrast agents

Magnetic resonance imaging (MRI) is a noninvasive technique that provides three-dimensional images of anatomic structures and that are often used as contrast agents between healthy and pathological tissues [61, 62]. Gadolinium (Gd^{3+}) chelators are commonly-used contrast agents and recent studies have determined that CD2.6D15 and CD2.7E15 have a high affinity for Gd^{3+} and they can be therefore used as Gd^{3+} chelators in magnetic resonance imaging (MRI) [63]. These protein-based MRI contrast agents have shown that metal-binding selectivity and affinity can be improved by utilizing the design approach [63].

6.3 Results and discussion

6.3.1 CD2.7E15.52I expression optimization

Engineered proteins that were created using a designed approach (CD2.7E15 and variants) has been previously expressed and purified using affinity chromatography but typically exhibited a common problem with low yields. Protein purification from inclusion bodies was an excellent option to further optimize the purification and in that way improve the proteins' yield at a low cost. This technique advantageously reduces protein degradation and improves proper refolding which typically enhances protein's activity.

CD2.7E15.52insertion (CD2.7E15.52I) was expressed in *E. coli* as a fusion engineered protein in a pET20b vector. Initial expression of this protein was attempted using four cell strains BL21, BL21(DE3), BL21(DE3)pLysS, and tuner(pLacI), but only BL21(DE3)pLysS and tuner(DE3)pLacI produced expression, with the latter resulting in the best expression. This protein was subsequently expressed at two temperatures 37 ^{o}C and 30 ^{o}C overnight in order to

further optimize CD2.7E15.52I expression. The expression growth curve shows that the cell culture reached an OD of ~1.5 when it was expressed at 37 ^{o}C (8 h expression) and it reached an OD of ~ 2.0 when it was expressed at 30 ^{o}C (16 h expression) (Figure 56A). The SDS gels (Figure 56B) show that a thicker expression band is observed when the protein is expressed at 30 ^{o}C overnight compared to that observed when the protein was expressed at 37 ^{o}C. Therefore, lower temperatures, such as 30 ^{o}C increased the expression level. Tuner(DE3)pLacI is IPTG dependent and for that reason the IPTG concentration used for induction was also optimized. The best results were obtained when protein expression was induced with 500 µL IPTG per liter.

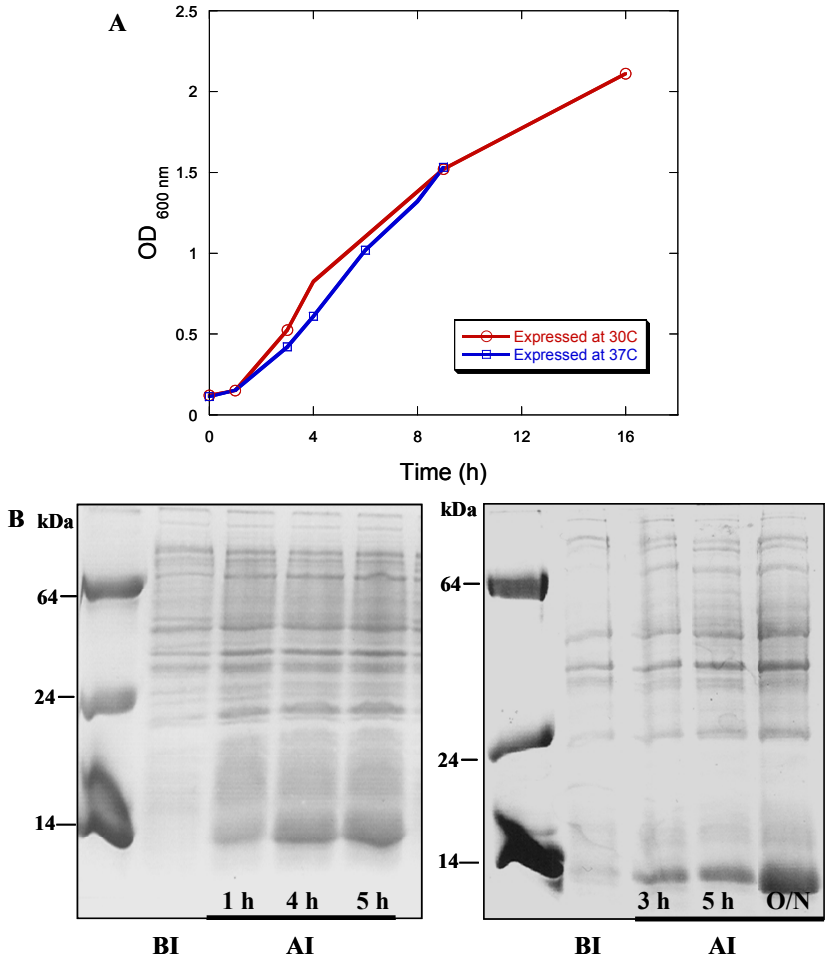


Figure 56. (A) CD2.7E15.52.Insertion growth curve at two different temperatures 37 ^{o}C (blue) 30 ^{o}C (red). (Vector: pET20b, LB media, tuner(DE3)pLacI). (B) CD2.7E15.52.Insertion expression in *E. coli* at 37 ^{o}C (left) and 30 ^{o}C overnight (right).

6.3.2 CD2.7E15.52I purification

This protein was purified using an 8 M urea refolding method followed by ion exchange chromatography. The purification SDS gels clearly show the refolding process step by step when the protein was expressed at 30 ^{o}C overnight (left) and at 37 ^{o}C (right) (Figure 57). The first two lines in both gels show the pellet (P) and the supernatant (Sup) after the cells were broken using French Press. It can be observed from the gel (Figure 57) that most of the protein is in the pellet form. The pellet was washed using a detergent (Titron X-100) in order to remove impurities present in the pellet, such as cell debris or remains. The following three lanes in each gels show the pellet and supernatant after the pellet was dissolved in 8 M urea and followed dilution in 4 M urea. It is evident that the protein was completely solubilized from the pellet to the supernatant form (Figure 57). This SDS gels also show the supernatant after it was dialyzed in 10 mM Tris-HCl buffer, pH 7.4. The purification produced better results when the protein was expressed at 30 ^{o}C overnight.

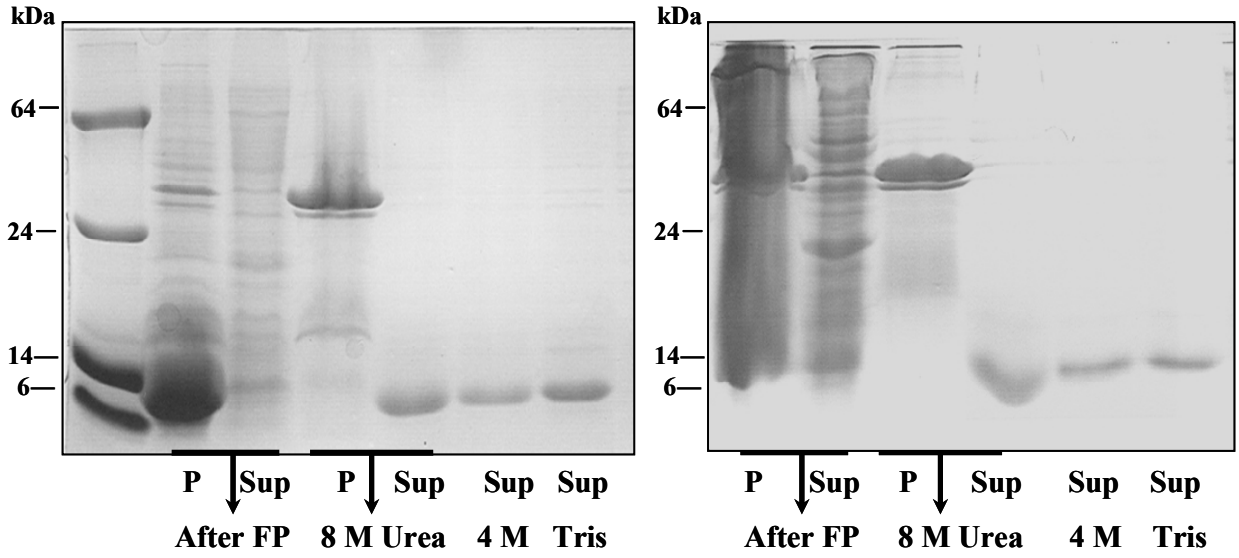


Figure 57. CD2.7E15.52.Insertion purification using urea refolding method. Protein was expressed at 30 ^{o}C overnight (left). Protein was expressed at 37 ^{o}C (right).

Dialyzed sample was further purified using ion exchange chromatography. Protein concentration was calculated using the extinction coefficient of CD2.WT (11,700 M^{-1} cm^{-1}) and its UV-vis absorbance at 280 nm. The final yield was ~20 mg/L.

6.3.3 CD2.7E15 expression

The same approach used to express CD2.7E15.52I was also implemented to express CD2.7E15. This protein was successfully expressed as a tag-less protein in *E. coli* tuner(DE3)pLacI at 30 ^{o}C overnight (Figure 58). The expression level was not as high as that observed for CD2.7E15.52I but it was enough to obtain pure protein. Another difference was that this protein expressed unknown proteins besides our targeted protein as a strong dark background was observed in the expression SDS gel.

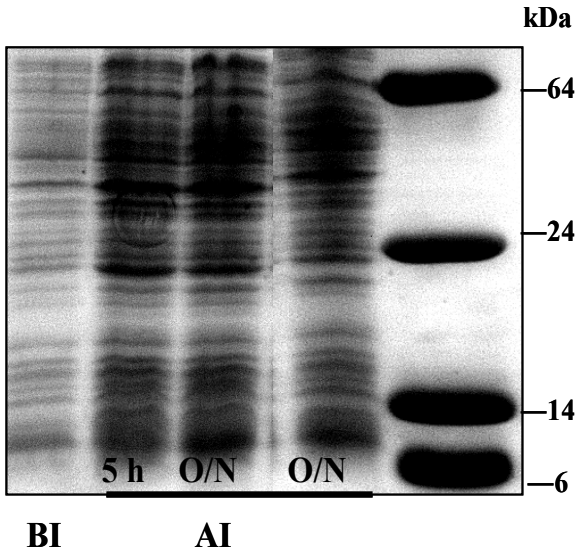


Figure 58. CD2.7E15 expression in *E. coli* at 30 ^{o}C overnight (vector: pET20b, LB media, tuner(DE3)pLacI).

6.3.4 CD2.7E15 purification

The same approach used to purify CD2.7E15.52I was also implemented to purify CD2.7E15. The purification was therefore performed using an 8 M urea refolding method

(Figure 59).The first two lines show the pellet (P) and the supernatant (Sup) after the cells were broken using French Press. It can be observed (Figure 59) that most of the protein is the pellet form. The pellet was washed using a detergent (Titron X-100) in order to remove impurities. The following three lanes in each gels show the pellet and supernatant after the pellet was dissolved in 8 M urea and then diluted to 4 M urea. It is clear that the protein was completely solubilized from the pellet to the supernatant form. This SDS gel also shows the supernatant after it was dialyzed in 10 mM Tris-HCl buffer, pH 7.4. The purity of this protein was not as high as that observed for CD2.7E15.52I since a series of bands above that of our targeted protein were observed in the SDS expression and purification gel. Protein was further purified using ion exchange chromatography after it was dialyzed in 10 mM Tris-HCl buffer. The additional bands observed in Figure 59 were completely removed after ion exchange FPLC. Figure 60 illustrates the FPLC profile and SDS gel confirming the high purity of the purified protein.

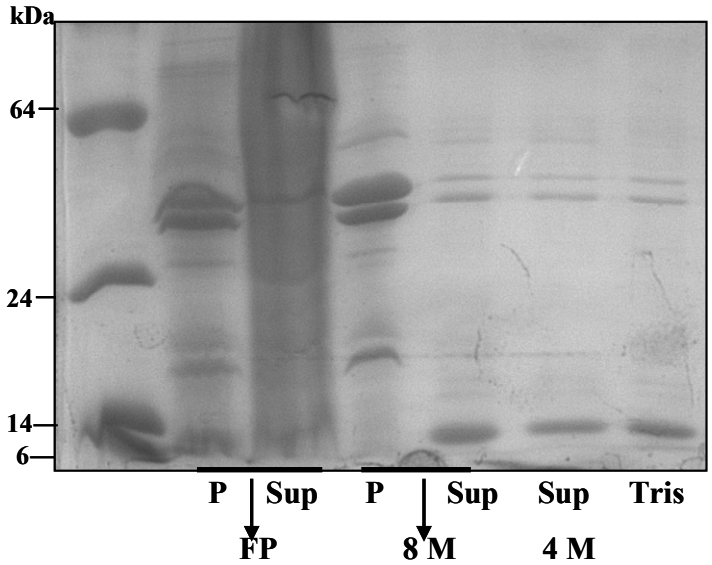


Figure 59. CD2.7E15 purification using urea refolding method.

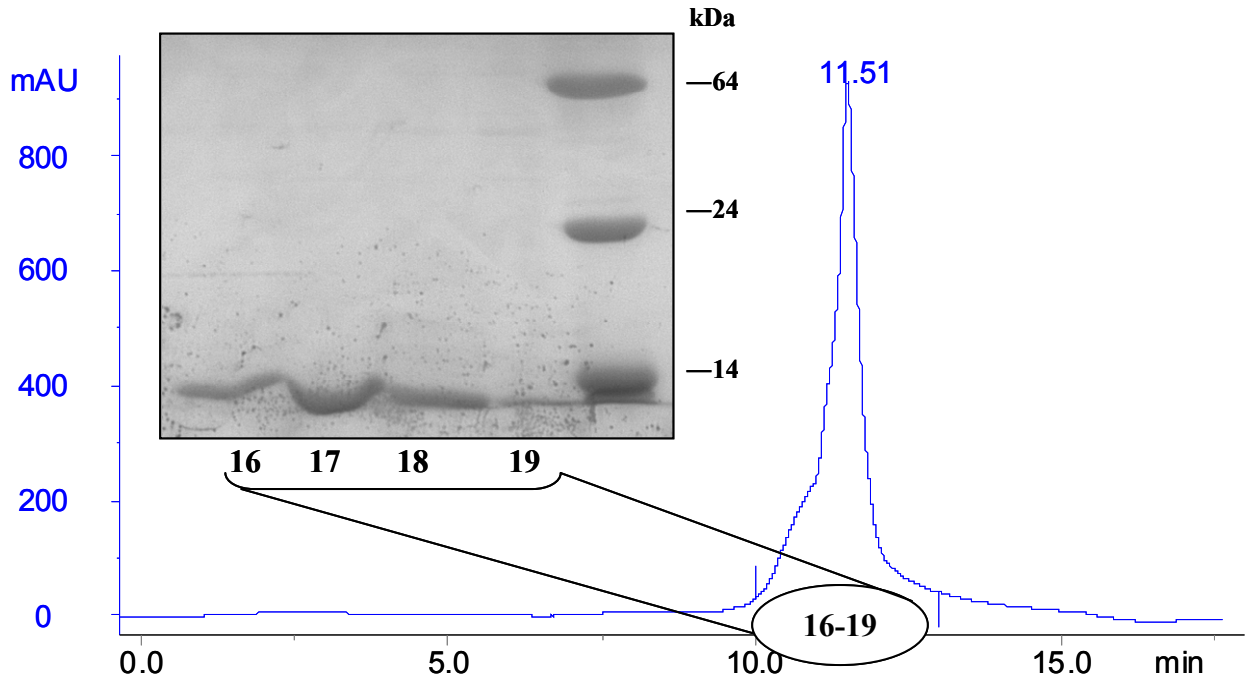


Figure 60. CD2.7E15 ion exchange chromatography and SDS gel showing the collected relative fractions from the FPLC.

Protein concentration was calculated using the extinction coefficient of CD2.WT (11,700 M^{-1} cm^{-1}) and its UV-vis absorbance at 280 nm. The final yield was ~10mg/L.

6.4 Summary

The design approach allows the investigation of local charges and the factors that might contribute to Ca^{2+} binding. Two designed engineered proteins CD2.7E15.52I and CD2.7E15 were successfully expressed in *E. coli* tuner(DE3)pLacI at 30 ^{o}C overnight. Tuner(DE3)pLacI have shown to form a high number of inclusion bodies during protein expression. This cell strain contains a mutation in *lacY* (lac permease) gene allowing over protein expression in the presence of a high IPTG concentration. Over protein expression therefore facilitates inclusion body formation and sometimes this is the only way to obtain protein. Protein expressed in inclusion bodies can be purified using refolding methods. These engineered proteins were actually purified from inclusion bodies utilizing a refolding strategy followed by ion exchange chromatography.

The final yields were ~20 mg/L for both engineered proteins but further expression and purification optimization procedures are still under investigation. These designed proteins have been used as protein-based MRI agents and they have been also used to perform metal-binding and affinity analyses.

7. FINAL CONCLUSIONS AND SIGNIFICANCE OF THIS THESIS

Ca^{2+} is an essential element in living organisms since it regulates a great number of biological and cellular signaling pathways. Ca^{2+} concentrations in extracellular and intracellular environments are balanced and stabilized by a select group of Ca^{2+}-binding proteins. This vital group of proteins usually has two or more Ca^{2+}-binding sites that interact with each other upon Ca^{2+} binding. They are able to undergo conformational changes upon Ca^{2+}-binding and these conformational changes are then translated into metabolic pathways. Additionally, they are able to sense intracellular and/or extracellular Ca^{2+} concentration changes or imbalances and take action in order to maintain a proper equilibrium in both environments.

The investigation of Ca^{2+}-binding sites is essential in order to understand the biochemical properties of Ca^{2+}-binding proteins and further, to comprehend and determine the mechanisms behind this reciprocal Ca^{2+}-protein interaction. Three approaches have been developed in Dr. Yang's laboratory in order to study Ca^{2+}-binding sites located in Ca^{2+}-binding proteins. The basic principle of these approaches is the prediction of Ca^{2+}-binding sites using computer algorithms, which provides a rational basis for grafting and/or designing them into host proteins. In the grafting approach, individual Ca^{2+}-binding sites were inserted into host proteins (e.g. CD2 or EGFP) in order obtain metal-binding affinities and site-specific properties of Ca^{2+}-binding proteins. In the design approach, site-directed mutagenesis was used to design Ca^{2+}-binding sites into host proteins. This approach has permitted the investigation of factors that contribute to local Ca^{2+}-binding and the relationship between net charges located in the coordination sphere of these sites. In the subdomain approach, Ca^{2+}-binding proteins have been dissected into different subdomains containing 2-3 Ca^{2+}-binding sites. This approach allowed the study of the interaction

between Ca^{2+}-binding sites, their conformational changes, and cooperative interaction upon Ca^{2+} binding.

A variety of techniques (e.g. CD, MS, NMR, metal-binding titrations, and X-ray crystallography) have been applied in order to test the predicted Ca^{2+}-binding sites and to understand important properties and characteristics of Ca^{2+}-binding proteins. The techniques used in these studies require high protein concentration and purity, and that is why expression and purification of engineered proteins play essential roles to further conduct those experiments.

The chosen expression host for engineered protein production was *E. coli*. This is the most widely used bacterial host for protein expression, but sometimes expression and purification of engineered proteins can be very challenging because the insertion of foreign sequences into a host protein might result in stability and conformational changes that affect protein yield and purity. Other challenges include low expression levels, formation of inclusion bodies due to improper folding, and proteolytic degradation.

In this thesis, the significance of expression and purification of engineered proteins as initial steps to further investigation of the properties of Ca^{2+}-binding proteins is discussed with respect to optimization of expression and purification to obtain high protein yield and purity. Logically, further experimentation is limited by the availability of proteins of sufficient quality and quantity.

In previous chapters, I have shown that expression and purification of engineered Ca^{2+}-binding proteins are highly influenced by a great number of factors, such as expression temperature, cell strains, vectors, and inducer concentrations. Additionally, there are an extended number of purification techniques and fusion technologies that can be employed to purify proteins based on the protein's physiochemical properties, such as their isoelectric points,

molecular weights, polarity, and/or affinity to ligands. Thus, purification of engineered proteins was performed using different techniques depending upon their characteristics.

Proteins engineered using a grafting approach in which an individual Ca^{2+}-binding motif was inserted into CD2.D1 (CD2-WT, CD2.STIM1, CD2.STIM1.Mut, and CD2.RubCa) were expressed in *E. coli* BL21(DE3) at 37 ^{o}C and further purified by affinity chromatography since they contained a GST-tag, followed by ion-exchange chromatography. Additionally, another set of engineered proteins that had EGFP as the host protein (WT-EGFP and EGFP-172-III-172) were also expressed in *E. coli* BL21(DE3) but the expression temperature was reduced to 30 ^{o}C in order to increase chromophore formation. These proteins were purified using affinity chromatography since they contained a 6x-his-tag. The yields for these proteins were between 20-70 mg/L (**Chapter 3**).

CaM is an intracellular Ca^{2+}-binding protein that was expressed in *E. coli* BL21(DE3) at 37 ^{o}C in LB media and at 30 ^{o}C overnight in ^{15}N labeled minimal media. A lower temperature was used for the minimal media because it required a longer incubation time since it contained fewer nutrients compared to LB media. CaM was purified using a phenyl-sepharose column and hydrophobic interaction chromatography, since this protein undergoes a conformational change upon Ca^{2+} binding that exposes hydrophobic patches to the surface. The final yields were 74.8 mg/L and 57.8 mg/L when the proteins were expressed in LB and ^{15}N labeled minimal media respectively. CaM ^{15}N labeled protein was further utilized to perform NMR studies (**Chapter 4**).

The CaSR is another important Ca^{2+}-binding protein that has been studied. Five Ca^{2+}-binding sites in the ECD of this receptor were predicted and two of these sites (3 and 5) were grafted into CD2.D1. These two sites were also expressed in *E. coli* BL21(DE3) at 37 ^{o}C and purified using GST-tag affinity and ion exchange chromatography. An additional set of mutants

were also expressed and purified using the same approaches. The final yields for these proteins were ~7 mg/L. Metal-binding studies performed by Dr. Huang have determined the metal-binding affinity of these two Ca^{2+}-binding sites (**Chapter 5**).

A subdomain approach was also utilized to study the CaSR in which the ECD of this receptor was dissected into three subdomains containing the predicted Ca^{2+}-binding sites. These proteins presented low expression levels, were highly insoluble, and/or were expressed as inclusion bodies. We first attempted to optimize the expression and purification of subdomain 1 by utilizing different vectors, *E. coli* cell strains, and temperatures. Subdomain 1 expression levels and solubility were highly enhanced when the protein was expressed at 30 ^{o}C and when rosetta(DE3)pLysS was used as the cell strain. These conditions were utilized to express and purify subdomain 2. Subdomain 3 was expressed in *E. coli* tuner(DE3)pLysS as inclusion bodies and they were purified using a refolding technique followed by affinity chromatography. The final yields were 7-36 mg/L. Additional efforts have been made to express and purify the ECD of the CaSR. This 73 kDa protein contains ~20 cysteine residues that have highly influenced protein expression and purification. The best cell strains that expressed this protein were rosetta-gami(DE3)pLysS and origamiB(DE3)pLysS since these strains contain mutation in thioredoxin and glutathione reductase genes facilitating disulfide bond formation as well as proper protein folding. The optimum expression and purification conditions are still under investigation since the protein presented high levels of degradation and precipitation. Further experiments performed by Dr. Yun Huang (e.g. CD and trp fluorescence) have shown that a conformational change is observed when saturating Ca^{2+} concentrations were added to the proteins. Additionally, metal-binding studies have shown that subdomain 1 exhibited a biphasic binding mechanism (**Chapter 5**).

Another set of engineered proteins (CD2.7E15 and CD2.7E15.52I) were created utilizing a design approach in which Ca^{2+}-binding sites were designed in CD2.D1 using site-directed mutagenesis. These proteins were expressed as tag-less proteins in *E. coli* tuner(DE3)pLacI as inclusion bodies and further purified using a refolding strategy followed by ion exchange chromatography. Protein expression and purification from inclusion bodies might be an efficient way to obtain active proteins with high yields and purity. CD2.7E15.52I has been used as a MRI contrast agent for tumor targeting (**Chapter 6**).

PUBLICATIONS AND MANUSCRIPTS IN PREPARATION

Huang Y., Zhou, Y., **Castiblanco, A.**, Yang, W., Brown, E. M., and Yang J. J. Multiple Ca^{2+}-binding sites in the extracellular domain of the Ca^{2+}-sensing receptor corresponding to cooperative Ca^{2+} response. **Biochemistry**, 2009. 48(2): p. 388-98.

Huang, Y., Zhou, Y., Yang, W., Butters, R., Lee, H. W., Li S., **Castiblanco, A.,** Brown, E. M., and Yang J. J. Identification and dissection of Ca^{2+}-binding sites in the extracellular domain of Ca^{2+}-sensing receptor. **J Biol Chem.**, 2007. 282(26): p. 19000-10.

Huang, Y., Zhou, Y., Wong H. C., Chen Y., **Castiblanco, A.,** and Yang J. J. A single EF-hand isolated from STIM1 forms dimer in the absence and presence of calcium. (*Under review*)

Huang, Y., Zhou, Y., Wong H. C., **Castiblanco, A.,** Chen Y., Brown E. M., and Yang J. J. Identification of calmodulin binding domain within the C-terminal region of Ca^{2+}-sensing receptor. (*In preparation*)

REFERENCES

1. McPhalen, C.A., N.C. Strynadka, and M.N. James, *Calcium-binding sites in proteins: a structural perspective.* Adv Protein Chem, 1991. **42**: p. 77-144.
2. Berridge, M.J., M.D. Bootman, and P. Lipp, *Calcium: a life and death signal.* Nature, 1998. **395**(6703): p. 645-8.
3. Dowd, D.R., *Calcium regulation of apoptosis.* Adv Second Messenger Phosphoprotein Res, 1995. **30**: p. 255-80.
4. Berridge, M.J., M.D. Bootman, and H.L. Roderick, *Calcium signalling: dynamics, homeostasis, and remodelling.* Nat Rev Mol Cell Biol, 2003. **4**(7): p. 517-29.
5. Van Eldik, L.J. and D.M. Watterson, *Calmodulin and signal transduction*. 1998, San Diego: Academic Press. xiv, 482 p.
6. Vogel, H.J., *Calcium-binding protein protocols*. Methods in molecular biology v. v. 172-173. 2002, Totowa, NJ: Humana Press.
7. Celio, M.R., T.L. Pauls, and B. Schwaller, *Guidebook to the calcium-binding proteins*. 1996, Oxford ; New York: Sambrook & Tooze Publication at Oxford University Press. xvi, 238 p.
8. Dadmun, M.D., *Computational studies, nanotechnology, and solution thermodynamics of polymer systems*. 2001, New York: Kluwer Academic/Plenum Publishers. xii, 178 p.
9. Yang, W., et al., *Structural analysis, identification, and design of calcium-binding sites in proteins.* Proteins, 2002. **47**(3): p. 344-56.
10. Huang, Y., et al., *Multiple Ca(2+)-binding sites in the extracellular domain of the Ca(2+)-sensing receptor corresponding to cooperative Ca(2+) response.* Biochemistry, 2009. **48**(2): p. 388-98.
11. Thakker, R.V., *Diseases associated with the extracellular calcium-sensing receptor.* Cell Calcium, 2004. **35**(3): p. 275-82.
12. Ye, Y., et al., *A grafting approach to obtain site-specific metal-binding properties of EF-hand proteins.* Protein Eng, 2003. **16**(6): p. 429-34.
13. Yang, W., et al., *Design of a calcium-binding protein with desired structure in a cell adhesion molecule.* J Am Chem Soc, 2005. **127**(7): p. 2085-93.
14. Yang, W., et al., *Rational design of a calcium-binding protein.* J Am Chem Soc, 2003. **125**(20): p. 6165-71.
15. Wilkins, A.L., W. Yang, and J.J. Yang, *Structural biology of the cell adhesion protein CD2: from molecular recognition to protein folding and design.* Curr Protein Pept Sci, 2003. **4**(5): p. 367-73.
16. Yang, J.J., et al., *Structural biology of the cell adhesion protein CD2: alternatively folded states and structure-function relation.* Curr Protein Pept Sci, 2001. **2**(1): p. 1-17.
17. Bodian, D.L., et al., *Crystal structure of the extracellular region of the human cell adhesion molecule CD2 at 2.5 A resolution.* Structure, 1994. **2**(8): p. 755-66.
18. Withka, J.M., et al., *Structure of the glycosylated adhesion domain of human T lymphocyte glycoprotein CD2.* Structure, 1993. **1**(1): p. 69-81.
19. Baneyx, F., *Protein expression technologies: current status and future trends*. 2004, Wymondham, Norfolk: Horizon Bioscience. x, 532 p.

20. Storz, G. and R. Hengge-Aronis, *Bacterial stress responses*. 2000, Washington, D.C.: ASM Press. xv, 485 p.
21. Lee, W., *Induction of Stress Protein Synthesis in E.Coli.* University of Nevada, Reno, 2005.
22. Kenneth, T., *Growth of Bacterial Populations.* University of Wisconsin-Madison Department of Bacteriology, 2002.
23. Mogk, A., M.P. Mayer, and E. Deuerling, *Mechanisms of protein folding: molecular chaperones and their application in biotechnology.* Chembiochem, 2002. **3**(9): p. 807-14.
24. Sorensen, H.P. and K.K. Mortensen, *Advanced genetic strategies for recombinant protein expression in Escherichia coli.* J Biotechnol, 2005. **115**(2): p. 113-28.
25. Sorensen, H.P. and K.K. Mortensen, *Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli.* Microb Cell Fact, 2005. **4**(1): p. 1.
26. Sanderson, M.R. and J. Skelly, *Macromolecular crystallography : conventional and high-throughput methods*. Oxford biosciences. 2007, Oxford: Oxford University Press. x, 281 p.
27. Endo, I. and Kagaku Kogaku Kyokai (Japan), *Bioseparation engineering : proceedings of an International Conference on Bioseparation Engineering: "Recovery and Recycle of Resources to Protect the Global Environment", organized under the Special Research Group on Bioseparation Engineering in the Society of Chemical Engineers, Japan, Nikko, Japan, July 4-7, 1999*. 1st ed. Progress in biotechnology. 2000, Amsterdam ; New York: Elsevier Science B.V. xi, 226 p.
28. *Novagen catalog*. 2002-2003.
29. Roe, S., *Protein purification techniques : a practical approach*. 2nd ed. The Practical approach series 244. 2001, New York: Oxford University Press. xvi, 262 p.
30. Janson, J.-C. and L. Rydén, *Protein purification : principles, high-resolution methods, and applications*. 2nd ed. 1998, New York: Wiley-Liss. x, 695 p.
31. Scopes, R.K., *Protein purification : principles and practice*. 3rd ed. Springer advanced texts in chemistry. 1994, New York: Springer-Verlag. xix, 380 p.
32. Meyers, R.A., *Encyclopedia of analytical chemistry: applications, theory, and instrumentation*. 2000, Chichester ; New York: Wiley.
33. Cutler, P., *Protein purification protocols*. 2nd ed. Methods in molecular biology. 2004, Totowa, N.J.: Humana Press. xii, 484 p.
34. Pellow, C.E., *Manual of practical medical and physiological chemistry*. 3rd ed. 1900, New York,: Appleton. xiv, 314 p.
35. Creighton, T.E., *Proteins : structures and molecular properties*. 2nd ed. 1993, New York: W.H. Freeman. xiii, 507 p.
36. Foster, J. and M. Sterman, *Urea weakens the protein structure, presumably by destruction of intramolecular hydrogen.* Am. Chem. Soc., 1956. **78**(15): p. 3656-3660.
37. Bennion, B.J. and V. Daggett, *The molecular basis for the chemical denaturation of proteins by urea.* Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5142-7.
38. Huang, Y., *Integration of Extracellular and Intracellular Calcium Signals: Roles of Calcium-Sensing Receptor (CaSR), Calmodulin, and Stromal Interaction Molecule 1 (STIM1)*. 2008.
39. Zhou, Y., *Exploring the role of calcium ions in biological systems by computational prediction and protein engineering*. 2007.

40. Stathopulos, P.B., et al., *Stored Ca2+ depletion-induced oligomerization of stromal interaction molecule 1 (STIM1) via the EF-SAM region: An initiation mechanism for capacitive Ca2+ entry.* J Biol Chem, 2006. **281**(47): p. 35855-62.
41. Stathopulos, P.B., L. Zheng, and M. Ikura, *Stromal interaction molecule (STIM) 1 and STIM2 calcium sensing regions exhibit distinct unfolding and oligomerization kinetics.* J Biol Chem, 2009. **284**(2): p. 728-32.
42. Liang, Y., J. Yao, and S. Gillam, *Rubella virus nonstructural protein protease domains involved in trans- and cis-cleavage activities.* J Virol, 2000. **74**(12): p. 5412-23.
43. Gerasimenko, J.V., et al., *Calcium uptake via endocytosis with rapid release from acidifying endosomes.* Curr Biol, 1998. **8**(24): p. 1335-8.
44. Zhou, Y., et al., *Identification of a Ca2+-binding domain in the rubella virus nonstructural protease.* J Virol, 2007. **81**(14): p. 7517-28.
45. Frey, T.K., *Molecular biology of rubella virus.* Adv Virus Res, 1994. **44**: p. 69-160.
46. Tsien, R.Y., *The green fluorescent protein.* Annu Rev Biochem, 1998. **67**: p. 509-44.
47. Heim, R., A.B. Cubitt, and R.Y. Tsien, *Improved green fluorescence.* Nature, 1995. **373**(6516): p. 663-4.
48. Zou, J., et al., *Developing sensors for real-time measurement of high Ca2+ concentrations.* Biochemistry, 2007. **46**(43): p. 12275-88.
49. Bergfors, T.M., *Protein crystallization : techniques, strategies, and tips : a laboratory manual*. IUL biotechnology series 1. 1999, La Jolla, Calif.: International University Line. xix, 306 p.
50. Ye, Y., et al., *Probing site-specific calmodulin calcium and lanthanide affinity by grafting.* J Am Chem Soc, 2005. **127**(11): p. 3743-50.
51. Chattopadhyay, N. and E.M. Brown, *Calcium-sensing receptor*. Endocrine updates 19. 2003, Boston: Kluwer Academic Publishers. xvii, 286 p.
52. Ward, B.K., et al., *Functional deletion of the calcium-sensing receptor in a case of neonatal severe hyperparathyroidism.* J Clin Endocrinol Metab, 2004. **89**(8): p. 3721-30.
53. Brown, E.M., et al., *Cloning and characterization of an extracellular Ca^{2+}-sensing receptor from bovine parathyroid.* Nature, 1993. **366**(6455): p. 575-80.
54. Egbuna, O.I. and E.M. Brown, *Hypercalcaemic and hypocalcaemic conditions due to calcium-sensing receptor mutations.* Best Pract Res Clin Rheumatol, 2008. **22**(1): p. 129-48.
55. D'Souza-Li, L., *The calcium-sensing receptor and related diseases.* Arq Bras Endocrinol Metabol, 2006. **50**(4): p. 628-39.
56. Drueke, T.B., *Modulation and action of the calcium-sensing receptor.* Nephrol Dial Transplant, 2004. **19**(5): p. V20-26.
57. Gao, Z.G. and K.A. Jacobson, *Keynote review: allosterism in membrane receptors.* Drug Discov Today, 2006. **11**(5-6): p. 191-202.
58. Huang, Y., et al., *Identification and dissection of Ca(2+)-binding sites in the extracellular domain of Ca(2+)-sensing receptor.* J Biol Chem, 2007. **282**(26): p. 19000-10.
59. Maniccia, A.W., et al., *Using protein design to dissect the effect of charged residues on metal binding and protein stability.* Biochemistry, 2006. **45**(18): p. 5848-56.
60. Battey, J.F., et al., *Molecular cloning of the bombesin/gastrin-releasing peptide receptor from Swiss 3T3 cells.* Proc Natl Acad Sci U S A, 1991. **88**(2): p. 395-9.

61. Tyszka, J.M., S.E. Fraser, and R.E. Jacobs, *Magnetic resonance microscopy: recent advances and applications.* Curr Opin Biotechnol, 2005. **16**(1): p. 93-9.
62. Woods, M., D.E. Woessner, and A.D. Sherry, *Paramagnetic lanthanide complexes as PARACEST agents for medical imaging.* Chem Soc Rev, 2006. **35**(6): p. 500-11.
63. Yang, J.J., et al., *Rational design of protein-based MRI contrast agents.* J Am Chem Soc, 2008. **130**(29): p. 9260-7.