Summer 8-11-2011

Optimization of Expression and Purification Methods for the Study of Protein-Based Magnetic Resonance Imaging Contrast Agents

Natalie White
Georgia State University, nmaor1@student.gsu.edu

OPTIMIZATION OF EXPRESSION AND PURIFICATION METHODS FOR THE STUDY OF PROTEIN-BASED MAGNETIC RESONANCE IMAGING CONTRAST AGENTS

by

NATALIE WHITE

Under the Direction of Dr. Jenny J. Yang

ABSTRACT

Magnetic Resonance Imaging instruments rely on a contrast agent to provide high-resolution images of tissues *in vivo*. However, current clinical contrast agents are hindered by low relaxivity and fast correlation time, necessitating high injection dosages. These concerns, among others, have driven the development of a class of protein-based contrast agents (ProCAs), by design of lanthanide binding sites into a scaffold protein. ProCA1 has a higher reported relaxivity and dosage efficiency than current contrast agents. In this study, expression and Glutathione-S-Transferase purification procedures were optimized, and a refolding method for rapid production of ProCA1 has been developed to enable studies of conformation, metal binding, relaxivity, and *in vivo* applications. Several ProCA1 variants with 4-5 charged ligand residues were shown to have strong gadolinium binding affinity (K_d of 10^{-12} M) and metal selectivity. Several options to improve ProCA1 have been explored, including addition of a polyethylene chain or a bombesin tag.

INDEX WORDS: Magnetic Resonance Imaging, Contrast agents, Relaxivity, Gadolinium, GST, Refolding method, PEGylation.

OPTIMIZATION OF EXPRESSION AND PURIFICATION METHODS FOR THE STUDY OF PROTEIN-BASED MAGNETIC RESONANCE IMAGING CONTRAST AGENTS

by

NATALIE WHITE

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

Master of Science

in the College of Arts and Sciences

Georgia State University

2011

OPTIMIZATION OF EXPRESSION AND PURIFICATION METHODS FOR THE STUDY OF PROTEIN-BASED MAGNETIC RESONANCE IMAGING CONTRAST AGENTS

by

NATALIE WHITE

Committee Chair: Dr. Jenny J. Yang

Committee: Dr. Giovanni Gadda

Dr. Gangli Wang

Dr. Zhi-Ren Liu

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2011

DEDICATION

There is no way I could have gotten through this program without the constant, continued support from G-d, my parents and husband, my in-laws, my grandparents, my sister, my sister-in-laws, my best friends, and my community. Thank you to my parents, Yaacov and Lauri Maor, for laying the foundation for my love of science since I was a child, for pushing me to take on challenges, and for encouraging me to reach for the stars. Thank you Dovid White, my husband, for helping me fulfill my goals this year and showing dedicated, unwavering support in the face of endless chatter and concern from me, for sitting through all my practice talks, and for keeping me company in the lab. I couldn't have asked for a better partner. To my in-laws, thank you for always having an encouraging word and believing I could do it. To my grandparents, Saba Louis and Ruta, that you for always being supportive. To my sister, Joni Maor, who has listened to me speak about my project for hours during long car rides and dinnertime. To my sister-in-laws for always rooting me on. To my best friends, Tzina Fishman, Shoshana Elon, and Rachel Weintraub, thank you for always sticking by my side, for being my cheerleaders, and for showing interest in my project. I feel immensely thankful to have been surrounded by such a strong support team.

ACKNOWLEDGMENTS

I would like to thank my committee members who have supported me in the past, and continue to guide me. Special gratitude to Dr. Yang for the past four years and for this amazing opportunity. Thank you for allowing me to join the lab during my freshman year in college, investing in me, and making sure no opportunity for a challenge, competition, scholarship, presentation/ poster, or conference passes by. Beyond science, you have mentored my growth as a person throughout these years, and have encouraged me to never fall short of my dreams. That is a tremendous gift, and I owe you infinite gratitude. Thank you to Julian Johnson, Dr. Shunyi Li, Johnny Chen, Dr. Jin Zuo who have given so much of their time to teaching me skills and guiding me throughout my undergraduate years. Thank you to Jen Ngo, with whom the GST studies were conducted as undergraduates. Thank you to Dr. Jie Jiang, David Xue, and Fan Pu for helping me with the fluorescence/ competition titrations and relaxivity and for allowing me to ask a myriad of questions until I understood MRI concepts throughout my graduate studies. Thank you to Dr. Mike Kirberger who has been a mentor and has helped me with fluorescence studies. Thank you to Ada Tang for helping me draw depth from the enzymatic assay data and for studying enzyme concepts together. Thank you to Hing Wong, who ran the NMR samples shown in this thesis. Thank you to the rest of the Yang group (previous and current members), not only for insightful suggestions which have helped me tweak and optimize the data, but also for the many interesting discussions we've had in appreciation of our different cultures and languages. Thank you to the department of Chemistry for opening doors for me in so many ways and for supporting me throughout my studies. I am forever grateful! This work is supported by grants from National Institute of Health (NIH) given to Dr. Jenny Yang.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS……v

LIST OF TABLES……..xi

LIST OF FIGURES……...xii

LIST OF ABBREVIATIONS……..…xxi

CHAPTER 1: INTRODUCTION..... 1

1.1 Principles of Magnetic Resonance Imaging 1

1.2 Contrast Agents are Needed to Overcome Low Sensitivity in MRI..... 9

1.3 Contrast Agent Relaxivity is Influenced by Several Parameters..... 10

1.4 Criteria for Design and Development of High-quality MRI Contrast Agents..... 12

1.5 Design of Protein-Based Contrast Agents 17

1.5.1 Engineered Proteins and ProCA1 17

1.5.2 Choice of Protein Host..... 19

1.5.3 Charged Variants of ProCA1 as Suitable Candidates for Contrast Agents 22

1.6 Expression Systems: Factors Affecting Protein Expression..... 22

1.7 Purification Systems 25

1.7.1 Comparison of Affinity Tag Purification Systems 25

1.7.2 Problems with GST Purification..... 29

1.7.3 Method of Refolding the Expressed Protein without Affinity Tag 31

1.8 Objectives of Thesis..... 33

CHAPTER 2: OPTIMAL CONDITIONS FOR GST TAG EXPRESSION AND PURIFICATION 36

2.1 Background about GST and Affinity Tag in Protein Purification 36

2.2 Materials and Methods 39

2.2.1 Cloning 39

2.2.2 Transformation 39

2.2.3 Expression of Wild-type CD2, CD2.7E15, and CD2.6D31 GST Fusion Proteins 40

2.2.4 Purification of GST Fusion Proteins by Affinity Chromatography 42

2.2.5 Bradford Concentration Assay (BCA) and GST Enzymatic Activity Assay 43

2.2.6 GST Beads Binding Assay 45

2.2.7 SDS-PAGE 45

2.3 Results and Discussion 45

2.3.1 Expression of CD2.7E15, CD2.6D31, and Wild-type CD2 45

2.3.1.1 The Effect of Temperature on Protein Expression 45

2.3.1.2 The Effect of Cell Strain on Expression 46

2.3.2 Estimation of Insoluble and Soluble Protein Ratio of Expressed Proteins by BCA Assay 48

2.3.3 The Relationship Between the GST Enzymatic Activity and Beads Binding Capability of the GST Fusion Protein 50

2.4 Conclusions 54

CHAPTER 3: OPTIMIZATION OF A TAG-LESS PURIFICATION PROCEDURE TO IMPROVE THE YIELD OF ENGINEERED PROTEIN 56

3.1 Introduction 56

3.2 Materials and Methods 57

3.2.1 Cloning 57

3.2.2 Transformation 57

3.2.3 Expression of His-7E15 and 7E15 59

3.2.4 Optimization of His-tagged 7E15 by Adjustment of Temperature and IPTG Concentration at Induction and Purification with Arginine 59

3.2.5 SDS-PAGE gel 60

3.2.6 Purification with Arginine Followed by Denaturing Conditions 62

3.2.7 Purification Using Denaturing Conditions 63

3.2.8 Determining Protein Concentration 65

3.3 Results and Discussion 67

3.3.1 Expression of Tag-less 7E15 67

3.3.2 Optimization of His-7E15 Expression 70

3.3.3 Purification of Tag-less 7E15 with Arginine Solubilization Followed by Denaturing Conditions 74

3.3.4 Purification of Tag-less 7E15 Under Denaturing Conditions 81

3.4 Conclusions and Future Work 87

CHAPTER 4: PROTEIN-BASED CONTRAST AGENTS 89

4.1 Site-Directed Mutagenesis to Create 7E15 Charged Variants 89

4.2 Methods and Materials 91

4.2.1 Cloning and Transformation 91

4.2.2 Over-expression 91

4.2.3 Purification Via Urea Refolding Method 92

4.2.4 Anion Exchange Chromatography Purification Based on Salt Gradient 94

4.2.5 Cation Exchange Chromatography Purification Based on pH Gradient 94

4.2.6 Far Ultraviolet Circular Dichroism 94

4.2.7 Determination of K_d of Charged Variants to Tb^{3+} Using Fluorescence Resonance Energy Transfer (FRET) Assay 95

4.2.8 Determination of Affinity of 7E15 Charged Variants to Gd^{3+} and Zn^{2+} by Competition with Molecular Probes Dyes 97

4.2.9 Relaxivity Measurements 99

4.3 Results and Discussion 99

4.3.1 Expression of CD2.7E15 Charged Variants 99

4.3.2 Purification, Mass Spectra, and Circular Dichroism of Analysis CD2.7E15 Charged Variants 101

4.3.3 Determination of K_d of 7E15 Charged Variants to Tb^{3+} by FRET Titration 107

4.3.4 Determination of Affinity of 7E15 Charged Variants to Gd^{3+} by Competition with Molecular Probes Dye Fluo-5N 113

4.3.5 Determination of Affinity of 7E15 Charged Variants to Zn^{2+} by Competition with Molecular Probes Dye FluoZin-1 125

4.3.6 Relaxivity of CD2.7E15 Charged Variants 131

4.4 Summary and Significance 138

CHAPTER 5: ENHANCEMENT OF PROCA1 BY PEGYLATION AND PEPTIDE TAGGING 140

5.1. Introduction 140

5.1.1 PEGylation 140

5.1.2 Enhancement of CD2.7E15 Contrast Agents with Peptide Tags 144

5.2 Materials and Methods 144

5.2.1 Cloning of Mutants 144

5.2.2 Expression of K3CR, K5R, K5166R, K6691R, K45C, K51C, K66C, and His-CD2.7E15-bom 146

5.2.3 Purification of K3CR, K5166R, K6691R, K51C, and K66C 147

5.2.4 Purification of His-CD2.7E15-bom 148

5.2.5 PEGylation of K51C and K66C 148

5.3 Results and Discussion 149

5.3.1 Expression and Purification of K3CR, K5R, K5166R, and K6691R 149

5.3.2 Expression, Purification, and PEGylation of K51C and K66C 153

5.3.3 Expression and Purification of His-CD2.7E15-bom 162

5.4 Conclusions and Future Work 166

CHAPTER 6: SIGNIFICANCE OF THIS WORK..... 167

CHAPTER 7: REFERENCES..... 170

LIST OF TABLES

Table 1.1 Comparison of affinity tags contrasting capacity and cost to obtain 10 mg 28

Table 1. 2 Summarizes proteins generated for projects throughout this thesis and their variations, the purpose of their variations, and the chapters in which they are discussed..... 34

Table 2.1 Summary showing the expression temperature and cell strain required to produce soluble fusion protein for wild-type CD2 and its variants…….55

Table 3.1 Summary of conditions found in chapter 2 with an added emphasis on conditions for inclusion body formation……...56

Table 3.2 Arginine solubilization temperature relative to concentration of soluble protein shows room temperature is favored for soluble protein……77

Table 4.1 Summary of the residues in 7E15 charged variants; red refers to positive, blue refers to negative, and black refers to neutral amino acids [15]……..90

Table 4.2 Shows absorbance ratio 280/260 nm, molar absorptivity, molar mass, final concentration, and average final yield in milligrams per liter of expressed 7E15 charged variants. 103

Table 4.3 Summary of the approximate K_d values for 7E15 in 10 mM Tris-HCl, pH 7.4 and 7E15N in 20 mM PIPES, 10 mM KCl, pH 6.8. Final protein concentration was 5 µM, excitation wavelength was 282 nm, and emission wavelength was 500-600 nm..... 112

Table 4.4 Summary of the K_d values of 7E15 charged variants to Gd^{3+} (5 µM Fluo-5N- 5 µM Gd^{3+}, in 20 mM PIPES, 10 mM KCl, pH 6.8), Tb^{3+} (5 µM 7E15 or 7E15N, in 10 mM Tris-HCl, pH 7.4 and 20 mM PIPES, 10 mM KCl, pH 6.8, respectively), Zn^{2+} (10 µM FluoZin-1-10 µM Zn^{2+}, in 10 mM Tris-HCl, pH 7.4), and Ca^{2+}. Error bars refer to average standard deviation from

three separate trials. All work concerning Ca^{2+} K_ds can be found in Dr. Anna Wilkins Maniccia's dissertation [37]. 139

T**able 5. 1** Summarizes the different class 1 and class 2 mutations of CD2.7E15 for PEGylation with arginine or cysteine agents. PEG represents the PEGylated residue, R represents the arginine mutations which are spared from PEGylation, and C represents cysteine mutations which are PEGylated. Also, the current status of each protein is summarized……...142

Table 5. 2 Designed forward (for) and reverse (rev) primers for cloning of K45C, K51C, and K66C. 145

LIST OF FIGURES

Figure 1.1 The magnetic moment and the angular momentum for spinning protons (A) Shows the magnetic moment, μ. (B) Shows the angular momentum, J. Taken from [3]. 3

Figure 1.2 Process of Relaxation. (A) Random alignment of hydrogen atoms in the absence of an applied magnetic field. (B) In the presence of an applied magnetic field, B_0, hydrogen nuclei adopt a spin either parallel or anti parallel to B_0. (C) When an RF pulse B_0 is applied, the magnetic moment of the nuclei deviates away from B_0, and (D) subsequently realigns with B_0 following removal of the RF pulse, causing the nuclei give off energy in the form of an RF signal. Taken from [6]..... 5

Figure 1.3 Diagrams illustrate simultaneous, but different processes of relaxivity (A) R_1 relaxivity (B) R_2 relaxivity. Taken from [5]. 8

Figure 1.4 Cartoon illustration of parameters affecting inner and outer/second sphere relaxivity, including water exchange rate, molecular rotational correlation time, electronic relaxation time of gadolinium, and ionic distance between protons and gadolinium; sphere represents water molecule. Taken from [10]. 11

Figure 1.5 Commercial contrast agents approved by the FDA include Magnevist, Dotarem, Omniscan, ProHance, Multi-Hance, and Gadovist. Taken from [13]. 14

Figure 1.6 Representation of contrast agent models. (i) Small chelating molecules such as Gd-DTPA with τ_R exceeding 100 ps. (ii) One or more Gd-DTPA or Gd-DOTA molecules covalently bound to a macromolecule still has fast τ_R due to internal mobility. (iii) Protein with metal binding site provides rigid binding framework with high affinity for gadolinium, a slow τ_R observed for the Gd^{3+}-protein unit as a whole. Taken from [1]..... 16

Figure 1.7 (A) Relaxivity values of gadolinium-bound ProCA1 are significantly higher than relaxivity values of Gd-DTPA. (B) Injection of contrast agent injection into a 26 g mouse produce enhanced MR images of organs (arrows). The left side shows pre-injection, and the right side shows 40 minutes post injection of 50 μL of 1.2 mM Gd^{3+}-CA1.CD2. Taken from [1]..... 18

Figure 1.8 Rat cell adhesion protein CD2.d1 (1hng) is chosen as a host protein for protein-based contrast agent design. Two tryptophan residues at positions 7 and 32 are shown in red. 21

Figure 1.9 Pymol illustration of design of wild-type CD2 variants. (A) CD2.6D31 showing structure of La^{3+}/Ca^{2+} binding pocket formed by four coordinating ligand residues (E29, D31, D41, D43). (B) La^{3+}/Ca^{2+} binding protein CD2.7E15 with five coordinating ligands (E15, E56, D58, D62, D64) (left) has pentagonal bipyramidal arrangement of the metal binding site (right). 21

Figure 1.10 Novagen expression vectors pGEX-2T and pET-20b; pGEX-2T contains a..... 24

Figure 1.11 Models to incorporate affinity tags to target proteins. (1) Fusion protein containing native protein attached to an affinity tag via a linker region which has a sequence for endoprotease cleavage. (2) In this model the affinity tag enhances the solubility of the protein (examples are GST, MBP). (3) Fusion protein with exopeptidase removal of affinity tag. (4) Fusion protein where a solubility and folding region is placed between the native protein and affinity tag. Taken from [22]. 26

Figure 1.12 Expression and purification of GST fusion protein. (A) Shows high yield following expression 5 hours after induction with IPTG. (B) Following purification, a large amount of fusion proteins does not bind to GST beads and passes through the columns (in WST) and a low yield is seen at the fraction which eluted with glutathione. 30

Figure 1.13 Schematic of refolding method..... 32

Figure 2. 1 Glutathione-S-Transferase enzyme exists as a dimer found in eukaryotes, and each monomer is 26 k Da. The active site of GST (enlarged) contains an active cysteine residue at position 10. For efficient binding to the Glutathione-S-Sepharose 4B affinity column, the GST tag on the fusion protein must be well-folded with its thiol group (on cysteine 10) positioned correctly. 38

Figure 2.2 Summary of expression procedure. 41

Figure 2.3 GST Detection Module (GE healthcare) [29]..... 44

Figure 2.4 Expression gels showing the final point 5 hours after induction for each cell strain and temperature, growth curves, and wet cell pellet weights per 2 L expressed cell culture for CD2.7E15, CD2.6D31, and wild-type CD2. (A) Expression gel for CD2.7E15 shows that lower temperatures (25 °C or 30 °C) may be more effective for expression as bands are thicker by 8-10%. (B) Growth curve for CD2.7E15 shows a gradual, steady growth for all cell strains. (C) Wet cell pellet weights is highest at 37 °C especially for Tuner cell strain. (D) Expression gel for CD2.6D31 shows no significant thickness among bands (less than 4% difference) but a more efficient growth with Rosetta pLysS. (E) Growth curve for CD2.6D31 shows a gradual, steady growth for all cell strains especially Rosetta pLysS. (F) Highest weight for cell pellet is seen at 30°C. (G) Expression gel for wild-type CD2 shows thicker bands at 37 °C by 4-9%. (H) Growth curve for wild-type CD2 shows a gradual, steady growth for all cell strains especially. (I) Cell pellet weights are highest at 37 °C excluding Rosetta pLysS..... 47

Figure 2.5 Estimation of inclusion bodies in 1 g cell culture among wild type CD2 and two variants at expression temperatures 37 °C, 30 °C, and 25 °C with cell strains Tuner, Rosetta pLysS, and BL21 (DE3). (A) Shows CD2.7E15 inclusion body level is highest at 37 °C

expression by approximately 10-40% more than in expression at 30 °C or 25 °C. (B) Shows CD2.6D31 inclusion body level is 10-20% higher with use of Tuner than with use of Rosetta pLysS cell strain. (C) In wild-type CD2, the differences in inclusion body production across temperatures and cell strains did not exceed 10%. 49

Figure 2. 6 Shows concentration of CDNB-Glutathione (product) formation over time for the first 100 seconds plotted and fitted linearly. (A) CD2.7E15. (B) CD2.6D31. (C) CD2WT. The slope is equal to the initial velocity (V_0)..... 51

Figure 2.7 The initial velocity (V_0) of GST fusion protein calculated from the slope of the graph of product concentration (mM) versus time (s) (Figure 2.6). (A) Shows that enzymatic activity is highest for CD2.7E15 expressed at low temperatures. (B) Show that enzymatic activity is highest for CD2.6D31 expressed with Rosetta pLysS cell strain regardless of the temperature. (C) Shows that enzymatic activity is highest for wild-type CD2 expressed at 37 ºC regardless of the cell strain. Error bars refer to average standard deviation from three separate trials. 52

Figure 2.8 Beads binding capability assay of GST fusion protein monitored by SDS-PAGE (left) and intensity ratio (right) indicating binding capability is shown to correspond with the GST enzymatic activity. (A and B) In CD2.7E15 the high intensity bands from protein expressed at lower temperatures are thicker than the higher temperatures by up to 80%. (C and D) For CD2.6D31, the highest intensity bands came from proteins expressed with Rosetta pLysS cell strain. (E and F) In wild-type CD2, the highest intensity bands were seen at high temperature expression, which were thicker by approximately up to 70%. 53

Figure 3.1 Plasmid vector map of pET30a. Taken from[30]……58

Figure 3.2 Shows the effect of adding SDS to proteins prior to running through a gel: negative charges from SDS disrupt positive charges by side chains of protein and the protein becomes

negatively charged and linearized, which enables it to migrate to the positive pole during electrophoresis. Taken from [31]. 61

Figure 3.3 Summarizes the purification procedures via non-denaturing and denaturing conditions outlined in sections 2.2.5 and 2.2.6. 64

Figure 3.4 Typical fluorescence absorbance spectra of protein containing tryptophan residues with a maximum at 280 nm and a typical "shoulder" (arrow). 66

Figure 3.5 Expression of 7E15. (A) Growth curve indicative of exponential increase in optical density of BL21 (DE3) pLysS prior to induction and curve levels off post induction. (B) SDS-PAGE showing gradual production of protein with the passage of time, the thickest bands found after overnight growth. 68

Figure 3.6 Expression of 7E15 with different temperature conditions. (A) Expression curve (above) and SDS-PAGE (below) show protein has been expressed at 37 °C for 4 hours. (B) Expression curve (above) and SDS-PAGE (below) showing improved expression when the temperature is lowered to 30 °C and the protein is expressed overnight. 69

Figure 3.7 SDS-PAGE and optical density curve of the final point of expression of His-7E15 at 37 °C (4 hours) with BL21 (DE3) cells. The optical density curve is exponential, yet the gel shows a low level of expression after induction with IPTG concentrations 0.2-0.8 mM. 71

Figure 3. 8 SDS-PAGE and optical density curve of the final point of expression of His-7E15 at 37 °C (4 hours) with BL21 (DE3) cells. The optical density curve shows exponential growth, and the gel shows a thick band indicative of a high level of expression after induction with 1 mM IPTG. 72

Figure 3.9 Purification of His-7E15 by size exclusion chromatography shows protein in fractions 11 and 12 and a DNA peak in fraction 8 (confirmed by DNA gel). 73

Figure 3.10 Purification of tag-less7E15 with arginine solubilization, followed by denaturing conditions. Following cell lysis, protein is in the cell pellet, which is washed with Triton X-100. Protein is dissolves poorly in 2 M arginine and following denaturation is refolded and pure (10 mM Tris samples). 75

Figure 3.11 Arginine solubilizaton samples sent to 1D-NMR..... 79

Figure 3.12 UV spectrum of protein after arginine/ urea refolding method shows reliability of method as protein refolds back to its native conformation resembling CD2. The total yield from 2.3 L expressed cell culture is shown below (257.8 mg).. 80

Figure 3.13 Purification of tag-less 7E15 under denaturing conditions. Protein was found in cell pellet after cell lysis, and was found in the supernatant post-solubilization with urea. The cell pellet after urea shows that solubilization was effective, as little protein remained insoluble. 82

Figure 3.14 Chromatograms from FPLC purification of 7E15 show high yield. (A) Peak obtained in fractions 17-20 from a 10 mL injection into the SP column yields 17.3 mg protein. (B) Peak from a 10 mL injection into the Q column yields 25.9 mg protein in fractions 15 and 16. The arrows from the gel band correspond to the peaks on the chromatogram. 84

Figure 3.15 Structural studies of 7E15 post refolding purification. (A) CD spectrum of 7E15 shows a similar secondary structure conformation as seen in CD2 [36]. (B) Tryptophan fluorescence spectrums of 7E15 purified by different methods. The arginine/urea method (shown in green) and the urea method (shown in red) show tertiary structure resemblance to GST purification (shown in blue).. 86

Figure 4.1 Summary of the urea refolding method……93

Figure 4.2 Expression of 7E15 (top), 7E15N (middle), and 7E15E and 7E15Q (bottom). (A) Optical density curve showing steady increase of cells at 37 °C with tag-less vector pET20b and

induction point. (B) SDS-PAGE gel showing gradual thickening of 7E15-variant bands after induction. 100

Figure 4.3 Purification of 7E15 charged variants with SDS- PAGE gel of refolding method (top left) followed by SDS- PAGE gel and chromatogram of ion exchange chromatography (bottom right)..... 102

Figure 4.4 Mass spectra of 7E15 charged variants obtained with ABI 4800 MADLI TOF/TOF mass spectrometer instrument..... 104

Figure 4.5 Secondary structure studies by CD. (A) 7E15 with EGTA, Ca^{2+}, and Gd^{3+}. (B) 7E15 with EGTA, Ca^{2+}, and Gd^{3+}. (C) 7E15N with EGTA, Ca^{2+}, and Gd^{3+}. (D) 7E15N and 7E15E each with EGTA, Ca^{2+}, and Gd^{3+}..... 106

Figure 4. 6 Cartoon illustrating donor emission (tryptophan) and acceptor absorption (terbium) spectral overlap. Taken from [41]..... 108

Figure 4.7 Fluorescence intensity of Tb^{3+} as a function of wavelength. Titration with 5 µM CD2.7E15 in 10 mM chelex Tris, pH 7.4, excitation at 282 nm, emission 500-600 nm, maximum at 545 nm (fitted to give an approximate K_d in µM). 110

Figure 4.8 Fluorescence intensity of Tb^{3+} as a function of wavelength. Titration with 5 µM CD2.7E15N in 20 mM PIPES, 10 mM KCl, pH 6.8, excitation at 282 nm, emission 500-600 nm, maximum at 545 nm (fitted to give an approximate K_d in µM). 111

Figure 4.9 Chemical structure of Fluo-5N, pentapotassium salt Molecular Probes Invitrogen dye. Taken from [42]. 114

Figure 4.10 Saturation test. Maximum fluorescence intensity peak corresponds to Fluo-5N- Gd^{3+} complex, while minimum fluorescence intensity peak corresponds to Fluo-5N..... 115

Figure 4.11 Titration of Fluo-5N-gadolinium with EDTA. 116

Figure 4.12 Competition assay with 10 µM Fluo-5N and 10 µM gadolinium in 10 mM Tris-HCl, pH 7.4 with CD2.7E15 (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials). 117

Figure 4.13 Average plot of 7E15-Gd^{3+} following assay of 10 µM Fluo-5N to 10 µM gadolinium in 10 mM Tris-HCl, pH 7.4 indicated K_{app} of 1.74 and a calculated K_d of 6.62×10^{-13}..... 118

Figure 4.14 Competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15 (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials)..... 119

Figure 4.15 Competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15N (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials). 120

Figure 4.16 Competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15Q (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials). 121

Figure 4.17 Competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15E (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials). 122

Figure 4.18 Average plots of competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15 and charged variants to determine K_d. (A) CD2.7E15. (B) CD2.7E15N. (C) CD2.7E15Q. (D) CD2.7E15E. 123

Figure 4.19 Overlaid average plot of competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15 and charged variants to determine K_d. 124

Figure 4. 20 Competition assay with 10 µM FluoZin-1 and 10 µM zinc in 10 mM Tris, pH 7.4 with CD2.7E15 (added in 8 µM increments until saturation) to determine K_d; excitation at 495 nm, emission range 500-650 nm (two trials). 126

Figure 4. 21 Competition assay with 10 µM FluoZin-1 and 10 µM zinc in 10 mM Tris, pH 7.4 with CD2.7E15E (added in 8 µM increments until saturation) to determine K_d; excitation at 495 nm, emission range 500-650 nm (two trials). 127

Figure 4. 22 Competition assay with 10 µM FluoZin-1 and 10 µM zinc in 10 mM Tris, pH 7.4 with CD2.7E15N (added in 8 µM increments until saturation) to determine K_d; excitation at 495 nm, emission range 500-650 nm (two trials). 128

Figure 4. 23 Competition assay with 10 µM FluoZin-1 and 10 µM zinc in 10 mM Tris, pH 7.4 with CD2.7E15Q (added in 8 µM increments until saturation) to determine K_d; excitation at 495 nm, emission range 500-650 nm (two trials). 129

Figure 4.24 Overlaid average plot of competition assay with 10 µM FluoZin-1 and 10 µM zinc in 10 mM Tris-HCl, pH 7.4 with CD2.7E15 and charged variants to determine K_d. 130

Figure 4. 25 Graphs of T_1 and T_2 for CD2.7E15 and the charged variants show change in relaxation time with increasing protein to gadolinium ratio in 10 mM HEPES, pH 7, 25 °C, 1.4 T. (A) T1 (right) and T2 (left) graphs of CD2.7E15. (B) T1 (right) and T2 (left) graphs of CD2.7E15N. (C) T1 (right) and T2 (left) graphs of CD2.7E15Q. (D) T1 (right) and T2 (left) graphs of CD2.7E15E. 133

Figure 4. 26 Relaxivity of CD2.7E15 and charged variants with a gadolinium concentration of 0.05 mM, in 10 mM HEPES, pH 7, 25 °C, at 1.4 T. Plot of reciprocal T values against protein concentration is fitted linearly to determine saturation point. The red, dashed line represents the point where the protein to gadolinium ratio is 1:1 (A) CD2.7E15. (B) CD2.7E15Q. (C) CD2.7E15E. (D) CD2.7E15N..... 135

Figure 4. 27 Summary of r_1 and r_2 of CD2.7E15 charged variants in 10 mM HEPES, pH 7, 25 °C, 1.4 T shows comparable relaxivities to each other, which are significantly higher than the relaxivity of Gd-DTPA. 137

Figure 5. 1 Cartoon summarizing the different class 1 mutations of CD2.7E15 for PEGylation with arginine agents relative to the metal binding site of the protein……..143

Figure 5.2 Optical density graphs and expression gels for expression in pET20b and BL21 (DE3) at 30 °C after induction overnight show expression for K3CR, K6691R, and K5166R. (A) K3CR. (B) K6691R. (C) K5R. (D) K5166R. 150

Figure 5.3 Purification of K3CR, K5166R, and K6691R. (A) K5166R. (B)K6691R.(C) K3CR. The gels correspond to the tallest peaks. 152

Figure 5.4 Expression of K51C and K66C with BL21 (DE3) cell strain and pET20b vector at 37°C. (A) Optical density curve indicating is exponential. (B) SDS-PAGE showing presence of protein during expression..... 154

Figure 5.5 Purification of K51C and K66C. (A) K66C. (B) K51C. M-marker; 1- supernatant after French press shows no protein as it is all in the cell pellet; 2- cell pellet after French press shows protein representing inclusion bodies; 3- supernatant after 2% Triton X-100 does not contain protein; 4- cell pellet after Triton X-100 contains cell pellet that is cleaner than before wash; 5-supernatant after solubilization with 8 M urea shows the protein in the supernatant; 6-

cell pellet after 8 M urea shows very little insoluble protein after refolding; 7- supernatant after 10 mM Tris buffer shows clean product ready for injection into FPLC. 156

Figure 5.6 Purification of K51C with SP column (flow rate 2 mL/min) 158

Figure 5.7 Purification of K66C with SP column (flow rate 2 mL/min) 159

Figure 5.8 Iodine-stained gel depicting PEGylation of K51C and K66C. M-marker; 1-K51C before PEGylation; 2-K51C after PEGylation; 3- K66C before PEGylation; 4-K66C after PEGylation with (methyl-PEG_{12})$_3$-PEG_4-Maleimide reagent. 161

Figure 5.9 Expression of His-7E15-bom at 37 °C in BL21 (DE3) cells with 1 mM IPTG at induction 163

Figure 5.10 Size Exclusion chromatography shows almost no product of His-CD2.7E15-bom when fractions 7-9 are visualized on SDS-PAGE. 165

LIST OF ABBREVIATIONS

AI	After induction
AS	After sonication
BAB	Beads After Binding
BBB	Beads Before Binding
BI	Before induction
BL	BL21 (DE3)
Bom	Bombesin tag
CA	Contrast Agent
CD	Circular dichroism
CP	Cell pellet
E.coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
Elu	Elution
FRET	Fluorescence Resonance Energy Transfer
GST	Glutathione-S-Trasferase
His	Histidine tag
IB	Inclusion bodies
IPTG	Isopropyl-β-D-thiogalactopyranoside
K_d	Dissociation constant
LB	Luria-Bertani
NSD	Nephrogenic systemic dermopathy
NMR	Nuclear Magnetic Resonance
MW	Molecular weight
OD	Optical density
ON	Overnight
PEG	Polyethylene glycol
ProCA1	Protein-based contrast agents, class I
r_1/r_2	Relaxivity
RF	Radiofrequency pulse
RP	Rosetta PLysS
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SN	Supernatant
T	Tuner
T-X-100/Wash	2% Triton X-100
Wst	Waste

CHAPTER 1: INTRODUCTION

1.1 Principles of Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is one of the most crucial techniques for medical diagnostic imaging because of its capability to provide high resolution images of soft tissues *in vivo*. MRI is noninvasive and uses no harmful radiation, and is therefore advantageous over other imaging techniques such as X-ray and positron emission tomography. MRI produces three dimensional images useful for diagnostics by recording discrepancies in relaxation time among neighboring, but physiologically different tissues [1; 2].

Magnetic resonance relies on the magnetic properties of the hydrogen atom, which possesses an one proton and has a nonzero spin. This allows hydrogen to interact with and be influenced by both external magnetic fields and radio waves. An MRI scan is able to detect signals from hydrogen atoms most abundantly found in the form of water molecules, unlike other compounds such as ^{12}C, ^{16}O, and ^{14}N, which are also present in biological tissues, but are uninfluenced by MRI [3].

The magnetic properties of hydrogen and other atoms are believed to originate from the spin properties of the nuclei, each characterized by a magnetic moment (μ) and angular momentum (J) related to mass and charge [4]. The magnetic moment describes how much torque an atom can exert on the external magnetic field (Figure 1.1 A), while the angular momentum arises from the spinning particles (Figure 1.1 B). Equation 1.1, in which μ, τ_{max}, and B_0 correspond to magnetic moment, torque, and magnetic field strength respectively, shows magnetic moment relationships [3]:

$$\mu = \tau_{max} / B_0 \qquad \text{(Equation 1.1)}$$

Equation 1.2, in which J, m, v and r correspond to the angular momentum, mass, velocity, and radius, shows the angular momentum relationship [3]:

$$J=mvr \qquad \text{(Equation 1.2)}$$

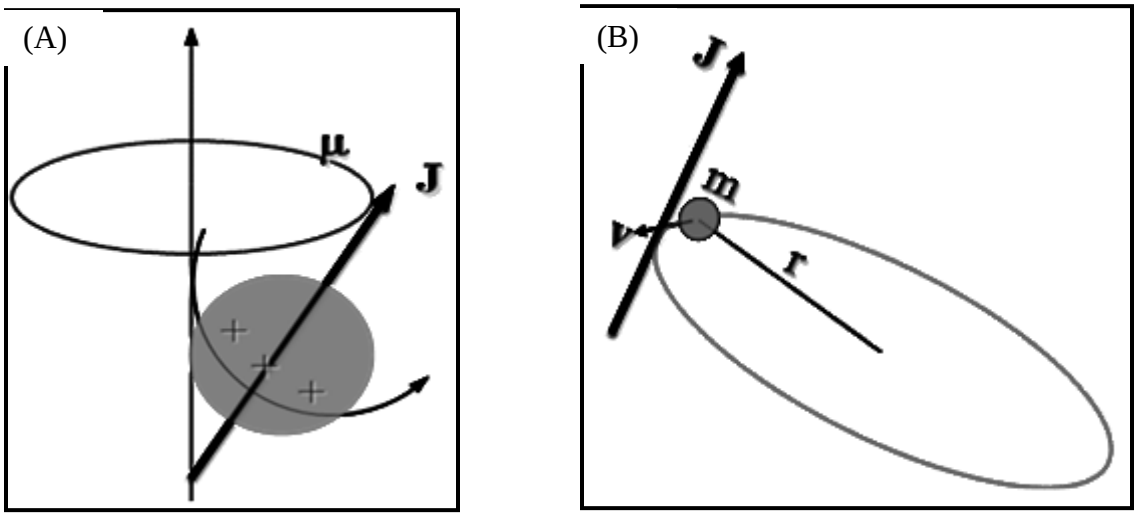


Figure 1.1 The magnetic moment and the angular momentum for spinning protons (A) Shows the magnetic moment, μ. (B) Shows the angular momentum, J. Taken from [3].

Both the proton and neutron in hydrogen each have a spin of $\pm$ ½ (magnetic quantum number); yet the proton has a charge while the neutron has none. When the number of protons equals the number of neutrons, as in most atoms, there is no net magnetization, because the protons and neutrons align in such a way that their spins and magnetizations cancel. But in atoms such as ^{1}H, cancellation does not occur because the unpaired electron produces a constantly changing magnetic field around the atom, due to vibrational and rotational motions [5].

Normally, hydrogen has no net magnetization because the nuclei are in a low energy state with random orientation in the absence of a strong magnetic field (Figure 1.2 A). However, hydrogen nuclei in water molecules will adopt a new axis of rotation that is either parallel or anti parallel to an applied external magnetic field (Figure 1.2 B).

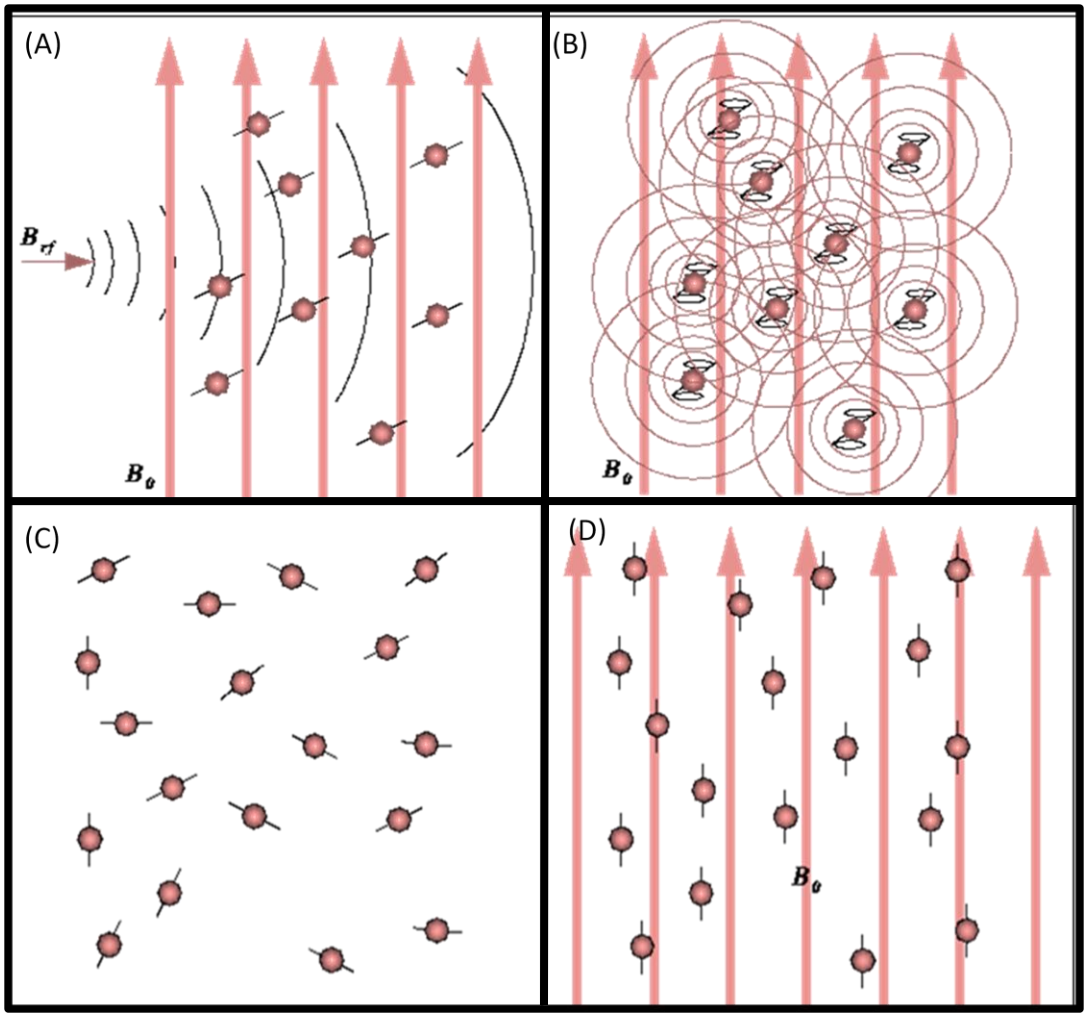


Figure 1.2 Process of Relaxation. (A) Random alignment of hydrogen atoms in the absence of an applied magnetic field. (B) In the presence of an applied magnetic field, B_0, hydrogen nuclei adopt a spin either parallel or anti parallel to B_0. (C) When an RF pulse B_0 is applied, the magnetic moment of the nuclei deviates away from B_0, and (D) subsequently realigns with B_0 following removal of the RF pulse, causing the nuclei give off energy in the form of an RF signal. Taken from [6].

If the sample is then subjected to a pulse of energy with high radiofrequency (RF) perpendicular to the direction of the magnetic field, it will absorb some of this energy, causing the magnetic moment to deviate from the magnetic field (Figure 1.2 C). When the RF pulse is removed, the nuclei begin to precess, or rotate in a motion similar to "[tracing] the walls of a cone" (Figure 1.2 D). The effect of this motion is a release of energy (RF) from the sample as the atoms return from a high energy to a low energy state, the frequency of which is called the Larmor frequency (ν) [5].

The Larmor frequency is directly proportional to the strength of the magnetic field B_0 multiplied by the gyromagnetic ratio γ (a nuclei- dependent constant) [5; 7; 8].

$$\nu = \gamma B_0 \qquad \text{(Equation 1.3)}$$

The nuclei in a sample of ^{1}H emit signals of identical frequency, thus the Larmor frequency is said to depend solely on the strength of the magnetic field. An MRI of a chosen magnetic field strength records the signal emitted as the sample "relaxes", or returns to its initial orientation relative to the magnetic field. Among other atoms with unpaired electrons such as ^{13}C, ^{19}F, ^{23}Na, ^{31}P, ^{1}H is most frequently used in NMR because of its natural abundance, sensitivity, and high Gyromagnetic ratio compared with ^{13}C and ^{15}N [5; 7].

MRI measures the time it takes for the perturbed protons in a sample to reorient with the magnetic field, which varies by environment and tissue type. The process of proton relaxation occurs via two simultaneous mechanisms: spin-lattice (longitudinal) relaxation (r_1) and spin-spin (transverse) relaxation (r_2). The spin-lattice or longitudinal relaxation time (T_1) describes the time it takes for a perturbed proton to return to its original orientation in the magnetic field. The environment in which the nuclei are in, or the lattice, interacts with these high energy nuclei and through vibrational and rational motions causes them to return to a low energy state.

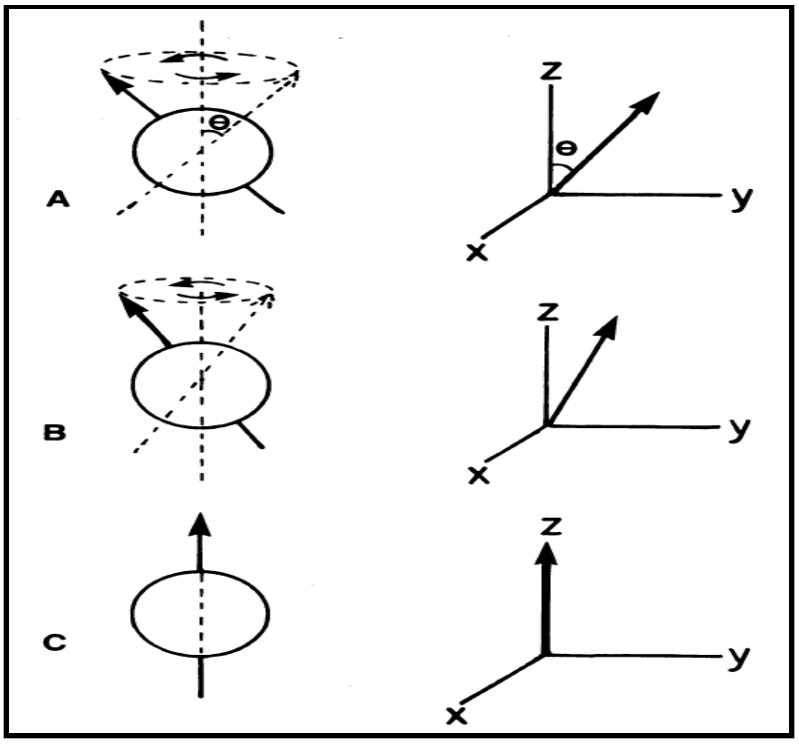
Longitudinal relaxation for a precessing proton is illustrated in Figure 1.3A. The proton is observed to precess at an angle θ following the initial RF pulse. As the proton loses energy gained from the RF pulse, θ decreases and net magnetization realigns with the applied field (*z*-axis). Equation 1.4 describes the time it takes for the magnetization (M) to decay; M_0, t, and T_1 correspond to the initial magnetization, time, and longitudinal relaxation time, respectively [5].

$$M = M_0 (1 - e^{-t/T1}) \qquad \text{(Equation 1.4)}$$

At the same time the nuclei realign with the field, they all resume their natural tendency to spin, but the spinning quickly becomes out of phase. This occurs because although the atoms are precessing at the same frequency their quantum states are different [5]. Spin-spin relaxation time (T_2), known as transverse relaxation time, measures the process of a nucleus going from a high to a low energy state, which is a less prominent signal than the T1 signal. Figure 1.3B shows that as the sample precesses, the magnetization vector begins to separate into vector components in the xy plane, or dephase. Finally the vectors cancel each other out in all directions as the spin-lattice relaxation also decreases. Equation 1.1.5 describes how the intensity (I) of the signal drops as dephasing occurs over time; I_0, t, and T_2 correspond to the initial intensity, time, and transverse relaxation time, respectively [5; 9].

$$I = I_0 (1 - e^{-t/T2}) \qquad \text{(Equation 1.5)}$$

(A)



(B)

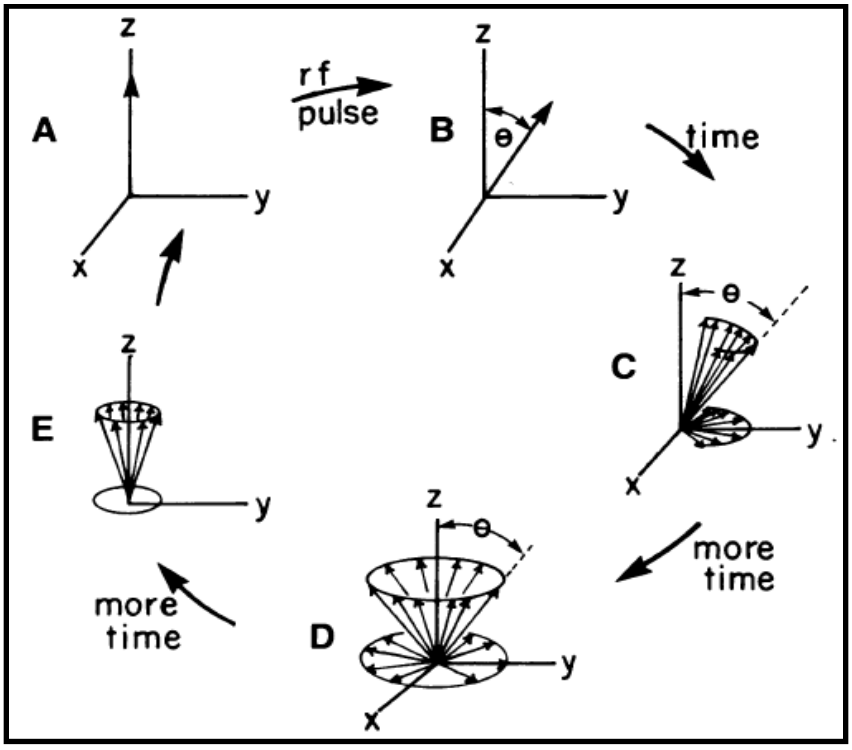


Figure 1.3 Diagrams illustrate simultaneous, but different processes of relaxivity (A) R_1 relaxivity (B) R_2 relaxivity. Taken from [5].

1.2 Contrast Agents are Needed to Overcome Low Sensitivity in MRI

One of the greatest limitations in MRI is low sensitivity. While MRI instruments are capable of attaining images, many rely on a contrast agent to enhance the image considerably and to increase the sensitivity and specificity of MRI [10]. Contrast agents are typically paramagnetic substances, which means they have unpaired electrons capable of being influenced by a magnetic field. A contrast agent is capable of catalytically increasing the relaxation rate of surrounding water molecules by decreasing T_1 and T_2, thereby increasing the degree of contrast on an image. This ability is called the "relaxivity," r_1 or r_2, of a contrast agent. Relaxivity, or how well a contrast agent is able to change the proton relaxation at a 1 mM concentration, is a measure of the effectiveness of the contrast agent [10]; high-quality contrast agents with high relaxivity are able to reduce the signal to noise ratio considerably.

Among the lanthanides, gadolinium has seven unpaired electrons, a high magnetic moment, long electronic relaxation times, and is capable of exerting the strongest effect on T_1 [11]. Gadolinium-based contrast agents are typically small molecules that chelate gadolinium in order to avoid the toxic effects of free metal [1]. After the sample is treated with contrast agent, the relaxivity, relative to the concentration of contrast agent, is shown in Equation 1.6 (adapted from source [8]).

$$R_{1,2} = \frac{\left({}^{1}/_{T_{1,2}Sample}\right) - \left({}^{1}/_{T_{1,2}\,Buffer}\right)}{[Gd^{3+}\ or\ Contrast\ Agent]} \qquad \text{(Equation 1.6)}$$

There are two types of contrast agents; those that that lengthen the relaxation rate of T_1, or T_1-agents (brightening agents), and those that alter the relation rate of T_2, or T_2-agents

(darkening agents). The most common T_1-agent is a gadolinium complex, capable of producing a positive contrast by reducing T_1. T_2-agents include iron-oxide agents capable of producing a negative contrast [10]. Typically, T_1-agents are used because of its availability.

1.3 Contrast Agent Relaxivity is Influenced by Several Parameters

Water molecules in the vicinity of gadolinium are affected by its paramagnetic influence. "The origin of relaxation is the dipole-dipole interactions between the proton nuclear spins and the fluctuating local magnetic field that results from the paramagnetic metal center" [10]. Components that make up the observed relaxivity originate from inner sphere and outer sphere relaxation (Figure 1.4). Protons directly coordinated with the gadolinium center contribute to inner sphere relaxation, while "bulk solvent water molecules" contribute to outer sphere relaxation [10]. Parameters that affect relaxation are water exchange rate (K_{ex} or $1/\tau_m$), radial distance between protons and metal center, rotational diffusion (τ_R), hydration number of the spheres (q), and the electronic properties of the gadolinium ion (T_{1e}). Water exchange between surrounding water molecules and coordinated water molecules must be a fast process for high relaxivity; molecular rotational correlation time must be slow; hydration number must be as high as possible without loss of affinity between gadolinium and its complex. Development of improved contrast agents can be achieved by manipulating these parameters, which are related by Equation 1.7, where $1/T_{c1}$ is equal to the inverse Larmour frequency [10]:

$$1/T_{c1} = 1/\tau_R + 1/T_{1e} + 1/\tau_m \qquad \text{(Equation 1.7)}$$

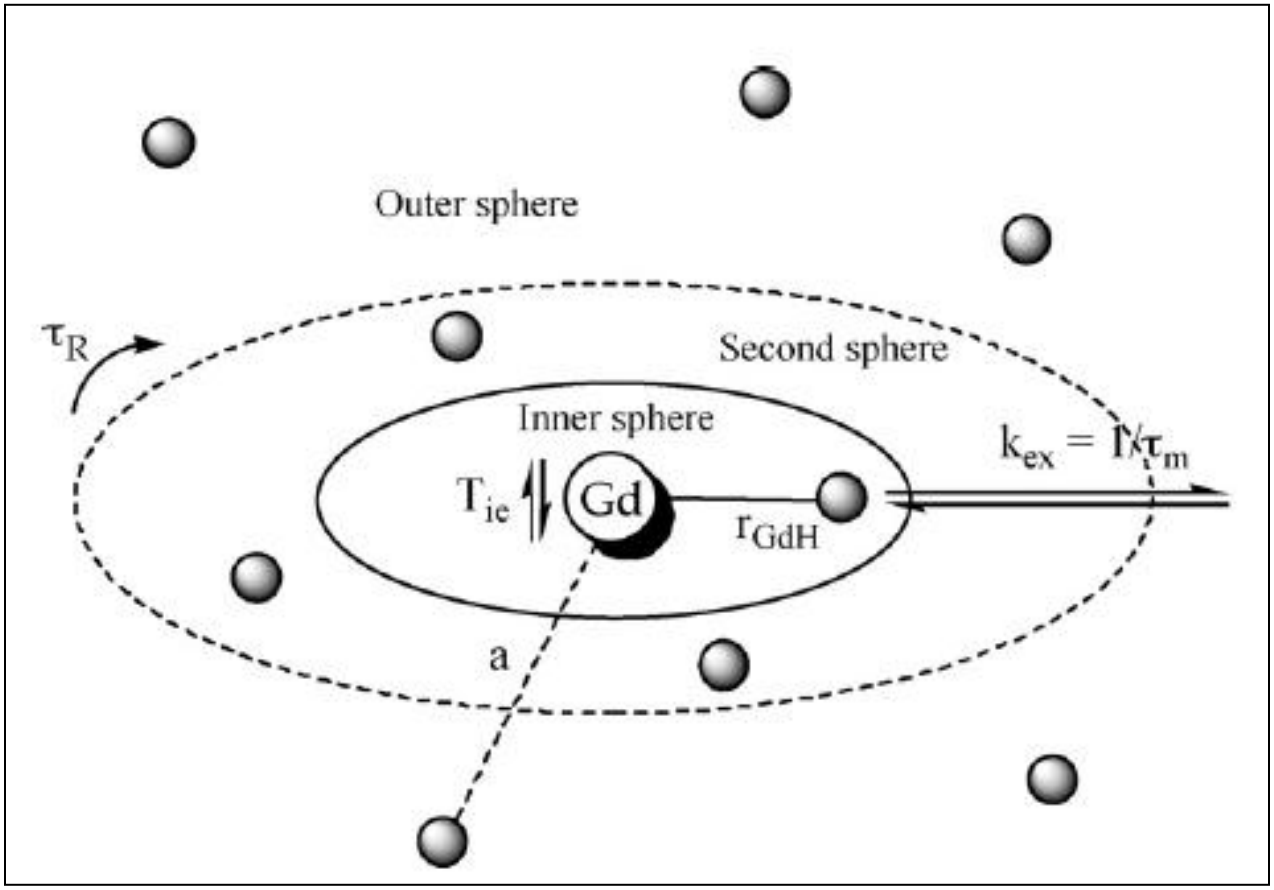


Figure 1.4 Cartoon illustration of parameters affecting inner and outer/second sphere relaxivity, including water exchange rate, molecular rotational correlation time, electronic relaxation time of gadolinium, and ionic distance between protons and gadolinium; sphere represents water molecule. Taken from [10].

Optimization of the water exchange rate, the molecular correlation time, and the hydration number of the inner sphere are ways to improve the effectiveness of contrast capability.

1.4 Criteria for Design and Development of High-quality MRI Contrast Agents

Current clinical agents are small molecules that chelate gadolinium, mostly polycarboxyamides. While they are easily synthesized, their reported relaxivities are 5 $mM^{-1}s^{-1}$ and their effectiveness is noticeable only at high dosages, 300-500 mM. The retention time is short in tissues (less than 30 minutes). Additionally, the sensitivity is low and these agents cannot detect molecular biomarker expression. Therefore, repeated doses are needed to obtain a clear image. Because of their fast molecular correlation time of 100 ps and a slow inner sphere water exchange rate, the relaxivity of commercially used contrast agents, such as Magnevist, are said to be compromised, as the relaxivity falls short of the theoretically possible value of 100 $mM^{-1}s^{-1}$. Also, these contrast agents are considered blood pool agents which means they lack target specificity [1; 2]. In addition, recent concerns have arisen over the renal toxicity of these contrast agents, with increased incidences of nephrogenic systemic dermopathy (NSD) reported after a gadolinium MRI in patients with preexisting kidney conditions [12]. For these reasons there is a pressing need for a new type of contrast agent.

A high-quality contrast agent is sought after, which will display high relaxivity, high affinity for lanthanides over physiological metals, *in vivo* stability, minimal toxicity, longer retention time in tissues for reduced dosage frequency and for sufficient imaging time, and a small-sized chelator for quick renal filtration from the body. Increasing the relaxivity is beneficial in reducing the dosage necessary to provide a high quality image, thus an increase in dosage efficiency. High affinity for lanthanides will prevent binding of physiological metals and prevent toxicity caused by free circulating gadolinium. In addition, they should have strong metal

selectivity over excess amounts of physiological metal ions such as calcium, zinc and magnesium. Lastly, they should be easily produced in large-scale amounts with minimal steps of preparation.

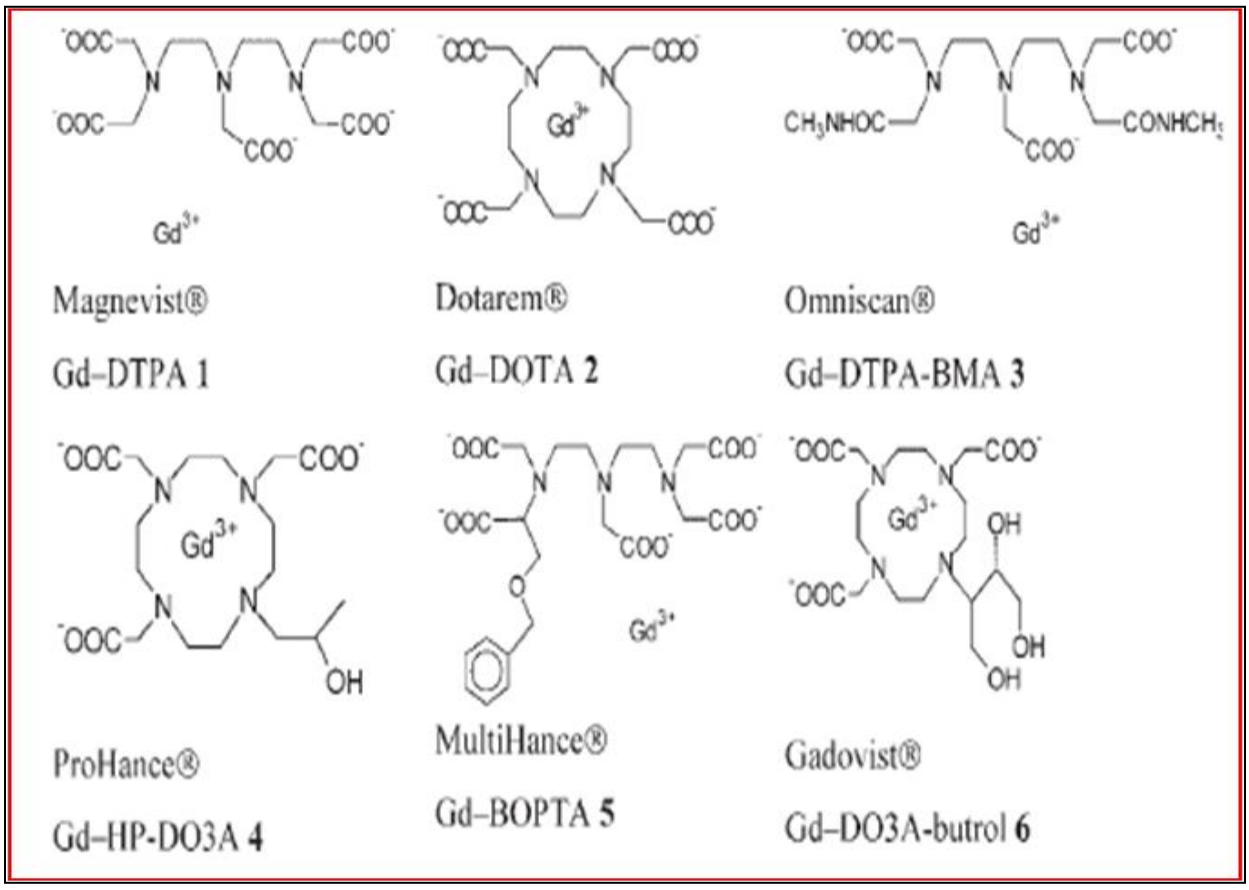


Figure 1.5 Commercial contrast agents approved by the FDA include Magnevist, Dotarem, Omniscan, ProHance, Multi-Hance, and Gadovist. Taken from [13].

As mentioned earlier, one approach to improving the relaxivity is to adjust the parameters that affect it. One parameter that has been addressed in recent research is reducing the fast molecular rotational correlation time of currently used contrast agents (Figure 1.6 i) [8]. The strategy is to covalently attach two chains of Gd-DTPA or Gd-DOTA molecules to larger macromolecules such as linear polymers, dendrimers, liposomes, viral capsids, or proteins. While these molecules in fact exhibit a larger molecular rotational correlation time, the relaxivity has not shown to increase appreciably compared to commercial contrast agents. This is because intrinsic internal mobility is high and the water exchange rate is restricted (Figure 1.6 ii). As a matter of fact, the size of these molecules, 10-100 nm, will further hinder proper filtration by the kidneys, an already existing toxicity concern with gadolinium contrast agents.

The strategy of designing a lanthanide binding site into a protein has been used in our laboratory to create a novel class of contrast agents (ProCA1) that addresses the limitations of current contrast agents. ProCA1 exhibits higher relaxivity, optimized pharmacokinetics and biodistribution, quick renal excretion, and lower dose efficiency. The protein provides a rigid binding framework with high affinity for gadolinium, and limitations such as internal mobility are eliminated; rotational correlation time is observed for the Gd^{3+}-protein unit as a whole (Figure 1.6 iii) [1].

.

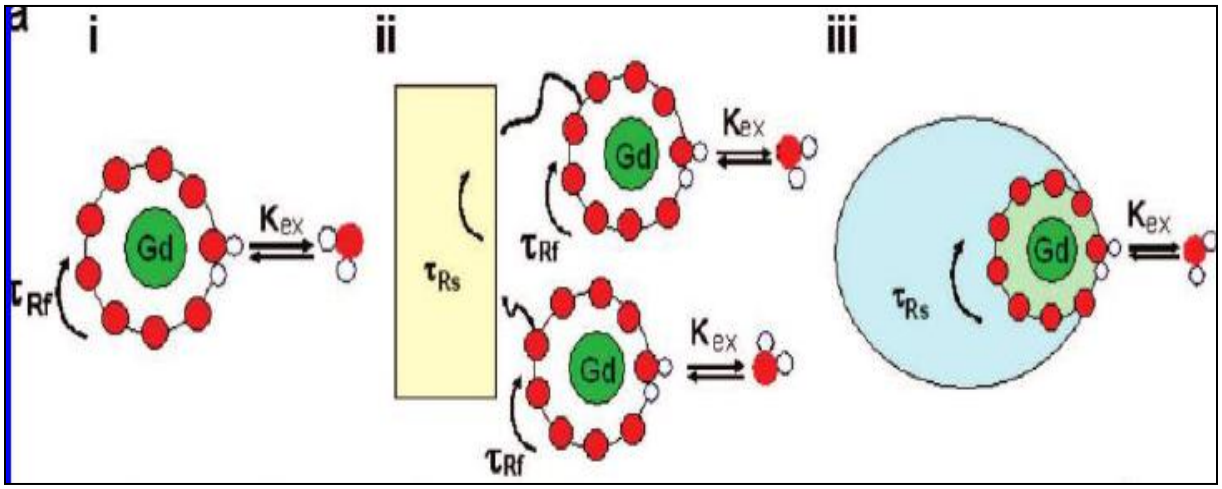


Figure 1.6 Representation of contrast agent models. (i) Small chelating molecules such as Gd-DTPA with τ_R exceeding 100 ps. (ii) One or more Gd-DTPA or Gd-DOTA molecules covalently bound to a macromolecule still has fast τ_R due to internal mobility. (iii) Protein with metal binding site provides rigid binding framework with high affinity for gadolinium, a slow τ_R observed for the Gd^{3+}-protein unit as a whole. Taken from [1].

1.5 Design of Protein-Based Contrast Agents

1.5.1 Engineered Proteins and ProCA1

Engineered proteins are gaining recognition for serving a broad range of interests in research, including elucidating the structure and function relationship of proteins, the development of analyte sensors, protein drugs, vaccines, and diagnostic probes. Dr. Jenny Yang's laboratory has used a novel design approach to introduce a metal binding pocket in the first domain of scaffold protein, CD2.d1, through protein engineering. The result is a class of protein-based contrast agents (ProCA1) that have shown improved relaxivity over Gd-DTPA. In contrast to commercial contrast agents a longer correlation time of 10 ns is observed in ProCA1 rather than 100 ps, because of a larger molecular weight (11.2 kDa versus 743 Da in Gd-DTPA). Gd-ProCA1 has an r_1 of 117 $mM^{-1}s^{-1}$ and an r_2 of 129 $mM^{-1}s^{-1}$ compared with Gd-DTPA which has an r_1 of 5.4 $mM^{-1}s^{-1}$ and an r_2 of 8 $mM^{-1}s^{-1}$ in the same field strength (1.5 Tesla) (Figure 1.7 A). A 40 fold lower dosage than Gd-DTPA is needed for enhancement of kidney images (Figure 1.7 B). ProCAs also has a high metal binding affinity in the pM range, which decreases the risks of free metal toxicity associated with NSD. The small size of CD2.d1, 2 nm, allows for easy renal excretion and intravascular distribution [1].

(A)

CA Class	Compounds (ligand residues)	r_1 ($mM^{-1}s^{-1}$)	r_2 ($mM^{-1}s^{-1}$)	B_0 (T)
Designed proteins	Gd^{3+}-CA.CD2	117	129	1.5
	(E15/E56/D58/D62/D64)	48	88	3.0
		6	50	9.4
Small compound	Gd-DTPA	5.4	8	1.5
		4.2	6.8	3.0

(B)

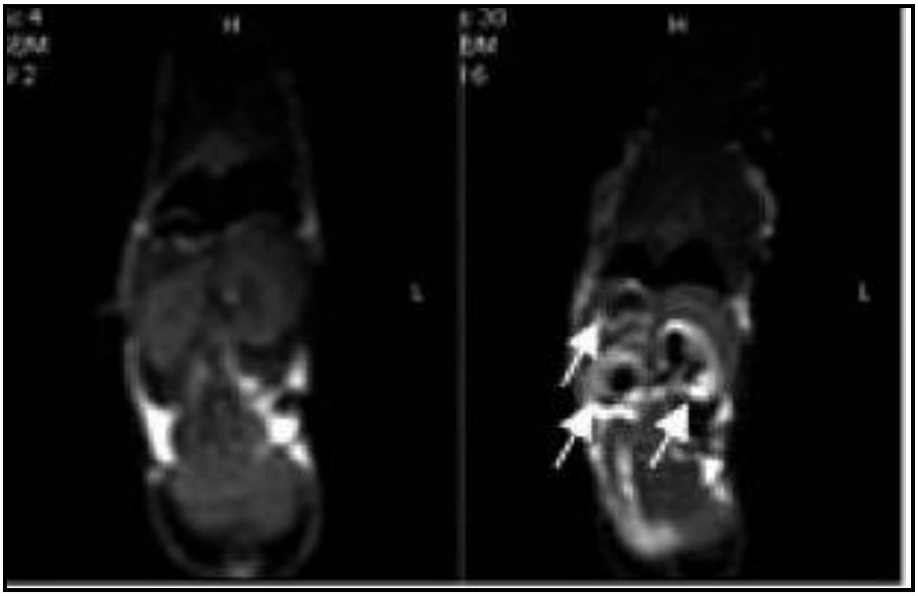


Figure 1.7 (A) Relaxivity values of gadolinium-bound ProCA1 are significantly higher than relaxivity values of Gd-DTPA. (B) Injection of contrast agent injection into a 26 g mouse produce enhanced MR images of organs (arrows). The left side shows pre-injection, and the right side shows 40 minutes post injection of 50 μL of 1.2 mM Gd^{3+}-CA1.CD2. Taken from [1].

1.5.2 Choice of Protein Host

Several different high coordination Ca^{2+} /La^{3+} binding sites have been designed in domain 1 of wild-type CD2 to understand key determinants for Ca^{2+} binding and Ca^{2+} dependent conformational change. CD2 is a cell adhesion molecule found on the surface of T-cells and natural killer cells and is part of the signal transduction pathway. It is composed of nine β-strands, forming two layers that give rise to a typical immunoglobulin fold. Along with Ig domains, CD2 has a transmembrane region, and a short cytoplasmic tail involved in protein-protein interactions (Figure 1.8) [14].

CD2.d1 was chosen as a host because of its high stability over a range of pH values 2-10 and salt concentrations 0-4 M. Additionally, it tolerates mutations for metal binding without significant change in globular structure, an important feature as a scaffold for contrast agents. CD2.d1 has two active tryptophan residues at positions 7 and 32, which can be used as spectroscopic probes for structure and metal binding studies using Tb^{3+}-Tryptophan sensitive FRET. Extensive structural studies have been performed using NMR and x-ray (Figure 1.8).

In the past, the Yang group has used two approaches to study calcium and lanthanide binding: the grafting approach, and the design approach. The grafting approach has allowed us to determine the calcium affinity to each of the four loops in calmodulin individually. The idea is to graft, or insert each of four calcium binding loops into CD2.d1, a rigid scaffold protein that does not bind calcium. This has allowed the determination of calcium binding affinity of each loop of calmodulin independently. This approach has been used to design calcium sensors as well.

The design approach allows the study of a single metal binding site in a host protein by directly mutating the residues within the protein. This eliminates the difficulties with multi-site

systems and metal-metal interaction, allowing us to understand the role of key factors such as net charge and ligand type on calcium binding affinity within a well-defined system.

The first domain of CD2 has been mutated using the design approach on the β-strand B to introduce a metal binding site formed by five coordinating ligands (E15, E56, D58, D62, D64). CD2.7E15, the novel protein, contains a class 2 calcium binding site with pentagonal bipyramidal arrangement of the ligands. CD2.7E15 differs from wild-type CD2 by 3 amino acids, utilizing 2 residues directly from CD2 [15]. Another mutant, CD2.6D31 has a calcium binding site formed by four coordinating ligand residues (E29, D31, D41, D43) (Figure 1.9).

Previous studies have shown that the two CD2 variants 7E15 and 6D31 exhibit different protein stability and conformational properties. The introduction of calcium ligand residues for 7E15 was found to lower the melting temperature (T_m) from 61 °C (seen in the wildtype) to 41°C, while retaining the wild type conformation in the presence and absence of Ca^{2+}. On the other hand, CD2.6D31 is unfolded in the absence of calcium. It retains its native like conformation upon binding of calcium [16].

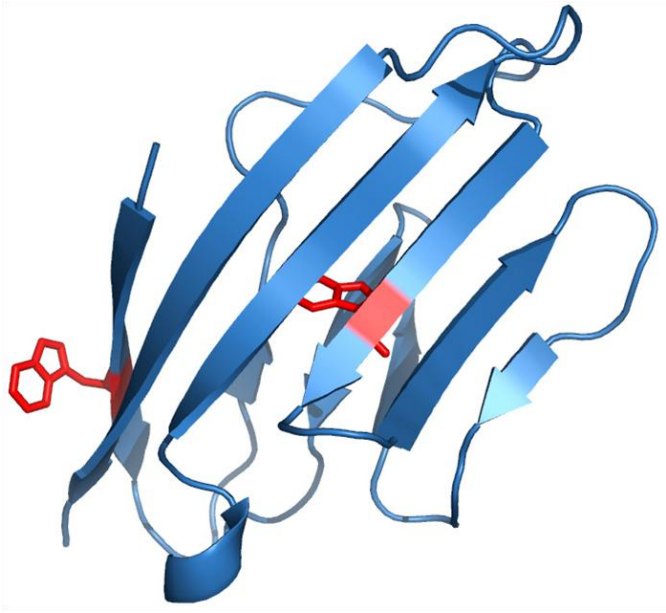


Figure 1.8 Rat cell adhesion protein CD2.d1 (1hng) is chosen as a host protein for protein-based contrast agent design. Two tryptophan residues at positions 7 and 32 are shown in red.

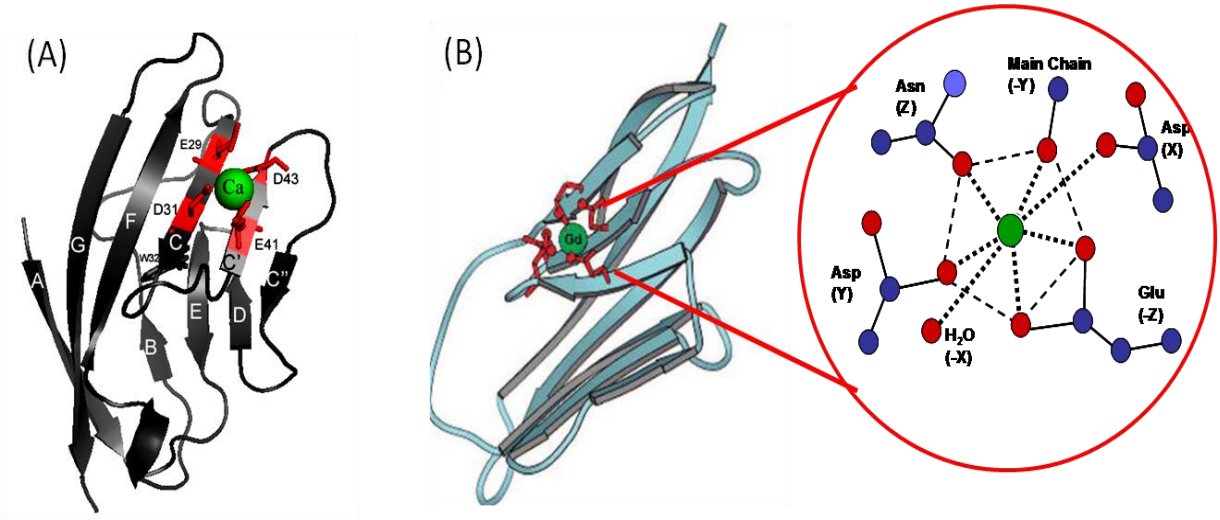


Figure 1.9 Pymol illustration of design of wild-type CD2 variants. (A) CD2.6D31 showing structure of $La^{3+/}Ca^{2+}$ binding pocket formed by four coordinating ligand residues (E29, D31, D41, D43). (B) La^{3+}/Ca^{2+} binding protein CD2.7E15 with five coordinating ligands (E15, E56, D58, D62, D64) (left) has pentagonal bipyramidal arrangement of the metal binding site (right).

1.5.3 Charged Variants of ProCA1 as Suitable Candidates for Contrast Agents

Charged variants of ProCA1 (or 7E15) were created by Dr. Anna Wilkins Maniccia utilizing the design approach with site-directed mutagenesis [17]. One of the intentions of creating 7E15 charged variants was to study whether varying the charge or coordination residue type in the metal binding pocket would affect calcium binding compared to the template, 7E15, within a fixed binding site [15]. Drs. Anna Maniccia and Wei Yang have shown that calcium binding affinity is increased as the charge in the binding pocket is increased. CD2.7E15 has a strong metal binding affinity for Tb^{3+} (K_d~10^{-10} M), as shown by FRET, and a weak calcium binding affinity (K_d of 0.15 mM). In later work, CD2.7E15 was shown to be an excellent MRI contrast agent with a strong Gd^{3+} binding affinity (K_d~10^{-10} M), as demonstrated with a metal buffer system, and significantly improved relaxivity [1].

It is not clear whether other charged variants are good candidates for MRI contrast agents. Among several available charged variants created, three -5 and -4 charged variants have a strong potential as MRI contrast agents. Residue 64 in 7E15 was changed to glutamate (called EEDDE in [15;17] or 7E15E throughout this paper) to create a -5 pair, 7E15/7E15E that differs by a methylene group. This residue was changed to glutamine or aspargine (called EEDDQ and EEDDN in [15;17] or 7E15Q and 7E15N throughout this paper), a -4 pair that also differs by a methylene group. One of the objectives of this thesis is to examine Tb^{3+}, Gd^{3+}, and Zn^{2+} binding with different charged variants and to examine the suitability of these proteins as the next contrast agents (Chapter 4).

1.6 Expression Systems: Factors Affecting Protein Expression

Large-scale production of the engineered proteins is crucial for structural analysis, functional analysis, therapeutic and *in vivo* studies. It is essential for the project of protein-based

Magnetic Resonance Imaging contrast agent carried out in Dr. Yang's laboratory. For example, an injection into a mouse for *in vivo* studies may require 5.6 mg of pure protein, and many metal binding titration experiments require up to 1 mL of highly concentrated protein (500 μM) for efficient trials. The first step to increasing the production of proteins is optimizing the expression conditions, which is crucial for preventing aggregation, degradation, and improper folding commonly seen in recombinant proteins.

Protein expression is highly dependent on the cell strain and conditions such as temperature, vector, medium, and the protein of interest. Bacterial strains versus mammalian have been shown to tolerate recombinant protein and to synthesize and fold protein well. Bacterial expression may result in well folded protein or an inclusion body depending on the expression levels. At high temperatures, *E. coli* tends to replicate quickly, producing more protein than the bacterial cells can fold at once. Under these conditions, high energy polypeptide intermediates will interact in a hydrophobic manner, thus aggregating to form a dense inclusion body. This dense structure of improperly folded, unprocessed protein will be more likely to form at higher temperatures, depending on cell strain and protein.

The expression vector contains a promoter site for expression that can affect the efficiency of expression. The pGEX-2T vector is designed to contain a GST tag attached near the N-terminal of the target protein through a short peptide chain for specific endoprotease cleavage [18]. Both pGEX-2T and pET-20b vectors contain a T7 promoter system, which is similar to the lac operon in *E. coli* (Figure 1.10). Depending on the cell strain, the promoter may be IPTG-sensitive. The vector pGEX-2T contains a GST tag which binds to the affinity column for purification and increases the solubility of the protein, while pET-20b does not [19]. Additionally, protein expression level and folding is a property of the fusion protein of interest.

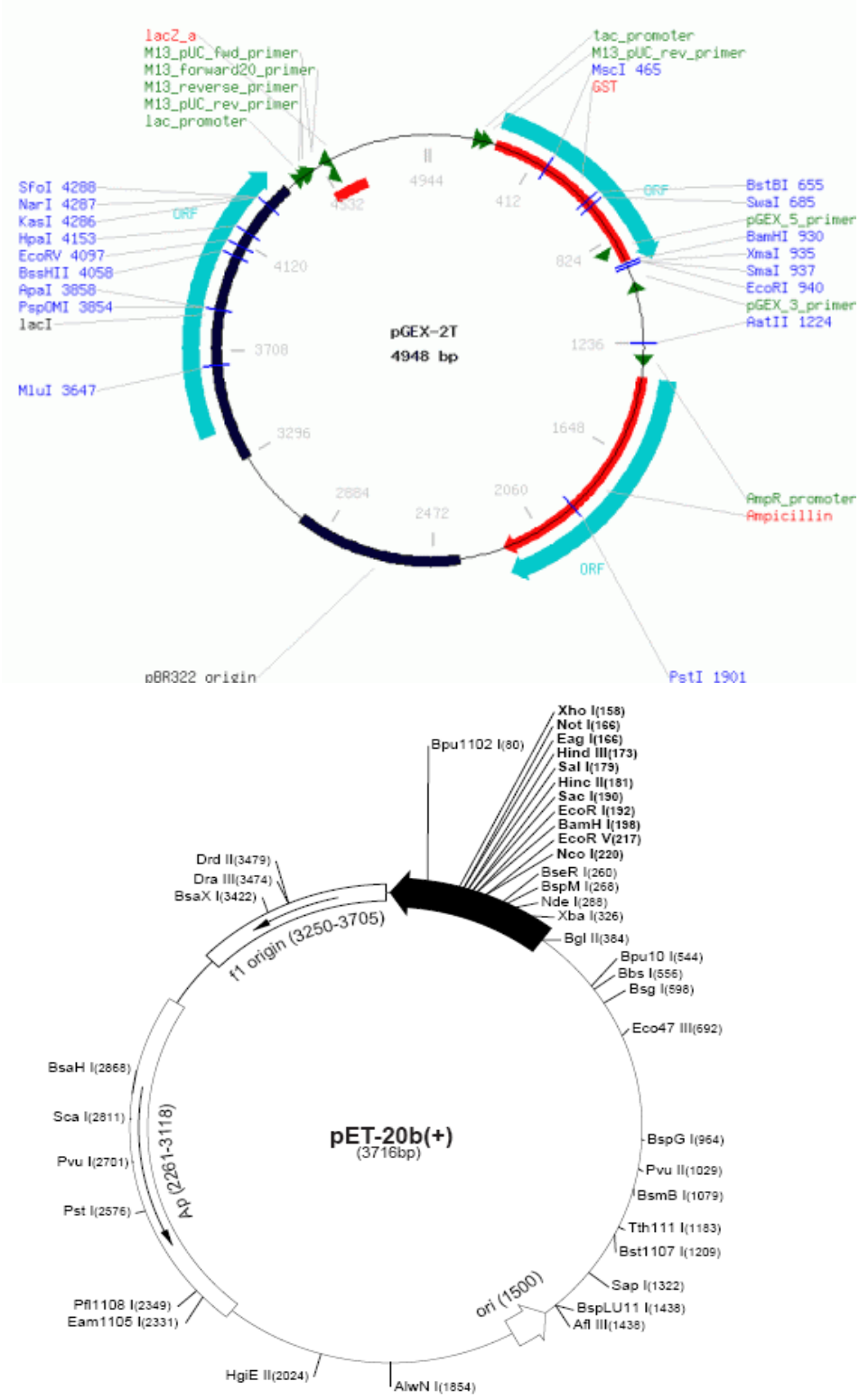


Figure 1.10 Novagen expression vectors pGEX-2T and pET-20b; pGEX-2T contains a GST purification tag while pET20b does not. Taken from [20; 21].

1.7 Purification Systems

1.7.1 Comparison of Affinity Tag Purification Systems

To achieve rapid expression and maximum purification of the recombinant proteins, affinity tags have been developed to fuse with proteins. Different models have been developed to incorporate tags. In the first model, a native protein is linked with an affinity tag by a linker region that has a sequence for endoprotease cleavage by thrombin. The second model is similar to the first except the affinity tag increases the overall solubility of the protein. Examples include Glutathione-S-Transferase (GST) and Maltose Binding Protein. In the third model, a fusion protein consists of an affinity tag removable by exopeptidase activity. In the fourth model the native protein is bound to a solubility and folding partner, fused N-Terminal to the target protein [22].

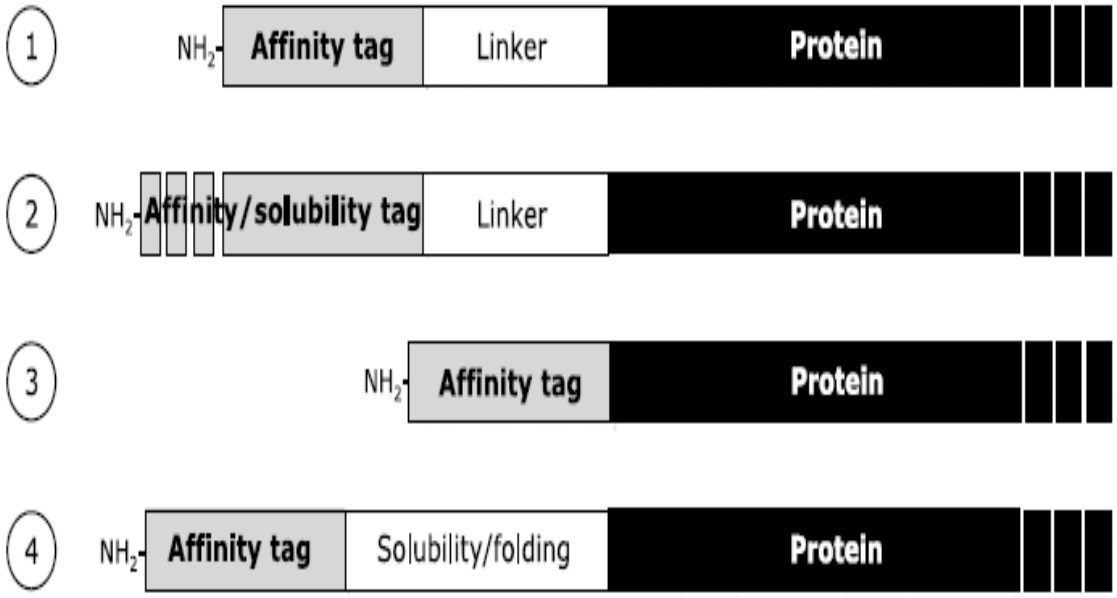


Figure 1.11 Models to incorporate affinity tags to target proteins. (1) Fusion protein containing native protein attached to an affinity tag via a linker region which has a sequence for endoprotease cleavage. (2) In this model the affinity tag enhances the solubility of the protein (examples are GST, MBP). (3) Fusion protein with exopeptidase removal of affinity tag. (4) Fusion protein where a solubility and folding region is placed between the native protein and affinity tag. Taken from [22].

Several affinity tags have been developed according to the above models including Maltose Binding Protein (MBP), Calmodulin Binding Peptide (CBP), Glutathione-S-Transferase (GST), Histidine (His), and FLAG. Use of MBP or CBP is based on affinity amylase or calmodulin, respectively, which are placed on chromatography resins. His tag is a sequence of 6 histidine residues (vector pET30a) that can bind nickel resins via high affinity of imidazole to nickel; His tags have known to improve folding of protein. GST is an enzyme that binds glutathione on affinity columns; GST tags have been used to increase solubility of small proteins expressed in *E.coli* [22, 18]. FLAG fusion protein is purified via anti-FLAG antibody binding to the FLAG sequence. GST and His tags are top choices for large scale expression and purification due to high capacity and low cost (Table 1.1) [22].

Table 1.1 Comparison of affinity tags contrasting capacity and cost to obtain 10 mg pure protein shows His and GST tags are top choices for large scale production of recombinant proteins. Taken from [18].

Tag	Size (aa)	Resin	Eluting agent	Source	Capacity	Cost/10 mg
MBP	396	Amylose	Maltose	Biolabs	3 mg/ml	$12
HIS	6	Talon	Imidazole	Clontech	5–14 mg/ml	$18
		Ni–NTA	Imidazole	Qiagen	5–10 mg/ml	$21
GST	218	GSH–Sepharose	Glutathione	Amersham	10 mg/ml	$36
CBP	28	Calmodulin affinity	EGTA	Stratagene	2 mg/ml	$114
STR (Strep II)	8	Strep-Tactin–Sepharose	Desthiobiotin	IBA	50–100 nmol/ml	$293
FLAG	8	Anti-FLAG M2 MAb agarose	FLAG peptide	Sigma	0.6 mg/ml	$1,045
HPC	12	Anti-Protein C MAb matrix	EDTA	Roche	2–10 nmol/ml	$4,983

1.7.2 Problems with GST Purification

In many cases, high expression level of GST fusion proteins is evident, while low yield is seen following purification (2-5 mg/ L cell culture). After sonication, protein in the supernatant does not bind well to the GST column, and a large portion is found in the waste. There is reason to believe that some or most of the protein expressed may be misfolded or denatured, rendering it unable to bind to affinity columns. Low yield of purification despite high expression is thought to occur due to problems during expression, such aggregation, degradation, and improper folding of fusion proteins (Figure 1.12).

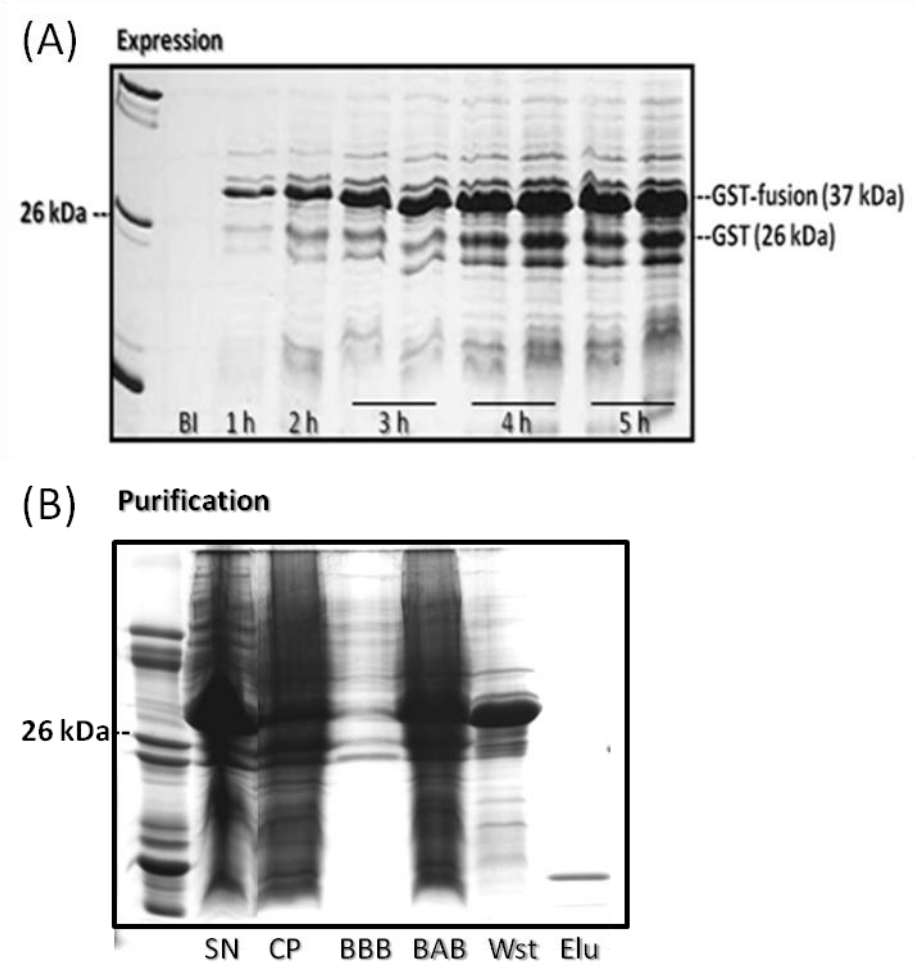


Figure 1.12 Expression and purification of GST fusion protein. (A) Shows high yield following expression 5 hours after induction with IPTG. (B) Following purification, a large amount of fusion proteins does not bind to GST beads and passes through the columns (in WST) and a low yield is seen at the fraction which eluted with glutathione.

1.7.3 Method of Refolding the Expressed Protein without Affinity Tag

While considerable effort was devoted to optimizing GST purification, the option of refolding of CD2.7E15 expressed as inclusion body without GST tag was investigated as a possibility to increase protein yield. The premise of these studies is that domain 1 of CD2 has shown to have a stable IgG-like fold and can be refolded well [23]; CD2.7E15 differs from its parent protein by 3 mutations and is likely to behave similarly.

Over-expression in *E.coli* in general has been shown to maximize the production of protein. As well, it was found from the GST studies that CD2.7E15 has a propensity to form high amounts of inclusion bodies under specific expression conditions which consists mostly of protein. The advantage of forming an inclusion body is protection from protease degradation, which would be likely because of protein size (11 kDa).

Additionally, several disadvantages of using GST tags have driven the search for this novel purification method. The use of thrombin during purification to cleave the Arg-Gly bond between GST and CD2.7E15 is not only expensive, but also reduces the yield of protein and results in excess cleaving [24]. Also the size of GST (26 kDa) compared to CD2.7E15 results in *E. coli* resources to be directed mostly at GST. Furthermore, previous studies have indicated a drastic increase in yield of His-Tag-7E15 proteins (166 mg/ L cell culture) using the refolding method (Julian Johnson, unpublished data).

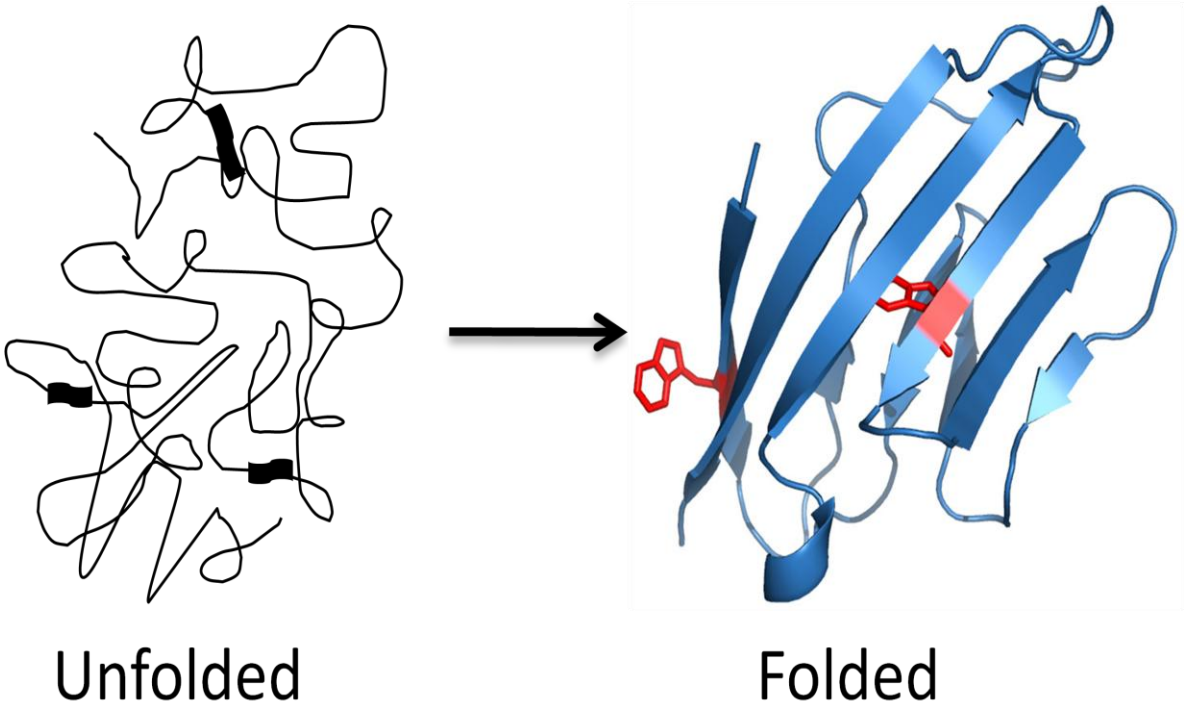


Figure 1.13 Schematic of refolding method.

1.8 Objectives of Thesis

As discussed in section 1.4, scaled up production of the CD2.7E15 and CD2.7E15 charged variants will provide a sufficient amount of protein necessary for structural and functional analysis, as well as therapeutic and *in vivo* studies, the first steps of studying these proteins as future clinical contrast agents. The objective of this thesis is to examine expression and purification conditions that will maximize the production of these contrast agents, then to determine suitability of engineered proteins as contrast agents via structural, metal binding, and relaxivity studies.

In the effort to obtain engineered proteins for biochemical and structural study and *in vivo* application, the proteins wild type CD2, CD2.7E15, CD2.6D31 were expressed as a fusion protein with the Glutathione-S-Transferase affinity tag under three different temperatures and three different cell strains, and were purified based on affinity column chromatography. The objective is to determine the optimal conditions for expression and GST purification (chapter 2). Then, CD2.7E15 expressed with a His tag and without a tag will be used to study the effectiveness of purification via the refolding method. A tag-less purification protocol is established as the most efficient purification method and is used thereon to purify all proteins (chapter 3). High purification yield is obtained using the refolding method, and CD2.7E15 and charged variants are studied as suitable contrast agents, using conformational and metal binding studies as well as relaxivity (chapter 4). Insertion of a PEG chains or peptide tags (bombesin) on various sites on ProCA1 are studied as possible ways to enhance protein-based contrast agents (chapter 5). Table 1.2 summarizes all the proteins studied in this thesis and their purpose.

Protein	Variation	Purpose of variation	Chapter(s) discussed
CD2.7E15/7E15	3 mutations of wild-type CD2 residues N15, L58, and K64 to E15, D58, and D64	To create a -5 charged binding pocket for the study of metal binding as contrast agents	2,3,4
CD2.6D31	2 mutations of wild-type CD2 residues R31 and K43, to D31 and D43	To create a -4 charged binding pocket for the study of metal binding as contrast agents	2
CD2.7E15E	1 mutation of CD2.7E15 from residue D64 to E64	To create a -5 charged analogue of CD2.7E15 for the study of metal binding	4
CD2.7E15N	1 mutation of CD2.7E15 from residue D64 to N64	To create a -4 charged analogue of CD2.7E15 for the study of metal binding	
CD2.7E15Q	1 mutation of CD2.7E15 from residue D64 to Q64		

Table 1. 2 Summarizes proteins generated for projects throughout this thesis and their variations, the purpose of their variations, and the chapters in which they are discussed.

Protein	Variation	Purpose of variation	Chapter(s) discussed
K3CR	3 lysine resides from the C terminal are mutated to arginines (K51, K66, K91R), and the remaining lysine residues are PEGylated (K43, K45, K47).	To test the effect of PEGylation in multiple locations near the binding site on relaxivity.	5
K5R	5 lysine resides are mutated to arginines (K43, K47, K51, K66, K91R), and the remaining lysine residues are PEGylated (K45).		
K5166R	2 lysine resides are mutated to arginines (K51 and K66), and the remaining lysine residues are PEGylated (K43, K45, K47k and K91).		
K6691R	2 lysine resides are mutated to arginines (K66 and K91), and the remaining lysine residues are PEGylated (K43, K45, K47 and K51).		
K51C	1 lysine residue (K51) is mutated to cysteine and is PEGylated by a cysteine reagent.	To test the effect of PEGylation in a single location near the binding site on relaxivity.	
K66C	1 lysine residue (K66) is mutated to cysteine and is PEGylated by a cysteine reagent.		

Table 1.2 (continued).

CHAPTER 2: OPTIMAL CONDITIONS FOR GST TAG EXPRESSION AND PURIFICATION

2.1 Background about GST and Affinity Tag in Protein Purification

As discussed in section 1.7.1, the use of affinity tags such as GST is an effective choice to mass produce recombinant proteins because of high levels of expression and enhanced solubility of the fusion protein in the cytoplasm. However, as pointed out in section 1.7.2, high expression of fusion protein is not a guarantee that high yield is achieved. Some or most of the protein is believed to aggregate in the cytoplasm as inclusion bodies. It is believed that expression levels of soluble protein are dependent on conditions such as temperature and cell strain. Studies by Waldo et al with green fluorescent proteins (GFP) illustrate that properties of the N-terminal fusion tag will affect the folding of GFP, which serves as a folding reporter [25]. The opposite effect, where the influence of engineered protein on the folding of the N-terminal protein, has never been examined. In past studies, varying the expression conditions such as temperature or cell strain has had a direct influence over the levels of soluble versus insoluble CD2.7E15 produced. We hypothesize that controlling the expression conditions, or providing favorable expression conditions to obtain properly folded CD2.7E15, will influence the folding of its fusion partner, GST. Proper folding of GST is essential to maximizing the outcome during purification.

In this study, we have chosen two engineered calcium binding proteins of CD2 with different folding properties as a fusion protein to test this hypothesis. We have investigated the expression of these CD2 variants and wild type under different temperatures and cell lines. Examination of SDS-PAGE, and BCA protein assay, is performed to monitor the cell growth and the expression ratio of soluble and insoluble proteins. This study examines the propensity of

inclusion body formation at three common expression temperatures: 37°C, 30°C, 25°C. Additionally, the binding capacity and ratio of soluble and insoluble protein is determined.

The GST enzymatic activity is monitored using the DTNB assay. GST enzyme exists as a dimer found in eukaryotes; each monomer is 26 kDa. The active site of GST contains a G and H site for Glutathione and electrophilic substrates respectively, which has an active cysteine residue at position 10. This special cysteine forms covalent disulfide bonds with the thiol group of the substrate glutathione, and foreign matter in the cell. The GST assay is reliant on the fact that GST joins CDNB and Glutathione in the active site and can be used to assess activity of the enzyme from different expression conditions. The binding capability of these fusion proteins to affinity columns is examined and was found to correlate with the GST assay.

For efficient binding to the Glutathione-S-Sepharose 4B affinity column, the GST tag on the fusion protein must be well-folded with its thiol group (on cysteine 10) positioned correctly. Unfortunately, in many cases, despite of high expression level of the fusion proteins, low yield of purification was reached largely due to the problems related to improper folding of engineered proteins as well as GST-tag.

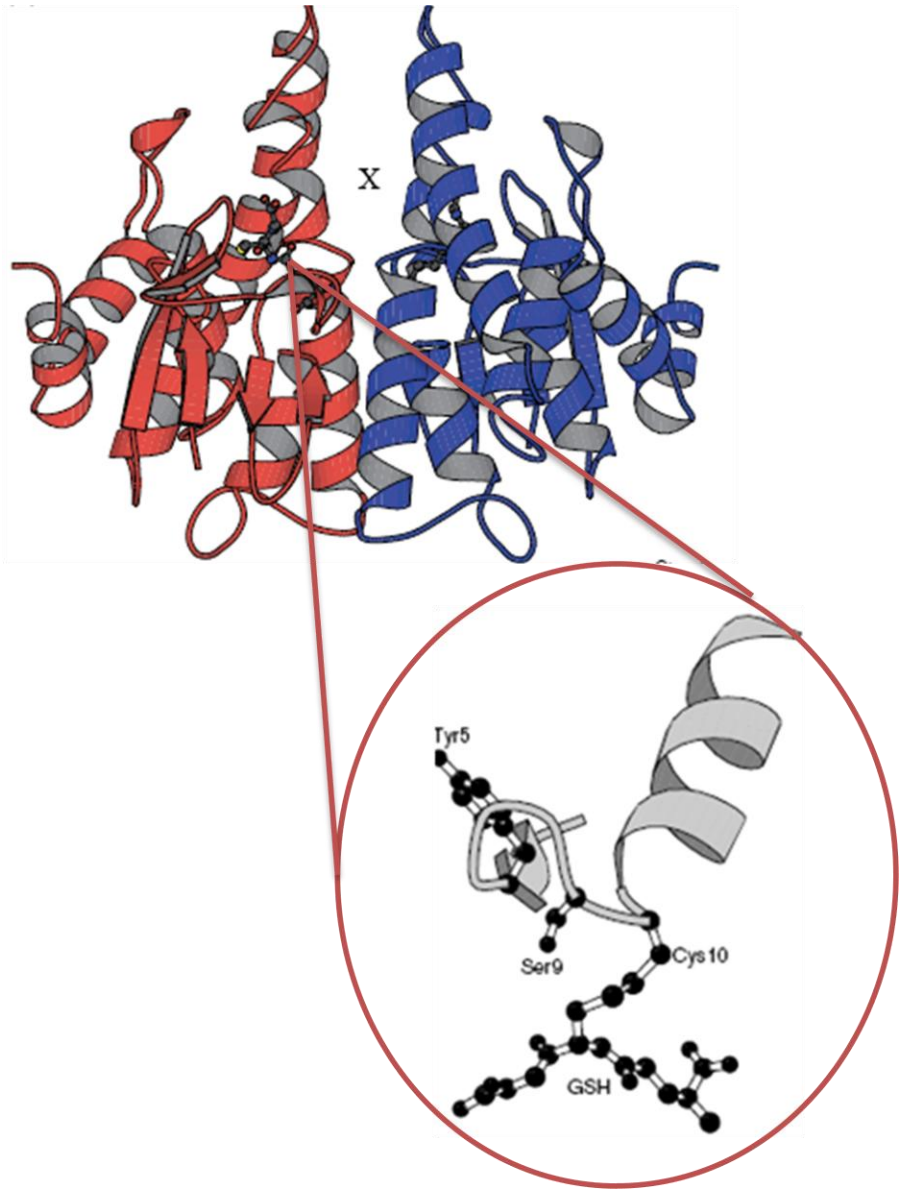


Figure 2. 1 Glutathione-S-Transferase enzyme exists as a dimer found in eukaryotes, and each monomer is 26 k Da. The active site of GST (enlarged) contains an active cysteine residue at position 10. For efficient binding to the Glutathione-S-Sepharose 4B affinity column, the GST tag on the fusion protein must be well-folded with its thiol group (on cysteine 10) positioned correctly.

The results indicate that GST binding capability for protein purification is largely dependent on the folding properties of the fusion protein. When expressed under defined, optimal conditions, well-folded native protein was found to promote a proper conformation in the fusion partner, GST, whose folding and activity are keys to a high purification yield.

2.2 Materials and Methods

2.2.1 Cloning

Cloning was conducted following established methods for GST-fusion proteins [26; 27].

2.2.2 Transformation

E. coli strains BL21 (DE3), Rosetta PLysS, and Tuner (Strategene, La Jolla, CA) were used for expression of proteins CD2.7E15, CD2.6D31, and wild type CD2. The vector plasmid pGEX-2T (GE healthcare) was used, which was designed to contain a Glutathione-S-Transferase (GST) tag DNAs. Domain 1 of CD2 and its variants DNAs were inserted into the vector as a GST fusion protein with a GST enzyme attached near the N-terminal of the protein through a short peptide chain that has a cleavage site for thrombin.

Heat shock transformation was utilized to uptake the vectors containing the engineered DNA into *E. coli* cells. First, 0.75 µL of DNA was combined with 50 µL of competent cells, and the solution was iced for 30 minutes to allow diffusion of DNA through the 20% glycerol solution bathing competent cells. Next, the mixture was dropped in a 42 °C water bath for precisely 90 seconds to allow DNA uptake by the competent cells. The cells were then promptly returned to the ice bucket for 2 minutes, after which they were given 75 µL LB medium and incubated for 30 minutes at 37 °C. The cells were then streaked with a sterile triangle onto an

agar plate containing ampicillin and were allowed to incubate overnight at 37 °C, upside down. An additional plate containing cells without DNA was incubated as a negative control.

2.2.3 Expression of Wild-type CD2, CD2.7E15, and CD2.6D31 GST Fusion Proteins

A bacterial colony was inoculated and incubated overnight in LB medium with ampicillin (2.5 mM). The cell pellets were collected by centrifugation (Sorvall RC SB Plus Centrifuge, ThermoScientific, Bruswick, NJ) at 4.410 x 10^3 *g* in 5 °C for 15 minutes. The cell pellets were diluted 1:20 with LB medium with ampicillin (2.5 mM), and were incubated at 37 °C in a shaker (Shaker Controlled Environment Incubator).

The optical density (OD) of each sample at 600 nm was recorded every 30 minutes using a UV spectrophotometer (Shimadzu Scientific Instruments, Norcross, GA). When the OD reached 0.6 for BL21 (DE3) and Rosetta PLysS, and 0.5 for Tuner cells, Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the cultures. The BL21 (DE3) and Rosetta pLysS strains required only 0.25 mM IPTG, while the Tuner strain required 0.5 mM. The shakers were adjusted to the desired temperature either 37°C, 30°C, or 25°C, for the duration of expression. The cell pellets were collected by centrifugation at 3.842 x 10^3 *g* for 20 minutes after the cultures grew about 5 hours. The cell pellets were stored at -20 °C until purification. KaleidaGraph software (Synergy Software, Reading, PA) was used to plot the optical density versus time plots. Image J software (Version 1.37) was used to quantify band thickness as a result of expression on the 15% acrylamide SDS-PAGE gels.

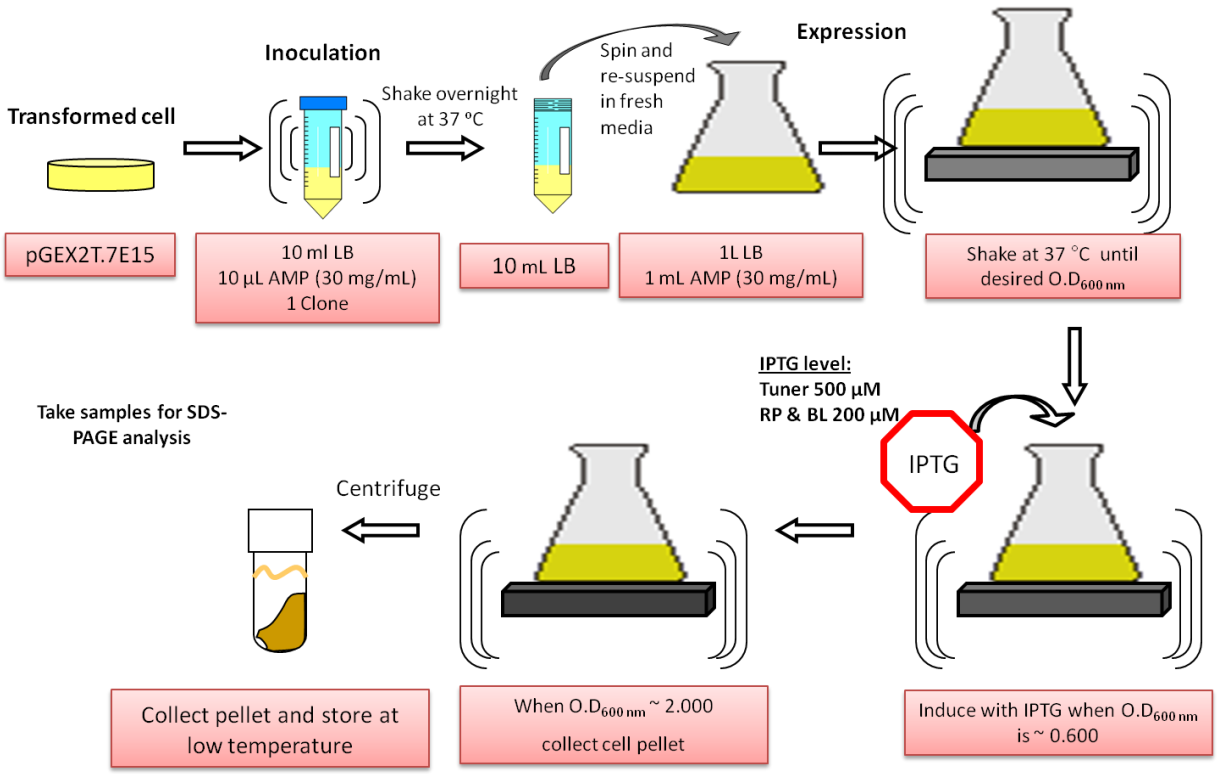


Figure 2.2 Summary of expression procedure.

2.2.4 Purification of GST Fusion Proteins by Affinity Chromatography

Cell pellet containing expressed protein was prepared by resuspension in 5 mL of lysis buffer (1% N-lauroylsarcosine sodium Tris-HCl, pH 7.0, 5 mM DTT, 5 μM AEBSF, and 25 U/mL benzonase nuclease (Novagen)). The mixture was sonicated six times at intervals of 30 seconds each, and centrifuged to separate debris at 4.410 x 10^3 *g* for 5 min. The supernatant was retained and filtered with a 0.45 μM filter.

Glutathione-S-Sepharose 4B affinity columns (GE healthcare) with capacity of 2 mg/ 1mL of resin, were treated with 9 M glutathione (GSH) for cleansing and rinsed with phosphate buffered saline (PBS).

The filtered supernatant was applied to affinity columns three times for binding to occur between the GST tag and the beads. In order to cleave the GST tag from the protein, 20 U of thrombin (GE healthcare) is added to each column, which is then placed on an agitator at room temperature for four hours. The target protein was eluted from the column with PBS buffer, followed by filtration with a 0.45 μM filter. In some case, the thrombin was not added to the affinity resin and the GST tag was not cleaved from the fusion protein. Once the fusion protein was bound to the column, the whole unit was eluted with highly concentrated glutathione by competition.

Protein was then further purified by Superdex G75 gel filtration and HiTrap SP cation exchange FPLC (GE healthcare).

2.2.5 Bradford Concentration Assay (BCA) and GST Enzymatic Activity Assay

The BCA Protein Assay Kit (Thermo Scientific, Lot Number JD 121926) was used in order to quantify the ratio of soluble and insoluble protein in solution in accordance with the protocol.

To examine the native folding of GST, we have monitored the GST activity using the GST Detection Module as detailed in the protocol (Amersham Biosciences, GE Healthcare). The GST activity assay is reliant on the fact that the GST enzyme binds glutathione and the electrophilic substrate Chloro-dinitrobenzene (CDNB) [28]. When fully oxidized GSH donates electrons to electrophilic CDNB, the absorbance at 340 nm increases. This assay measured the absorbance at 340 nm after GSH and CDNB were added to protein solution. The absorbance over the initial 100 seconds was observed, and over 20 points were plotted and fitted. The absorbance did not exceed 0.8. A negative control with buffer only was checked first to verify the absence of activity. The activity of GST, or the initial velocity, was found by calculating the slope of product concentration [P] (Equation 2.1) over time, where ε is the molar absorptivity, and l is the path length.

$$[P] = A_{340nm}/(\varepsilon * l) \qquad \text{(Equation 2.1)}$$

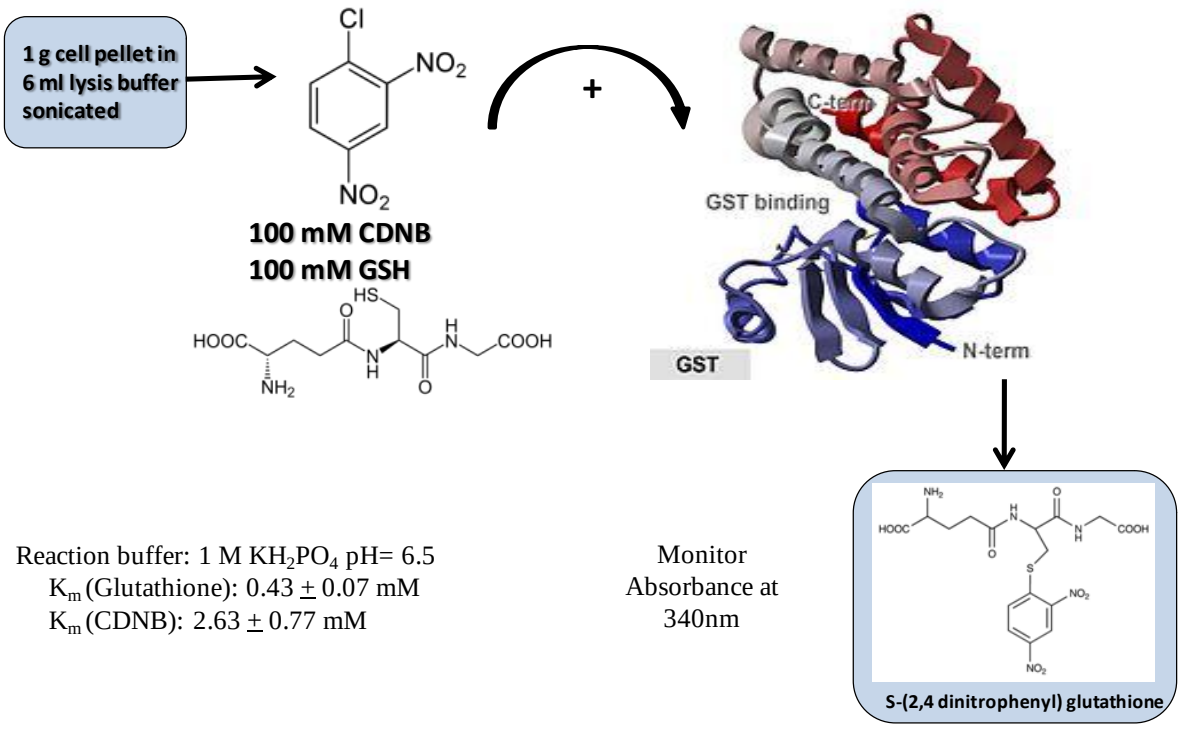


Figure 2.3 GST Detection Module (GE healthcare) [29].

2.2.6 GST Beads Binding Assay

The GST beads binding assay was performed by mixing a 2:1 ratio of Glutathione-S-Sepharose 4B beads with sample. Samples of the beads were drawn after the beads were bound with the protein and washed with phosphate saline buffer. Further analysis by an SDS-PAGE gel and Image J software provided a measure of the bead capacity by a comparison of the band intensity ratios.

2.2.7 SDS-PAGE

Prior to induction and hourly post induction, 1 mL samples were collected from the expression flasks for SDS-PAGE. As well, samples of the supernatant and cell pellet post sonication, and the beads before and after binding were collected for SDS-PAGE. All samples were centrifuged at 9.757x 10^3 *g* for 1 minute in an eppendorf centrifuge and the supernatant was discarded. To each sample, a 1:10 ratio of SDS-PAGE sample buffer comprised of 5% ß-mercaptoethanol was added. Samples were boiled at 100 °C for 10 minutes in a water bath.

2.3 Results and Discussion

2.3.1 Expression of CD2.7E15, CD2.6D31, and Wild-type CD2

2.3.1.1 The Effect of Temperature on Protein Expression

The expression of wild-type CD2 and two variants at three different temperatures and cell strains five hours after induction with appropriate levels of IPTG is seen in Figure 2.3. CD2.7E15 was expressed efficiently at 25 °C or 30 °C, as shown by the intensity of the SDS-PAGE. Expressions at 30 °C resulted in bands that were denser by 8-10% compared to 37°C. At higher temperatures, cleavage of fusion protein is seen. Usually, the SDS-PAGE depicts two

bands for the expressed fusion protein, one of which is lower than 34 k Da. The cleavage seems to diminish as the temperature is lowered (Figure 2.3 A, B and C).

SDS-PAGE depicting the expression of CD2.6D31 shows bands that were not significantly diverse from each other. The difference in density among the three temperatures did not exceed 4%. The same cleavage of fusion protein during expression was noted with higher temperatures, and band density decreased with temperature (Figure 2.3 D, E and F).

Wild-type CD2 was expressed in higher levels at 37 °C versus 25 °C. The SDS-PAGE shows bands that are darker at 37 °C by 4-9%, depending on the cell strain, that indicate the presence of the fusion protein (Figure 2.3 G, H and I).

2.3.1.2 The Effect of Cell Strain on Expression

The three commonly used cell strains of Tuner, BL21 (DE3), and Rosetta pLysS, were tested for optimal expression. Although CD2.7E15 expressed in Tuner cells produced the densest bands on the SDS-PAGE, cleavage of the affinity tag was observed more than in other cell strains. BL21 (DE3) and Rosetta pLysS produced the expression with the lowest amount of fusion protein cleavage. CD2.6D31 and wild-type CD2 produced the densest bands with Rosetta pLysS or BL21 (DE3) cell strains (Figure 2.3)

The difference in expression was apparent among the three fusion proteins. CD2.7E15 expressed in higher amounts than wild-type CD2 and CD2.6D31 in all cell strains. In fact, SDS-PAGE shows band densities of CD2.7E15 greater than wild-type CD2 by 10.6%, and CD2.6D31 by 22.8%. Overall, CD2.6D31 expressed in the lowest amount among the three fusion proteins, followed by wild-type CD2, and CD2.7E15.

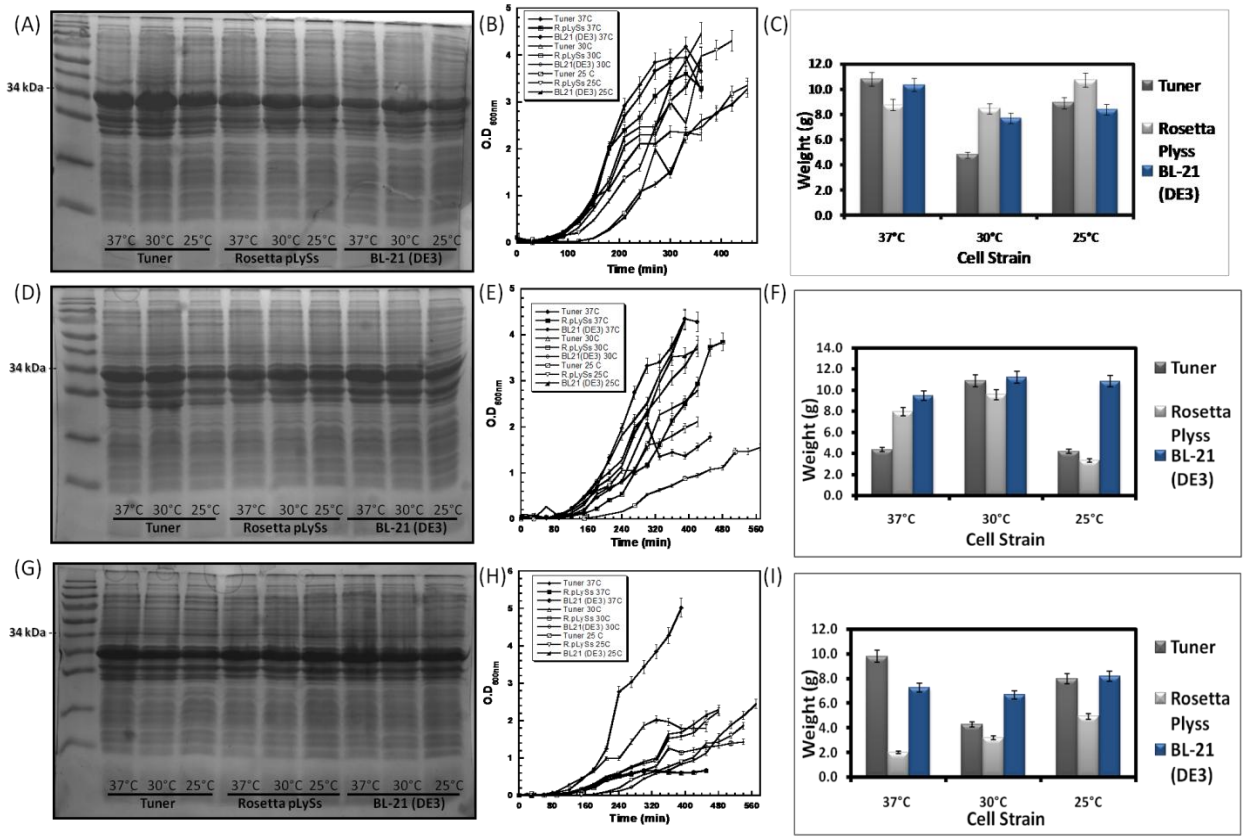


Figure 2.4 Expression gels showing the final point 5 hours after induction for each cell strain and temperature, growth curves, and wet cell pellet weights per 2 L expressed cell culture for CD2.7E15, CD2.6D31, and wild-type CD2. (A) Expression gel for CD2.7E15 shows that lower temperatures (25 °C or 30 °C) may be more effective for expression as bands are thicker by 8-10%. (B) Growth curve for CD2.7E15 shows a gradual, steady growth for all cell strains. (C) Wet cell pellet weights is highest at 37 °C especially for Tuner cell strain. (D) Expression gel for CD2.6D31 shows no significant thickness among bands (less than 4% difference) but a more efficient growth with Rosetta pLysS. (E) Growth curve for CD2.6D31 shows a gradual, steady growth for all cell strains especially Rosetta pLysS. (F) Highest weight for cell pellet is seen at 30°C. (G) Expression gel for wild-type CD2 shows thicker bands at 37 °C by 4-9%. (H) Growth curve for wild-type CD2 shows a gradual, steady growth for all cell strains especially. (I) Cell pellet weights are highest at 37 °C excluding Rosetta pLysS.

2.3.2 Estimation of Insoluble and Soluble Protein Ratio of Expressed Proteins by BCA Assay

The relative fractions of expressed fusion proteins as a soluble form or as an insoluble form in inclusion body were examined by measuring protein concentrations of 1 g cell pellets dissolved in lysate buffer or in the cell pellets using BCA protein assay (as described in section 2.2.5). Figure 2.4 shows that expression of CD2.7E15 at 37 °C tends to favor the production of inclusion bodies by approximately 10-40% more than expression of this protein at 30 °C or 25°C. CD2.7E15 contains the highest level of inclusion bodies at 37 °C due to its especially high expression level in at this temperature. In CD2.6D31 expression, Tuner cell strain tends to produce the largest amounts of inclusion bodies, 10-20% more than in other cell strains. For this protein, Rosetta pLysS cell strain tends to produce less cell pellet. In wild-type CD2, the differences in inclusion body production across temperatures and cell strains were not as pronounced, none exceeding over a 10% difference. Decreasing expression temperature to 25 °C results in more soluble form of expressed proteins especially for CD2.7E15 and CD2.6D31.

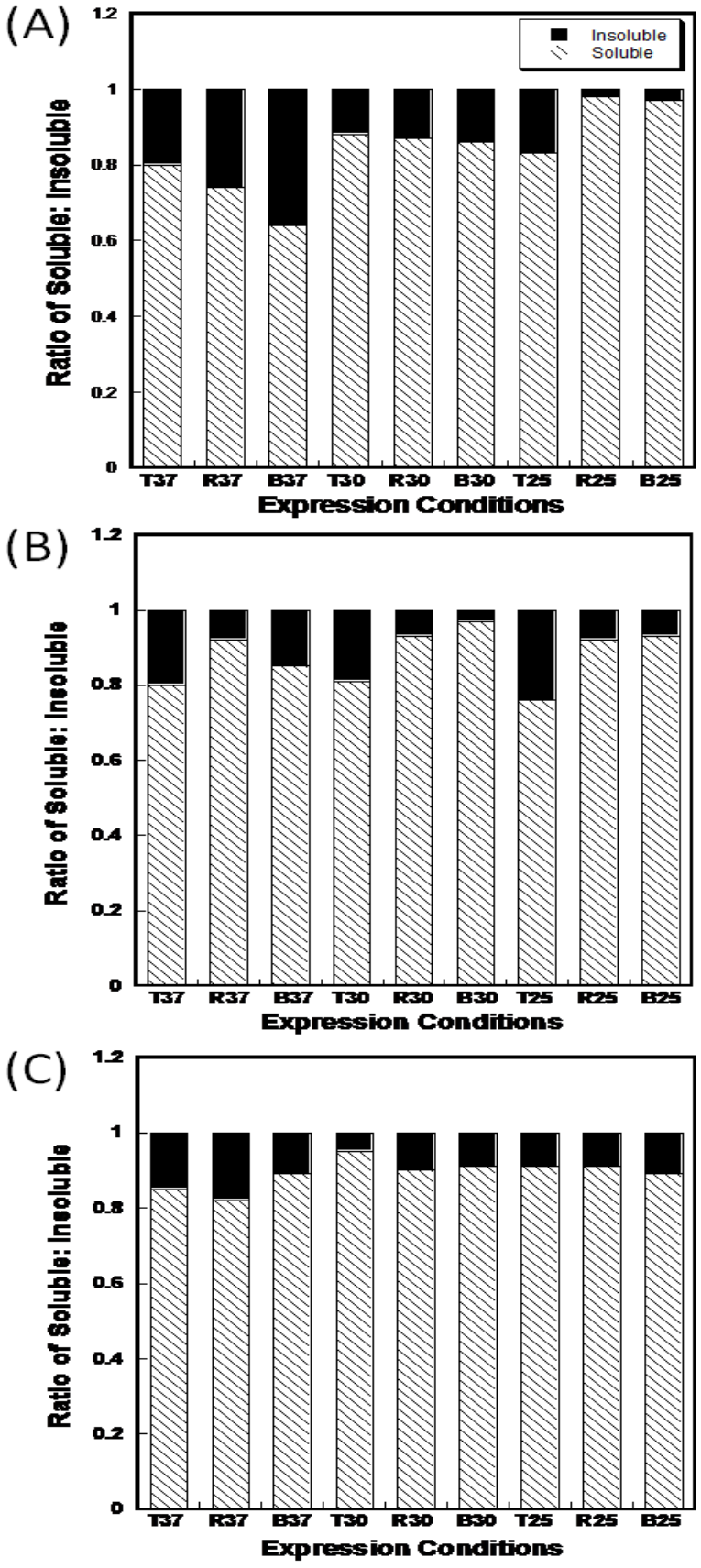


Figure 2.5 Estimation of inclusion bodies in 1 g cell culture among wild type CD2 and two variants at expression temperatures 37 °C, 30 °C, and 25 °C with cell strains Tuner, Rosetta pLysS, and BL21 (DE3). (A) Shows CD2.7E15 inclusion body level is highest at 37 °C expression by approximately 10-40% more than in expression at 30 °C or 25 °C. (B) Shows CD2.6D31 inclusion body level is 10-20% higher with use of Tuner than with use of Rosetta pLysS cell strain. (C) In wild-type CD2, the differences in inclusion body production across temperatures and cell strains did not exceed 10%.

2.3.3 The Relationship Between the GST Enzymatic Activity and Beads Binding Capability of the GST Fusion Protein

In this study, the folding of GST is examined by measuring the enzymatic activity of the GST using the GST Enzymatic Assay (as detailed in section 2.2.6). Well-folded GST, which exhibited high activity in this assay, was found to bind to Glutathione-S-Sepharose 4B beads most efficiently.

CD2.7E15 expressed at lower temperatures was found to have the highest enzymatic activity indicated by the fastest initial velocity ($mM^{-1}s^{-1}$) (Figure 2.5 A). Cell strains like Tuner seemed to express the most highly active GST, as well as BL21 (DE3) at 25 ºC. The SDS-PAGE of beads binding assay shows thick bands at the low temperature end up to 80% higher than the high temperature end, which means that the most active GST bound most efficiently to the affinity beads (Figure 2.7 A, B).

Expression of CD2.6D31 specifically with the cell strain Rosetta pLysS regardless of temperature produced the most highly active GST (Figure 2.5 B). The SDS-PAGE gel depicting the beads binding assay correlates with the enzymatic activity: the most protein bound to the beads came mostly from the Rosetta pLysS samples (Figure 2.7 C and D). This was seen for all temperatures of expression especially at 30 ºC.

Figure 2.5 C shows that for wild-type CD2 expressed at 37 ºC, highly active protein was produced in cell lines like Tuner and BL21 (DE3). Such protein bound to the beads in greatest amounts, as seen on the SDS-PAGE (Figure 2.7 E and F). Protein expressed at 30 ºC with Rosetta pLysS cell strain was also found to have the highest activity. The highest intensity bands were seen at high expression temperatures, which were thicker by approximately up to 70%. Wild-type CD2 was found to expresses highly active protein at temperatures of 37 ºC across all cell strains.

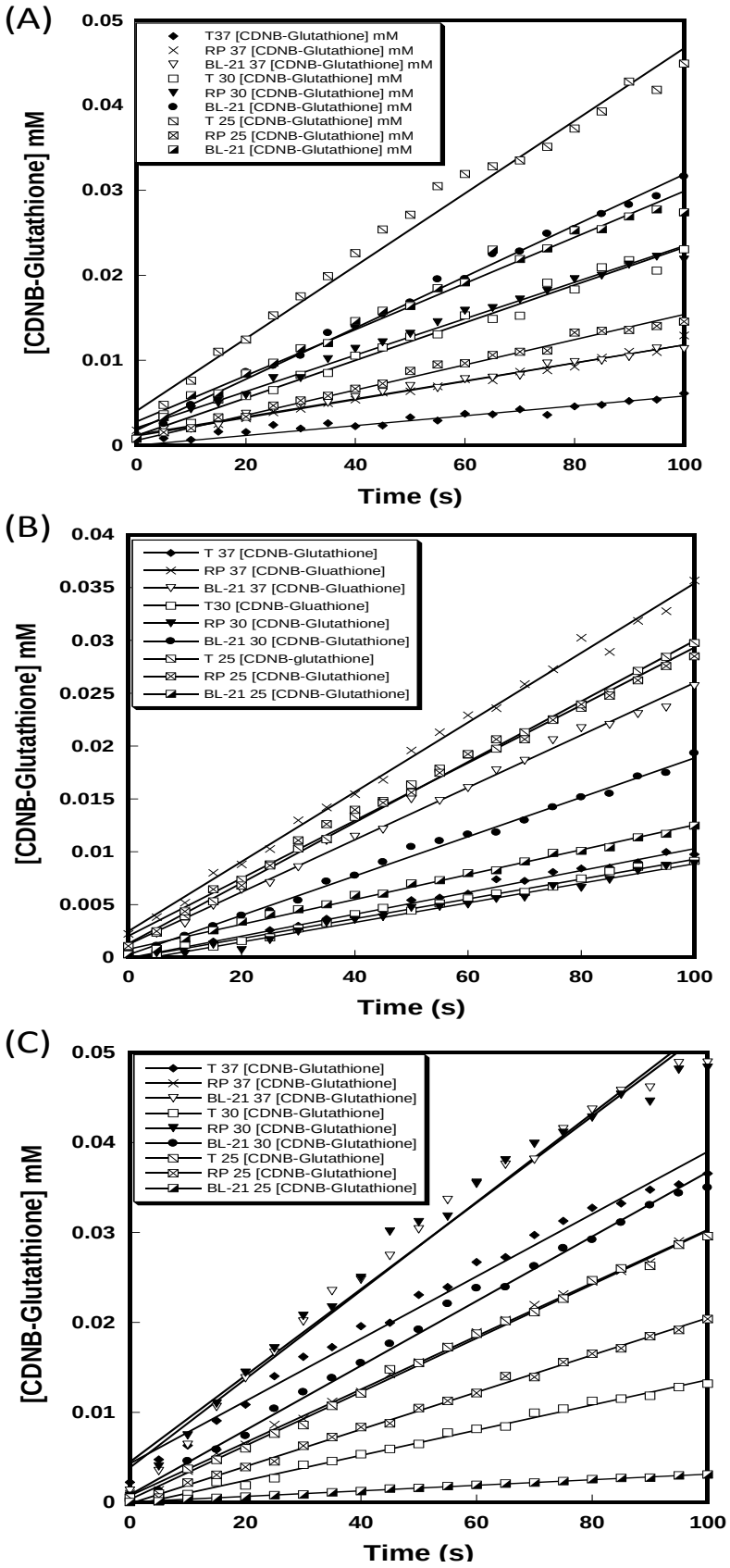


Figure 2. 6 Shows concentration of CDNB-Glutathione (product) formation over time for the first 100 seconds plotted and fitted linearly. (A) CD2.7E15. (B) CD2.6D31. (C) CD2WT. The slope is equal to the initial velocity (V_0).

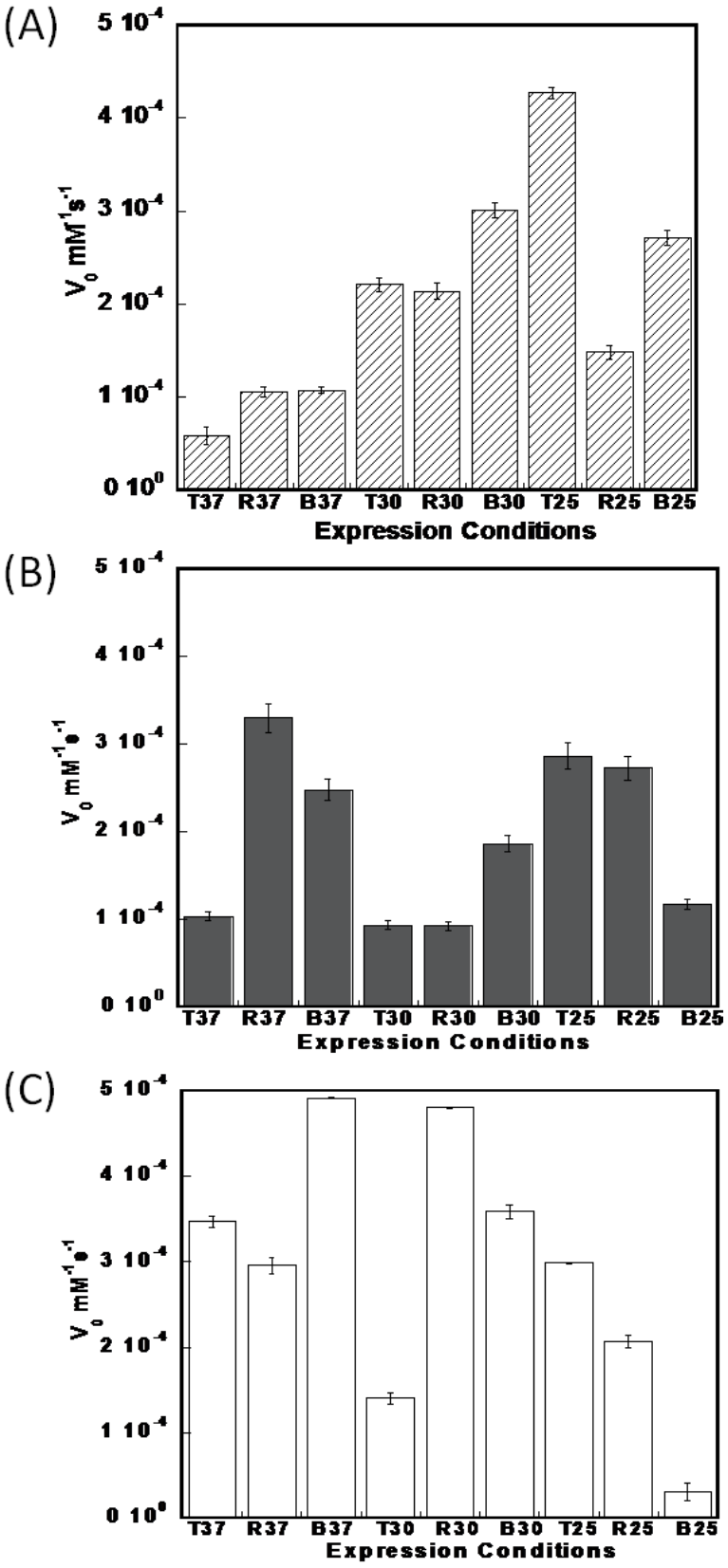


Figure 2.7 The initial velocity (V_0) of GST fusion protein calculated from the slope of the graph of product concentration (mM) versus time (s) (Figure 2.6). (A) Shows that enzymatic activity is highest for CD2.7E15 expressed at low temperatures. (B) Show that enzymatic activity is highest for CD2.6D31 expressed with Rosetta pLysS cell strain regardless of the temperature. (C) Shows that enzymatic activity is highest for wild-type CD2 expressed at 37 ºC regardless of the cell strain. Error bars refer to average standard deviation from three separate trials.

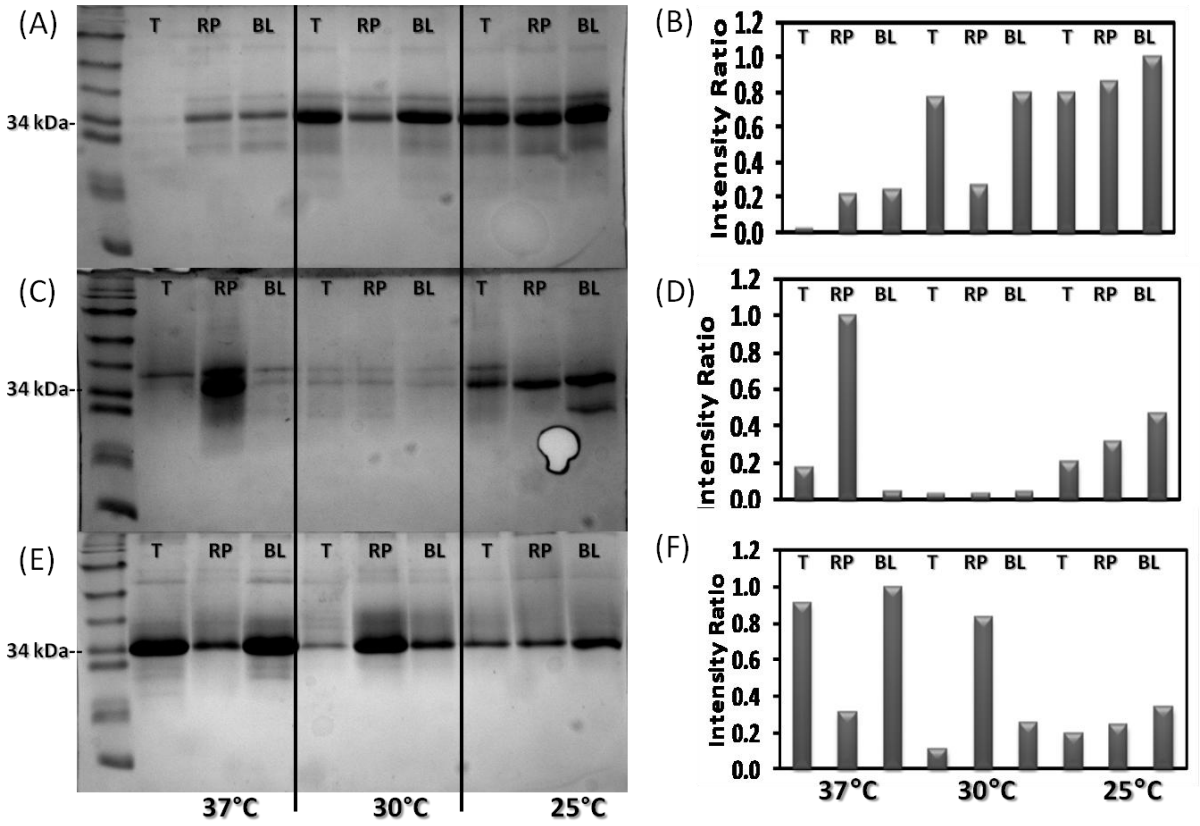


Figure 2.8 Beads binding capability assay of GST fusion protein monitored by SDS-PAGE (left) and intensity ratio (right) indicating binding capability is shown to correspond with the GST enzymatic activity. (A and B) In CD2.7E15 the high intensity bands from protein expressed at lower temperatures are thicker than the higher temperatures by up to 80%. (C and D) For CD2.6D31, the highest intensity bands came from proteins expressed with Rosetta pLysS cell strain. (E and F) In wild-type CD2, the highest intensity bands were seen at high temperature expression, which were thicker by approximately up to 70%.

2.4 Conclusions

Optimal conditions were determined for the expression of each CD2 variant. Cell growth was found to be optimal at 37 °C for all three cell strains, as the replication of *E. coli* is maximal. Yet, the quantity of soluble protein produced during expression is variable, depending on the temperature and cell strain. CD2.7E15 was found to express the highest quantity of soluble, active protein at low temperatures with cell strain BL21 (DE3). Tuner was found to form high amounts of inclusion bodies in this protein. The optimal level of soluble and active CD2.6D31 was expressed with Rosetta pLysS, regardless of chosen temperature. Wild type CD2 was expressed in highest quantities of soluble form at higher temperatures, such as 37ºC. Cell strain preference for this protein was BL21 (DE3) and Tuner. Overall, Tuner cell strain was found to produce the highest level of inclusion bodies for CD2.7E15 and CD2.6D31, and to cleave the fusion protein at 37 ºC and 30 ºC.

The folding of the GST-affinity is largely dependent on the engineered protein folding. Ultimately, engineered protein folding is affected by the expression temperature and cell strain. Purification was carried out successfully, and correlations between beads binding assay and GST enzymatic activity indicate that high enzymatic activity in GST resulted in larger quantity of fusion protein binding to the affinity columns. CD2.7E15 expressed at low temperatures and in BL21 (DE3) is well folded, and thus the affinity tag is well folded and biologically functional. The same is true for CD2.6D31 and for wide type CD2 that are expressed in the optical conditions mentioned (Table 2.1). On the whole, purification level is highly linked with GST folding, which is tied back to the expression conditions. This information is a powerful tool for future work with these proteins. As seen in chapter 3, the tag-less purification method was optimized based on these findings.

Table 2.1 Summary showing the expression temperature and cell strain required to produce soluble fusion protein for wild-type CD2 and its variants.

7E15

	37°C	30°C	25°C
Tuner	Insoluble	Insoluble	Soluble
Rosetta pLySs	Insoluble	Soluble	Soluble
BL-21 (DE3)	Insoluble	Soluble	Soluble

6D31

	37°C	30°C	25°C
Tuner	Insoluble	Insoluble	Insoluble
Rosetta pLySs	Soluble	Soluble	Soluble
BL-21 (DE3)	Soluble	Soluble	Soluble

CD2WT

	37°C	30°C	25°C
Tuner	Soluble	Soluble*	Soluble*
Rosetta pLySs	Soluble	Soluble*	Soluble*
BL-21 (DE3)	Soluble	Soluble*	Soluble*

* Wild-type CD2 is soluble under these conditions, but the yield is low

CHAPTER 3: OPTIMIZATION OF A TAG-LESS PURIFICATION PROCEDURE TO IMPROVE THE YIELD OF ENGINEERED PROTEIN

3.1 Introduction

The focus of chapter 2 was to optimize expression and purification of engineered protein using the GST system. Although this was performed successfully, the section 1.7.2 above had briefly mentioned the limitations of this method, and section 1.7.3 detailed that purification using a refolding method has shown increased yield in 7E15 and His-tagged 7E15. The refolding method is not effective with GST-tagged proteins because of the size of GST and the cysteine content. The need to establish a rapid method that would generate high yield of engineered protein for future study is addressed in this chapter. As mentioned in section 1.8, the purpose of this chapter is to detail the expression and purification experiments that led up to the establishment of the refolding method implemented in future chapters. Expression conditions found in chapter 2 to generate insoluble protein were used to generate inclusion bodies which were then refolded (Table 3.1). The objective is to use the refolding method to generate high yield of 7E15.

Table 3.1 Summary of conditions found in chapter 2 with an added emphasis on conditions for inclusion body formation.

7E15

	37°C	30°C	25°C
Tuner	Insoluble	Insoluble	Soluble
Rosetta pLySs	Insoluble	Soluble	Soluble
BL-21 (DE3)	Insoluble	Soluble	Soluble

3.2 Materials and Methods

3.2.1 Cloning

Cloning of 7E15 and 7E15-bom was performed by Lixia Wei (Dr. Zhi-Ren Liu's laboratory). Protein sequence was removed by restriction enzymes from the pGEX-2T vector and inserted into a pET30a vector, which contains a 6-histidine purification tag, between BamH1 and EcoR1 restriction sites.

3.2.2 Transformation

E.coli strains BL21 (DE3), BL21, BL21 (DE3) pLysS, and Tuner (Strategene, La Jolla, CA) were used for expression of proteins 7E15, His-7E15-bom, and 7E15-52I. The vector plasmids pET20b (tag-less) and pET30a (designed to contain a 6 histidine residue affinity tag) (GE healthcare) were used. Heat shock transformation was utilized to uptake the vectors containing the engineered DNA into *E. coli* cells. DNAs of 7E15 and 7E15-bom were cloned and inserted into the vectors. First, 0.75 µL of DNA was combined with 50 µL of competent cells, and the solution was placed in ice for 30 minutes to allow diffusion of DNA through the 20% glycerol solution bathing competent cells. Next, the mixture was dropped in a 42 °C water bath for precisely 90 seconds to allow DNA uptake by the competent cells. The cells were then promptly returned to the ice bucket for 2 minutes, after which they were given 75 µL LB medium and incubated for 30 minutes at 37 °C. The cells were then streaked with a sterile triangle onto an agar plate containing antibiotic and were allowed to incubate overnight at 37 °C, upside down. An additional plate containing cells without DNA was incubated as a negative control.

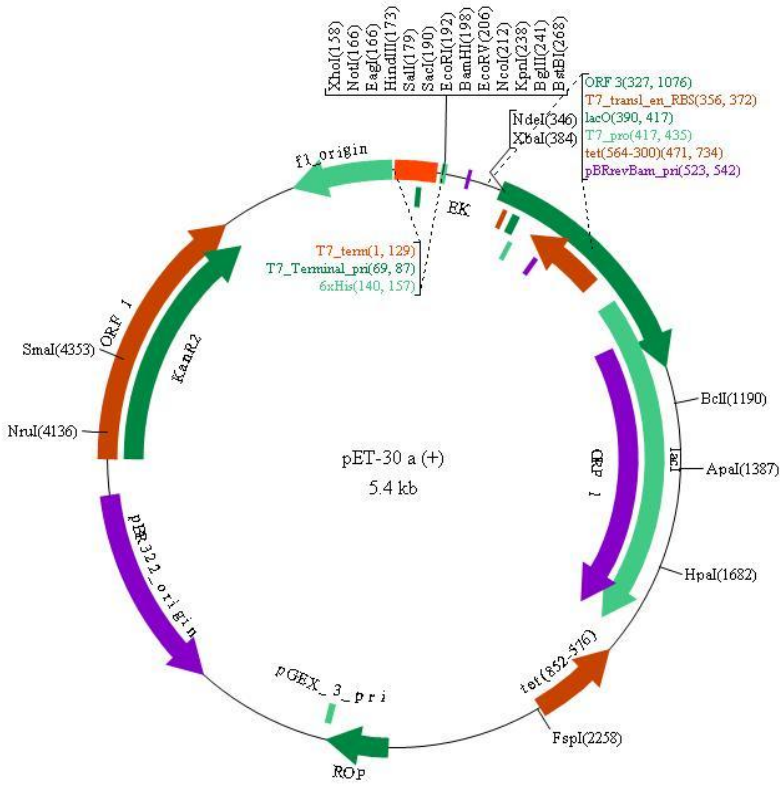


Figure 3.1 Plasmid vector map of pET30a. Taken from [30].

3.2.3 Expression of His-7E15 and 7E15

A bacterial colony was inoculated and incubated overnight in LB medium and 2.5 mM ampicillin or kanamycin, and centrifuged (Sorvall RC SB Plus Centrifuge, ThermoScientific, Bruswick, NJ) 4.410 x 10^3 *g* for 15 minutes at 5°C. The cell pellet was diluted 1:20 with LB medium and appropriate and transferred to a flask and incubated at 37 °C in a shaker (Shaker Controlled Environment Incubator).

The optical density of each sample at 600 nm was recorded every 30 minutes using UV spectrophotometer (Shimadzu Scientific Instruments, Norcross, GA). When the optical density (OD) reached 0.6 for BL21 (DE3), Rosetta PLysS, and BL21 (DE3) pLysS, and 0.5 for Tuner cells, flasks were induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG). The BL21 (DE3), BL21 (DE3) pLysS, and Rosetta pLysS flasks required only 0.25 mM IPTG, while the Tuner flasks required 0.5 mM. The shakers were adjusted to the desired temperature either 37 °C or 30°C, for the duration of cell growth. The flasks were collected and centrifuged at 3.842 x 10^3 *g* for 20 minutes when the flasks have grown 4 hours or overnight. The cell pellet collected and stored at -20 °C until purification. KaleidaGraph software (Synergy Software, Reading, PA) was used to plot the optical density versus time plots. Image J software (Version 1.37) was used to quantify band thickness as a result of expression on the 15% acrylamide SDS-PAGE gels.

3.2.4 Optimization of His-tagged 7E15 by Adjustment of Temperature and IPTG Concentration at Induction and Purification with Arginine

As described earlier (section 1.6.1), several factors were found to affect expression level. In this section the temperature and IPTG level were adjusted to optimize the parameters for the previously low expression level of His-7E15. The expression protocol detailed in sections 2.2.3 was carried out as stated with a pET30a vector. At the point of induction some flasks received

0.2, 0.4, 0.6, 0.8 and 1 mM IPTG. The effectiveness of expression was evaluated by band intensity on the SDS-PAGE gel. Expression at 37 °C and 30 °C took place 4 hours after induction.

The protein was purified. Cell pellet was resuspended three times in PBS containing 1% Triton x-100, and centrifuged for 30 minutes. The cell pellet was dissolved in 1M arginine and left on the shaker overnight at 4°C. The following day, the mixture was centrifuged to separate insolubilzed protein and the supernatant was found in the supernatant. Size exclusion chromatography was performed and SDS-PAGE samples were collected of the supernatants and cell pellets after sonication, washing, and solubilization.

3.2.5 SDS-PAGE gel

Prior to induction and hourly post induction, expression samples were collected, centrifuged at 9.757x 10^3 *g* for 1 minute in an eppendorf centrifuge and the supernatant was discarded. To each eppendorf, a 1:10 ratio of sample buffer comprised of 5% ß-SH EtOH was added to create a reducing environment and sodium dodecyl sulfate (SDS) was added to denature the protein. Addition of SDS causes the protein to become negatively charged and linearized. Subsequently, separation on the gel by migration to the positive pole is a representation of size and not shape or charge [31]. Samples were boiled 10 minutes to dissociate hydrogen bonds and denature the protein, so it would be easily detectable on a 15% acrylamide SDS-PAGE gel (Sigma).

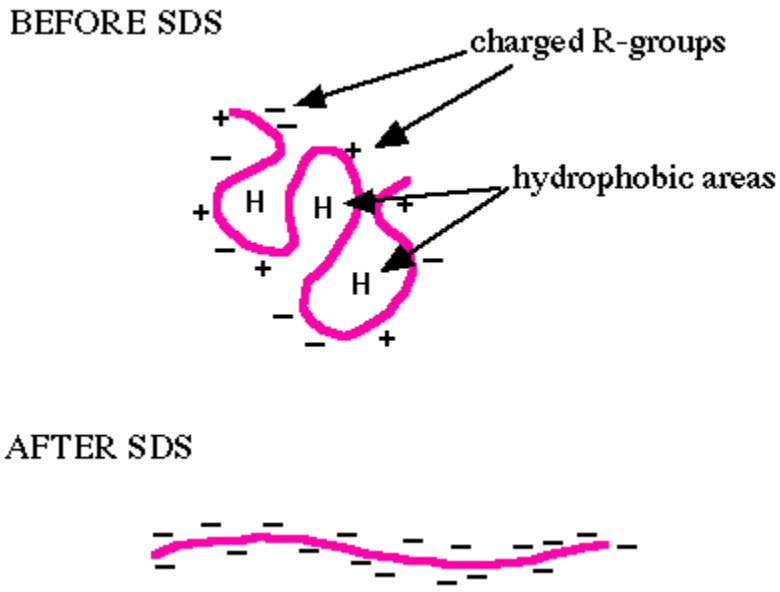


Figure 3.2 Shows the effect of adding SDS to proteins prior to running through a gel: negative charges from SDS disrupt positive charges by side chains of protein and the protein becomes negatively charged and linearized, which enables it to migrate to the positive pole during electrophoresis. Taken from [31].

3.2.6 Purification with Arginine Followed by Denaturing Conditions

Frozen cell pellet was thawed, suspended in PBS, benzonase nuclease, DTT and serine protease inhibitors, and sonicated 6 times, 20 seconds each time in five minute intervals. Sonication is a technique that breaks the bacterial cell walls in order to extract expressed protein trapped in the cytoplasm of the cell. High energy sound waves are applied to a sample in a water bath (for even distribution) which causes breakage of the cell membrane [32]. Cells were further broken by French press, which is another method that passes a sample through a high pressure chamber. The release of sample back to atmospheric pressure through the valve causes rupture of the cell wall, without the heat associated with sonication [32]. The mixture was then centrifuged at 1.2947×10^4 *g* for 30 minutes and supernatant was discarded. The cell pellets were then washed three times with 2% Triton X-100, a detergent used to clear debris, with centrifugation at 1.2947×10^4 *g* for 20 minutes between each washing. Cell pellet was observed as it turned whiter after every wash, and the supernatant was discarded each time. Next, the cell pellets were dissolved in 30 ml of 2 M arginine (pH 8.0), a solubilizer, and stirred overnight at a temperature of 4 °C, 25 °C, or 37 °C to completely solubilize the cell pellets.

The following morning, dialysis was performed twice for 2 hours with 10 mM Tris buffer (pH 7.4), and once overnight. Subsequently, the solution was centrifuged at 1.2947×10^4 *g* for 20 minutes, keeping the supernatant this time, and the pH was adjusted to 7.0. Then, 8 M urea was added to the protein and it was placed on an agitator overnight. Dialysis with 4 M urea was performed the next morning for 4 hours in order to refold the protein. Dialysis was performed with 2 M urea for 4 hours. For overnight dialysis, the buffer was changed to 10 mM Tris. The following day, the protein was dialyzed again with 10 mM Tris 2 times for 2 hours, and pH was lowered below 3.0 by HCl. Protein was then manually filtered with a 0.45 µM filter, and then

injected into a cation exchange column in the FPLC. Pure protein was then concentrated with Nitrogen Pressure in an Amicon concentrator (depending on the final concentration). This method is similar to studies done by Umetsu [33].

3.2.7 Purification Using Denaturing Conditions

Post analyzing the results from the purification method described in the previous section, a second method was developed, omitting the step with arginine. Frozen cell pellet was thawed, suspended in PBS, benzonase nuclease, DTT and serine protease inhibitors, and sonicated 6 times, 20 seconds each time in five minute intervals. Cells were further broken by French press. The mixture was then centrifuged at 1.2947×10^4 *g* for 30 minutes and supernatant was discarded. After the cell pellet is washed with Triton x-100 and centrifuged, it is dissolved in 8 M urea at room temperature for 2 hours. The protein is then placed in dialysis with 4 M urea for 2 hours at 4 °C overnight, followed by 2 M urea for 2 hours. Next, the protein is placed in dialysis 3 times with 10 mM Tris for 2 hours and is further purified by either the cation exchange (having lowered the pH and filtered the solution) or anion exchange (Q column) FPLC.

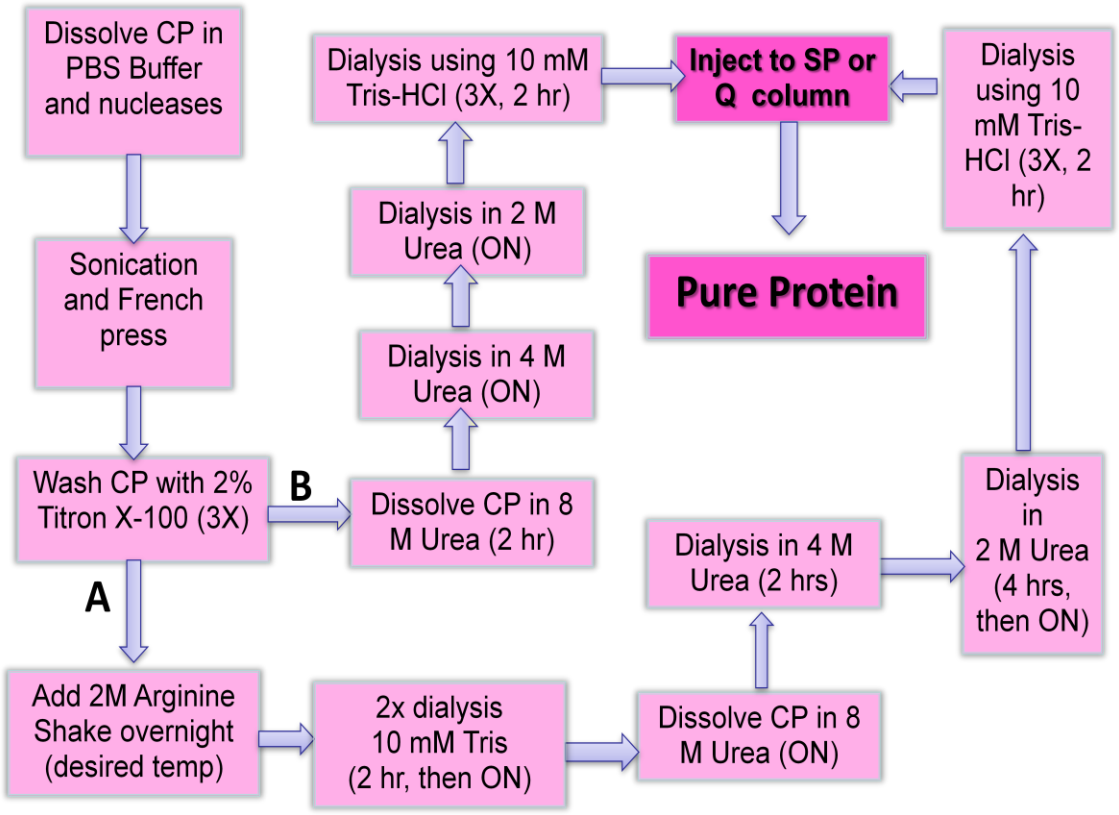


Figure 3.3 Summarizes the purification procedures via non-denaturing and denaturing conditions outlined in sections 2.2.5 and 2.2.6.

3.2.8 Determining Protein Concentration

The concentration of purified protein was determined by using a spectrofluorimeter (Shimadzu) to measure the UV absorbance in wavelength range of 350-220 nm. Two quartz cells (1 cm) filled with 800 µL nanopure water were measured from 350-220 nm to create a baseline from which absorbance differences would be recorded. One of these cells was left while the other was emptied and filled with protein. If the absorbance exceeded 2.0, which is outside the detection limits, the reading is inaccurate and a 1:16 dilution was usually performed to re-measure the sample.

CD2 and its variants have two active tryptophan residues capable of an absorption spectrum at 280 nm. When the absorbance spectrum resembles that shown in Figure 3.4, with a typical maximum peak at 280 nm and a "shoulder" (shown by the arrow), it is an indication that the protein has a good tertiary structure similar to the parent protein (CD2). Furthermore, this spectrum can indicate the level of DNA present in the sample. DNA has a typical absorbance at 260 nm. If the absorbance ratio 260/280 nm is greater than 1, the level of DNA is high and will interfere with protein function.

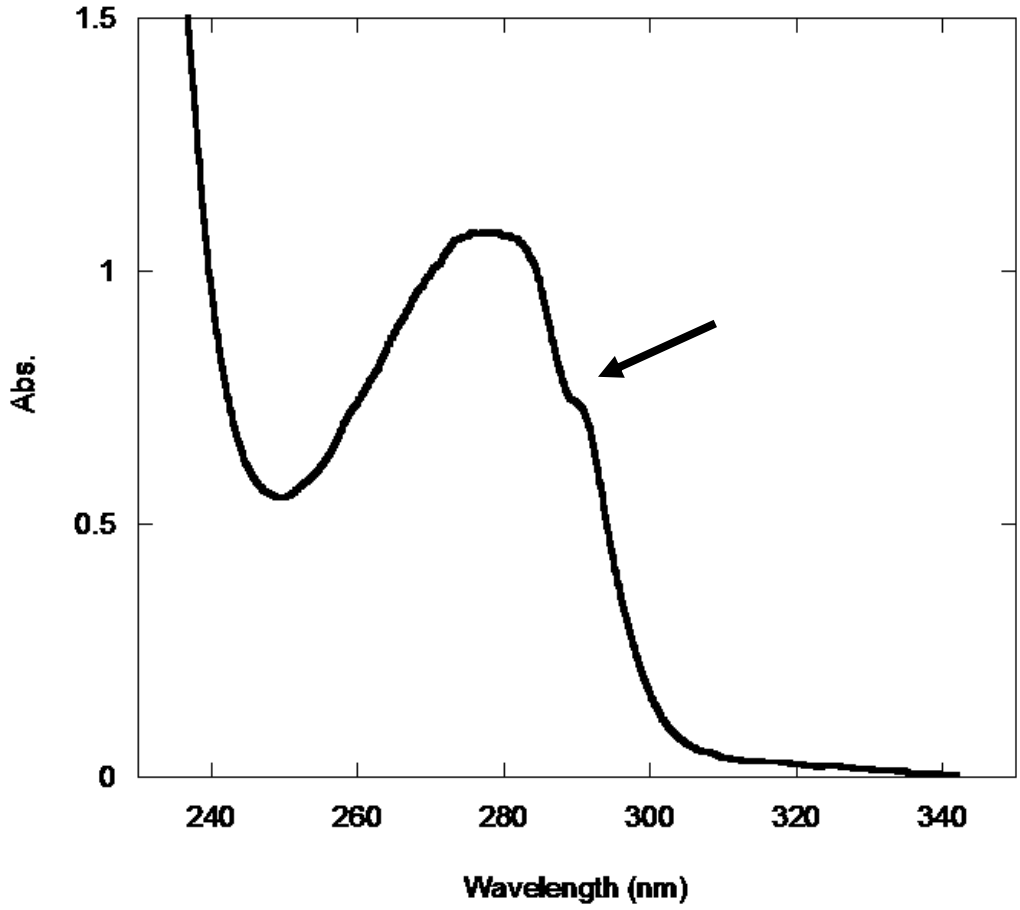


Figure 3.4 Typical fluorescence absorbance spectra of protein containing tryptophan residues with a maximum at 280 nm and a typical "shoulder" (arrow).

Protein concentration is obtained by subtracting the absorbance at 350 nm from the absorbance at 280 nm and then using Beer's law with a calculated extinction coefficient.

$$A = \varepsilon cl \qquad \text{(Equation 3.1)}$$

Using online data base Protein Calculator [34] and experimentally, the molar extinction coefficient (ε) was calculated based on the content of tryptophans, tyrosines, and disulfide bonds in the protein; for wild-type CD2, which lacks disulfide bonds, Protein Calculator estimates $\varepsilon = 13940\ M^{-1}cm^{-1}$. Experimentally, the value for CD2 and 7E15 (and its tagged version His-7E15), which has mutations that do not involve aromatic residues, is 11700 $M^{-1}\,cm^{-1}$. The coefficient for His-7E15-bom was determined by Julian Johnson in his thesis to be 16500 $M^{-1}cm^{-1}$ [35].

3.3 Results and Discussion

3.3.1 Expression of Tag-less 7E15

The expression of 7E15 was performed with tag-less vector pET20b at 30 °C overnight in LB medium with *E. coli* cell strain BL21 (DE3) pLysS. Overnight expression at a low temperature was chosen to enable the formation of inclusion bodies without the effect of degradation. The optical density graph is exponential to the point of induction, at about 0.600 at 600 nm. At induction, the cells received 0.25 mM IPTG and expressed overnight, until the harvest point. The SDS-PAGE shows that before induction, basal expression of other proteins was low, as is expected of pLysS cell strain. After the addition of IPTG, the engineered protein was produced as optical density increased, demonstrating a successful expression outcome (Figure 3.5). Figure 3.6 shows evidence that the conditions used for expression were optimal. At 30 °C overnight induction, the expression bands are thicker than at 37 °C for 4 hours.

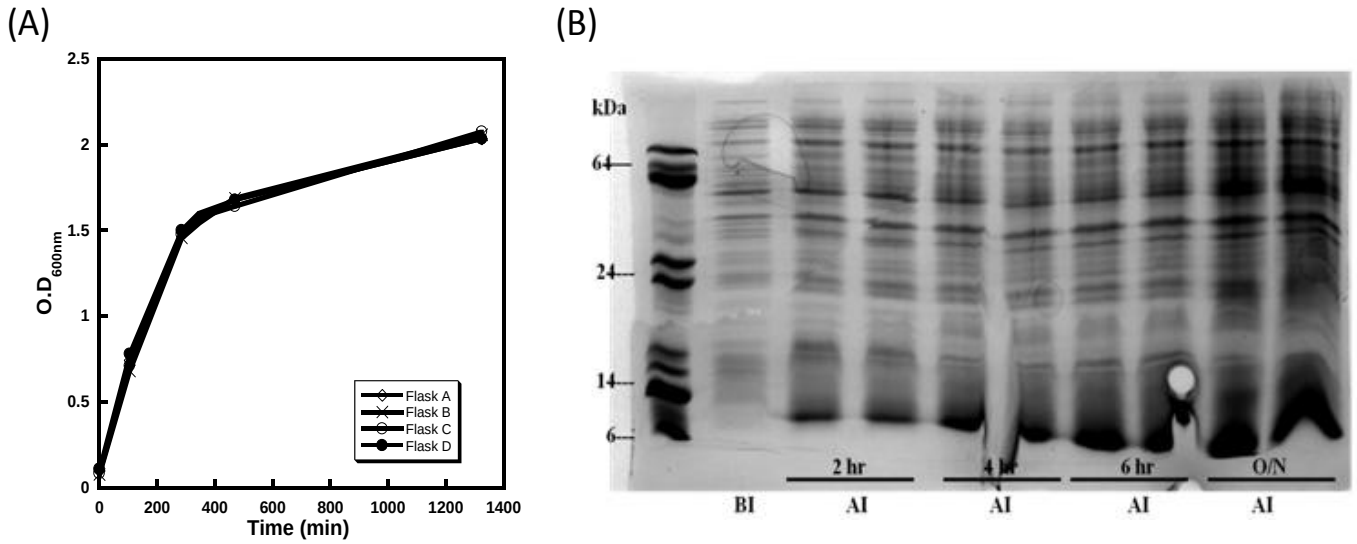


Figure 3.5 Expression of 7E15. (A) Growth curve indicative of exponential increase in optical density of BL21 (DE3) pLysS prior to induction and curve levels off post induction. (B) SDS-PAGE showing gradual production of protein with the passage of time, the thickest bands found after overnight growth.

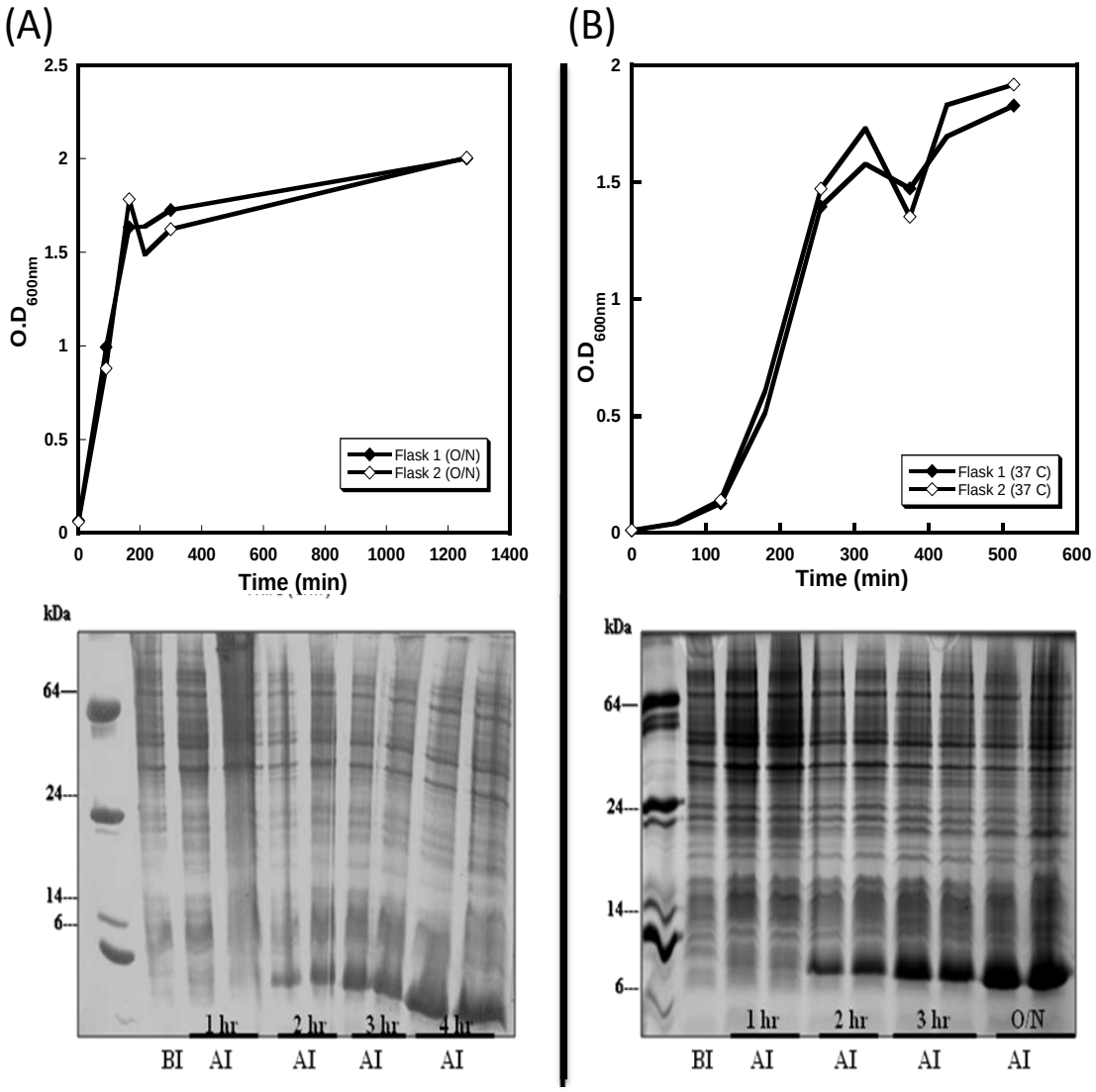


Figure 3.6 Expression of 7E15 with different temperature conditions. (A) Expression curve (above) and SDS-PAGE (below) show protein has been expressed at 37 °C for 4 hours. (B) Expression curve (above) and SDS-PAGE (below) showing improved expression when the temperature is lowered to 30 °C and the protein is expressed overnight.

3.3.2 Optimization of His-7E15 Expression

The expression level of His-7E15 was optimized by increasing the concentration of IPTG administered at induction. Although the optical density curve after 4 hours of expression with BL21 (DE3) cells at 37 °C is exponential, a low level of expression is seen after induction with IPTG concentrations of 0.2 mM, 0.4 mM, 0.6 mM, and 0.8 mM (Figure 3.7). However, expression under the same conditions except with IPTG concentration of 1 mM yielded thick bands on the SDS-PAGE after induction indicative of a high expression level (Figure 3.8). The expression product was mostly in inclusion body, especially because the expression level is high and because the temperature is high (37°C). This is evidenced on SDS-PAGE by the presence of protein in the cell pellet after the cell lysing step in purification rather than in the supernatant. The expression yield is higher than in the pGEX-2T vector. A possible reason for increased yield after raising the IPTG level in the pET30a vector may be tied to the affinity of IPTG to the repressor protein unique to each vector. Purification with size exclusion chromatography shows protein in fractions 11 and 12, while the peaks corresponding to fractions 7 through 9 are DNA. The yield was 44 mg per 1 gram of cell pellet.

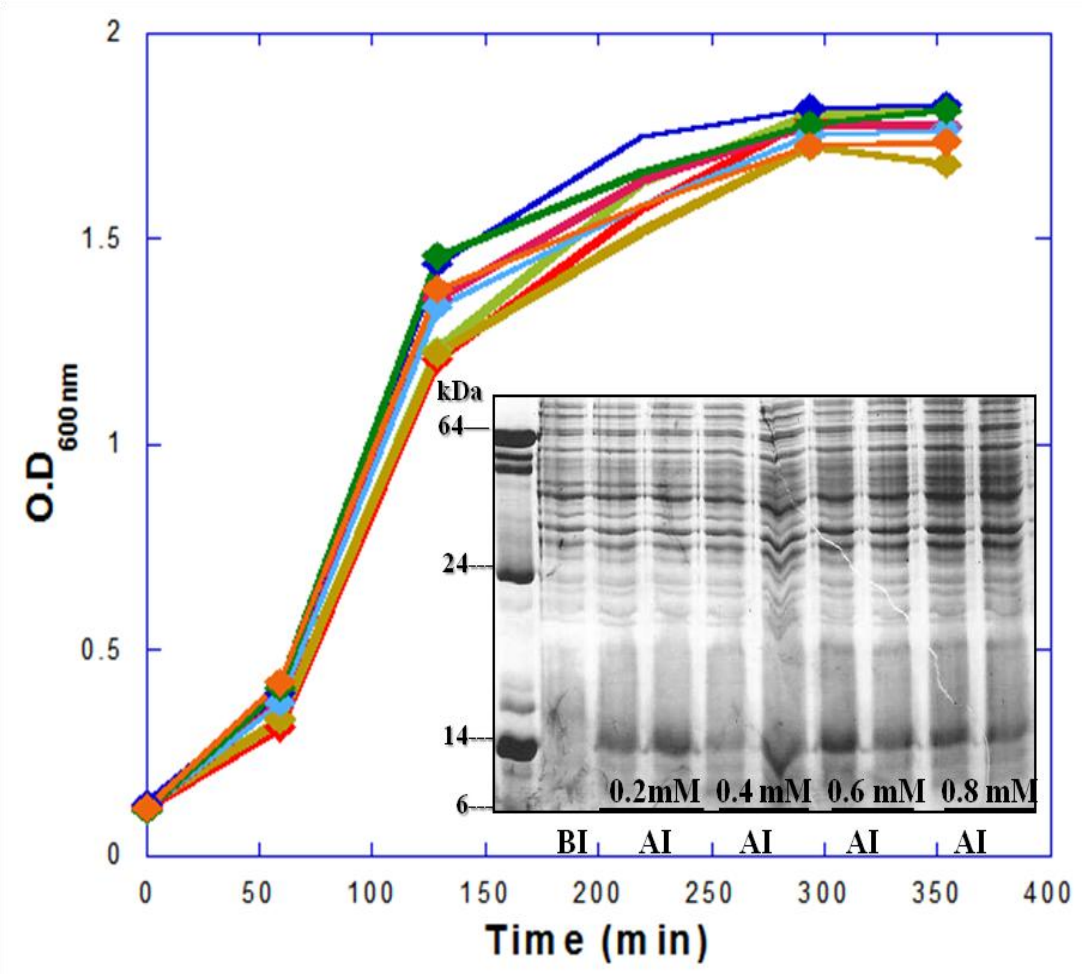


Figure 3.7 SDS-PAGE and optical density curve of the final point of expression of His-7E15 at 37 °C (4 hours) with BL21 (DE3) cells. The optical density curve is exponential, yet the gel shows a low level of expression after induction with IPTG concentrations 0.2-0.8 mM.

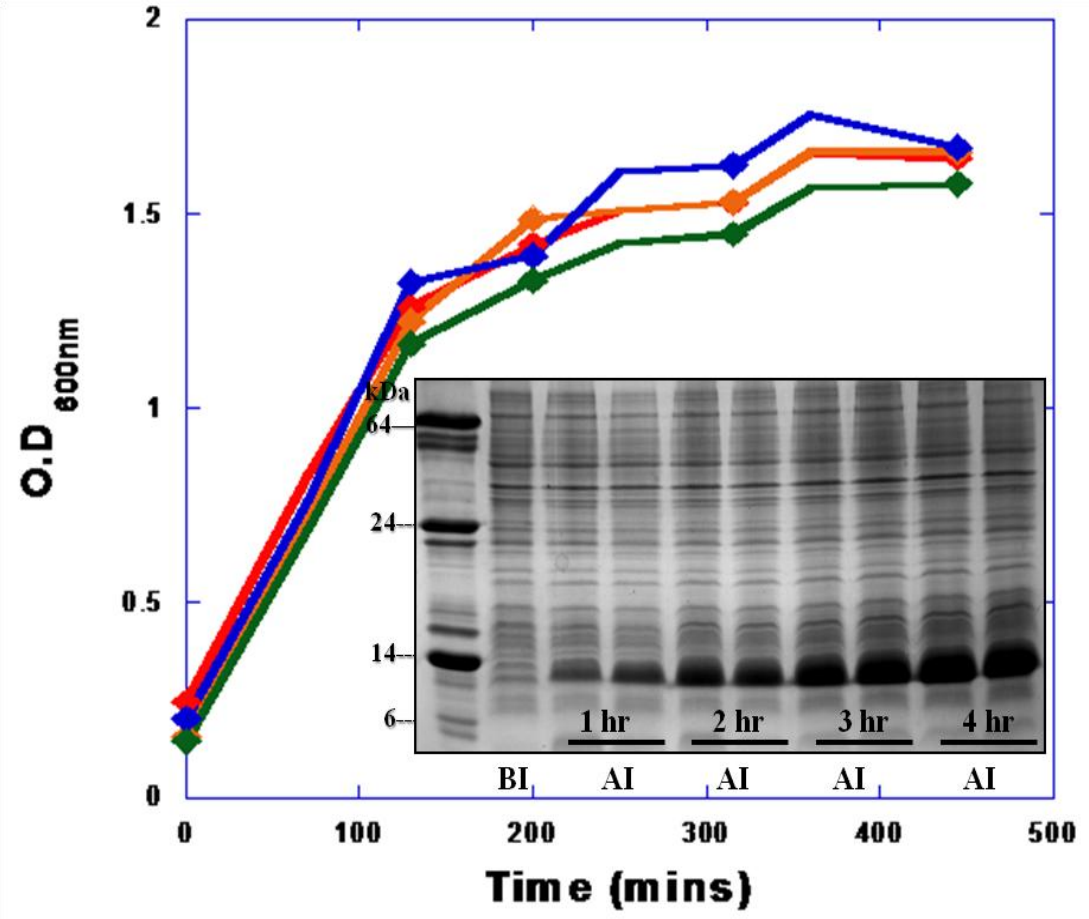


Figure 3. 8 SDS-PAGE and optical density curve of the final point of expression of His-7E15 at 37 °C (4 hours) with BL21 (DE3) cells. The optical density curve shows exponential growth, and the gel shows a thick band indicative of a high level of expression after induction with 1 mM IPTG.

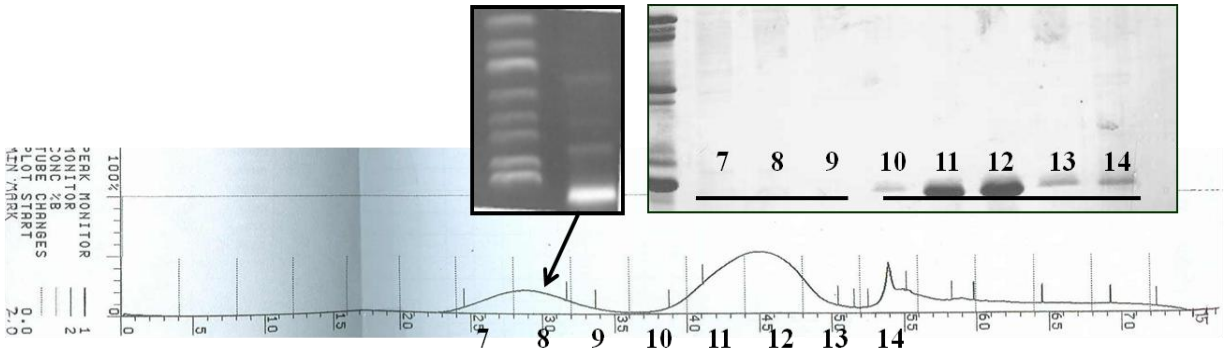


Figure 3.9 Purification of His-7E15 by size exclusion chromatography shows protein in fractions 11 and 12 and a DNA peak in fraction 8 (confirmed by DNA gel).

3.3.3 Purification of Tag-less 7E15 with Arginine Solubilization Followed by Denaturing Conditions

Cell pellet was subjected to French press followed by sonication for cell lysis. The inclusion bodies formed during expression were found in the cell pellet after centrifugation, and were washed with 2% Triton X-100, a detergent to remove lipid-soluble waste. After solubilization with 2 M arginine, a large quantity of the protein remains in the cell pellet and is not solubilized well. When the supernatant is solubilized in 8 M urea and placed in dialysis with 4 M urea, followed by 2 M urea, followed by 0 M urea (10 mM Tris buffer), there are less impurities on the SDS-PAGE band representing the protein (Figure 3.10).

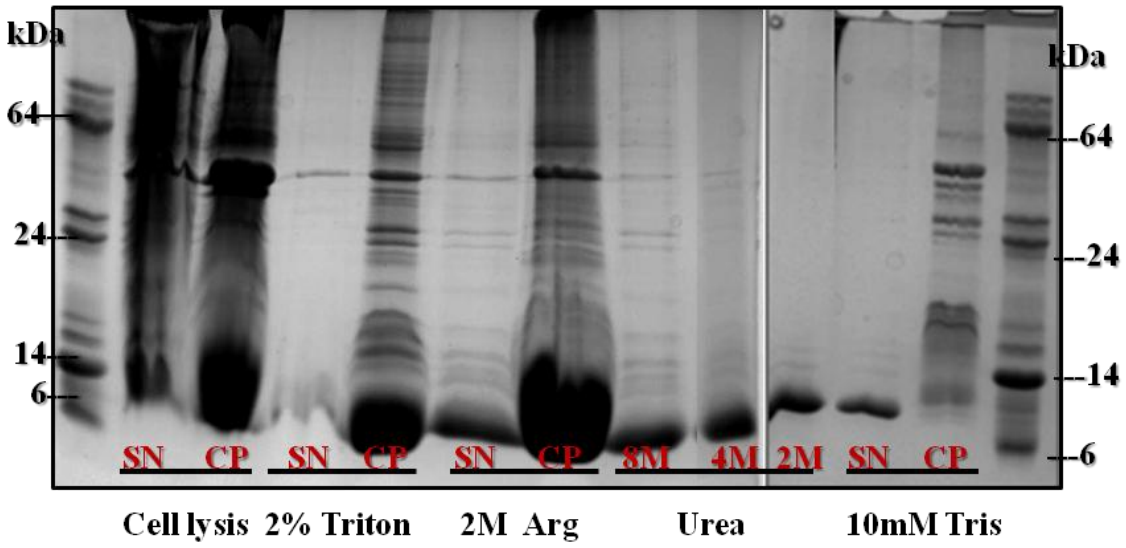


Figure 3.10 Purification of tag-less7E15 with arginine solubilization, followed by denaturing conditions. Following cell lysis, protein is in the cell pellet, which is washed with Triton X-100. Protein is dissolves poorly in 2 M arginine and following denaturation is refolded and pure (10 mM Tris samples).

Temperature-dependent arginine solubilization was tested in attempt to optimize protein solubilization in 2 M arginine. Cell pellet was solubilized in arginine in different temperatures and the effect of temperature on the quantity of solubilized protein was observed. To this end, cell pellet was purified and solubilized with arginine at 4°C, room temperature, and 37 °C for each temperature. The concentration of soluble protein was measured following solubilization and was found be to highest for room temperature, followed by 4 °C and then 37°C.

Table 3.2 Arginine solubilization temperature relative to concentration of soluble protein shows room temperature is favored for soluble protein.

Arginine Solubilization Temperature	Absorbance 280nm	Concentration (µM)
37°C	0.6588	47.26
25°C	0.7198	51.64
4°C	0.6362	45.64

Samples of each temperature were sent to 1D NMR and were run to examine protein structure. Tryptophans at positions 7 and 32, along with a tyrosine residue at position 76, are visible, typical findings in CD2 that suggest correct structure. At room temperature these peaks are seen with greatest resolution (Figure 3.11). A UV scan was performed to test the protein structure after refolding and to test the reliability of the refolding method since it is an unconventional way to purify protein (Figure 3.12). The tryptophan fluorescence spectrum below shows protein refolds back to its native conformation resembling CD2 (Figure 3.15). The total yield from 2.3 L expressed cell culture is shown below (257.8 mg). Per 1 g cell pellet, 11.68 mg protein were obtained, remarkably higher than the yield from GST purification by at least 10 fold (Figure 3.12).

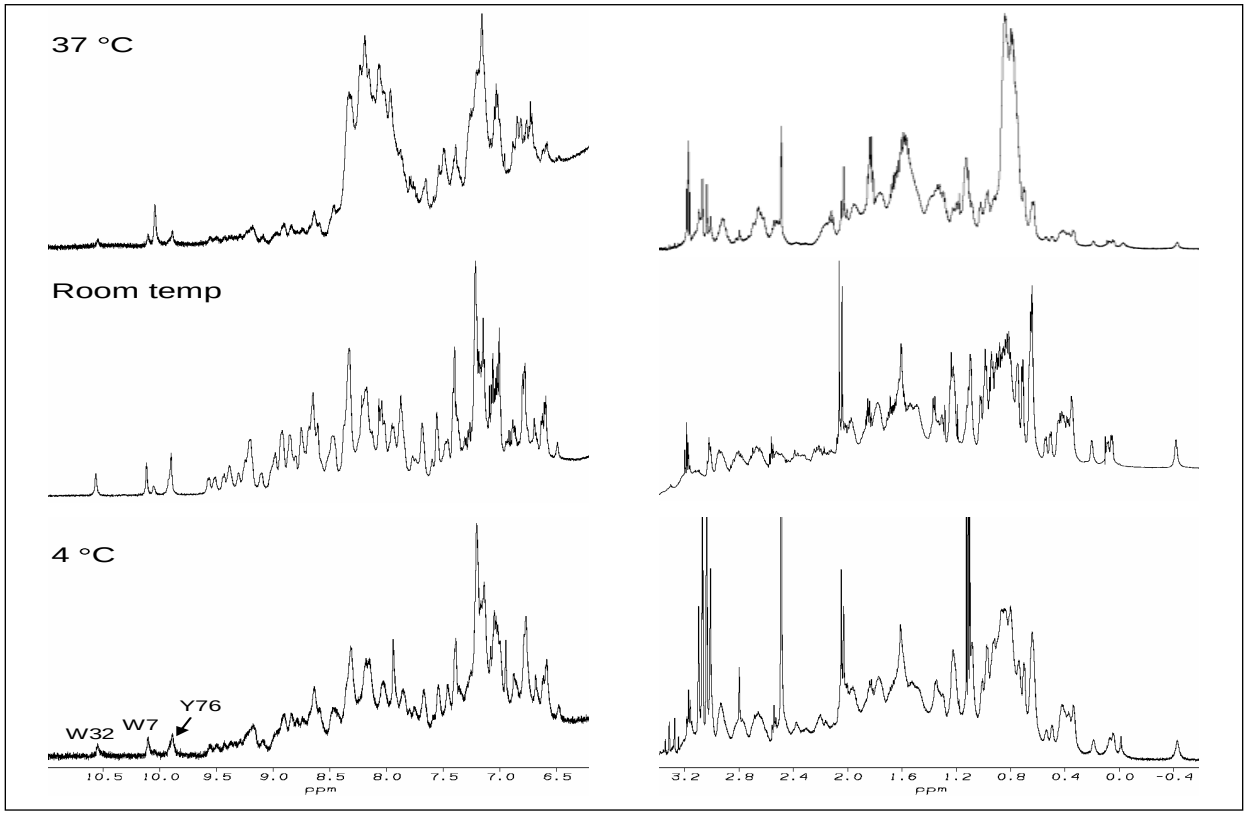
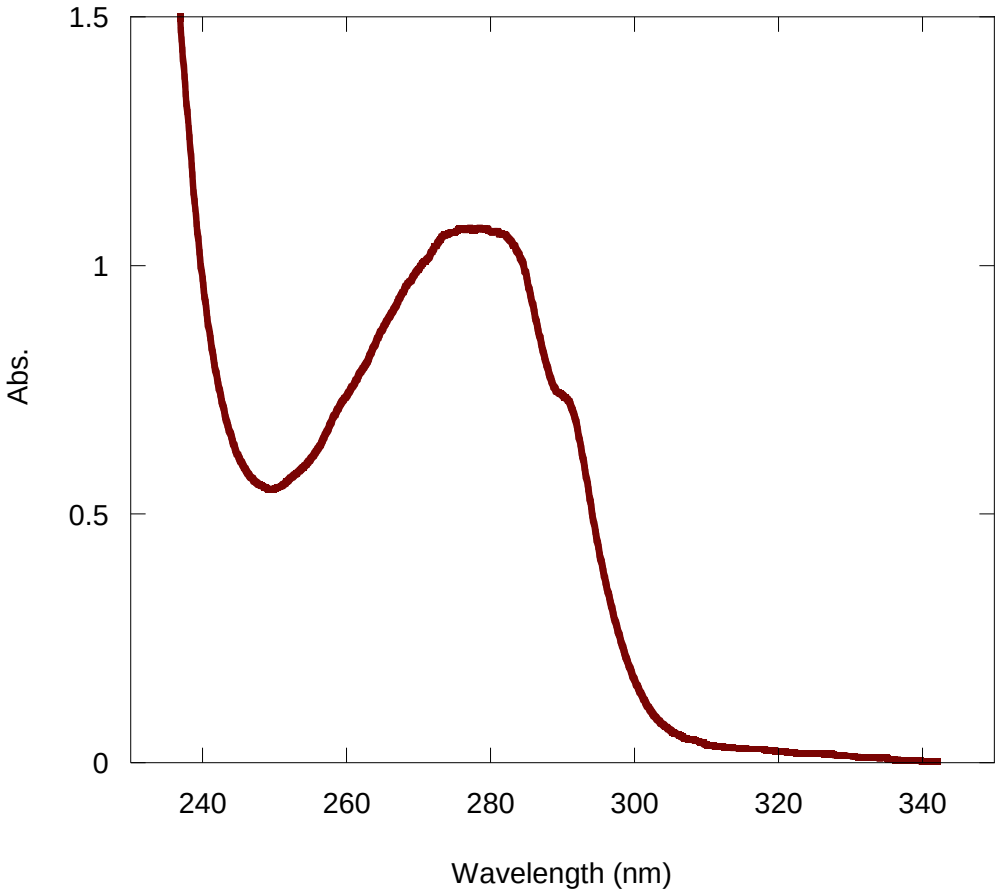


Figure 3.11 Arginine solubilizaton samples sent to 1D-NMR



Abs	Con. µM	Volume (ml)	Mass (mg)
1.068	91.32	250.0	257.8

Figure 3.12 UV spectrum of protein after arginine/ urea refolding method shows reliability of method as protein refolds back to its native conformation resembling CD2. The total yield from 2.3 L expressed cell culture is shown below (257.8 mg).

3.3.4 Purification of Tag-less 7E15 Under Denaturing Conditions

Cell pellet was purified omitting the solubilization step of arginine. After cell lysis the protein was found in the cell pellet as expected. Solubilization in urea was effective, and a small quantity of protein was found in the cell pellet after solubilization, in contrast to arginine solubilization, which left most of the protein in the cell pellet (Figure 3.13). The protein was refolded and soluble after dialysis with Tris.

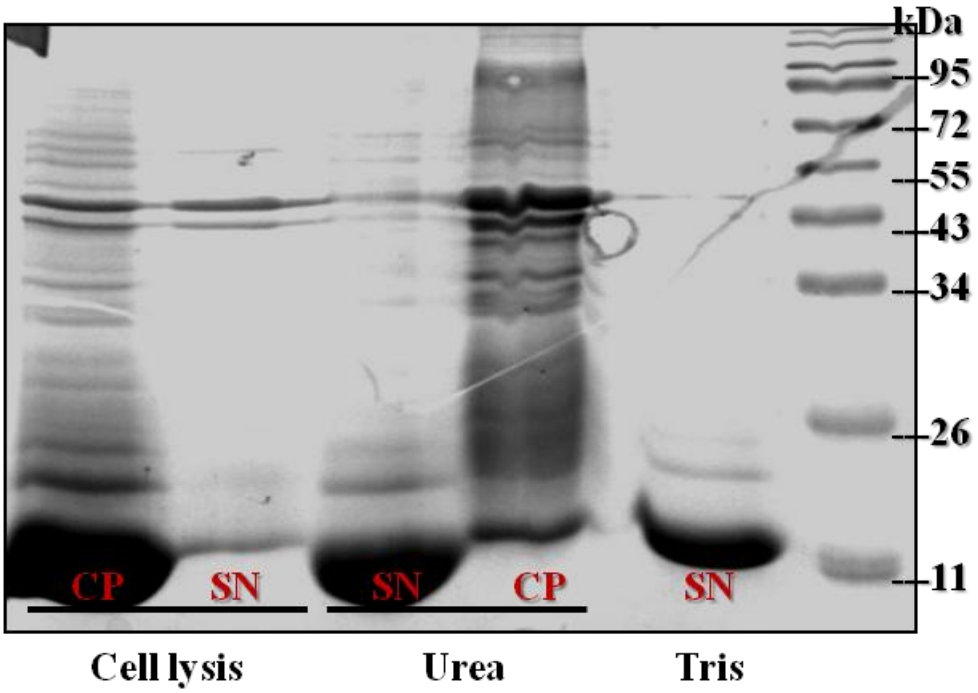


Figure 3.13 Purification of tag-less 7E15 under denaturing conditions. Protein was found in cell pellet after cell lysis, and was found in the supernatant post-solubilization with urea. The cell pellet after urea shows that solubilization was effective, as little protein remained insoluble.

Additional purification in SP and Q columns shows higher yields than obtained per liter of GST purification (Figure 3.14). The SP column gave 17.3 mg protein per 10 mL injection, which is theoretically 156 mg per liter of expressed cell culture. The Q column gave 25.9 mg per 10 mL injection, which is theoretically 234 mg per liter of expressed cell culture.

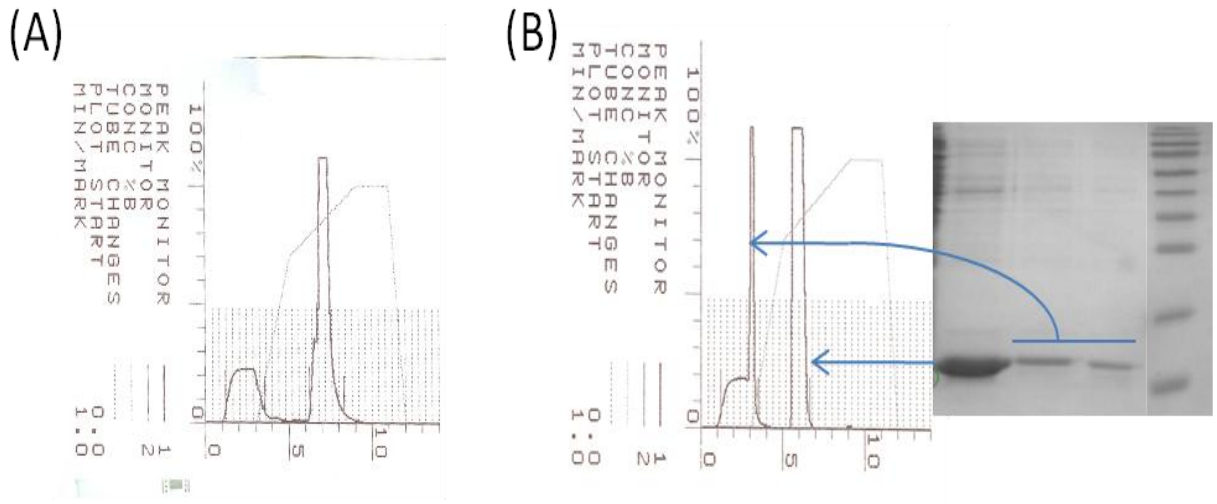


Figure 3.14 Chromatograms from FPLC purification of 7E15 show high yield. (A) Peak obtained in fractions 17-20 from a 10 mL injection into the SP column yields 17.3 mg protein. (B) Peak from a 10 mL injection into the Q column yields 25.9 mg protein in fractions 15 and 16. The arrows from the gel band correspond to the peaks on the chromatogram.

A CD spectrum and tryptophan fluorescence spectrum are the strongest evidence that the protein regains its secondary and tertiary native structure after denaturation with urea and that the refolding method is reliable. The CD structure of refolded 7E15 shows a typical β sheet conformation with a minimum point at 219 nm. This secondary structure is similar to what is typically seen in the CD structure of CD2 [36]. Additionally, compared to protein after GST purification, which has been shown to refold correctly by previous experimentation, the tryptophan fluorescence spectrums of both refolding purification methods are similar, indicating tertiary structure resemblance. A total of 266 mg of protein were obtained from purification. For every gram of cell pellet 11.77 mg protein were extracted via refolding, which is higher than arginine/urea purification and significantly higher than GST purification (Figure 3.15).

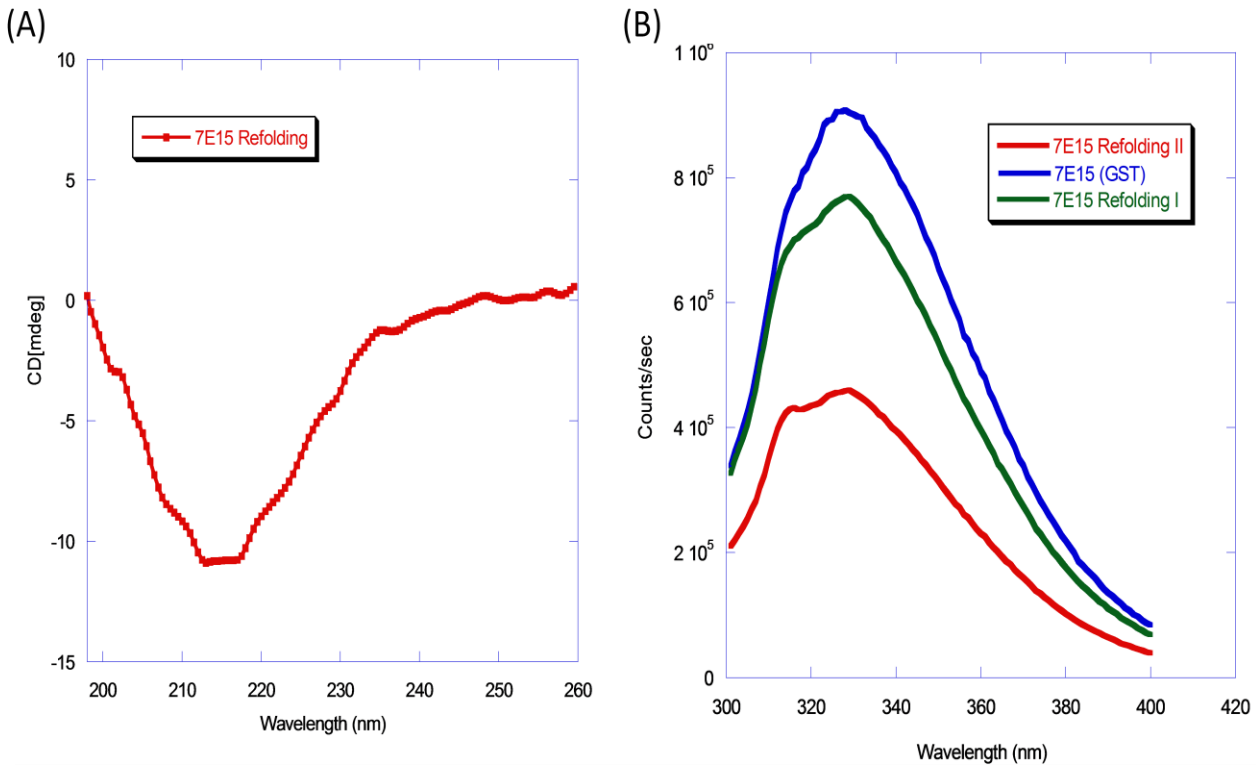


Figure 3.15 Structural studies of 7E15 post refolding purification. (A) CD spectrum of 7E15 shows a similar secondary structure conformation as seen in CD2 [36]. (B) Tryptophan fluorescence spectrums of 7E15 purified by different methods. The arginine/urea method (shown in green) and the urea method (shown in red) show tertiary structure resemblance to GST purification (shown in blue).

3.4 Conclusions and Future Work

A rapid, effective method for generating high yield of 7E15 has been established by generating and purifying inclusion bodies. The theoretical yield for a liter of expressed cell culture has risen impressively to 156 mg -234 mg pure protein from 1 L expressed cell pellet. The first step, over-expression of 7E15 as inclusion bodies, has been performed successfully under conditions established in chapter 2. Expression of His-tagged 7E15 has been optimized by adjusting the IPTG level to 1 mM.

Purification of obtained inclusion bodies via two methods was tested. The first method was solubilization in arginine followed by denaturation in urea. The objective was to test whether solubilizing the protein first would boost the level of soluble product after denaturation. The protein did not solubilize well in arginine, as indicated by SDS-PAGE, and only a small quantity of protein was extracted. Solubilization of protein in arginine seemed to depend on temperature, with more soluble amounts of protein at room temperature than at either extreme (37 °C or 4 °C). The temperature of solubilization in arginine was raised to room temperature to increase the level of solubilized protein. Purification by direct immersion of cell pellet in 8 M urea at room temperature showed a high level of solubilized protein. Protein which was not solubilized was centrifuged and separated.

The yields from arginine/ urea and urea were 11.68 mg/ 1 g cell pellet and 11.77 mg/ 1 g cell pellet, respectively, both of which are an improvement from traditional GST purification. Additionally, structural studies with CD and tryptophan fluorescence indicate that refolding methods can reliably generate protein that does not lose its native conformation.

Although purification yield from urea is not significantly greater than the yield from arginine/urea, it is still advisable to use the urea method for several reasons. First, optimization

of a rapid and efficient purification method includes removing extra steps which may cause loss of protein. Furthermore, the contribution of solubilized protein from the arginine step to the total yield is not significant enough to warrant an extra purification step. In fact, loss of protein in the cell pellet and low solubilization are reasons to omit this step.

CHAPTER 4: PROTEIN-BASED CONTRAST AGENTS

4.1 Site-Directed Mutagenesis to Create 7E15 Charged Variants

Charged variants of ProCA1 (or 7E15) were created by Dr. Anna Wilkins Maniccia utilizing the design approach with site-directed mutagenesis [17]. One of the intentions of creating 7E15 charged variants was to study whether varying the charge or coordination residue type in the metal binding pocket would affect calcium binding compared to the template, 7E15, within a fixed binding site [15]. Drs. Anna Maniccia and Wei Yang have shown that calcium binding affinity is increased as the charge in the binding pocket is increased. CD2.7E15 has a strong metal binding affinity for Tb^{3+} (K_d~10^{-10} M), as shown by FRET, and a weak calcium binding affinity (K_d of 0.15 mM). In later work, CD2.7E15 was shown to be an excellent MRI contrast agent with a strong Gd^{3+} binding affinity (K_d~10^{-10} M), as demonstrated with a metal buffer system, and significantly improved relaxivity [1].

It is not clear whether other charged variants are good candidates for MRI contrast agents. Among several available charged variants created, three -5 and -4 charged variants have a strong potential as MRI contrast agents. Residue 64 in 7E15 was changed to glutamate (called EEDDE in [15;17] or 7E15E throughout this paper) to create a -5 pair, 7E15/7E15E that differs by a methylene group. This residue was changed to glutamine or aspargine (called EEDDQ and EEDDN in [15;17] or 7E15Q and 7E15N throughout this paper), a -4 pair that also differs by a methylene group. The objective of this chapter is to examine Tb^{3+}, Gd^{3+}, and Zn^{2+} binding with different charged variants and to examine the suitability of these proteins as the next contrast agents.

Table 4.1 Summary of the residues in 7E15 charged variants; red refers to positive, blue refers to negative, and black refers to neutral amino acids [15].

Residues						
Protein	**15**	**56**	**58**	**62**	**64**	**Charge**
CD2WT	**N**	**E**	**L**	**D**	**K**	**0**
7E15	**E**	**E**	**D**	**D**	**D**	**-5**
7E15E	**E**	**E**	**D**	**D**	**E**	**-5**
7E15Q	**E**	**E**	**D**	**D**	**Q**	**-4**
7E15N	**E**	**E**	**D**	**D**	**N**	**-4**
EENDN	**E**	**E**	**N**	**D**	**N**	**-3**
NENDN	**N**	**E**	**N**	**D**	**N**	**-2**

4.2 Methods and Materials

4.2.1 Cloning and Transformation

The genes for 7E15 and its charged variants were cloned with PCR by Dr. Anna Wilkins Maniccia in accordance with the protocol in her dissertation [37]. The transformation procedure took place as described in section 2.2.2. The exceptions are that BL21 (DE3) *E.coli* competent cells and vector pET-20b, designed to contain no affinity tag for purification, were used instead.

4.2.2 Over-expression

A bacterial colony was inoculated and incubated overnight in LB medium with 2.5 mM ampicillin. Flasks containing 1 L LB medium and 2.5 mM ampicillin received 50 mL of the culture grown overnight, and the 1 L flasks were incubated at 37 °C in a shaker (Shaker Controlled Environment Incubator).

The optical density (OD) of each sample at 600 nm was recorded every 30 minutes using UV spectrophotometer (Shimadzu Scientific Instruments, Norcross, GA). When the OD reached 0.6 for BL21 (DE3), Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to cultures. The BL21 (DE3) strain required 0.25 mM IPTG. The shakers were adjusted to 37 °C for the duration of expression. The cell pellets were collected by centrifugation at 4.410×10^3 *g* for 20 minutes after the cultures have grown overnight, or when the cell population began to level off. The cell pellets were stored at -20 °C until purification. KaleidaGraph software (Synergy Software, Reading, PA) was used to plot the optical density versus time plots. Image J software (Version 1.37) was used to quantify band thickness as a result of expression on the 15% acrylamide SDS-PAGE gels.

4.2.3 Purification Via Urea Refolding Method

Purification was performed using the refolding method as described in section 2.2.7 with modification. Cell pellet from expression (approximately 5 g) was suspended in lysis buffer containing PBS. Benzonase nuclease (60 µL) and serine inhibitor PMSF (200 µL) were added. The cell pellet was broken using sonication (6 times for 20 seconds each), French press (1500 psi), or cell disruption. The mixture was then centrifuged (S34 rotor) at 1.2947 x 10^4 *g* for 30 minutes and supernatant was discarded. The cell pellets were then suspended in 20 mL of 2% Triton X-100, and centrifuged at 1.2947 x 10^4 *g* for 20 minutes between each of three washings. Next, the cell pellet was dissolved in 25 mL of 8 M urea (pH 7.3) and stirred overnight at 4 °C to completely solubilize the cell pellet.

The solublized cell pellet was centrifuged for 20 minutes at 1.2947 x 10^4 *g*. PBS was then added drop-wise for a final concentration of 4 M urea and the solution was stirred for 2 hours. Next, the solution was placed in dialysis for 4 hours with 2 M urea, followed by addition of proteases (200 µL) and overnight dialysis with 2 L 10 mM Tris-HCl buffer (pH 7.4). The dialysis bag has pores which allow molecules with molecular weight over 3.5 k Da to remain in the bag. The dialysis buffer was changed to 10 mM Tris the following morning, which was replaced to gradually remove urea from the solution. Subsequently, the solution was centrifuged at 1.2947 x 10^4 *g* for 20 minutes, and the supernatant was kept.

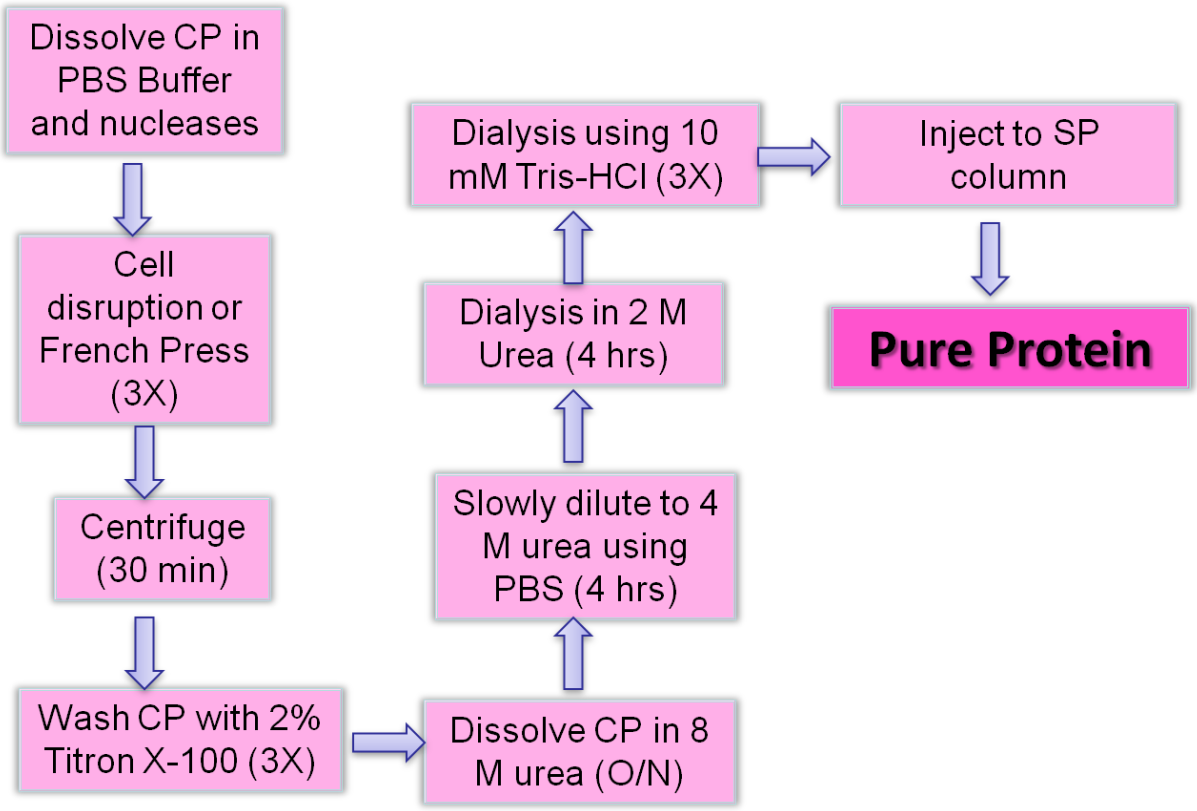


Figure 4.1 Summary of the urea refolding method

4.2.4 Anion Exchange Chromatography Purification Based on Salt Gradient

The Q-column was used in an ÄKTA prime FPLC instrument (Amersham Biosciences, Piscataway, NJ). The supernatant was filtered using a 0.45 µM manual filter. Protein was injected into the column, and the flow rate was 1-2 ml/min. The binding buffer, buffer A (pH 8.5) contained no salt while the elution buffer, buffer B (also pH 8.5), contained 1M NaCl. The column was washed with buffer A over three column volumes (CV), followed by a gradually increasing percentage of buffer B to 70% over 9 CV. Lastly, the column was washed with buffer B over 1-2 CV. The protein eluted in 1.5 mL fractions.

4.2.5 Cation Exchange Chromatography Purification Based on pH Gradient

The conditions for the SP column were as follows. The pH of the protein was lowered to 3.0 by adding HCl, and the protein was filtered manually with a 0.45 µM filter. The protein was injected and the flow rate was 1-2 ml/min. Buffer A has a pH of 4.0 while Buffer B has a pH of 8.0.

4.2.6 Far Ultraviolet Circular Dichroism

The purpose of this experiment was to observe the effect of adding calcium on the conformation of 7E15 variants with a circular dichroism (CD) instrument (JASCO, J-810 spectropolarimeter). Protein samples were diluted with 20 mM Tris (pH 7.4) buffer to a final concentration of 50 µM. The first sample contained 50 uM protein and 2 mM EGTA in 200 uL buffer. The second sample contained 50 µM protein and 2 mM Ca^{2+} in 200 uL buffer. The samples were placed in a 2 cm UV cell and scanned from 190-260 nm. Ten scans were averaged at 50 nm per minute. The spectra were corrected by subtraction from buffer samples.

4.2.7 Determination of K_d of Charged Variants to Tb^{3+} Using Fluorescence Resonance Energy Transfer (FRET) Assay

The binding affinity of CD2.7E15 and its variants to terbium was determined using spectra obtained from FRET assay with a SC-500 PTI spectrofluorimeter. Two sets of titrations each were performed in triplicate with 7E15 and 7E15N. The first set was performed in 10 mM chelex Tris-HCl, pH 7.4, while the second set was performed in 20 mM PIPES, 10 mM KCl, pH 6.8. In both cases, the final protein concentration was 5 µM, and the cuvette had 1 cm path length quartz cell. Incubation time upon addition of more terbium was 5 minutes. The excitation was set to 282 nm, a tryptophan excitation wavelength, while the emission was set to 500-600 nm. The integration time was 0.2 seconds. To reduce Raleigh scattering, a glass plate was used as a filter. The excitation and emission slit widths were 1.0 and 1.8 mm, respectively.

The K_d of 7E15 to terbium was calculated after correcting the baseline with Kaleidagraph by subtracting free terbium from bound terbium. The data was fitted according to the Equation 4.20, as published, which was derived from the following equations where P: Protein; M: Metal Ion; MP: Complex of protein and metal ion; F: Fraction of complex in total protein [38]:

$$MP \leftrightarrow M + P \quad \text{(Equation 4.1)}$$

$$K_d = \frac{[M][P]}{[MP]} \quad \text{(Equation 4.2)}$$

$$F = \frac{[MP]}{[P]_T} = \frac{[MP]}{[P] + [MP]} \quad \text{(Equation 4.3)}$$

[MP] is the only species that fluoresces and gives off intensity. F defines the formation of [MP] based on how much metal is added. The fraction increases based on how much complex is formed.

$$[MP] = [P]_T F \quad \text{(Equation 4.4)}$$

$$([P] + [MP])F = [MP] \quad \text{(Equation 4.5)}$$

$$[P] = \frac{[MP]}{F} - [MP] \quad \text{(Equation 4.6)}$$

$$[P] = [MP] * (\frac{1}{F} - 1) \quad \text{(Equation 4.7)}$$

Using Equation 4.4:

$$[P] = [P]_T F * (\frac{1}{F} - 1) \quad \text{(Equation 4.8)}$$

$$[P] = [P]_T * (1 - F) \quad \text{(Equation 4.9)}$$

$$[M] = [M]_T - [MP] \quad \text{(Equation 4.10)}$$

$$[M] = [M]_T - [P]_T F \quad \text{(Equation 4.11)}$$

$$K_d = \frac{[[P]_T*(1-F)]\,[[M]_T-[P]_T F\,]}{[P]_T F} \quad \text{(Equation 4.12)}$$

$$K_d = \frac{([M]_T-[P]_T F)([P]_T*(1-F))}{[P]_T F} \quad \text{(Equation 4.13)}$$

$$K_d = \frac{([M]_T - [P]_T F)(1 - F)}{F} \quad \text{(Equation 4.14)}$$

$$K_d F = ([M]_T - [P]_T F)(1 - F) \quad \text{(Equation 4.15)}$$

$$K_d F = [M]_T - [M]_T F - [P]_T F + [P]_T F^2 \quad \text{(Equation 4.16)}$$

$$0 = [M]_T + (-[M]_T - [P]_T - K_d)F + [P]_T F^2 \quad \text{(Equation 4.17)}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad \text{(Equation 4.18)}$$

$$F = \frac{([\mathrm{P}]_\mathrm{T} + [\mathrm{M}]_\mathrm{T} + \mathrm{K_d}) - \sqrt{(-[P]_T - [M]_T - K_d)^2 - 4([P]_T[M]_T)}}{2([P]_T)} \quad \text{(Equation 4.19)}$$

K_d can be estimated by curve fitting of F~$[M]_T$ with a constant protein concentration.

The equation we use in Kaleidagraph is:

$$F = \frac{((M2 + M0 + M1) - \sqrt{(M2 + M0 + M1)^2 - (4 * M2 * M0)}}{2 * M2} \quad \text{(Equation 4.20)}$$

Where M1=K_d; M2=$[P]_T$; M0= $[M]_T$

F is proportional to intensity of fluorescence.

$$F = \frac{[MP]}{[P]_T} = \frac{[MP]}{[P] + [MP]} \quad \text{(Equation 4.21)}$$

In the fitting curve, fluorescence intensity is accounted for by multiplying each side by 10^5. The equation becomes:

$$F = M3 * \left[(M1 + M2 + M0) - \sqrt{(M1 + M2 + M0)^2 - (4 * M1 * M0)}\right] / 2 * M1 \quad \text{(Equation 4.22)}$$

Where M1= $[P]_T$; M2= K_d; M3= Maximum Intensity, i.e 6 x 10^5 [38].

4.2.8 Determination of Affinity of 7E15 Charged Variants to Gd^{3+} and Zn^{2+} by Competition with Molecular Probes Dyes

The Gd^{3+} binding affinity of charged variants was tested using a competition assay with fluorescent dye Fluo-5N, which has a published Gd^{3+} K_d of 3.8 $\pm$ 0.2 x 10^{-12} (Molecular Probes, Invitrogen). Experiments were modeled after previously published data [39]. The titrations were performed in duplicate using a SC-500 PTI spectrofluorimeter. The emission scans were from 500-650 nm and a 1 cm path length cell was used. The excitation wavelength was 488 nm, and the excitation slit widths ranged from 0.31-0.36 mm while emission slit widths ranged from 0.41-0.46 mm. The integration was 0.2 seconds, the step size was 1, and the incubation time was 5 minutes between each scan. The glass slide was left out of the fluorimeter. The buffer conditions were 10 mM Tris-HCl, pH 7.4 or 20 mM PIPES, 10 mM KCl, pH 6.8. The Fluo-5N and Gd^{3+} concentrations were 10 µM or 5 µM, diluted with buffer.

Fluo-5N and Gd^{3+} were mixed together in a 1:1 ratio for all experiments. Gd^{3+} was competed out with by gradual addition of 7E15 charged variants in 20 µM intervals until saturation, a final concentration of 150 µM. The effect of dilution was accounted for and calculated using excel. The emission scan showed a maximum fluorescence intensity at 519 nm, which was fitted and plotted versus protein concentration using Kaleidagraph to obtain the apparent dissociation constant, K_{app}, an approximation of K_d that does not include the binding affinity of Fluo-5N and Gd^{3+}. The 1:1 fitting equation used is a modified version of Equation 4.22, where M1= $[P]_T$; M2= K_d; M3= Maximum Intensity, i.e 6 x 10^5:

$$F = 1 - \left\{ M3 * \left[(M1 + M2 + M0) - \sqrt{(M1 + M2 + M0)^2 - (4 * M1 * M0)} \right] / 2M1 \right\}$$

(Equation 4.23)

The K_d value was obtained using the following published equation, where K_{d2} is the protein affinity, K_{d1} is the Gd^{3+}- Fluo-5N affinity [39]:

$$K_{d2} = K_{app} * \frac{K_{d1}}{K_{d1} + [Fluo - 5N]_T} \qquad \text{(Equation 4.24)}$$

As a control experiment, a titration with EDTA was performed. As well, a titration was performed where the intensity of Fluo-5N (minimum intensity) and the intensity of Gd^{3+}-Fluo-5N (maximum intensity) were graphed to characterize the intensity range of the experiment.

The zinc affinity of charged variants was tested using a competition assay with fluorescent dye FluoZin-1, which has a published Zn^{2+} K_d of 8.2 µM (Molecular Probes, Invitrogen). The buffer conditions were 10 mM Tris-HCl pH 7.4, the excitation slit widths ranged from 0.25-0.30 mm, while emission slit widths were 0.50. The emission was recorded from 500-650 nm while the excitation wavelength was 495 nm. The protein was added in 8 µM

intervals until saturation, a final concentration of 125 µM (approximately). The effect of dilution was accounted for and calculated using excel.

4.2.9 Relaxivity Measurements

T_1 and T_2 measurements were performed with a Bruker Minispec and an mq60 NMR analyzer, at 25 °C, at a field strength of 1.4 T or 60 MHz. The protein concentration was varied against a fixed gadolinium (50 µM). The buffer was 10 mM HEPES, pH 7. Protein concentration versus $1/T_1$ or $1/T_2$ was plotted in KaleidaGraph and samples with protein to gadolinium ratio of 1:1, 1.5:1, 2:1, and 3:1, the saturation point of the relaxivity plot, were used to determine the average relaxivity. T_1 and T_2 were fitted using Equations 4.25 [46] and 4.26 [46], respectively, where M_0 represents the initial magnetization. Relaxivity was calculated using Equation 1.6.

$$Y = M_0 * e^{-t/T_1} \qquad \text{(Equation 4.25)}$$

$$Y = M_0 * \left(1 - 2e^{-t/T_2}\right) \qquad \text{(Equation 4.26)}$$

4.3 Results and Discussion

4.3.1 Expression of CD2.7E15 Charged Variants

CD2.7E15 charged variants, 7E15, 7E15N, 7E15E, 7E15Q were expressed in LB medium using BL21 (DE3) *E. coli* cell strain and tag-less vector pET20b at 37 °C overnight. The conditions were selected with the intention of producing inclusion bodies for purification. When the optical densities reached 0.600 at 600 nm, the flasks were induced with 250 µM IPTG (Figures 4.2 A). Following overnight expression, they were harvested. The SDS-PAGE gels show thickening bands with the passage of time, which illustrates that protein was formed during expression and specifically after induction with IPTG (Figures 4.2 B).

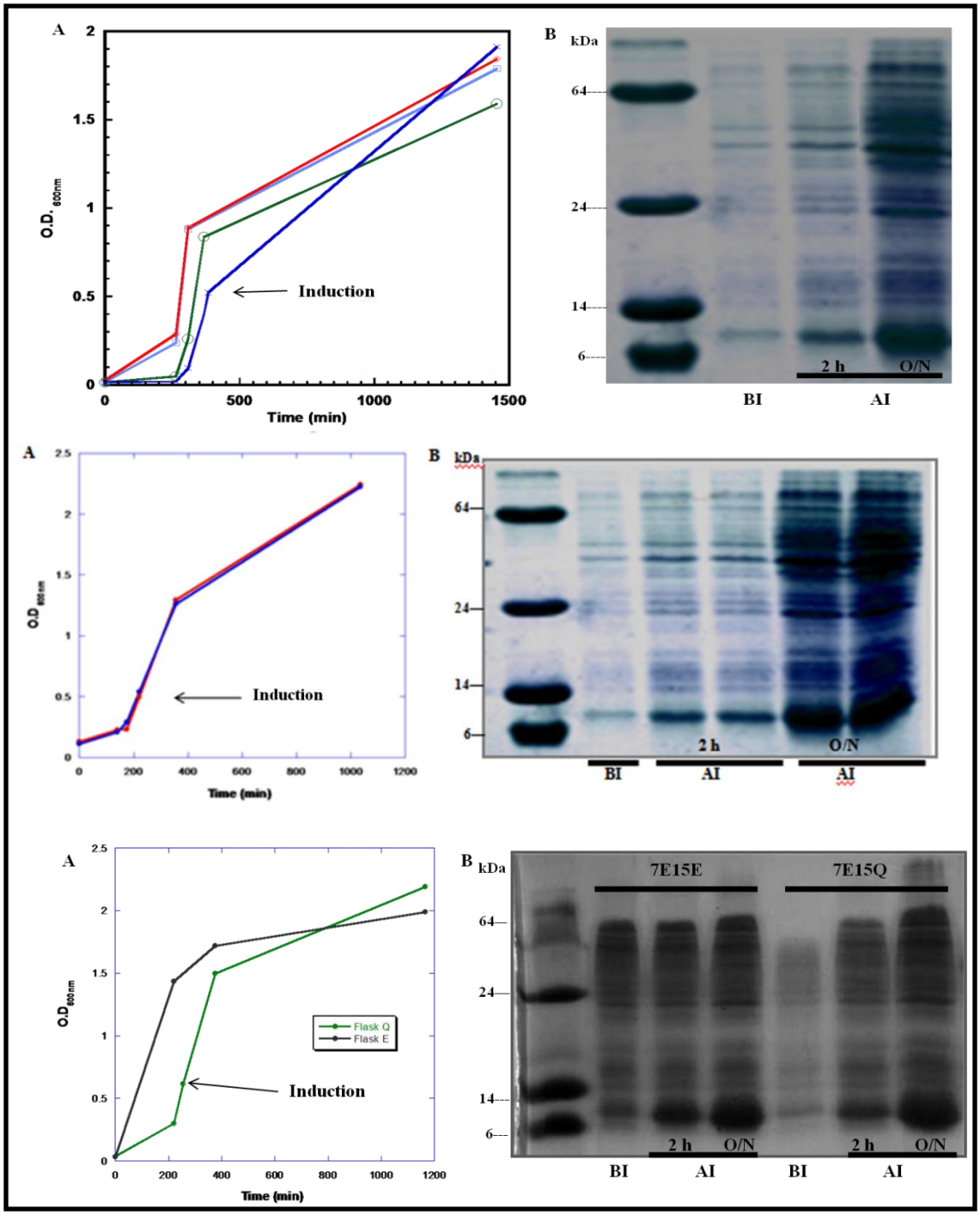


Figure 4.2 Expression of 7E15 (top), 7E15N (middle), and 7E15E and 7E15Q (bottom). (A) Optical density curve showing steady increase of cells at 37 °C with tag-less vector pET20b and induction point. (B) SDS-PAGE gel showing gradual thickening of 7E15-variant bands after induction.

4.3.2 Purification, Mass Spectra, and Circular Dichroism of Analysis CD2.7E15 Charged Variants

CD2.7E15 charged variants were purified using the refolding method followed by cation exchange chromatography. SDS-PAGE gel shows that following French press, proteins were in cell pellet, due to inclusion body formation during expression. Following wash with 2% Triton X-100, proteins remained in the cell pellet and a negligible quantity was lost in the supernatant (discarded). Addition of 8 M urea dissolved the pellets, and proteins were found in the supernatant. Gradual removal of urea resulted in refolding proteins in the soluble form, and following dialysis in 10 mM Tris-HCl (pH 7.4) and centrifugation, most protein was found in supernatant, and any misfolded protein was found in the cell pellet (discarded). Protein was injected into the cation exchange column (FPLC system) after adjusting the pH to 3.0 with HCl. The peak in fractions 29-31 was found to contain protein with high purity (confirmed by SDS-PAGE) (Figure 4.3). The final average concentrations and yields (in mg) for 7E15 charged variants are shown in Table 4.2. The proteins were analyzed with an ABI 4800 MALDI TOF/TOF mass spectrometer by Yanyi Chen to verify refolding after purification. The experimental masses shown on the MS spectra are within less than 2% error than the calculated weights (Figure 4.4).

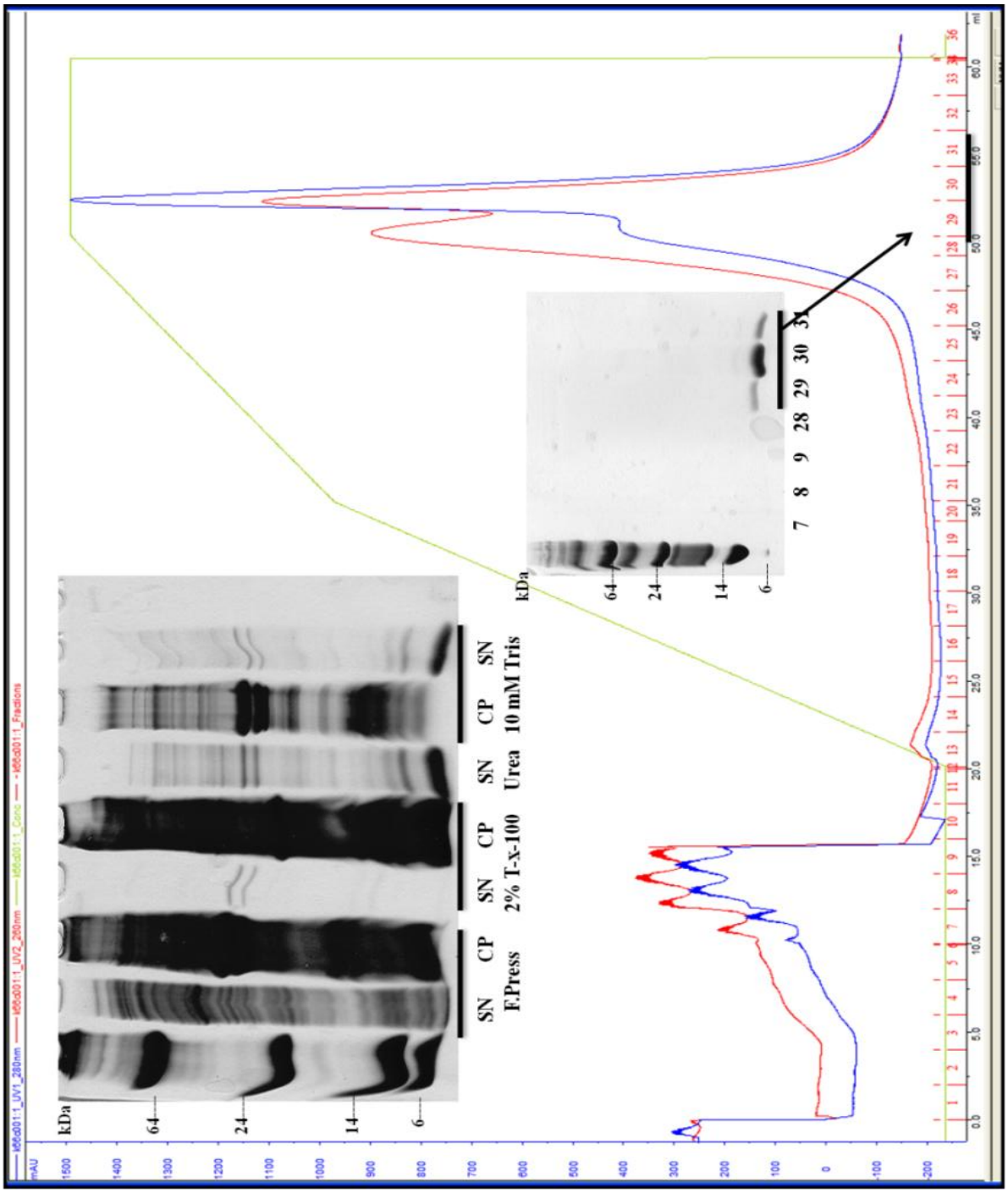


Figure 4.3 Purification of 7E15 charged variants with SDS- PAGE gel of refolding method (top left) followed by SDS- PAGE gel and chromatogram of ion exchange chromatography (bottom right).

Table 4.2 Shows absorbance ratio 280/260 nm, molar absorptivity, molar mass, final concentration, and average final yield in milligrams per liter of expressed 7E15 charged variants.

Protein	Abs 280/260 nm	ε ($M^{-1}cm^{-1}$)	Molar mass (Da)	C_f (μM)	Yeild (mg)
7E15	1.618	11700	11146.5	434	36.3 ± 1.0
7E15E	1.214	11700	11145.5	424	35.4 ± 6.2
7E15N	1.642	11700	11159.5	545	45.6 ± 3.4
7E15Q	1.595	11700	11160.5	531	44.4 ± 5.1

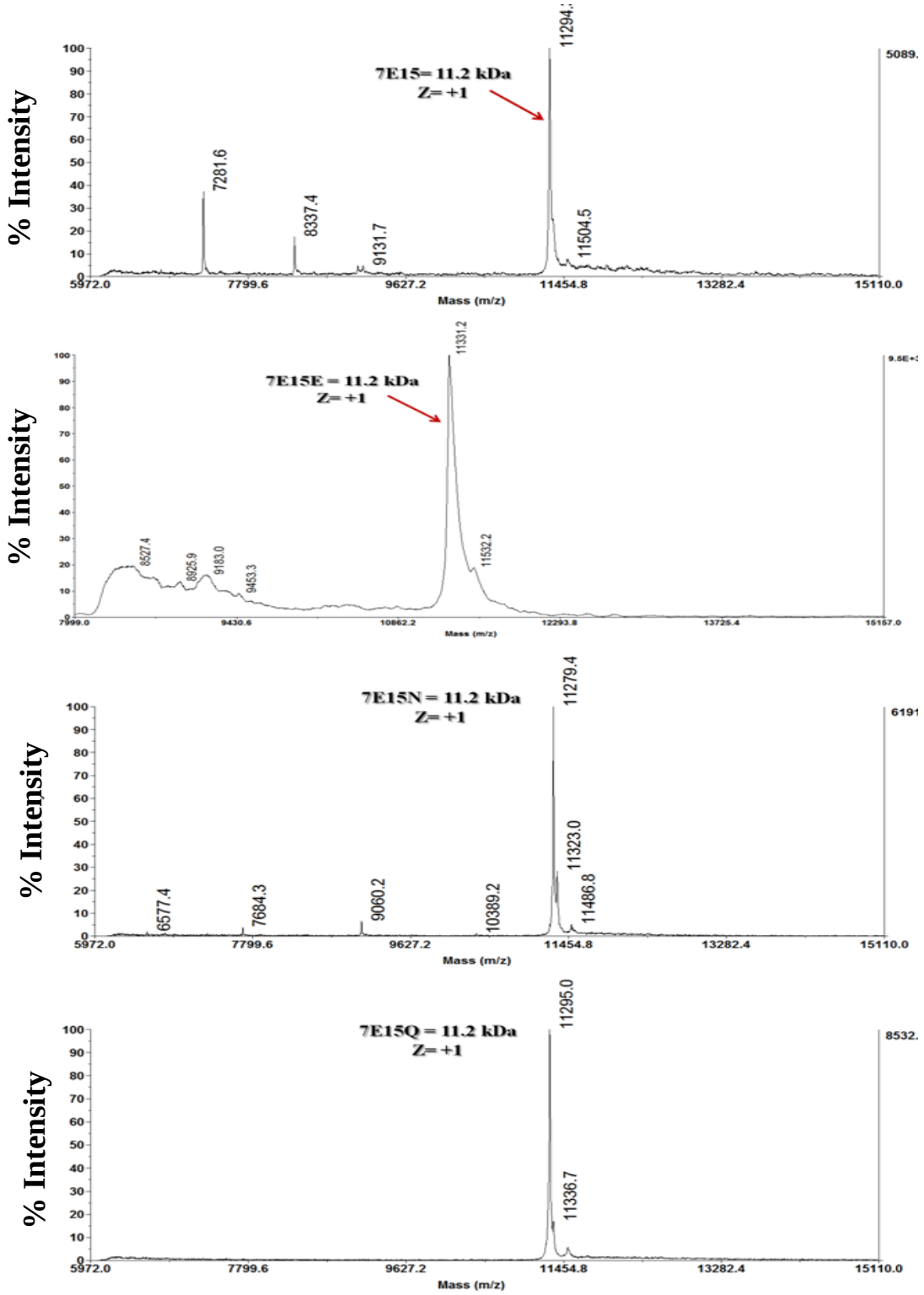


Figure 4.4 Mass spectra of 7E15 charged variants obtained with ABI 4800 MADLI TOF/TOF mass spectrometer instrument.

Studies of secondary and tertiary conformation of charged variants upon binding to metals such as calcium and lanthanides were studied extensively by Dr. Anna Wilkins Maniccia with CD, FRET, and NMR, and details have been published [40]. The simplified study below, modeled by Dr. Maniccia's work, was performed to verify the secondary structure of purified charged variants and to study the effect of adding calcium and gadolinium. The CD spectra below show a negative peak at 219 nm, implying a β-sheet secondary conformation typically exhibited by CD2. Upon adding calcium, there is no significant change in conformation, as shown in previous studies [40]. When gadolinium is added a slight change in conformation observed, due to precipitation (Figure 4.5 A, B, C, D).

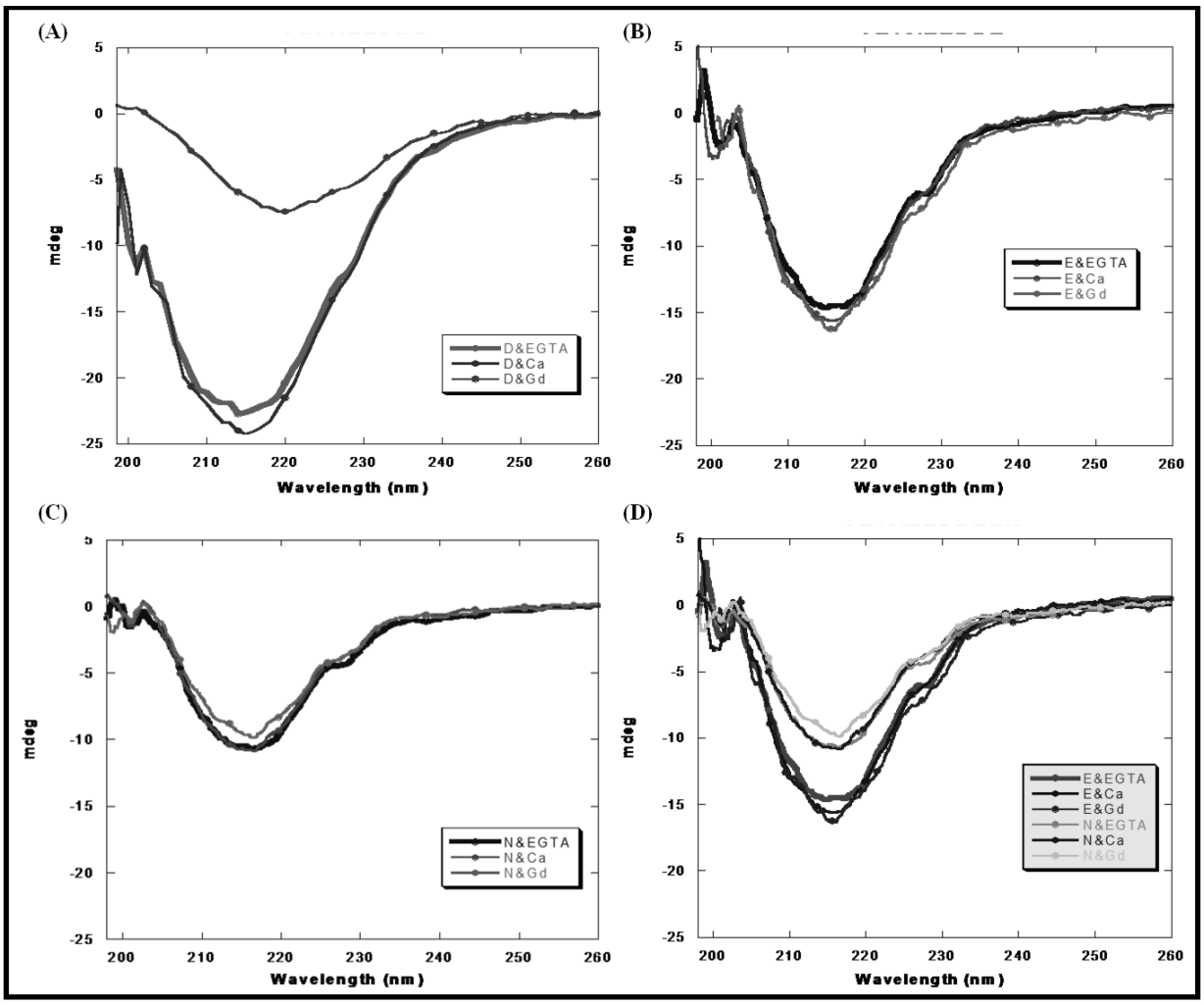


Figure 4.5 Secondary structure studies by CD. (A) 7E15 with EGTA, Ca^{2+}, and Gd^{3+}. (B) 7E15 with EGTA, Ca^{2+}, and Gd^{3+}. (C) 7E15N with EGTA, Ca^{2+}, and Gd^{3+}. (D) 7E15N and 7E15E each with EGTA, Ca^{2+}, and Gd^{3+}.

4.3.3 Determination of K_d of 7E15 Charged Variants to Tb^{3+} by FRET Titration

Metal binding studies were performed to determine the binding affinity of the charged variants to terbium. Terbium and calcium have similar atomic radii (~180 pm) and behave similarly, except calcium is spectroscopically inactive and binding cannot be studied with this method. FRET method utilizes the fact that tryptophan can transfer the energy from excitation to terbium, because the excitation of tryptophan alone does not produce a recordable signal. At 282 nm, tryptophan is excited and emits the energy at a wavelength of 340 nm, which is the wavelength range that terbium absorbs energy. Between 500 and 600 nm, terbium emits the energy absorbed from tryptophan at 545 nm and 595 nm (two of four possible wavelengths).

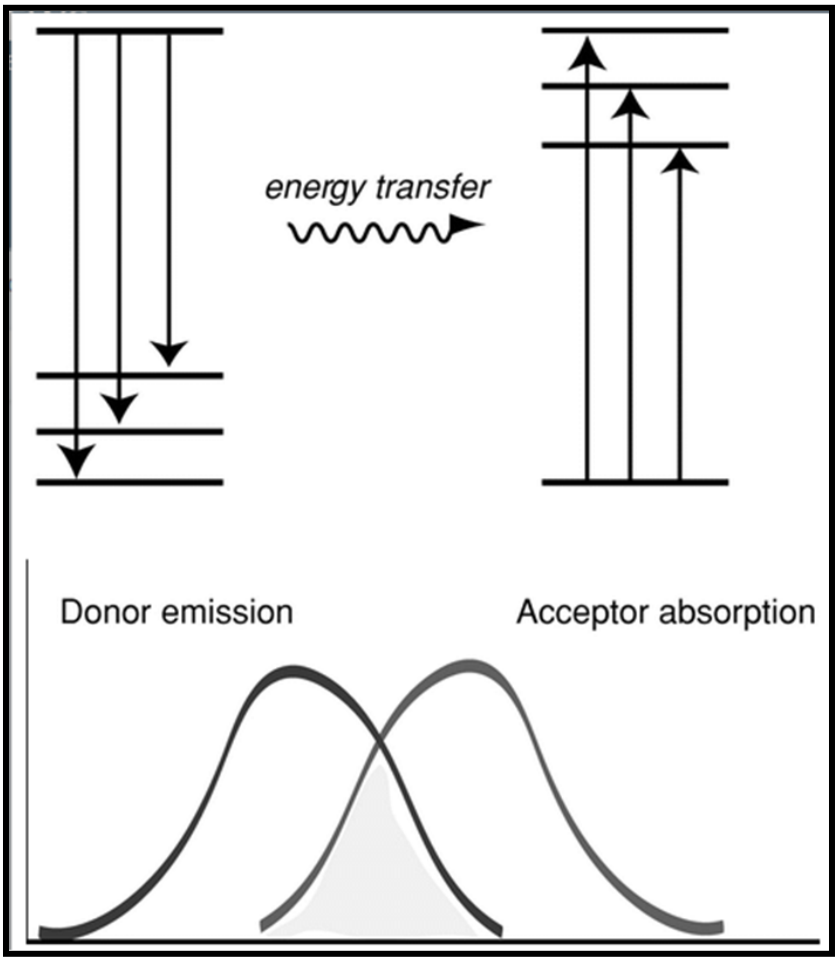


Figure 4. 6 Cartoon illustrating donor emission (tryptophan) and acceptor absorption (terbium) spectral overlap. Taken from [41].

The fluorescence of the 7E15-terbium complex increases with increasing terbium concentration until saturation (~20 µM), when no more protein is available to bind terbium. The maximum at 545 nm is fitted. The curves were fitted using the equation described in the methods portion (Equation 4.20) and 1:1 binding is clearly seen (Figures 4.7. and 4.8.). Table 4.3 summarizes the approximate K_d values in µM, 4.45$\pm$ 0.65 for 7E15, and 8.97 $\pm$ 0.87 for 7E15N; previous results show a binding affinity of CD2.7E15 to terbium of 0.4 $\pm$ 0.2 µM [40].

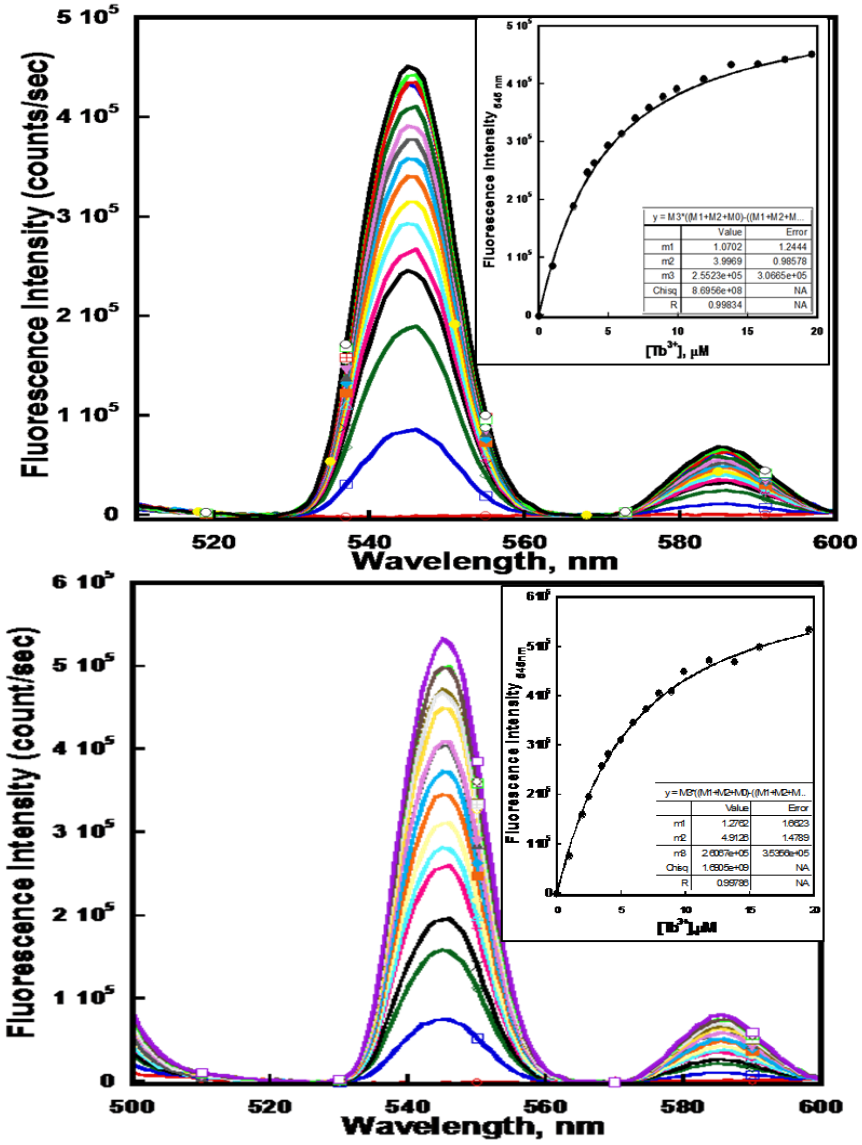


Figure 4.7 Fluorescence intensity of Tb^{3+} as a function of wavelength. Titration with 5 µM CD2.7E15 in 10 mM chelex Tris, pH 7.4, excitation at 282 nm, emission 500-600 nm, maximum at 545 nm (fitted to give an approximate K_d in µM).

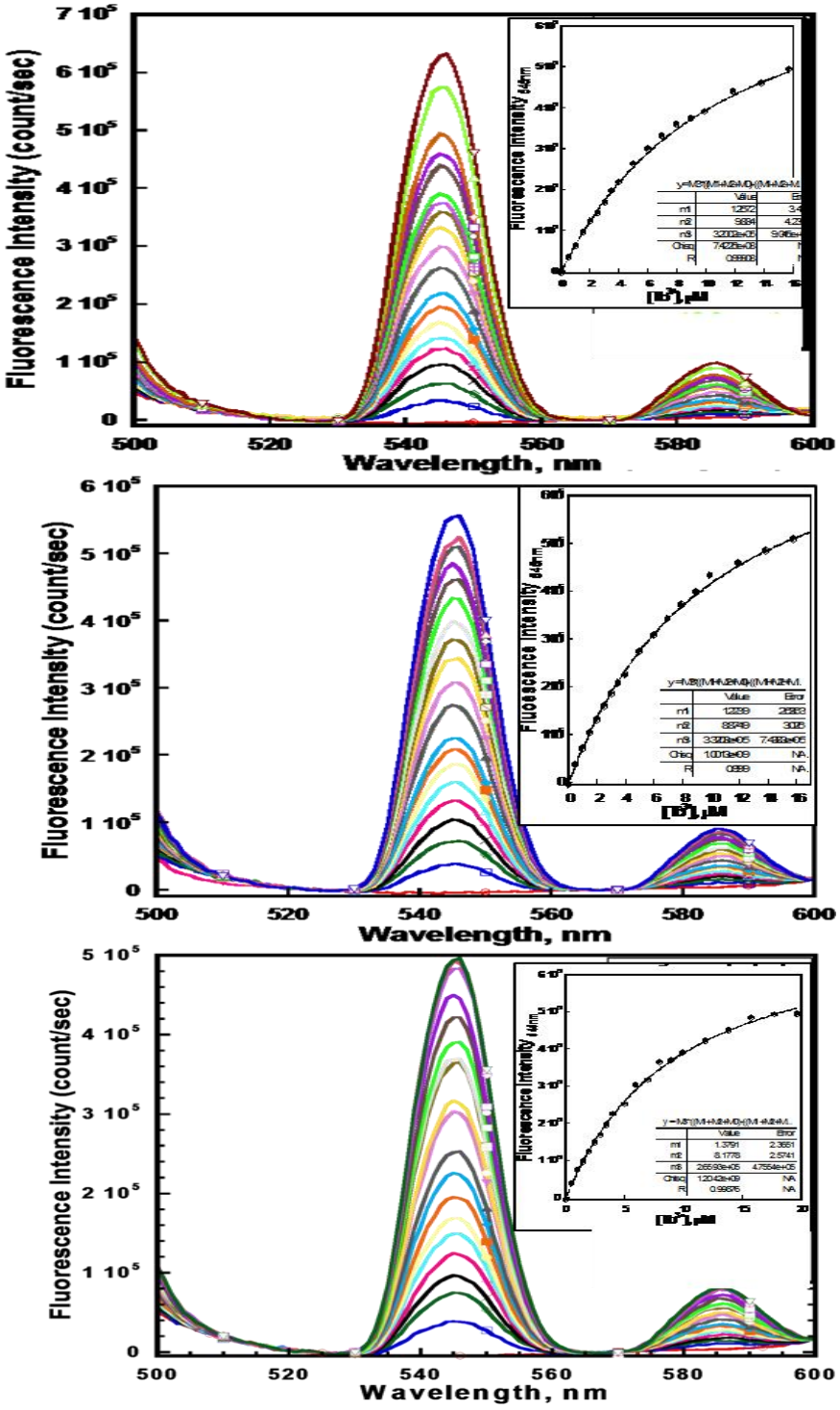


Figure 4.8 Fluorescence intensity of Tb^{3+} as a function of wavelength. Titration with 5 μM CD2.7E15N in 20 mM PIPES, 10 mM KCl, pH 6.8, excitation at 282 nm, emission 500-600 nm, maximum at 545 nm (fitted to give an approximate K_d in μM).

Table 4.3 Summary of the approximate K_d values for 7E15 in 10 mM Tris-HCl, pH 7.4 and 7E15N in 20 mM PIPES, 10 mM KCl, pH 6.8. Final protein concentration was 5 µM, excitation wavelength was 282 nm, and emission wavelength was 500-600 nm.

7E15 Variant	Approximate K_d (µM)
7E15	4.45 ± 0.65
7E15N	8.97 ± 0.85

4.3.4 Determination of Affinity of 7E15 Charged Variants to Gd^{3+} by Competition with Molecular Probes Dye Fluo-5N

The binding affinity of 7E15 charged variants to gadolinium was measured using fluorescent dye Fluo-5N pentapotassium salt, which has a published K_d to gadolinium of $3.8 \pm 0.2 \times 10^{-12}$ and a high sensitivity. The dye itself has fluorescence intensity due to aromatic components within its structure (Figure 4.9). A saturation test was performed during which the dye was saturated with gadolinium. The maximum intensity peak corresponded to Fluo-5N-Gd^{3+} while minimum intensity peak corresponded to Fluo-5N (Figure 4.10). As well, a titration of Fluo-5N-Gd^{3+} with EDTA was performed as a control, as EDTA has a K_d to gadolinium of approximately 10^{-18} [39] (Figure 4.11).

Following this, titration of protein gradually into 1:1 ratio of Fluo-5N-Gd^{3+} showed a decrease in fluorescence intensity as saturation is approached. This is because the gadolinium is competed out by the protein and 7E15-Gd^{3+} does not have a fluorescence signal. The first trial in duplicate with 7E15 took place with 10 µM Fluo-5N and 10 µM Gd^{3+} in conditions of 10 mM Tris-HCl pH 7.4 (Figure 4.12). An average plot of all trials shows that the K_{app} is 1.74; calculated K_d value is 6.62×10^{-13} (Figure 4.13). The other charged variants were tried using these conditions, but the plots were not fitted well because of incomplete saturation due to precipitation (but no hydroxide formation).

The next set of experiments was performed with a lower concentration of dye and metal, 5 µM Fluo-5N and 5 µM Gd^{3+}, in 20 mM PIPES, 10 mM KCl, pH 6.8. Fluorescence intensity versus wavelength was graphed. The fitted curve at 519_{max} nm displaying K_{app} (fitted with Equation 4.24) is shown on the top right of each graph (Figure 4.14, 4.15, 4.16, 4.17). The average fitted plot for each charged variant with K_{app} is shown (Figure 4.18). The K_d to

gadolinium for each charged variant was calculated using Equation 4.23 and the results indicate that CD2.7E15 has the lowest affinity to gadolinium while CD2.7E15Q has the highest affinity to gadolinium (Table 4.4).

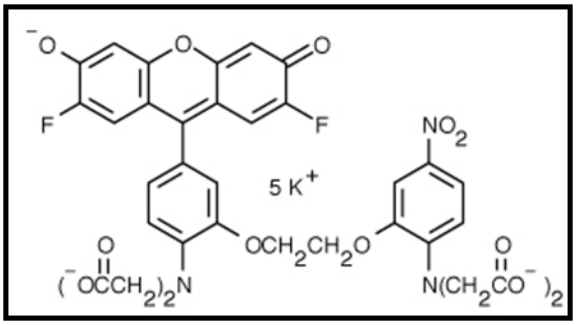


Figure 4.9 Chemical structure of Fluo-5N, pentapotassium salt Molecular Probes Invitrogen dye. Taken from [42].

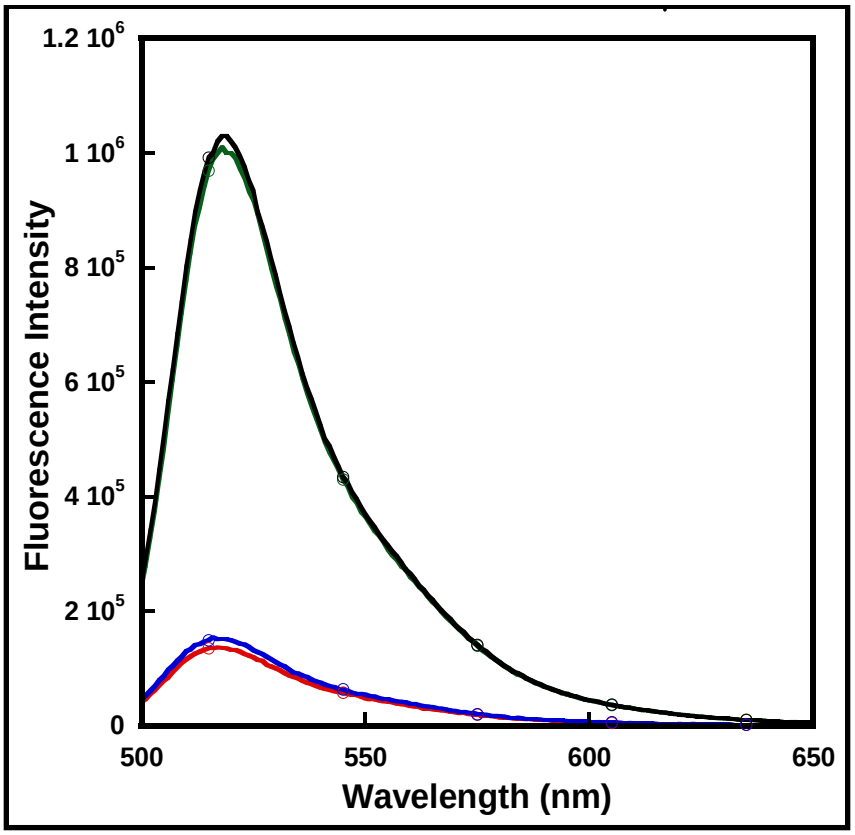


Figure 4.10 Saturation test. Maximum fluorescence intensity peak corresponds to Fluo-5N- Gd^{3+} complex, while minimum fluorescence intensity peak corresponds to Fluo-5N.

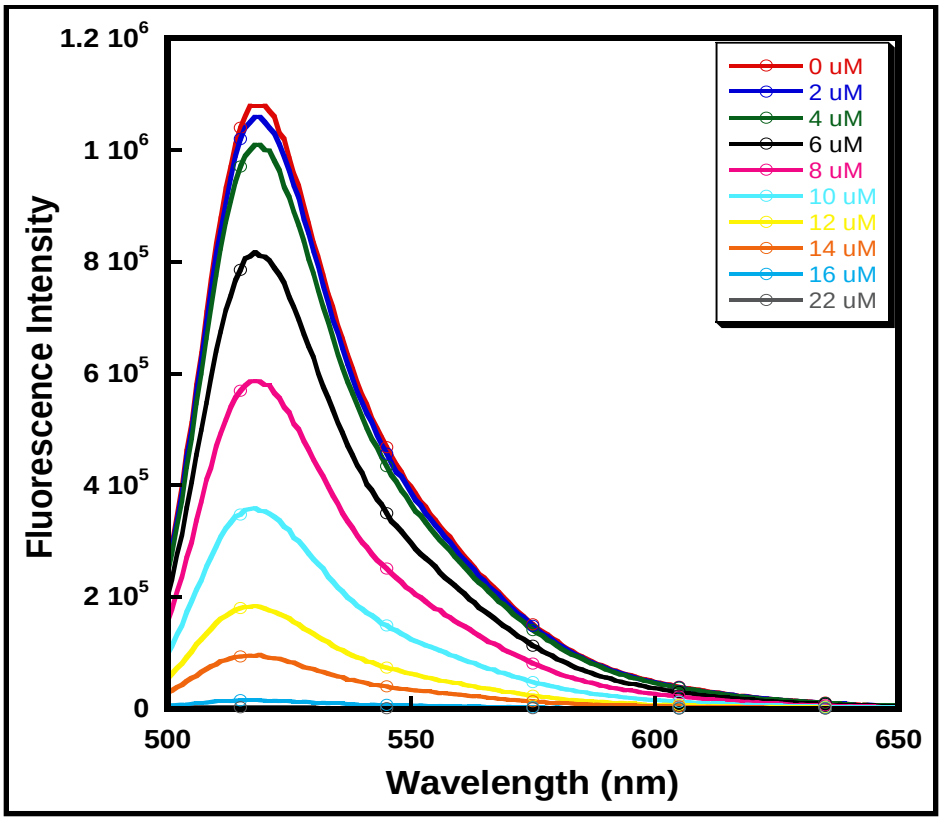


Figure 4.11 Titration of Fluo-5N-gadolinium with EDTA.

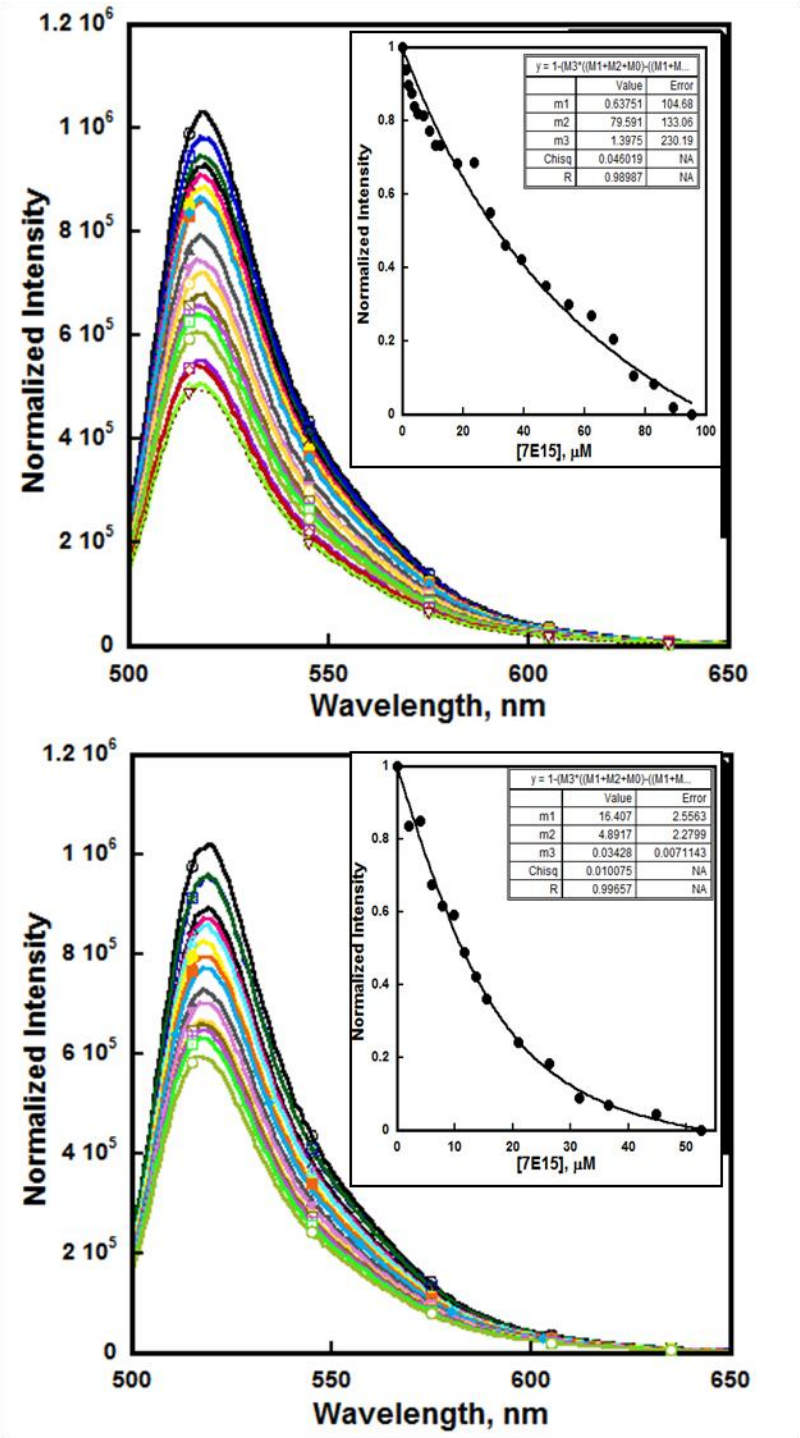


Figure 4.12 Competition assay with 10 µM Fluo-5N and 10 µM gadolinium in 10 mM Tris-HCl, pH 7.4 with CD2.7E15 (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials).

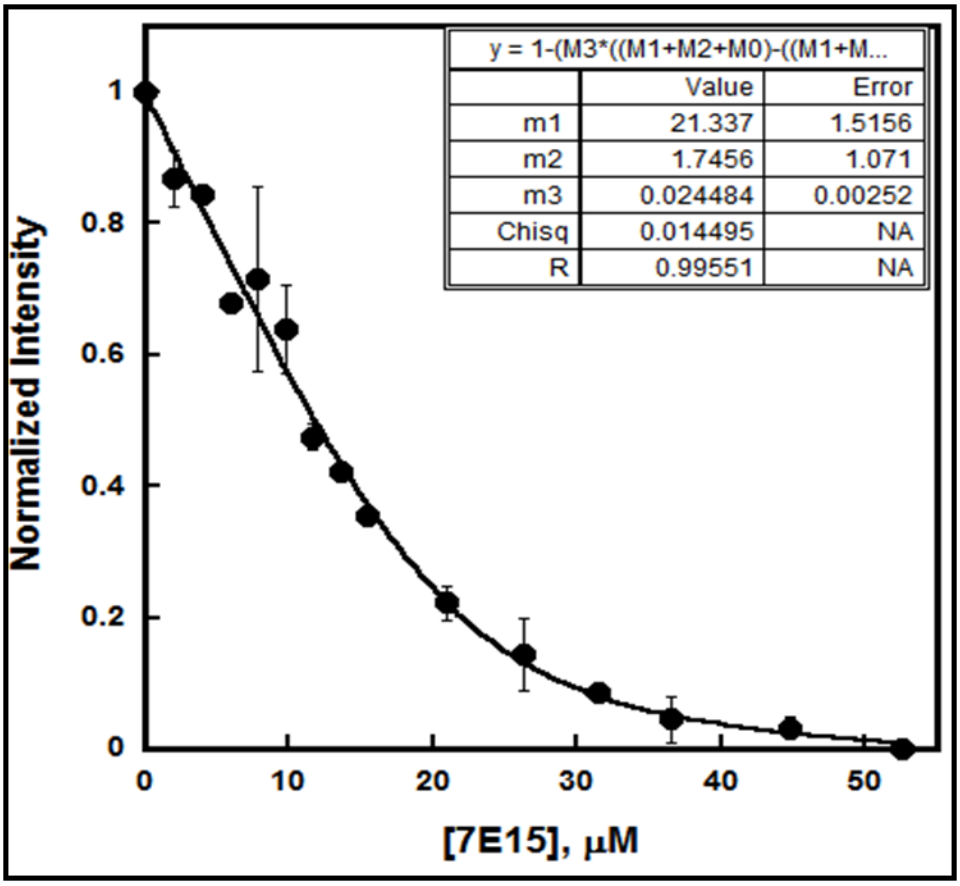


Figure 4.13 Average plot of 7E15-Gd^{3+} following assay of 10 µM Fluo-5N to 10 µM gadolinium in 10 mM Tris-HCl, pH 7.4 indicated K_{app} of 1.74 and a calculated K_d of 6.62 x 10^{-13}.

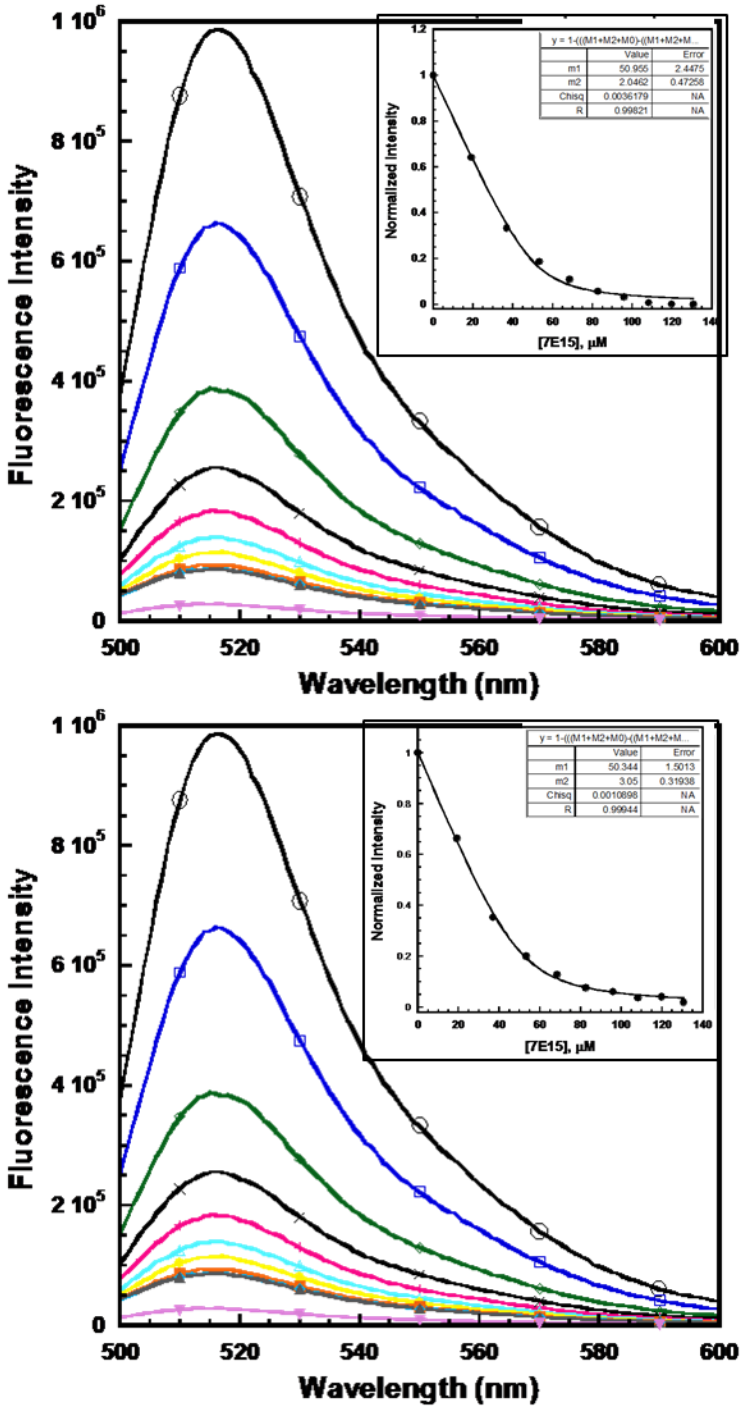


Figure 4.14 Competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15 (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials).

.

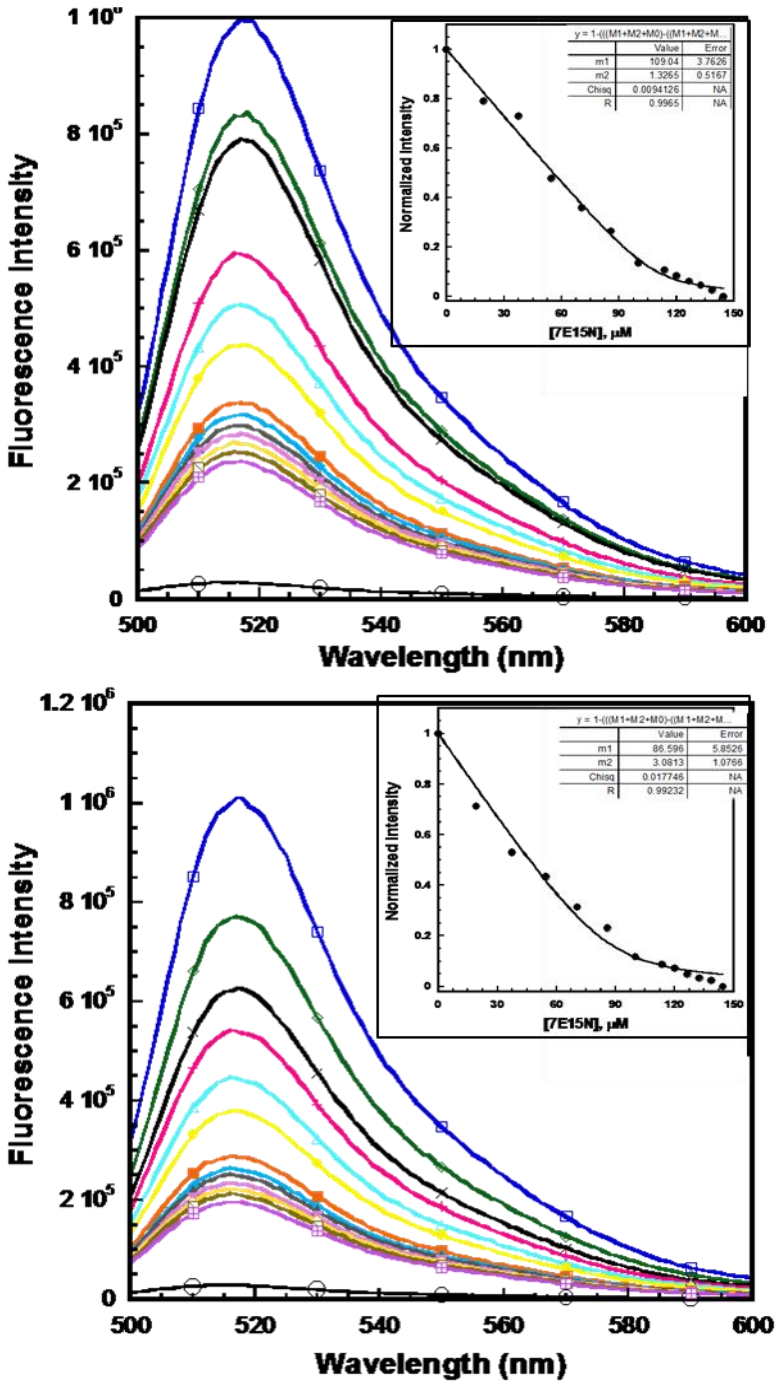


Figure 4.15 Competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15N (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials).

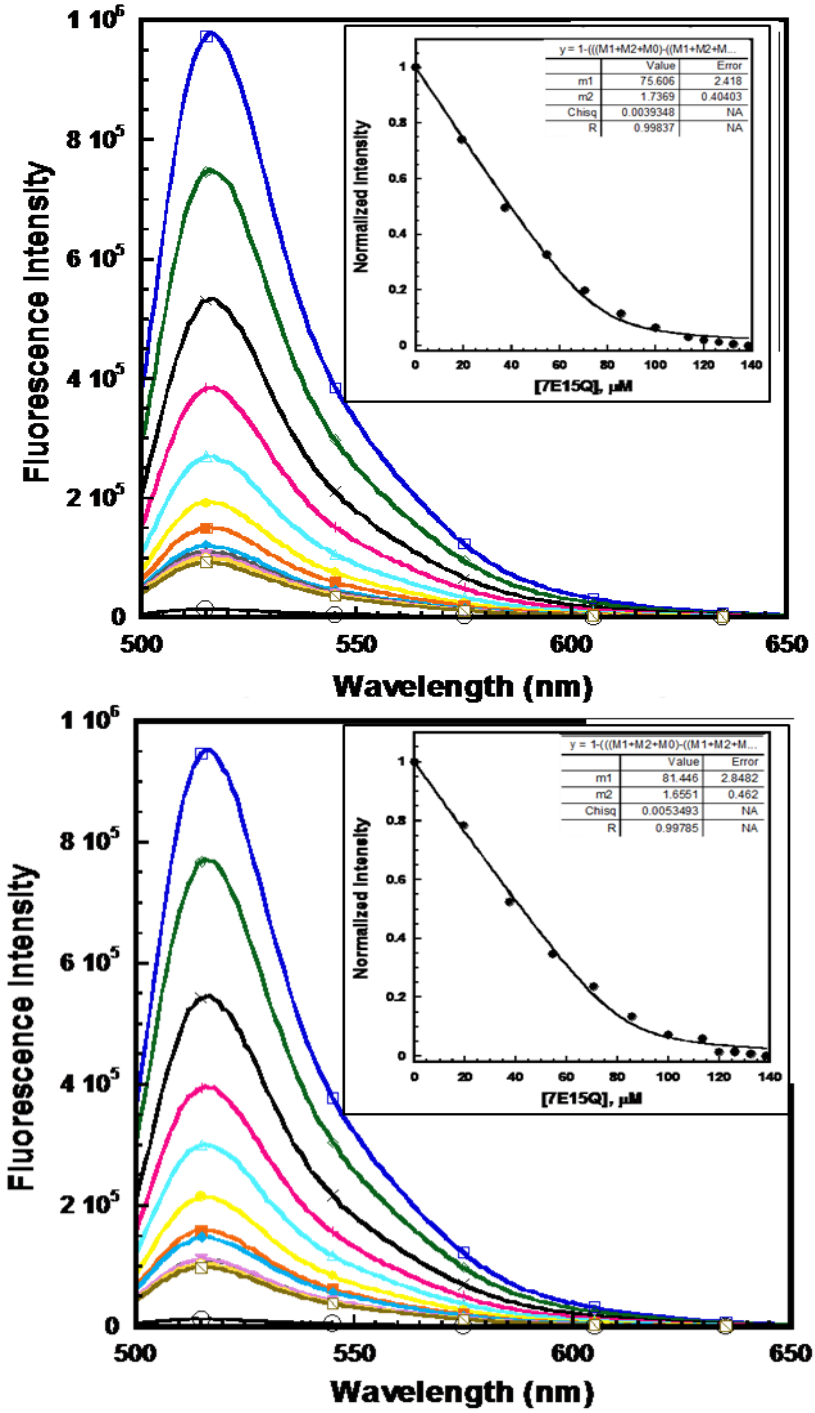


Figure 4.16 Competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15Q (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials).

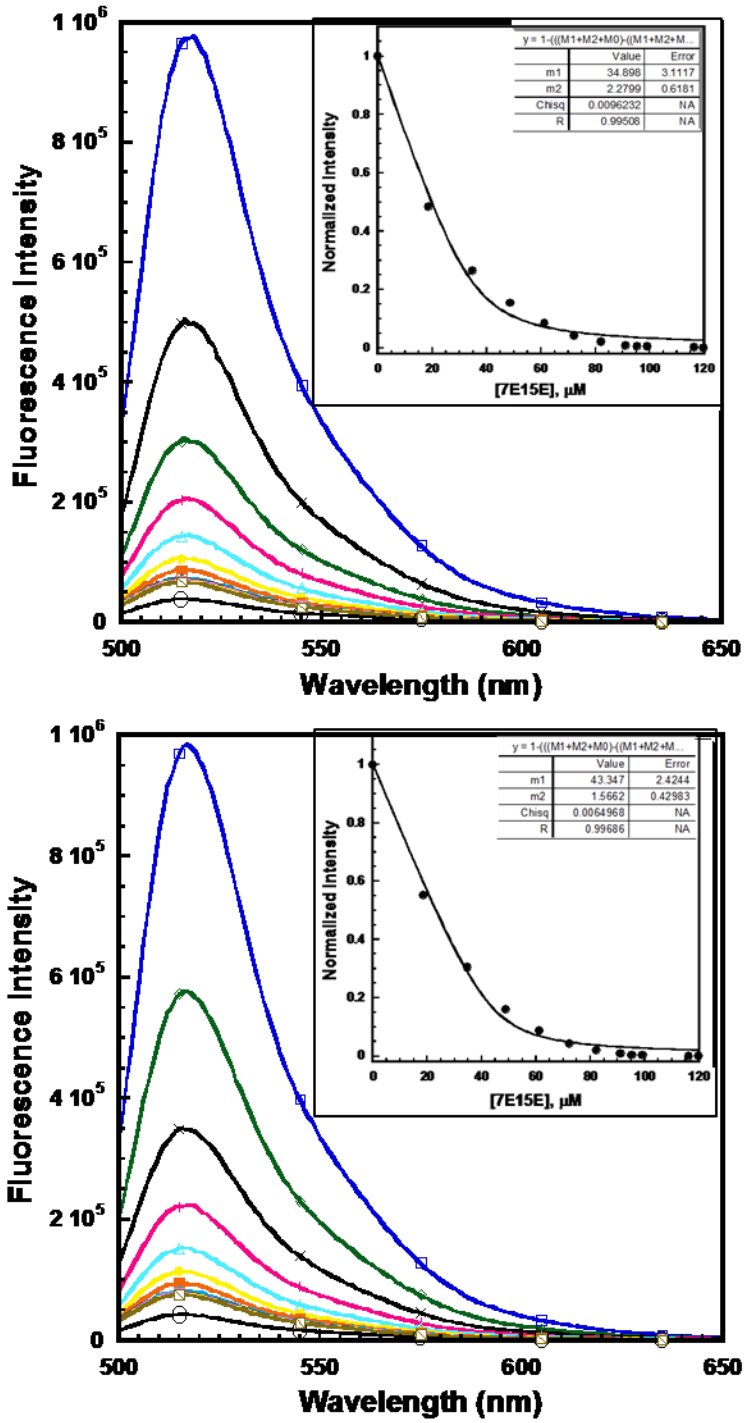


Figure 4.17 Competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15E (added in 20 µM increments until saturation) to determine K_d; excitation at 488nm, emission range 500-650 nm (two trials).

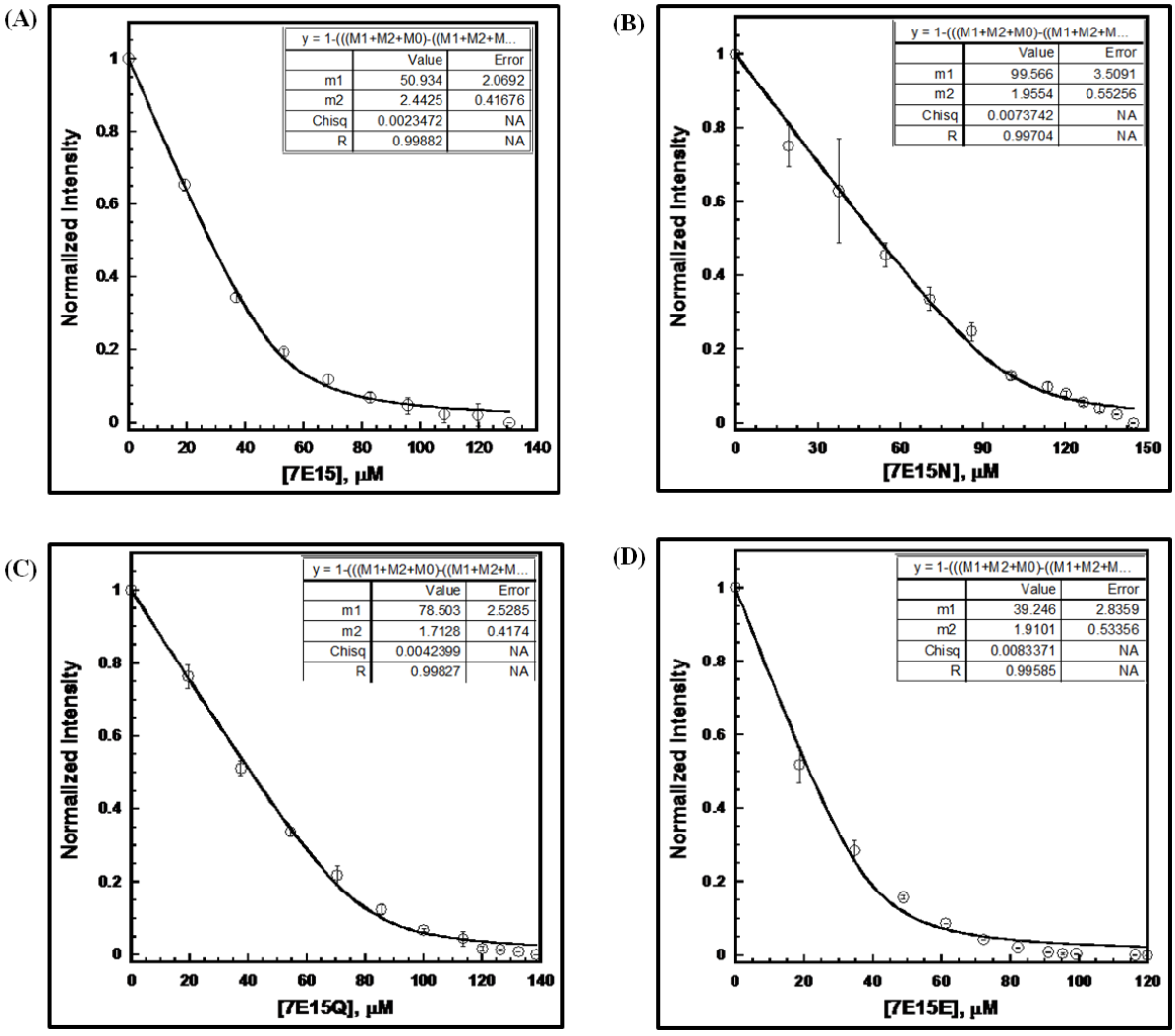


Figure 4.18 Average plots of competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15 and charged variants to determine K_d. (A) CD2.7E15. (B) CD2.7E15N. (C) CD2.7E15Q. (D) CD2.7E15E.

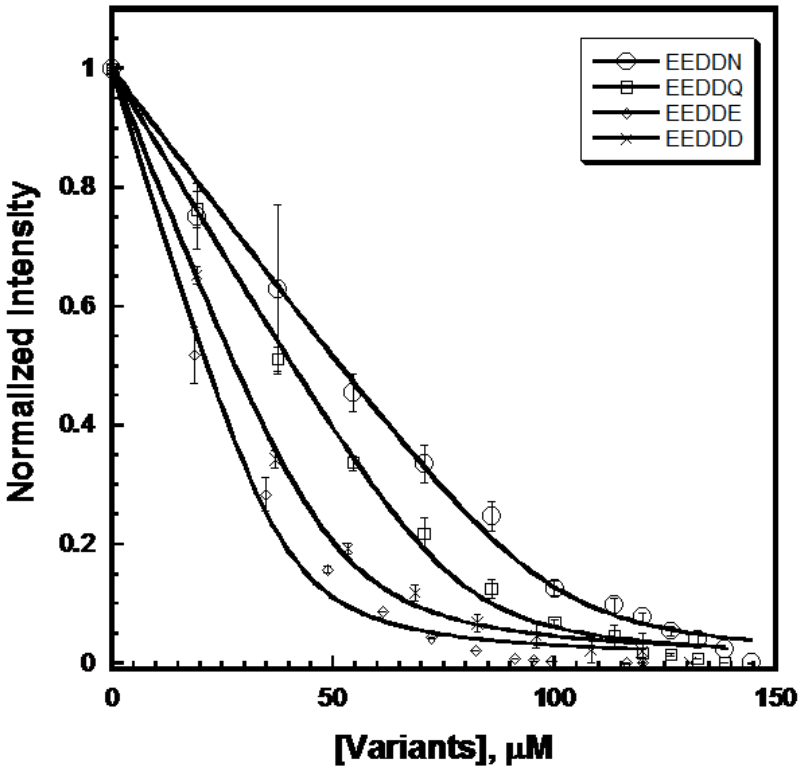


Figure 4.19 Overlaid average plot of competition assay with 5 µM Fluo-5N and 5 µM gadolinium in 20 mM PIPES, 10 mM KCl, pH 6.8 with CD2.7E15 and charged variants to determine K_d.

4.3.5 Determination of Affinity of 7E15 Charged Variants to Zn^{2+} by Competition with Molecular Probes Dye FluoZin-1

The Zn^{2+} binding affinity to CD2.7E15 was determined by competition assay with 10 µM Fluoin-1 and 10 µM Zn^{2+}, in 10 mM Tris, pH 7.4. Fluorescence intensity versus wavelength was graphed and the point at maximum wavelength 545 nm was used to fit the curve and obtain the K_{app} (fitted with Equation 4.24), which is shown on the top right of each graph (Figure 4.20, 4.21, 4.22, 4.23). The average fitted plot of each charged variant with K_{app} is shown (Figure 4.24). The fitting is not optimal because the concentration of the dye and zinc are high and incomplete saturation occurs even with high concentrations of protein. The binding affinity of the CD2.7E15 charged variants to zinc was calculated using Equation 4.23. Results indicate that CD2.7E15N has the lowest affinity to zinc while CD2.7E15Q has the highest affinity to zinc (Table 4.4).

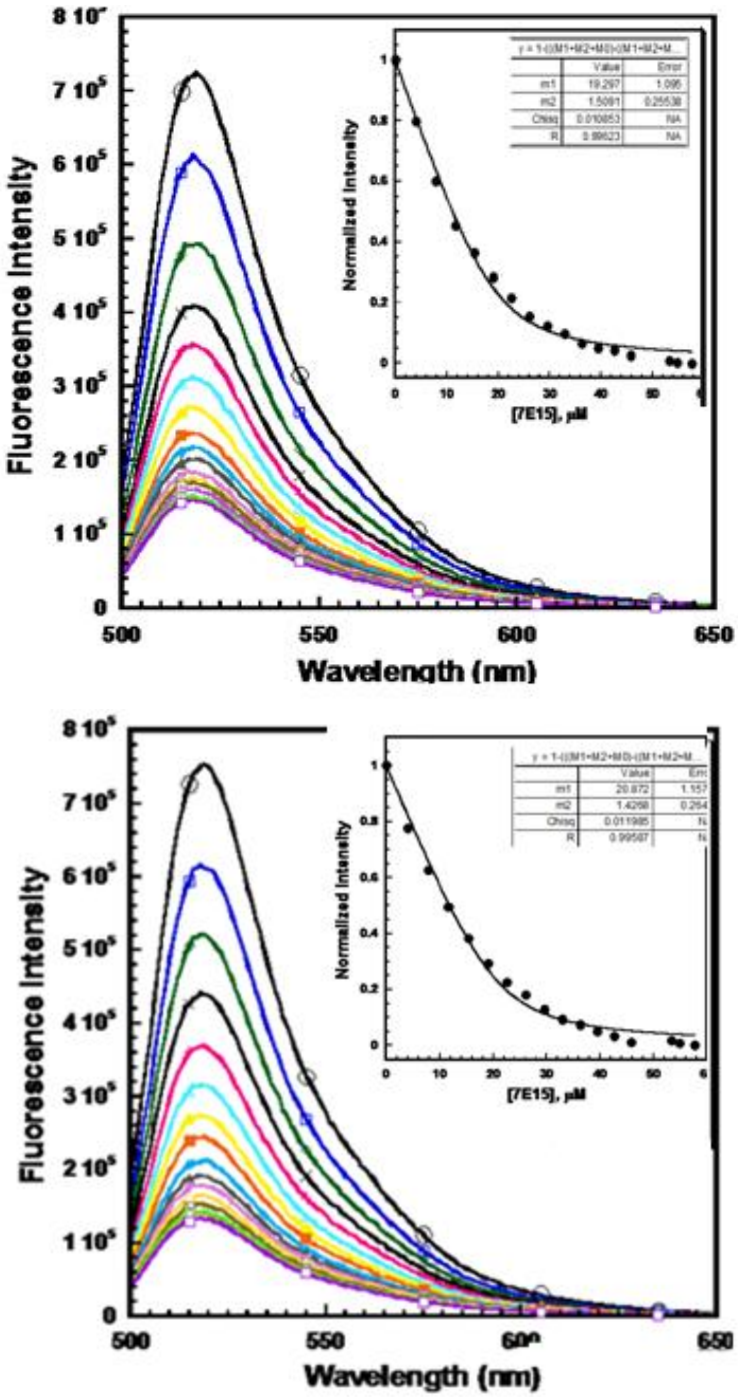


Figure 4. 20 Competition assay with 10 µM FluoZin-1 and 10 µM zinc in 10 mM Tris, pH 7.4 with CD2.7E15 (added in 8 µM increments until saturation) to determine K_d; excitation at 495 nm, emission range 500-650 nm (two trials).

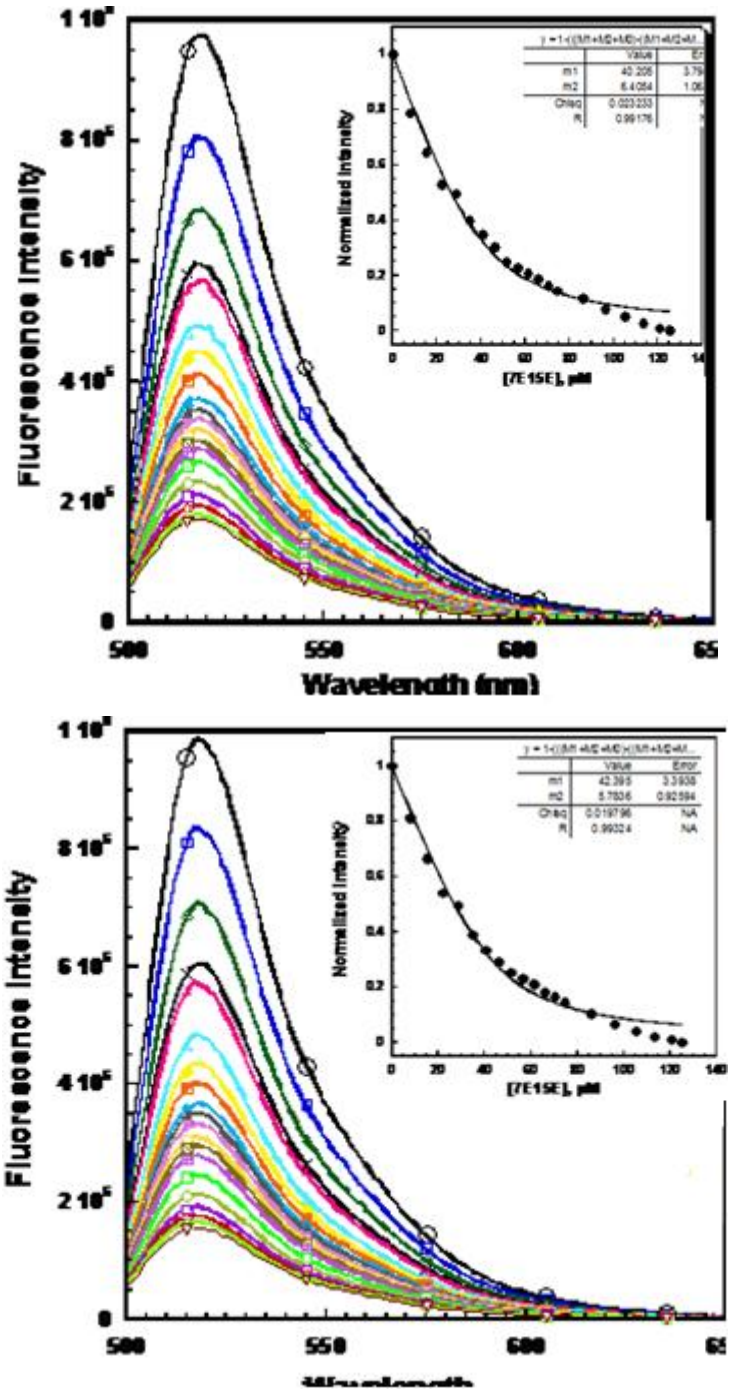


Figure 4. 21 Competition assay with 10 µM FluoZin-1 and 10 µM zinc in 10 mM Tris, pH 7.4 with CD2.7E15E (added in 8 µM increments until saturation) to determine K_d; excitation at 495 nm, emission range 500-650 nm (two trials).

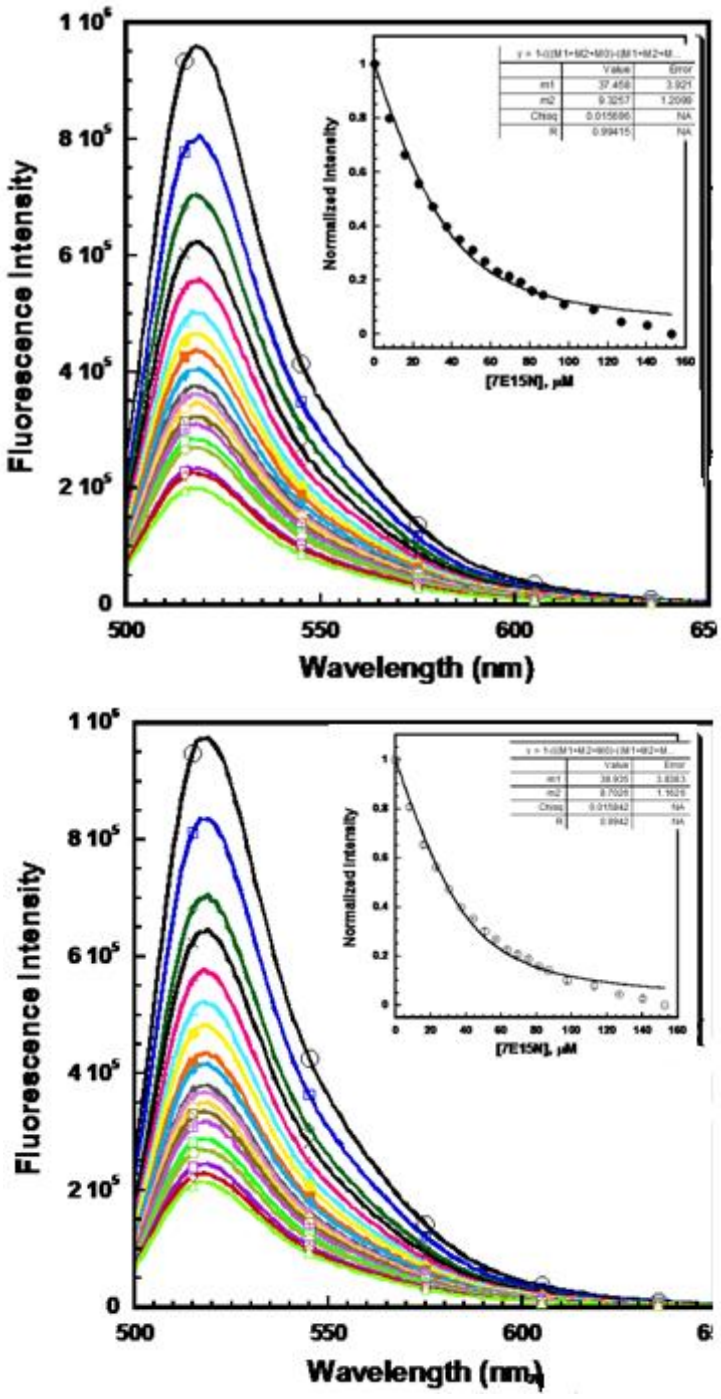


Figure 4. 22 Competition assay with 10 μM FluoZin-1 and 10 μM zinc in 10 mM Tris, pH 7.4 with CD2.7E15N (added in 8 μM increments until saturation) to determine K_d; excitation at 495 nm, emission range 500-650 nm (two trials).

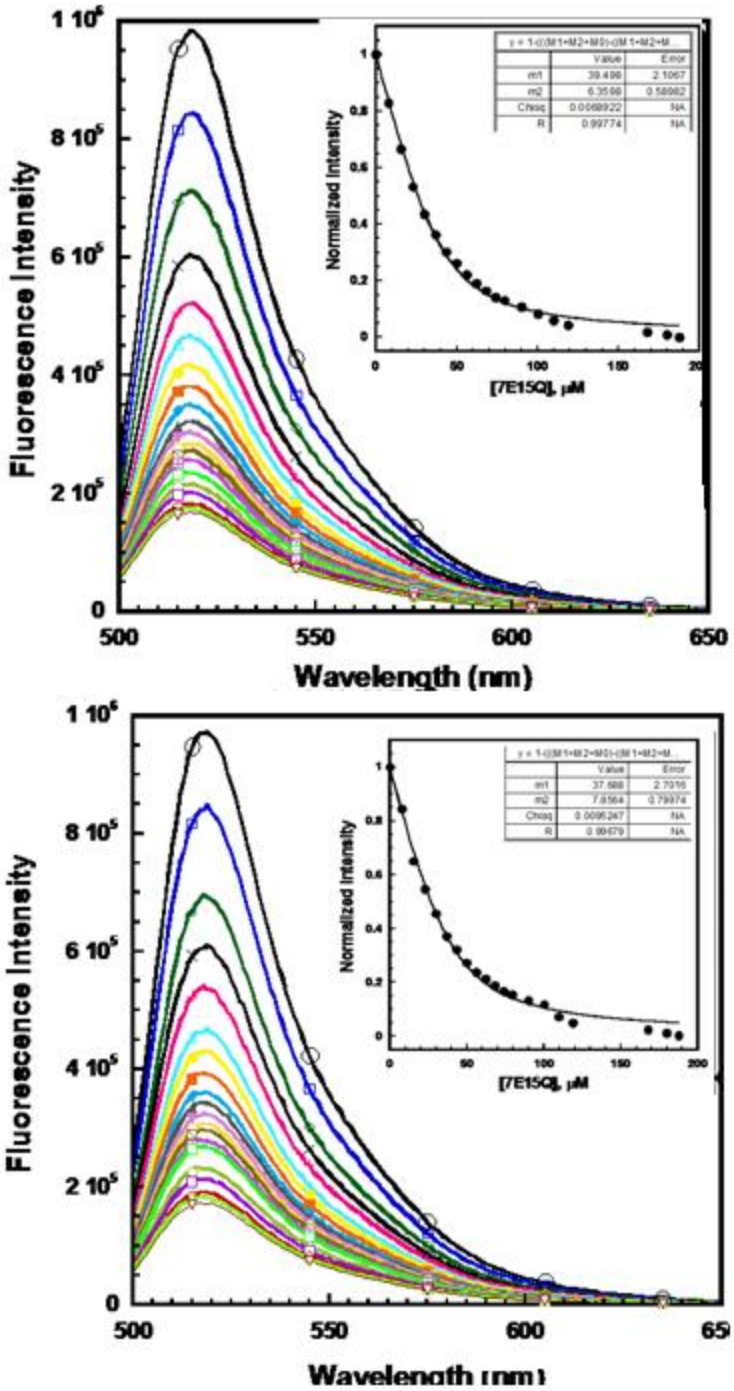


Figure 4. 23 Competition assay with 10 µM FluoZin-1 and 10 µM zinc in 10 mM Tris, pH 7.4 with CD2.7E15Q (added in 8 µM increments until saturation) to determine K_d; excitation at 495 nm, emission range 500-650 nm (two trials).

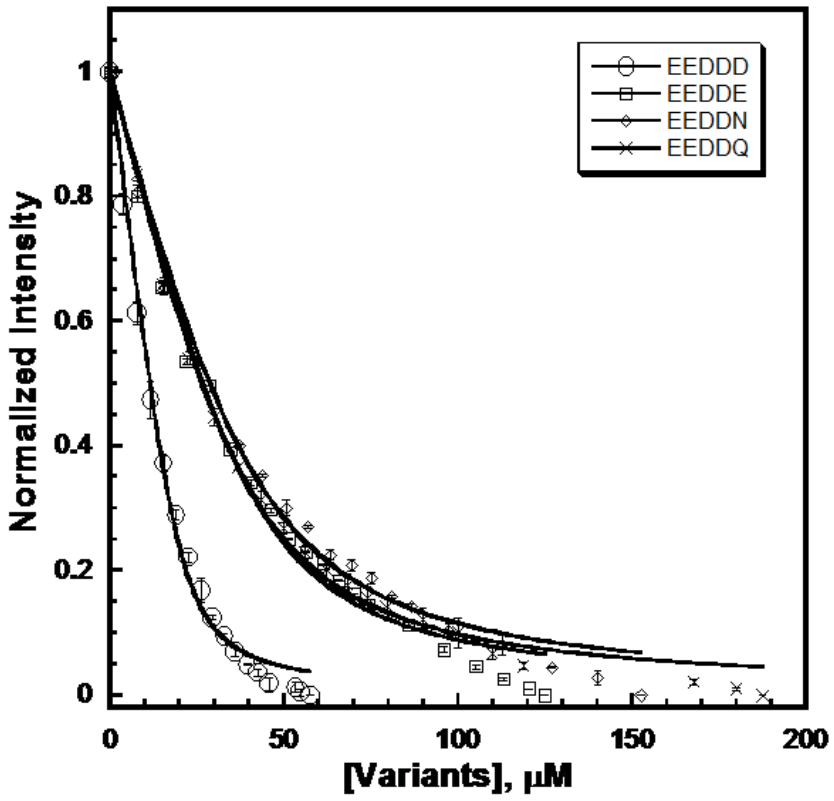
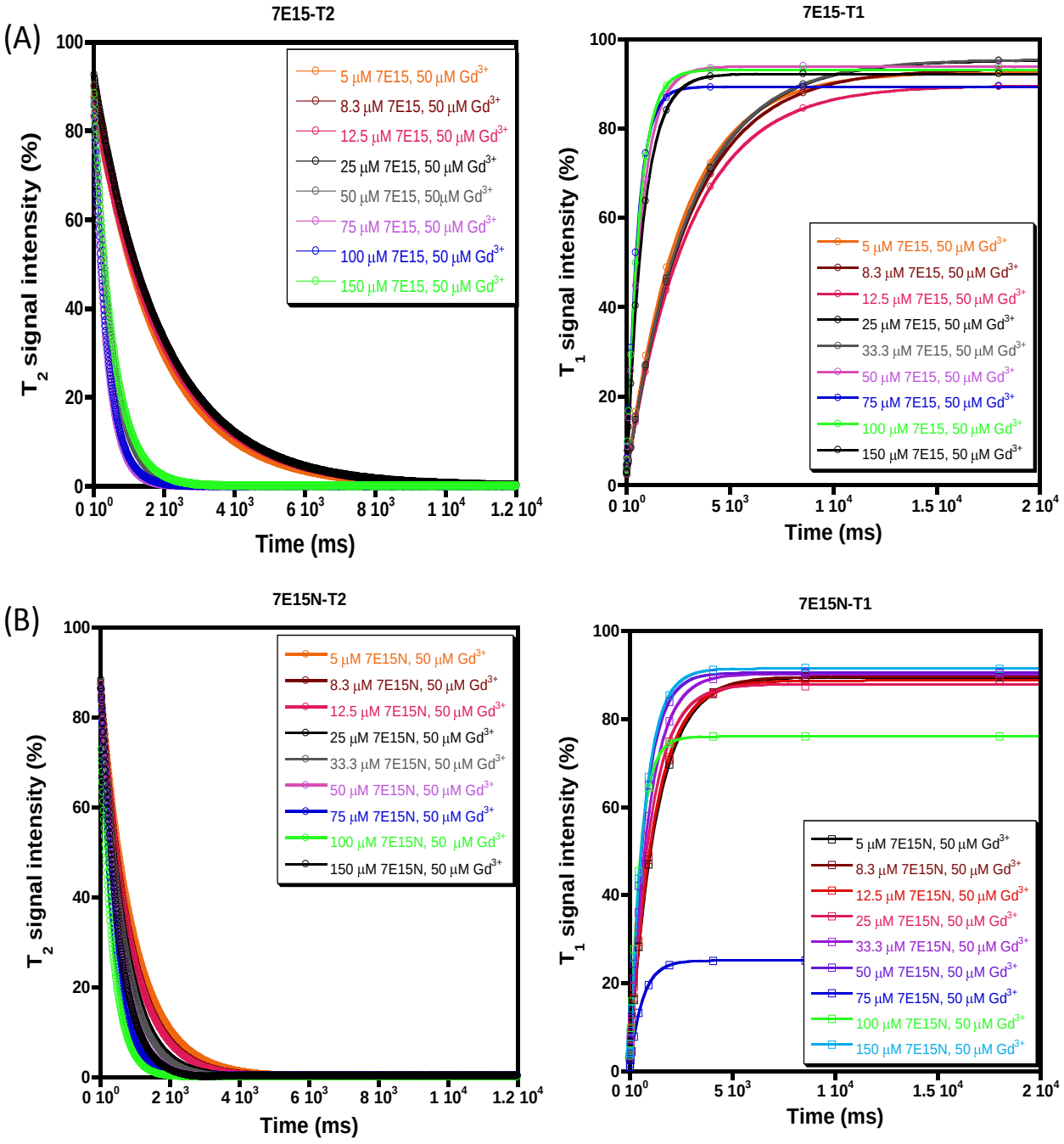


Figure 4.24 Overlaid average plot of competition assay with 10 µM FluoZin-1 and 10 µM zinc in 10 mM Tris-HCl, pH 7.4 with CD2.7E15 and charged variants to determine K_d.

4.3.6 Relaxivity of CD2.7E15 Charged Variants

The relaxivity of CD2.7E15 and charged variants was measured by changing the protein concentration against a fixed concentration of gadolinium. This method highlights the 1:1 binding mode in CD2.7E15 and the charged variants. Figure 4.25 A-D shows the change in relaxation time with increasing protein to gadolinium ratio. Relaxation time decreases as the ratio of protein to gadolinium increases, until the point of protein saturation with gadolinium. In Figure 4.25 B, the T_1 signal of 75 µM 7E15N is significantly lower than the others due to instrumental error.


(A)
7E15-T2
T2 signal intensity (%)
Time (ms)
5 μM 7E15, 50 μM Gd3+
8.3 μM 7E15, 50 μM Gd3+
12.5 μM 7E15, 50 μM Gd3+
25 μM 7E15, 50 μM Gd3+
50 μM 7E15, 50μM Gd3+
75 μM 7E15, 50 μM Gd3+
100 μM 7E15, 50 μM Gd3+
150 μM 7E15, 50 μM Gd3+
7E15-T1
T1 signal intensity (%)
Time (ms)
5 μM 7E15, 50 μM Gd3+
8.3 μM 7E15, 50 μM Gd3+
12.5 μM 7E15, 50 μM Gd3+
25 μM 7E15, 50 μM Gd3+
33.3 μM 7E15, 50 μM Gd3+
50 μM 7E15, 50 μM Gd3+
75 μM 7E15, 50 μM Gd3+
100 μM 7E15, 50 μM Gd3+
150 μM 7E15, 50 μM Gd3+
(B)
7E15N-T2
T2 signal intensity (%)
Time (ms)
5 μM 7E15N, 50 μM Gd3+
8.3 μM 7E15N, 50 μM Gd3+
12.5 μM 7E15N, 50 μM Gd3+
25 μM 7E15N, 50 μM Gd3+
33.3 μM 7E15N, 50 μM Gd3+
50 μM 7E15N, 50 μM Gd3+
75 μM 7E15N, 50 μM Gd3+
100 μM 7E15N, 50 μM Gd3+
150 μM 7E15N, 50 μM Gd3+
7E15N-T1
T1 signal intensity (%)
Time (ms)
5 μM 7E15N, 50 μM Gd3+
8.3 μM 7E15N, 50 μM Gd3+
12.5 μM 7E15N, 50 μM Gd3+
25 μM 7E15N, 50 μM Gd3+
33.3 μM 7E15N, 50 μM Gd3+
50 μM 7E15N, 50 μM Gd3+
75 μM 7E15N, 50 μM Gd3+
100 μM 7E15N, 50 μM Gd3+
150 μM 7E15N, 50 μM Gd3+

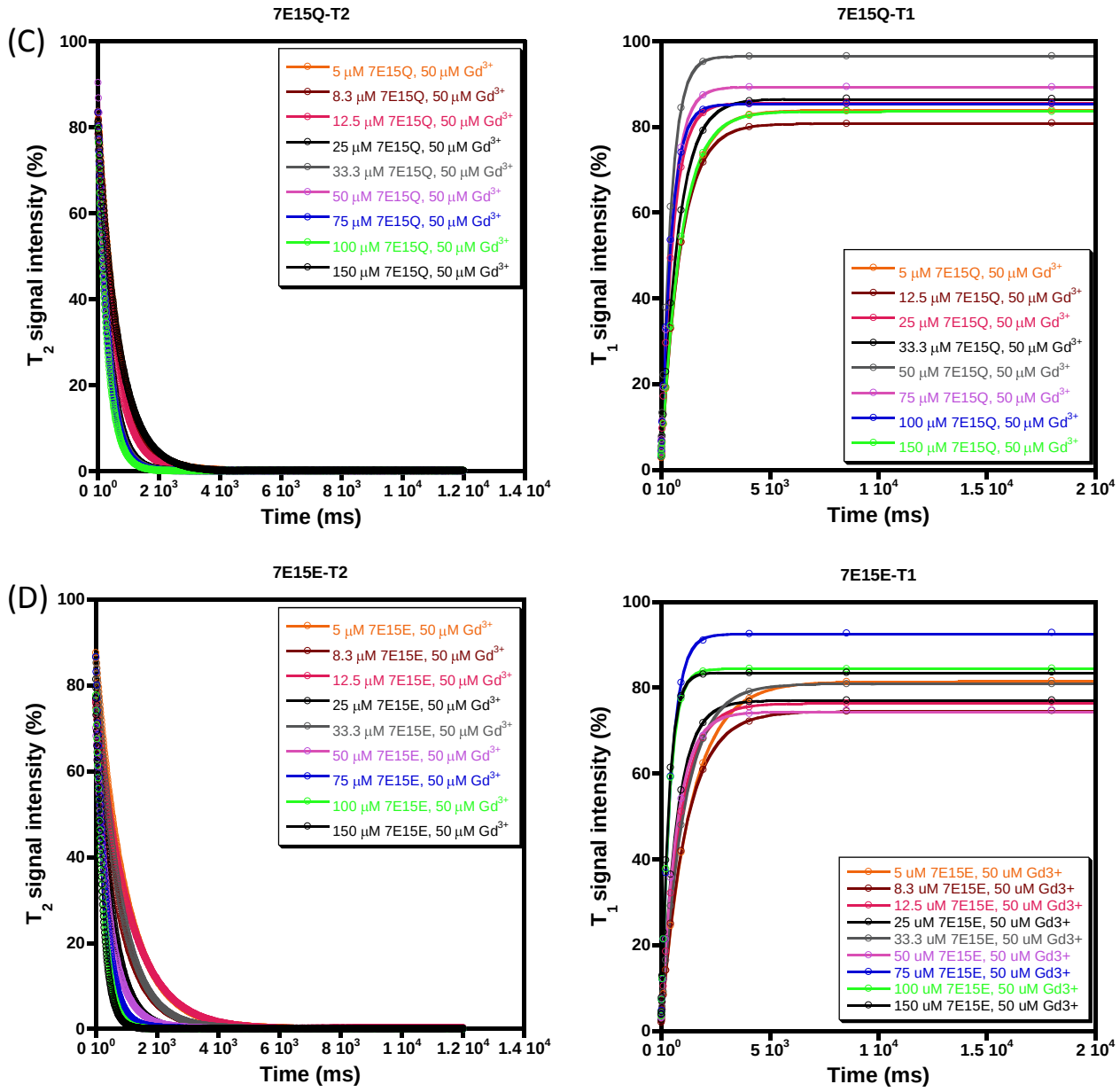


Figure 4. 25 Graphs of T_1 and T_2 for CD2.7E15 and the charged variants show change in relaxation time with increasing protein to gadolinium ratio in 10 mM HEPES, pH 7, 25 °C, 1.4 T. (A) T1 (right) and T2 (left) graphs of CD2.7E15. (B) T1 (right) and T2 (left) graphs of CD2.7E15N. (C) T1 (right) and T2 (left) graphs of CD2.7E15Q. (D) T1 (right) and T2 (left) graphs of CD2.7E15E.

Protein concentrations were then plotted against reciprocal T values (relaxivity), which were calculated using Equation 1.6 (Figure 4.26), and fitted linearly. The point where the lines cross indicates the concentration at which saturation was observed. Since CD2.7E15 and the charged variants exhibit 1:1 binding, the graph began to saturate when the protein concentration reached 0.05 mM (in Figures 4.26 B and D) and continued to saturate as the protein to gadolinium ratio increased.

In Figure 4.26 A, the saturation point is reached when the protein to gadolinium ratio is 1.5:1. Past this point, the relaxivity values decrease sharply. This could be due to several reasons. First, although the protein was placed in dialysis with chelex Tris-HCl prior to this experiment, it is possible that small quantities of calcium are still present and are bound to the metal site (calcium and lanthanides have similar ionic radii). This may have prevented the gadolinium from saturating all the protein molecules and the concentration of protein at the saturation point is erroneously high. Additionally, during purification, it is possible that a portion of the protein misfolded during refolding. This implies that two species of protein with different folding properties and binding capabilities to gadolinium are present in the sample. A CD or tryptophan spectrum would be needed to confirm this.

In Figure 4.26 C, the curve begins to saturate when the protein concentration is 0.05 mM and then continues to increase. This could be due to incomplete saturation. Additionally, non-specific binding of gadolinium to other residues in CD2 is possible. Because of their configuration in space, residues D25, D27, D71, and E99 create another potential binding site for non-specific binding of calcium or other metals. Another possibility is experimental error.

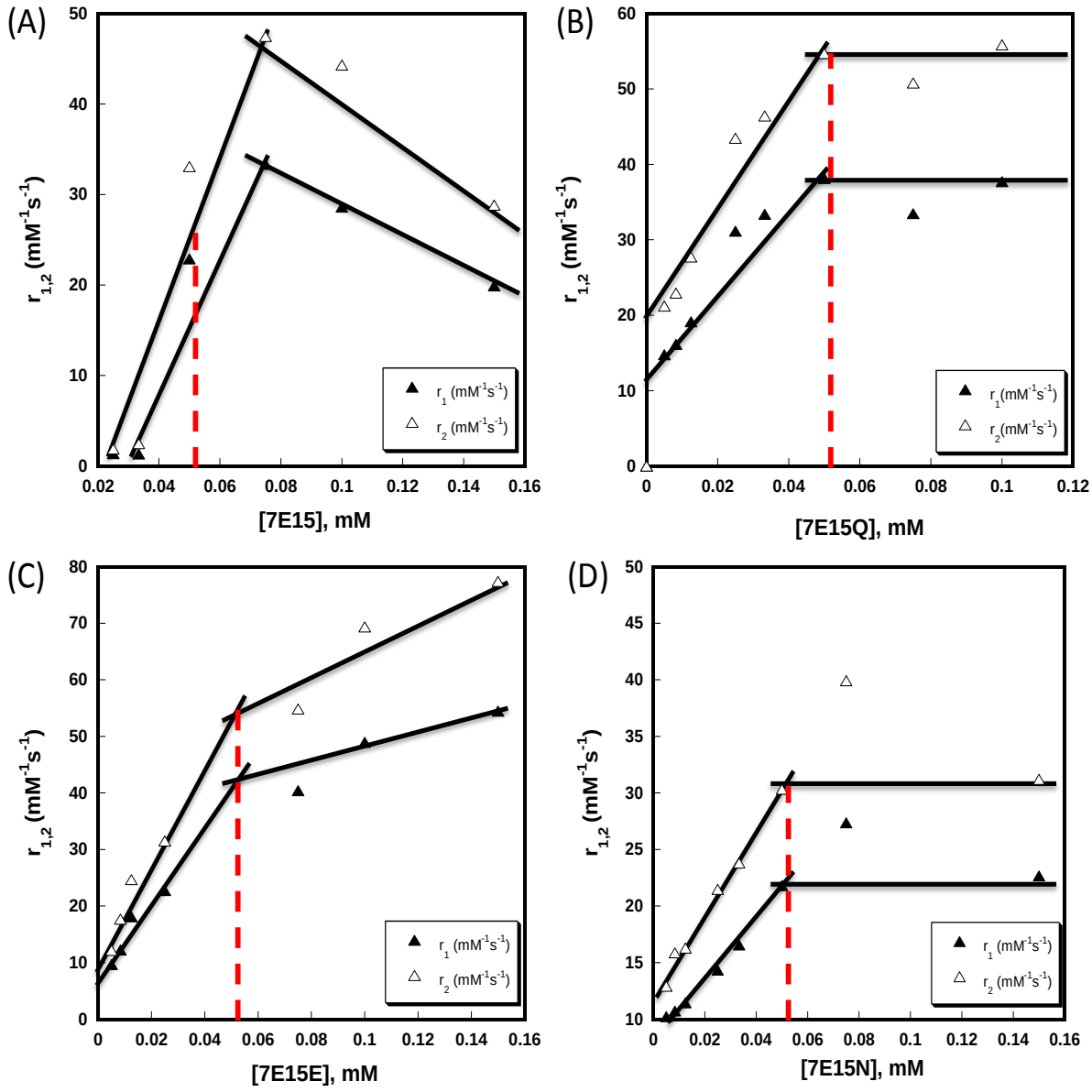


Figure 4. 26 Relaxivity of CD2.7E15 and charged variants with a gadolinium concentration of 0.05 mM, in 10 mM HEPES, pH 7, 25 °C, at 1.4 T. Plot of reciprocal T values against protein concentration is fitted linearly to determine saturation point. The red, dashed line represents the point where the protein to gadolinium ratio is 1:1 (A) CD2.7E15. (B) CD2.7E15Q. (C) CD2.7E15E. (D) CD2.7E15N.

Typically, the point where the protein to gadolinium ratio is 1:1 and above is averaged to find the relaxivity. However, because results need to be repeated and optimized, the relaxivity value at saturation point was used to determine r_1 and r_2 for each protein. The r_1 and r_2 values of CD2.7E15 were 22.9 $mM^{-1}s^{-1}$ and 33.1 $mM^{-1}s^{-1}$, respectively. The r_1 and r_2 values of CD2.7E15E were 22.7 $mM^{-1}s^{-1}$ and 33.9 $mM^{-1}s^{-1}$, respectively. The r_1 and r_2 values of CD2.7E15N were 21.9 $mM^{-1}s^{-1}$ and 30.4 $mM^{-1}s^{-1}$, respectively. The r_1 and r_2 values of CD2.7E15Q were 39.5 $mM^{-1}s^{-1}$ and 55.4 $mM^{-1}s^{-1}$, respectively. The relaxivity values for CD2.7E15 and the charged variants are within the same. Overall, the r_1 and r_2 values were significantly greater than the relaxivity of Gd-DTPA (Figure 4.27).

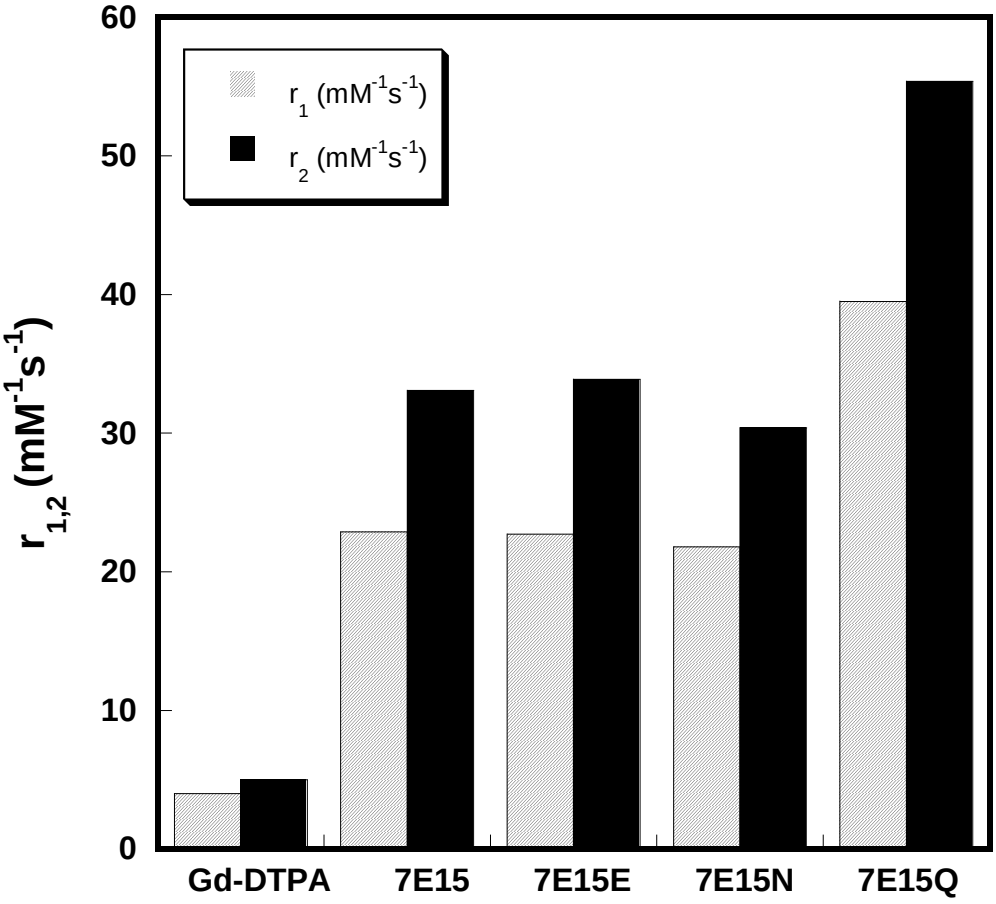


Figure 4. 27 Summary of r_1 and r_2 of CD2.7E15 charged variants in 10 mM HEPES, pH 7, 25 °C, 1.4 T shows comparable relaxivities to each other, which are significantly higher than the relaxivity of Gd-DTPA.

4.4 Summary and Significance

The Gd^{3+} binding affinities of charged variants were overall larger than the Zn^{2+} and Ca^{2+} binding affinities. (Figures 4.19 and Figure 4.24, and Table 4.4). This suggests that CD2.7E15 and its charged variants will bind to metals such as Gd^{3+} with significantly higher affinity than to physiological metals. Strong binding affinities of these contrast agents to Gd^{3+} is important to reduce toxic effects of free Gd^{3+}. It is important to note that the zinc binding affinity is not fitted optimally. However, the data is comparable to previously published data [1], in which the binding affinity of CD2.7E15 (in log K_a) is 6.72, compared to the data in Table 4.4. The metal selectivity of Gd^{3+} over Zn^{2+} or Ca^{2+} is significantly higher (Table 4.4). Additionally, the relaxivity values for CD2.7E15 charged variants are similar to each other, and overall they are significantly higher than the relaxivity of Gd-DTPA. The relaxivity results are preliminary and in the future should be repeated. Among the charged variants the difference in binding affinity is small and would be more pronounced if more charged variants such as -3 and -2 were tested. Still, -4 and -5 charged variants are suitable candidates for the next contrast agents.

Table 4.4 Summary of the K_d values of 7E15 charged variants to Gd^{3+} (5 µM Fluo-5N- 5 µM Gd^{3+}, in 20 mM PIPES, 10 mM KCl, pH 6.8), Tb^{3+} (5 µM 7E15 or 7E15N, in 10 mM Tris-HCl, pH 7.4 and 20 mM PIPES, 10 mM KCl, pH 6.8, respectively), Zn^{2+} (10 µM FluoZin-1-10 µM Zn^{2+}, in 10 mM Tris-HCl, pH 7.4), and Ca^{2+}. Error bars refer to average standard deviation from three separate trials. All work concerning Ca^{2+} K_ds can be found in Dr. Anna Wilkins Maniccia's dissertation [37].

	K_d Gd^{3+} (pM)	K_d Tb^{3+} (µM)*	K_d Zn^{2+} (µM)	K_d Ca^{2+} (mM)	Log K_a Gd^{3+}	Log K_a Zn^{2+}	Log K_a Ca^{2+}	Log (KGd^{3+}/KZn^{2+})	Log (KGd^{3+}/KCa^{2+})
7E15	2.31 ± 0.55	5.53±1.92	2.85±0.21	6.0±3.0	11.64	5.55	2.22	6.09	9.41
7E15E	1.45±0.39	--	3.33±0.50	--	11.84	5.48	--	6.36	11.84
7E15N	1.48±0.97	8.98±0.86	4.08±0.40	25.0±4.0	11.83	5.39	1.60	6.44	10.23
7E15Q	1.30±0.05	--	0.69±0.03	14.00±0.03	11.89	6.16	1.85	5.72	10.03

CHAPTER 5: ENHANCEMENT OF PROCA1 BY PEGYLATION AND PEPTIDE TAGGING

5.1. Introduction

5.1.1 PEGylation

PEGylation, or addition of a PEG (Polyethylene glycol) chain of various sizes to proteins, has drawn attention in therapeutic protein research as an option to increase the benefits and development of protein agents. Proteins with PEG chains retain their biological function while benefiting from added molecular weight provided by this chain, which decreases renal filtration and prolongs retention in the blood. Also, the protein surface is "masked" by this chain, which reduces the immunologic response incurred from introducing the protein into a biological system. Another benefit is an increase in the solubility of the protein and the biodistribution due to the hydrophilic PEG chain. Furthermore, the protein is protected from protease degradation [43; 44; 45]

In our laboratory, PEGylation has been performed on members of ProCA1 and has been noted to increase the relaxivity and biocompatibility of these agents; PEGylation of ProCA1 has been predicted to change the water relaxation rate (unpublished data). The questions raised in this chapter are whether PEGylation can be further optimize relaxivity of ProCA1, by varying the site of PEGylation among the six lysine residues in CD2.7E15 (positions 43, 45, 47, 51, 66, 91), and the effect of site-specific PEGylation relative to the location of the metal binding site. The hypothesis is that PEGylation near the binding site will increase the relaxivity of ProCA1 by increasing the water number and molecular rotational correlation time. This chapter will focus on preliminary work with PEGylation of 2 classes of mutated CD2.7E15. The first class, class 1, consists of K3CR, K5R, K5166R, and K6691R. K3CR contains 3 mutations from lysine to

arginine, specifically on the 3 lysine residues closest to the C terminal in positions 51, 66, 91. K5R has 5 mutations from lysine to arginine on all lysines except residue 45. K5166R and K6691R have lysine to arginine mutations on sites 51 and 66, and 66 and 91, respectively (Table 5.1). Changing the lysine to an arginine group spares the arginine from PEGylation so the remaining lysine groups are PEGylated. The second class of CD2.7E15 mutations, class 2, includes mutations from lysine to cysteine. They include K45C, K51C, and K66C, which each have one mutation from lysine to cysteine at sites 45, 51, and 66, respectively. The PEGylation agent targets the cysteine for PEGylation (Table 5.1).

Table 5. 1 Summarizes the different class 1 and class 2 mutations of CD2.7E15 for PEGylation with arginine or cysteine agents. PEG represents the PEGylated residue, R represents the arginine mutations which are spared from PEGylation, and C represents cysteine mutations which are PEGylated. Also, the current status of each protein is summarized.

Protein	43	45	47	51	66	91	Expressed	Purified	PEGylated
K3CR	PEG	PEG	PEG	R	R	R	✓	✓	
K5R	R	PEG	R	R	R	R	✓ (poor)		
K5166R	PEG	PEG	PEG	R	R	PEG	✓	✓	
K6691R	PEG	PEG	PEG	PEG	R	R	✓	✓	
K51C	K	K	K	PEG	K	K	✓	✓	✓
K66C	K	K	K	K	PEG	K	✓	✓	✓

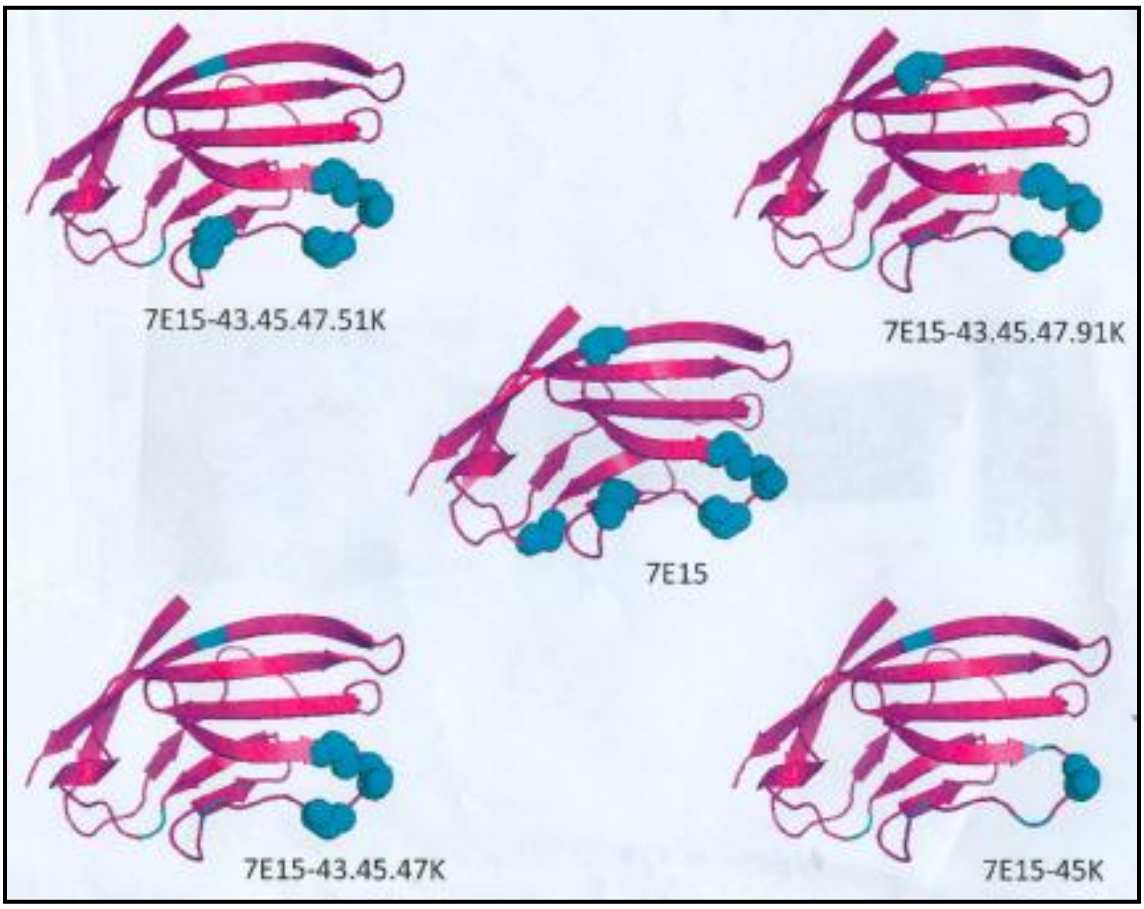


Figure 5. 1 Cartoon summarizing the different class 1 mutations of CD2.7E15 for PEGylation with arginine agents relative to the metal binding site of the protein.

5.1.2 Enhancement of CD2.7E15 Contrast Agents with Peptide Tags

The versatility of CD2.7E15 as a protein construct enables the addition of specific peptide tags which may target over-expressed receptors found on cancerous cells. This approach enhances the specificity of the contrast agents, which currently are limited to detection of material with diameter range above 1-2 mm. Preliminary studies (shown here) show expression and purification results from His-CD2.7E15-bom, bombesin tagged 7E15. Bombesin is a gastrin releasing peptide analog that targets gastrin-releasing peptide receptors, which are over-expressed in high levels on the surface of many cancerous cells [2]. This leads to enhancement of the image detectable by MRI. Published work shows insertion of a gastrin releasing peptide (GRP) (linker and 10 residues from the C terminal of GRP) into different regions in GST-tagged CD2.7E15[2].

5.2 Materials and Methods

5.2.1 Cloning of Mutants

Cloning of K3CR, K5R, K5166R, and K6691R was performed by Yanyi Chen. Cloning of K45C, K51C, K66C was performed under the guidance of Shenghui Xue. First, forward and reverse primers, 21-29 nucleotides long, were designed and ordered (from IDT) with the appropriate mutation for each protein and considering the melting temperature (T_m) (Table 5.2).

Table 5. 2 Designed forward (for) and reverse (rev) primers for cloning of K45C, K51C, and K66C.

		T_m (°C)
7E15K66C_for	Ttggacatatgtaatctgacaagagatg	68
7E15K66C_rev	Gtctccatttgcgtcgatctcaa	68
7E15K51C_for	Ttttttgtgttcgggagcatttgagat	66
7E15K51C_rev	Ggcttcatcttccttttaaactcg	68
7E15K45C_for	Aaaaggtgtatgaagccttttttgaaatc	68
7E15K45C_rev	Aaactcggcaaccagggtgct	66

Several samples with primer and DNA of CD2.7E15 underwent PCR (Techne TC 3000). First, the strands were denatured at 94 °C for 2 min, then the primers were allowed to bind to the DNA at 55 °C (2 min), followed by polymerase extending the primers at 72 °C (10 min). This was repeated for 30 cycles. Following the PCR reaction, a DNA gel was used to extract the DNA with a Gel Extraction Kit (Qiagen).

The DNA was then phosphorylated and ligated. Phosphorylation took place by adding 30 µL DNA to reaction buffer (5 µL), enzyme (1 µL), ATP (2 µL) and H_2O (12 µL), and incubating at 37 °C (30 min), followed by 65 °C (20 min). The DNA was then ligated by addition of 1 µL ligase and 2 µL ATP and stored at 4 °C overnight.

Following phophorylation and ligation, the DNA samples were inserted into DH5α cells, and were purified using the Miniprep kit (Qiagen). The DNA from each sample was sequenced and the samples containing the mutation were transformed and kept for expression and purification.

Cloning for His-CD2.7E15-bom was performed by Lixia Wei.

5.2.2 Expression of K3CR, K5R, K5166R, K6691R, K45C, K51C, K66C, and His-CD2.7E15-bom

A bacterial colony was inoculated and incubated overnight in LB medium with 2.5 mM ampicillin. Flasks containing 1 L LB medium and 2.5 mM ampicillin received 50 mL of the culture grown overnight, and the 1 L flasks were incubated at 37 °C in a shaker (Shaker Controlled Environment Incubator).

The optical density (OD) of each sample at 600 nm was recorded every 30 minutes using UV spectrophotometer (Shimadzu Scientific Instruments, Norcross, GA). When the OD reached 0.6 for BL21 (DE3), Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the flasks. The

BL21 (DE3) strain required 0.25 mM IPTG. The shakers were adjusted 30 °C for the duration of expression for K3CR, K5R, K5166R, K6691R, and remained 37°C for K51C, K66C, and His-CD2.7E15-bom. The cell pellets were collected by centrifugation at 4.410 x 10^3 *g* for 20 minutes after overnight expression. The cell pellets were stored at -20 °C until purification. KaleidaGraph software (Synergy Software, Reading, PA) was used to plot the optical density versus time plots. Image J software (Version 1.37) was used to quantify band thickness as a result of expression on the 15% acrylamide SDS-PAGE.

5.2.3 Purification of K3CR, K5166R, K6691R, K51C, and K66C

Purification was performed using the refolding method as described in section 2.2.7 with modification. Cell pellet from expression (approximately 5 g) was suspended in lysis buffer containing PBS. Benzonase nuclease (60 µL) and serine inhibitor PMSF (200 µL) were added. The cell pellet was broken using sonication (6 times for 20 seconds each), French press (1500 psi), or cell disruption. The mixture was then centrifuged (S34 rotor) at 1.2947 x 10^4 *g* for 30 minutes and supernatant was discarded. The cell pellets were then suspended in 20 mL of 2% Triton X-100, and centrifuged at 1.2947 x 10^4 *g* for 20 minutes between each of three washings. Next, the cell pellet was dissolved in 25 mL of 8 M urea (pH 7.3) and stirred overnight at 4 °C to completely solubilize the cell pellet.

The solublized cell pellet was centrifuged for 20 minutes at 1.2947 x 10^4 *g*. PBS was then added drop-wise for a final concentration of 4 M urea and the solution was stirred for 2 hours. Next, the solution was placed in dialysis for 4 hours with 2 M urea, followed by addition of proteases (200 µL) and overnight dialysis with 2 L of 10 mM Tris-HCl buffer (pH 7.4). The dialysis bag has pores which allow molecules with molecular weight over 3.5 k Da to remain in

the bag. The dialysis buffer was changed to 10 mM Tris the following morning, which was replaced to gradually remove urea from the solution. Subsequently, the solution was centrifuged at 1.2947 x 10^4 *g* for 20 minutes, and the supernatant was kept.

5.2.4 Purification of His-CD2.7E15-bom

Cell pellet was dissolved in lysis buffer, sonicated, and centrifuged. The cell pellet was kept and dissolved in 50 ml PBS containing 1% Triton x-100, and was then centrifuged for 30 minutes at 1.2947 x 10^4 *g*. This was repeated in triplicate. The cell pellet was dissolved in 1 M arginine, and the mixture was left on the agitator at 4 °C overnight. The following day, centrifugation was performed (1.2947 x 10^4 *g* for 30 minutes) and the supernatant containing solubilized protein was collected. Size Exclusion chromatography was performed. SDS-PAGE was used to image the supernatants and cell pellets after sonication, washing, and solubilization.

5.2.5 PEGylation of K51C and K66C

Protein (100 µM) was placed in dialysis with PBS and pumped with nitrogen gas. A reducing agent, TCEP, was added in 1:10 ratio. The nitrogen gas was stopped and the sample was placed on the shaker at 4°C. The protein was concentrated (Amicon) to a fourth of its initial volume, and PEG-Cys reagent ((methyl-PEG_{12})$_3$-PEG_4-Malemide) was added in a 1:10 ratio. Nitrogen gas was pumped through the system, which was sealed and placed at 4 °C overnight. SDS-PAGE was used to image protein pre-PEGylation and post-PEGylation and the gel was stained with iodine to determine if PEGylation was successful. The gels was first washed in water, then with $BaCl_2$ (15 min), then with iodine (4 min), and lastly with water. The excess protein was removed using the Q column.

5.3 Results and Discussion

5.3.1 Expression and Purification of K3CR, K5R, K5166R, and K6691R

K3CR, K5R, K5166R, and K6691R were expressed in tag-less vector pET20b with BL21 (DE3) cell strain at 37 °C overnight. The expression graph is exponential. When the optical density (OD) reading reached 0.600 at 600 nm, the flasks received 0.2 mM IPTG (Figure 5.2). Following overnight expression the flasks were harvested. SDS-PAGE shows band thickening after induction indicating expression of K3CR (Figure 5.2 A), K6691R (Figure 5.2 B), and K5166R (Figure 5.2 D). No expression was seen for K5R (Figure 5.2 C).

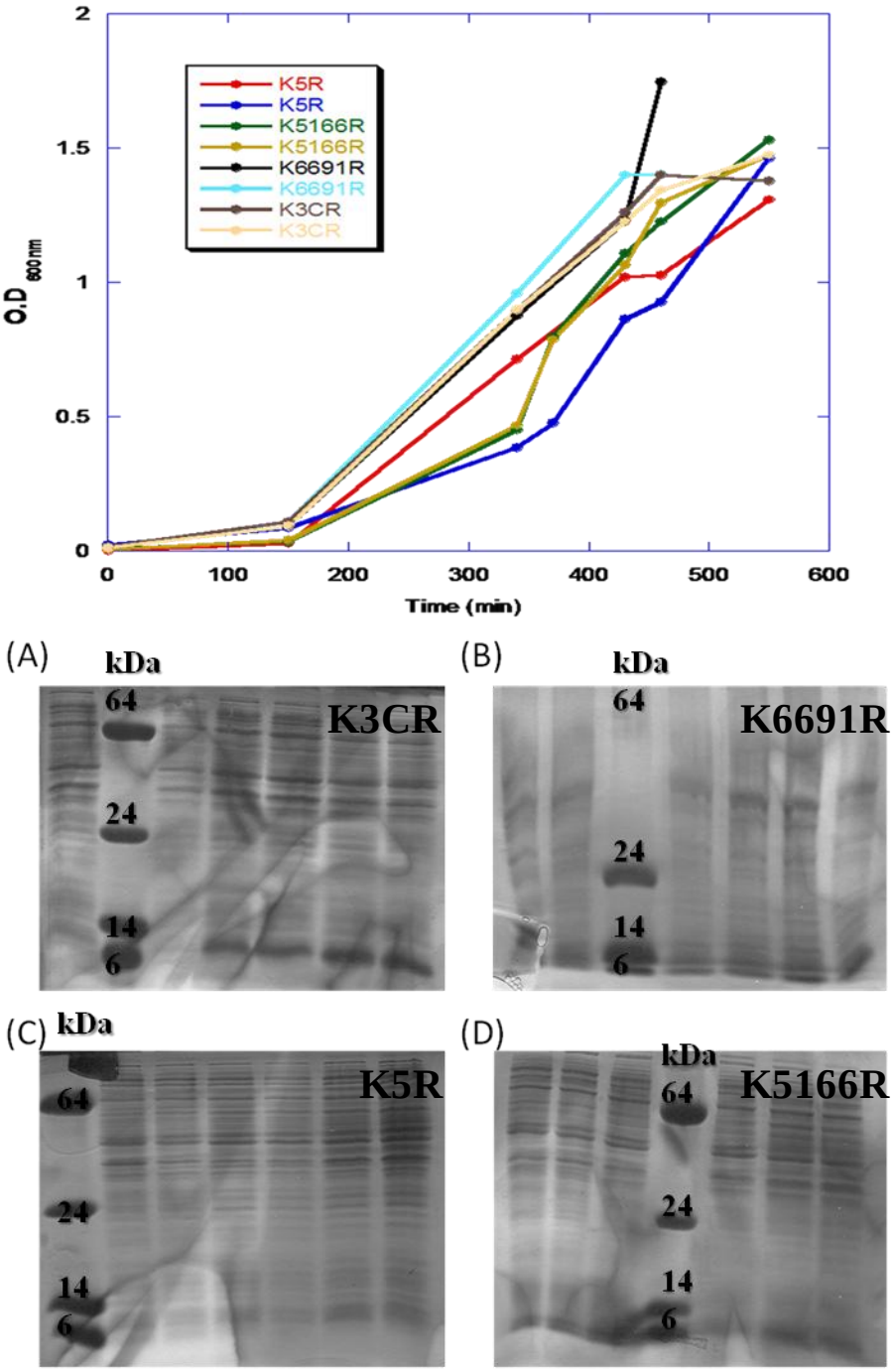


Figure 5.2 Optical density graphs and expression gels for expression in pET20b and BL21 (DE3) at 30 °C after induction overnight show expression for K3CR, K6691R, and K5166R. (A) K3CR. (B) K6691R. (C) K5R. (D) K5166R.

The refolding method was used to purify K3CR, K5166R, and K6691R. Following French press, the proteins were found in inclusion bodies (cell pellets). They were washed with 2% Triton X-100 to remove lipid-soluble dirt, and were dissolved in 8 M urea. The concentration of urea was reduced gradually to 4 M, then to 2 M, and finally 0 M (10 mM Tris-HCl buffer). Following centrifugation, the protein was found in the supernatant and after filtration was injected into the SP column. The gels show protein in fractions 13 and 17 (Figure 5.3 A, B, and C). The concentrations were measured using UV. For K3CR, the yield was 1.2 mg. For K5166R, the yield was approximately 3.5 mg, and for K6691R, the yield was 1.7 mg. A terbium titration verified that the conformation was stable (not shown).

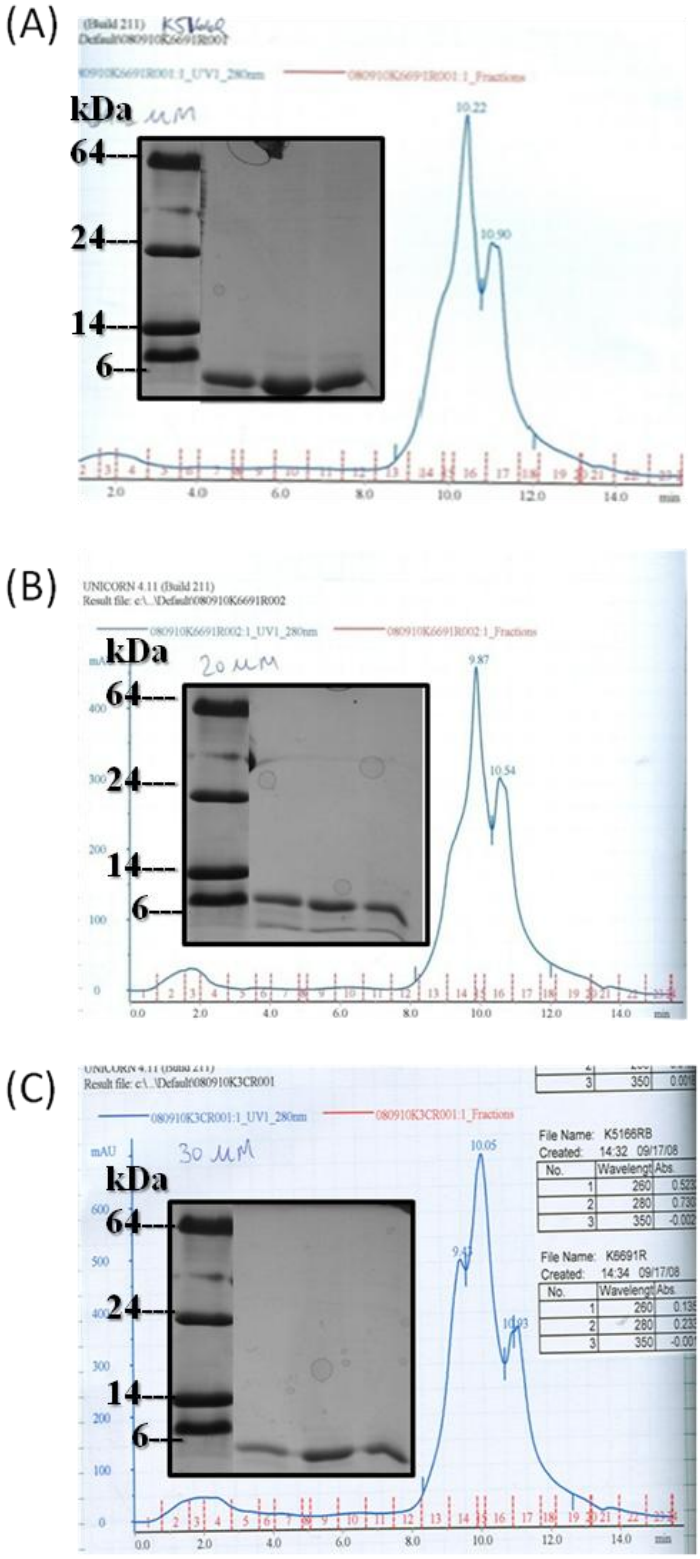


Figure 5.3 Purification of K3CR, K5166R, and K6691R. (A) K5166R. (B)K6691R.(C) K3CR. The gels correspond to the tallest peaks.

5.3.2 Expression, Purification, and PEGylation of K51C and K66C

K51C and K66C were expressed successfully with BL21 (DE3) cell strain and vector pET20b at 37 °C for 4 hours. Expression graphs demonstrate exponential rise in OD. At an OD of 0.600 at 600 nm, the flasks received 0.25 mM IPTG. SDS-PAGE shows thickening bands with the passage of time, illustrating that the proteins were expressed successfully during expression, especially after induction (Figure 5.4). The flasks were harvested after 4 hours.

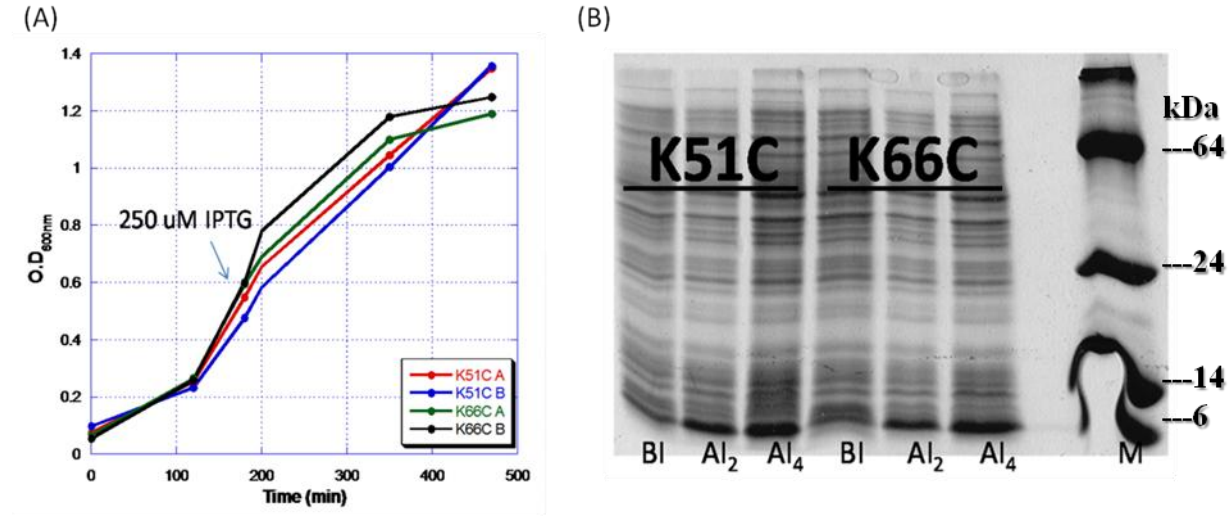


Figure 5.4 Expression of K51C and K66C with BL21 (DE3) cell strain and pET20b vector at 37°C. (A) Optical density curve indicating is exponential. (B) SDS-PAGE showing presence of protein during expression.

Purification was performed via the refolding method described in section 5.2.2. Successful purification is shown by SDS-PAGE. The supernatant after French press lacks protein, which is present the cell pellet; rather the cell pellet after French press shows protein representing inclusion bodies from expression. Following wash with 2% Triton X-100, the cell pellet looks cleaner and no protein is lost in the supernatant. After dissolving in 8 M urea the protein in the supernatant, and the cell pellet after 8 M urea shows very little insoluble protein after refolding. After 10 mM Tris buffer, the supernatant shows clean product ready for injection into FPLC (Figure 5.5).

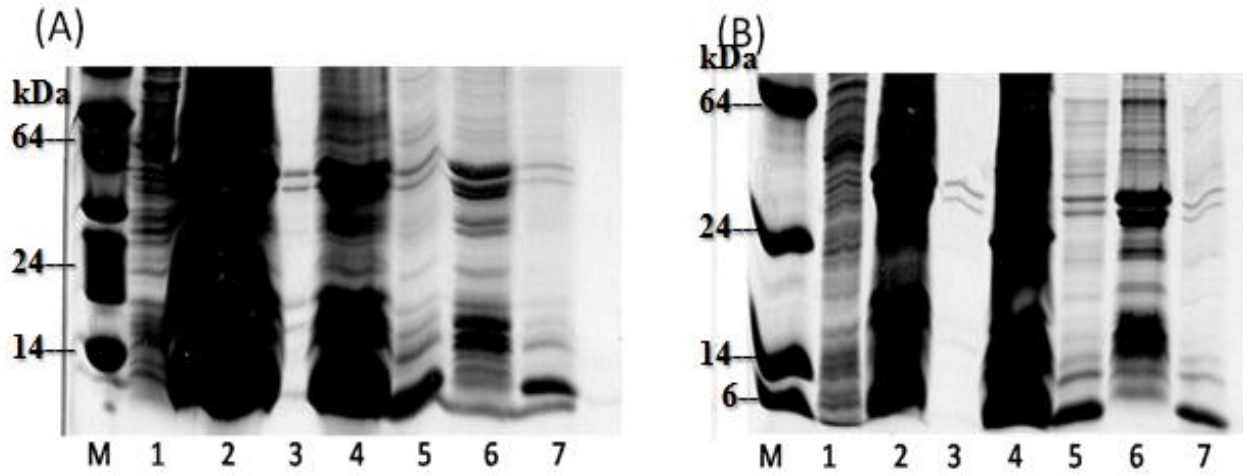


Figure 5.5 Purification of K51C and K66C. (A) K66C. (B) K51C. M-marker; 1- supernatant after French press shows no protein as it is all in the cell pellet; 2- cell pellet after French press shows protein representing inclusion bodies; 3- supernatant after 2% Triton X-100 does not contain protein; 4- cell pellet after Triton X-100 contains cell pellet that is cleaner than before wash; 5-supernatant after solubilization with 8 M urea shows the protein in the supernatant; 6-cell pellet after 8 M urea shows very little insoluble protein after refolding; 7- supernatant after 10 mM Tris buffer shows clean product ready for injection into FPLC.

The protein (supernatant after 10 mM Tris) was injected into the SP column, following adjustment of the pH of proteins to 3.0 with HCl. The peak in fractions 29-31 was found to contain protein with high purity (confirmed by SDS-PAGE) (Figure 5.6 and 5.7). The yield of K51C was 6.7 mg/ 10 mL injection and the yield for K66C was 5.2 mg/ 10 mL injection.

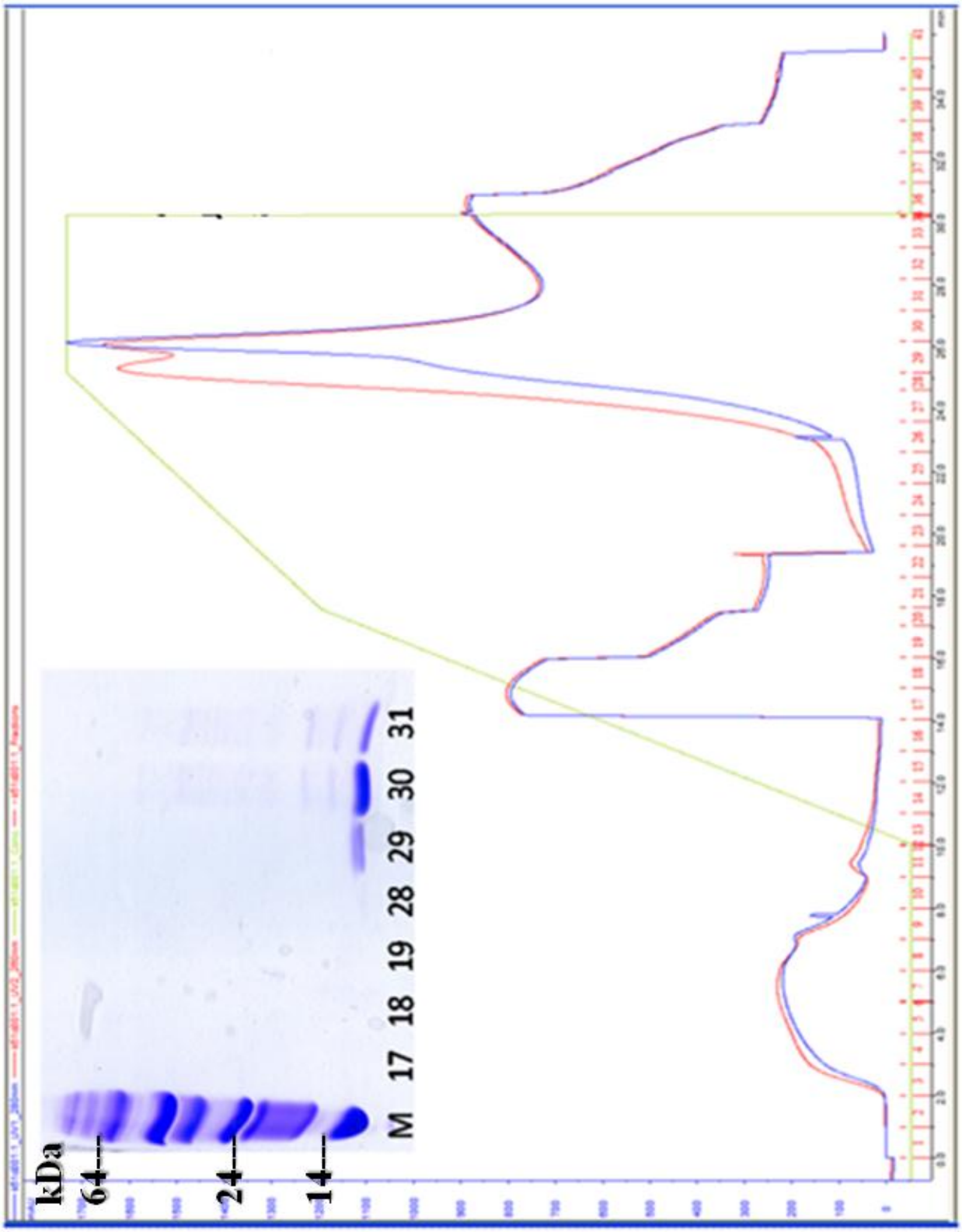


Figure 5.6 Purification of K51C with SP column (flow rate 2 mL/min)

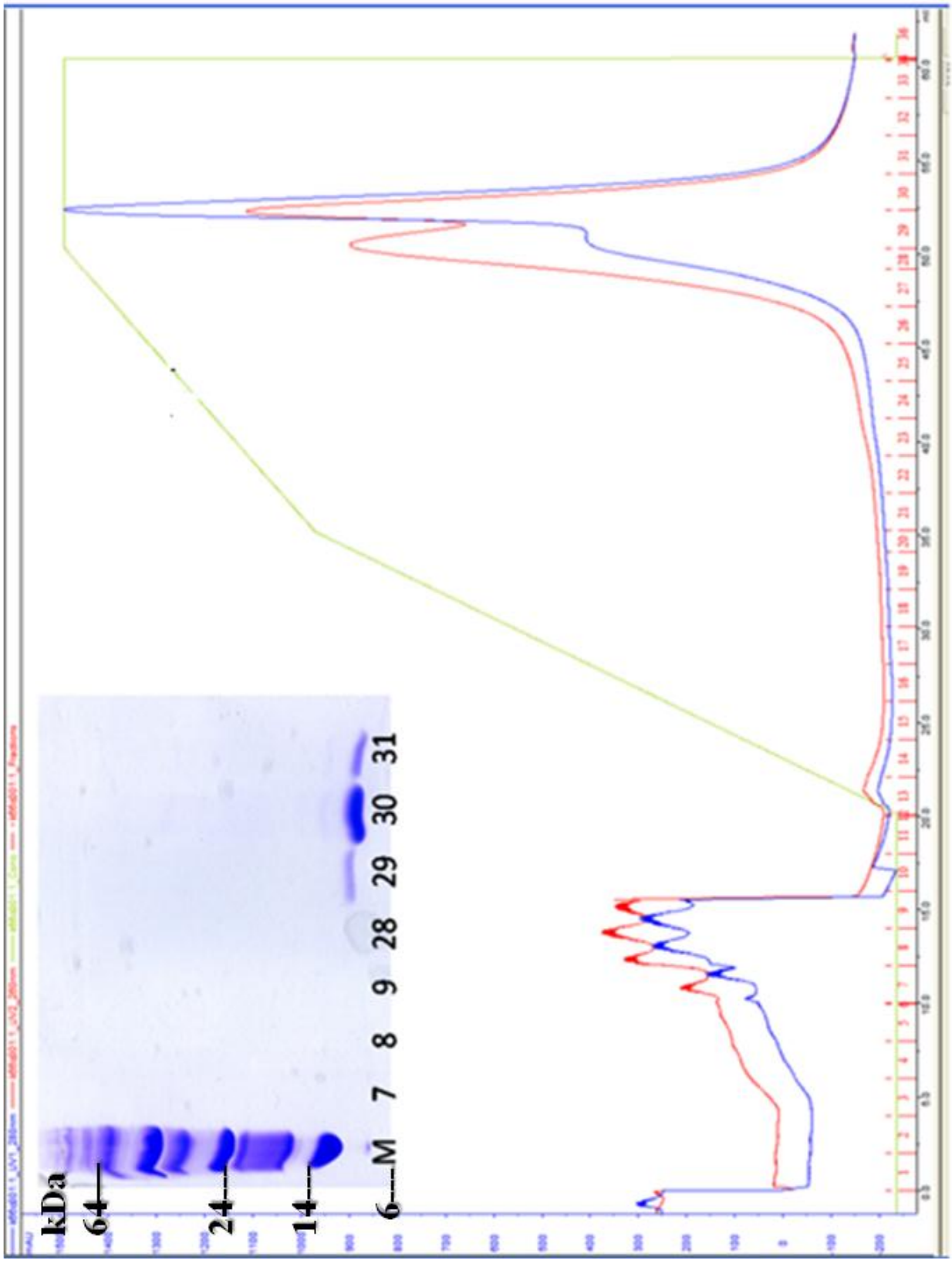


Figure 5.7 Purification of K66C with SP column (flow rate 2 mL/min)

Following purification 100 µM protein was PEGylated. Iodine-stained gel shows successful PEGylation of K51C and K66C (Figure 5.8). After PEGylation, the protein was injected into the Q column to remove excess reagent, because lowing the pH for the SP column resulted in precipitation. Following injection into the Q column, the protein was contaminated with DNA and was incubated with fresh benzonase nuclease for 4 hours at room temperature.

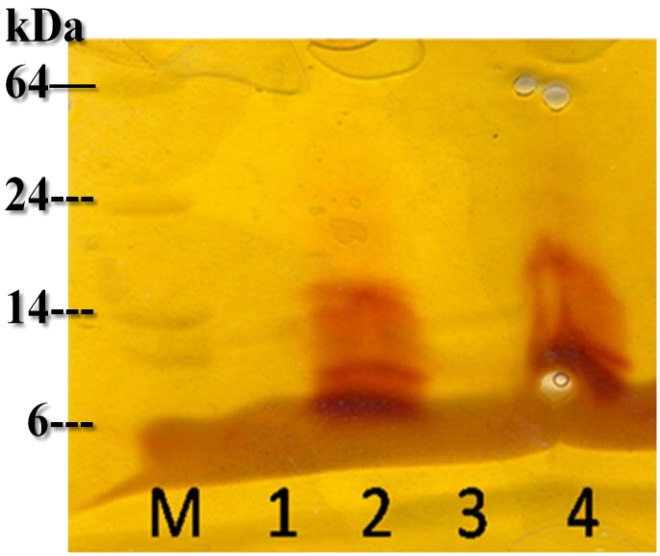


Figure 5.8 Iodine-stained gel depicting PEGylation of K51C and K66C. M-marker; 1-K51C before PEGylation; 2-K51C after PEGylation; 3- K66C before PEGylation; 4-K66C after PEGylation with (methyl-PEG_{12})$_3$-PEG_4-Maleimide reagent.

5.3.3 Expression and Purification of His-CD2.7E15-bom

Expression of His-CD2.7E15-bom was performed at 37 °C in BL21 (DE3) cells. Upon induction the flasks received 1 mM IPTG, discussed in section 3.3.2 as the optimal amount. SDS-PAGE shows an increase in protein production after induction (Figure 5.9).

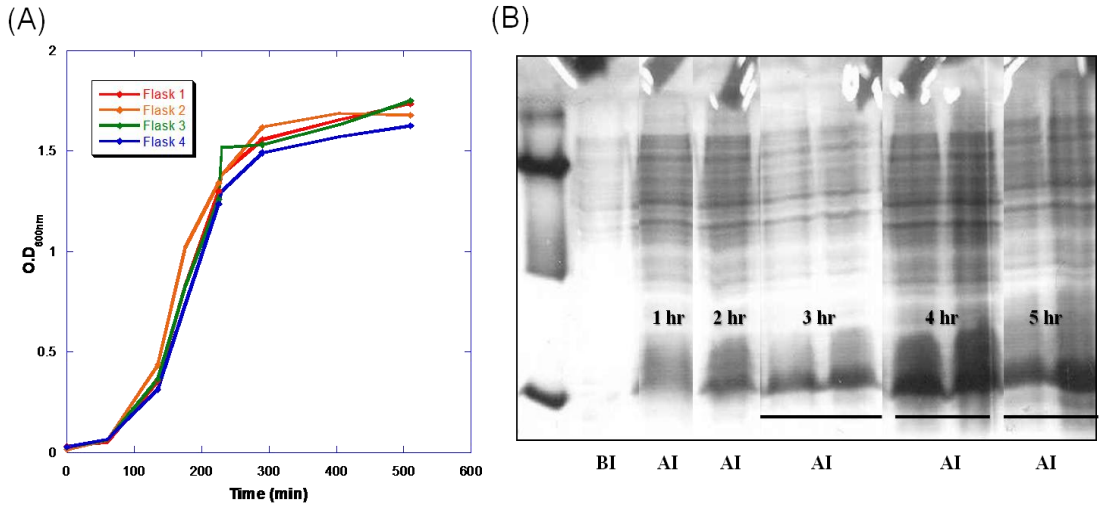


Figure 5.9 Expression of His-7E15-bom at 37 °C in BL21 (DE3) cells with 1 mM IPTG at induction

Purification of His-CD2.7E15-bom was not optimal. After sonication, SDS-PAGE showed that the protein was in an inclusion body. Cell pellet was partially dissolved in 50 ml 1% Triton x-100, and was then centrifuged for 30 minutes at 1.2947 x 10^4 *g*, for a total of 3 washing steps. However, after cell pellet was placed in 1M arginine, placed on the agitator overnight at 4°C, and centrifuged the following day, SDS-PAGE showed low solubilization of the protein in arginine, represented by a thin band in the supernatant. Size Exclusion chromatography was performed which shows almost no protein on the SDS-PAGE.

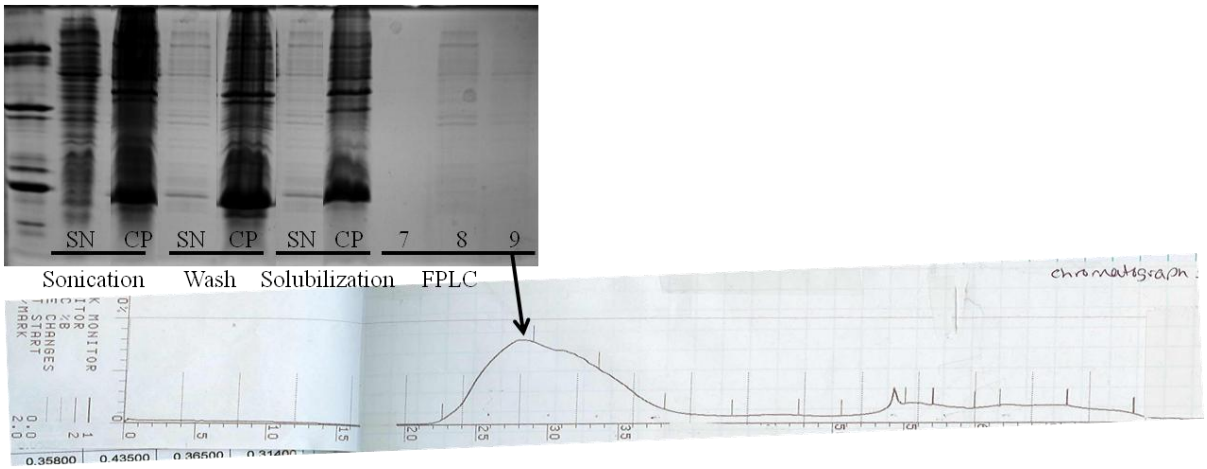


Figure 5.10 Size Exclusion chromatography shows almost no product of His-CD2.7E15-bom when fractions 7-9 are visualized on SDS-PAGE.

5.4 Conclusions and Future Work

Modification of contrast agents by PEGylation or addition of peptide chains has been predicted to enhance the effectiveness of the contrast agents. Preliminary studies show that optimization of expression, purification, and PEGylation is needed. Expression of arginine mutants (class 1) was low, yet successful for K3CR, K5166R, and K6691R. Yields after purification were low (1.7-3.2 mg) because expression was low. K5R did not express well, and conditions for expression need to be optimized. After purification yield is increased, then a sufficient amount of protein will be available for PEGylation, relaxivity studies, and structural examination as contrast agents.

Expression and purification of cysteine mutants (class 2) was successful with yields as high as 5.2-6.7 mg per 10 mL injection. PEGylation was successful; additional dosages of benzonase nuclease may help to reduce contamination by DNA for further optimization of the protocol. Once contamination is removed, the PEGylated protein can be used further in structural and relaxivity studies as a contrast agent.

Preliminary expression and purification studies of His-CD2.7E15-bom showed successful expression with an optimized level of IPTG, yet purification in low yield. Later studies performed by Lixia Wei show higher purification yield using a GST tag. Furthermore, the protein has been studied as a contrast agent bonded to a GRP targeting sequence (instead of the analog, bombesin) and has shown high relaxivity as a contrast agent [2].

CHAPTER 6: SIGNIFICANCE OF THIS WORK

MRI is an excellent diagnostic tool, providing a wealth of information about tissues *in vivo* and without the use of radiation. Contrast agents used to enhance MRI have several limitations that have led to the search of new contrast agent. We propose a protein-based contrast agent, designed by inserting a metal binding site into CD2. Scaled up production of these proteins is necessary for structural, functional, therapeutic, and *in vivo* analysis, studies which are necessary to bring the contrast agents to a clinical level. The objective of this thesis was to examine expression and purification conditions for CD2.7E15 and variants, because these conditions are vital to maximizing the production of CD2.6D31, CD2.7E15, and CD2.7E15 charged variants with the least time and effort. Determination of high efficiency expression and purification method allows further evaluation of various engineered proteins as suitable contrast agents by relaxivity and structural studies.

To begin the protein mass-production studies, wild type CD2, CD2.7E15, CD2.6D31 were chosen as models to reevaluate the old expression and purification procedures. They were expressed as a fusion protein with the Glutathione-S-Transferase (GST) affinity tag, with three temperatures and three cell strains, and were purified using affinity column chromatography. Optimal expression conditions were determined, which were different for each CD2 variant. Furthermore, the quantity of soluble protein produced during expression was variable for each protein, and depended on the temperature and cell strain as well.

Next, GST purification was examined in light of expression. Affinity tag GST folding was found to largely depend on the engineered protein folding (which depended on expression). Correlation between beads binding assay and GST enzymatic activity indicate that high enzymatic activity in GST resulted in larger quantity of fusion protein binding to the affinity

columns. High purification level is achievable when the quality of GST folding is high, which is tied back to the expression conditions. This information helped determine which conditions are conducive to producing soluble protein and which encouraged inclusion body formation in the three engineered proteins. This was a powerful database for later selection of expression conditions for CD2.7E15 charged variants (chapter 2).

A rapid, effective method for generating high yield of CD2.7E15 has been established by generating inclusion bodies (chapter 2) and testing two ways to purify them. The first included solubilization in arginine followed by denaturation in urea. However, the protein did not solubilize well in arginine, as indicated by SDS-PAGE, and only a small quantity of protein was extracted. Purification by direct immersion of cell pellet in 8 M urea at room temperature showed a high level of solubilized protein. Structural studies with CD and tryptophan fluorescence provide evidence that suggests that the refolding method can reliably generate protein that does not lose its native conformation. The theoretical yield for a liter of expressed cell culture has risen impressively to 156 mg -234 mg pure protein from 1 L expressed cell pellet, a tremendous improvement from traditional GST purification (chapter 3).

Using this high-yield purification method, it was possible to generate large quantities of protein for structural, metal binding, relaxivity studies, *in vivo* studies (future), and determination of the suitability of charged variants of ProCA1 as contrast agents. Charged variants had a higher selectivity for Gd^{3+} over Zn^{2+} and Ca^{2+}, suggesting that these proteins will bind to metals such as Gd^{3+} and lanthanides with much higher affinity than to physiological metals. A strong binding affinity to Gd^{3+} is an important quality for a contrast agent to reduce the toxic effects of free Gd^{3+}. Strong binding is advantageous to enhance imaging as well. The metal selectivity of Gd^{3+} over Zn^{2+} or Ca^{2+} was also found to be significantly higher. Among the different charged

variants, the difference in binding affinity was small. The relaxivities of CD2.7E15 and the charged variants were similar to each other but overall were significantly higher than the relaxivity of Gd-DTPA. We conclude that -4 and -5 charged variants are suitable candidates for the next contrast agents based on metal selectivity and high relaxivity (chapter 4).

Research on improving contrast agents is focused on insertion of PEG chains or peptide tags (bombesin) in proteins. These options were studied with ProCA1 as possible ways to enhance protein-based contrast agents. Preliminary studies show that future work is needed to further develop this approach. Arginine mutants (K3CR, K5166R, and K6691R) were expressed, albeit in low quantities. The mutants were purified successfully in low yield (1.7-3.2 mg), but the yield requires optimization for continued studies of this protein by PEGylation, relaxivity studies, and structural examination. Expression and purification of cysteine mutants was successful with yields as high as 5.2-6.7 mg per 10 mL injection. Additionally, PEGylation was successful, except that the product was contaminated by DNA. Optimization of the purification procedure will allow the PEGylated protein to be further used in structural and relaxivity studies. Preliminary expression and purification studies of His-CD2.7E15-bom led way for the protein to be studied as a contrast agent bonded to a GRP targeting sequence (Lixia Wei) [2], which has shown high relaxivity as a contrast agent (chapter 5). The studies in this thesis have contributed to the growing field of MRI research by providing more candidates for protein-based contrast agents.

CHAPTER 7: REFERENCES

[1]J.J. Yang, J. Yang, L. Wei, O. Zurkiya, W. Yang, S. Li, J. Zou, Y. Zhou, A.L. Maniccia, H. Mao, F. Zhao, R. Malchow, S. Zhao, J. Johnson, X. Hu, E. Krogstad, Z.R. Liu, Rational design of protein-based MRI contrast agents. J Am Chem Soc 130 (2008) 9260-9267.

[2]L. Wei, S. Li, J. Yang, Y. Ye, J. Zou, L. Wang, R. Long, O. Zurkiya, T. Zhao, J. Johnson, J. Qiao, W. Zhou, A. Castiblanco, N. Maor, Y. Chen, H. Mao, X. Hu, J.J. Yang, Z.R. Liu, Protein-based MRI contrast agents for molecular imaging of prostate cancer. Mol Imaging Biol 13 (2011) 416-423.

[3]J. Voyvodic, Basic Principles of MRI

[4]W.M. Faulkner, Basic Principles of MRI. (1996).

[5]W.R. Hendee, C.J. Morgan, Magnetic resonance imaging. Part I--physical principles. West J Med 141 (1984) 491-500.

[6]B. Mackiewich, Basic Principles of MRI (Available from http://www.cs.sfu.ca/~stella/papers/blairthesis/main/node11.html).

[7]J. Voyvodic, Basic Physical Principles of MRI.

[8]P. Caravan, Strategies for increasing the sensitivity of gadolinium based MRI contrast agents. Chem Soc Rev 35 (2006) 512-523.

[9]Biosciences, Nuclear Magnetic Resonance Spectroscopy Theoretical Principles. (Available from http://teaching.shu.ac.uk/hwb/chemistry/tutorials/molspec/nmr1.htm).

[10]W.-T.W. Kannie Wai-Yan Chan, Small molecular gadolunium (III) complexes as MRI contrast agents for diagnostic imaging Coordination Chemistry Reviews 251 (2007) 2428-2451.

[11]W. Krause, Contrast Agents I Magnetic Resonance Imaging Springer-Verlag Berllin: Heidelberg, 2002.

[12]W. Parker, Alonso Gadolinium MRI Contrast Dye Side Effect Injuries Nephrogenic Systemic Fibrosis/ Nephrogenic Fibrosing Dermopathy, in.

[13]S. Laurent, L.V. Elst, R.N. Muller, Comparative study of the physicochemical properties of six clinical low molecular weight gadolinium contrast agents. Contrast Media Mol Imaging 1 (2006) 128-137.

[14]A.L. Wilkins, W. Yang, J.J. Yang, Structural biology of the cell adhesion protein CD2: From molecular recognition to protein folding and design. Curr Protein Pept Sc 4 (2003) 367-373.

[15]A.W. Maniccia, W. Yang, J.A. Johnson, S. Li, H. Tjong, H.X. Zhou, L.A. Shaket, J.J. Yang, Inverse tuning of metal binding affinity and protein stability by altering charged coordination residues in designed calcium binding proteins. PMC Biophys 2 (2009) 11.

[16]S. Li, W. Yang, A.W. Maniccia, D. Barrow, Jr., H. Tjong, H.X. Zhou, J.J. Yang, Rational design of a conformation-switchable Ca2+- and Tb3+-binding protein without the use of multiple coupled metal-binding sites. FEBS J 275 (2008) 5048-5061.

[17]A.W. Maniccia, Analyzing the effects of net charge and ligand type to metal binding affinities for calcium-binding proteins by design (2005).

[18]J. Arnau, C. Lauritzen, G.E. Petersen, J. Pedersen, Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expr Purif 48 (2006) 1-13.

[19]S.C. Makrides, Strategies for achieving high-level expression of genes in Escherichia coli. Microbiol Rev 60 (1996) 512-538.

[20]Novagen, Available from http://www.genomex.com/vector_maps/pet20b+.pdf.

[21]Amersham, Available from http://www.lablife.org/p?a=vdb_view&id=g2.fpzDkSJdEuqykKDjECRqRQzAQdI (Plasmid vector map).

[22]J.J. Lichty, J.L. Malecki, H.D. Agnew, D.J. Michelson-Horowitz, S. Tan, Comparison of affinity tags for protein purification. Protein Expr Purif 41 (2005) 98-105.

[23]A. Carroll, W. Yang, Y. Ye, R. Simmons, J.J. Yang, Amyloid fibril formation by a domain of rat cell adhesion molecule. Cell Biochem Biophys 44 (2006) 241-249.

[24]J.Y. Chang, Thrombin specificity. Requirement for apolar amino acids adjacent to the thrombin cleavage site of polypeptide substrate. Eur J Biochem 151 (1985) 217-224.

[25]G.S. Waldo, B.M. Standish, J. Berendzen, T.C. Terwilliger, Rapid protein-folding assay using green fluorescent protein. Nat Biotechnol 17 (1999) 691-695.

[26]W. Yang, L.M. Jones, L. Isley, Y. Ye, H.W. Lee, A. Wilkins, Z.R. Liu, H.W. Hellinga, R. Malchow, M. Ghazi, J.J. Yang, Rational design of a calcium-binding protein. J Am Chem Soc 125 (2003) 6165-6171.

[27]W. Yang, A.L. Wilkins, Y. Ye, Z.R. Liu, S.Y. Li, J.L. Urbauer, H.W. Hellinga, A. Kearney, P.A. van der Merwe, J.J. Yang, Design of a calcium-binding protein with desired structure in a cell adhesion molecule. J Am Chem Soc 127 (2005) 2085-2093.

[28]W.H. Habig, M.J. Pabst, W.B. Jakoby, Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. J Biol Chem 249 (1974) 7130-7139.

[29]Glutathione-S-Transferase, Image available from http://en.wikipedia.org/wiki/Glutathione_S-transferase.

[30]E. Biosciences, pET-30a (+) map. Available from http://www.biovisualtech.com/bvplasmid/pET-30_a_%28+%29.htm.

[31]M. Campbell, SDS-PAGE (PolyAcrylamide Gel Electrophoresis), in, 2001.

[32]R.J. Simpson, Large-scale extraction of recombinant proteins from bacteria. Cold Spring Harb Protoc 2010 pdb prot5484.

[33]M. Umetsu, K. Tsumoto, S. Nitta, T. Adschiri, D. Ejima, T. Arakawa, I. Kumagai, Nondenaturing solubilization of beta2 microglobulin from inclusion bodies by L-arginine. Biochem Biophys Res Commun 328 (2005) 189-197.

[34]C. Putnam, Protein Calculator in, 2006

[35]J.A. Johnson, The Production of Designed Potential Protein Contrast Agents and Their Encapsulation in Albumin Microspheres. (2008).

[36]J.J. Yang, Y. Ye, A. Carroll, W. Yang, H.W. Lee, Structural biology of the cell adhesion protein CD2: alternatively folded states and structure-function relation. Curr Protein Pept Sci 2 (2001) 1-17.

[37]A.W. Maniccia, Analyzing the Effects of Net Charge and Ligand Type to the Metal Binding Affinities for Calcium-binding Proteins by Design (2005).

[38]W. Yang, T. Tsai, M. Kats, J.J. Yang, Peptide analogs from E-cadherin with different calcium-binding affinities. J Pept Res 55 (2000) 203-215.

[39]J.J. Yang, J. Yang, L. Wei, O. Zurkiya, W. Yang, S. Li, J. Zou, Y. Zhou, A.L. Maniccia, H. Mao, F. Zhao, R. Malchow, S. Zhao, J. Johnson, X. Hu, E. Krogstad, Z.R. Liu, Rational design of protein-based MRI contrast agents. J Am Chem Soc 130 (2008) 9260-9267.

[40]A.W. Maniccia, W. Yang, S.Y. Li, J.A. Johnson, J.J. Yang, Using protein design to dissect the effect of charged residues on metal binding and protein stability. Biochemistry-Us 45 (2006) 5848-5856.

[41]G.D. Scholes, Long-range resonance energy transfer in molecular systems. Annu Rev Phys Chem 54 (2003) 57-87.

[42]Invitrogen, (Available from http://www.invitrogen.com/site/us/en/home/).

[43]F.M. Veronese, Peptide and protein PEGylation: a review of problems and solutions. Biomaterials 22 (2001) 405-417.

[44]B. Plesner, C.J. Fee, P. Westh, A.D. Nielsen, Effects of PEG size on structure, function and stability of PEGylated BSA. Eur J Pharm Biopharm.

[45]F.M. Veronese, G. Pasut, PEGylation, successful approach to drug delivery. Drug Discov Today 10 (2005) 1451-1458.

[46]J.P. Hornak, The Basics of NMR, available from http://www.cis.rit.edu/htbooks/mri/inside.htm