Optimizing the Production Parameters of Engineered Protein-Based MRI Contrast Agents

Rose Auguste
Georgia State University

Follow this and additional works at: http://scholarworks.gsu.edu/chemistry_theses

Recommended Citation
Auguste, Rose, "Optimizing the Production Parameters of Engineered Protein-Based MRI Contrast Agents." Thesis, Georgia State University, 2014.
http://scholarworks.gsu.edu/chemistry_theses/61

This Thesis is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
OPTIMIZING THE PRODUCTION PARAMETERS OF ENGINEERED PROTEIN-BASED MRI CONTRAST AGENTS

By

ROSE AUGUSTE

Under the Direction of Dr. Jenny J. Yang

ABSTRACT

Magnetic Resonance Imaging is a noninvasive, complex imaging modality that uses contrast agents to enhance sensitivity and resolution for a clear image enhancement. To optimize the relaxation and targeting properties, our lab has created several protein-based contrast agents, ProCAs. To enable various biophysical characterizations and preclinical studies, we need to optimize expression and purification of the engineered ProCAs. ProCA1 variants are cultured in various cell strains during a tag-less method. ProCA32 with multiple binding sites is also successfully purified. It has a strong metal-binding capability with a relaxivity ($r_2$: 27 ± 0.54; $r_1$: 20.6 ± 0.20 mM$^{-1}$s$^{-1}$) that is significantly greater than the clinically used contrast agent, diethylene triamine pentaacidic acid, DTPA ($r_1$: 3.8 mM$^{-1}$s$^{-1}$). Our findings in optimizing expression and purification conditions, structural conformation, metal binding, and relaxivity measurements of these ProCAs will facilitate in vitro and in vivo studies needed to move forward into the pre-clinical then clinical phase.

INDEX WORDS: Expression, Tag-less refolding method, GST-tag refolding, Gadolinium, Metal binding, Relaxivity, Structural Conformation
OPTIMIZING THE PRODUCTION PARAMETERS OF ENGINEERED PROTEIN-BASED MRI CONTRAST AGENTS

by

ROSE AUGUSTE

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master Thesis in the College of Arts and Sciences Georgia State University 2014
OPTIMIZING THE PRODUCTION PARAMETERS OF ENGINEERED PROTEIN-BASED MRI CONTRAST AGENTS

by

Rose Auguste

Committee Chair: Jenny J. Yang

Committee: Donald Hamelberg
            Siming Wang
            Robert Wohlhueter
            Jun Yin

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
May 2014
DEDICATION

I dedicate this work to my family and friends for their unwavering support and motivation that helped me maintain a positive attitude throughout my years of research. A very special thanks goes to my parents: Joël and Andrèle Auguste, for their constant encouragement, helping me during my rehearsal practice and preparing mentally beforehand for my thesis defense. I am very grateful to my younger sisters, Adèle and Arielle, for providing the laughter, the ice breaking moments of relaxation so badly needed at time of intense work throughout my journey in the graduate program. I want to acknowledge as well my extended family members, my godparents, Frantz and Gaëtane Auguste, and my aunts, tatie Shirley, tatie Gege, tatie Marjorie, tatie Mirna, my long list of cousins, for the mental support and constant encouragement to move forward with what I am doing. I would like to finish this dedication by giving an immense thank to the matriarchs of my family, my grandmothers, Rose Claire Auguste who passed away long before I was born, whose name I bear and who inspired me through her past life, and Leanne Zamor, who I love greatly and is currently living in Port-au-Prince. These two women’s stories have been a constant inspiration for continuing my career and life goal to do greater good in life; both women have been unassuming in their own way assisting others and have stayed humble in the process. My continuous appreciation again goes to my parents, for being the rock and anchor of my life, for motivating me during the times when I feel low and to embrace me no matter the situation. I carry on this immense gratitude to all of these people in my family and my friends, Liza Ngo, and Faith Emmanuel, who stood by me throughout this journey that I’ve embarked upon since summer 2011.
ACKNOWLEDGEMENTS

Great gratitude goes to Dr. Jenny J. Yang for fostering the opportunity to experience research applications linked to the field of biochemistry and biophysics. In addition, I appreciate the efforts my advisor, Dr. Yang, has taken time in pushing me to perform to the best of my ability and in overcoming my limitations in terms of oral presentation and writing skills. I am sincerely honored and grateful to have been a part of this interesting MRI contrast agent project. I would like to acknowledge the committee members of my research project: Dr. Hamelberg, Dr. Wang, Dr. Wohlhueter, and Dr. Yin, for the helpful editing suggestions and the additional writing insight for my thesis paper. In addition, I greatly appreciate my committee members for their patience waiting for my thesis draft and taking time out of their busy schedule to attend the oral presentation of my thesis defense. I thank my mentors, Natalie White, Fan Pu, and Dr. Jingjuan Qiao for their guidance, knowledgeable advice, experience and expertise provided in this study. I give great thanks to every group member in the lab that have volunteered their personal time to explain or clarify certain lab procedures, namely the handling of SDS-PAGE (Dr. Yanyi Chen and Juan Zou, FPLC (Dr. You Zhou, Shenghui Xue, Dr. Chen Zhang, and Dr. Jingjuan Qiao), UV-Vis spectrophotometry, and Fluorescence spectroscopy (Dr. Shenghui Xue, Fan Pu, and Dr. Yanyi Chen). I give a gracious thank you to all of Dr. Yang’s current and past research group members for the support and help: Dr. Yusheng Jiang, Jie Feng, Xueyun Liu, Dr. Jie Jiang, and Florence Reddish. A kind thank you goes to the following new lab members: Li Zhang, Anvi Patel, Fantashia Goolsby, Mani Salarian, Chaky, Corrie Purser, Rakshya Gorkhali, Shanshan Tan, and Cassie Miller, who joined during the two years I have been pursuing my Master’s Thesis. A special thanks goes to Cassie, Anvi, and Corrie, for giving me pointers on presenting and helping me prepare for my oral defense.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xvii</td>
</tr>
<tr>
<td><strong>1. INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Medical Imaging in Cancer Research</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1 <em>Imaging modalities applied in medical research</em></td>
<td>1</td>
</tr>
<tr>
<td>1.2 Principles and Applications in MRI</td>
<td>4</td>
</tr>
<tr>
<td>1.2.1 <em>A brief history on the development of MRI</em></td>
<td>4</td>
</tr>
<tr>
<td>1.2.2 Relaxivity Theory</td>
<td>4</td>
</tr>
<tr>
<td>1.2.3 Characteristics of current contrast agents used for MRI</td>
<td>6</td>
</tr>
<tr>
<td>1.2.4 The toxicity level and nephrogenic systemic fibrosis relationship with GBCAs</td>
<td>6</td>
</tr>
<tr>
<td>1.3 Development of the Novel Protein-Based Contrast Agents (ProCAs) for MRI</td>
<td>11</td>
</tr>
<tr>
<td>1.3.1 <em>Criteria to consider prior to applying the design approach for protein-based Contrast agents</em></td>
<td>11</td>
</tr>
<tr>
<td>1.3.2 Model for the contrast agent: Background on the scaffold protein</td>
<td>11</td>
</tr>
<tr>
<td>1.3.3 Design approach for the novel ProCA1 variants</td>
<td>12</td>
</tr>
<tr>
<td>1.3.4 The Grafting approach for promotion of the ProCA1 targeting capability</td>
<td>13</td>
</tr>
<tr>
<td>1.4 Amount of Protein Needed After Optimizing Protein Expression and Purification</td>
<td>19</td>
</tr>
<tr>
<td>1.5 Overall Purpose of this Study</td>
<td>21</td>
</tr>
</tbody>
</table>
1.5.1 Production of these novel protein-based MRI Contrast Agents ........................................21

1.6 Overview of this Thesis ........................................................................................................23

2. EXPRESSION AND REFOLDING PURIFICATION OF TAG-LESS PROCA1 VARIANTS IN VARIOUS E. COLI CELL STRAINS ........................................................................................................27

2.1 Introduction ..........................................................................................................................27

2.2 Methods ................................................................................................................................32

2.2.1 Cloning ..............................................................................................................................32

2.2.2 Transformation ..................................................................................................................32

2.2.3 Small and large-scale expression .......................................................................................33

2.2.4 Tag-less refolding method ...............................................................................................33

2.2.5 FPLC-anion exchange Q – column ..................................................................................34

2.2.6 Quantitative analysis of the 15 % SDS-PAGE through Image J software .........................34

2.2.7 Ultraviolet-visible absorbance spectra and agarose gel electrophoresis applications 35

2.3 Results and Discussion ........................................................................................................40

2.3.1 Host cell strain influence on the expression of tag-less charged ProCA1 variants ......40

2.3.2 The denatured-refolding purification of tag-less charged ProCA1 variants ....................43

2.3.3 The refolded tag-less ProCA1 charged variants isolated by the FPLC Q-column anion

exchanger ..................................................................................................................................44

2.4 Conclusion and Future Work ...............................................................................................50

3. OPTIMIZATION OF EXPRESSED AND PURIFIED GST-TAGGED PROCA1 VARIANTS .........................52

3. 1 Introduction ..........................................................................................................................52
3.2 Methods ................................................................................................................................. 59
  3.2.1 Transformation .................................................................................................................... 59
  3.2.2 Small-scale and large-scale expression of GST-tag ProCA1 variants................................. 59
  3.2.3 GST-tag gravity column purification .................................................................................... 59
  3.2.4 Refolding of GST-tagged ProCA1 variants ......................................................................... 60
  3.2.5 FPLC - cation exchange SP - column and ultrafiltration application ............................... 61
  3.2.6 Quantifying the SDS-PAGE results of the GST-tag purification ....................................... 61
  3.2.7 Western-Blot and SDS-PAGE ............................................................................................ 62

3.3 Results and Discussion ........................................................................................................ 67
  3.3.1 Expression of GST-tagged CD2.WT, ProCA1, and ProCA1 variants ............................... 67
  3.3.2 Purification of GST-tag CD2.WT, ProCA1, and ProCA1 variants .................................... 80

3.4 Conclusion and Future Work ............................................................................................... 99
  3.4.1 Summary of the small and large-scale expression for all ProCA1 variants ........................ 99
  3.4.2 The summarized GST-tag purification of charged and/or targeted, ProCA1 variants 102

4. CONFORMATIONAL ANALYSIS, METAL BINDING, AND RELAXIVITY PROPERTIES OF THE ENGINEERED PROCA1S ................................................................................................................................. 106
  4.1 Introduction ................................................................................................................................ 106

  4.2 Methods .................................................................................................................................. 108
    4.2.1 Intrinsic tryptophan fluorescence emission spectra .......................................................... 108
    4.2.2 Far Ultraviolet Circular Dichroism ..................................................................................... 108
    4.2.3 ProCA1 variants binding to Tb$^{3+}$ through an NTA-Tb$^{3+}$ buffer system ..................... 109
4.2.4 FLuo-5N competition assay through fluorescence spectroscopy ........................................113

4.2.5 Relaxivity measurement ...................................................................................................114

4.3 Results and Discussion ........................................................................................................117

4.3.1 The conformational analysis of CD2 mutants using Far UV-CD ........................................117

4.3.2 Intrinsic tryptophan emission of 7E15Q and ProCA1.affi342 targeted to HER2/neu 121

4.3.3 NTA-Tb³⁺ buffer system fluorescence titration ..................................................................123

4.3.4 FLuo-5N competition assay ..............................................................................................127

4.3.5 The summary of the metal-binding characteristics for all ProCA1 variants ......................127

4.3.6 Relaxivity measurements for ProCA1 variants ..................................................................130

4.4 Conclusion and Future work ...............................................................................................135

5. PURIFICATION AND CHARACTERIZATION OF PROCA32 FROM A LARGE SCALE FERMENTATION ......................................................137

5.1 Introduction ..........................................................................................................................137

5.2 Methods ................................................................................................................................139

5.2.1 Cloning ...............................................................................................................................139

5.2.2 Purification procedure of ProCA32 ..................................................................................139

5.3 Results and Discussion .........................................................................................................140

5.3.1 The purification of ProCA32 from expression through fermentation ...............................140

5.3.2 The study on the metal-binding capability of ProCA32 ..................................................149

5.4 Conclusion and Future work ...............................................................................................157

6. SIGNIFICANCE AND MAJOR FINDINGS OF THIS THESIS ............................................................................................................158
LIST OF TABLES

Table 1.1 Comparing the advantages and disadvantages of current imaging techniques [2-5]. ...........3
Table 1.2 Summary of the amount of protein needed for each biophysical analysis. .........................20
Table 1.3 The summary of ProCA variants used in this study [20, 29, 34-36].................................26
Table 2.1 A summary of the properties for the four *E.Coli* cell strains [37, 48-51] ...........................31
Table 2.2 Comparing the advantages and challenges for the refolding purification method [40-43, 52, 53].................................................................................................................................31
Table 3.1 A summary of the UV-Vis absorbance ratio and yield (mg/ 1 L culture) for each eluted or concentrated samples from the GST-tag and refolding method of the charged ProCA1 variants. ........89
Table 3.2 A summary of the UV-Vis absorbance ratio and yield (mg/L) for each eluted or concentrated samples from the GST-tag and refolding method of the targeted ProCA1 variants. ..................98
Table 3.3 The summary of the small-scale and large-scale expression of GST-tag ProCA1 variants. ....101
Table 4.1 A summary of the percent secondary structure construct of CD2.WT, 7E15Q and ProCA1.affi1907 targeted to EGFR based on the far UV CD spectra .........................................................120
Table 4.2 A summary of the dissociation constants of Tb$^{3+}$ and Gd$^{3+}$ toward the ProCA1 variants [22, 34].................................................................................................................................129
Table 5.1 The summarized relaxivity values and $K_d$ value for the metal binding study for ProCA32. ..156
LIST OF FIGURES

Figure 1.1 A timeline history leading to the establishment of clinical MRI applications [6] ......................... 8
Figure 1.2 Variations of the clinically used gadolinium-based contrast agents (GBCAs) for MRI application system [15] .................................................................................................................. 9
Figure 1.3 The relaxivity process of excited proton nuclei during MRI scans. ........................................... 10
Figure 1.4 Schematic representation of the factors that affect a contrast agent’s relaxation properties. ................................................................................................................................. 15
Figure 1.5 The alignment of the ribbon surface between human and rat derived CD2.D1. ....................... 16
Figure 1.6 The binding pocket for the single site mutation of ProCA1 variants. ................................. 17
Figure 1.7 The different factors involved when using affibody as a multimodal probe for MRI scan. 18
Figure 1.8 A scheme of the overall goal in this thesis study................................................................. 22
Figure 2.1 The general factors to consider for the expression and purification systems[46, 47]........... 30
Figure 2.2 The transformation scheme along side the pET-20b vector containing the T7 promoter site. .......................................................................................................................... 36
Figure 2.3 The expression scheme represents the stepwise method used for the tag-less ProCA1s’ overexpression in the four different E.coli competent cell strains ........................................... 37
Figure 2.4 The denaturing and refolding steps are described in a detailed step-by-step method for the tag-less purification of each ProCA1s . ........................................................................... 38
Figure 2.5 The FPLC-Q column preparation scheme and programmed salt gradient. ......................... 39
Figure 2.6 The growth curve for the small-scale expression of ProCA1 in a) BL21 (DE3), Tuner, and BL21 (DE3) pLysS (in 10 mL LB medium volume) b) with the corresponding 15% SDS-PAGE, and c) In following, the low expression intensity for all three cell strains ..................................................... 41
Figure 2.7 The large-scale expression of ProCA1 variants and the corresponding growth curve for each cell strain.................................................................42

Figure 2.8 The 15 % SDS-PAGE for all ProCA1 variants are demonstrated for the tag-less refolding method........................................................................................................45

Figure 2.9 The 7E15E variant is isolated through a FPLC system using an anion exchanger column (Q-column)........................................................................46

Figure 2.10 Ion exchange FPLC Chromatograms of dialysis and sonication samples for the following proteins: a) 7E15E flask 1 ........................................................................47

Figure 2.11 The overall yield for each variant is summarized in correspondence to the cell strain condition........................................................................................................48

Figure 2.12 Size-exclusion chromatography used for isolation of impurities from the concentrated samples post refolding purification method........................................................................................................49

Figure 3.1 The schematic representation of the GST-tag ProCA1 and ProCA1.affi342..........................56

Figure 3.2 The specific parameters for optimizing the expression yield of GST-tagged ProCA1 variants. ........................................................................................................................................................................................................................................57

Figure 3.3 The outlined steps for optimizing the parameters in the purification system of the GST-tagged ProCA1 variants........................................................................................................................................................................................................................................58

Figure 3.4 The commercially available pGEX-2T plasmid vector [55] has a GST promoter site before the Eco RI site. ........................................................................................................................................................................................................................................64

Figure 3.5 The scheme for GS-4B column purification of GST-tagged ProCA1 variants. ..................65

Figure 3.6 The refolding procedure of the inclusion body (IB) of GST-tagged ProCA1 variants..........66

Figure 3.7 The small-scale expression of GST-tagged CD2.7E15E, CD2.7E15N, and ProCA1 ..........71

Figure 3.8 The large-scale expression of GST-tagged CD2.D1 wild type........................................72
Figure 3.9 The large-scale expression of GST-tagged ProCA1 and CD2.7E15E at varying cell strains, IPTG concentration, and/or induction temperature. .................................................................73
Figure 3.10 The large-scale expression of GST-tag CD2.7E15N and CD2.7E15Q at varying IPTG concentration and induction temperature. .................................................................74
Figure 3.11 The small-scale expression of GST-tag ProCA1.ZHER:342 and ProCA1.ZEGFR:1907 ....................77
Figure 3.12 The large-scale expression of GST-tagged ProCA1.ZHER2:342 at varying cell strains, IPTG concentration, and induction temperature. .................................................................78
Figure 3.13 The large-scale expression of GST-tag ProCA1.ZEGFR:1907 at varying cell strains, IPTG concentration, and induction temperature. .................................................................79
Figure 3.14 The GST-tag purification of ProCA1, CD2.7E15E, CD2.7E15N, and CD2.7E15Q with modifications in the cell lysis buffer .............................................................................................................85
Figure 3.15 A summary of the intensity ratio for each sample from the GST-tag purification of ib) ProCA1, ia) 7E15E, ii) 7E15N, and iii) 7E15Q with the indicated lysis buffer conditions ......................86
Figure 3.16 The GST-tag refolding purification of CD2.7E15E, CD2.7E15N, and CD2.7E15Q .............87
Figure 3.17 An example of the FPLC chromatograph for CD2.7E15Q (pH: 3.80) through the pH gradient in the SP column .............................................................................................................88
Figure 3.18 The GST-tag purification of ProCA1.affi342 and ProCA1.affi1907 with modifications in the cell lysis buffer .............................................................................................................93
Figure 3.19 A summary of the intensity ratio for each sample from the GST-fused purification of ProCA1.affi342 and ProCA1.affi1907 with modifications in the cell lysis buffer. .........................94
Figure 3.20 The GST-tag refolding purification of ProCA1.affi342 and ProCA1.affi1907 ......................95
Figure 3.21 A summary of the intensity ratio for each sample from the GST-tag refolding purification of ProCA1.affi342 and ProCA1.affi1907 .................................................................................................96
Figure 3.22 The FPLC chromatograph of ProCA1.affi1907 through the Q column based on a salt-gradient. .................................97

Figure 3.23 The summary of the GS-4B purification of each ProCA1 variants; charged and targeted. 105

Figure 4.1 The schematic representation of NTA-Tb$^{3+}$ buffer system .................................................................115

Figure 4.2 The schematic representation of FLuo-5N competition assay .................................................................116

Figure 4.3 Comparing the secondary structure of CD2.WT versus ProCA1 variants versus ProCA1.affi1907 through Far UV-CD. ........................................................................................................119

Figure 4.4 The effect of the metal ions on the secondary structure of a) 7E15Q and b) ProCA1.affi1907. ..............................................................................................................................................120

Figure 4.5 shows the intrinsic tryptophan emission spectra of ProCA1 and ProCA1.affi342 compared with the free Trp in 10 mM HEPES, pH 7.0. ........................................................................................................122

Figure 4.6 The determination of the free terbium concentration (μM) range between NTA-Tb$^{3+}$ and ProCA1-Tb$^{3+}$ buffer system ........................................................................................................125

Figure 4.7 Fluorescence emission for 7E15E, ProCA1.affi.EGFR, 7E15N, and 7E15Q in NTA-Tb$^{3+}$ buffer system ..............................................................................................................................................126

Figure 4.8 Relaxivity time (T$^{1,2}$, ms) and rate (r$^{1,2}$, mM$^{-1}$ s$^{-1}$) for ProCA1 charged variants .................132

Figure 4.9 The average relaxivitys, r$^{1,2}$ (mM$^{-1}$ s$^{-1}$), of the charged ProCA1 variants with 1:1 ratio concentration with Gd$^{3+}$ ........................................................................................................................................133

Figure 4.10 The averaged relaxivitys, r$^{1,2}$ (mM$^{-1}$ s$^{-1}$), of targeted ProCA1.affibody variants at varying protein to metal ratio concentrations ........................................................................................................134

Figure 5.1 The cartoon of the holo form rat-alpha parvalbumin with the indicated two EF-hand sites. ..................................................................................................................................................138

Figure 5.2 The purification gel of ProCA32 based on the purification method from Shenghui Xue’s dissertation. ........................................................................................................................................143
Figure 5.3 The isolated DNA/nucleotide contamination is successful as indicated by the agarose gel

Figure 5.4 The UV-Vis spectra indicate successful isolation of the protein of interest within the corresponding peak fractions.

Figure 5.5 The 15% SDS-PAGE of the isolated ProCA32 to the other contaminants.

Figure 5.6 Comparing the absorbance ratio between each fraction range for ProCA32.

Figure 5.7 The final yield corresponding to ~ 180 g from the 2 liter aliquot pellet of fermented ProCA32.

Figure 5.8 Fluorescence emission for 7E15Q and ProCA1.affi.EGFR in Fluo-5N competition assay.

Figure 5.9 The fluorescence Intensity of NTA-Tb³⁺ buffer system with ProCA32.

Figure 5.10 The different relaxivity values for ProCA32 in different expression/purification procedure.

Figure 5.11 The saturated a) T₁, T₂ (ms) values and r₁,r₂ (mM⁻¹s⁻¹) values for the fermented ProCA32 with a fixed Gd³⁺ (50 μM). 
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>After Induction</td>
</tr>
<tr>
<td>BAB-OC</td>
<td>Beads After Binding Outside the Column</td>
</tr>
<tr>
<td>BAB-IC</td>
<td>Beads After Binding Inside the Column</td>
</tr>
<tr>
<td>BAE</td>
<td>Beads After Elution</td>
</tr>
<tr>
<td>BL</td>
<td>BL21 (DE3)</td>
</tr>
<tr>
<td>BLpLsS</td>
<td>BL21 (DE3) pLysS</td>
</tr>
<tr>
<td>CD2.D1</td>
<td>Cell Cluster of Differentiation 2, domain 1</td>
</tr>
<tr>
<td>CP</td>
<td>Cell Pellet</td>
</tr>
<tr>
<td>DNA/RNA</td>
<td>Deoxyribonucleic acid/ Ribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EGFR</td>
<td>Epidermal Growth Factor Receptor</td>
</tr>
<tr>
<td>EGTA</td>
<td>Ethylene glycol tetraacetic acid</td>
</tr>
<tr>
<td>GS-4B</td>
<td>Glutathione Sepharose- 4 beads</td>
</tr>
<tr>
<td>GST/GSH</td>
<td>Glutathione-s-transferase/Glutathione</td>
</tr>
<tr>
<td>HER2/neu</td>
<td>Human Epidermal Growth Factor Receptor 2</td>
</tr>
<tr>
<td>IB</td>
<td>Inclusion Body</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropylthio-β-galactoside</td>
</tr>
<tr>
<td>LB broth</td>
<td>Luria-Bertani broth</td>
</tr>
<tr>
<td>NTA</td>
<td>Nitrilotriacetic acid</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>ProCA</td>
<td>Protein-based MRI Contrast Agent</td>
</tr>
<tr>
<td>Q- column</td>
<td>Quaternary ammonium anion exchanger column</td>
</tr>
<tr>
<td>RpLsS</td>
<td>Rosetta-gami (DE3) pLysS</td>
</tr>
<tr>
<td>SN</td>
<td>Supernatant</td>
</tr>
<tr>
<td>SP-column</td>
<td>Sulfopropyl cation exchanger column</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

1.1 Medical Imaging in Cancer Research

Cancer is among the leading causes of death worldwide. It follows that preclinical and clinical applications are essential for characterizing new drug developments that can be used for the purpose of cancer treatment, early and effective detection, and accurate prognosis prior to therapy. Mutations cause uncontrolled proliferation of cell growth to other regions inside the body and result in metastasis. Essentially, these malignant cells are known as tumors that cause varying forms of cancer.

An accurate determination of the patient’s prognosis is essential for a clinician to make the decision for treatment options. The patient goes through a specified treatment therapy based on the progression or stage at which the cancer has developed in the body. Essentially, treatment and diagnosis rely on the accurate prognosis of the particular disease. Accurately identifying the level of the cancer helps clinicians have a better diagnosis for treatment and essential suppression of further cancer growth by continuous therapy [1]. There are several stages of a cancer’s progression. It is essential to locate any tumors that consist of $>10^9$ cells in diameter at the early stage of cancer and are treatable in the premalignant stage [1]. The later stages essentially indicate that the cancer has spread to other regions from the primary tumor site and, at this point, treatment is most challenging. Thus, it is critical to monitor disease progression at an early stage, and molecular imaging allows this procedure in the process of cancer prognosis and treatment.

1.1.1 Imaging modalities applied in medical research

In the medical field, various instrumentations are applied for imaging. These various modalities are used for obtaining anatomical and functional information on the biological process within the body. There are medical techniques used to indicate and monitor the progression of tumor cell growth. The essence of cancer diagnosis is the use of imaging applications that obtain a level of sensitivity and
efficacy for detecting cancerous versus non-cancerous tissue. Some of the most advanced clinical imaging instruments are positron emission tomography (PET), computed tomography (CT), single photon emission computed tomography (SPECT), optical imaging, ultrasound, magnetic resonance spectroscopy, and magnetic resonance imaging (MRI). These modalities are complementary to one another in the sense that each possesses benefits and limitations to the level of functional and anatomical information that can be achieved (Table 1.1). They are among some of the techniques applied for the identification and localization of tumors in order to determine a proper prognosis for further clinical treatments.

Unlike PET and CT, MRI is a noninvasive medical imaging modality that is not dependent on the use of ionizing radiation. This application results in a 3 dimensional image of the soft tissue when exposed to a particular magnetic field strength. Both anatomical and functional information is achieved with the use of MRI. Each imaging modality holds a different level of sensitivity for detection of tumor cells. For instance, PET and SPECT have high sensitivity and depth penetration but use irradiation exposure during image scanning with isotopic labeled agents. Similar to MRI, ultrasound, optical imaging and MRS are also non-ionizing imaging tools. While ultrasound bio microscopy is widely used for in vivo applications, it has limitations for body depth and resolution [2]. MRI provides soft tissue enhancement, and the important idea is to differentiate the non-cancerous tissue from the cancerous ones in an effective manner. In this case, the sensitivity is low, so a contrast agent is used to increase the sensitivity and resolution of the MRI signal.
Table 1.1 Comparing the advantages and disadvantages of current imaging techniques [2-5].

<table>
<thead>
<tr>
<th>Imaging Technique</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ultrasound (US)</td>
<td>Real-time imaging, external and internal (endoscopy) application</td>
<td>Possible tissue heating from transducer, gas production in tissues</td>
</tr>
<tr>
<td>Optical Imaging</td>
<td>High-throughput screening for target confirmation</td>
<td>Low depth penetration (&lt;1cm), whole body imaging not possible</td>
</tr>
<tr>
<td>Magnetic Resonance Imaging (MRI)</td>
<td>3D spatial resolution and soft tissue contrast without depth limitation</td>
<td>Sensitivity dependent on molecular contrast agent for detection</td>
</tr>
<tr>
<td>Computed Tomography (CT)</td>
<td>High spatial resolution and unlimited depth penetration</td>
<td>Radiation exposure, poor soft tissue contrast</td>
</tr>
<tr>
<td>Positron emission Tomography (PET)/Single Photon Emission Computed Tomography (SPECT)</td>
<td>Hybrid tool with MR or CT for anatomical info</td>
<td>Limited to low spatial resolution, radiation exposure</td>
</tr>
<tr>
<td>Magnetic Resonance Spectroscopy (MRS)</td>
<td>Whole body imaging, no ionizing radiation</td>
<td>Low sensitivity in micromolar range</td>
</tr>
</tbody>
</table>
1.2 Principles and Applications in MRI

1.2.1 A brief history on the development of MRI

Magnetic resonance imaging (MRI) was formally known as “nuclear magnetic resonance (NMR) imaging”. The name was changed to make it less intimating. Jean Baptist Joseph Fourier is among the key pioneers of contemporary NMR and MRI techniques [6]. MRI operates on the same basic principles as NMR, the only difference is that a spectrum is obtained for NMR, while MRI outputs a set of 3 dimensional, spatially resolved images [7]. In 1975, Richard Ernst was able to apply Fourier’s famous mathematical formula and obtain image construction from the MRI [6]. MRI has evolved dramatically since the initial design in 1980. Scientists Tesla, Lamor, and Rabi paved the way for the use of NMR imaging, now commonly known as MRI. The clinical MRI field is progressing with new ideas for pre-clinical and clinical applications [8-10].

1.2.2 Relaxivity Theory

MRI signals are obtained from the water protons, provided by the body, when exposed to a primary magnetic field ($B_0$) and radiofrequency (rf) pulse. Specifically, the human body constitutes of ~70 % water, or ~90 M of hydrogen atoms provided by these water molecules [11]. MRI depends on the relaxation of the single proton nuclei for an image signal output. These hydrogen atoms have nuclei with multiple spin states, which hold a magnetization. During an MRI scan, these nuclei present in the various tissues inside the body, and will align with the available magnetic field, causing an overall net magnetization. During this process, the nuclei’s vector is at either spin-up or spin-down direction in alignment with the magnetic field ($B_0$). The applied radio-frequency pulse then generates a rotational torque ($\tau_r$) on the nuclei’s vector. This process then causes precession of the nuclei spin, that is a spiral motion along the applied rf angle, between the spin up and spin down directions. This process is dependent on the magnetic field strength ($B_0$), gyromagnetic property of the nuclei ($\gamma$), the Lamor
frequency ($\omega$) and is represented mathematically in Equation 1-1. This initial information gives a clear understanding on how the MR signal is achieved as an image. The low sensitivity of the MRI results in a low enhancement between the different tissue contrasts. Thus, a contrast agent (CA) is used to overcome this image enhancement limitation. The process by which a contrast agent is administered and relaxes the water’s proton molecules is relaxation rate. The contrast agent in complex form with a metal ion in solution demonstrates rotational correlation time ($\tau_r$), longitudinal ($r_1$), and transverse ($r_2$) relaxation rate characteristics. $T_1$ and $T_2$ relaxation time values each correspond to the reciprocal of the longitudinal and transverse relaxation rates. Due to the added CA, a better image contrast is observed for the soft tissues. Each tissue has different levels of water and so they will observe different relaxation rate and contrast enhancement.

\[
\omega = \gamma B_0
\]

Equation 1-1

Relaxivity is observed when relaxation rate difference is measured as a function of time. Specifically, this process is the difference in relaxivity demonstrated after the exposure of the CA. Relaxation time is an important parameter that occurs in two parts: 1) the nuclei spins de-phasing along the x-y plane is called the spin-spin, or $T_2$ transverse relaxation, 2) the nuclei spin aligning back with the initial applied magnetic field ($B_0$) along the z-y plane is known as $T_1$ longitudinal relaxation. The main factors that effect relaxivity are as demonstrated in Figure 1.4 a as a scheme. These factors are as follow: 1) the water proton exchange dynamic, 2) the contrast agent rotational tumbling rate, 3) the inner- and outer- sphere water molecules in the surrounding environment, 4) the metal ion and water oxygen ionic bond distance, 5) the construct and chemical characteristic of the chelating agent. These factors are all considered and correlate to the criteria for the design approach for our lab’s novel protein-based MRI contrast agent.
1.2.3 Characteristics of current contrast agents used for MRI

MRI has a high dependence on contrast agents (CAs) for obtaining clear images. The clinically approved CA, Gd\(^{3+}\)-DTPA (diethylene triamine pentaacetic acid), has a bound lanthanide ion that leads to increased enhancement of the MRI. The developed contrast agents are specifically designed for Gd\(^{3+}\) binding to form a Gd\(^{3+}\)-complex. There is a variety of metal-based CAs used for MRI that has either a paramagnetic or ferromagnetic property. The paramagnetic-based CAs demonstrates favorable T\(_1\)-weighted imaging, which means that the image enhancement of soft tissue organs (light) compared to muscle and bone (dark) are brighter after obtaining the MRI scan. Thus, tumor cell localization during pre-clinical imaging is facilitated based on the light to dark intensity. Iron oxide and nanoparticle (NP) based agents are examples of ferromagnetic-based CAs developed for use in vivo [12]. T\(_2\)-weighted iron NPs demonstrate the opposite enhancement effect on an MRI scan that is less desired for medical imaging, and the iron nanoparticle agents pose a challenge for maintaining stability during synthesis [13].

The lanthanide compounds, specifically gadolinium (Gd\(^{3+}\)), are important for obtaining a reasonably high relaxation rate. The focus of this study addresses gadolinium-based contrast agents. Rotational correlation time (\(\tau_{\text{rot}}\)) and relaxation rates are two important components that contribute to a clear MRI. In addition, these parameters are directly proportional. Free Gd\(^{3+}\) are highly toxic to the body.

1.2.4 The toxicity level and nephrogenic systemic fibrosis relationship with GBCAs

Nephrogenic systemic fibrosis (NSF) is a disease caused by the exposure of high levels of free metal ion in the body. Specifically, a study has followed up on the claim that this disease is mainly present in patients who suffer from chronic kidney disease or kidney failure. These specific patients are more prone to develop the symptoms associated with this disease. A symptom involves the discoloration of skin and organs in the body such as heart, lungs, and the central nervous system. Some studies associated the development of the disease and patients with renal failure that are administered
a dosage of small GBCA. The link between NSF development and GBCAs dosage brought an emphasis on safety by the FDA. The patients who suffer from kidney failure are administered other type of CAs or a lower dosage then the clinically used dosage for obtaining an MRI scan. Other precaution in which the FDA has developed is in regulating the pharmacokinetics and bio-distribution of the Gd-based contrast agents. Once administered in vivo. This same study also indicates that from a large group of test subjects, the results were inconclusive in terms of supporting the initial claim that renal impaired patients are prone to develop NSF due to the administration of GBCAs. Thus, the correlation between the dosage amount of GBCAs and patients with kidney failure are on a going surveillance since the initial outbreak cases in 2001.

Specifically, GBCAs are composed of a small organic compound bound to gadolinium ion. The CA provides a clear difference in the image contrast due to an increase relaxation rate. Gd\(^{3+}\) ions interact with the water molecules present in the surrounding soft tissue, and this interaction gives off a relaxation rate in the form of image enhancement. However, Gd\(^{3+}\)\textsuperscript{-}-DTPA and other small chelator based MRI contrast agents shown in Fig 1.2 have a relatively low relaxation time (\(r_1; 3.8 \text{ mM}^{-1} \text{s}^{-1}\)) due to their small molecular size. The molecular imaging contrast agent database (MICAD) lists twelve different derivative based agents used for MRI [14]. Thus, our lab has designed a protein based contrast agent, (ProCA), which possesses a higher relaxation rate compared to the CAs currently in use clinically.
In 1891, Tesla coil is developed, Nikola Telsa also developed the rotating magnetic field where the unit strength of a magnetic field is known as 1 Telsa = 1 Newton/Ampere-meter.

Early 1900’s, Lamor equation states \( \omega = \gamma B_0 \)
In 1924, Gerlach and Stern demonstrated the quantum effect from a magnetic moment of silver atoms through a molecular beam deflection.

In the 1930- mid 1940, Isidor Rabi detected and measured single states of rotation of atoms and molecules, and determined the nuclei magnetic moment.

In 1967, Ganssen patented a work involved in whole-body NMR tool for blood flow measurement using small coils.

In 1980, the prototype for whole-body imaging was developed by Edelstein et al. and the imaging time acquires ~5 min.
In 1983, published work from Leon Axel’s group showed the 3D nature of MRI through imaging of the cardiovascular anatomy.

In 1973-1974, Lauterbur and Mansfield described using NMR signals for spatial localization to obtain 2D imaging.
In 1975, Ernst ‘s publication on NMR signals for image reconstruction.

In 1971, Raymond Damadian demonstrated the ability to measure \( T_1 \) and \( T_2 \) relaxation times of rat tumor tissue and that cancer tissue holds longer relaxation time vs. normal tissue.

In 1984, Initial published data on Gd\(^{3+}\)- enhancement in clinical MRI.
In 1986, Imaging time was modified to 5 sec with minor change in image resolution.

Figure 1.1 A timeline history leading to the establishment of clinical MRI applications [6].
Figure 1.2 Variations of the clinically used gadolinium-based contrast agents (GBCAs) for MRI application system [15].

<table>
<thead>
<tr>
<th>Name</th>
<th>Formula</th>
<th>$r_1$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnevist®</td>
<td>Gd–DTPA 1</td>
<td>$3.6 \text{ mM}^{-1}\text{s}^{-1}$</td>
</tr>
<tr>
<td>Dotarem®</td>
<td>Gd–DOTA 2</td>
<td>$3.5 \text{ mM}^{-1}\text{s}^{-1}$</td>
</tr>
<tr>
<td>Omniscan®</td>
<td>Gd–DTPA-BMA 3</td>
<td>$3.8 \text{ mM}^{-1}\text{s}^{-1}$</td>
</tr>
<tr>
<td>ProHance®</td>
<td>Gd–HP-DO3A 4</td>
<td>$3.6 \text{ mM}^{-1}\text{s}^{-1}$</td>
</tr>
<tr>
<td>MultiHance®</td>
<td>Gd–BOPTA 5</td>
<td>$4.8 \text{ mM}^{-1}\text{s}^{-1}$</td>
</tr>
<tr>
<td>Gadovist®</td>
<td>Gd–DO3A-butrol 6</td>
<td>$3.7 \text{ mM}^{-1}\text{s}^{-1}$</td>
</tr>
<tr>
<td>OptiMARK®</td>
<td>Gd–EOB-DTPA 7</td>
<td>$5.30 \text{ mM}^{-1}\text{s}^{-1}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$4.6 \text{ mM}^{-1}\text{s}^{-1}$</td>
</tr>
</tbody>
</table>
Figure 1.3 The relaxivity process of excited proton nuclei during MRI scans. The two proton nuclei relaxation processes are b) $T_1$ longitudinal relaxation time and $T_2$ transverse relaxation time occur after a) an applied radiofrequency in presence of a magnetic field [7, 16][17]. b) The vector of the proton nuclei de-phase out of the lattice or surrounding along the x-y plane as a function of time. This process is $T_2$ spin-spin relaxation (purple). The proton nuclei rotate along the z - y plane in the lattice then gradually re-align with the applied magnetic field ($B_0$) along the z - axis (green). This process is known as spin-lattice relaxation. Both relaxation processes are important to understand the physics of how an MRI signal is transmitted as an enhanced image. The contrast agent acts as a catalysis to slow down the tumbling rate of the water protons once it is administered into the patient’s body. This process causes the relaxation rate to increase and a more distinct image contrast between the various types of soft tissues and organs are enhanced.
1.3 Development of the Novel Protein-Based Contrast Agents (ProCAs) for MRI

The limitation of the MRI technique is the use of a CA as previously mentioned. Specifically, the CAs provided has to fulfill certain requirements such as: low toxicity, high permeability, strong targeting capability, low dosage, and high relaxivity for efficient analysis of the image enhancement. In considering these characteristics, our lab has developed novel designed protein based contrast agents for MRI applications. Furthermore, these novel ProCAs are divided into various generations but only generation one and three are discussed in this study. ProCA1 and ProCA3 are generated from different protein scaffold. Each of the scaffold proteins have unique characteristic in terms of sustaining conformational stability prior to multiple site mutagenesis.

1.3.1 Criteria to consider prior to applying the design approach for protein-based Contrast agents

Relaxation process is dependent on many factors. The involvement of each of these factors is illustrated in Figure 1.4 and they are noted as essential criteria. The criteria to optimizing the image enhancement properties of the novel designed contrast agent involve the following: 1) high relaxation rate, 2) strong metal binding to reduce toxicity associated with nephrogenic systemic fibrosis (NSF), 3) low immunogenicity, 4) enhanced detection of tumor cells and tissues, 7) good pharmacokinetics in vivo. Water number value is an important parameter when studying relaxation measurement.

1.3.2 Model for the contrast agent: Background on the scaffold protein

CD2.D1 originates from the T-cell receptor in the immunoglobulin response system [18]. This protein maintains conformational stability at varying pH and salt concentrations, which indicates a high tolerance to mutation. Previous studies on the use of rat CD2.D1 as a model for designing calcium binding protein confirm this conformational stability towards mutations [19]. It has two tryptophan residues, one buried in the hydrophobic core at position 32 and one exposed in the hydrophilic surface
at position 7 (Figure 1.4 b). This protein consists of a β-sheet topology with an immunoglobulin
superfamily (IgSF) domain [18]. It poses no complications in terms of folding due to no cysteine residues
present in the loops.

**1.3.3 Design approach for the novel ProCA1 variants**

ProCA1 derives from a metal binding pocket inserted on the surface of the first domain of the
rat cell cluster of differentiation 2 (rCD2.D1). Figure 1.5 b summarizes the following four engineered
protein binding sites: CD2.7E15, CD2.7E15E, CD2.7E15N, and CD2.7E15Q. The following terms referring
back to the mutations will be used throughout this thesis: 7E15, 7E15E, 7E15N, and 7E15Q. These
mutations are designed by a former research colleague, Dr. Anna W. Mannicia, with the intention of
analyzing the effect of charge numbers and ligand types at the metal binding pocket on protein folding,
stability and metal binding affinity [20]. 7E15 has mutations at positions 15, 58, and 64. The designed
ProCAs are T₁-weighted MRI agents due to the high positive contrast enhancement during T₁
longitudinal proton relaxivity measurement. Thus, the engineered binding pocket has a pentagonal bi-
pyramidal coordination confirmation with Gd³⁺ (Figure 1.5 a). In addition, the net charge of the binding
site is indicated for CD2.7E15 (-5), CD2.7E15E (-5), CD2.7E15N (-4), and CD2.7E15Q (-4) (Figure 1.5 b).
The binding affinity of a metal-protein complex determines the strength of the residue side chains
interaction with the ion [21]. The original aspartic acid residue is mutated to glutamic acid, named
7E15E, following aspartate (7E15N) and glutamate (7E15Q). 7E15 and 7E15E hold a - 5 charge while
7E15N and 7E15Q have a - 4 charge. The different charge of the binding pocket is essential for
comparing the binding affinity to the metal ions during fluorescence titration and relaxation rate
measurement. The binding pocket coordination geometry is described as a pentagonal bipyramidal for
the Gd³⁺ ion complex with the five-ligand side chains. It is assumed that this coordination is sTable but
theoretically decreases in affinity depending on the metal ion atomic radii present. For instance,
lanthanide metals Ln³⁺ (1.94 Å) and Tb³⁺ (1.81 Å) have similar atomic covalent radius to Gd³⁺ (1.84 Å),
while zinc (1.2 Å) has a lower atomic covalent radius but higher electronegativity. The other physiological metals such as Ca$^{2+}$, Mg$^{2+}$, and K$^{+}$ have similar atomic radii as well in comparison to Gd$^{3+}$. These ionic radii comparisons help in obtaining an understanding on the binding affinity or dissociation constant (Kd) value obtained for each protein variants during fluorescence titration. These metal binding studies are discussed in detail in the results and discussion of Chapter 4.

We must optimize expression and purification methods to increase production yield and further continue the study with a detailed biophysical analysis (e.g. metal binding affinity, relaxivity measurement) on the purified ProCA samples. ProCAs are expressed and purified for further conformational analysis. The expression behavior of the mutants is analyzed in this study when expressed in varying E. Coli cell strains. The sole purpose is to obtain an optimized condition for expression and purification of ProCAs as efficiently as possible, so they can be used for further analysis in detecting cancer cell markers such as HER2 when injected into mice for pre-clinical trials [22].

Abada et al. demonstrated that a macro cyclic gadolinium-based contrast agent with phosphate ligands has a two-fold higher relaxivity ($r_1$: 8.5 mM$^{-1}$s$^{-1}$) compared to the CA with carboxylate ligands ($r_1$: 4.2 mM$^{-1}$s$^{-1}$) [23]. Such results give rise to questions involving whether change in ligand may have an effect on the relaxation rate properties of the contrast agents. For our initial design approach of the binding pocket, the ligands are the side chains of three aspartic acids and two glutamic acids, which make a -5 charge on the beta sheet surface (i.e. B, D, E) of what is known as CD2.7E15 or ProCA1 (Figure 1.4 b). It is also very interesting to see the effect of electrostatic interactions or different ligand residues on these properties of designed protein contrast agents. Table 1.2 lists each created variant that is related to this study and their corresponding design rationale.

**1.3.4 The Grafting approach for promotion of the ProCA1 targeting capability**

Human epidermal growth factor receptor (EGFR) and type 2 (HER2/neu, ErbB2) are defined as biomarkers commonly found and expressed on various cancerous cells, specifically breast and ovarian
cancer [24-26]. Our lab has established a targeted contrast agent by engineering an affibody protein
domain (~7.5 KDa, 3 helices) for targeting the tumor receptors (Figure 1.7 d) [22]. Affibody is utilized as a
probe for image diagnostic and therapy, and mimics antibody-targeting capability, but they are derived
from a phage display library from the Z domain of staphylococcal protein A (SPA) [27]. Two varying
mutants are analyzed in this study. HER2 targeted contrast agent called ProCA1.affiZHER342 and EGFR
targeted contrast agents called ProCA1.affiZEGFR1907 (Table 1.2). A lab colleague, Dr. Jingjuan Qiao
grafted these affibody-targeting peptides to the c-terminal of ProCA1. A five-residue linker joins the two
moieties. The GGSGG linker is chosen because it allows sufficient movement and prevention of steric
hindrance between the two moieties. These developed targeted protein contrast agents have improved
the sensitivity and specificity compared to the conventional FDA-approved/clinically-applied CAs such as
Gd-DTPA. The various in vivo studies demonstrate the successful increase in relaxivity and image
enhancement, as well as endothelial penetration and even bio distribution of the ProCAs [16, 22, 28-29].
In contrast, other derived targeted bio-based contrast agents such as antibody, micelles, nanoparticles,
iron oxide, do not demonstrate such even bio distribution capability due to their size. In Chapter 4 of
this thesis, I will report our effort to improve expression and purification of these variants and
biophysical characterization of these variants that are important for future in vivo applications.
Figure 1.4 Schematic representation of the factors that affect a contrast agent’s relaxation properties. 
a) The various water protons present in the surrounding environment interact with the gadolinium ion during administration of the gadolinium-based contrast agent. B) ProCA demonstrating a spin lattice relaxation rate at a higher range than Clinical CAs during the increased field strength (MHz). c) The various pegylation with lysine residues on ProCA1 surface promotes low immunogenicity and increased permeability and half-life in the blood vessels. Figures referenced from [29].
Figure 1.5 The alignment of the ribbon surface between human and rat derived CD2.D1.

The alignment of the ribbon cartoon of a) hCD2d1 (purple, generated from Pymol through pdb code: 1HNF) and b) rat CD2.D1 (blue, generated from Pymol through pdb code: 1HNG). Trp 7 and Trp 32 are highlighted yellow [30]. c) ProCA1 has a binding pocket designed specifically for forming a complex with gadolininium cation. d) The clustal protein sequence alignment for human/rat CD2 and CD2.7E15 indicates the conserved residues maintained from both species [31].
Figure 1.6 The binding pocket for the single site mutation of ProCA1 variants. The listed variant has it’s a) corresponding charge, residue labeled, and is generated using PyMOL software [30]. b) The Table summarizes the cartoon representation of the enlarged binding pocket region with its five residues compared to the wild type parent protein (CD2.D1).
The different factors involved when using affibody as a multimodal probe for MRI scan. The overall scheme of a) the signal transduction pathway that are mediated by different tyrosine kinase receptors. b) Specific heterodimer formation between the EGFR and HER family and the binding of the targeted drug (Ertotinib) for the down regulation of cancer cell proliferation. c) The in vivo targeting capability of cancer xenograft mice models with the HER2 binding site cartoon on alpha 1 and alpha 2 helices (generated by PyMoL, PDB code: 2KZI) [30]. The designed model of ProCA1.affi342-m pegylated with PEG40 for an improved multimodal probe [26, 32-33].
1.4 Amount of Protein Needed After Optimizing Protein Expression and Purification

The novel engineered protein-based contrast agents (ProCAs) are established for MRI usage. The production of these ProCAs is in need for a progressive analysis of these proteins in vivo. It typically takes 10 to 14 years for a newly developed drug to pass through clearance for use at the clinical phase [3]. The basic steps to taking a drug into the clinical phase involve rigorous preliminary studies. These studies refer to the large scale-up production of these engineered ProCAs. The expression and purification conditions are the main focus for the production parameters. These methods are crucial for establishing an optimal protocol that can transition into the industry for the production of the novel ProCAs. Clinical imaging is the long-term goal and it entails using ~ 0.10 mmol/kg subject weight of 0.5 M solution dosage for human MRI imaging depending on the contrast agent and the subject’s weight. Prior to transitioning into the pre-clinical then clinical phase, the ProCAs are expressed then purified based on the selected methods mentioned in Chapter 2 and Chapter 3. Table 1.2 summarizes the typical protein amount needed for each biophysical application. The main instruments that are applied in this thesis are fluorescence emission spectroscopy, circular dichroism, and relaxometer. The total protein weight is indicated at the milligram amount for the specific biophysical techniques used for the analysis of the function and structure conformation of the purified ProCAs. This indicates that a microgram to milligram amount of these ProCA1 variants are crucial for continuous structural and functional characteristics, and relaxation measurements for in vitro and in vivo applications (Chapter 4).
Table 1.2 Summary of the amount of protein needed for each biophysical analysis. The results (Chapter 4) are obtained based on these highlighted biophysical techniques (orange).

<table>
<thead>
<tr>
<th>Biophysical Technique</th>
<th>Protein needed*</th>
<th>Total Protein Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMR</td>
<td>100 - 500 µM of 500 µL (high purity 95-99 %)</td>
<td>2.80 mg (for 500 µM)</td>
</tr>
<tr>
<td>Fluorescence spectroscopy (titrations)</td>
<td>5 - 50 µM of 1000 µL</td>
<td>0.56 mg (for 50 µM)</td>
</tr>
<tr>
<td>Mass spectrometry</td>
<td>10 µM of 10 µL</td>
<td>1.12 µg (for 10 µM)</td>
</tr>
<tr>
<td>Circular Dichroism</td>
<td>2 - 20 µM of 400 µL - 1000 µL</td>
<td>0.22 mg (for 20 µM)</td>
</tr>
<tr>
<td>Relaxometer</td>
<td>100 - 500 µM of 300 µL - 1000 µL</td>
<td>5.60 mg (for 500 µM)</td>
</tr>
<tr>
<td>Animal experiment (mouse)</td>
<td>100 µL for 5 mM ProCAs per 20 g mouse High purity (99 %)</td>
<td>5.60 mg (for 5 mM ProCA)</td>
</tr>
</tbody>
</table>

*These values are based on the conventional approach procedure applied in the lab.
1.5 Overall Purpose of this Study

1.5.1 Production of these novel protein-based MRI Contrast Agents

The expression and purification conditions are altered and the results are analyzed. The result that is best suitable for progress at the large industrial scale can be devised based upon the troubleshooting of parameters during expression and purification. The optimized conditions for purification are analyzed for best bioactivity, protein stability, low toxicity and low contaminants. Basically, each conditions changed during purification are screened for better results. Specifically, the best results that are suitable for this study involve achieving low contaminants, high yield, and suitable functional ProCAs at a low cost and less time consuming manner. Comparing the tag-less and glutathione-S transferase (GST)-tag purification systems are amongst the main direction to which we can determine what purification system will satisfy for the efficient and high quality production of these novel based engineered ProCAs. This study only focuses on expression and purification of the rat CD2.D1 scaffold protein. Eventually, human-derived CD2.D1 variants will be analyzed and the established optimized conditions can be applied to those variants as future work. A standard optimal expression and purification method that follows a current good manufacturing practice (cGMP)-like procedure must be obtained as well. In addition, we must be able to transition into an industrial large-scale method for application in vivo, specifically large animal experiments (i.e. rats, dogs, monkeys).
Figure 1.8 A scheme of the overall goal in this thesis study.
1.6 Overview of this Thesis

In this thesis, the production of the previously designed and engineered variants of ProCA1 from Dr. Anna W. Maniccia and Dr. Jingjuan Qiao’s dissertation, and ProCA32 variant from Dr. Shenghui Xue’s dissertation are discussed in detail in Chapters 2-5. Specifically, the objective of Chapters 2 and 3 is to seek optimal conditions through: change in parameters during expression and purification methods for obtaining high yield, low toxicity, high purity, stable bioactivity of protein of interest through an effective, low cost procedure. This applies as well to all other ProCAs that will be mentioned throughout this study. In Table 1.3, each charged variant tackled in this study is summarized with its associated mutation, purpose for the mutation, and related Chapters.

In Chapter 2 of this thesis, expression conditions for tag-less ProCA1 charged variants are analyzed by change in host cell strain. Competent cell strains used are BL21 (DE3), BL21 (DE3) pLysS, Tuner, and Rosetta-gami (DE3) pLysS. Henceforth, these cell strains are named BL21, BL21 (DE3) pLysS, Tuner, and Rosetta-gami pLysS throughout Chapter 2 and 3. The tag-less expression method demonstrates low molecular weight at ~ 11.2 KDa for each ProCA1 charged variants, which are listed below in Table 1.2. Chapter 2 describes the effect of recombinant protein production is dependent on the used host cell strain. Each cell strain has a deletion mutation, which gives them a specified characteristic during expression system. For instance, ProCA1 charged variants expressed in BL21 (DE3) pLysS cell strain demonstrate lower production at higher molecular weight compared to BL21 and Tuner. Tuner is dependent upon the concentration of inducer added during the expression and this phenomenon is demonstrated from the expression gels and final yield post FPLC-Q column isolation (Section 2.3 of Chapter 2). Further isolation of the concentrated pure samples is passed through a size exclusion high-pressure column but the yield is decreased dramatically.

The expression and purification conditions of GST-tag charged ProCA1 charged variants and targeted ProCA1.affibody variants are tackled in Chapter 3 of this study. The following parameters for
expression system are analyzed: a) the temperature post induction, b) the inducer (IPTG) concentration, and c) the host *E.coli* cell strain. The following parameters for the GST-tag purification system are analyzed: a) 100 mg lysozyme/per 1 L culture, b) use larger diameter column, c) use 10 mM DTT and 0.05 % triton X-100, and d) use of a refolding method on inclusion body GST-tag ProCA1 charged variants and ProCA1.affibody variants. Summarized expression and purification data for GST-tag ProCA1 charged variants through change in conditions/parameters are listed under Section 3.3 of Chapter 3. The yield from each host cell strain and purification conditions for all variants charged and targeted is summarized in Section 3.3.

In addition to obtaining an optimal condition for expression and purification, the structural and functional characteristics are summarized in Chapter 4. The secondary and tertiary structural conformation stability of the ProCAs is tested through monitoring the secondary structure with circular dichroism and tertiary hydrophobic folding with tryptophan-fluorescence emission spectroscopy. The functional characteristics involve determining the metal-binding affinity of the binding pocket to the lanthanide metals Gd$^{3+}$ and Tb$^{3+}$. The metal binding study for Tb$^{3+}$ is observed through the following two fluorescence methods: a) a buffer system titration method, and b) a competition assay using FLuo-5N dye. The summarized Kd for each of the titration methods for ProCA1 variants targeted and charged are compared in a Table that is available in Section 4.3 of Chapter 4. The competition method is applied for Gd$^{3+}$ binding affinity determination with the use of a Fluo-5N Invitrogen commercial dye. Furthermore, the relaxation properties are measured with change in ProCA1 variants to Gd$^{3+}$ concentration ratio, and analysis of relaxation rate change in relation to ProCA1 variants concentration change with the fixed Gd$^{3+}$ concentration. The ProCA32 variant has a different scaffold protein and two binding pockets compared to ProCA1 variants mentioned previously.

The conventional bench work purification system of the fed batch fermentation for ProCA32 is analyzed and compared to two other systems. Specifically, the conventionally expressed ProCA32 and
fed-batch fermented ProCA32 pellets are compared through the typical bench work purification system. In addition, the ProCA32 pellet from fed-batch fermentation is purified through the use of a specified PEG concentration. Essentially, all three tactics for purification are compared in terms of overall yield after FPLC-Q column isolation. The relaxation properties at varying ratio of the ProCA32 to Gd³⁺ concentration are also compared for each sample obtained from the individual expression/purification systems. The purified ProCA32 from fed-batch fermentation is analyzed for structural and functional stability through an NTA-buffer system titration and a competition assay with FLuo5N titration using Fluorescence spectroscopy. Chapter 6 describes the goals achieved; the challenges, and the major findings from this study overall.
Table 1.3 The summary of ProCA variants used in this study [20, 29, 34-36]. All ProCAs of the rat scaffold protein CD2.D1. All ProCA1 charged variants are referred as: ProCA1, 7E15E, 7E15N, and 7E15Q for the design approach in the binding pocket. The grafting approach at the c-terminus of ProCA1 is represented based on ProCA1.affi342 targeted towards HER2/neu and ProCA1.affi1097 targeted towards EGFR cancer biomarkers. The derived ProCA32 represents the 3rd generation of designed protein based contrast agents and holds two binding pocket compared to the first generation ProCA1 that has one binding pocket design on its scaffold’s surface.

<table>
<thead>
<tr>
<th>Rat scaffold Protein variants</th>
<th>Mutation</th>
<th>Purpose for mutation</th>
<th>References for mutation</th>
<th>Related Chapters in this Thesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD2.7E15 or (7E15) or (ProCA1)</td>
<td>N15E,L58D,K64D</td>
<td>Designing a -5 charged pocket for relaxivity and metal binding study</td>
<td>(Maniccia, A. W. 2005) (White, N. 2011)</td>
<td>2 to 4</td>
</tr>
<tr>
<td>CD2.7E15E or 7E15E, (E)</td>
<td>N15E,L58D,K64E</td>
<td>Designing a -5 charged pocket for relaxivity and metal binding study</td>
<td>(Maniccia, A. W. 2005) (White, N. 2011)</td>
<td>2 to 4</td>
</tr>
<tr>
<td>CD2.7E15N or 7E15N, (N)</td>
<td>N15E,L58D,K64N</td>
<td>Designing a -4 charged pocket for relaxivity and metal binding study</td>
<td>(Maniccia, A. W. 2005) (White, N. 2011)</td>
<td>2 to 4</td>
</tr>
<tr>
<td>CD2.7E15Q or 7E15Q, (Q)</td>
<td>N15E,L58D,K64Q</td>
<td>Designing a -4 charged pocket for relaxivity and metal binding study</td>
<td>(Maniccia, A. W. 2005) (White, N. 2011)</td>
<td>2 to 4</td>
</tr>
<tr>
<td>CD2.7E15.ZHER2:342 or (ProCA1.affi.ZHER2:342) or (ProCA1.affi342)</td>
<td>L9M,Q11N,Q17A, A18L,W24G,T25Q, S27K (for affibody peptide)</td>
<td>Design for targeting of HER2/neu cancer cell biomarker</td>
<td>(Qiao, J. 2011)</td>
<td>3 to 4</td>
</tr>
<tr>
<td>ProCA32</td>
<td>S56D (on CD binding pocket),F103W</td>
<td>Design of mutated binding pocket for optimizing ionic binding of ligands to metal ion</td>
<td>(Xue, S. 2013)</td>
<td>5</td>
</tr>
</tbody>
</table>
2. EXPRESSION AND REFOLDING PURIFICATION OF TAG-LESS PROCA1 VARIANTS IN VARIOUS E. COLI CELL STRAINS

2.1 Introduction

The general factors involved in the production of our recombinant ProCA1 variants are: 1) plasmid/vectors, 2) competent host cell strain, and 3) cultured media for the expression system, 4) method for the extraction of protein, 5) collecting the soluble (supernatant)/insoluble (inclusion body) form for the desired purification system. The correlation between each of these criteria is summarized through a scheme in Figure 2.1. The chosen vector depends in part on the planned purification technique, such as pET with 6-Histidine tag (His-tag), or pGEX system (GST-tag), or pMAL for maltose binding protein tag, all involving a tag purification method for efficient protein yield. My focus in this Chapter is specifically directed at the tag-less expression and denaturing - refolding purification methods of the ProCA1 charged variants, while the fusion-tag system is discussed later (Chapter 3) as a comparison.

The parameter in the expression procedure tested for optimal cultivation results is the change in the host cell strain for tag-less ProCA1 charged variants. Specifically, 7E15E, 7E15N, and 7E15Q are expressed and purified, followed by the calculated yield (mg/L). E.coli is a commonly used bacterial host cell strain, and there are many commercially available competent cells derived from E.coli. The use of the bacterial cell strains gives a high throughput of the protein of interest (ProCAs) in a short time limit. The host E.coli cell strain is the production machinery used for expressing the protein of interest. The complex mechanism taken place inside the E.coli cell depends on several factors that indirectly play a role during the over-expression of our recombinant protein. Some of these factors involved in the expression system are: the inducer concentration, temperature set after induction, and the host cell strain. The host cell strain is the condition I draw my focus when analyzing for the best optimal
expression level of tag-less ProCA1 variants.

The four cell strains are as follows: BL21 (DE3), BL21 (DE3) pLysS, Tuner, and Rosetta-gami (DE3) pLysS are examined in the LB broth medium. Table 2.2 summarizes the specific mutation and characteristic of each of these listed cell strains. Based on Natalie White’s thesis, Tuner cell strain promotes high inclusion body yield of the protein of interest, which is compatible for the refolding purification method [34]. BL21 (DE3) is the host cell that allows for the expression of plasmids containing a T7 promoter site. The same cell strain with the added pLysS mutation activates the T7 lysozyme, which allows for low basal expression of the protein prior to induction. This particular host cell also prevents the high expression of toxic proteins to allow for successful translation of the protein to occur once cells are induced in the luria-broth media. Rosetta-gami pLysS holds a similar characteristic to BL21 (DE3) pLysS, but in addition it allows for promotion of rare tRNAs for translation of rare codons (i.e. AGG, AGA, Arg, AUA Ile, CUA Leu, CCC Pro, GGA Gly) [37]. I utilized this information and tested if there is a correlation between the expression level of the protein and the yield of the protein after the purification process. The expression conditions that give a higher ratio of the protein in the inclusion body is most suitable for the tag-less refolding purification system.

There are multiple large-scale expression and purification systems that are used for the production of potential agents to be used in medical applications [38]. For instance, Cox et al. devised a more effective three step purification method for recA genome of E.coli, which plays a vital role in DNA replication, annealing and regulation of DNA mutation/repair where the purification system only involves three steps: 1) lysing the cell in specified buffer, 2) collecting /solubilizing pellets, and 3) isolation through a series of chromatographic steps (i.e. phosphocellulose and single stranded DNA cellulose) that involve no gradient buffer elution [39]. The purification approach that I focused on is in the refolding of the designed ProCA1 variants that are cultured with a pET20b vector. Through a set of trial and error, previous lab members have established the current refolding protocol for these ProCA1
variants that are mainly found in the inclusion body after cell lysis [34].

The focus of this Chapter is on the tag-less expression and refolding purification methods. The tag-less purification method involves recovering the misfolded ProCA1s through a refolding system. The pellets are treated with harsh denaturant at a high concentration for optimal unfolding of the ProCA1s. The protein’s tertiary confirmation is recovered through a step-wise dialysis and dilution process. There are current studies that focus on similar refolding purification protocols [40-45]. The advantages and challenges when utilizing the refolding purification system are compared in Table 2.3. One drawback from applying this purification procedure is possible presence of nucleotide contaminants and aggregation during the refolding process. Although a high concentration and yield of the protein can be recovered through the denaturation-refolding system, aggregation still poses a problem when refolding the protein back to its native tertiary structure. Another challenge facing the denaturing-refolding purification method is the recovery of the protein’s native functionality as a chelating agent. Solving the drawbacks during the recovery of the protein from the inclusion body is crucial for effective production and reproducible protein yield. So we raise the question: How can we promote the increase in the production of these ProCA1s variants and move forward with the refolding purification? These challenges involved in the refolding purification system can be overcome through troubleshooting of important parameters as demonstrated by current studies. Essentially, our attention is the expression system level and how it affects the purification system in terms of protein yield (mg/L).

Theoretically, the low expression level of the recombinant protein can give low yield after purification. We want to determine which cell strain gives the highest ProCA1 charged variants yield (mg/L), lowest contamination for DNA/nucleotides, and the solubility level of each ProCA1 charged variants during cell lysis prior to the denatured-refolded purification system. Our effort is directed at analyzing the expression level of our ProCA1 variants during cultivation and then proceeds through the effective recovery of its tertiary structural conformation after the refolding purification method. The
overall work in this Chapter consists of the following: a) the small-scale expression of ProCA1 in two different cell strains, b) the large-scale expression of 7E15E, 7E15N, and 7E15Q in varying *E.coli* cell strains, c) the refolding purification method applied to the large scale expressed variants, and d) The UV-Visible absorbance spectrum and final yield in milligrams from each cell strain expression system.

Figure 2.1 The general factors to consider for the expression and purification systems[46-47].
Table 2.1 A summary of the properties for the four *E.Coli* cell strains [37, 48-51]

<table>
<thead>
<tr>
<th><em>E.Coli</em> cell strain</th>
<th>Mutant</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21 (DE3)</td>
<td>λDE3</td>
<td>A mutant derived from BL21, it activates the T7 promoter site during induction.</td>
</tr>
<tr>
<td>BL21 (DE3) pLysS</td>
<td>pLysS</td>
<td>A mutant to BL21 (DE3), the pLysS mutation allows release of T7-lysozyme which causes low expression of unwanted proteins prior to inducing.</td>
</tr>
<tr>
<td>Rosetta-gami pLysS</td>
<td><em>trxB/gor, pLysS</em></td>
<td>A mutant derived from Origami, it promotes low basal expression level through the release of T7-lysozyme, allows for disulfide formation, and universal translation of proteins with rare codons by generating rare tRNAs.</td>
</tr>
<tr>
<td>Tuner</td>
<td><em>LacZY deletion, Lac permease (lacY)</em></td>
<td>A mutant to BL21, it promotes easy permeability of the inducer in the cell. It depend on the IPTG concentration.</td>
</tr>
</tbody>
</table>

Table 2.2 Comparing the advantages and challenges for the refolding purification method [40-43, 52-53].

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>• High yield of recovered protein for</td>
<td>• Dilution of protein solution</td>
</tr>
<tr>
<td>biophysical studies such as NMR, Relaxivity</td>
<td>• Prone for high level of contaminants (i.e.</td>
</tr>
<tr>
<td>measurement</td>
<td>lipids, DNA/RNA fragments)</td>
</tr>
<tr>
<td>• Good for proteolytic stability, protein of</td>
<td>• Time consuming ( takes 3 days to a week for</td>
</tr>
<tr>
<td>interest is in the inclusion body</td>
<td>required refolding)</td>
</tr>
<tr>
<td>• Low cost effective</td>
<td>• Use of high concentrated denaturants (8M urea, 6 M</td>
</tr>
<tr>
<td></td>
<td>GdmCl) for solubility, Aggregation during refolding</td>
</tr>
<tr>
<td></td>
<td>• Change in the stability of the protein of interest</td>
</tr>
<tr>
<td></td>
<td>during refolding.</td>
</tr>
</tbody>
</table>
2.2 Methods

2.2.1 Cloning

The charged ProCA1 variants (from Chapter 2 to 4) cloning are described according to the protocol in Dr. Anna Wilkin’s dissertation [20]. The targeted ProCA1 variants to HER2 peptide (from Chapter 3 to 4) cloning are described according to the protocol in Dr. Jingjuan Qiao’s dissertation [36].

2.2.2 Transformation

To begin this method, the following E. Coli cell strains: (BL21 (DE3), Rosetta-gami (DE3) (DE3) pLysS, BL21 (DE3) pLysS, and/or Tuner, were allowed to thaw in an ice bucket. The CD2.7E15 (or ProCA1) DNA plasmid with a pET-20b (for Chapter 2) or pGEX-2T (for Chapter 3) vector was also thawed in the ice bucket. The competent cell (50 µL) was then mixed with the DNA plasmid (0.50 µL) in an empty autoclaved eppendorf tube, and then left on ice for thirty minutes. In following, a water bath at 42 ºC was controlled. The eppendorf tube containing the mixture was transferred into the water bath for 90 seconds, and then quickly transferred back on ice for two minutes. This crucial heat shock method possessed an impact on the cell strain’s membrane wall. The cell membrane was vulnerable in this step and as a result the foreign DNA plasmid that holds the recombinant DNA sequence of the protein of interest was present in the cytoplasm of the competent host cell. Prepared autoclaved LB Broth Miller medium (75 µL) was added to the cell/DNA mixture and placed in the Isotemp incubator (Fisher Scientific) for 30 minutes at 37 ºC. After obtaining the incubated tube, the mixed solution (50 µL) was placed in the center of the luria-broth agar with ampicillin antibiotic plate and spread with the use of a sterile triangle. The plate was then placed for incubation at 37 ºC for overnight. The next morning the plate was observed for any colony growth and placed in the 4 ºC refrigerator with parafilm. Transformation methods is summarized as a scheme in Figure 2.2 for all expressed 7E15 and charged variants. Note: all proteins expressed in this Chapter 2 involved the Tag-less vector pET-20b.
2.2.3 Small and large-scale expression

Inoculation was the first step in the expression methods for CD2.7E15 wild type. This step was quick and the flask (1000mL capacity) was placed in the shaker (Branstead Lab-line, ~200 rpm, temperature 37 ºC, Max Q 5000), which was used to spread the nutrients and oxygen within the available space in the flask for overnight (o/n). The second step involved the transfer of the overnight culture into the two larger one Liter Luria-Bertani media flasks (2800 mL capacity) for the large-scale system while a 10 mL volume of fresh media is used in sterile plastic falcon tubes (50 mL capacity) for the small-scale system. The O.D. readings were recorded with the use of a UV-Vis spectrophotometer (UV-1601, Shimadzu Scientific Instruments Inc., Columbia, MD) at a fixed wavelength (λ = 600 nm). The bacteria cell pellets were collected on the same day after induction for four hours through the use of a low speed centrifuge (Sorvall Model RC5Bplus, MFD by Kendro Laboratory products, Newton, CT). The collected cell pellets are then placed in the - 20 ºC freezer until further application for the purification process. All charged variant proteins are expressed in large one liter flasks left overnight at 30 ºC in following the collection of the proteins the next morning. Expression of charged variants in BL21 (DE3) involved inoculation o/n at 37 ºC in sterile plastic based falcon tubes (50 mL capacity).

2.2.4 Tag-less refolding method

Re-suspending the cell pellet from the expression methods in a prepared lysis buffer was the initial step to the denaturing-refolding purification methods for 7E15 and it's charged variants. The sonication step soon followed and it was repeated five to six times for 30 seconds with a five-minute interval with cooling on ice. The bacteria cells were treated with a constant vibrational force, which lyses the cell wall and exposed the protein of interest to an unstable environment. Thus, it was important to keep the solution under ice as the sample overheats during the protein extraction process from the host cell. The solution is then centrifuged at high speed (17 Krpm, Sorvall RC-5B Refrigerated Superspeed Centrifuge, Du Pont Instrument, Newton, CT) for 30 minutes. The cell pellets are washed with 2 % Triton
In following the protein in the pellet are denatured by prepared 8 M urea (OMnipur min 99.5 %, EMD Chemicals Inc. Gibbstown, NJ) o/n. The next morning the protein is centrifuged and supernatant (SN) is transferred into dialysis bags. The protein refolding process is executed with treatment of 2 M Urea then 10 mM Tris buffer placed inside a 2000 mL plastic beaker for overnight. The solutions are collected and stored in sterile falcon tubes at -20 °C until the next purification process (Fig 2.3).

**2.2.5 FPLC- anion exchange Q – column**

Fast Protein Liquid Chromatography (FPLC, General Electric, #INV-907, made in Sweden) is a chromatographic process in which the protein in solution is exposed to a salt gradient (Q-column) program (Figure 2.5). The program system was initially washed with ultra pure deionized water (ddH$_2$O). The 10 mM HEPES (Free acid, molecular biology grade, 99.7% purity, EMD Biosciences, Inc. LaJolla, CA) buffer (pH ≈ 8.0) and 10 mM HEPES, 1M NaCl (Crystals, GR ACS, EMD Chemicals Inc. Gibbstown, NJ) buffer (pH ≈ 8.1) are prepared and used as Buffer A and Buffer B, respectively. The column is gradually washed with the salt buffer so protein is eluted out of the column as the percent salt concentration increases gradually as demonstrated in the results (Figure 2.9 and 2.10).

**2.2.6 Quantitative analysis of the 15 % SDS-PAGE through Image J software**

Since the majority of the expression and purification results were mainly qualitative, we utilized Image J software (1.47v, Rasband, Wayne, NIH, USA) to quantitatively measure the protein’s band intensity. Essentially, the intensity of the band was measured and this value was used for the analysis of the small and large-scale expression results. The gel results for the expression (2.3.1 and 3.3.1) and purification (3.3.2) were quantified through a normalized intensity ratio value. The bands that demonstrated the low intensity ratio represented a low expression in correspondence to the dark pixels of the imaged SDS-PAGE.
2.2.7 Ultraviolet-visible absorbance spectra and agarose gel electrophoresis applications

Analytically, the purified protein solutions were analyzed for the contents of DNA contamination through a UV-VIS spectrum generated on the computer (UV-probe version 2, Pharma spec. 1700, Shimadzu scientific Instrumentation Inc. Newton, CT). In addition, agarose gel was used to indicate any presence of DNA in the protein solution (6 µL of SyBr indicator, 10 µL of HiLo marker, 10 µL prepared supercoiled DNA marker, at 80 volts). Quantitatively, the UV-VIS absorbances were recorded for each concentrated protein solution, in order to determine the final protein concentration \( C_F \) experimentally. Nanopure water (ddH\(_2\)O) (800 µL) was added to the quartz cells. The baseline was run initially prior to adding the protein sample in the cell. In following, auto zero was used on the absorbance value prior to adding a desired amount of the concentrated protein into the cell in the sample holder position. Shortly after the display of the spectrum, the absorbance maximum value at 280 nm was multiplied by the dilution factor in order to obtain \( C_F \) with Equation 2-1. The theoretical molar absorptivity \( (13940 \, \text{M}^{-1}\text{cm}^{-1}) \) was determined through a protein web calculator database while the experimental molar absorptivity was determined in our lab \( 11700 \, \text{M}^{-1}\text{cm}^{-1} \) [54]. The experimental molar absorptivity value was used in the Beer Lambert Equation for obtaining \( C_F \).

\[
A = \epsilon \ell \, C_F \quad \text{Equation 2-1}
\]

Note: All samples collected during expression and purification methods were run through 15% SDS-PAGE and recorded as data.
Figure 2.2 The transformation scheme along side the pET-20b vector containing the T7 promoter site.

IPTG is an inducer that activates the T7 promoter site located in a) pET20b (+) vector [55], which promotes the overexpression of the recombinant DNA fragment of ProCA1s during the induction process. b) The initial transformation of the recombinant DNA of ProCA1s is inserted through a heat shock method into the desired host cell strain.
Figure 2.3 The expression scheme represents the stepwise method used for the tag-less ProCA1s’ overexpression in the four different *E.coli* competent cell strains.
Figure 2.4 The denaturing and refolding steps are described in a detailed step-by-step method for the tag-less purification of each ProCA1s. Four major steps highlight the procedure for the established denaturing-refolding purification of the tag-less ProCA1 charged variants. Abbreviations: cell pellet (CP), supernatant (SN), overnight (o/n).
Figure 2.5 The FPLC-Q column preparation scheme and programmed salt gradient. The column is pre-washed with buffer A to remove any non-specific binding from previous usage. The protein of interest, ProCA1s, presumably elute between 30% and 70% NaCl in two-milliliter fraction volumes.
2.3 Results and Discussion

2.3.1 Host cell strain influence on the expression of tag-less charged ProCA1 variants

Initially, a small-scale expression system measures the efficiency of three clones grown on the agar plate for each cell strains. The cell strains BL21 (DE3), and Tuner are examples of the small-scale expression system for ProCA1. All clones for BL21 (DE3) and Tuner demonstrate similar expression intensity (Figure 2.6). The cell growth increases exponentially with time (minutes) for variants expressed in BL21 (DE3) and Tuner in the small-scale expression system.

In the case of the large-scale expression system, it is relatively the same for all ProCA1 variants aside from the host cell strain. The four E.coli cell strains have different necessities and demands during the expression process. Thus, as a result different levels of expression are demonstrated in Figure 2.7. Specifically, 7E15E and 7E15N demonstrate an exponential decay in Rosetta-gami (DE3) pLysS post-induction point range, which is between 500 - 1500 minutes. Similarly, 7E15N and 7E15Q show a decrease in optical density after the induction point but 7E15N decreases at a larger fold than 7E15Q during cultivation in BL21 (DE3) pLysS. The gradual increase in O.D. readings for the charged variants in BL21 (DE3) gradually indicates a bacteria cell growth throughout the expression process. The low optical density value between the one-hour and the overnight induction conditions affect 7E15E and 7E15N expressed in Rosetta-gami (DE3) pLysS, suggested a decrease in cell growth. The growth curve and the 15% SDS-PAGE expression gels (Figure 2.7 a and b) support this occurrence. It was expected that an increase in band thickness and intensity would be present for expression of the proteins after induction. The results show otherwise for 7E15E and 7E15N, in Rosetta-gami (DE3) pLysS; and 7E15N and 7E15Q, in BL21 (DE3) pLysS; after the induction process.

Expression of an unknown protein is indicated between 68 and 26 KDa markers from the one liter 7E15Q post-induction sample that is cultivated in Rosetta-gami (DE3) pLysS (Figure 2.7 a). Thin
bands of this unknown protein are also present at the same KDa position for 7E15E and 7E15N post-induction samples that are cultured in Rosetta-gami (DE3) pLysS. The expression process is low for the charged variants in BL21 (DE3) cell strain, while low intensity bands are displayed for the proteins in Rosetta-gami (DE3) pLysS. Little bacteria cell growth is indicated for proteins expressed in Rosetta-gami (DE3) pLysS. Some experimental factors are considered for the influence on the low cultivation level in Rosetta-gami (DE3) pLysS and BL21 (DE3) pLysS. In addition, the expression level for both E.coli strains contains a T7 lysozyme; promoted from the pLysS marker, which hinders the excessive cell proliferation prior to the induction step. This reasoning supports the exponential decay for some of the charged variants expressed in BL21 (DE3) pLysS and Rosetta-gami (DE3) pLysS.

Figure 2.6 The growth curve for the small-scale expression of ProCA1 in a) BL21 (DE3), Tuner, and BL21 (DE3) pLysS (in 10 mL LB medium volume) b) with the corresponding 15% SDS-PAGE, and c) In following, the low expression intensity for all three cell strains.
Figure 2.7 The large-scale expression of ProCA1 variants and the corresponding growth curve for each cell strain.

a) The successful and reproducible expression level is indicated for all ProCA1 variants growth in the Tuner cell strain. b) Only 7E15E and 7E15Q are overexpressed at an optical density ($2 \rightarrow 2.5$) value compared to 7E15N (~1 O.D.) during the growth in BL21 (DE3) pLysS. The overexpression of ProCA1 variants is suitable in BL21 (DE3) while unsuccessful expression level is indicated in Rosetta-gami (DE3) pLysS with the corresponding intensity ratio. The protein of interest is highlighted in the respective colored box (red for Tuner, orange for BL21 (DE3), and yellow for BL21 (DE3) pLysS) at a ~11.2 KDa molecular weight. The expression condition for all 3 host cell strains is the same (0.20 mM IPTG, at 25 °C). ProCA1 charged variants have highest expression intensity ratio (1.00 for 7E15E, 0.84 for 7E15N, 1.33 for 7E15Q) in Tuner cell strain compared to the other expression ratio levels quantified based on the pixel density of the protein band at ~11.2 KDa position.
2.3.2 The denatured-refolding purification of tag-less charged ProCA1 variants

Each liter of expressed ProCA1 charged variants is treated through a set of denaturation and refolding steps for purification from the inclusion body. The summarized gel samples during the denaturing-refolding steps are demonstrated (Figure 2.8) for all charged variants expressed in BL21 (DE3), BL21 (DE3) pLysS, and Tuner. The cell pellets cultured (Figure 2.8) are lysed and the band intensity of the protein of interest is visually compared in the supernatant (SN) and pellet (P). In BL21 (DE3), there is a large intensity for all charged variants after the final refolding step in 10 mM HEPES dialysis buffer. In addition, little ProCA1 charged variants are recovered after cell lysing since there was low expression intensity in BL21 (DE3) pLysS and Rosetta-gami pLysS (Figure 2.8 b, c). In BL21 (DE3) pLysS, there is a large contamination range at the high molecular weight but it is separated from the protein of interest after the denaturing step in 8 M urea for overnight (Figure 2.8 b). Only 7E15E is successfully purified from this host cell strain, where 7E15E (14.13 mg) has a higher yield than 7E15Q (0.16 mg) (Figure 2.9 b). The ProCA1s ratio from supernatant to inclusion body after cell lysis varies for each ProCA1 charged variants. 7E15E demonstrates a favorable intensity ratio in the pellet form while 7E15N and 7E15Q demonstrate a favorable ratio in the supernatant form. These results suggest that the expression conditions promoted the solubility for 7E15N and 7E15Q that are both -4 charged, compared to insoluble 7E15E with a -5 charge. A second trial with the second liter demonstrates a similar phenomenon but only for high solubility enhancement in 7E15Q, while 7E15E and 7E15N were in the high insoluble ratio. These results rule out that the charged pocket has an effect on the conditions set during expression in Tuner. Since there is a predominant ratio of ProCA1s in the supernatant and pellet after cell lysis, both supernatant and refolded inclusion body are passed through the FPLC-Q column for further isolation from contaminants (Figure 2.10).
2.3.3 The refolded tag-less ProCA1 charged variants isolated by the FPLC Q-column anion exchanger

The chromatographs (Figures 2.9 - 2.10) display the isolation of the charged variants through the anion exchange column. 15 % SDS-PAGE confirms protein presence in peak two for all charged variants. In addition, the UV-VIS spectra data are consistent with these results. The 15% SDS-PAGE of the concentrated protein samples indicates successful protein isolation for 7E15E in BL21 (DE3) as an example of protein isolation post refolding (Figure 2.9). Both dialysis samples and supernatant sonication samples are run through the FPLC Q-column for flask one charged variants expressed in Tuner. In the FPLC process for charged variants expressed in Tuner, the ProCA1 variants were isolated for flask one. 7E15E protein is present in the isolated peak two (Figure 2.10 a). The 7E15N protein is eluted in fractions of peak one and the same result is obtained for 7E15Q. These results from the chromatogram and 15 % SDS-PAGE indicate an unsuccessful separation for 7E15N and 7E15Q in the SN sonication sample. Sonication samples have DNA and other biological contaminants that may cause the protein of interest to not elute with an increase in salt concentration throughout the FPLC application. A high protein yield from two liters for each charged variants; 7E15E (19.30 mg/1L), 7E15N (61.48 mg/1L), and 7E15Q (72.20 mg/1L), are displayed (Figure 2.11 c) for expression in Tuner.

Other components are present in the protein sample after concentration of peak two for 7E15E and 7E15N expressed in BL21 (DE3). 7E15Q shows the lowest DNA presence based upon the absorbance max ratio at 260 nm / 280 nm (Figure 2.11 a). For further purification to remove unwanted proteins from the concentrated samples of 7E15E, we used a size-exclusion or gel filtration chromatography. As shown in Figure 2.12, our protein of interest with a ~11.2 KDa was eluted out in fractions contained in peak three. The percent intensity of the gel band is lower compared to the initial concentrated sample (Figure 2.11 a). Tables in Figure 2.11 summarize the yield of the different variants with BL21 (DE3) pLysS and Rosetta-gami (DE3) (DE3) pLysS, E. Coli cell strains. A low yield for CD2.7E15E (14.18 mg),
CD2.7E15N (17.04 mg), and CD2.7E15Q (12.56 mg) were calculated (Figure 2.11 a) for expression in BL21 (DE3).

Figure 2.8 The 15 % SDS-PAGE for all ProCA1 variants are demonstrated for the tag-less refolding method. All protein variants are successfully recovered from the inclusion body post refolding method but all display impurities at the high molecular weight region for a) BL21 (DE3) condition; while Rosetta-gami (DE3) (DE3) pLysS confirms that no protein was produced during overexpression; b) only 7E15E is recovered at high yield compared to 7E15Q and 7E15N from overexpression in the BL 21 (DE3) pLysS condition; c) the over expression of 7E15Q and 7E15N are both in the supernatant (SN) post sonication while 7E15E is mainly in the pellet (P). This leads to low recovery of 7E15N and 7E15Q from the CP while 7E15E is recovered successfully in the SN.
Figure 2.9 The 7E15E variant is isolated through a FPLC system using an anion exchanger column (Q-column). 7E15E sample expressed from BL21 (DE3) demonstrates a high mAU in peak indicated by the arrow. The sample demonstrates nucleotide contamination successfully isolated after the program gradient reaches 100 % salt concentration.
Figure 2.10 Ion exchange FPLC Chromatograms of dialysis and sonication samples for the following proteins: a) 7E15E flask 1
Conditions are as follow: HiTrap 5 mL Q-column with cation beads, Pressure: 0.3 to 0.5 MPa, Flow rate: 3 mL/min Detector: UV lamp 280 nm shows the presence of protein in peak 2 (for 7E15E dialysis sample)(confirmed by SDS-PAGE gels)
Figure 2.11 The overall yield for each variant is summarized in correspondence to the cell strain condition.  
a) BL21 (DE3) cell strain gives successful yield for all three variants; b) BL21 (DE3) pLysS cell strain gives a successful yield for 7E15E only while 7E15Q has a relatively low yield; c) Tuner cell strain gives a relatively high yield for all three variants compared to the other three cell strains mentioned for the expression system.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Cell Strain</th>
<th>Yield (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD2.7E15E [BL21 (DE3)]</td>
<td>14.18</td>
<td></td>
</tr>
<tr>
<td>CD2.7E15N [BL21 (DE3)]</td>
<td>17.04</td>
<td></td>
</tr>
<tr>
<td>CD2.7E15Q [BL21 (DE3)]</td>
<td>12.56</td>
<td></td>
</tr>
<tr>
<td>CD2.7E15E [BL21 (DE3) pLysS]</td>
<td>14.13</td>
<td></td>
</tr>
<tr>
<td>CD2.7E15N [BL21 (DE3) pLysS]</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>CD2.7E15Q [BL21 (DE3) pLysS]</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>CD2.7E15E [Tuner]</td>
<td>19.30</td>
<td></td>
</tr>
<tr>
<td>CD2.7E15N [Tuner]</td>
<td>61.48</td>
<td></td>
</tr>
<tr>
<td>CD2.7E15Q [Tuner]</td>
<td>72.20</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.12 Size-exclusion chromatography used for isolation of impurities from the concentrated samples post refolding purification method.
2.4 Conclusion and Future Work

The overall work in this Chapter are as follow: a) the small-scale expression of ProCA1 in two different cell strains, b) the large scale expression of 7E15E, 7E15N, and 7E15Q in varying *E.coli* cell strains, c) the refolding purification method applied to the large scale expressed variants, and d) The UV-Visible absorbance spectrum and final yield in milligrams from each cell strain expression system.

Rosetta-gami (DE3) pLysS cell strain has the lowest yield in the protein expression system and consequently no reasonable amounts of ProCA1 variants are obtained after cell lysing for the purification. Loss of protein in the purification process that results in low yield is due to discarded sonication samples during the purification method. Thus, as an optimizing condition, all supernatant samples and cell pellet samples are conserved during the purification method for expressed variants in Tuner. All waste should be conserved until the confirmation of protein presence by SDS-PAGE and UV-VIS absorbance spectrum as a keynote. Little is known about the ProCA1s expression mechanism after cell lysis. Surprisingly, Tuner promotes solubility in the cytoplasm at a high protein band intensity after cell lysing so further optimization in Tuner cell strains for ProCA1s variants should be observed in terms of change in IPTG concentration. Since Tuner promotes high expression in proportion to the increase IPTG concentration, varying the IPTG concentrations during expression at a small scale is beneficial for future analysis.

Protein expression and purification in Tuner indicates highest yield for all ProCA1s. Conclusions based on expression and purification results seem to convey that BL21 (DE3) and Tuner induce expression of the proteins of interest along with another protein of a higher molecular weight. The nutrient rich LB medium may cause the co-expression of this unknown contaminant, thus other media systems are considered as future work. BL21 (DE3) pLysS is the only cell strain that promotes low expression of this high molecular weight contaminant. Further analysis is suggested on the expression results of the charged variants by varying the temperature, medium system, and IPTG concentration for
the expression system at a small-scale. Another consideration for the parameters in the purification system is: changing the denaturant and using a lower concentration of that particular denaturant. In addition, the determination of the folding ability of each variant in relation to the change in denaturant concentration may be useful for continued optimization in the refolding method.
3. OPTIMIZATION OF EXPRESSED AND PURIFIED GST-TAGGED PROCA1 VARIANTS

3.1 Introduction

Glutathione S-transferase (GST) fusion system is widely used for large-scale protein production [56]. This fusion system is chosen due to several advantages. First, it provides a purification tag for generated protein. Second, the GST-tag at the N-terminal can be cleaved by thrombin during purification. Third, GST helps protein fold correctly. We’ve taken the developed recombinant DNA sequence of engineered ProCA1 and cloned it into a pGEX-2T vector containing a glutathione S-transferase (GST)-tag site. GST is a 26 KDa enzyme with a binding site that catalyzes the reduced glutathione (GSH) (Figure 3.1) [57]. The monomer form of E.Coli GST is composed of two domains: 1) thioredoxin-like domain with the GSH binding site (G site) and 2) the α-domain with the hydrophobic binding site (H-site) (Figure 3.1 b) [58]. A known residue determined by Wang et al. have shown that histidine at position 106 (His\textsuperscript{106}) contributes to the catalytic activity in the binding site while cysteine at position 10 predominantly contributes to the redox reaction of thiol/disulfide bond formation [58]. Thrombin is a ~ 32 KDa enzyme that recognizes a specific peptide sequence (e.g. Leu-Val-Pro-Arg-Gly-Ser), and the cleavage site has been modified for further improvement of the protein yield [59]. Thrombin is preferred due to its effective cleavage of the GST-tag compared to other known cleaving enzymes such as PreScission protease and Factor Xa. Applying the GST moiety to the N-terminus of ProCA1 variants promotes an equivalent co-expression of ProCA1 variants. Essentially, ProCA1 tagged to the GST moiety will subsequently co-express inside the host cell during cultivation. One study worked with GST-tagged Interferon proteins, which naturally misfold as inclusion body in vivo; they were successfully purified at 100-mg/L yield based on a set optimal production and purification method using the GST-tag fusion system [59]. In comparison, ProCA1 holds a sTable structural conformation during expression at specified conditions. One challenge during expression is the co-expression of free GST present in the cultivation medium causes low (< 5 mg/L) yield after purification.
In this Chapter, we aimed to overcome these challenges by analyzing different conditions for establishing an optimal production and purification procedure that can generate a high yield output of our final products ProCA1 variants using the GST-tag fusion system. The first part of this study is to screen the expression conditions as a process to determine the optimal GST-tagged cultivation of charged and targeted ProCA1 variants (Figure 3.2 a). The second part of this study is to find the optimal lysis buffer during purification for promoting a higher yield of the cleaved or untagged ProCA1 variants (Figure 3.2 b).

Sections 3.3.1.1 and 3.3.1.2 discuss the troubleshooting analysis of the small and large-scale expression of GST-tagged ProCA1 variants. The small-scale system focuses on increasing the expression level of GST-tagged ProCA1, 7E15E, and 7E15N, ProCA1.ZHER2: 342, and ProCA1.ZEGFR by changing host cell strains. An advantage to the small-scale expression system is that it utilizes limited materials for data collecting and results are concluded effectively since three different colonies can be chosen for the expression of each charged variant. Later in this Section, the large-scale expression discusses the impacts of the change in temperature, IPTG concentration, and/or host cell strains have on the expression for CD2.WT, ProCA1, 7E15E, 7E15N, 7E15Q, ProCA1.ZHER2: 342, and ProCA1.ZEGFR. ProCA1.ZHER2: 342 and ProCA1.ZEGFR are designed ProCA1 variants with a targeting moiety that binds to HER2/neu and EGFR, respectively, on the surface of cancer cells. This novel designed targeted ProCA1 variant is analyzed for production through various expression conditions such as cell strain, post induction temperature change and inducer concentration. The cell strains involved in this analysis are listed as following: BL21 (DE3), BL21 (DE3) pLysS, and Tuner. The results for all ProCA1 variants production are then summarized in a Table for ease of interpreting the data.

The GST-tag purification method is executed after determining the optimal condition in the expression system for both charged and targeted ProCA1 variants (Section 3.3.2.1). The biggest challenge for the purification of these GST tagged proteins is the partial folding of HST tag in the cell
lysate which may hinder the effective binding of GST to GSH substrate. In no particular preference to the expression intensity level, the cultivated pellets are suspended in the lysis buffer with optimized conditions (Figure 3.2 b). Buffer A, which has 1 % sarcosyl cationic detergent, represents the conventional lysis conditions currently in use in the lab. As an effort to increase the solubility, binding, and oxidation state of the enzymatic pocket on the GST moiety, we modified buffer A by adding 0.05% triton x-100, a nonionic detergent, increasing the EDTA concentration, and adding DTT reducing reagent. This modification is named cell lysis buffer B. Studies have shown that sarcosyl/triton x-100 detergent ratio of 1:2 benefits in efficient GST catalytic activity and binding to the immobilized GSH substrate in an affinity chromatography column [60-61]. In addition, the chelating factor of EDTA also influence in the solubilization of the protein fusion by promoting peptidoglycan cleavage of the outer cell membrane wall [62]. This reaction weakens the bonds to hold the cell wall intact and makes it easier for protein to excrete out of the host cell during the lyses process. The mechanical and biological methods for cell lysis are considered. The protein of interest is cultured in the E.Coli host cell and it is crucial that a particular method and time duration for cell lysis are effectively applied for favorable GST-fusion protein solubility into the supernatant. Sonicator, french press, and cell disruptor are the three commonly used instruments in the lab for the protein extraction from the host cell. These mechanical methods have a drawback, which is producing heat to the protein sample. This effect may cause partial unfolding of the protein and then aggregates as inclusion body. Thus, the longer the cell is disrupted under harsh conditions, the more likely that the protein of interest develops into inclusion body. The goal of this study is to identify the best condition in the cell lysis buffer that is most beneficial for protein extraction into the supernatant environment and that promotes effective enzymatic activity of the GST-tag moiety for an increased yield production after Glutathione Sepharose-4 Beads column purification (Figure 3.2 b).
Both supernatant and refolded inclusion body GST-tagged proteins are applied to the GS-4B column during the purification for achieving an increased yield. Due to a large amount of the fusion protein present as inclusion body after cell lysis, one focus is directed at decreasing the percent of GST-tagged ProCA1 present as inclusion body. Völkel et al successfully recovered inclusion body into folded, soluble, and bioactive protein with full-length by using a denaturing-refolding method with a His-tag then followed by a GST-tag purification system [63]. Essentially, I applied a similar procedure to recover the initial GST-tagged proteins from the insoluble inclusion body form.
Figure 3.1 The schematic representation of the GST-tag ProCA1 and ProCA1.afl342. The N-terminal fusion is a) the typical tag design used for GST-fusion partners. b) Cartoons of all three moieties displayed in this scheme are: GST bound to GSH, CD2.7E15 free form (ProCA1), and ZHER2: 342-affibody peptide targeted towards HER2/neu biomarkers. The following are the pdb codes used to generate the cartoon for GST-GSH (1AQW), CD2.7E15 (1hng), and affibody ZHER2: 342 (2KZI) through PyMOL[30].
The specific parameters for optimizing the expression yield of GST-tagged ProCA1 variants. The selected three parameters are essential for the expression yield of the GST-tagged ProCA1 variants. The inducer concentration refers to the IPTG amount (mM) added to each one liter LB medium with treatment of 0.10 mM ampicillin. The temperature adjustment during and after the adding of IPTG is important for the optical density of the media, which is directly related with the protein yield. There are various E. Coli host cell strains (Table 2.2) with specific mutations that enable the synthesis of the recombinant DNA sequence of our ProCAs within the cytoplasm. The growth of the cell strains depends on the temperature environment, specified IPTG concentration, and the characteristic of the protein of interest that is synthesized in the host cell strain.
Figure 3.3 The outlined steps for optimizing the parameters in the purification system of the GST-tagged ProCA1 variants. Both ProCA1 charged / targeted variants are tested in this optimizing scheme that aims to increase the yield of each protein of interest. Specifically, modification of lysis buffer A conditions is important to the protein yield after GS-4B column purification. DTT, Triton X-100, and increased EDTA are added as a modification to lysis buffer A and relabeled as lysis buffer B. Lysozyme is then added to test whether the less harsh lysing process of the cell will generate less aggregation of the ProCA1 as inclusion body and this modification is labeled lysis buffer C.
3.2 Methods

3.2.1 Transformation

The following E. Coli cell strains, (BL21 (DE3), BL21 (DE3) pLysS, and/or Tuner, were allowed to thaw in an ice along with the CD2.7E15 (or ProCA1) variant DNA plasmids/ pGEX-2T vector. This was proceeded the same for all other five ProCA1 variants, both charged and targeted. Further details of this procedure are available in Chapter 2 experimental Section 2.2.2. Note: all proteins expressed in this Chapter 3 use the GST-tag vector pGEX-2T (Figure 3.3).

3.2.2 Small-scale and large-scale expression of GST-tag ProCA1 variants

The small- scale expression was conducted in small falcon tubes (50 mL capacity and 15 mL capacity) rather than glass flask. The expressed samples are collected and 15% SDS-PAGE was used for analysis of purity. Refer to the experimental Section 2.2.3 of Chapter 2 for the small and large scale expression details.

3.2.3 GST-tag gravity column purification

Cell pellets of all variants were collected after overexpression and thawed on ice at room temperature for an hour. The cell pellet was then suspended in a lysis detergent (20 mL) along with 2 mM Mg$^{2+}$ (100 μL) and benzonase nuclease (1 μL) for DNA cleavage, named lysis buffer A. The lysis buffer with the protein was constantly mixed in a falcon tube with a vortex. Once the solution was homogeneous then the solution was applied through either a cell disrupter (3X) or French press (3X) until the solution was clear and less viscous in appearance. The solution was placed in centrifuged tubes for sedimentation for 30 minutes at 17K rpm speed. The supernatant and cell pellet were taken for SDS-PAGE analysis. The supernatant was filtered with a 0.45 μm membrane filter and then passed through a gravity affinity column with glutathione-coated sepharose beads (~5 mL). The supernatant was passed through the column twice then rinsed with phosphate buffered saline (1X PBS) until the beads turned
clear white. After the binding step, 200 μL of stock thrombin (1 mg/mL) was added to the column suspended in 1X PBS. The final thrombin concentration is 0.04 mg/mL. Then, the column was shaked at 4 °C for overnight cleavage. The next day, the column was placed on the shaker at room temperature for another hour. The protein solution was then eluted out of the column with consecutive shaking in PBS buffer in following a second elution. The eluted samples were collected for SDS and UV-Vis absorbance were measured at 280 nm.

For a similar protocol which involves in the use of lysozyme labeled as lysis buffer C, the procedures are as following: First, the cell pellet was thawed with lysate buffer C by adding 1-1.5 mM MgCl$_2$, 5 mM DTT, 1% N-lauryl Sarcosine and one Tablet of protease inhibitor per 100 mL lysis buffer volume. Second, a magnetic stir bar was used to suspend the cell pellet in the lysis solution instead of vortex. Third, lysozyme was added at a ratio depending on volume of cell pellet used (300 mg per 3L culture medium) so 100 mg of cell pellet from each liter of medium. Fourth, lysate solution of cell pellets with the lysozyme was left on ice for 1 to 2 hours. Sonication for 1 minute may be used if needed. Finally, the sample was centrifuged at high speed and supernatant was collected and filtered. The sample was (clear and not viscous in appearance after this step. The volume started with 30 mL and doubles to 70-80 mL after all the above steps. Then the supernatant sample was passed through the GS-4B column. The total volume of sample is 6 to 10 times of the column volume because of the column capacity and efficacy.

### 3.2.4 Refolding of GST-tagged ProCA1 variants

The over-expressed cells containing the GST-tag ProCA1 variants were lysed by mechanical methods (i.e. sonication, French press, cell disruption). After lysating the cell, the insoluble form of the protein aggregated as inclusion body (IB). This insoluble form was washed with the lysis buffer B consisting of 10 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 5 mM DTT. After centrifugation at 17 Krpm, the pellet was solubilized in buffer containing 20 mM Na$_3$PO$_4$, pH 7.0, 8 M urea, and 100 mM 2-
mercaptoethanol. The dissolving step is to stir the sample at room temperature for an hour in the mentioned buffer. The sample was centrifuged again and the supernatant was collected at this stage of the purification. The supernatant was diluted to two fold the original volume of the buffer (10 mM Na$_3$PO$_4$, pH 7.0, 4 M urea, and 50 mM 2-mercaptoethanol) and left stirring either overnight or 4hrs at 4 °C. The sample was dialyzed to remove urea and β- mercapto-ethanol. Then the sample will be filtered and loaded into the GS-4B column by using the established protocol.

**3.2.5 FPLC - cation exchange SP - column and ultrafiltration application**

Fast Protein Liquid Chromatography (FPLC, GE, and #INV-907, made in Sweden) was a chromatographic process in which the protein in solution was exposed to a pH gradient system using a SP cation exchange column. The FPLC system was first cleaned with ultra-pure deionized water (ddH$_2$O). Then 50 mM Sodium Acetate (NaOAC) buffer (pH ≈ 4.0) and 50 mM Tris buffer (pH ≈ 8.0) were prepared and used as Buffer A and Buffer B, respectively. The column was gradually washed with the high pH buffer B so protein was eluted out of the column and separated from other contaminants and thrombin protease from the GST-tag column purification. The isoelectric point (pI) values of the variants are similar and are intended to elute out at pH 7.2- pH: 7.4 (Figure 3.16, 3.21). Amicon ultrafiltration pressure cell was used to increase the concentration of the purified protein solution. Inert Nitrogen gas applied pressure to the protein solution, thus causing elution of smaller molecules in solution. The membrane filter capacity (< 3.0 KDa) prevents the protein of interest and other biological molecules with a larger molecular weight from filtering out as waste.

**3.2.6 Quantifying the SDS-PAGE results of the GST-tag purification**

The SDS-PAGE results were quantified by taking the intensity ratio of the supernatant against the other sample band intensities (Figure 3.10 and Figure 3.11) through an Image J software (1.47v, Rasband, Wayne, NIH, USA). The supernatant from lane 2 in the GST-tag purification for ProCA1, 7E15E,
7E15N, and 7E15Q were normalized to 1 and were used as a reference to compare with the other intensity values taken from the other samples throughout the purification process. Essentially, the initial intensity value for the supernatant was divided by itself and all other intensity values from the other list of samples. The indicated range below the graph demonstrates how the density of the samples compares to the intensity of the supernatant. As a disclaimer: each band represents a 10-μL-volume sample loaded into the SDS-PAGE. Thus, it is assumed that all protein band intensity come from an equivalent volume that are quantified using image J software. The higher the ratio to one indicates that there is a similar or higher band density compared to the supernatant and vice versa.

As a general note, I must state that the summarized intensity ratios quantified in Figure 3.10, Figure 3.11, Figure 3.15, and Figure 3.17, have no error bars since the intensity ratio is based on one band intensity value. The purpose of the band intensity charts is to identify the actual percent of the protein in the supernatant versus the inclusion body form after cell lyses. In addition, the percent of soluble form of the protein is also compared to the GSH-GST-ProCA1 complex amount inside the column, the GST-ProCA1 complex, and the GST-free amount after the column is treated with phosphate buffer saline (1X PBS). The percent ratio determines the amount of bound GST-tagged ProCA1 variants that contributes to the final yield of the protein eluted out of the column after enzymatic cleavage with thrombin. Several of these factors are analyzed by comparing the intensity ratio of each band based on the set lysis buffer conditions. Basically, the detailed analysis of these intensity ratios indicate whether the change in the lysis buffer conditions have an impact on each factor such as the bound/free form of GST-ProCA1 amount to the beads.

3.2.7 Western-Blot and SDS-PAGE

Western blot assay was used to locate the position of ProCA1 in the gel. Regular 15 % SDS-PAGE gel was run with the samples. Then self-generated rabbit antibody against rat ProCA1 was used as the primary antibody. AP labeled polyclonal antibody was used as the secondary antibody. The 15 % SDS-
PAGE was transferred for two hours or overnight followed by the blocking and antibody incubation steps. Finally, the membrane was treated with enhancer and substrate for imaging. The instructions were similar to Fan Pu’s established protocol based on the commercially available procedure [64].
Figure 3.4 The commercially available pGEX-2T plasmid vector [55] has a GST promoter site before the Eco RI site.
Figure 3.5 The scheme for GS-4B column purification of GST-tagged ProCA1 variants. All ProCA1 variants are purified by using the four steps illustrated above for the GST-tag purification method and the sample abbreviations are indicated on the right panel of the scheme for clarification. Each eluted ProCA1 mutant is combined and passed through an SP-cation exchange column with an adjusted pH at ~ 3 to 4 to remove excess thrombin and other possible contaminants after GS-4B column purification.
Figure 3.6 The refolding procedure of the inclusion body (IB) of GST-tagged ProCA1 variants. This procedure is divided into four steps where the GST-tag ProCA1s are denatured then refolded back into the soluble form. The recycled supernant holds unbound GST-tag ProCA1 that is collected as flow through from the initial GS-4B column purification. This sample is labeled as beads after binding outside the column (BAB-OC). The sample is treated with steps 3 to 4 to serve as a recovery of the GST moiety’s catalytic activity to bind to the reduced glutathione on the beads in the column.
3.3 Results and Discussion

3.3.1 Expression of GST-tagged CD2.WT, ProCA1, and ProCA1 variants

3.3.1.1 Reliability of quantifying the protein bands using Image J software

The molecular weight of the GST-tagged ProCA1s is at a high position on the gel where there is more congestion. Some uncertainties to this is that the current yield in the expression intensity may be including other protein bands due to the fact that other background is present on the gel. This arbitrary or error is not necessarily relevant to the final purification yield. During purification, a final single protein band is observed as opposed to the heavy set of protein band on the expression gel’s background. Quantifying the protein band from the expression samples are estimated by the assumed protein molecular weight between 68 and 26 KDa protein markers.

3.3.1.2 The host cell strain effect on the small-scale expression of GST-tagged/charged ProCA1s

During the expression procedure in the small-scale system, exponential increases of optical density were observed (Figure 3.6 a and b) for ProCA1, 7E15E, and 7E15N. The small-scale expression is used to detect a clone or transformant that has the highest expression. In this system, 7E15E demonstrates highest expression level in terms of the intensity ratio and is used as a reference against the other intensity values obtained from the other protein samples. BL21 (DE3) pLysS is highest in expression level for GST-tag 7E15E, while BL21 (DE3) (79 %) and Tuner (92 %) demonstrate lower expression level. For the case of 7E15N, the intensity values are fairly similar in range percentage as observed in BL21 (DE3) (93 %), BL21 (DE3) pLysS (91 %) and Tuner (90 %). There is no conclusive indication that either of the three cell strains tested is most suitable for best cultivation level, solely based on the similar intensity values. Therefore GST-tag 7E15N can be successfully expressed at a similar cultivation level in either of the three cell strains mentioned earlier. For the case of GST-tag 7E15E, there is an indication that GST-fusion protein is expressed with a highest cell culture level in BL21 (DE3) pLysS.
The increased expression level in BL21 (DE3) is predominant for GST-tagged ProCA1 as compared to the low intensity value in BL21 (DE3) pLysS (90 %).

3.3.1.3 The effect of host cell strain, post-induction temperature, and IPTG concentration on the expression yield of CD2.WT

The optical density value varies as a function of time monitored at 600 nm during the large-scale expression in a one-liter LB medium. For GST-CD2.WT, there is a wide range in optical density value post induction time, which is between 1000 to 1500 minutes (Figure 3.7). This phenomenon suggests possible growth difference for each one-liter flask. Tuner cells demonstrate the most saturation in terms of protein band present within the 26 to 68 KDa regions. In following, the percent intensity is measured in correspondence to the 15% SDS-PAGE of the samples collected prior to collecting the expressed cells. The host cell strain involved in the large-scale expression of CD2.WT is BL21 (DE3) and Tuner. The highest intensity percentage with cells expressed in Tuner (0.30 mM IPTG at 30 °C overnight) is used as the intensity reference. The percent intensity value for cells cultured in BL21 (DE3) (0.30 mM IPTG) increases during overnight induction at 30 °C, and 25 °C (76 %, 72 %), respectively. Low expression is indicated during a five-hour induction period at 37 °C (58 %) in the same cell strain and inducer concentration. A longer induction time shows high expression of GST-tagged CD2.WT.

3.3.1.4 The effect of host cell strain, post-induction temperature, and IPTG concentration on the expression yield of ProCA1

GST-tagged ProCA1 demonstrates the overall highest optical density value with 0.30 mM IPTG at 25 °C in flask 3 (Figure 3.8 a). The expression conditions are separated by cell strain (BL21 (DE3), BL21 (DE3) pLysS, Tuner), IPTG concentration (0.10 to 0.35 mM), and temperature post induction (25 °C, 37 °C) for this large-scale expression. The optical density values are all calculated based on the diluted overnight culture samples to satisfy the optical density range recorded by the UV-vis spectrophotometer.
(0.1 to 1). The dilution factor is then used to determine the actual O.D. These values range from ~3 to ~16 for overnight expression at 25 °C while there is a narrow range from ~2 to ~6 optical density post induction at 37 °C. It is typical to decrease the temperature during overnight cultivation to promote gradual expression growth of the protein in the cytoplasm. The cell cultured at high temperature demonstrates a lower optical density, which may suggest low expression of the cells containing GST-tagged ProCA1. The protein bands are demonstrated at ~37.2 KDa for expression with 0.30 mM IPTG at 37 °C in BL21 (DE3) cells and used as the reference which shows highest overall expression (Figure 3.8 c, e). Expression levels fluctuate from a decreased to increased then back to a decreased intensity ratio with respect to the increased IPTG concentration at a fixed temperature (37 °C) (i.e. 0.10 mM, 66%, 0.20 mM 59 %, 0.30 mM, 100%, and 0.35 mM, 83%) in BL21 (DE3). GST-tagged ProCA1 is suitable for high expression at high IPTG concentration during an overnight induction period at 25 °C (i.e. 0.10 mM, 46 %, and 0.30 mM 54 %). At a 25 °C post induction temperature, the cultivation level increased by 74 % in BL21 (DE3) pLysS with a low IPTG concentration (0.10 mM) as compared to a decrease in BL21 (DE3). In the same expression conditions, Tuner cells shows highest culture level compared to BL21 (DE3) (Figure 3.8 c, e).

3.3.1.5 The effect of post-induction temperature, and IPTG concentration for 7E15E, 7E15N, and 7E15Q

GST-tagged 7E15E, 7E15N, and 7E15Q are all expressed in same E.coli host cell, BL21 (DE3). The overall highest culture for all three charged variants is shown in a 0.20 mM IPTG concentration during overnight induction at 25 °C. There is fluctuation with the expression level intensity ratio versus the IPTG concentration. The increased intensity ratio from 0.10 mM (58%) to 0.20 mM (100%) is observed but a sharp decrease occurs at the higher inducer concentrations (i.e. 0.30 mM, 72%, and 0.35 mM, 34 %) in 25 °C temperature for GST-tagged 7E15E. In terms of the fixed inducer concentration (0.10 mM), there is
no significant change in the expression level as the intensity ratios are similar in value (i.e. 25 °C, 58 %, and 37°C, 53 %) during the different cultivation temperatures (Figure 3.8 f). Similar to GST-tagged 7E15E, the expression output of GST-tagged 7E15N increased from 73 % (0.10 mM) to 100 % (0.20 mM) but then decreased from 77 % (0.30 mM) to 48 % (0.35 mM). At the constant IPTG concentration (0.10 mM), the expression level decreased from 73 % (25 °C) to 58 % (37 °C) with respect to the induction temperature (Figure 3.9 e). For GST-tagged 7E15Q, as the inducer concentration decreased, i.e. 0.35 mM, 0.30 mM, 0.20 mM; the intensity ratio increased (49 %, 70 %, 100 %) respectively, during a 25 °C overnight expression temperature (Figure 3.9 f).
Figure 3.7 The small-scale expression of GST-tagged CD2.7E15E, CD2.7E15N, and ProCA1.

a) The small-scale expression of CD2.7E15E and CD2.7E15N/pGEX-2T in BL21 (DE3), Tuner, and BL21 (DE3) pLysS shows an increased exponential growth as demonstrated by the optical density curve. b) As a control, the optical density in the small-scale expression of ProCA1/pGEX-2T in BL21 (DE3) pLysS and BL21 (DE3) are plotted as a function of time (min). c) The 15 % SDS-PAGE demonstrates the band intensity at ~37.2 KDa for CD2.7E15E, CD2.7E15N, and d) ProCA1. e) The percent band intensity are averaged and compared for each charged variants with the corresponding expressed cell strain for 7E15E, 7E15N, and f) ProCA1. Abbreviations are as follow: BL21 (DE3) (BL), Tuner (T), BL21 (DE3) pLysS (BP), 7E15E (E), 7E15N (N). Disclaimer: the quantified protein band intensity ratio base on the pixel density are a rough estimation due to the estimated molecular weight for each ProCA1 variants at ~37.2 KDa.
Figure 3.8 The large-scale expression of GST-tagged CD2.D1 wild type
a) The large-scale expression of CD2.WT/pGEX-2T in BL21 (DE3), and Tuner shows an increased exponential growth as demonstrated by the optical density curve. b) The 15% SDS-PAGE demonstrates the band intensity at ~38 KDa for all expression conditions. c) Each percent band intensity is compared with the corresponding expressed cell strain, temperature and IPTG concentration (mM). Abbreviations are as follow: BL21 (DE3) (BL), Tuner (T), overnight (o/n), wild type (WT). Disclaimer: the quantified protein band intensity ratio base on the pixel density are a rough estimation due to the estimated molecular weight for each ProCA1 variants at ~ 38 KDa
Figure 3.9 The large-scale expression of GST-tagged ProCA1 and CD2.7E15E at varying cell strains, IPTG concentration, and/or induction temperature.

a) The large-scale expression of ProCA1/pGEX-2T and b) CD2.7E15E/pGEX-2T in BL21 (DE3) shows an increased exponential growth as demonstrated by the optical density curve. c) The 15% SDS-PAGE for ProCA1 and d) E demonstrates the band intensity at ~37.2 KDa for all expression conditions. e) Each percent band intensity for ProCA1 and f) E are compared with the corresponding expressed cell strain, temperature and IPTG concentration (mM). Abbreviations are as follow: BL21 (DE3) (BL), BL21 (DE3) pLysS (BLpLsS), Tuner (T), overnight (o/n).

Disclaimer: the quantified protein band intensity ratio base on the pixel density are a rough estimation due to the estimated molecular weight for each ProCA1 variants at ~37.2 KDa
Figure 3.10 The large-scale expression of GST-tag CD2.7E15N and CD2.7E15Q at varying IPTG concentration and induction temperature. 

a) The large-scale expression of CD2.7E15N/pGEX-2T and b) CD2.7E15Q/pGEX-2T in BL21 (DE3) shows an increased exponential growth as demonstrated by the optical density curve. c) The 15% SDS-PAGE for N and d) Q demonstrates the band intensity at ~37.2 KDa for all expression conditions. e) Each percent band intensity for N and f) Q are compared with the corresponding expressed cell strain, temperature and IPTG concentration (mM). Abbreviations are as follow: BL21 (DE3) (BL), overnight (o/n). Disclaimer: the quantified protein band intensity ratio base on the pixel density are a rough estimation due to the estimated molecular weight for each ProCA1 variants at ~ 37.2 KDa.
3.3.1.6 The effect of host cell strain on the small-scale expression yield of GST-tagged HER2–targeted ProCA1

As discussed previously, the small-scale expression system allows for multiple trial analyses. Thus, expression conditions in terms of cell strains can be troubleshoot. The expression of the targeted ProCA1 is monitored simultaneously in both BL21 (DE3) and BL21 (DE3) pLysS. Both targeted ProCA1 variants display an increased exponential growth in accordance for all three trials corresponding to the two host cell strains. The final optical density ranges from 1.5 to 2 at 600 nm wavelength post overnight growth at 25 °C with a 5 mL LB media. In following, the percent intensity from the bands displayed on the 15% SDS-PAGE gels is compared and the cell strain which promotes the highest expression is observed for ProCA1.ZHER2: 342 and ProCA1.ZEGFR: 1907. The small-scale expression for the GST-tag targeted ProCA1.affi1907 is cultured at a high density in BL21 (DE3) (Figure 3.10). ProCA1.affi.HER2: 342 in the contrary is expressed in BL21 (DE3) (93 %) and BL21 (DE3) pLysS (94 %) at a similar intensity amount to ProCA1.affi1907 expressed in BL21 (DE3) pLysS.

3.3.1.7 The effect of host cell strain, post-induction temperature, and IPTG concentration on the large-scale expression yield of GST-tagged HER2/neu ProCA1 (ProCA1.affi342)

In the large-scale expression system, the overall highest intensity ratio is represented in the following expression condition for GST-tag ProCA1.affi342: 0.35 mM IPTG at 37 °C in BL21 (DE3) cells (Figure 3.11). Gradual increase in the intensity ratio (66, 70, 100 %) is observed in respect to the increase IPTG concentration (0.10, 0.30, 0.35 mM) with the exception of a gradual 20 % decrease observed at the highest IPTG concentration (i.e. 0.40 mM, BL21 (DE3), 80 %). An increase in the induction time at a lower induction temperature as compared to the shorter induction time period at the higher induction temperature demonstrate a decreased intensity ratio for cells expressed in BL21 (DE3) at a fixed 0.10 mM IPTG concentration. In the contrary, there is highest intensity ratio at 37 °C (96
% compared to the overnight 25 °C (68 %) induction period for GST-tag ProCA1.affi342 expressed in BL21 (DE3) pLysS. The same intensity ratio is observed for GST-tag ProCA1.affi342 expressed at 37 °C in Tuner (70 %) (0.20 mM IPTG) and BL21 (DE3) (0.30 mM IPTG).

3.3.1.8 The effect of host cell strain, post-induction temperature, and IPTG concentration on the large-scale expression yield of GST-tagged EGFR ProCA1 (ProCA1.affi1907)

The highest cultivation level is observed when GST-ProCA1.affi1907 is expressed in the following conditions: 0.10 mM IPTG, at 37°C in BL21 (DE3) pLysS. In terms of the fixed BL21 (DE3) cell strain, the increased intensity ratio is proportional to the increased IPTG concentration, i.e. 0.10 mM (46%) to 0.35 mM (93 %); but a sharp decrease is observed at the highest 0.40 mM (36 %) IPTG concentration (Figure 3.12). In terms of a fixed 0.10 mM IPTG concentration, the expression intensity level decreased, i.e. 54 % (at 25 °C) to 46 % (at 37 °C), with respect to the increased induction temperature in BL21 (DE3) cells. In the contrary, an increased expression intensity as a function of an increased induction temperature is observed in BL21 (DE3) pLysS cells. These results are consistent with those obtained for expression of GST-tagged ProCA1.affi342 and ProCA1.affi1907. GST-tag ProCA1.affi1907 increased in expression by 77 % in Tuner cells at 37 °C with 0.20 mM IPTG (Figure 3.12 c).
Figure 3.11 The small-scale expression of GST-tag ProCA1.ZHER: 342 and ProCA1.ZEGFR:1907

a) The small-scale expression of ProCA1.ZHER: 342 and ProCA1.ZEGFR:1907 PGEX-2T in BL21 (DE3), and BL21 (DE3) pLysS shows an increased exponential growth as demonstrated by the optical density curve. b) The 15% SDS-PAGE demonstrates the band intensity at ~44 KDa for both ProCA1.ZHER:342 and ProCA1.ZEGFR:1907. c) The percent band intensity are averaged and compared for each charged variants with the corresponding expressed cell strain along side the standard error bar. Abbreviations are as follow: BL21 (DE3) (BL), BL21 (DE3) pLysS (BLpLsS), ProCA1.ZHER2:342 (HR342 or ZHER2:342), ProCA1.ZEGFR:1907 (EGFR-1907 or ZEGFR:1907). Disclaimer: the quantified protein band intensity ratio base on the pixel density are a rough estimation due to the estimated molecular weight for each ProCA1 variants at ~ 44 KDa
Figure 3.12 The large-scale expression of GST-tagged ProCA1.ZHER2:342 at varying cell strains, IPTG concentration, and induction temperature.

a) The 15% SDS-PAGE demonstrate the band intensity at ~44 KDa for all expression conditions.  
b) The large-scale expression of ProCA1.ZHER2: 342/pGEX-2T in BL21 (DE3), BL21 (DE3) pLysS, and Tuner shows an increased exponential growth as demonstrated by the optical density curve.  
c) Each percent band intensity is compared with the corresponding expressed cell strain, temperature and IPTG concentration (mM). Abbreviations are as follow: BL21 (DE3) (BL), BL21 (DE3) pLysS (BLpLsS), Tuner (T), overnight (o/n). Disclaimer: the quantified protein band intensity ratio base on the pixel density are a rough estimation due to the estimated molecular weight for each ProCA1 variants at ~ 44 KDa.
Figure 3.13 The large-scale expression of GST-tag ProCA1.ZEGFR: 1907 at varying cell strains, IPTG concentration, and induction temperature.

a) The 15% SDS-PAGE demonstrate the band intensity at ~44 KDa for all expression conditions. b) The large-scale expression of ProCA1.ZEGFR: 1907/pGEX-2T in BL21 (DE3), BL21 (DE3) pLysS, and Tuner shows an increased exponential growth as demonstrated by the optical density curve. c) The percent band intensity is compared with the corresponding expressed cell strain, temperature and IPTG concentration (mM). Abbreviations are as follow: BL21 (DE3) (BL), BL21 (DE3) pLysS (BLpLsS), Tuner (T), overnight (o/n). Disclaimer: the quantified protein band intensity ratio base on the pixel density are a rough estimation due to the estimated molecular weight for each ProCA1 variants at ~ 44 KDa.
3.3.2 Purification of GST-tag CD2.WT, ProCA1, and ProCA1 variants

3.3.2.1 The effect of lysis buffer conditions on the GST-tag ProCA1 charged variants

The following 15 % SDS-PAGE is a compilation of the results for the GST-tag purification (Figure 3.13). Each column represents the samples taken based on the set lysis buffer conditions. For instance, the supernatant sample and the inclusion body sample represent the soluble and insoluble form, respectively, of the protein after cell disruption. The specific intensity ratio for these bands are compared (Figure 3.14). The sample lane of the protein solution for beads after binding outside the column (BAB-OC) represents the combination of the free contaminants, the GST-tagged ProCA, and the GST-free components in lane four (Figure 3.13). The sample lane of the protein solution for beads after binding inside the column (BAB-IC) indicates the density amount of bound GST-tag ProCA1 charged variants to the glutathione coated beads. There is successful binding of the GST-tag proteins with the beads based on the demonstrated protein band at ~ 37.2 KDa. In addition, the free eluted GST moiety at 26 KDa suggest that successful binding of the GST-tag proteins occurred during the purification process. Lane 8 demonstrates the eluted free GST after the column is treated with a high concentration of free glutathione in the reduced form. Overall, each GST-tag ProCA1 variant has a band density at ~37.2 KDa from the corresponding lane numbers (1, 2, 4, and 5) and are quantified based on their intensity (Figure 3.14). ProCA1, CD2.7E15E, CD2.7E15N, and CD2.7E15Q are eluted out successfully as indicated by the protein band at ~ 11.2 KDa in lane 7. The strongest band intensity represents the concentration of successful eluted free ProCA1 charged variants for the lysis buffer with lysozyme and the established method used from previous results in Natalie White’s Thesis [34]. There is successful purification of all ProCA1 charged variants treated with the three different cell lysis buffers. This is an important concept to grasp when analyzing the compilation of these SDS-PAGE results (Figure 3.13).

GST-7E15E binds successfully to the beads (BAB-IC, 110 %) but a portion of the fusion protein is left unbound (BAB-OC, 57 %). 67% of the GST-7E15E is unfolded or aggregated as inclusion body after
cell lysis. As a result, only 27 % of 7E15E is eluted after enzyme cleavage with thrombin. The low percentage in elution is a result of inefficient binding of the GST-tag protein during the binding process, which consists of the beads being treated in a consecutive wash with the supernatant. All these results correspond to the GST-7E15E sample treated with lysis buffer A. One factor to consider is the increase in the binding efficiency between the beads and the GST-tag protein by extending this binding process. Other considerations are further discussed in the summarized Section for the GST-tag purification of both charged and targeted ProCA1 variants (Figure 3.22). In the lysis buffer B condition, GST-7E15E is in the inclusion body compared to the supernatant. A decrease in the beads binding to GST-7E15E is suggested as a result of the percentage between BAB-Outside Column sample (99 %) and BAB-Inside Column sample (54 %) in reference to the supernatant intensity ratio. The eluted GST-free moiety (51 %) represents the percent bound of the GST-GSH complex that was left post the enzymatic cleavage process. These results suggest that binding is not effective after the minor modification of some parameters in the lysis buffer A. GST-7E15E was not treated with lysis buffer C, so GST-ProCA1 is used as an example of a charged variant with a -5 charged binding pocket, and the results are analyzed in comparison to GST-7E15N and GST-7E15Q which carry a -4 charged binding pocket. GST-ProCA1, treated with lysis buffer C, has less aggregation or unfolding as inclusion body (43 %). The binding efficiency between the beads and the fusion protein increased when comparing the percent intensity for samples labeled as BAB-OC (88 %) and BAB-IC. Although there is an increase to form the GSH-GST-ProCA1 complex inside the column, only approximately half of the protein is eluted (53 %) after enzyme cleavage with respect to the initial protein amount present in the supernatant sample. The eluted GST-free (29 %) represents the successful binding of GSH-GST-ProCA1 prior to enzyme cleavage. There is a possibility that leftover GSH-GST is present inside the column because of the low percent intensity of the sample labeled BAE compared to the elution.
The soluble and insoluble portion of GST-7E15N (IB, 93 %) and GST-7E15Q (96 %) are similar in intensity value after cell lysing with treatment of the lysis buffer A. The binding efficiency for the GST-tag proteins to bind onto the GS 4 beads are analyzed by comparing the intensity ratio for samples labeled as BAB-OC and BAB-IC. The GST-proteins bind successfully, but a similar portion of the unbound protein fusion is present in the flow though after the binding process. BAB-OC (101 %, 52 %) and BAB-IC (84 %, 52 %) samples demonstrate a small difference in intensity ratio for GST-7E15N and GST-7E15Q, respectively. This observation suggests that binding is effective to a certain extent, meaning that we need to consider the beads capacity limit for complex formation to occur with the GST-tag moiety of the protein during the binding process. For the case of cell lysing with the lysis buffer B, GST-7E15N and GST-7E15Q have similar soluble/insoluble ratio when comparing the SN and IB intensity ratios. The intensity ratio increases for BAB-OC sample while it decreases for BAB-IC. GST-7E15Q (BAB-IC, 87 %) forms complex with GSH at a similar intensity ratio range as compared to GSH-GST-7E15N (BAB-IC, 76 %) complex formation. The GST-free elution sample (BAE, 79 %) has a comparable intensity ratio to the labeled BAB-IC sample, and BAE shows the efficiency of the complex formation (GSH-GST-7E15N) before enzyme cleavage. In addition, enzyme cleavage of the GSH-GST-7E15Q complex gives a low percent intensity ratio for the GST-free moiety. Three possibilities to the lower GST free percent intensity are: the low GSH concentration present in the elution buffer, the pH adjustment of the elution buffer, and the thrombin concentration for enzyme cleavage. GST-7E15N and GST-7E15Q, treated with the lysis buffer C, illustrate similar solubility/insoluble intensity ratio as compared to the same protein variants treated with the lysis buffer A.
3.3.2.2 The denatured-refolded influence on the GS-4B purification of ProCA1 charged variants

The insoluble portion of 7E15E, 7E15N, and 7E15Q are treated through a refolding method, and the binding capability of the refolded sample is passed through the GS-4 beads. The results are compiled in the form of each sample lane intensity displayed on the 15 % SDS-PAGE (Figure 3.15). All three GST-ProCA1 charged variants demonstrate similar weak binding capability to the glutathione-coated beads. Lane 1 to 4 demonstrate the successful refolding of GST-tag 7E15N and 7E15Q based on the consistency in the protein band intensity at 37.2 KDa. In following, GST-tag 7E15N and 7E15Q interestingly has similar unbinding to the column by comparing lane 6 and 7 similar band intensity. The quantified intensity ratio further supports these visual observations of the SDS-PAGE results (Figure 3.12). Evidently, there is a low binding interaction between the glutathione reduced form and the GST-7E15N and GST-7E15Q as indicated by the light protein band in lane 8 and 9. In the contrary of GST-tag 7E15N and 7E15Q, GST-tag 7E15E demonstrates a low solubility after treatment with 8 M urea for one hour. The clear difference in protein band intensity between the supernatant in lane 1 and the inclusion body in lane 2 supports these observations for GST-tag 7E15E. All refolded GST-tag proteins that are recovered at the soluble state demonstrate a low binding activity towards the reduced glutathione present on the sepharose beads in the gravity column. Some factors to consider are 1) the expression yield in terms of band intensity for the proteins prior to cell lysis, 2) the state of the glutathione coated beads in the column, 3) the refolded GST-ProCA enzymatic activity, and 4) the pH condition of the column. The overall yield obtained after purification is depending of the initial yield gathered during overexpression of the GST-tagged ProCA1s. The concentration of the salt, ionic detergent, denaturant, and pH play an important role in the case of refolding these GST-tagged ProCA1s.
3.3.2.3 ProCA1 charged variants isolated by the FPLC cation exchanger SP-column

The concentration of each protein elution fractions are monitored through the UV-Vis spectrophotometer and then prepared for further isolation through the FPLC SP-column system. Each protein sample is adjusted to a low pH (4.00 - 3.00) for elution based on a pH gradient program. The protein at this specific acidic pH is expected to bind to In addition, the protein’s original binding pocket holds acidic residues (aspartic and glutamic acid). The side chains of these residues are protonated due to the acidic condition of the sample and for an effective binding to the SP-cation exchanger column. All ProCA1 charged variants elute out at pH: 7.20 to pH: 7.40 (Figure 3.16). The desired fractions that demonstrate the highest mAU in that pH range are collected and then passed through an ultrafiltration system. Table 3.1 summarized the protein final concentration and yield is determined based on the UV-Vis spectra and these samples are used for the conformational analysis.

3.3.2.4 Comparing the yield (mg/L) after the GS-4B purification of ProCA1 charged variants

The purpose of compiling the UV-Vis absorbance ratio is to compare the level of nucleotide presence in correspondence to lysis buffer. Second, the final concentration (μM) and the yield (mg/L) based on the elution concentration are compared based on the lysis buffer system. A detail analysis on which lysis buffer Best contributes for high protein yield is analyzed for each ProCA1 charged variant (Table 3.1). Furthermore, the concentration obtained for the refolding attempt of the insoluble portion of GST-tag variants after cell lysing is compared as well with the previous lysis buffer system. 7E15E (9 mg/L) and 7E15N (4 mg/L) show highest yield in lysis buffer B, while 7E15Q (10 mg/L) has highest yield in buffer A. ProCA1 (29 mg/L) demonstrates high yield based on the lysozyme content in the buffer C. A small yield is obtained for 7E15E (2 mg/L), 7E15N (1 mg/L), and 7E15Q (0.47 mg/L) after refolding the GST-tag protein from the inclusion body in buffer B. As a result, a small percent amount of the proteins are recovered back into the soluble state. The intensity ratio for the samples taken during the refolding of the GST-tag proteins from the inclusion body is supported in Figure 3.15.
Figure 3.14 The GST-tag purification of ProCA1, CD2.7E15E, CD2.7E15N, and CD2.7E15Q with modifications in the cell lysis buffer.
The following are the 15% SDS-PAGE samples taken for each ProCA1 charged variant: a) the conventional cell lysis buffer (1% N-Lauralsarcosyl, 1 mM EDTA, 0.50 mM PMSF), b) the modified cell lysis buffer (1% N-lauralsarcosyl, 5 mM EDTA, 6.7 mM PMSF, 10 mM DTT, 0.05 % Triton-X 100), and c) buffer A condition with lysozyme (1 mg per 1 liter cell cultured) for i) CD2.7E15E, ii) ProCA1, iii) CD2.7E15N, and iv) CD2.7E15Q.
Figure 3.15 A summary of the intensity ratio for each sample from the GST-tag purification of ib) ProCA1, ia) 7E15E, ii) 7E15N, and iii) 7E15Q with the indicated lysis buffer conditions.

Each columns represent the a) lysis buffer A, b) lysis buffer B, and lysis buffer C with the protein band intensity ratio for each sample collected during the GS-4B purification system.

Intensity ratio: 1 : similar density to SN, ≤ 1 : lower density to SN, ≥ 1 : higher density to SN
Figure 3.16 The GST-tag refolding purification of CD2.7E15E, CD2.7E15N, and CD2.7E15Q
There is a summary of the intensity ratio for each sample from the GST-tag refolding purification of a) CD2.7E15E, b) CD2.7E15N, and c) CD2.7E15Q.
Figure 3.17 An example of the FPLC chromatograph for CD2.7E15Q (pH: 3.80) through the pH gradient in the SP column. A low pH range (~3 to 4) is adjusted to isolate our 7E15Q from thrombin (< ~1 mg/mL), reducing reagent (DTT), and detergents (1 % sarcosyl & Triton X-100). The above chromatograph indicates the purity of the protein sample after the GS-4B column purification, and the fraction sample collected shows one distinct protein band on the 15 % SDS-PAGE affirming the stability, no degradation, and purity level of the sample protein after isolation from thrombin.
Table 3.1 A summary of the UV-Vis absorbance ratio and yield (mg/1 L culture) for each eluted or concentrated samples from the GST-tag and refolding method of the charged ProCA1 variants.

<table>
<thead>
<tr>
<th>Lysis Buffer</th>
<th>$A_{260nm}/A_{280nm}$</th>
<th>$C_i$ ($\mu$M) w/ Dilution factor</th>
<th>Total Yield* (mg/L)</th>
<th>(mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.556</td>
<td>115</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>$B^\dagger$</td>
<td>0.834</td>
<td>112</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>$B^\ddagger$</td>
<td>0.912</td>
<td>82</td>
<td>2</td>
<td>3.7</td>
</tr>
<tr>
<td>C</td>
<td>0.698</td>
<td>340</td>
<td>29</td>
<td>15</td>
</tr>
<tr>
<td>A</td>
<td>0.780</td>
<td>41</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>$B^\dagger$</td>
<td>1.055</td>
<td>70</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>0.533</td>
<td>227</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>$B^\ddagger$</td>
<td>1.103</td>
<td>114</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>0.551</td>
<td>116</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>$B^\dagger$</td>
<td>1.347</td>
<td>17</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>C</td>
<td>0.529</td>
<td>258</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>$B^\ddagger$</td>
<td>1.440</td>
<td>9</td>
<td>0.47</td>
<td>9</td>
</tr>
</tbody>
</table>

$\dagger$ The final samples are from elution fractions collected out of the GS-4B column.

$\ddagger$ The refolded GST-tagged ProCA1s from the inclusion body

* The calculated total yield is based on the final concentration sample from one GST-tag purification

Lysis Buffer: A) conventional method  B) Add 0.05 % Triton X-100, 5 mM EGTA  C) Add Lysozyme

Net charge of binding pocket: 7E15E ***7E15N 7E15Q
3.3.2.5 The effect of lysis buffer conditions on the purification of GST-tagged ProCA1s targeted

This Section consists of the GST-tagged ProCA1.affi targeted variants that are purified within a set of lysis buffers. Each buffer as mentioned earlier is tested to see which one promotes highest solubility to insolubility intensity ratio after cell lysis. In addition, the binding efficiency between the GST-tag proteins and the GSH coated sepharose beads are observed during the purification step. Each lane on the 15 % SDS-PAGE represents the samples taken during the GS-4B column purification step. The specific bands displayed in a red box are quantified using the Image J software. The two parts to the purification results of ProCA1.affi342 and ProCA1.affi1907 are: a) the soluble portion or SN of the protein that is passed through the GS-4B column, b) the insoluble form of the protein that is collected for possible refolding through the use of the lysis buffer B. These results are demonstrated on the 15 % SDS-PAGE (Figure 3.17, 3.19). Initially, the first part into obtaining the eluted ProCA1.affibodies demonstrate different outcomes for binding capability and elution output when comparing the buffer lysis conditions. These results are analyzed further in a quantitative manner for both part A and B.

Each cultured cell pellets that contain GST-ProCA1.affi342 or GST-ProCA1.affi1907 is purified in subsequent different lysis buffers. The results are discussed through detailed analysis of the measured intensity ratios for each of these two variants based on the individual gel band results with correspondence to the used lysis buffer. Consistency with the protein extraction from the host cell is observed in the intensity ratio for GST-ProCA1.affi342 between the SN and IB. So there is an equal amount of the protein present as soluble and insoluble form after cell lyses when using buffer A. The same case is observed when using buffer B. Increasing the chelator, EDTA, concentration (mM) and adding detergent does not increase the folding of the GST-protein complex into the soluble form after cell lyses for ProCA1.affi342. In the case of using buffer C, more inclusion body is observed when adding lysozyme prior to cell lyses with sonication/cell disruption. This particular condition may not necessarily contribute to high inclusion body; other factors such as the expression conditions may have an indirect
impact on this outcome. Following this phenomenon, there is less GSH-GST-ProCA1.affi342 present inside the GS-4B column. For the case of ProCA1.affi1907 targeted to EGFR, insoluble portion is lower in percent intensity during the usage of buffer lysis A (59 %) and C (48 %), while there is no significant change in the insolubility of the protein when using buffer lysis B. There is a large amount of GSH-GST-ProCA1.affi.EGFR after using the lysis buffer B (61 %) and C (85 %) while there is a low GSH-GST-protein formation (28 %) in the column after using cell lysis buffer A.

3.3.2.6 The denatured-refolded influence on the GS-4B purification of targeted ProCA1s

Part B of the purification results involves the analysis of the refolded inclusion body collected during the use of the lysis buffer B for both GST-ProCA1.affi.HER2:342 and GST-ProCA1.affi.EGFR. The supernatant that is recycled after the washing step during the initial GS-4B column purification is refolded alongside the inclusion body and the results are compared in the 15 % SDS-PAGE (Figure 3.19) and the protein bands’ intensity ratio (Figure 3.20). The assumption is that the non-bond GST protein may be inactive or have some misfolding of the thiol group although it is present in the soluble form. There is darker band intensity for the refolded GST-ProCA1.affis from the inclusion body form compared to the refolding step of the recycled washout or supernatant. These observations are taken from a qualitative into a quantitative approach as the intensity ratios are measured for indication of the percent density of the protein of interest present in each sample (Figure 3.19). GST-ProCA1.affi.HER2:342 decreases in density during the refolding step and only 43 % intensity ratio is passed through the GS-4B column as compared to the referenced SN treated in 8 M urea for one hour while GST-ProCA1.affi.EGFR has a constant intensity ratio during the refolding process of the inclusion body. The increased solubility is most prominent for GST-ProCA1.affi1097 during inclusion body refolding. For the case of refolding the recycled washout or supernatant, GST-ProCA1.affi.HER2 and ProCA1.affi.EGFR (83 %) both demonstrate a consistence in percent intensity based on the final sample (SN, 10 mM Tris) obtained prior to GS-4B column purification. As a result, ProCA1.affi.HER2:342 (55 %)
and ProCA1.affi1907 (47%) showed a 50% staining recovery in the elution sample. The yield is obtained after the purification and then compared for both protein variants. The goal is to refold the insoluble form into a soluble form then treat the refolded supernatants through the GS-4B column for further purification. This is the process that is applied for essential back-up recovery of additional protein as gaged by its ability to bind to the column and to increase the yield after purification. The compared yield is determined using the UV-Vis absorbance maxima and final concentrated volume of the purified samples (Table 3.2).

### 3.3.2.7 Comparing the yield (mg/L) after the GS-4B purification of targeted ProCA1s

The protein sample is isolated and a spectrum (Figure 3.18) with the specified fraction peak is collected based on the SDS-PAGE. Table 3.2 summarizes the yield measured for each ProCA1.affi variant with the corresponding final volume (mL). The GST-ProCA1.affi.HER2 and GST-ProCA1.affi1907 demonstrate drastically different elution yield. In the case of refolding, it is evident that GST-ProCA1.affi.HER2: 342 (0.05 mg/L, 1 mg/L) have a lower yield compared to GST-ProCA1.affi.EGFR (7 mg/L, 4 mg/L) after the denatured-refolding from the supernatant or inclusion body sample, respectively. This observation suggests that GST-ProCA1.affi1907 is most susceptible to the refolding process and increase capability of the GST moiety to catalyze the reduced glutathione on the beads for complex formation with GSH when passed through the GS-4B column.
Figure 3.18 The GST-tag purification of ProCA1.affi342 and ProCA1.affi1907 with modifications in the cell lysis buffer.
Figure 3.19 A summary of the intensity ratio for each sample from the GST-fused purification of ProCA1.affi342 and ProCA1.affi1907 with modifications in the cell lysis buffer.
Figure 3.20 The GST-tag refolding purification of ProCA1.affi342 and ProCA1.affi1907.
Figure 3.21 A summary of the intensity ratio for each sample from the GST-tag refolding purification of ProCA1.affi342 and ProCA1.affi1907.

### Intensity ratio: 1 : similar density to SN, ≤ 1 : lower density to SN, ≥ 1 : higher density to SN
Figure 3.22 The FPLC chromatograph of ProCA1.affi1907 through the Q column based on a salt-gradient.
Table 3.2 A summary of the UV-Vis absorbance ratio and yield (mg/L) for each eluted or concentrated samples from the GST-tag and refolding method of the targeted ProCA1 variants.

<table>
<thead>
<tr>
<th>Lysis Buffer</th>
<th>$A_{260nm}/A_{280nm}$</th>
<th>$C_f$ (µM) w/ Dilution factor</th>
<th>Total Yield* (mg/L)</th>
<th>(mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.749</td>
<td>13</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>B†</td>
<td>1.266</td>
<td>3</td>
<td>0.23</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>1.048</td>
<td>53</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>B†**</td>
<td>1.266</td>
<td>3</td>
<td>0.05</td>
<td>2</td>
</tr>
<tr>
<td>B†***</td>
<td>0.908</td>
<td>9</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>A</td>
<td>0.687</td>
<td>13</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>B†</td>
<td>0.633</td>
<td>43</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>C</td>
<td>1.338</td>
<td>23</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>B†**</td>
<td>0.657</td>
<td>58</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>B†***</td>
<td>0.808</td>
<td>29</td>
<td>4</td>
<td>15</td>
</tr>
</tbody>
</table>

Lysis Buffer: A) conventional method, B) Add 0.05 % Triton X-100, 5 mM EGTA, C) Add Lysozyme, D) refolding of IB E) refolding of SN (BAB-OC)

† The final samples are from elution fractions collected out of the GS 4 beads column
★ The refolded GST-tagged ProCA1s targeted from the inclusion body
♦ The refolded GST-tagged ProCA1s targeted from the SN flow-through (BAB-OC) sample

* The calculated total yield is based on the final concentration sample from one GST-tag purification
3.4 Conclusion and Future Work

This Chapter involved expressing and purifying the various ProCA1 variants that are charged and/or targeted. The most effective approach used for optimizing the production of these proteins is through a series of troubleshooting the expression and purification in different set conditions. This Chapter takes part in two series: a) the optimization of expressed GST-tagged ProCA1 variants, and b) the optimization of purified GST-tagged ProCA1 variants through a gravity GS-4B column. The initial step to producing the recombinant proteins is via a fed-batch/bench work expression system. This task takes about a week for completion and the cultured cells that contain the expressed protein of interest is the end result prior to the purification step. The essential outcome from the expression and purification are discussed further into two parts. Part A discussed the summary obtained from the optimization of the small and large-scale expression system for GST-tag ProCA1 variants, while part B focuses on the summarized results from the gravity GS-4B column purification.

3.4.1 Summary of the small and large-scale expression for all ProCA1 variants

The optimal conditions are summarized with reference to the fixed conditions as listed in Table 3.3. In the small-scale expression system, all charged variants cultured in the smaller LB culture volume exhibit no significant difference in the expression level of the GST-tagged proteins with the varied cell strains. The three sets of fixed parameters in the large-scale expression are between 1) E.coli cell strain, 2) post-induction temperature, and 3) inducer concentration (i.e. IPTG). All the GST-tagged proteins are expressed depending on these sets of fixed conditions. During expression in the same E.coli cell strain and post induction temperature and time, all ProCA1 variants demonstrate similar culture growth levels with 0.20 mM IPTG. Similarly, the affibody-targeted variants show the same expression level with 0.35 mM IPTG during treatment of the same E.coli strain and temperature conditions. ProCA1 is the set control for the expression and it demonstrates a similar expression behavior as the targeted variants for expression at a high temperature. During a fixed temperature/cell strain: a successful optimizing
condition is indicated for the high expression level of GST-tagged 7E15E, 7E15N, 7E15Q (0.20 mM IPTG at 25 °C in BL21 (DE3), GST-tag ProCA1.affi342 and ProCA1.affi 1907 (0.35 mM IPTG at 37 °C in BL21 (DE3), following ProCA1 (0.30 mM IPTG, at 37 °C in BL21 (DE3). At a fixed IPTG /temp, successful high expression level is present between BL21 (DE3) pLySs and BL21 (DE3) for GST-tag ProCA1 affibodies. At a fixed IPTG /cell strain, successful high expression level is present at 25°C and 37°C in BL21 (DE3) pLySs/BL21 (DE3) for GST-tag ProCA1 affibodies, and GST-tag 7E15E, and 7E15N. Results suggest that at low temp with a longer cultivation time for the 3 charged variant proteins’ (7E15E, 7E15N, and 7E15Q) expression level is most favorable while a shorter cultivation time at high temp is suitable for the targeted ProCA1 variants (ProCA1.affi342 & ProCA1.affi1907). As future plan, we aim to modify the trials further for expression on the small scale with a set control for IPTG, or temperature, and correlate those findings with the current large-scale results.
Table 3.3 The summary of the small-scale and large-scale expression of GST-tag ProCA1 variants. The small-scale expression represented on the left hand panel of the Table refers to the effect of the host cell strain while the large-scale expression represented on the right hand panel refers to the host cell strain (green), temperature (purple), and IPTG (yellow) concentration.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Small-scale Expression</th>
<th>Protein</th>
<th>Expression Parameters (constant)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>E. Coli Cell Strain, Temp. (°C)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tuner (0.30 mM, 30 °C)</td>
</tr>
<tr>
<td>CD2.WT</td>
<td>BL21 (DE3)/BL21 (DE3) pLysS</td>
<td>ProCA1 0.30 mM (BL21 (DE3), 37 °C)</td>
<td>I) BL21 (DE3) pLysS (0.10 mM, 25 °C o/n)</td>
</tr>
<tr>
<td>ProCA1</td>
<td>BL21 (DE3) pLysS/Tuner 7E15E 0.20 mM (BL21 (DE3), 25 °C o/n)</td>
<td>N/A 25 °C o/n, 37 °C (0.10 mM, BL21 (DE3))</td>
<td></td>
</tr>
<tr>
<td>7E15E</td>
<td>BL21 (DE3) pLysS/Tuner 7E15N 0.20 mM (BL21 (DE3), 25 °C o/n)</td>
<td>N/A 25 °C o/n (0.10 mM, BL21 (DE3))</td>
<td></td>
</tr>
<tr>
<td>7E15N</td>
<td>BL21 (DE3) pLysS/Tuner 7E15Q 0.20 mM (BL21 (DE3), 25 °C o/n)</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>7E15Q</td>
<td>BL21 (DE3) pLysS/Tuner ProCA1.affi342 0.35 mM (BL21 (DE3), 37 °C)</td>
<td>I) BL21 (DE3) (0.10 mM, 25 °C o/n)</td>
<td>I) 37 °C (0.10 mM, BL21 (DE3)pLysS)</td>
</tr>
<tr>
<td>ProCA1.affi342</td>
<td>BL21 (DE3)/BL21 (DE3) pLysS ProCA1.affi1907 0.35 mM (BL21 (DE3), 37 °C)</td>
<td>I) BL21 (DE3) pLysS (0.10 mM, 25 °C)</td>
<td>I) 37 °C (0.10 mM, BL21 (DE3)pLysS)</td>
</tr>
<tr>
<td>ProCA1.affi1907</td>
<td>BL21 (DE3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.4.2 The summarized GST-tag purification of charged and/or targeted, ProCA1 variants

The purification of the ProCA1 variants is optimized through troubleshooting in the three sets of lysis buffer conditions. The breaking of the bacteria cell wall is essential for excretion of the recombinant protein of interest, and in doing so; the protein is exposed to harmful agents present in the lysate after cell lysis. In addition, the viscosity of the lysate sample, the solubility, and bioactive form of the GST-ProCA1 variants is important to proceed into the GS-4B column. In pointing out these key factors, the essential idea for optimizing this purification system is to modify the lysis buffer for promoting an increase in the solubility of the tag-protein for potential higher recovery into the soluble state after cell lysis. I hypothesize that a high ProCA1 variants ratio are folded partially due to the method of cell lysing; as a result they are present in the soluble form and a potential higher yield is obtained after the GS-4B column purification. Treating the extracted protein with different cell lysis buffer conditions tested this hypothesis. The overall results convey otherwise that the lysis buffer condition has a small impact on the overall protein folding during protein extraction from the host cell and the final yield measured after GS-4B column purification. The change in the components within the lysis buffer posed some change in the solubility and binding efficiency of the GST-ProCA1 variants but these results varied with the particular ProCA1 variant. The -5 and -4 charged variants are compared in terms of the soluble/insoluble and unbound/bound form during GS-4B column purification.

It is more prominent that -5 charged variants demonstrate a high percent intensity ratio of the soluble and bound form (lysis buffer A and lysis buffer C). There was no significant change between the soluble and insoluble form based on the similar intensity ratio exhibited by 7E15E (lysis buffer B). In addition 7E15E demonstrates a high percent intensity of the unbound form (99 %) compared to the bound form (54 %) during the purification. The measured yield (lysis buffer B, 9 mg/L) is 3 fold higher than the yield obtained after treatment of the lysis buffer A (3 mg/L). Based on these conclusions, the lysis buffer A promotes a higher soluble/insoluble ratio for GST-7E15E without the added DTT, high
EGTA concentration, triton x-100, and lysozyme. 7E15N and 7E15Q are -4 charged variants and their results are compared to the -5 charged variants.

There is no change in the soluble to insoluble ratio for 7E15N in the lysis buffer A and C (1.08,1.00), respectively, while buffer B has a lower ratio indicating an increase in the inclusion body formation. The high increased soluble/insoluble ratio is present in 7E15E that is treated with lysis buffer A, 7E15N (lysis buffer A, C, and D), 7E15Q (lysis buffer A, C, D). For the targeted protein variants, ProCA1.affi342 demonstrates high intensity ratio in lysate buffer A, B, and D for the soluble/insoluble ratio. ProCA1.affi1907 shows high intensity ratio in lysate buffer A and C. The ultimate goal is to obtain the highest ratio of the protein of interest in the supernatant and for this protein to demonstrate high binding interaction inside the column. This can be detected by the increase ratio values for SN/IB and BAB-IC/BAB-OC.

The soluble form is achieved through the use of lysis buffer A, C, and D, where the conditions in the buffer successfully demonstrate the high intensity ratio of soluble/insoluble form of 7E15E, ProCA1, 7E15N and 7E15Q. In the case of the targeted protein buffer A, B, D demonstrates high soluble/insoluble ratio of ProCA1.affi342, while buffer A and C demonstrate highest soluble/insoluble ratio of ProCA1.affi1097. Not only is the increased solubility of importance in this analysis, but also achieving a high binding/unbinding form inside the column is important for potential higher output of the protein of interest after enzymatic cleavage with thrombin. The increase in bound/unbound form is present in the following buffers with the corresponding protein: 7E15E (A), 7E15N (C), 7E15Q (C), ProCA1.affi342 (A), ProCA1.affi1907 (C). In terms of the yield there is no significant correlation demonstrated with the highest intensity for both SN/IB and BAB-IC/BAB-OC. These results are essentially preliminary and suggest that further analysis is needed for effective conclusions to be made.

The current conclusion that can be made based on these preliminary results is that modifying the current lysate buffer conditions in A into lysate buffer B is essential for obtaining a high yield as
demonstrated for 7E15E (9 mg/L) and 7E15N (4 mg/L). In following, we must consider the discrepancy to the data obtained because each cultured pellets come from varying expression conditions as 7E15Q, ProCA1.affi342, and ProCA1.affi1907 demonstrate a highest yield in different lysate buffer conditions, A (10 mg/L), C (5 mg/L), and D (7 mg/L). Thus, further analysis into modifying the conditions such as increasing the percent concentration of the triton X-100 detergent and verifying the pH of the buffer throughout the cell lysis process are some essential criteria to consider for future experimental plans. Further analysis is needed for accurate interpretation of the data collected because these results are a summary of single trials. As future work, the experiments will be scaled-down for ease of executing the experiment in triplicates for each corresponding lysis buffer condition. The end result will be gathered data that is easily re-producible, more effective and reliable.
Figure 3.23 The summary of the GS-4B purification of each ProCA1 variants; charged and targeted. Each protein variant is indicated with their intensity ratio of the bound/unbound form (grey/yellow) and the soluble/insoluble form (blue/purple) in a) the GS-4B purification and b) refolding of IB. c) Each yield is indicated with their corresponding protein variant and lysis buffer condition while the summarized yield (mg/L) is measured using the UV-Vis absorbance maximum at 280 nm for GS-4B purification and d) refolding of GST-tagged ProCA1s.
4. CONFORMATIONAL ANALYSIS, METAL BINDING, AND RELAXIVITY PROPERTIES OF THE ENGINEERED PROCA1S

4.1 Introduction

Chapter 4 focuses on the detailed structural conformation, metal binding and relaxation analysis of the expressed and purified ProCA1 variants, both targeted and charged. These analyses are important for the prospective clinical application. In following, ProCA1 variants are derived from the CD2.D1, or CD2.WT; that is an adhesion protein with a secondary β-sheet structural conformation (Figure 1.3). There are several spectroscopic techniques such as NMR, CD, and fluorescence that can be used to investigate the conformational changes of the ProCA1 variants. Specifically, tryptophan fluorescence emission and far ultraviolet (UV) CD spectroscopy are used among these applications. The secondary structure of the designed contrast agents after purification is investigated using far UV CD. Conformational change upon metal binding with the physiological (such as Ca\(^{2+}\), Mg\(^{2+}\), K\(^{2+}\), and Zn\(^{2+}\)) and lanthanide metals (e.g. La\(^{3+}\), Tb\(^{3+}\), and Gd\(^{3+}\)) can also be examined by this method.

The ProCA1 variants and its wild type have two distinct tryptophan residues (i.e. Trp 7 and Trp 32). Trp 32 is buried within the hydrophobic core and Trp 7 is exposed on the hydrophilic surface of rat CD2.D1. This known structural characteristic is used to monitor the tertiary conformational analysis of the ProCA1 variants after the purification method. The tryptophan fluorescence emission signal is used to detect the folding of the variants with indication of the blue shift of the fluorescence maximum at the specified wavelength (350 nm) due to the buried tryptophan. The well-folded CD2.WT protein has a short wavelength emission maximum at ~ 325 nm where a blue shift occurs from the standard free L-tryptophan emission at 350 nm. The ProCA1 variants have a metal binding pocket and specific competition assays are developed for the analysis of metal binding interaction of the ProCA1 variants to the lanthanide metals (Tb\(^{3+}\), Gd\(^{3+}\)). Commercially available dyes are used as probes for detection of the binding affinity for lanthanide metals of the engineered binding proteins. The determined dissociation
constant can indicate whether the engineered binding pocket is functioning properly in terms of the stability of the metal-protein complex. Accordingly, the goals of this Chapter are to examine conformational properties, determine the metal binding affinity, and the relaxivity values of protein variants after purification.
4.2 Methods

4.2.1 Intrinsic tryptophan fluorescence emission spectra

The secondary conformation of all ProCA1 variants was studied through the measurement of the tryptophan fluorescence emission. Measuring the tryptophan fluorescence emission signal will determine whether the purified protein holds a compact globular structure. The emission intensity peak between 300 - 350 nm demonstrated the folding of this protein. The hydrophobic core of the protein is intact if a blue shift is demonstrated at a lower wavelength (~325 nm). This result assured that the protein holds a stable secondary structure. The parameters involved exciting at 280 nm and measuring the emission at a range of 300 nm to 500 nm.

4.2.2 Far Ultraviolet Circular Dichroism

The CD experiment was measured using Jasco-810 spectropolarimeter. All samples were prepared at a set concentration with a stock volume (300 uL - 1 mL) then placed in the 1 mm path length quartz cuvette. Spectra were measured from a starting wavelength at 260 nm and a final at 190 nm with an accumulation of 5 or 15. To determine the percent composition of the secondary structure of the experimental spectra, each protein spectrum were fitted through a set of analysis programs (e.g. SELCON3, CONTINLL, CDSSTR, K2D) provided by DICHROWEB server [65-66]. The best graphical output correlates with the percent beta strand and helix results from the protein data bank (Pdb) file of CD2.WT, (1hng). All exported raw data were converted using Equation 4-1 from mdeg to mean residue ellipticity (M.R.E.) as a function of the wavelength (nm). The best fitting was averaged in terms of the percentage in helix, strand, and turns for the secondary structure analysis determined by each program. CD2.WT is used as a control (Figure 4.5).
\[ \theta \text{ (deg} \times \text{cm}^2 \times \text{dmol}^{-1}) = \frac{\text{Ellipticity (mdeg)}}{I \text{ (cm)} \times C \text{ (M)} \times n \times 10} \]  
\text{Equation 4-1} 

, where \( I \) is the path length, \( C \) is the [ProCA], and \( n \) is the residue number of the ProCA.

### 4.2.3 ProCA1 variants binding to Tb\(^{3+}\) through an NTA-Tb\(^{3+}\) buffer system

Metal binding capability of the protein of interest was analyzed through a series of metal binding assays through the use of a spectrophotometer (photon technology international, (PTI) FEliX32 software) but the initial determination of the appropriate dynamic range for free terbium (Tb\(^{3+}\)) is essential for the correct binding affinity between the ProCA1 variants and the lanthanide metal (Figure 4.7). Once the dynamic range of free Tb\(^{3+}\) is established for ProCA1, the metal binding behavior towards Tb\(^{3+}\) is examined for each ProCA1 variants at the exception of ProCA1.affi342 due to the low stock concentration available. A buffer system involving nitrilotriacetic acid (NTA) chelating agent is used to compete with ProCA1-Tb\(^{3+}\) complex. Decreased fluorescence intensity occurs during this buffer titration system. The dissociation constant of free terbium to ProCA1 variant is determined based on the fractional change or relative fluorescence intensity, \( \Delta f \), as a function of the log base concentration of free Tb\(^{3+}\) (M) (Equation 4-2).
We use the buffer system because it assures that a reasonable dissociation constant can be determined after the data processing through elimination of background noise. There are two sets of buffer conditions used for the NTA-Tb\(^{3+}\) buffer system. Buffer A had 20 mM PIPES, 100 mM KCl, at pH: 6.8, with 500 μM NTA, 500 μM Tb\(^{3+}\) and 20 μM ProCA1 variants. Buffer B is the same sample as buffer A with the exception of no protein present in solution. Buffer B is an experimental control where no distinct change in fluorescence intensity is observed during each titration points. This assures that the actual decrease in fractional change is due in part to the release of free Tb\(^{3+}\) in solution as ProCA1-Tb\(^{3+}\) complex formation decreases. The 100 mM NTA stock is titrated into both buffer systems simultaneously with the varied dilution factors for each titration points. Fluorescence emission was measured in the light region, and thus a range from 200 nm to 600 nm was expected for the fluorescence spectra. Both sample buffer A and B was scanned at 280 nm excitation and 500 nm to 600 nm emissions. A decrease in fluorescence intensity is observed at 545 nm that is the emission of Tb\(^{3+}\) total present in the sample. Tb\(^{3+}\) free in the sample is calculated based on the difference between Tb\(^{3+}\) total and Tb\(^{3+}\)-NTA total in sample buffer A (Equation 4-6). The Tb-NTA total in the sample during each titration points is represented as the fractional change during the titration process (Equation 4-5). NTA

\[
\Delta F = \frac{F - F_{\min}}{F_{\max} - F_{\min}} = \frac{[M]_{\text{Free}}}{[M]_{\text{Free}} + K_d}
\]

Equation 4-2

\[
\Delta f = \frac{m_1 m^2}{(m_1 m^2 + m_0 m^2) m_3}
\]

where \(m_1\) is IC\(_{50}\) or Kd (i.e. 10\(^{-12}\)), \(m_2\) is binding mode (1:1), \(m_3\) is maxima intensity

Equation 4-3
has a dissociation constant ($K_d$) of $10^{-12}$ M for Tb$^{3+}$ and the $K_d$ of ProCA1-Tb$^{3+}$ complex has been previously determined to be at $10^{-12}$ M. So the decrease in fluorescence intensity is normalized as the fractional change or relative fluorescence intensity (Equation 4-3) and plotted against the log based value of the calculated free Tb$^{3+}$. The determined dissociation constant is represented as $m_1$ or the IC$_{50}$ based on the hill Equation for the curve fit of the data points (Equation 4-5). Please refer to Dr. Shenghui Xue’s dissertation for further details of this procedure.
\[ K_{d_{Tb-ProCA1}} = \frac{[Tb^{3+}]_{\text{Free}} \cdot [NTA]_{\text{Free}}}{[Tb^{3+}-NTA]} \]  

Equation 4-4

\[ x = [Tb^{3+}-NTA]_{\text{Total}} = \frac{b \pm (b^2 - 4a^c)^{1/2}}{2a} \]

; where \( x \): fractional change (f), \( a = [NTA]_{\text{Total}} \), \( b^2 = ([Tb^{3+}]_{\text{Total}} + [TA]_{\text{Total}} + K_{d_{Tb-NTA}})^2 \), \( c = [Tb^{3+}]_{\text{Total}} \)  

Equation 4-5

\[ [Tb^{3+}]_{\text{Free}} = \frac{K_{d_{Tb-NTA}} \cdot [Tb^{3+}-NTA]}{[NTA]_{\text{Free}}} = [Tb^{3+}]_{\text{Total}} - [Tb-NTA]_{\text{Total}} \]  

Equation 4-6
4.2.4 FLuo-5N competition assay through fluorescence spectroscopy

Another competition assay used for determination of the dissociation constant of ProCA1 variants to Gd$^{3+}$ is the use of Fluo-5N dye, where the ProCA1 variants are competing with the dye to bind to Gd$^{3+}$ in a specified buffer environment with 100 mM KCl at low pH. We devised a competition assay with the use of a commercially available dye named FLuo-5N [67]. The dye acts as a chelator for Gd$^{3+}$, and protein of interest is titrated into the buffer system containing the dye-Gd$^{3+}$ complex. Through titration, the intensity decreases as the ProCA1 variants compete out the Gd$^{3+}$ from the dye. Essentially, the relative fluorescence intensity is plotted as a function of the increased protein concentration throughout the titration process. The curve fit for the processed data points was executed through Kaleidagraph software using Equation 4-7 based on a 1:1 binding mode. The apparent $K_d$ is initially measured and represents the combined intermediate complex form between Fluo5N----Gd$^{3+}$ - ProCA1. The determined $k_{app}$ is $m_2$ and all known variables needed for determining the actual $k_d$ is calculated through Equation 4-8. The dissociation constant ($K_d$) for Gd$^{3+}$ binding to ProCA1.affi1907 or CD2.7E15Q is calculated based on the known $k_d$ of the dye to Gd$^{3+}$ and the apparent dissociation constant ($K_{app}$) from the curve fit using Equation 4-7. Further details of this method can be found in Natalie White’s thesis.

\[
\Delta f = 1 - \{ m_3 \cdot \frac{[(m_1 + m_2 + m_0) - ((m_1 + m_2 + m_0)^2 - (4 \cdot m_1 \cdot m_0)]^{1/2}}{2 \cdot m_1} \}
\]

where $m_3$: [M]$_{Total}$, $m_1$: [P]$_{Total}$, $m_2$: $K_{app}$, $m_3$: maximum Intensity (i.e. $1 \times 10^6$ counts, 1/sec)  

\[
K_{d2} = K_{app} \cdot \frac{K_{d1}}{K_{d1} + [Fluo - 5N]_T}
\]

Equation 4-7

Equation 4-8
4.2.5 Relaxivity measurement

The $T_1$ and $T_2$ relaxation time delay were measured using the 60 MHz relaxometer (minispec ProFiler, Bruker Corporation). The relaxation rate is determined based on the following formula:

$$R_{1,2} = \frac{\left(\frac{1}{T_{1,2 \text{Sample}}} - \frac{1}{T_{1,2 \text{ Buffer}}} \right)}{[Gd^{3+} \text{ or Contrast Agent}]}$$

Equation 4-9

where [Gd$^{3+}$] symbolizes the gadolinium metal concentration and [Contrast Agent] represents the measured $T_1$ and $T_2$ values of the buffer only (i.e. 10 mM HEPES at pH: 7.0).
Figure 4.1 The schematic representation of NTA-Tb³⁺ buffer system.

a) The NTA sample is added subsequently (μL increments) into buffer A and B. The NTA chelator has a Kd: 10⁻¹² M towards Tb³⁺. As a result, the complex form of the protein-metal concentration decreases in buffer A while NTA-metal complex increases in concentration in both buffer A and B. Both sample buffer A and B is excited at a fixed 282 nm while the maxima intensity is emitted at 545 nm. b) The mechanism in buffer A is demonstrated as excess NTA causes a decrease in the Tb³⁺ ’s fluorescence intensity.
Figure 4.2 The schematic representation of FLuo-5N competition assay.

a) The ProCA1 sample is added subsequently (μL increments) into buffer A. As a result the complex form of protein-metal concentration increases while FLuo-5N-metal complex decreases in concentration in buffer A.  
b) The excitation is fixed at 494 nm while the maxima intensity is emitted at 515 nm for the mechanism that occurs in buffer A during the measurements of FLuo-5N dye’s fluorescence intensity.
4.3 Results and Discussion

4.3.1 The conformational analysis of CD2 mutants using Far UV-CD

After expressing and purifying the protein of interest discussed in Chapters 2 and 3, specified techniques are applied for further structure and functional analysis of the purified sample of that protein. Specifically, in this Section, the structural characteristics of the ProCA1 variants are analyzed through far ultraviolet-visible circular dichroism and tryptophan fluorescence emission spectra.

The 7E15Q (pET-20b, tag-less) demonstrates a different CD spectra compared to 7E15Q (pGEX-2T , GST-tag) (Figure 4.3 a). The two proteins come from different purification methods, and the results are surprisingly different in terms of minima values. 7E15Q expressed as a fusion with GST demonstrates similar spectra to the CD2.WT also expressed as a fusion with GST. This phenomenon suggests that the conformation is different for ProCA1 variants purified through a tag system compared to tag-less purification. There are some factors to consider for a plausible explanation of this occurrence, which are as follows: a) the purification method, b) the protein concentration, c) purity of the buffer system and or the purity of the protein samples. As mentioned previously, the same protein may maintain a different folding conformation suggesting the drastic change in CD spectra through different purification methods.

4.3.1.1 Metal binding influence on the secondary structure of 7E15Q and ProCA1

The example of the CD spectra overlay demonstrates the conformational change observed during exposure of varying metal ions (Figure 4.4). The physiological metal does not change the secondary structure of 7E15Q (Figure 4.4 a). The holo and apo form of 7E15Q have no change in the secondary structure when comparing the spectra overlay. To confirm these findings, a CD curve fit of the raw data is fitted through the use of an analysis program named CONTINLL. These results are
summarized in Table 4.1. CD2 WT (x-ray) refers to the pdb file results of the percent secondary conformation, which is 1 % helix and 53 % strand. The results indicate that 7E15Q (pET20b, tag-less method) as apo form is at 99 % helix while the holo form with Ca^{2+} has 2% helix and 49 % strand conformation.

There is drastic change in the minima at ~ 210 nm -215 nm, compared to the minima at ~222 nm in the presence of both lanthanide and physiological metals for ProCA1.affi1907 (Figure 4.4 b). The highest minima at the 215 nm wavelengths indicate conformational change in the ProCA1.affi1907 secondary structure with Gd^{3+} and Mg^{2+}. There is a similar decrease in ellipticity for ProCA1.affi1907 in the presence of Zn^{2+} and Ca^{2+}. The drastic change in ellipticity minima is due in part to the high metal concentration and low pH 6.8 value. Metals dissociate easily in water in the presence of low pH, and this acidic environment is considered during the analysis of the secondary structure of ProCA1.affi1907. The radius size of the particular metal ion is essential for appropriate comparison as well as the solubility in the presence of protein solution. Only terbium ion demonstrated precipitation during the presence of the protein sample, which explains the small change in ellipticity compared to the other metal ions. One consideration is the determination of the solubility product constant during sample preparation of the stock metal ions, in addition to a smaller protein to metal ion concentration ratio of 1:2 instead of the 1:200 ratio for a more distinct change in ellipticity minimum.
Figure 4.3 Comparing the secondary structure of CD2.WT versus ProCA1 variants versus ProCA1.affi1907 through Far UV-CD.

Each CD spectrum represents an average of 15 CD scans placed on a) the prepared samples of 50 μM 7E15Q, 20 mM EGTA and the expressed 7E15Q/pGEX-2T demonstrate similar minima to CD2.WT/pGEX-2T at ~216 nm representative of the secondary β-sheet conformation in a 1 mm cell at a far UV wavelength (200-260 nm). b) The far ultraviolet CD spectra overlay of ProCA1.affi1907-2 mM EGTA demonstrates minima at a range between 200 nm to 230 nm with CD2.WT as a control.
Figure 4.4 The effect of the metal ions on the secondary structure of a) 7E15Q and b) ProCA1.affi1907.

a) There is a slight change in mean residue ellipticity (M.R.E., deg*cm$^2$/dmol) values between the overlaid spectra, which indicates no significant β-sheet conformational change imposed by the protein-calcium solution. b) Comparing the M.R.E. value for ProCA1.affi1907 (10 μM) based on the presence of the varying metals (2 mM), which is 1:200 ratio based on the protein to metal concentration.

Table 4.1 A summary of the percent secondary structure construct of CD2.WT, 7E15Q and ProCA1.affi1907 targeted to EGFR based on the far UV CD spectra.

Each protein sample has the corresponding percent helix, strand, turns, and unordered. These values are generated from CONTINLL analysis program. The determined percent values that correlate to the known values provided by the pdb file of CD2.WT (X-ray) are highlighted in yellow.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Helix %</th>
<th>Strand %</th>
<th>Turns %</th>
<th>Unordered%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD2.WT (X-ray)</td>
<td>1</td>
<td>53</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>CD2.WT</td>
<td>2</td>
<td>48</td>
<td>21</td>
<td>N/A</td>
</tr>
<tr>
<td>7E15Q/pET20</td>
<td>99</td>
<td>0</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>7E15Q/pET20_Ca$^{2+}$</td>
<td>2</td>
<td>49</td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td>7E15Q/pGEX-2T</td>
<td>2</td>
<td>83</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>EGFR</td>
<td>0</td>
<td>70</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EGFR-Ca$^{2+}$</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EGFR-Mg$^{2+}$</td>
<td>0</td>
<td>38</td>
<td>0</td>
<td>62</td>
</tr>
<tr>
<td>EGFR-Zn$^{2+}$</td>
<td>0</td>
<td>53</td>
<td>42</td>
<td>5</td>
</tr>
<tr>
<td>EGFR-Tb$^{3+}$</td>
<td>98</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>EGFR-Gd$^{3+}$</td>
<td>99</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
4.3.2 Intrinsic tryptophan emission of 7E15Q and ProCA1.affi342 targeted to HER2/neu

The tryptophan fluorescence emission demonstrates a small blue shift present compared to the free-tryptophan emission spectra for ProCA1.affi1907 (Figure 4.6 d). This phenomenon demonstrates that the sample has a partially exposed hydrophobic core in the non-denatured environment. In addition, there is a similar maxima emission to the free L-tryptophan emission spectrum at ~352 nm for ProCA1.affi1907 exposed in 8 M urea and 6 M guanidine hydrochloride. Although the UV-Vis spectra indicate the similar maxima absorbance at 280 nm, there is no indication that the ProCA1.affi1907 holds a compact hydrophobic core based on the tryptophan emission spectra at ~348 nm.

The fluorescence tryptophan emission is similar for ProCA1 (pET-20b tag-less) and ProCA1 (pGEX-2T, GST-tag) (Figure 4.6 a,b). These results suggest that the tertiary conformation of the protein is intact due to the blue shift phenomenon. The shift at the shorter wavelength (~325 nm) demonstrates a tertiary conformation representative of the buried tryptophan residue in the hydrophobic core. Some samples are treated with denaturant and used as a control and compared with the sample in the non-denatured environment. This technique is applied to ensure that a well-folded protein is present in the purified solution. Guanidine hydrochloride or urea at high concentrations causes the 350 nm tryptophan emission similar to the L-Trp free that is used as a control. Further tests are considered for analysis of the CD spectra based on the different methods of purification applied to the protein of interest. One-dimensional NMR can be considered as a spectroscopic technique for further work on the structural conformational analysis of these ProCA1 variants.
Figure 4.5 shows the intrinsic tryptophan emission spectra of ProCA1 and ProCA1.affi342 compared with the free Trp in 10 mM HEPES, pH 7.0.
The fluorescence intensity spectra overlay for tryptophan emission at ~325 nm of the buried Trp 32 in a) ProCA1/pET-20b and b) pGEX-2T demonstrates a blue shift compared to L-Trp free at 350 nm.
4.3.3 NTA-Tb³⁺ buffer system fluorescence titration

The charged variants are analyzed through a two-set buffer system where NTA is used as the chelator that competes with ProCA1-Tb³⁺ complex for binding with the lanthanide metal ion. This chelator possesses a smaller dissociation constant ($10^{-12}$ M) for Tb³⁺ compared to ProCA1 variants. Ideally, the binding affinity of the ProCA1 variants toward the lanthanide ion should be small at the nano- to pico-molar range. After analyzing the binding capability of the ProCA1 variants with saturated terbium concentration at a wide range (Figure 4.7), the dynamic range for the free terbium concentration is determined within 300 μM to 1000 μM of the total terbium concentration for effective dissociation constant determination using the Hill Equation. This reasoning is the basic principle behind the determination of the specified free terbium present in solution, which is in the specific dynamic concentration range of $10^{-11}$ M and $10^{-13}$ M during a curve fitting.

CD2.7E15N and CD2.7E15Q have an assumed -4 net binding pocket charge based on the two aspartic acid and two glutamic acid residues in their designed metal binding site. These mutations differ at one site of the binding pocket, and the metal binding capability for Tb³⁺ ion is analyzed through a devised competition assay with the use of a fluorometer. Terbium ion exhibits a fluorescence emission maxima at ~ 545 nm through the energy transfer from the tryptophan residues from 7E15N or 7E15Q that emit at 325 nm, where the fluorescence intensity increases. Buffer B demonstrates the NTA -Tb³⁺ complex formation spectra, and buffer A displays the NTA competing with Tb³⁺-ProCA1 variants complex in solution during acquisition of each buffer samples. The intensity is decreased gradually as NTA is titrated into the 7E15N-Tb³⁺ or 7E15Q-Tb³⁺ complex samples (Figure 4.8 e,f). During this process, NTA competes out the Tb³⁺ from the ProCA1 variant’s binding pocket. At these decreased titration points, there is less fluorescence energy exchange between the tryptophan residues and Tb³⁺ ion. Tb³⁺ free concentration increases in solution and forms complex with the high NTA present in the sample A buffer. As a result, there is decreased fluorescence intensity. Theoretically, the more NTA required
competing out the Tb$^{3+}$ from the ProCA1 variant’s binding pocket, then the stronger binding the ProCA1 variants demonstrate for binding to Tb$^{3+}$. Both 7E15N and 7E15Q demonstrate a similar curve fit that means that both mutations demonstrate a similar binding characteristic to Tb$^{3+}$.

In following, 7E15E and ProCA1.affi.EGFR have a -5 charged binding pocket. Both variants demonstrate similar binding characteristic to Tb$^{3+}$ just as the -4 charged variants. The curve fit for these -5 charge variants are similar and demonstrates a lower binding affinity then the -4 charged variants. Essentially, the -5 charged variants demonstrate a stronger binding affinity to Tb$^{3+}$ based on the results summarized in Table 4.2. The spectra for buffer A and buffer B demonstrate the gradual decrease in fluorescence intensity during an increase in NTA concentration. Free Tb$^{3+}$ concentration essentially decreases as NTA-Tb$^{3+}$ concentration increase, while 7E15E-Tb$^{3+}$ and ProCA1.affi.EGFR-Tb$^{3+}$ complex decreases in concentration during the titration process.
Figure 4.6 The determination of the free terbium concentration (μM) range between NTA-Tb\(^{3+}\) and ProCA1-Tb\(^{3+}\) buffer system. The competition assay consist of a) buffer system b: 50 mM HEPES, 150 mM KCl, 1 mM NTA and b) buffer system a: 50 mM HEPES, 150 mM KCl, 1 mM NTA, 30 μM 7E15N-Tb\(^{3+}\) buffer system based on the fluorescence emission spectra overlay of the total terbium concentration (μM).
Figure 4.7 Fluorescence emission for 7E15E, ProCA1.affi.EGFR, 7E15N, and 7E15Q in NTA-Tb^{3+} buffer system. The measured fluorescence intensity of a) buffer B: 20 mM PIPES, 100 mM KCl, 500 μM Tb^{3+}, pH: 6.8, no protein presence as baseline correction for 7E15E; buffer A: 20 mM PIPES, 500 μM Tb^{3+}, pH: 6.8, with protein at 20 μM for 7E15E; alongside the curve fit for the -5 charged binding pocket proteins: c) 7E15E and d) ProCA1.affi.EGFR, and the -4 binding pocket proteins: e) 7E15N and f) 7E15Q, based on the Hill Equation.

* The calculated dissociation constant values (K_d) are from the averaged three individual samples (three trials).
4.3.4 Fluo-5N competition assay

In the case of determining the binding affinity of ProCA1 variants toward Gd$^{3+}$ ions, we devised a competition assay that uses the Fluo-5N dye during titration of the protein in solution. Fluo-5N dye is used as a probe that binds to Gd$^{3+}$ at a low $10^{-12}$ M binding affinity for Gd$^{3+}$. The protein is titrated in the buffer system that contains the dye-Gd$^{3+}$ complex. Fluo-5N dye gives emission maxima at ~515 nm and the mean relative intensity is plotted as a function of the protein concentration (μM). The intensity decreases gradually as the protein competes out Gd$^{3+}$ from the dye-Gd$^{3+}$ complex until saturation occurs where there is unchanged fluorescence intensity. Only dissociation constant for 7E15Q and ProCA1.affi1907 are demonstrated using the competition assay. Both variants differ in binding affinity due to the values obtained for the $K_d$, which is the dissociation constant of ProCA1 variant to Gd$^{3+}$. The apparent dissociation constant demonstrated by the curve fit for both 7E15Q (2.53 x $10^{-5}$ M) and ProCA1.affi.EGFR (1.81 x $10^{-5}$ M). The dissociation constant of 7E15Q-Gd$^{3+}$ (9.61 x $10^{-11}$ M) and ProCA1.affi.EGFR-Gd$^{3+}$ (6.87 x $10^{-11}$ M) complex formation is calculated through Equation. ProCA1.affi.EGFR shows a stronger binding affinity towards the Gd$^{3+}$, which is an essential criterion as a Gd$^{3+}$ based contrast agent compared to the charged 7E15Q. Since similar charged variants demonstrated the same binding characteristics towards Tb$^{3+}$, then it is presumed that charged variants with a -5 binding pocket can demonstrate a similar characteristic for Gd$^{3+}$ ion.

4.3.5 The summary of the metal-binding characteristics for all ProCA1 variants

The summary of these findings on the variants, charged and targeted, is compared in terms of binding dissociation constant for both Tb$^{3+}$ and Gd$^{3+}$ (Table 4.1). The summarized binding study establishes that the -5 and -4 charged variants demonstrate a strong binding affinity for both Tb$^{3+}$ and Gd$^{3+}$. The specific site mutation to changing the net charge in the binding pocket assures that results charge of the binding pocket does play an essential role in the metal-binding interaction based on the
summarized results. It is safe to assume that the high charge-binding pocket is favorable for stronger lanthanide ion binding. Furthermore, the ProCA1.affi.EGFR is the only targeted variant, which demonstrates the lowest dissociation constant for both Tb$^{3+}$ and Gd$^{3+}$. Thus, the designed charged variant with a -5 charge and design targeting capability is favorable for use as a contrast agent.
Table 4.2 A summary of the dissociation constants of Tb\(^{3+}\) and Gd\(^{3+}\) toward the ProCA1 variants [22, 34].

<table>
<thead>
<tr>
<th>Protein (net charge of the binding pocket)</th>
<th>(K_{\text{app}}) (M)</th>
<th>(K_{d2}) Gd(^{3+}) (M)</th>
<th>(K_{d,\text{Tb}^{3+}}) (M) *</th>
</tr>
</thead>
</table>
| CD2.7E15N (-4)                            | N/A                    | \(1.48 \pm 0.56 \times 10^{-12}\) \(*\) &nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&n
| CD2.7E15Q (-4)                            | \(2.53 \pm 1.40 \times \text{10}^{-5}\) \(*\) | \(9.64 \pm 5.31 \times \text{10}^{-11}\) \(*\) &nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&n
| ProCA1 (-5) \(*\)                          | N/A                    | \(2.31 \pm 0.32 \times \text{10}^{-12}\) \(\cdot\) | \(5.53 \pm 1.11 \times \text{10}^{-6}\) |
| CD2.7E15E (-5)                             | N/A                    | \(1.45 \pm 0.23 \times \text{10}^{-12}\) \(*\) &nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&n
| ProCA1.affi342-m(-5)\(\$\)                | \(9.82 \times \text{10}^{-6}\) | \(1.86 \times \text{10}^{-12}\) | N/A |
| ProCA1.affi1907 (-5)                       | \(1.81 \pm 1.28 \times \text{10}^{-5}\) \(\odot\) | \(6.87 \pm 0.49 \times \text{10}^{-11}\) \(\odot\) | \(7.53 \pm 0.38 \times \text{10}^{-13}\) |

*Results are based on 3 trials (\(\ast\)) or 2 trials (\(\odot\)) with the corresponding standard error (\(\pm\)).

*Bold-italic: results adapted from Natalie White’s thesis (\(*\)) and Jingjuan Qiao’s dissertation(\(\$\)).
4.3.6 Relaxivity measurements for ProCA1 variants

The large-scale expression and the GST-tag purification of 7E15E, 7E15N, and 7E15Q from Chapter 3 are treated with a chelator resin in a dialysis for overnight in order to eliminate calcium and other free ions present in the protein solution. Each protein sample are then collected and saturated with the desired fixed GdCl₃ concentration for relaxivity measurements. Specifically, 7E15E, 7E15N, and 7E15Q have one binding pocket designed to act as a chelating agent for Gd³⁺ ion. These measurements are plotted as a function of the protein concentration. Theoretically, the observed phenomena is a sharp increase in \( r_1, r_2 \) value then a gradual decrease following saturation of the \( r_1, r_2 \) values as the protein concentration increases. 7E15E and 7E15Q demonstrate a similar curve pattern for \( r_1, r_2 \) and \( T_1, T_2 \) as a function of the protein concentration. 7E15N displays a sharp increase in \( r_1, r_2 \) then an elevated increase in these values soon follow with respect to the increase protein concentration (Figure 4.11 e). In the case of the \( T_1, T_2 \) values, 7E15N demonstrates a saturation with respect to the increase protein concentration (Figure 4.11 b). Essentially, the relaxation time values are inversely proportional to the relaxation rate values. Thus, the compared curves for each ProCA1 charged variants should be a reciprocal plot of one another assuming there are no experimental errors. Specifically, CD2.7E15E- Gd³⁺ complex decreases then saturates in \( T_1, T_2 \) values while the same sample complex increases then saturates in \( r_1, r_2 \) values.

As discussed in Chapter 1, section 1.3.3, the ProCA1 variants have one engineered binding pocket with negatively charged amino acid side chains that form an ionic bond with Gd³⁺. Through the formation of this complex in aqueous solution, in the presence of a magnetic field; relaxation time values are measured using a minispec relaxometer with a 1.4 Telsa. The change in relaxivity values is tested based on the one to one ratio of the protein to Gd³⁺ concentration as demonstrated in the bar graph (Figure 4.12). Each value is labeled in correspondence to each protein variant. CD2.WT is used as a
negative control. Since it has no mutations, the parent protein demonstrates the lowest relaxation rate values as expected for ratio 1:1 with a higher relaxivity value. In terms of numbers, there are no significant difference in the relaxivity values of the 1:1 ratio for all three charged variants. Overall, 7E15Q demonstrates the highest \( r_1 \) and \( r_2 \) values for 1:1 ratio based on the values. In following, all the relaxivity values measured based on the 1:1 ratio fall under a similar range for all charged variants. Thus, further analysis is needed for a concise conclusion to be achieved.

ProCA1.affi1907 targeted to EGFR and ProCA1.affi.HER2: 342 targeted to HER2/neu demonstrate similar \( r_1 \) and \( r_2 \) values when comparing these two ProCA1.affibodies. It is expected that these targeted variants display similar relaxivity values due to the fact that they both have the same metal-binding pocket. The results (Figure 4.13) reassure that although the targeting peptide differs in each of these two variants, there is an unchanged effect in the behavior of the binding pocket. In addition, the linker that connects the grafted peptide and the scaffold ProCA1 enables the ProCA1 moiety that belongs to the variant to move freely without being hindered or restricted by the grafted peptide on the c-terminal of ProCA1. Essentially, there is no particular difference in the relaxivity values when the ratio of the protein to Gd\(^{3+}\) concentration is set as 1:1. Thus, an equal ratio of protein and free gadolium is favorable for the ProCA1 variants during preparation of the complex ProCA1-Gd\(^{3+}\) for further \textit{in vivo} study. In addition, a concise dosage amount can be used for both targeted variants during animal experiment.
Figure 4.8 Relaxivity time ($T_{1,2}$, ms) and rate ($r_{1,2}$, mM$^{-1}$s$^{-1}$) for ProCA1 charged variants. The relaxivity values in terms of time a), b), c) ($T_{1,2}$, ms) and d), e), f) rate ($r_{1,2}$, mM$^{-1}$s$^{-1}$) as a function of the concentration of 7E15E, 7E15N and 7E15Q, with a fixed Gd$^{3+}$ (50 μM, 80 μM, 50 μM) respectively.
Figure 4.9 The average relaxivitys, $r_{1,2}$ (mM$^{-1}$s$^{-1}$), of the charged ProCA1 variants with 1:1 ratio concentration with Gd$^{3+}$.

The following conditions in terms of ratio are: ProCA1 variants (30 μM) and Gd$^{3+}$ (30 μM) ----ratio 1:1.
Figure 4.10 The averaged relaxivities, $r_{1,2}$ (mM$^{-1}$s$^{-1}$), of targeted ProCA1.affibody variants at varying protein to metal ratio concentrations.

The following conditions in terms of ratio are: ProCA1.affi.HER2:342 (50 μM), ProCA1.affi.EGFR (50 μM), and Gd$^{3+}$ (50 μM).
4.4 Conclusion and Future work

The designed ProCA1 variants are analyzed for their structural beta sheet conformation and functional characteristic as contrast agents through a series of biophysical applications. Such instruments consist of far UV CD, intrinsic tryptophan fluorescence emission for the secondary and tertiary conformation. In following, the functional characteristics as a lanthanide chelator are demonstrated through two sets of fluorescence competition assays and the relaxivity measurements. The results assure that after the purification method, the ProCA1 charged variant, 7E15Q holds its secondary β-sheet conformation. Specifically, the structural analysis of 7E15Q purified from different purification methods demonstrates a difference in minima and predicted helix and strand percentage (Figure 4.5) for the far UV CD spectra. This phenomenon indicates that there may be some secondary conformational change with proteins purified with the tag-less versus GST-fusion purification procedure for 7E15Q. Further analysis is needed for the other ProCA1 variants in order to make a concise conclusion about the effect each purification method has on the proteins secondary conformation. The intrinsic tryptophan fluorescence emission spectra indicate that through different purification systems, ProCA1 holds an intact hydrophobic core. Thus, the difference in purification system has no drastic effect on the tertiary conformation of ProCA1. It can be presumed that other charged variants demonstrate this similar stability in tertiary conformation because all demonstrate similar relaxivity values. Therefore, there is no significant difference in the metal binding characteristics to Gd$^{3+}$ when measuring the relaxivity values between the three charged variants (7E15E, 7E15N, and 7E15Q). Similarly, 7E15Q and ProCA1.affi.EGFR demonstrate dissociation constant towards Gd$^{3+}$ at $10^{-11}$ M range. It is presumed that the charged variants and targeted ProCA1.affi.EGFR variant demonstrate similar binding characteristic towards Gd$^{3+}$ based on the relaxivity and FLuo-5N-Gd$^{3+}$ fluorescence measurements. Interestingly, Tb$^{3+}$ dissociation constant is similar for CD2.7E15Q and 7E15N that have a -
4 binding pocket charge, while 7E15E and ProCA1.affi.EGFR demonstrate lowest $K_d$ towards Tb$^{3+}$. The fluorescence emission spectra for FLuo-5N competition system need to be repeated since a large standard error is demonstrated on the curve fit. Thus, the calculated $K_{app}$ and $K_{d2}$ are considered for re-evaluation. In addition, further experiments for determining the other ProCA1 charged variants would be executed as future work. This goal is set towards making a concise conclusion on the impact that the different charges in the binding site has on the dissociation constant for Gd$^{3+}$. 
5. PURIFICATION AND CHARACTERIZATION OF PROCA32 FROM A LARGE SCALE FERMENTATION

5.1 Introduction

As discussed in Chapter 1, Yang lab has developed a novel set of designed ProCAs with the goal to proceed into the preclinical and clinical application. ProCA32 is a part of the 3\textsuperscript{rd} generation of these engineered contrast agents and holds two canonical EF hand-binding sites (contained in helices EF and CD) and one additional EF-hand motif AB, which holds a mutation. The parent protein scaffold is parvalbumin, which is a calcium binding protein. The natural function of this protein is to act as a cytosolic Ca\textsuperscript{2+} buffer [68]. This protein differs with ProCA32 by one site mutation within the CD binding pocket (S56D). Furthermore, an additional mutation (F103W) is made close to the CD binding pocket for effective fluorescence emission analysis. The two binding sites on the parent protein and the advantage of using this scaffold protein will effectively increase relaxation property. The characteristic of this scaffold protein demonstrates stable serum stability, has low toxicity and behaves as a ratiometric T\textsubscript{1}/T\textsubscript{2} weighted contrast agent. The interaction with gadolinium and other metals are tested through a series of biophysical applications in Dr. Shenghui Xue’s dissertation. His results suggest that this contrast agent has a high metal selectivity with gadolinium. This protein scaffold holds a stable conformation during high temperature conditions. The advantages of this third generation of ProCA, specifically ProCA32 mutation, are directed for large-scale production. Specifically, the cultivation of this protein is run through a GMP-like procedure that involves the fermentation process of a 15-liter culture volume. A portion of the cell pellets generated from this over-expression is analyzed through relaxivity measurements and NTA-Tb\textsuperscript{3+} buffer system. Initially, the fermented ProCA32 is purified by a set protocol that efficiently utilizes streptomycin antibiotic that causes precipitation of unwanted contaminants such as nucleotide fragments following cell lysis. The purpose of this Chapter is to successfully purify the
fermented protein and analyze the metal-binding characteristic of this contrast agent through NTA-Tb$^{3+}$ buffer system and relaxivity measurements.

Figure 5.1 The cartoon of the holo form rat-alpha parvalbumin with the indicated two EF-hand sites. This image is generated using PyMOL software with the Pdb code: 1S3P (pink), and 1RWY (yellow) [30]. The EF-hand binding sites are indicated by the red arrows with Ca$^{2+}$ represented as green sphere.
5.2 Methods

5.2.1 Cloning

The cloning of ProCA32 variant is described according to the protocol in Dr. Shenghui Xue's dissertation [69].

5.2.2 Purification procedure of ProCA32

The following are the steps to the purification of the 180 g cell pellet collected from fermentation: Thaw cell pellet (~ 180 g), Aliquot into 6 individual 50 mL falcon tubes, Add 40 mL PBS or preferred buffer (i.e. 10 mM HEPES, pH: 7.0 or 10 mM Tris). Add 90 μL of PMSF (100 mM stock concentration) and benzonase nuclease (1 μL). Suspended cell pellet in this lysis buffer by vortex or stirring prior to sonication. Once a homogenous mixture was achieved– sonication (5X) each sample for 30 sec. with 5 min interval time--- obtaining a lysis solution. The extraction of the protein was done by lysis solution with cell disruptor for three times (3X). Centrifugation was set at a high speed (17,000 rpm) then boiled the supernatant (SN) in a water bath for ten minutes. Centrifuged the solution along with the white precipitant formed at low speed (6,000 rpm) using the bench-top centrifuge (MR 1812) for ten minutes. Kept the supernatant and discarded the pellet, then added 3 % streptomycin to precipitate DNA and nucleic acid fragments. The solutions with streptomycin were placed at 4 °C for overnight. Centrifuged solution at 6,000 rpm for ten minutes. Heated the sample by placing the supernatant in a water bath for another ten minutes, and centrifuged again for an additional ten minutes. Collected supernatant and transferred into dialysis bag and placed in 10 mM HEPES buffer overnight (change buffer twice). This procedure was developed and repeated as described in Dr. Shenghui Xue's dissertation [69].
5.3 Results and Discussion

5.3.1 The purification of ProCA32 from expression through fermentation

As mentioned earlier, ProCA32 is fermented at a large 15 L volume but the purification is based on the 2 Liter aliquot of the cell pellets. The aliquot pellet holds approximately 180 g of overexpressed protein. One fraction of these cell pellets is broken down through a series of steps for the purification. In Figure 5.1, the 15 % SDS-PAGE summarizes each sample of the soluble and insoluble form of each samples related to the series of steps in the purification protocol. The expression is fairly high as indicated by the band intensity. The initial step in purifying the protein is the separation of the soluble and insoluble form through centrifugation. ProCA32 has a relatively high ratio in the supernatant versus inclusion body, which means that the larger ratio of the protein is present in the supernatant samples as demonstrated in lane 1, 3, 5-10. All lanes containing the soluble form of the protein indicate an intense band. In the second and proceeding steps for the purification of ProCA32, heating and treatment with antibiotic called streptomycin is key for removal of unwanted proteins and nucleotide fragments. ProCA32 is a uniquely devised protein that withstands a stable structural conformation in the presence of high temperature. Thus, boiling the sample is an efficient and practical manner for removal of unwanted and unstable protein. Further isolation with the nucleotide and other impurities are achieved through FPLC- anion exchange Q column with a salt gradient program. The higher the percent salt concentration increases, the lower affinity the ProCA32 has with binding to the Q column and through competition with the high ion present in the buffer system, the protein of interest is eluted out by 2 mL fractions.

Initially, the FPLC system runs the set program that initiates the injection of the protein sample. The milli-absorbance units (mAU) of the peak fractions are detected by the UV- lamp at 280 nm, once the sample passes through the instrument system. These peaks are viewed in the FPLC program as a function of the buffer volume (mL). A 10 mL injection volume is set for each run to avoid oversaturation
of the 5 mL Q-column, mA readout, and increase resolution on the FPLC. An overall set of 37 runs is achieved for ProCA32 as a result. The initial FPLC run demonstrates 5 peaks. Each fragment is analyzed for the detection of ProCA32 through 15% SDS-PAGE results (Figure 5.5). Peak two and three are the main regions in which the protein of interest is present between the 35% to 75% salt concentration. The DNA gel demonstrates the successful isolation of nucleotide fragments post FPLC in peak five. Peak two and three are kept separated until all the fraction runs are gathered.

The UV-Vis absorbance spectra for the fractions combinations (15-17, 19-21, 23-31) indicate the presence of ProCA32 based on the 280 nm maxima and the distinct shoulder (this characteristic has been also observed in Dr. Shenghui Xue’s study) at ~295 nm (Figure 5.3). The 18th FPLC run has a higher resolution as demonstrated by the separated middle peak two and three. As we go through each consecutive FPLC run, the model chromatograph indicates four distinct peaks where each are eluted out based on the salt gradient. It’s presumed that nucleotide contamination is present in the highest salt concentration based on the agarose gel demonstrating the high concentration of DNA for fractions 42 to 48 (Figure 5.5). Thus, peak four contains nucleotide contamination that has been successfully isolated from the protein sample. ProCA32 sample (400 mL) from dialysis took approximately nine days to isolate by FPLC.

The collection of SDS-PAGE demonstrates the samples for each combined fractions throughout all the 37 FPLC runs. All the highlighted fragment combinations indicate that a significant concentration of ProCA32 is present after running the FPLC. The common peak fractions for ProCA32 mainly fall in the fragment number 17 to 31 as indicated by the protein bands at the ~12 KDa mark. Each samples are run as an effective way of screen which combine fractions contains the clear protein band at the ~12 KDa mark. All combined fragment samples are then individual scanned through a UV-Vis spectrophotometer for comparing the level of nucleotide contaminants possibly present after FPLC isolation. The 260 nm to 280 nm absorbance maxima for each combined fractions and the corresponding FPLC run number are
quantified and compared based on a ratio in Figure 5.5. The gel samples demonstrated in Figure 5.4 correlate to the same samples with the compared absorbance ratio from Figure 5.5. The identification of the level of nucleotide contamination through UV-Vis absorbance is used as a way to quantify the approximate concentration level of nucleotide, other than using the DNA agarose gel. Essentially, all the combined fractions that initially indicated the presence of ProCA32 are further analyzed by 260 nm/280 nm absorbance ratios prior to ultrafiltration. The numerator represents the absorbance maxima at 260 nm, representing the nucleotides, while the denominator of the ratio represents the absorbance maxima at 280 nm for ProCA32. Since all sets of the combined fractions demonstrate a higher ratio to 0.50, the set ratio range to be analyzed is between 0.50 and 1.50. In following, the individual set of combined fractions is separated into two categories based on the ratio level that is lowest or highest to 1.50. Overall, ten of the twenty-nine combined fragment sets indicated ratios within the specified range. These samples are combined and concentrated under inert gas by ultrafiltration to a final 20 mL volume. The overall yield obtained from after concentrating the initial 400 mL volume of dialysis ProCA32 is 768 mg per ~2L (Figure 5.6). Thus, theoretically, 384 mg/L is effectively isolated and purified within a month span. Furthermore, the purified samples are analyzed for metal-binding characteristic.
Figure 5.2 The purification gel of ProCA32 based on the purification method from Shenghui Xue’s dissertation. The 15% SDS-PAGE contains samples taken throughout the purification process of the fermented ProCA32. Each lane are numbered and are represented as follow: 1) Supernatant after cell disruption, 2) inclusion body after cell disruption, 3) supernatant after heating for 10 minutes, 4) inclusion body after heating for 10 minutes, 5) supernatant after adding streptomycin from sample before overnight and 6) after overnight, 7) supernatant after cell disruption for the 2\textsuperscript{nd} suspended pellets, 8) supernatant after adding streptomycin for the sample after overnight, 9) supernatant after dialysis twice overnight for 1\textsuperscript{st} and 10) 2\textsuperscript{nd} pellet suspension.
Figure 5.3 The isolated DNA/nucleotide contamination is successful as indicated by the agarose gel. Each specified fractions in the red box indicate the labeled samples 1 to 5. The sample fraction labeled in peak 5 demonstrates a strong DNA/nucleotide contaminant that is eluted last at the highest salt gradient percentage (~ > 70%).
Figure 5.4 The UV-Vis spectra indicate successful isolation of the protein of interest within the corresponding peak fractions. Run #18 is displayed above by the FPLC anion Q- column exchanger. The injection volume total into the Q Hitrap column is 10 mL of the ProCA32 solution. ProCA32 is mainly present in fractions 15 to 31, as the small shoulder (at ~ 295 nm) exhibits possible protein in solution. The two-distinct peaks 2 and 3 represent the holo and apo form of ProCA32, respectively. Fractions 32 to 36 do not exhibit this phenomenon.
Figure 5.5 The 15% SDS-PAGE of the isolated ProCA32 to the other contaminants. The nucleotide fragments demonstrated in lane 5 of a) the agarose gel. In following, the 15% SDS-PAGE indicates the combined fragment sets in correspondence to FPLC b) run #1-4, c) run #5-8, d) run #8-11, e) run #12-16, f) run #16-19. Each highlighted red combined fractions indicate the presence of ProCA32 at ~12 KDa.
Figure 5.6 Comparing the absorbance ratio between each fraction range for ProCA32. The A260 nm/A280 nm ratio is higher when the presence of nucleic acid is at a higher maxima value at 260 nm compared to the maxima absorbance value at 280 nm for ProCA32. The diluted samples are indicated with the dilution factor as (800 μL of total Protein and buffer to 50 μL) for Absorbance maximum at 280 nm. The highlighted red fragment sets represent the specified ratio range between 0.50 and 1.50 for the lowest possible amount of DNA/ProCA32 concentration in terms of absorbance maxima values.
Figure 5.7 The final yield corresponding to ~ 180 g from the 2 liter aliquot pellet of fermented ProCA32.
5.3.2 The study on the metal-binding capability of ProCA32

5.3.2.1 NTA-Tb$^{3+}$ buffer system of ProCA32

Two sets of biophysical techniques are used on the concentrated ProCA32 sample after the purification process described in Section 5.3.1. The metal binding capability of the site mutation in the EF hand site and the other binding site of ProCA32 are characterized based on the dissociation constant through an NTA-Tb$^{3+}$ buffer system. It is essential to characterize the Tb$^{3+}$ binding capability of the fermented ProCA32 and compare the findings to previous work established by Dr. Shenghui Xue. The protein stability is demonstrated by the low dissociation constant at $10^{-13}$ M range. Specifically, ProCA32 sample demonstrates dissociation constant ($5.82 \times 10^{-13}$ M) much weaker than the previous ($1.21 \times 10^{-23}$ M) determined results Dr. Shenghui Xue. There maybe two factors to consider for understanding the difference in dissociation constants. One plausible reason is that there are different buffer systems used, the previous results used 50 mM HEPES while the current results demonstrated (Figure 5.7) utilized a 20 mM PIPES buffer. The second reason is the difference in Tb$^{3+}$ and NTA concentration between the current result that uses 500 μM and the previously determined $K_d$ for ProCA32 used 1 mM for both NTA and Tb$^{3+}$. In a sense, the difference in NTA and Tb concentration is most reasonable since theoretically the determined $K_d$ is approximately half the $K_d$ value of that determined in Dr. Shenghui Xue ‘s dissertation. In addition, the protein concentration used is 20 μM compared to the 30 μM concentration of ProCA32 used in the previous studies. Thus, the dissociation constant observed is plausible for determining the binding capability of the two binding pockets to Tb$^{3+}$. 
Figure 5.8 Fluorescence emission for 7E15Q and ProCA1.affi.EGFR in Fluo-5N competition assay. The fluorescence emission spectra overlay for buffer system: 20 mM PIPES, 100 mM KCl, pH: 6.8, a) with 7E15Q or b) ProCA1.affi.EGFR; alongside the curve fit with the normalized relative fluorescence intensity as a function of [7E15Q], μM or d) [ProCA1.affi1907], μM. The determined $K_{app}$ represents the apparent dissociation constant of Fluo5N-Gd$^{3+}$-ProCA1 variant, while the $K_d$ represents the actual dissociation constant for ProCA1-Gd$^{3+}$ complex formation.
Figure 5.9 The fluorescence intensity of NTA-Tb$^{3+}$ buffer system with ProCA32.

The measured fluorescence intensity of a) buffer B: 20 mM PIPES, 100 mM KCl, 500 μM Tb$^{3+}$, pH: 6.8, no protein presence as baseline correction for ProCA32; b) buffer A: 20 mM PIPES, 500 μM Tb$^{3+}$, pH: 6.8, with protein at 20 μM for ProCA32; alongside the curve fit based on the Hill plot.
5.3.2.2 Relaxivity measurements of ProCA32

Relaxivity measurements are gathered as a way to identify the metal binding interaction in the presence of a magnetic field for ProCA32. There are two sets of relaxivity measurements. One experiment demonstrates the high relaxivity values at varying ProCA32 to Gd\textsuperscript{3+} concentration ratio, while the other experiment demonstrates the saturation curve as relaxivity is measured versus the increased ProCA32 concentration (μM range). Figure 5.8 shows the difference in relaxivity values for each set of purified ProCA32 samples. Sample one differs in the expression system used compared to sample two and three. Sample two differs in the purification system used compared to sample one and three. Basically, there is a significant difference in relaxivity values for each of these samples, which indicates that both the expression and purification systems have an impact on the metal binding characteristic of ProCA32. Sample one has a highest relaxivity values for 1:2 ratio compared to the 1:1 ratio. Similarly, sample two and sample three also have a higher relaxivity value for 1:2 ratio rather than the 1:1 ratio. These results coincide with the fact that two binding pockets are involved, thus the Gd\textsuperscript{3+} concentration needs to be increased compared to the protein concentration. Furthermore these results correlate with previous findings that ProCA32 variant demonstrate higher relaxivity values with a higher Gd\textsuperscript{3+} to protein concentration. Theoretically, only one binding pocket is occupied by the Gd\textsuperscript{3+} ion in the 1:1 ratio while at increased Gd\textsuperscript{3+} to ProCA32 concentration, there is an increase chance of both binding sites to be bound to Gd\textsuperscript{3+} in solution. Essentially, the relaxivity depends on the per Gd\textsuperscript{3+} ion present in solution for ProCA32. Sample one demonstrates the highest relaxivity values as compared to sample two and sample three for both ratios.

To proceed, the second set of relaxivity measurements is demonstrated in Figure 5.9. The relaxivity time and the relaxivity are plotted and compared side to side as a function of the increased ProCA32 concentration. Theoretically, as the protein concentration increases, the relaxivity increases up
to the 1:1 ratio then gradually reaches a saturation mark (Figure 5.9b). The relaxivity times ($T_{1,2}$, ms) are inversely proportional to the rate values. A sharp decrease in the time values is observed then an increase and saturation in those values follows, while increasing the protein concentration. The same stock protein sample is analyzed (Figure 5.9) but it demonstrates a lower relaxivity value during the 1:2 ratio. The change in the relaxivity values for ProCA32 sample may be caused by the sample preparation error while using different GdCl$_3$ stock concentration. This reasoning is plausible since the relaxivity values summarized in Table 5.1 are similar to the 1:1 ratio in Figure 5.8 for sample. Thus, the relaxivity values of ProCA32 are much higher with two occupied Gd$^{3+}$ binding sites than that of one occupied Gd$^{3+}$ binding site.
Figure 5.10 The different relaxivity values for ProCA32 in different expression/purification procedure. A comparison between the relaxivity values of the different purification methods for ProCA32 demonstrates a higher relaxivity value for the 60 X 1 Liter flask/streptomycin sample compared to both fermentation samples. The conditions involved are: ratio 1:2 --- 50 μM ProCA32 and 100 μM Gd^{3+}, ratio 1:1 --- 100 μM ProCA32 and 100 μM Gd^{3+}. 
Figure 5.11 The saturated a) $T_1, T_2$ (ms) values and $r_1, r_2$ (mM$^{-1}$s$^{-1}$) values for the fermented ProCA32 with a fixed Gd$^{3+}$ (50 μM).
Table 5.1 The summarized relaxivity values and $K_d$ value for the metal binding study for ProCA32.

<table>
<thead>
<tr>
<th>Protein</th>
<th>$K_{d,Tb^{3+}}$ (M)</th>
<th>$r_2$ (mM$^{-1}$s$^{-1}$)$^0$</th>
<th>$r_1$ (mM$^{-1}$s$^{-1}$)$^0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProCA32$_{fermented}$</td>
<td>$5.82 \times 10^{-13}$</td>
<td>$27 \pm 0.54$</td>
<td>$20.6 \pm 0.20$</td>
</tr>
</tbody>
</table>

$^0$ The mean relaxivity rate values at 1:2 ratio (ProCA32, 50 $\mu$M : Gd$^{3+}$, 100 $\mu$M) from Figure 5.10b.
5.4 Conclusion and Future work

There is successful purification of ProCA32 from a large cell amount (~ 180 g = 30 X 1L when considering that 6 X 1L = ~36 g). The 400 mL total of ProCA32 is obtained after dialysis to diffuse out all unwanted streptomycin and nucleic acid derivatives. Total volume achieved is consistent when considering the overall volume in which the protocol involves (2 x (40 mL x 6 tubes) = ~ 480 mL. The time consuming portion of the purification is isolating the unwanted protein and nucleotide by use of FPLC system. As future work, isolation technique using ion exchange column with a salt gradient is followed for further purification procedure. Agarose gel and UV-Vis absorption are some techniques applied for successful detection of nucleotide derivatives and obtain overall concentration of the sample labeled three. The expression and purification systems do play a major part in the proteins metal binding characteristic as supported by the difference in relaxivity values. The Kd value for Tb$^{3+}$ binding to ProCA32 (5.82 x 10$^{-13}$ M), $r_2$ (27 0.54 mM$^{-1}$s$^{-1}$), and $r_1$ (20.6 0.20 mM$^{-1}$s$^{-1}$) demonstrate the metal-binding interaction between the two binding sites and the lanthanide metals (Gd$^{3+}$ and Tb$^{3+}$).
6. SIGNIFICANCE AND MAJOR FINDINGS OF THIS THESIS

Overall, Yang lab has developed a novel class of protein-based contrast agents (ProCAs) that effectively increase the relaxation rate. Our goal is to fulfill all the necessary criteria involved in achieving the ideal relaxation rate in order to enhance the image. These ProCAs are administered to mice with various diseases models at a milligram amount for in vivo study. To promote the clinical application of ProCAs, the production of the agents with optimized biophysical characteristics and biological function is essential. In order to achieve this final goal of clinical application, there are six main questions that have been addressed in this thesis to determine a reliable and repeatable production protocol for rat derived ProCAs.

A. The Production of ProCAs:

1. What are the optimal expression conditions for the cultivation of each ProCA1 variants in the tag-less system?

2. What are the optimal conditions for the expression system of GST-tagged ProCA1 variants and whether each ProCA1 variant demonstrates a different expression pattern?

3. Whether the total yield (mg/L) of the denaturing-refolding purification method for the tag-less ProCA1s and GST-tagged ProCA1s is comparable or not?

4. How much impact of the lysis buffer condition has on ProCA1 variants’ yield of the GST-tag purification?

B. The structural conformation and metal-binding capability of ProCAs:

5. What is the relationship of metal binding affinity and proton relaxation values with structural conformation of each ProCA1 variant?
6. Does the ProCA32 mutant retain structural stability after a large-scale fermentation? What is the overall effect of the purification and cultivation methods on the proton relaxation property?

I. In Chapter 2, we focused on studying the influence of different *E. Coli* host cell strain on the expression yield of each ProCA1 charged variants including 7E15E, 7E15N, and 7E15Q. We also studied the correlation of the purification yield with the protein expression level.

The major findings for the tag-less expression and purification results indicate that all ProCA1 variants (i.e. 7E15E, 7E15N, 7E15Q) are successfully denatured then refolded with an increased yield (≥ 12 mg/L) from cultured *E.Coli* cell strains (BL21 (DE3) and Tuner). ProCA1 variants have highest expression in Tuner cell strain. Rosetta-gami (DE3) pLysS and BL21 (DE3) pLysS cell strains gave low to no expression of ProCA1 variants at the set IPTG concentration (20 mM) and post induction temperature (25 °C, o/n). Tuner cell strain promotes solubility of the extracted ProCA1 variants during cell lyses, and a future small-scale reproduction confirms reassure this finding based on the set IPTG and post induction temperature.

II. In Chapter 3, we focused on the influence that different *E. Coli* host cell strains, post induction temperatures, and/or inducer concentrations have on the expression yield of each ProCA1 charged variants, 7E15E, 7E15N, and 7E15Q. In addition, we raised the question whether or not the specific modifications in the lysis buffer will affect the purification yield of cleaved GST-tagged ProCA1 variants.

The major findings for GST-tag expression results suggest that the method at low temp with a longer cultivation time for the three charged variant proteins’ (7E15E, 7E15N, 7E15Q) expression level is most favorable while a shorter cultivation time at a high temperature is suitable for the targeted
ProCA1 variants (ProCA1.affi342 & ProCA1.affi1907). These results can be further tested to confirm that the grafted peptide may have some indirect influence on the expression behavior of the targeted ProCA1 variants as compared to the original ProCA1 charged variants. Two different purification systems, tag-less (Chapter 2) versus GST-tag (Chapter 3) were analyzed for the same ProCA1 variants and the yields were compared. The final result demonstrated that tag-less purification through the denatured-refolding process proves most beneficial in terms of the higher protein yield output in mg/L.

In the GST-tag purification, refolding, and modifying the current lysate buffer conditions in A into lysate buffer B is essential for obtaining a high yield as demonstrated by 7E15E (9 mg/L) and 7E15N (4 mg/L). There is some discrepancy to the data because 7E15Q, ProCA1.affi342, and ProCA1.affi1907 demonstrate a highest yield in the respective lysate buffer conditions, A (10 mg/L), C (5 mg/L), and D (7 mg/L). Essentially, the bottleneck for the expression and purification procedures needs to be examined further for the rat derived ProCAs so that these results can be transitioned and compared with the currently expressed and purified human derived ProCAs from results of other current lab researchers.

III. In Chapters 4 and 5, we analyzed the secondary structural conformation, the metal binding study, and relaxation property of both ProCA1 charged and targeted variants, and ProCA32.

Circular dichroism, fluorescence titration, and relaxivity measurement are major applications that are applied for a varied compilation of results on the structural conformation and metal binding behavior of the purified ProCA1 variant samples (Chapter 4). These biophysical applications require a set amount of protein stock concentration as listed in Table 2.1. In the metal binding study, the dissociation constant to Tb$^{3+}$ falls in the $10^{-12}$ - $10^{-13}$ M for -5 charged binding pocket and $10^{-12}$M for -4 charged binding pocket while the dissociation constant to Gd$^{3+}$ is similar for both -4 and -5, due to large error bar. The relaxivity measurements at 1.41 Tesla are consistent with previously collected data, and are to be repeated to determine reproducibility of data.
The significance of this thesis is directed towards the production phase of these novel engineered protein-based contrast agents (ProCAs) for MRI application. One major point is to seek optimal conditions through changing the parameters during expression and purification methods to obtain a high yield, high purity, and stable functioning ProCA chelator through an effective, low cost procedure. It is evident that the long term goal is to set a standard optimal expression and purification method that is favorable for current good manufacturing practice (cGMP)-like, able to transition into an industrial large scale method for application in vivo, specifically for further pre-clinical applications with large animal models (i.e. rats, dogs, monkeys).

The clinical application is to administer this novel designed protein-based contrast agent (ProCA) for MRI. The reasons for aiming our focus at developing MRI based agents are that 1) MRI is non-invasive, 2) has no depth limitation, and 3) provides high spatial and 3 dimensional view of soft tissues in the body. These characteristics set this modality apart from other widely used imaging tools used in medical research. An image enhancement is achieved based on the addition of a substance known as a contrast agent. However, the major barrier to extend MRI for molecular imaging of disease for early detection and prognosis is due to limitations associated to various clinical approved contrast agents such as Gd\(^{3+}\)-DTPA with low relaxivity and undesired pharmacokinetics and the high injection dosage amount (0.10 mmol/kg). Our designed MRI contrast agents with high relaxivity and biomarker targeting capability are expected to overcome the urgent medical needs and extend the application of MRI to early detection of various human diseases, monitor treatment effect and facilitate development of new drug and treatment strategy.
7. REFERENCES


17. EMRIC Experimental MRI Centre. Cardiff University. March 2014. [http://www.cardiff.ac.uk/biosi/researchsites/emric/basics.html](http://www.cardiff.ac.uk/biosi/researchsites/emric/basics.html)


30. PyMOL, *The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.*


64. abcam, *Western Blotting- A Beginner’s Guide.*


69. Xue, S., *Design of Novel Protein-based MRI Contrast agents with high relaxivity and stability for biomedical imaging*, in *Biology*. 2013, Georgia State University: Atlanta, GA.