

I. CRYSTALLIZATION OF CRBP II MUTANTS AS A PROBE FOR
UNDERSTANDING WAVELENGTH REGULATION
II. REENGINEERING AND CRYSTALLIZATION OF CELLULAR RETINOL BINDING
PROTEIN II (CRBP II) AS A FLUORESCENT TAG
III. SYNTHESIS OF GAMMA-CARBOXY GLUTAMIC ACID AND APPLICATIONS
TOWARDS THE SYNTHESIS OF CONANTOKINS

By

Camille Theonie Watson

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry - Doctor of Philosophy

2016

ABSTRACT

- I. CRYSTALLIZATION OF CRBP II MUTANTS AS A PROBE FOR UNDERSTANDING WAVELENGTH REGULATION
- II. REENGINEERING AND CRYSTALLIZATION OF CELLULAR RETINOL BINDING PROTEIN II (CRBP II) AS A FLUORESCENT TAG
- III. SYNTHESIS OF GAMMA-CARBOXY GLUTAMIC ACID AND APPLICATIONS TOWARDS THE SYNTHESIS OF CONANTOKINS

By

Camille Theonie Watson

- I. The mechanism by which humans are able to see colors has been an intriguing area of study over the past several decades. This interest comes from the fact that the different opsins bind to a single chromophore, retinal, as a protonated Schiff base and a wavelength range of 420 nm to 560 nm is observed. Different postulations have been put forward and mutagenic studies have been done on rhodopsin in an attempt to explain this phenomenon. Without crystallographic evidence, the results of these experiments proved to be inconclusive. Rhodopsin being a trans-membrane protein is very difficult to crystallize and give poor expression yields. As a result our lab has engineered a small cytosolic protein, Cellular Retinol Binding Protein II (CRBP II), as a rhodopsin mimic. Our studies, with the aid of high resolution crystal structures, have shown that chromophore conformation and complete burial of the chromophore in the binding pocket is essential for wavelength regulation.
- II. Fluorescent proteins tags are widely used in the field of biological sciences. Fluorescence based assays provide a means of probing protein localization, protein-protein interactions, protein expression and other biological processes without isolation from the cellular environment. Well renown Green Fluorescent Protein (GFP) and modified GFP's have been developed and are currently widely used as fluorescent tags. There are also non-protein fluorescent tags that are also currently being widely used; these include the SNAP and FLAsH tags. Even with the advent

of these fluorophoric protein tags there is still a lack of red and near-IR fluorescent proteins that are bright and photochemically stable. Since Cellular Retinol Binding Protein II (CRBP_{II}), has proven to be a robust system and binds all-*trans*-retinal in the nanomolar range, we decided to consider other molecules (fluorophores) that would bind as a protonated Schiff base in order to develop a new red fluorescent protein tag. Several chromophores have been used but to date two of them, Merocyanine and Julolidine, has proven to be suitable. We were able to get high resolution crystal structures of CRBP_{II} with Merocyanine, which provided a platform for analyzing the protein-fluorophore interactions to improve the quantum efficiency of the system.

III. Gamma-carboxyglutamic acid (Gla) is formed in humans as a post-translational modification of glutamic acid via the vitamin D carboxylase enzyme. This modification allows for tight binding of calcium ions, which allows for calcification of bones. This modified amino acid is also essential in the blood coagulation cascade processes. The Gla residues are also found in a family of peptides known as the conantokins, which are found in the venom of sea snails of the *Conus* genus. The metallo-zipper motif was first discovered in these conantokins. The 'metallo-zipper' motif provides a novel interface, which is dependent only on metal chelation. It was found that in the presence of metal ions such as Ca²⁺, Mg²⁺, Zn²⁺, and Cu²⁺, the structure of both Con-G and Con-T is changed from a linear structure to a helical structure. We have devised a new synthesis of gla and made attempts to synthesize different variants of the conantokins.

Dedicated to my loving husband and my family

TABLE OF CONTENTS

| | |
|--|------|
| LIST OF TABLES | viii |
| LIST OF FIGURES | x |
| KEY TO ABBREVIATIONS | xvii |
| CHAPTER I: INTRODUCTION | 1 |
| UNDERSTANDING THE MYSTERY OF COLOR VISION | 1 |
| I.1 WAVELENGTH REGULATION STUDIES ON MODEL COMPOUNDS | 12 |
| I.2 MUTAGENESIS STUDIES ON THE VISUAL RHODOPSINS | 21 |
| I.3 MUTAGENESIS STUDIES ON THE MICROBIAL RHODOPSINS | 27 |
| I.4 WAVELENGTH REGULATION DUE TO CONFORMATIONAL CHANGE | 31 |
| I.5 MODERN COMPUTATIONAL STUDIES ON WAVELENGTH REGULATION | 33 |
| I.6 STRATEGIES FOR SPECTRAL TUNING IN A RHODOPSIN MIMIC | 36 |
| I.7 AN OVERVIEW OF PROTEIN X-RAY CRYSTALLOGRAPHY | 41 |
| REFERENCES | 48 |
| CHAPTER II: ENGINEERING AND CRYSTALLIZATION OF THE 2ND GENERATION RHODOPSIN MIMIC, CRBP_{II} | 57 |
| II.1 AN INTRODUCTION: THE 1 ST GENERATION RHODOPSIN MIMIC | 57 |
| II.2 STUDY USING C15 AS THE CHROMOPHORE FOR WAVELENGTH TUNING | 62 |
| II.3 DEVELOPMENT OF THE 2 ND GENERATION RHODOPSIN MIMIC, CRBP _{II} | 65 |
| II.4 ENGINEERING OF THE 2 ND GENERATION RHODOPSIN MIMIC, CRBP _{II} | 67 |
| II.5 CRYSTALLIZATION OF THE 2 ND GENERATION RHODOPSIN MIMIC, CRBP _{II} | 69 |
| II.6 EFFECTS OF THE THR51 MUTATION ON WAVELENGTH REGULATION | 74 |
| II.7 EXPLORING THE EFFECTS OF MUTATIONS ON THE 6-s-cis AND 6-s-trans CONFORMATION OF THE CHROMOPHORE ALL-trans-RETINAL | 75 |
| II.8 MATERIALS AND METHODS | 80 |
| II.8.1 Mutagenesis | 80 |
| II.8.2 WT-CRBP _{II} gene sequence | 81 |
| II.8.3 WT-CRBP _{II} amino acid sequence | 81 |
| II.8.4 Primers used for mutagenesis | 81 |
| II.8.5 XL1-Blue competent cells preparation | 83 |
| II.8.6 Heat shock transformation | 85 |
| II.8.7 Protein expression and purification in pET-17b system | 85 |
| II.8.8 Crystallization Methods | 86 |

| | |
|--|-----|
| REFERENCES | 91 |
| CHAPTER III: DEVELOPMENT AND CRYSTALLIZATION OF CRBP_{II} DERIVATIVES AS FLUORESCENT PROTEIN TAGS | 94 |
| III.1 GREEN FLUORESCENT PROTEIN (GFP) AND ITS DERIVATIVES | 94 |
| III.2 OTHER FLUORESCENT TAGS | 99 |
| III.2.1 SNAP tag | 99 |
| III.2.2 FAsH tag | 101 |
| III.3 ENGINEERING CRBP _{II} INTO A FLUORESCENT PROTEIN TAG | 103 |
| III.4 CRYSTALLIZATION OF CRBP _{II} AS A FLUORESCENT PROTEIN TAG | 105 |
| III.4.1 CHARACTERIZATION OF DIFFERENT POLYENE FLUOROPHORE BOUND TO CRBP _{II} MUTANTS | 105 |
| III.4.2 EFFECT OF THE T51 V MUTATION ON THE QUANTUM EFFICIENCY OF CRBP _{II} MUTANTS BOUND TO MEROCYANINE POLYENE | 109 |
| III.4.3 EXPLORING THE EFFECTS OF DIFFERENT T51 MUTATIONS ON THE QUANTUM EFFICIENCY OF CRBP _{II} MUTANTS BOUND TO MEROCYANINE POLYENE | 114 |
| III.4.4 EFFECTS OF L117E MUTATION ON THE UTILITY OF CRBP _{II} MUTANTS BOUND TO MEROCYANINE POLYENE AS A FLUORESCENT TAG | 120 |
| III.4.5 CHARACTERIZATION OF JULOLIDINE BOUND TO CRBP _{II} MUTANTS AS FLUOROPHORIC pH SENSORS | 128 |
| III.5 MATERIALS AND METHODS | 130 |
| III.5.1 Mutagenesis | 130 |
| III.5.2 WT-CRBP _{II} gene sequence | 131 |
| III.5.3 WT-CRBP _{II} amino acid sequence | 131 |
| III.5.4 Primers used for mutagenesis | 131 |
| III.5.5 XL1-Blue competent cells preparation | 133 |
| III.5.6 Heat shock transformation | 134 |
| III.5.7 Protein expression and purification in pET-17b system | 135 |
| III.5.8 Crystallization Methods | 136 |
| III.5.9 Quantum efficiency determination of CRBP _{II} mutants bound to merocyanine polyene aldehyde | 141 |
| REFERENCES | 142 |
| CHAPTER IV: (I) TOTAL SYNTHESIS OF GAMMA-CARBOXYGLUTAMIC ACID AND APPLICATIONS TOWARDS PEPTIDE SYNTHESIS. (II) CRYSTALLIZATION OF NR2_a SUBUNIT WITH CONANTOKIN T (Con-T) | 147 |
| IV.1 AN INTRODUCTION | 147 |
| IV.2 PREVIOUSLY REPORTED SYNTHESSES OF L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla) | 149 |
| IV.3 SYNTHESIS OF L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla) VIA THE CHIRAL OXAZINONE TEMPLATE | 152 |
| IV.4 SYNTHESIS OF L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla) VIA THE GARNER'S ALDEHYDE TEMPLATE | 159 |
| IV.5 SYNTHESIS OF L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla) USING L-SERINE AS A CHIRAL TEMPLATE | 164 |

| | |
|---|-----|
| IV.6 SYNTHESIS OF CONOPEPTIDES USING L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla) | 168 |
| IV.7 EXPRESSION AND CRYSTALLIZATION OF NR2A WITH CONANTOKIN T (Con-T) | 177 |
| IV.8 MATERIALS AND METHODS | 183 |
| IV.8.1 Synthesis of L-gamma-carboxyglutamic acid (L-Gla) | 183 |
| IV.8.2 Peptide Synthesis | 198 |
| IV.8.2.1 Method used for the manual synthesis | 198 |
| IV.8.2.2 Method used to synthesize the peptides via Peptide Synthesizer | 199 |
| IV.8.3 HPLC run protocol | 203 |
| IV.8.4 Mass Spectra Data of the peptides synthesized | 204 |
| IV.8.5 Expression of NR2A with Conantokin T (Con-T) | 210 |
| IV.8.6 Cloning | 210 |
| IV.8.6.1 Plasmid DNA Double Digestion Protocol | 210 |
| IV.8.6.2 Plasmid DNA Ligation Protocol | 211 |
| IV.8.7 NR2A gene sequence | 212 |
| IV.8.8 NR2A amino acid sequence | 212 |
| IV.8.9 XL1-Blue competent cells preparation | 213 |
| IV.8.10 Heat shock transformation | 214 |
| IV.8.11 Mutagenesis | 214 |
| IV.8.12 Primers used for mutagenesis | 215 |
| IV.8.13 Protein expression and purification in pET-22b(+) and pET-28b systems | 215 |
| IV.8.14 Inclusion bodies purification | 216 |
| REFERENCES | 218 |

LIST OF TABLES

| | |
|--|-----|
| Table 1: Inductive effects in the absorption of retinal-PSB. | 17 |
| Table 2: Effects of different placement of positive charge on the absorption maxima of retinal PSB. | 18 |
| Table 3: Effect of polarizable solvents on retinal-PSB with different counterions. | 21 |
| Table 4: Sequence identity (below the diagonal) and sequence homology (above diagonal) between different visual rhodopsins. | 22 |
| Table 5: Table of mutants of Bovine Rhodopsin. | 23 |
| Table 6: Effect of solute anions on absorption maximum of rhodopsin E113 mutants. | 26 |
| Table 7: Characterization of different ring-locked analogues of retinal. | 32 |
| Table 8: CRABP II mutants with all- <i>trans</i> -retinal and all- <i>trans</i> C15-retinal. | 63 |
| Table 9: Effects of mutations on the different zones of the bound chromophore. | 73 |
| Table 10: Effects of T51 mutations. | 75 |
| Table 11: PCR reaction solutions. | 80 |
| Table 12: PCR temperature control cycles. | 80 |
| Table 13: Crystallization Conditions. | 87 |
| Table 14: Data Collection and refinement statistics. The ligand is all- <i>trans</i> -retinal in each case. | 89 |
| Table 15: Fluorescent characterization of retinal-PSB. | 106 |
| Table 16: Showing the distances of the aromatic residues in the mutants KL versus KL:T51V bound to merocyanine. | 112 |
| Table 17: Quantum efficiency of the different variants of the KL:T51X series. | 116 |
| Table 18: Quantum efficiency of the different variants of the KL:T51X:R58F series. | 116 |
| Table 19: PCR Reaction Solutions. | 130 |
| Table 20: PCR temperature control cycles. | 130 |

| | |
|---|-----|
| Table 21: Crystallization Conditions. | 137 |
| Table 22: Data Collection and refinement statistics. The ligand is merocyanine in each case. | 139 |
| Table 23: α_D Readings for VII' at different concentrations. Reported $\alpha_D = -50.3^\circ$. | 153 |
| Table 24: Showing the conditions used using different Lithium salts. | 167 |
| Table 25: Showing the conditions used using different cell lines, expression conditions, IPTG concentrations and expression times using pET22b(+). | 181 |
| Table 26: HPLC run for Con-G. | 203 |
| Table 27: HPLC run for Ala-Con-G | 203 |
| Table 28: Double Digestion Reaction solution. | 211 |
| Table 29: PCR Reaction solution. | 214 |
| Table 30: PCR temperature control cycles. | 215 |

LIST OF FIGURES

| | |
|--|----|
| Figure 1: Visible Spectrum. | 2 |
| Figure 2: The Human Eye. | 2 |
| Figure 3: a. Rod and Cone cells in the retina. b. Rhodopsins embedded in the disks in the photoreceptor cells. | 3 |
| Figure 4: Diagram of the photo transduction cascade. | 5 |
| Figure 5: a. Diagram showing the crystal structure of bovine rhodopsin (PDBcode:1F88). b. View of the binding site in bovine rhodopsin. | 7 |
| Figure 6: Schematic representation of the intermediates observed in the isomerization of 11- <i>cis</i> -retinal to all- <i>trans</i> -retinal. | 8 |
| Figure 7: a. Overlay of the crystal structure of bovine rhodopsin ground state (magenta) (PDB code: 1U19) and bathorhodopsin (green) (PDB code: 2G87). b. View of the chromophore binding site. | 9 |
| Figure 8: a. The overlay of the binding sites of bovine rhodopsin ground state (green) and meta II (magenta). b. Overlay of the crystal structures of bovine rhodopsin ground state (green) (PDB code: 1U19) and meta II (magenta) (PDB code: 3PQR). The short G α peptide is shown in purple. | 10 |
| Figure 9: 11- <i>cis</i> -retinal regeneration pathway in human retina. | 11 |
| Figure 10: Chemical conversion of all- <i>trans</i> -retinol (380 nm) into retinene cation (600nm). | 14 |
| Figure 11: Model compounds for studying the counterion effect on the absorbance maxima of retinal-PSB. | 16 |
| Figure 12: Model compounds used in the point charge theory studies. a. Opsin shift observed for bovine rhodopsin. b. Opsin shift observed for bacteriorhodopsin. | 19 |
| Figure 13: Crystal structure of bovine rhodopsin highlighting the residues that are important for blue shifting. | 24 |
| Figure 14: Diagram showing the six different rhodopsins found in a single Archaeon. | 28 |
| Figure 15: Binding cavity of bacteriorhodopsin. (PDB code: 1C3W). | 30 |

| | |
|---|----|
| Figure 16: Results of the computational analysis of 11- <i>cis</i> -retinal-PSB absorption with different torsion of the C6 – C7 bond highlighted. | 34 |
| Figure 17: Charge distribution of retinal-PSB. a. All- <i>trans</i> -retinal-PSB with numbered polyene carbons. Calculated charged state of the carbon atoms in the b. ground state and c. excited state. | 37 |
| Figure 18: a. The calculated bond distances for different retinal compounds. b. Resonance structures of cyanine dyes and retinal-PSB. | 39 |
| Figure 19: Electrostatic potential calculation. Electrostatic potential calculation (APBS suite) of blue, rod, green, and red opsin (calculations of blue, green and red are based on available homology models) projected on the Van der Waals surface of retinal (the electrostatic potential of the retinylidene chromophore was set to zero as to only illustrate the electrostatic contribution of the protein). The chromophore is divided into the three segments; the qualitative average score for each segment represents the overall electrostatic potentials that lead to wavelength regulation of each opsin. | 40 |
| Figure 20: The Electromagnetic Spectrum. | 42 |
| Figure 21: Diagram illustrating the relationship between the asymmetric unit, the unit cell and the entire crystal. | 43 |
| Figure 22: Triclinic unit cell, with edges, a , b , and c and angles α , β , and γ . | 43 |
| Figure 23: Diagrams illustrating a crystal diffracting the source beam into many discrete beams, each of which produces a distinct reflection on the film and the <i>hkl</i> designation of each reflection. | 44 |
| Figure 24: a. (210) planes in a two-dimensional section of the lattice. b. (230) planes in a two dimensional section of the lattice. | 45 |
| Figure 25: An illustration of the relationship between the diffraction data and the structure. | 47 |
| Figure 26: Crystal structure of WT-CRABPII with bound retinoic acid (PDB code: 1CBS). | 58 |
| Figure 27: Key interactions of the chromophore, all- <i>trans</i> -retinoic acid, in the binding pocket. | 59 |
| Figure 28: a. Hydrogen bonding interactions between R132K:Y134F-CRABPII and all- <i>trans</i> -retinal. b. The Bürgi-Dunitz trajectory. | 60 |

- Figure 29:** **a.** Structures of all-*trans*-retinoic acid and all-*trans*-retinal. **b.** Crystal structure of R132K:R111L:L121E. **c.** Overlay of the structures of WT-CRABP_{II}-retinoic acid (green, PDB code: 1CBS) and R132K:R111L:L121E – bound to the retinylidene PSB (cyan, PDB code: 2G7B). 61
- Figure 30:** Crystal structure of CRABP_{II} with all-*trans* C15-retinal analogue. 64
a. Crystal structure of R132K:R111L:L121E (KLE) with arginine 59 highlighted in green. **b.** The chemical structure of all-*trans*-retinal and all-*trans*-C15-retinal analogue. **c.** Crystal structure of R132K:R111L:L121E:R59W with all-*trans*-C15-retinal analogue with the Tryptophan at position 59 highlighted in yellow. **d.** The overlaid structures of R132K:R111L:L121E (KLE) with all-*trans*-retinal and R132K:R111L:L121E:R59W with all-*trans*-C15 retinal analogue.
- Figure 31:** Comparison of WT-CRBP_{II} and WT-CRABP_{II}. **a.** Overlaid crystal structures of WT-CRBP_{II} (purple, 2RCT) bound to all-*trans*-retinol (red) and WT-CRABP_{II} (green, 1CBS) bound to all-*trans*-retinoic acid (cyan). **b.** The space filling model of the region highlighted by the rectangle of CRABP_{II} (top) and CRBP_{II} (bottom) showing the ionone region. 66
- Figure 32:** Overlaid crystal structures of WT-CRBP_{II} with all-*trans*-retinol (green) and all-*trans*-retinal (magenta). The hydrogen bonding network is highlighted in both systems. 68
- Figure 33:** Proposed mechanism for proton transfer in single mutant Q108K. 69
- Figure 34:** Crystal structure of Q108K:K40L bound to all-*trans*-retinal with surrounding residues highlighted. 71
- Figure 35:** **a.** Protonated Schiff Base of all-*trans*-retinal. The chromophore is divided into three zones designated **I**, **II** and **III** to clarify the discussion of the different mutations. **b.** The possible stabilization of the protonated Schiff Base by Trp 106 and water mediated hydrogen bonding with Gln 4 is illustrated. 72
- Figure 36:** **a.** Chemical structure of the 6-*s*-cis and 6-*s*-trans conformations of retinal. 77
b. Overlay of the crystal structures of KLVCWLW-Q4H (green) and KLVCWLW-Q4H-L77T (magenta). Position 77 is highlighted in both cases. Distance between L77 and the gem dimethyl of the L77T structure is 1.8 Å which would result in unfavorable steric interactions.
- Figure 37:** Diagram illustrating the ‘bump and hole’ strategy that would favor the 6-*s*-cis and 6-*s*-trans conformations. 78

| | |
|--|-----|
| Figure 38: a. Crystal structure (left) and fluorophore (right) of EGFP (PDB code: 2Y0G). b. Molecular mechanism illustrating the maturation of the fluorophore in green fluorescent proteins. | 96 |
| Figure 39: Mechanism for the formation of the fluorophore and molecular structure of the different variants developed from DsRed. | 98 |
| Figure 40: Illustration of SNAP protein tag mechanisms. a. The mechanism of the DNA repair enzyme hAGT. b. General ways of labeling hAGT with a labeled benzyl guanine. c. Molecular Structure of WT-hAGT inhibitor CG. d. Molecular structure of labeled propargyl guanine. | 100 |
| Figure 41: Illustration of FAsH mechanism and three different organoarsenic compounds used for the different color FAsH system. | 102 |
| Figure 42: General structure of the fluorophores to be in the development of CRBP2 into a fluorescent tag. | 104 |
| Figure 43: Normalized UV-VIS and fluorescence spectra of retinal-PSB with <i>n</i> -butylamine and different CRBP2 mutants. | 106 |
| Figure 44: a. Structures of the different fluorophores. b. Resonance of the azulene polyene head group. c. Formation of a protonated Schiff base of azulene polyene with <i>n</i> -butylamine. | 108 |
| Figure 45: Characterization of merocyanine polyene. a. Formation of merocyanine polyene-PSB. b. Resonance structure of merocyanine polyene-PSB. c. UV-Vis spectra of merocyanine polyene, PSB and PSB formed with CRBP2 mutant KLVF (Q108K:K40L:T51V:R58F). d. Fluorescence spectra of merocyanine polyene. | 110 |
| Figure 46: Comparison of Q108K:K40L (KL) and Q108K:K40L:T51V (KL:T51V) bound to merocyanine polyene. a. Overlaid crystal structures of KL (green) and KL:T51V (blue). Space filling models of b. KL and c. KL:T51V showing the surrounding residues. | 112 |
| Figure 47: a. Water mediated hydrogen bonding network in Q108K:K40L (KL) between T51, T53 and main chain carbonyl of V62. b. Overlaid crystal structures of KL (green) and KL:T51V (blue). | 113 |
| Figure 48: a. Crystal structure of KL:T51V. b. Crystal Structure of the tetra-mutant KL:T51V:T53V. c. Overlaid crystal structures of KL:T51V (blue) and KL:T51V:T53V (magenta). | 115 |
| Figure 49: Overlaid structures of KL:T51V (green) and KL:T51V:R58F:L117E (cyan) highlighting the residues in the binding pocket. | 118 |

| | |
|---|-----|
| Figure 50: Overlaid structures of KL:T51V:R58F:L117E (blue) and KL:T51I:R58F:L117E (green) highlighting the residues in the binding pocket. | 119 |
| Figure 51: Overlaid structures of KL:T51A:R58F (magenta) and KL (green) highlighting the residues in the binding pocket. | 120 |
| Figure 52: Constructs of pEGFP-KLVF(CRBPII)-RB and pEGFP-RB. | 122 |
| Figure 53: Confocal pictures of a. pEGFP-KLVF(CRBPII)-RB construct. b. pEGFP-RB construct. EGFP was excited at 488 nm and KLVF-merocyanine polyene was excited at 594 nm. | 123 |
| Figure 54: a. Crystal structure of KL:T51V with residues L117 and T51V highlighted. b. Proposal for the mechanism of activation of the aldehyde by the introduction of an acidic residue at position 51 or 117. | 124 |
| Figure 55: Energy minimized model of KL with merocyanine aldehyde (cyan). The interactions between residues Q108K and T51 are highlighted. | 125 |
| Figure 56: UV-VIS spectra showing the effect of the L117E mutation on the kinetics of PSB formation with merocyanine polyene. Mutants are KL:T51V:R58Y:A33W and KL:T51V:R58W:A33W:L117E. | 126 |
| Figure 57: Overlaid structures of KL:T51V:R58F:L117E bound to merocyanine (cyan) and KLVCLW:Q4R (PDB ID: 4EEJ) bound to all- <i>trans</i> -retinal (green) highlighting the residues L117 and L117E. The distances between the polyene and the oxygen of L117E is 3.0 Å. The distance between the C20 methyl group and the oxygen of L117E is 3.1 Å. | 127 |
| Figure 58: Proposed mechanism for activation of merocyanine polyene via a proton transfer in the presence of L117E mutation. | 127 |
| Figure 59: UV-Vis spectra showing the pH titration of the mutant Q108K:K40L:T51V:R58F:L117E:Q4F (KLVF-L117E-Q4F). | 129 |
| Figure 60: Proposed mechanism for the carboxylation of glutamic acid. | 148 |
| Figure 61: Earliest reported synthesis of DL gamma-carboxyglutamic acid (Gla). | 149 |
| Figure 62: a. Synthesis of L-gamma-carboxyglutamic acid using the oxazinone Template. b. Synthesis of compound XI . | 150 |
| Figure 63: a. Synthesis of L-gamma-carboxyglutamic acid using the Garner's aldehyde template. b. Synthesis of Garner's aldehyde, compound XII . | 151 |
| Figure 64: ECCD of VII'. | 154 |

| | |
|---|-----|
| Figure 65: Synthesis of Garner's Aldehyde XII . | 159 |
| Figure 66: Synthesis of L-gammacarboxyglutamic acid (Gla) from XII . | 161 |
| Figure 67: Conditions used in an attempt to deprotect the acetonide. | 162 |
| Figure 68: Conditions used that successfully deprotected the acetonide. | 163 |
| Figure 69: Steps leading to the final product compound XVI . | 164 |
| Figure 70: Synthesis of L-Gla from L-serine as chiral template. | 165 |
| Figure 71: Mechanism of hydrolysis of the methyl ester. | 166 |
| Figure 72: Amino acid sequence of a few of the known Conantokins. | 170 |
| Figure 73: Diagram of the Synapse with the N-methyl-D-Aspartate Receptor (NMDAR). | 172 |
| Figure 74: Sequence of four modified members of the Conantokin family that were synthesized and the original sequence of Con-G. | 174 |
| Figure 75: Sequence of the shortened modified members of the Conantokin family that were synthesized. | 176 |
| Figure 76: Diagram of the N-methyl-D-Aspartate Receptor (NMDAR). TM (Transmembrane portion of the receptor); 1 (subunit 1, NR1); 2 (subunit2, NR2). | 178 |
| Figure 77: Diagram of the construct of the N-methyl-D-Aspartate Receptor (NMDAR) that was crystallized. Diagram shows the amine terminal domain (ATD), carboxyl terminal domain (CTD) and the transmembrane domain (TM). | 179 |
| Figure 78: pET22b(+) cloning /expression region. | 180 |
| Figure 79: Mass Spectra of the Con-G-tetramer (sequence IR γ K). Major peak corresponds to the [M + H] ⁺ ; Observed Mass = 589.3439; Expected Mass = 588.3231. | 204 |
| Figure 80: Mass Spectra of Con-G-13 mer (sequence $\gamma\gamma$ LQ γ NQ γ LIR γ K). Observed Mass = 939.4534 which corresponds to [M+2H] ²⁺ ; Expected Mass = 1876.7810. | 205 |
| Figure 81: Mass Spectra of Ala Con-G-13 mer (sequence $\gamma\gamma$ AA γ AA γ AAR γ K). Observed mass = 746.8721 which corresponds to the [M+2H] ²⁺ ; Expected Mass = 1491.5961. | 206 |

- Figure 82:** Mass Spectra of Con-G-8mer with Glutamate substituted for L-Gla (sequence NQELIREK). Observed mass $[M + H]^+ = 1029.50$;
Expected Mass = 1028.550. 207
- Figure 83:** Mass Spectra of Con-G-8mer with Proline substituted for L-Gla (sequence NQPLIRPK). Observed Mass $[M + H]^+ = 965.58$; Expected Mass = 964.571. 208
- Figure 84:** Mass Spectra of Ala-Con-G-8mer with Glutamate substituted for L-Gla (sequence AAEEAAREK). Observed Mass $[M + H]^+ = 845.42$;
Expected Mass = 844.429. 209
- Figure 85:** Mass Spectra of Ala-Con-G-8mer with Proline substituted for L-Gla (sequence AAPAARPK). Observed mass $[M + H]^+ = 781.420$;
Expected Mass = 780.450. 210

KEY TO ABBREVIATIONS

| | |
|-----------------------|------------------------|
| Å | Angstrom |
| ϵ | Extinction Coefficient |
| λ_{\max} | Maximum Wavelength |
| λ_{Em} | Emission Wavelength |
| λ_{Ex} | Excitation Wavelength |
| nm | Nanometer |
| M | Molar |
| μM | Micromolar |

Amino Acids

| | |
|--------|---------------|
| Ala, A | Alanine |
| Arg, R | Arginine |
| Asn, N | Asparagine |
| Asp, D | Aspartate |
| Cys, C | Cysteine |
| Gln, Q | Glutamine |
| Glu, E | Glutamic Acid |
| His, H | Histidine |
| Leu, L | Leucine |
| Lys, K | Lysine |
| Phe, F | Phenylalanine |

| | |
|---------|---|
| Pro, P | Proline |
| Thr, T | Threonine |
| Trp, W | Tryptophan |
| Tyr, Y | Tyrosine |
| Val, V | Valine |
| Amp | Ampicillin |
| bR | Bacteriorhopsin |
| Clm | Chloramphenicol |
| CRABPII | Cellular Retinoic Acid Binding Protein II |
| CRBPII | Cellular Retinol Binding Protein II |
| Con T | Conantokin T |
| Con G | Conantokin G |
| Da | Dalton |
| DMF | Dimethylformamide |
| DMSO | Dimethylsulfoxide |
| DNA | Deoxyribonucleic Acid |
| Equiv | Equivalent |
| EDTA | Ethylenediaminetetraacetic acid |
| FlAsh | Fluorescein Arsenic Hairpin Binder |
| FPLC | Fast Protein Liquid Chromatography |
| FQ | Fast Q, quaternary ammonium resin |
| GPCR | G-Protein Coupled Receptor |
| h | Hour |

| | |
|--------------------|---|
| HPLC | High Performance Liquid Chromatography |
| iLBP | Intracellular Lipid Binding Protein |
| IPTG | Isopropylthiogalactoside |
| Kan | Kanamycin |
| K _d | Dissociation constant |
| LB | Luria Broth |
| L | Liter |
| MALDI-TOF | Matrix Assisted Laser Desorption Ionization- Time of Flight |
| Mero | Merocyanine |
| mL | Milliliter |
| mM | Millimolar |
| mW | Milliwatt |
| μL | Microliter |
| nBuNH ₂ | n-Butylamine |
| nm | Nanometer |
| nM | Nanomolar |
| NMR | Nuclear Magnetic Resonance |
| Ni-NTA | Nickel nitrilotriacetic acid |
| NiSO ₄ | Nickel Sulfate |
| PAGE | Polyacrylamide Gel Electrophoresis |
| PBS | Phosphate saline buffer |
| PCR | Polymerase Chain Reaction |

| | |
|----------------------------------|----------------------------------|
| PSB | Protonated Schiff Base |
| QY | Quantum Yield |
| RA | Retinoic Acid |
| ReAsh | Resorufin Arsenic Hairpin Binder |
| Rt | Retinal |
| Rt, RT | Room Temperature |
| RPM | Revolutions per minute |
| SB | Schiff Base |
| SDS | Sodium Dodecyl Sulfate |
| t | Time |
| t _{1/2} | Maturation Half Time |
| THF | Tetrahydrofuran |
| UV | Ultraviolet Light |
| Vis | Visible Light |
| CRBP_{II} Mutants | |
| WT | Wild Type Protein |
| KL | Q108K:K40L |
| KL _V | Q108K:K40L:T51V |
| KL _{VF} | Q108K:K40L:T51V:R58F |
| KL _{VFE} | Q108K:K40L:T51V:R58F;L117E |
| NaCl | Sodium Chloride |
| Tris | Tris(hydroxymethyl)aminomethane |
| HCl | Hydrochloric acid |

| | |
|----------------------------------|---|
| NaH ₂ PO ₄ | Sodium Monophosphate |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| CHAPS | 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate |

CHAPTER I: INTRODUCTION

UNDERSTANDING THE MYSTERY OF COLOR VISION

How are we able to see and perceive different colors? The answer to such a question is not a simple one. If one should consider what is classified as ‘normal’ routine activities, these processes actually involve quite elaborate mechanisms. The visual process is one such process that is considered to be one of the most efficient mechanisms in humans.

The fascination with the concept of vision dates back to as early as Ancient Greece. Greek philosophers, Euclid (ca. 300 BC), Aristotle (330 BC), Democritus (ca. 350 BC), and Plato (ca. 450 BC) have all sought to gain insight in an attempt to explain this phenomenon.¹ Of all the theories put forward by the different philosophers, Aristotle’s theory is the most accepted to this day. Unlike his predecessors Aristotle relied on his observations and came to the conclusion that (sun) light is reflected by an object and is somehow then transmitted by a medium into the eye. This conclusion was in contradiction to the theories proposed by his predecessors who suggested that light is emanated from the eye to the object; this phenomenon is referred to as “*corpuscular emanation*”.² Over time this theory was improved and revisited by several well renowned scientists from Johannes Kepler who gave the first complete description of the optics of the eye to Isaac Newton to provided insight of the laws that govern light.^{3 4} The Planck- Einstein equation: $E = h\nu = h c/\lambda$, tells us that light of different colors corresponds to different energies. This is depicted in **Figure 1** which shows the different wavelengths corresponding to the different colors in the visible spectrum.

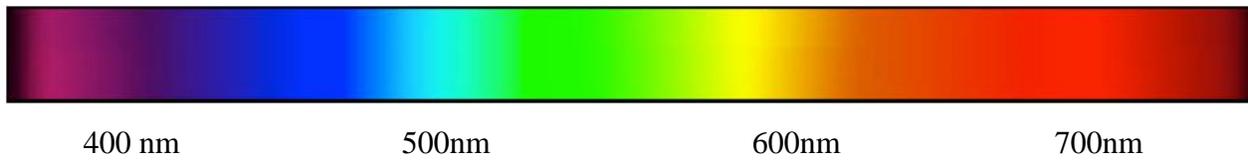


Figure 1: Visible Spectrum.⁶

A basic overview of the visual process is described as follows. Light energy enters the eye through the pupil and is focused by the lens on the retina **Figure 2**. The light is then absorbed by the photoreceptors located in the retina, which then leads to a complex cascade of events. The photons absorbed are converted into electrical signals that are eventually transmitted through the optic nerve to different areas of the brain.

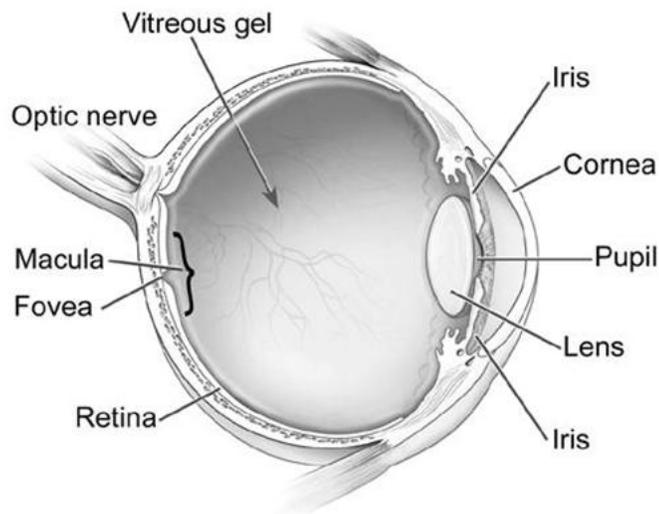


Figure 2: The Human Eye.⁷

In 1878, the photosensitive protein in rod cells, which is now known as rhodopsin, was isolated by F.W. Kuhne.⁵ Rhodopsin consists of two entities: the opsin, which is the 7- α -helix trans-membrane protein and the chromophore, 11-*cis*-retinal.⁸ The chromophore is covalently

bound to the opsin via a lysine residue located in the binding pocket as a protonated Schiff base.⁹

There are four different types of opsins present in the human eye. They are: the red, blue and green opsins found in the cone cells and are concentrated in the fovea of the human eye and the rod opsin, which is found in the rod cells of the eye and is responsible for dim light vision.¹⁰

There are about 100 million rod cells and 6 million cone cells in the human retina.¹¹ It can be concluded therefore that rod cells are more sensitive to light than the cone cells.

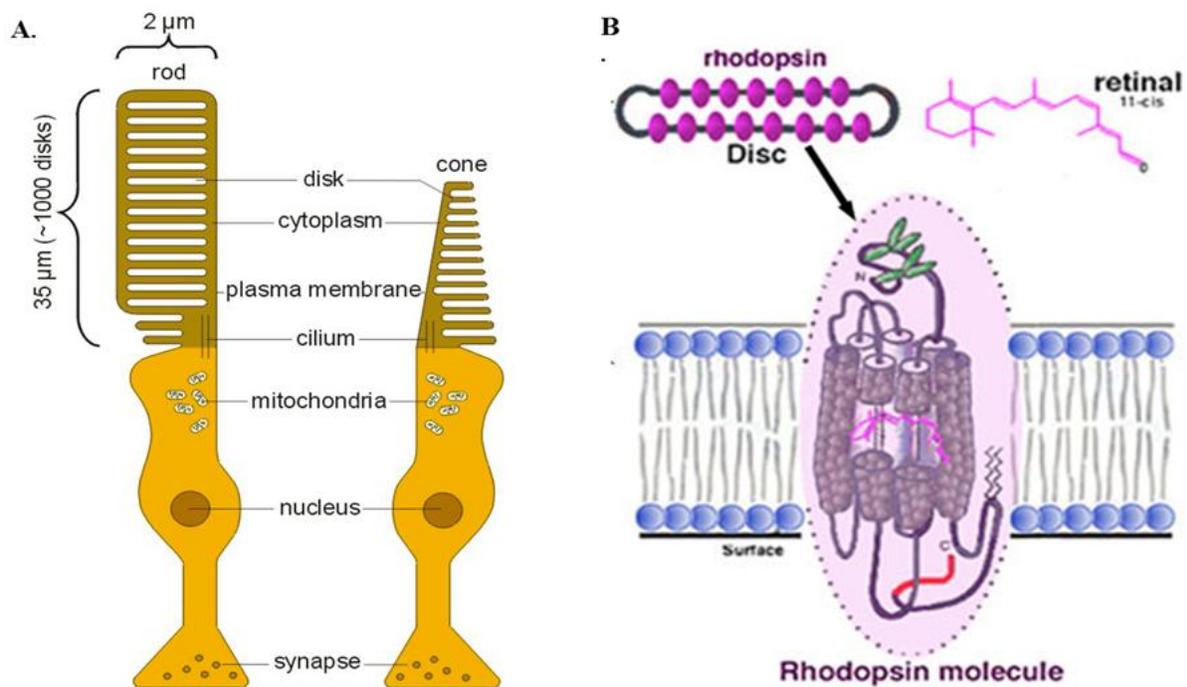


Figure 3: a. Rod and Cone cells in the retina. **b.** Rhodopsins embedded in the disks in the photoreceptor cells.^{12 13}

As shown in **Figure 3a**, the photoreceptor cells contain an inner and outer segment.¹⁴ The outer segment consists of several hundred phospholipid bilayer discs that are arranged in a parallel manner. Rhodopsin is located in these discs and accounts for ~50% of the disc's area. The orientation of the rhodopsin in these discs is shown in **figure 3b**. Each rod cell may contain

up to 1700 discs and each disc may contain up to 1.5 million rhodopsin molecules, this further explains the sensitivity of rod cells.¹⁵ The inner segment consists of the nucleus, mitochondria and the synaptic end.

Extensive mechanistic research into how the light activation of rhodopsin lead to color vision has been done. The human vision process begins with the covalent binding of 11-*cis*-retinal to the opsin forming an 11-*cis*-retinylidene.¹⁶ Upon light absorption, 11-*cis*-retinylidene isomerizes to all-*trans*-retinylidene. This isomerization results in a drastic conformational change in rhodopsin, which is the activated state.¹⁷ This activated form of rhodopsin has a high affinity for the protein transducin (**Figure 4**), a member of the G-protein family.

G proteins, also known as guanine nucleotide-binding proteins, are a family of proteins involved in transmitting chemical signals originating from outside a cell into the inside of the cell. G proteins function as molecular switches.¹⁸ Their activity is regulated by factors that control their ability to bind to and hydrolyze guanosine triphosphate (GTP) to guanosine diphosphate (GDP). When they bind GTP, they are 'on', and, when they bind GDP, they are 'off'. G proteins belong to the larger group of enzymes called GTPases. There are two classes of G proteins. The first functions as monomeric small GTPases while the second form functions as heterotrimeric G protein complexes. The latter classes of complexes are made up of alpha (α), beta (β) and gamma (γ) subunits. The beta and gamma subunits can form a stable dimeric complex referred to as the beta-gamma complex. G proteins located within the cell are activated by G protein-coupled receptors (GPCRs) that span the cell membrane.

Signaling molecules bind to the domain of the GPCR located outside the cell. An intracellular

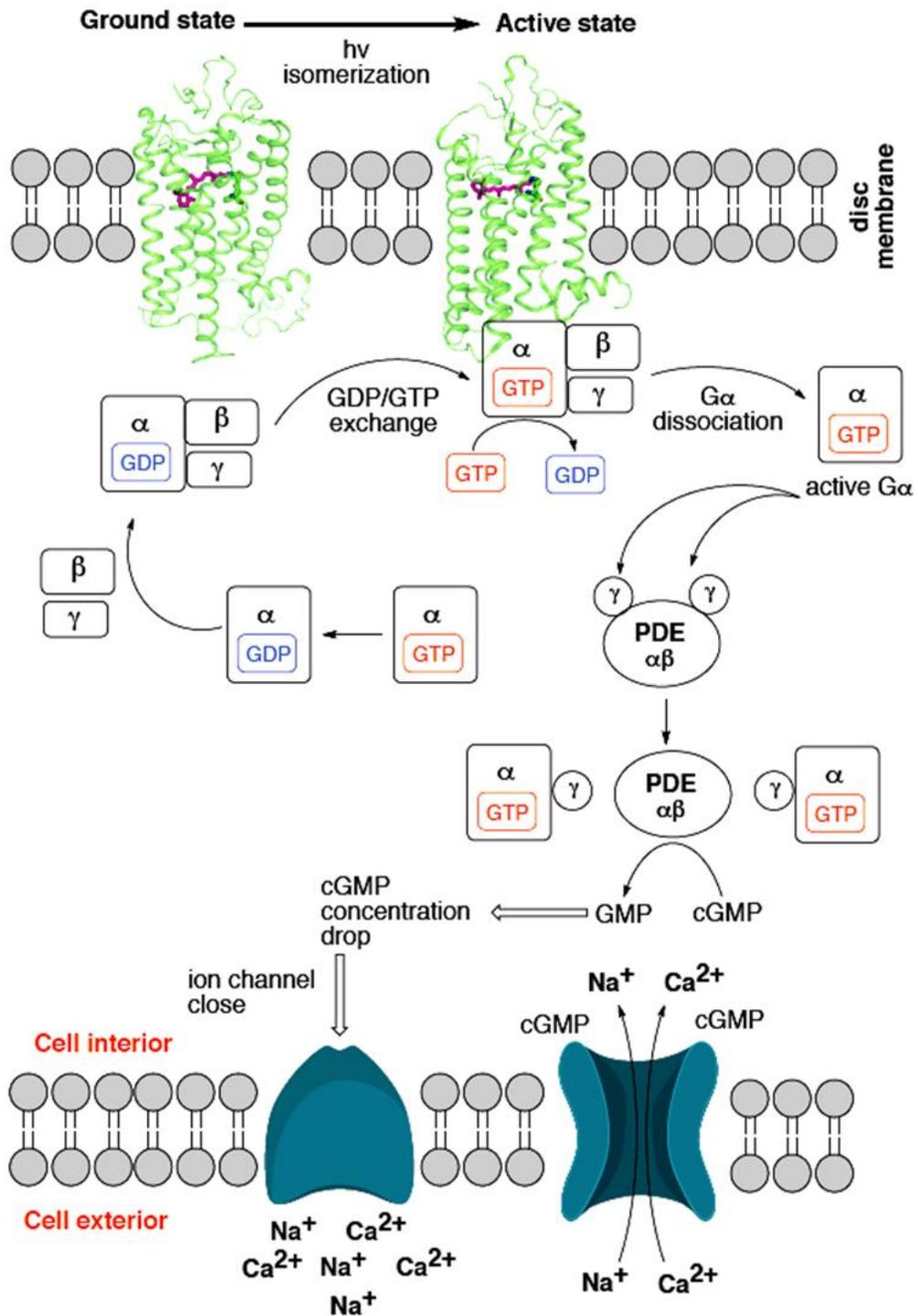


Figure 4: Diagram of the photo transduction cascade.

GPCR domain in turn activates a G protein. The G protein activates a cascade of further signaling events that finally results in a change in cell function. G protein-coupled receptor and G proteins working together transmit signals from many hormones, neurotransmitters, and other signaling factors. G proteins regulate metabolic enzymes, ion channels, transporter, and other parts of the cell machinery, controlling transcription, motility, contractility, and secretion, which in turn regulate diverse systemic functions such as embryonic development, learning and memory, and homeostasis.

When the activated rhodopsin binds to transducin, the alpha subunit of transducin exchanges the GDP for GTP (**Figure 4**). This results in the dissociation of alpha subunit from the beta-gamma subunits. The alpha subunit is now free and available to bind two gamma subunits of cyclic phosphodiesterase (PDE). This association results in the activation of the PDE to hydrolyze the phosphodiester bond of cyclic guanosine monophosphate (cGMP). The hydrolysis results in a decrease in the concentration of cGMP that closes the cGMP gated channel. This causes a hyperpolarization in the photoreceptor cells resulting in the creation of a signal for the associated neuron across the synaptic gap.¹⁹

Rhodopsin is a member of the G protein coupled receptor (GPCR) family. It has been a protein of interest for a long time. Having a crystal structure of rhodopsin would provide a platform for understanding the mechanisms of color vision and GPCR action.²⁰ The first crystal structure reported was that of bovine rhodopsin (**Figure 5**).²¹ This data prove to be quite insightful as it shows the chromophore is oriented in the plane of the membrane and perpendicular to the barrel to the 7-alpha-helices. Several key interactions were also observed.

These are, Lys 296, the residue with which the chromophore forms the protonated Schiff base; and Glu 113, which stabilizes the protonated Schiff base through a salt bridge.

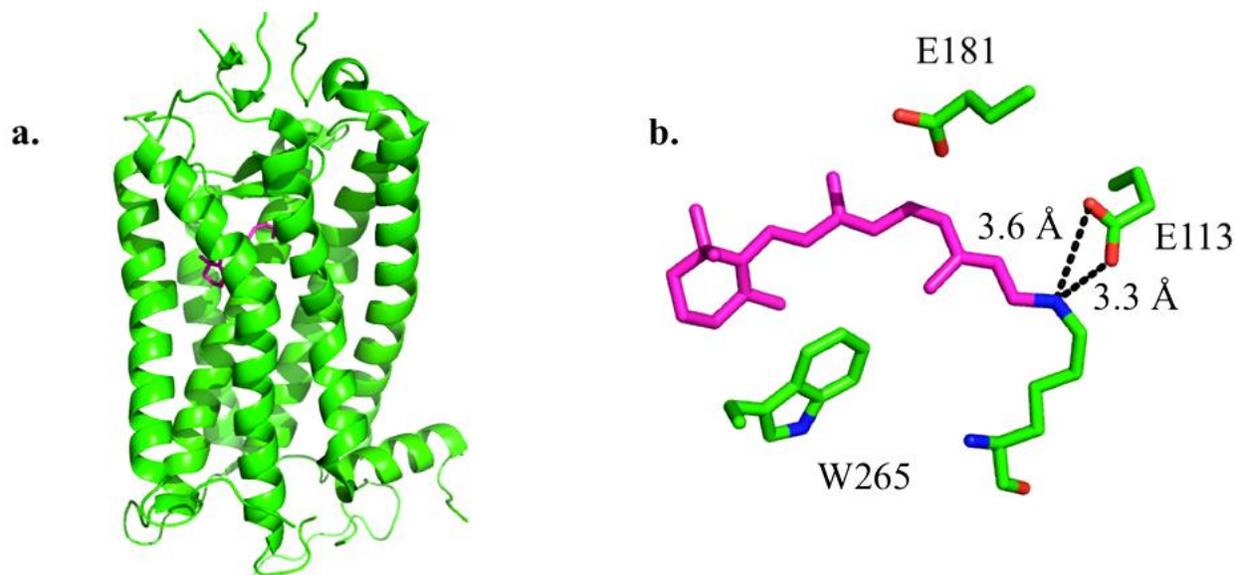


Figure 5: **a.** Diagram showing the crystal structure of bovine rhodopsin (PDB code: 1F88). **b.** View of the binding site in bovine rhodopsin.

This activation-deactivation process of rhodopsin involves several thermal/dark steps characterized by intermediates identified at low temperatures **Figure 6.**²²⁻²⁷ In the dark state, 11-*cis*-retinal binds rhodopsin, as illustrated in the crystal structure above.^{21 28-30} Within femtoseconds after irradiation, photorhodopsin is formed as a highly distorted all-*trans*-retinylidene,³¹ which then relaxes to the first isolatable intermediate, bathorhodopsin.^{32 33} The formation of lumirhodopsin immediately follows within nanoseconds, followed by the formation of metarhodopsin I, which occurs in microseconds and then metarhodopsin II, which occurs in milliseconds. The metarhodopsin II is the activated form of rhodopsin that triggers the photo transduction cascade. It was found that the Schiff base was unprotonated in the metarhodopsin II state. Following the metarhodopsin II activation, the Schiff base is reprotonated to give the

metarhodopsin III state. This finally led to the dissociation of the all-*trans*-retinal from rhodopsin via hydrolysis.

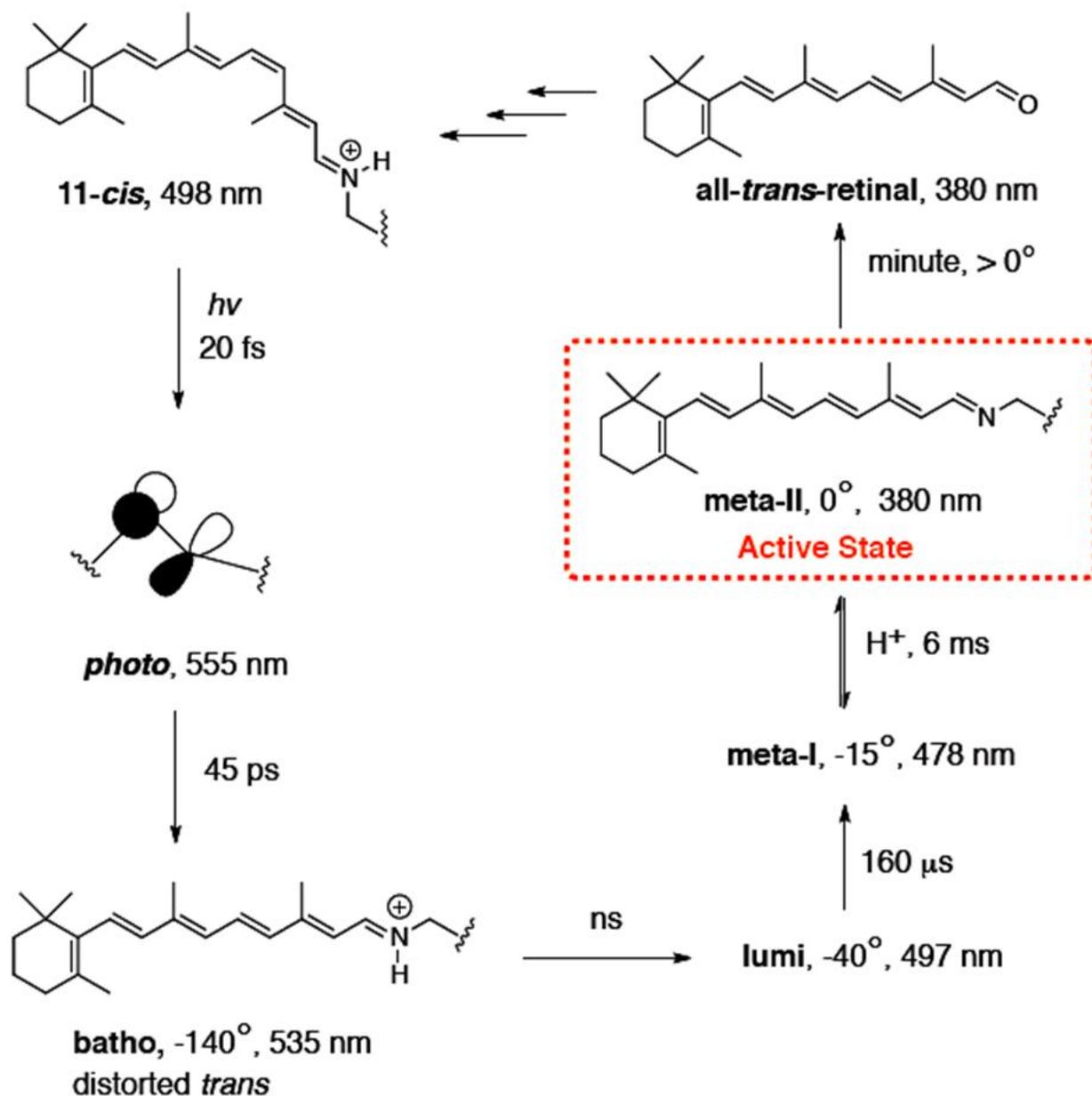


Figure 6: Schematic representation of the intermediates observed in the isomerization of 11-*cis*-retinal to all-*trans*-retinal.²²⁻²⁷

The efficiency of the *cis-trans* isomerization is due to the fact that 11-*cis*-retinal is very distorted. This distortion is as a result of the unfavorable steric interactions between the C13

methyl group and the C10 hydrogen of retinal. The crystal structure of bathorhodopsin was obtained by exposing the crystals of bovine rhodopsin to 488 nm light at low temperature (95 K).³⁴ The overlay of the crystal structures of both the ground state rhodopsin and bathorhodopsin shows that there is minimal protein movement. However, the isomerization of 11-*cis*-retinal results in the formation of a distorted *trans*-bond. Most of the energy absorbed from the light is trapped on this intermediate and serves as the energy required for the subsequent conformational changes of the protein. As shown in **Figure 6**, the λ_{\max} is 535nm, which is red shifted compared to the ground state with a λ_{\max} of 498nm. This bathochromic shift is a result of the formation of this higher energy intermediate, which results in an increase in the ground state energy.

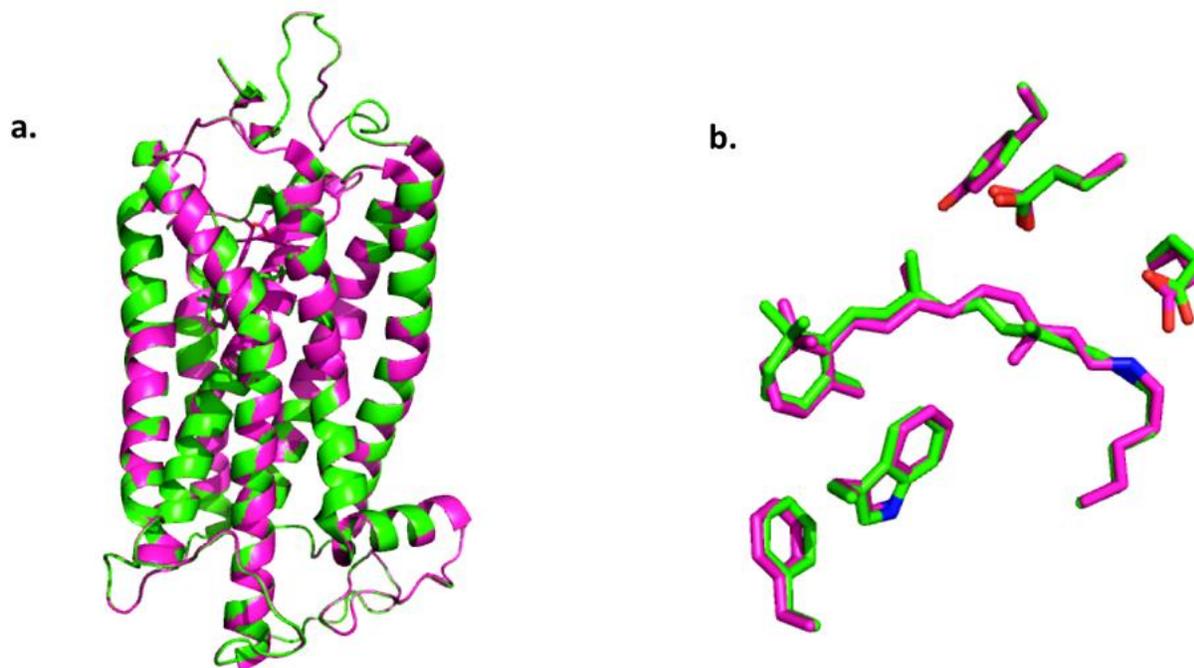


Figure 7: **a.** Overlay of the crystal structure of bovine rhodopsin ground state (magenta) (PDB code: 1U19) and bathorhodopsin (green) (PDB code: 2G87). **b.** View of the chromophore binding site.

This increase in the ground state energy lowers the energy gap between the ground state and the excited state, resulting in a red shift.

The meta II intermediate is recognized as the active state for the G-protein activation. Recently, crystal structures of meta II rhodopsin bound to all-*trans*-retinal and bound to a short C-terminal peptide from the G α subunit of transducin, have been published.^{35 36} These crystal structures show the interactions between the activated form of rhodopsin and the G α subunit of transducin.

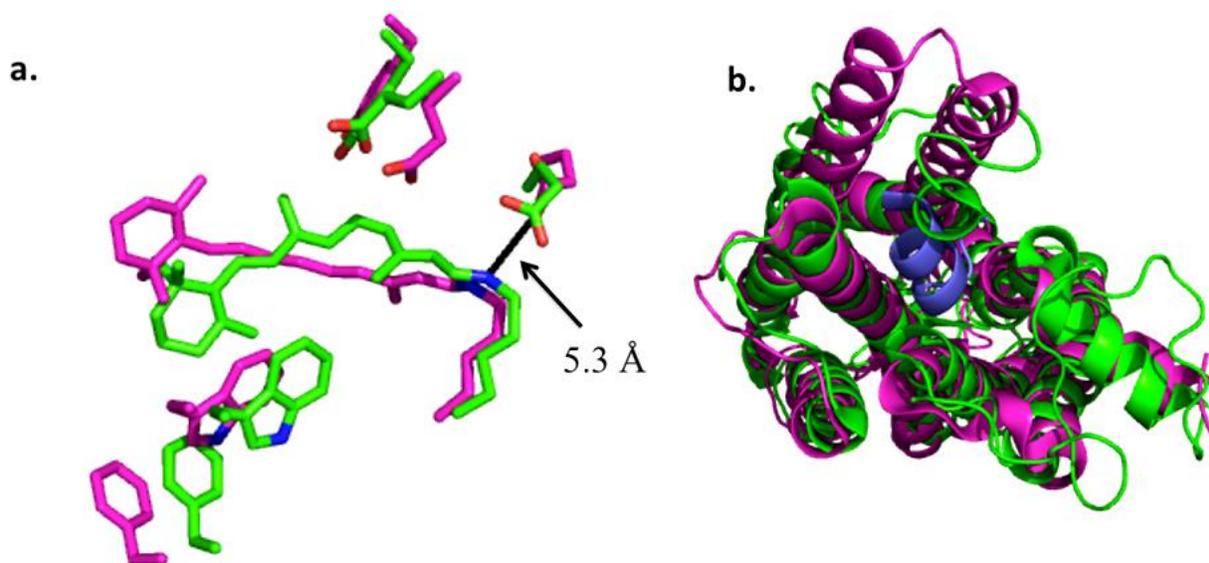


Figure 8: **a.** The overlay of the binding sites of bovine rhodopsin ground state (green) and meta II (magenta). **b.** Overlay of the crystal structures of bovine rhodopsin ground state (green) (PDB code: 1U19) and meta II (magenta) (PDB code: 3PQR). The short G α peptide is shown in purple.

As shown in figure I-8a the conformation of the chromophore is very different in the two states.

This conformational change in the chromophore causes the transmembrane helix 3 (TM3) to move out to allow for the binding of the G α peptide (**Figure 8b**). This phenomenon also explains why the Schiff base is deprotonated in the meta II state. The movement of TM3 results in the movement of the stabilizing counterion further away (5.3 Å) from the Schiff base and also results

in the binding pocket being more exposed, thus decreasing the pK_a of the retinal PSB in the meta II state.

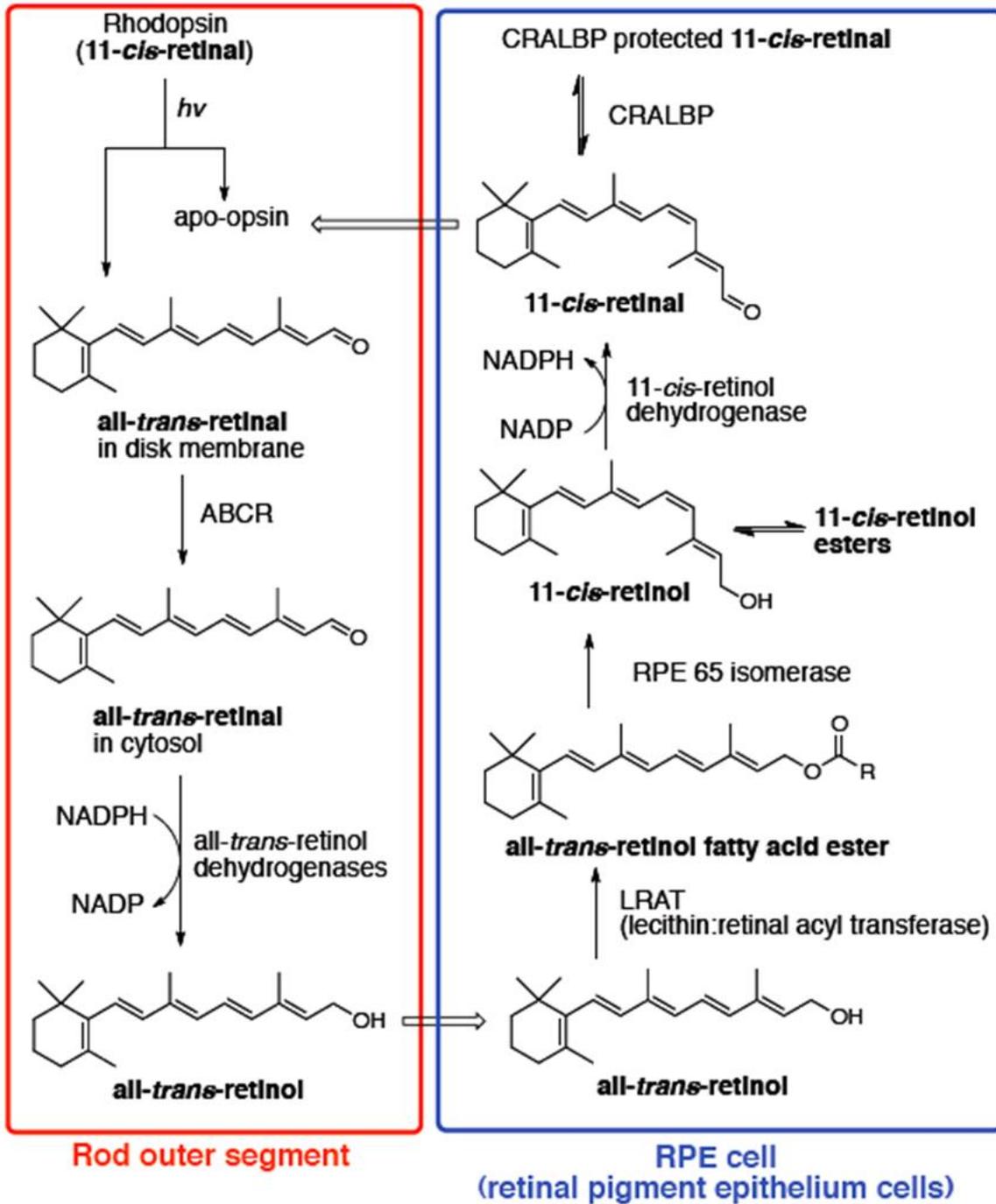


Figure 9: 11-cis-retinal regeneration pathway in human retina.³⁹

The lifetime of the meta II state is longer in rod rhodopsin when compared to the other rhodopsins, and therefore can activate more G-proteins.³⁷ This further explains the sensitivity of rod rhodopsin to dim light.

As mentioned before, after the meta II state the Schiff base is reprotonated and the all-*trans*-retinal is hydrolyzed due to instability of the all-*trans*-retinal complex. The 11-*cis*-retinal has to be regenerated from the all-*trans*-retinal for the visual cycle to continue. The regeneration of the 11-*cis*-retinal occurs in the epithelial cells of the retinal pigment.³⁸ **Figure 9** illustrates the regeneration of the photopigment. After the all-*trans*-retinal dissociates from the rhodopsin, it is transported to the cytosol by an ATP-binding cassette transporter (ABCR). Once in the cytosol, the all-*trans*-retinal is reduced to all-*trans*-retinol by all-*trans*-retinol dehydrogenases after which it is transported to the epithelial cells in the retinal pigment. In the epithelial cells, all-*trans*-retinol is esterified with palmitoyl acid by lecithin: retinal acyl transferase (LRAT). The all-*trans*-retinol ester is then isomerized and hydrolysed at the same time to 11-*cis*-retinol by RPE65. RPE65 is a protein that is found in the epithelial cells of the retinal pigment.³⁹ The energy required for the isomerization from *trans* to *cis* comes from the hydrolysis of the ester. The 11-*cis*-retinol is then oxidized to the corresponding aldehyde and is then taken up by the 11-*cis*-retinal binding protein, cellular retinal binding protein (CRALBP). 11-*cis*-retinal is then transported back to the outer segment of the rod cells and the vision cycle starts over again.

I.1 WAVELENGTH REGULATION STUDIES ON MODEL COMPOUNDS

The discovery of the red pigment by Kuhne (1877) to the discovery of trichromacy (1960s) took several decades. It was the development of mass spectrometry in the 1960s that allowed the

recording of the UV-Vis spectra of single cone and rod cells. The three types of cone rhodopsin are; blue rhodopsin (λ_{\max} ~420 nm), green rhodopsin (λ_{\max} ~530 nm) and red rhodopsin (λ_{\max} ~ 560 nm).^{40 41} The three rhodopsins span ~140 nm of the visual spectrum and collectively represent the foundation of human color vision. Different vertebrates have different photoreceptors as a result of their adaptation to their environment for survival.

The cone receptors in vertebrates use a single chromophore 11-*cis*-retinal but other chromophores like, 11-*cis*-dehydroretinal (has one more double bond than 11-*cis*-retinal), is found in some invertebrate species.^{42 43} The rhodopsins that use 11-*cis*-dehydroretinal as chromophore are called porphyropsin.

Deep water fish (deep sea elasmobranchs) have evolved to utilize more red shifted cone photoreceptors; for some these fishes the λ_{\max} reaches as far as 630 nm. This is a result of the fact that light with longer wavelength penetrates water more easily than light with shorter wavelength. In these species the ratio of the photoreceptors may change with the season but the total number of photoreceptors is constant.⁴⁴ Some species have evolved to change the chromophore used based on the environmental or physical changes. One such example is the amphibian species. When frogs undergo metamorphosis, the photoreceptor changes from porphyropsin (520 nm) to rhodopsin (500 nm). This enables them to adapt to changes in the environment.⁴⁵ The fact that these rhodopsins span ~140 nm of the visual spectrum by binding to a single chromophore, 11-*cis*-retinal, is still being meticulously investigated. 11-*cis*-retinal by itself in aqueous solution absorbs at 380 nm and when it forms a Schiff base (SB) with n-butylamine, it blue shifts to 360 nm. When the Schiff base gets protonated, the absorbance red

shifts 80 nm to 440 nm. Since all the opsins are more red shifted than the 380 nm observed for the SB, that would therefore suggest that the SB formed in the opsins with 11-*cis*-retinal are protonated.

In comparison to rod, red and green rhodopsins the protonated Schiff base (PSB) of retinal in solution is more blue shifted. This shift is referred to as the ‘opsin shift’.⁴⁶ Different model studies have been done and several hypotheses have been put forward to explain this ‘opsin shift’.

Based on what is known about polyene systems, the more delocalized π electrons present the more conjugated the system, which results in a red shift. In the case of the rhodopsins, it

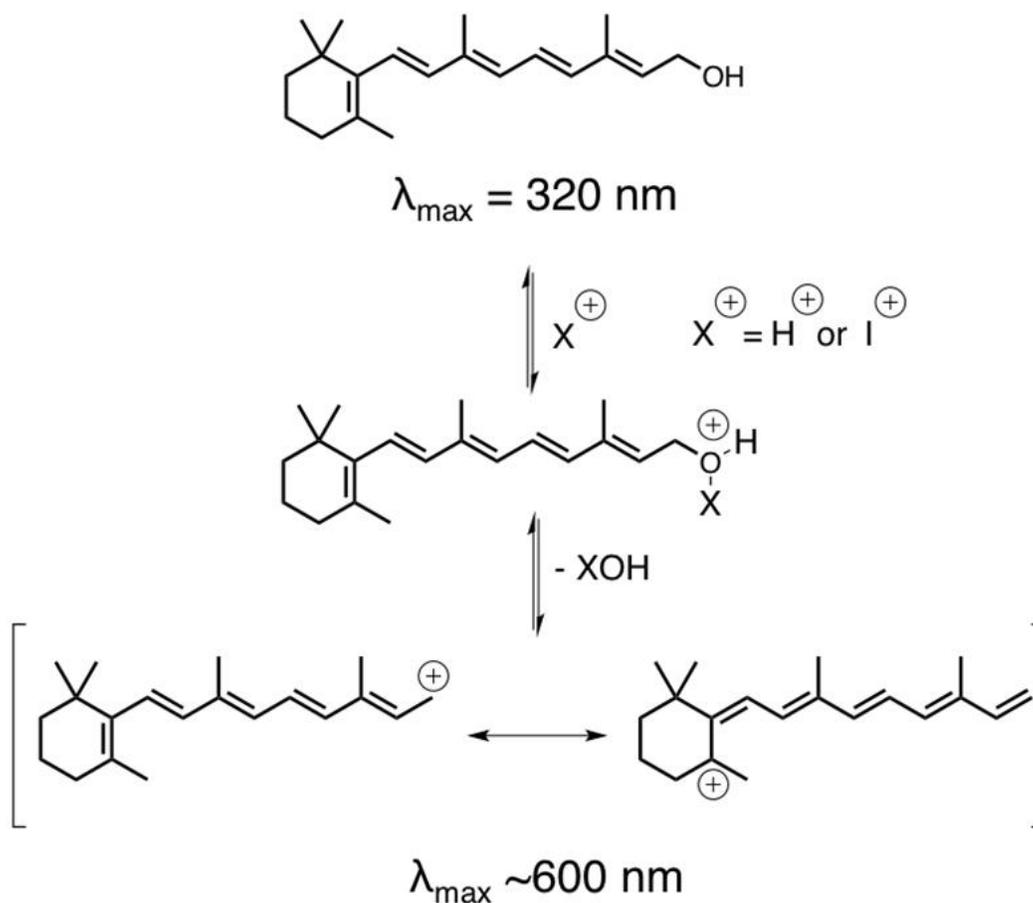


Figure 10: Chemical conversion of all-*trans*-retinol (380 nm) into retinene cation (600 nm).

is believed the protein-chromophore interactions dictate whether the positive charge of the PSB is delocalized along the polyene or localized in the iminium region. The delocalization of the positive charge along the polyene would result in red shift, while the localization of the positive charge in the iminium region would result in a blue shift. The delocalization of the positive charge is believed to be a large contributor to the opsin shift observed. The delocalization of the positive charge results in a decrease in the energy gap between the ground state and the excited state.

Since the difference of the absorbance of red rhodopsin and that of the retinal PSB in solution is 120 nm, many studies have been done to replicate this maximum red shift observed. Blatz (1971) showed that the conjugated cation derived from the dehydration of all-*trans*-retinol (320 nm) was red shifted (600 nm) in comparison to the neutral molecule (Figure I-10). The drastic change in the absorbance was attributed to the delocalization of the positive along the polyene.⁴⁷ This system represents the most delocalization of the positive charge in a polyene and suggests that the maximum absorption of a retinal PSB is around 600 nm.

It is evident that the retinene cation is not the retinylidene and so more studies would have to be done to more accurately predict the maximum absorption of the PSB. To achieve maximum delocalization the counterion was moved further away from the iminium. Blatz performed a model study of retinal-PSB with counterions that differed in size. The studies showed that the larger the counterion the larger the red shift. They could also determine the distance at which the counterion would have a negligible stabilizing effect on the iminium.⁴⁸ The maximum absorption predicted in this study was 580 nm. Honig and Ebrey also predicted that the isolated PSB would have an absorbance maximum of around 600 nm.⁴⁹ In a more recent

study, the absorbance of the retinal-PSB in vacuum was found to be 610 nm.^{50 51} In vacuum there is no counterion present and so maximum delocalization can be achieved. This experimental value correlates well with the previously observed absorption and shows that the importance of positive charge delocalization for red shifting.

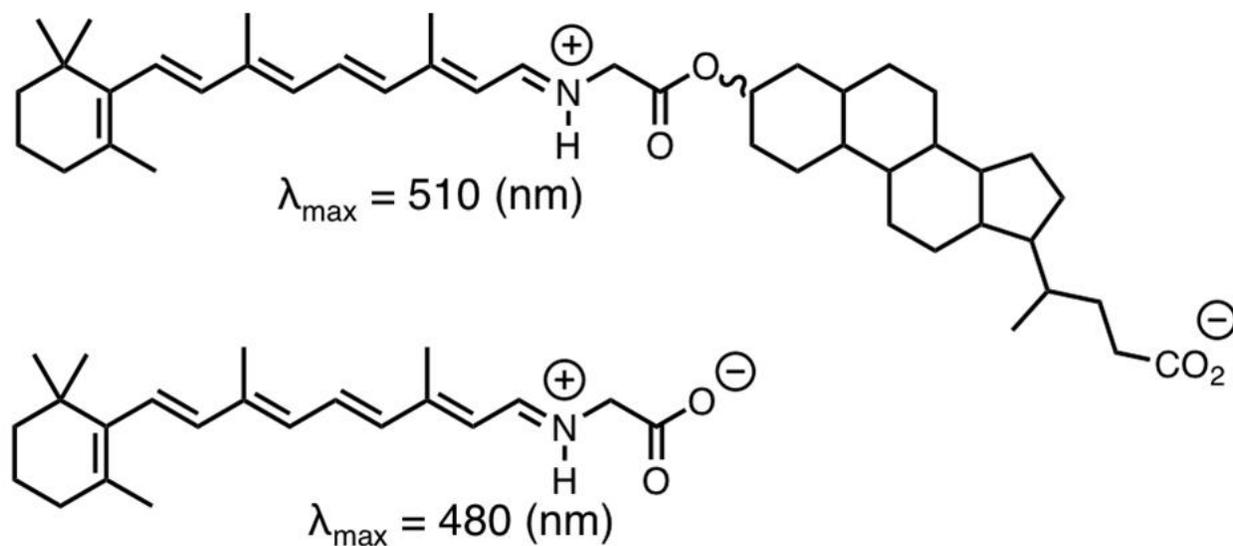


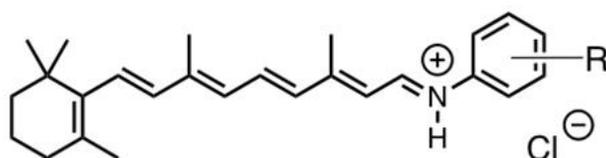
Figure 11: Model compounds for studying the counterion effect on the absorbance maxima of retinal-PSB.

Blatz has effectively shown the effect of the counterion in modulating the wavelength. Weakening the interactions of the counterion with the iminium could encourage positive charge delocalization and lead to red shift. Sheves and coworkers have shown that red shifting is observed when the counterion is placed further away from the iminium (**Figure 11**).⁵²

Electronic interactions between the opsins and retinal-PSB could play important roles in the opsin shifts observed in color rhodopsins. In 1967, Rosenberg used retinal-PSB models to study the electronic inductive effects on the absorption of the chromophore. **Table 1** shows the effects of electron withdrawing groups, which could promote delocalization of the positive

charge on the imine nitrogen through inductive effects, results in red shift.⁵³ In this study the author suggests that the protein environment in rhodopsin could have similar effects as the electron withdrawing groups.

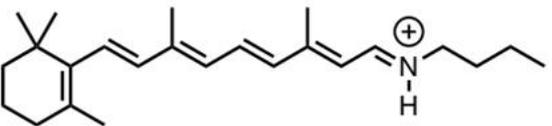
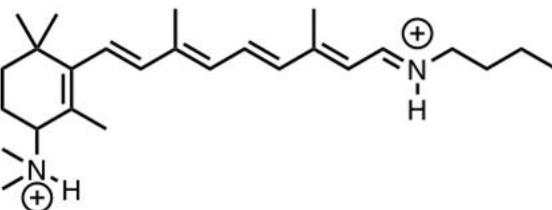
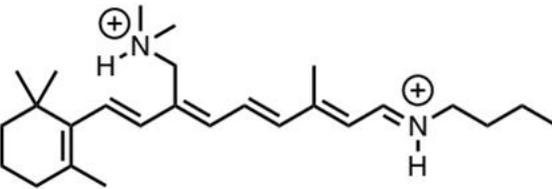
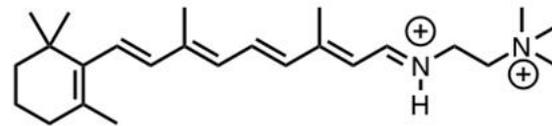
Table 1: Inductive effects in the absorption of retinal-PSB.



| -R | Para Substitution λ_{\max} in EtOH (nm) | Para Substitution λ_{\max} in CHCl ₃ (nm) | Meta Substitution λ_{\max} in EtOH (nm) |
|-------------------|---|--|---|
| -OCH ₃ | 505 | 522 | 508 |
| -H | 504 | 522 | 504 |
| -Cl | 512 | 530 | 514 |
| -CN | 533 | 564 | - |
| -NO ₂ | 534 | 574 | 522 |

It was shown by Sheves' group that the electronic effects between retinal-PSB and its surrounding environment could have a significant effect on the absorption, through model compound studies.⁵⁴ The most red shifted model compounds are the ones with the positively charged groups on the side of the PSB (**Table 2**). This will promote the charge delocalization across the polyene due to positive charge repulsion. However, when positive charges are placed along the polyene or on the ionone ring, a blue shift is observed due to the reduction in charge delocalization. This illustrates that different electronic environments created by the rhodopsins could affect the absorbance by either promoting or inhibiting delocalization.

Table 2: Effects of different placement of positive charge on the absorption maxima of retinal-PSB.

| Chromophore | λ_{\max} | λ_{\max} | λ_{\max} | λ_{\max} | λ_{\max} |
|---|------------------|------------------|------------------|--|---|
| | (nm) | (nm) | (nm) | (nm) | (nm) |
| | EtOH | TFE | TFIF | CH ₂ Cl ₂ (1.0 eq TFA) | CH ₂ Cl ₂ (1.0 M TFA) |
|  | 440 | 467 | 492 | 448 | 513 |
|  | 423 | 431 | 442 | 426 | 461 |
|  | 419 | 419 | 428 | 423 | 455 |
|  | 455 | 508 | 536 | 468 | 538 |

It is on the same premise that the point charge theory was based. The theory proposes that there are other negatively charged residues, besides the counterion, along the polyene that could interact with the chromophore and increase the delocalization of the positive charge.⁴⁶ The theory was put forward due to the close proximity of the counterion, which is necessary for stabilization of the PSB. This would mean that removal of the counterion could not be used to

explain the opsin shift observed. Since the counterion is still present, then negative point charges along the polyene could lead to more delocalization and a red shift.

It was proposed by Honig and Nakanishi that the location of the point charge for bovine rhodopsin and bacteriorhodopsin were different using model compounds, **Figure 12**.^{55 46} In bovine rhodopsin, the point charge was localized in the middle of the polyene, while in bacteriorhodopsin the point charge was localized near the ionone ring. This could be one of the reasons to explain why bacteriorhodopsin (560 nm) is more red shifted than bovine rhodopsin (500 nm).

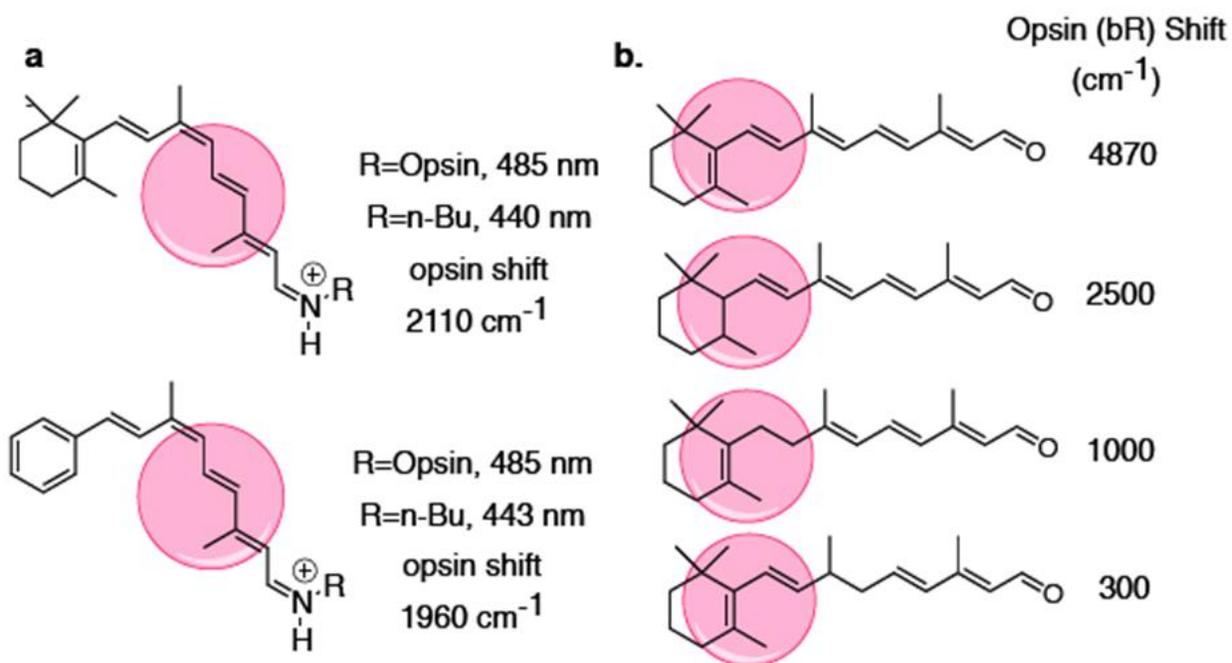


Figure 12: Model compounds used in the point charge theory studies. **a.** Opsin shift observed for bovine rhodopsin. **b.** Opsin shift observed for bacteriorhodopsin.^{55 46}

The result for bovine rhodopsin showed that changes in the ionone region structure has little effect on the opsin shift. This is in good agreement with the hypothesis that the point charge in bovine rhodopsin is localized in the middle of the polyene. The result for bacteriorhodopsin

showed that the most red shifted system is the one in which the conjugation is extended all the way to the ionone ring. If there is a break in the conjugation before the ionone ring then blue shifting is observed. This observation is also in agreement with the hypothesis that the point charge in bacteriorhodopsin is localized near the ionone ring.⁵⁶

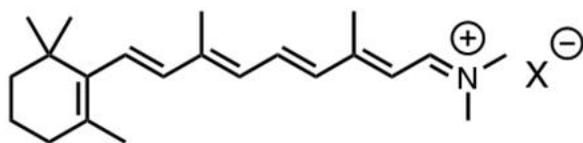
However, mutagenesis studies of the charged residues proved the point charge theory to be flawed as the dramatic changes in the opsin shift was not observed, except when the counterion was mutated.⁵⁷ The crystal structures of both bacteriorhodopsin and bovine rhodopsin show that there are no negatively charged residues in close proximity to the chromophore other than the counterion.^{21 54 58 59} If one considers negative dipoles instead of negative charges then the point charge theory could still hold true.

The effects of polarizability and permanent dipoles could play a significant role in the opsin shift. Irving proposed in the 1970's that induced dipoles could have an effect on the retinal-PSB absorption.⁶⁰ This can be attributed to the difference in the charge distributions of ground state and excited state of the retinal-PSB. This difference in charge distribution could greatly affect the polarity of a polarizable environment. Stryer and Mathies determined the dipole moments of the ground state and excited state retinal-PSB and they observed that there was a transfer of negative charge from the ionone ring region to the PSB upon excitation.⁶¹

The permanent dipoles of the amino acid side chains orient themselves in a manner that stabilizes the ground state to the largest extent possible. Therefore, when the retinal-PSB is excited the difference in the charge distribution results in the development of unfavorable electrostatic interactions. This phenomenon is due to the fact that the permanent dipoles are

unable to orient themselves in a manner to stabilize the excited state within the time scale of the excitation process.

Table 3: Effect of polarizable solvents on retinal-PSB with different counterions.



| Solvent | Dielectric Constant | λ_{\max} | λ_{\max} | λ_{\max} |
|-----------------|---------------------|---|----------------------------|-----------------------------|
| | | X = ClO ₄ ⁻ (nm) | X = I ⁻ (nm) | X = Br ⁻ (nm) |
| Ethyl ether | 4.33 | 451 | 445 | 442 |
| Methanol | 32.36 | 453 | 448 | 444 |
| Benzene | 2.28 | 474 | 455 | 451 |
| Chloroform | 4.80 | 481 | 485 | 477 |
| Dichloromethane | 9.08 | 496 | 489 | 484 |

This hypothesis was studied further by using solvents that provide different polarizable environments. The results showed that the more polarizable the solvent the more red shifted the PSB (Table 3).⁶²

I.2 MUTAGENESIS STUDIES ON THE VISUAL RHODOPSINS

Developments in the field of genetics and molecular cloning allowed for the comparison of the sequence of the different visual opsins. Nathans and coworkers reported the sequence of the genes encoding the red, green and blue rhodopsin for the first time in 1986.⁶³ The results showed

that blue rhodopsin is very different from red and green rhodopsins and green and red rhodopsins are quite similar, differing in only 15 amino acids.

Table 4: Sequence identity (below the diagonal) and sequence homology (above diagonal) between different visual rhodopsins.

| | % Rhodopsin | % Blue | % Green | % Red |
|------------------|------------------------|-------------------|--------------------|------------------|
| Rhodopsin | 100 | 75 | 73 | 73 |
| Blue | 42 | 100 | 79 | 79 |
| Green | 40 | 43 | 100 | 99 |
| Red | 41 | 44 | 96 | 100 |

Note: Sequence homology is the percentage of conserved residues with the same functionally equivalent physicochemical properties. Sequence identity is the percentage of identical residues in sequence alignment.

Nietz and coworkers carried out the sequence alignment on eight cone pigments with absorptions varying from 530 nm to 560 nm. From the alignment they concluded that three amino acids are likely be responsible for the difference in absorbance of red and green rhodopsin.⁶⁴

Mutagenesis studies by Sakmar and Oprian further supported this finding. Oprian mutated all the 15 residues where green rhodopsin is different from red according to the sequence alignment.⁶⁵ From the results obtained, it was concluded that seven residues are responsible for the difference between red and green rhodopsin, three of which were already proposed by Nietz.

Sakmar and coworkers also carried out mutagenesis studies by mutating the three residues proposed by Nietz in bovine rhodopsin. The results are shown in **Table 5**.

Table 5: Table of mutants of Bovine Rhodopsin.

| Mutation | λ_{\max} (nm) | Protein Shift (cm^{-1}) |
|-------------|-----------------------|------------------------------------|
| WT- bRho | 500 | 0 |
| A164S | 502 | 75 |
| F261Y | 510 | 400 |
| A269T | 514 | 550 |
| F261Y/A269T | 520 | 775 |
| F261Y/A164S | 512 | 400 |
| A269T/A164S | 514 | 550 |

The crystal structure of bovine rhodopsin shows that the three residues being considered for mutagenesis are located in the ionone ring region. It is expected that when these residues are mutated to polar residues that a red shift should be observed because these polar residues should aid the stabilization of the excited state thus leading to a red shift. However, as shown in **Figure 13**, the maximum red shift observed was 20 nm. This shift is not sufficient to explain the opsin shift if 120 nm observed for red rhodopsin.

From the crystal structures it is evident why the mutation A169S had very little effect on the retinal-PSB absorption. A164 is 9.9 Å away from the chromophore and because the coulombic force is inversely proportional to distance squared, the longer the distance the weaker the electrostatic interactions. Mutation of A269 and F261 resulted in a red shifted species since they are much closer to the ionone ring. This red shifting is still a bit surprising considering that the residues are 5 Å away from the chromophore. This is probably due to the low dielectric

constant in the hydrophobic environment of the binding pocket. The electric field generated from the dipole in this environment tends to have a long lasting effect over distance.

This study on bovine rhodopsin has illustrated the importance of changes in polarity of the protein environment and hydrophobicity on wavelength regulation.

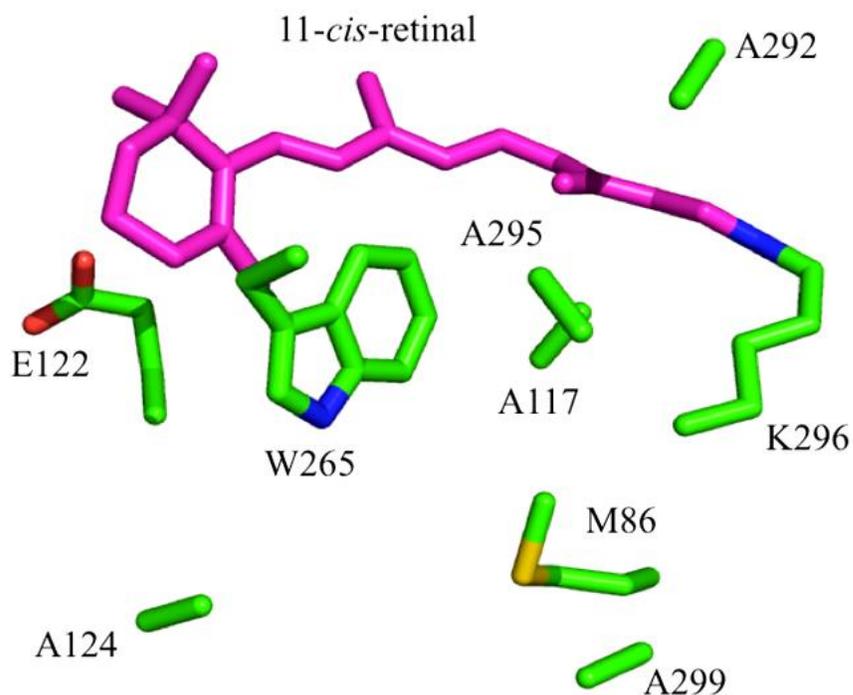


Figure 13: Crystal structure of bovine rhodopsin highlighting the residues that are important for blue shifting.

Sakmar and coworkers also carried out mutagenesis studies on rod rhodopsin. Nine mutations in all blue shifted rod rhodopsin from 500 nm to 438 nm.⁶⁶ **Figure 14** shows the positions that were mutated. The mutations made were as follows: A124T, A299C, A295S, A117G, A292S, G90S, E122L, M86L, W265Y. The results show that the introduction of negative polarity in the PSB region in addition to the removal of the negative polarity from the ionone ring region lead to blue shift.

However, no crystal structures were obtained for the mutants to verify that it is strictly an electronic change and not a structural change in the protein. It was also found that Trp 265 is critical for proper protein folding in red and green rhodopsin. In blue rhodopsin however, this residue is replaced by another aromatic residue, Tyrosine.⁶⁷ The mutation of Trp 265 to Tyrosine resulted in the most drastic blue shifting. This suggests that the Tryptophan plays a crucial role for red shifting in the red and green rhodopsins. The mechanism by which Trp 265 causes red shift and W265Y causes blue shift would be difficult to ascertain without crystallographic data.

Mutagenesis studies have also been done to elucidate the role of the counterion of the rhodopsin system. Oprian and coworkers indentified the counterion by mutating every aspartate and glutamate residue thought to be embedded in the membrane to see which one would affect PSB formation. It was found that mutating Glutamate at position 113 to Glutamine (E113Q) significantly reduced PSB formation. This is a result of lowering the pK_a of the PSB due to less stabilization.⁶⁸

Sakmar and coworkers later showed that the red shift caused by Glu 113 mutations is affected by the anions present in the solution, **Table 6**.⁷⁰ The studies showed that when Glu 113 is mutated to a neutral residue like Glutamine, hydrophobic residues, or even shortening the length of the side chain using aspartate, could lead to a 30 nm red shift.

In the case of the mutant E113A, the PSB can be stabilized by the solute anion. The cavity that is formed as a result of the mutation could allow for the passage of small anions such as fluoride into the binding cavity, which can then stabilize the PSB. This stabilization would result in a blue shift when compared to rhodopsin, which was observed. The larger anions such

as tartrate would experience more difficulty entering the binding cavity; this would result in a red shift, which was observed. Crystallographic data is necessary to verify this hypothesis.

Table 6: Effect of solute anions on absorption maximum of rhodopsin E113 mutants.

| Solution Ion | λ_{max} (nm) Rho | λ_{max} (nm) E113D | λ_{max} (nm) E113N | λ_{max} (nm) E113Q | λ_{max} (nm) E113A |
|-------------------------|---|---|---|---|---|
| Fluoride | 499 | 505 | 515 | 508 | 486 |
| Chloride | 499 | 510 | 520 | 496 | 506 |
| Bromide | 499 | 504 | 512 | 493 | 500 |
| Iodide | 501 | 510 | 519 | 504 | 507 |
| Formate | 499 | 507 | 520 | 488 | 496 |
| Acetate | 498 | 501 | 522 | 488 | 496 |
| Perchlorate | 498 | 508 | 520 | 500 | 510 |
| Tartrate | 497 | 509 | 522 | 510 | 525 |
| Citrate | 497 | 507 | 524 | 516 | 526 |
| Bezoate | 497 | 509 | 524 | 513 | 528 |
| Chloroacetate | 498 | 504 | 514 | 494 | 488 |
| Dichloroacetate | 498 | 504 | 524 | 501 | 509 |
| Trichloroacetate | 498 | 503 | 519 | 506 | 512 |
| Deionized water | 500 | 513 | 520 | 515 | 528 |

Similar observations were made using some of the cone visual pigments. The pigments were incubated with different solute anions, which resulted in different absorption maxima. The hypothesis put forward to explain the different absorption maxima is that different anions would

result if the formation of different hydrogen bonding networks in the PSB region.³⁰ Again crystal structures are necessary however, to prove this hypothesis.

In conclusion, mutagenesis studies have greatly enhanced our understanding of the visual pigment and factors that govern wavelength regulation. However, to hone in on the intricate protein-chromophore interactions, which are necessary for wavelength regulation, crystallographic data is needed. Comparison of the crystal structure of the native protein and the mutants would provide a wealth of information about the effects of the mutations on the protein, electronic or conformational, as well as the retinal-PSB.

I.3 MUTAGENESIS STUDIES ON THE MICROBIAL RHODOPSINS

Rhodopsin systems are not unique to higher organisms. Similar retinal binding rhodopsins are also found in lower organisms such as bacteria (bacteriorhodopsin, previously discussed) and archaea. The rhodopsins found in these organisms are used for other functions such as light activated proton or chloride pumping, and phototaxis.⁷⁰ Several types of rhodopsin can be found in a single organism with different absorption maxima. An example is an archaeon that was found to have six different rhodopsin with absorption wavelengths ranging from 483 nm to 578 nm.⁷¹ **Figure 14** illustrates the light dependent functions of the six rhodopsins found in the archaeon.

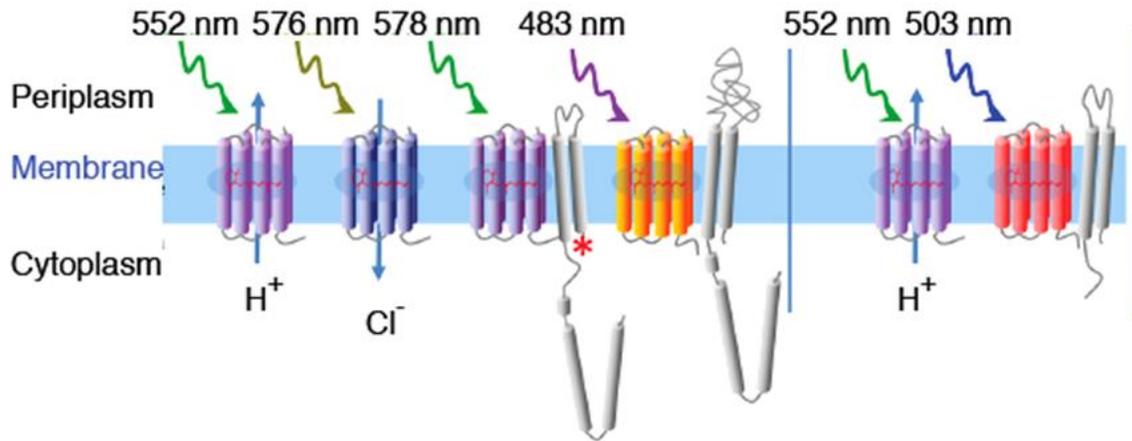


Figure 14: Diagram showing the six different rhodopsins found in a single Archaeon.⁷¹

Crystallization of these microbial rhodopsins proved to be very successful as crystal structures of bacteriorhodopsin, sensory rhodopsin II, and halorhodopsin, have been obtained.⁵⁸

⁷² ⁷⁴ These structures provide an excellent platform to probe wavelength regulation. The crystal structures revealed that there are a number of tryptophan residues lining the polyene in the binding pockets of the three rhodopsins. This observation led to the examination of the role of polarizable, electron rich aromatic residues as it relates to wavelength regulation.

The mutation of these Tryptophan residues to Phenylalanine resulted in a blue shift.⁷⁴ This can be the result of the fact that the tryptophan residues are better able to stabilize the positive charge through π -cation interactions. This can facilitate more delocalization of the positive charge along the polyene. Being polarizable, the tryptophan residues can also stabilize the excited state chromophore through induced dipoles, thus leading to red shift. Tryptophan being large and flat can also play a role in planarizing the chromophore and rigidifying the system, which can also contribute to the red shift. However, crystal structures of the mutants of

bacteriorhodopsin have not been crystallized to help better understand the contribution of the mutations to the opsin shift.

It was also found that the methionine residues that are close to the ionone ring were critical for the red shift observed in bacteriorhodopsin.⁷⁵ This could be attributed to the polarizability of the sulfur group and the close packing of the residues onto the chromophore contributing to the rigidity of the system. Crystal structures are essential to verify this hypothesis.

Mutagenesis studies on the counterion were also done on bacteriorhodopsin to probe the counterion effect on the absorption wavelength. As expected, mutation of Asp 85 to a neutral residue leads to a red shift.⁷⁶ It was interesting to observe a red shift at pH=2 for bacteriorhodopsin. It was believed that at pH=2 the aspartate is protonated and this allows for the delocalization of the positive charge along the polyene leading to red shift. However, the crystal structure of bacteriorhodopsin shows an elaborate hydrogen bonding network in the PSB region and so mutation of the counterion could result in drastic changes in the hydrogen bonding network. Crystal structures are crucial in order to understand the effects of this hydrogen bonding network on wavelength regulation.

The factors resulting in the absorption difference between sensory rhodopsin II (500 nm) and bacteriorhodopsin (560 nm) are still not clearly understood, even with the availability of the crystal structures. Mutagenesis of the 10 amino acids of sensory rhodopsin II to the corresponding amino acids in bacteriorhodopsin proved to be unsuccessful. A maximum shift of 24 nm was observed.⁷⁷ The residues mutated were about 5 Å away from the chromophore so the results indicate that residues further away from the chromophore could have an effect on the

spectral profile. This is not difficult to perceive as the long range electrostatic interactions are possible in a low dielectric, hydrophobic environment which is the case for rhodopsin.

In conclusion, the mutagenesis studies on microbial rhodopsins aided the investigation of factors responsible for wavelength regulation. However, the lack of crystallographic data has hampered the accurate elucidation of the effects of the mutations on the absorption and on protein-chromophore interactions.

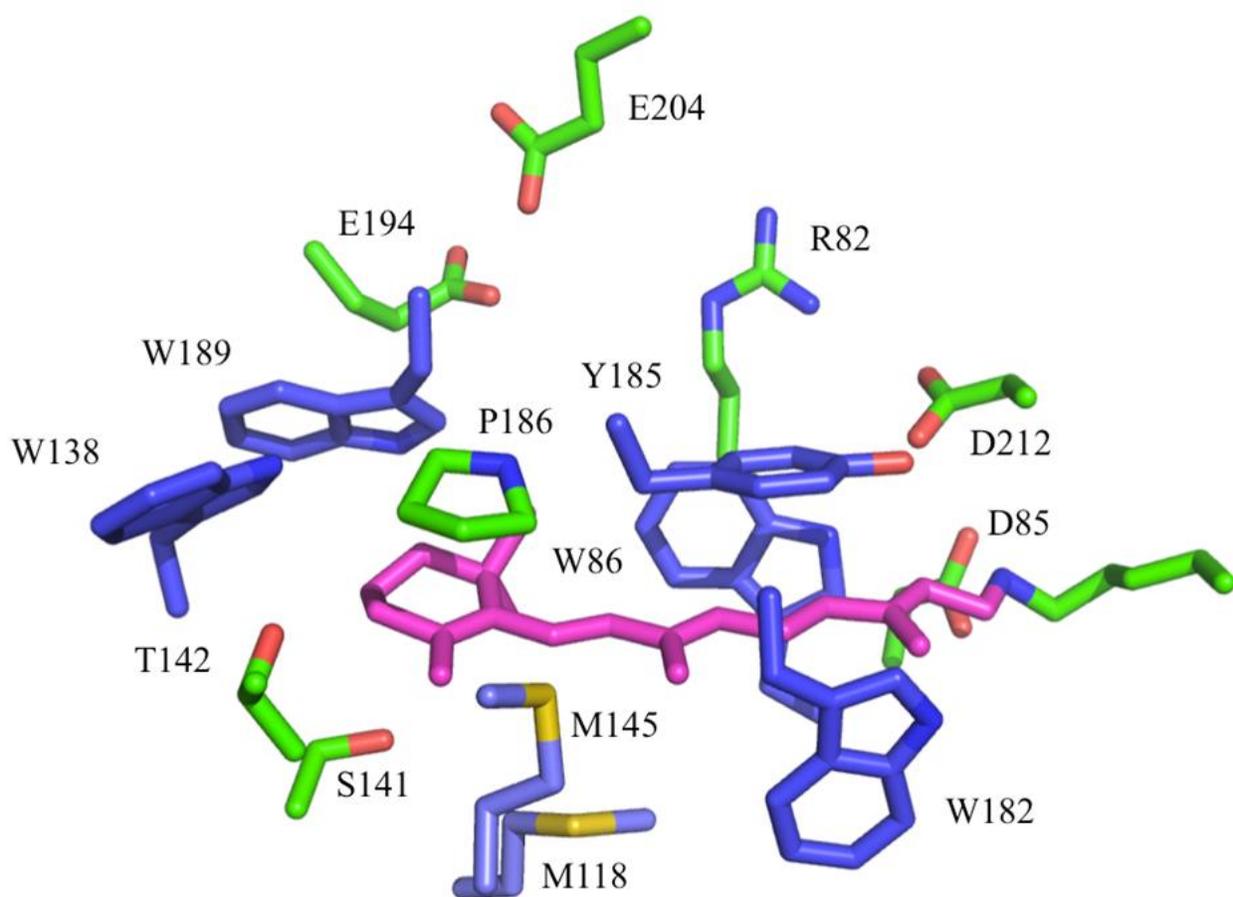


Figure 15: Binding cavity of bacteriorhodopsin. (PDB code: 1C3W).

I.4 WAVELENGTH REGULATION DUE TO CONFORMATIONAL CHANGE

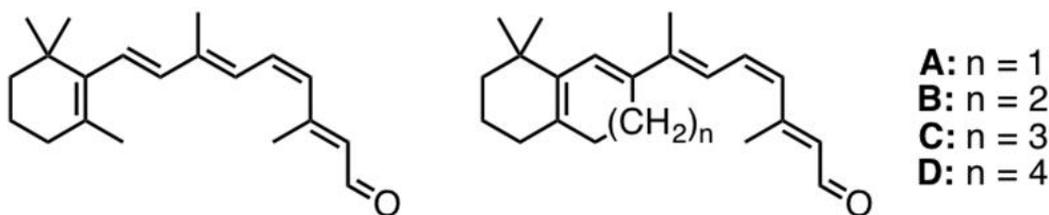
The conformation of the chromophore has been suggested to play a role in the observation of different absorption maxima of a retinal-PSB. For a polyene system, the more planar the molecule the better the conjugation and the more red shifted the absorption. On the contrary, if there is a reduction of the conjugation due to twisting about the single bonds of the polyene, then a blue shift will be observed. The postulate that relates this phenomenon to the rhodopsin systems is that the different absorption maxima observed are due to different protein-chromophore packing. This means that the more red shifted retinal-PSB would be planar and the blue shifted retinal-PSB would be more twisted.

Since there are few crystal structures of the rhodopsins systems, a comparison of the conformation of the chromophore in each rhodopsin cannot be made to support or refute this hypothesis.

The major site of conformational differentiation is the C6 – C7 single bond. The two conformations that result from the rotation of this bond are *6-s-cis* and *6-s-trans*. In the case of the *6-s-cis* conformer, the C6 – C7 bond is twisted to relieve the steric interactions between C5-methyl and C8-hydrogen. This results in less conjugation as the plane of the ionone ring is twisted away from the plane of the polyene. In the case of the *6-s-trans* conformer, the polyene is more flat and therefore more conjugated. The *6-s-cis* conformer is observed in bovine rhodopsin, while the *6-s-trans* conformer is observed in the microbial rhodopsins. This observation was proposed to be the reason why microbial rhodopsins are more red shifted than the visual rhodopsins.

This hypothesis was explored using different ring-locked retinal analogues in bovine rhodopsin systems.⁷⁸ The results shown in **Table 7** reveal that planarization of the C6- C7 bond leads to red shifting.

Table 7: Characterization of different ring-locked analogues of retinal.



| Chromophore | λ_{\max} (nm) Aldehyde | λ_{\max} (nm) PSB | λ_{\max} (nm) Rhodopsin | Opsin Shift (cm^{-1}) |
|----------------|-----------------------------------|------------------------------|------------------------------------|-------------------------------------|
| A | 422 | 506 | 539 | 1200 |
| B | 416 | 496 | 546 | 1850 |
| C | 386 | 457 | 503 | 2000 |
| D | 374 | 440 | 483 | 2000 |
| 11-cis-retinal | 377 | 440 | 498 | 2650 |

The comparison of the 11-*cis*-retinal aldehyde and PSB with the ring-locked compounds aldehyde and PSB showed that the opsin shift for the ring locked compounds is less than that observed for 11-*cis*-retinal. Consider compounds **A** and **B**, these compounds are the more planar and yet gave the least opsin shift. This would suggest that a portion of the opsin shift observed for rhodopsin bound to 11-*cis*-retinal is a result of the planarization of the C6 – C7 bond within the binding pocket. Since compounds **A** and **B** are already planar in solution, less planarization is required which resulted in a smaller opsin shift.

Garavelli and coworkers did calculations to determine the absorption of all-trans-retinal and retinal analogues PSBs in vacuum. The results show that the 6-*s-trans* conformer of the retinal PSB absorbs at 610 nm and the 6-*s-cis* conformer absorbs at 530 nm. This difference in absorption is due to the dihedral angle of the C6 – C7 bond.⁵¹ It is evident that conformation plays a role in understanding protein-chromophore interactions and so crystal structures are crucial to this end.

I.5 MODERN COMPUTATIONAL STUDIES ON WAVELENGTH REGULATION

Different hypotheses have been put forward regarding wavelength regulation. Each of these hypotheses has contributed significantly in trying to understand the intricacies of spectral tuning. However, even though there have been significant developments to this end, no single hypothesis has been put forward to fully explain wavelength regulation.

Over the years there have been remarkable developments in computational methods, which allows for the use of modeling studies in order to quantify the absorption maxima. Sakurai's group carried out computational studies to try to decipher the 10 nm red shift observed in halorhodopsin upon binding of chloride ions.⁷⁹ This phenomenon is opposite to what is expected as the placement of a negative charge in the PSB region should result in a blue shift. For the calculations, they set the ionizable groups within the interior of the protein to be charged if a stabilizing ion pair interaction is possible and all the ionizable groups that are exposed to the aqueous media to be neutral. Whenever the electronic state of the chromophore is the only consideration the expected blue shift was obtained from the calculations. However, if the induced polarization in the protein environment, which results in the delocalization of the positive charge towards the ionone ring region upon excitation, is considered then a red shift is

obtained from the calculations. These results provided the evidence of the role of electronic polarization on wavelength regulation.

Computational studies to determine the vertical excitation energy from S_0 to S_1 of the different twisted 6-*s-cis* conformations of 11-*cis*-retinal PSB were carried out by Garavelli and Olivucci in the gas phase.⁸⁰ It was found that the most twisted conformer was the most blue shifted. This was not surprising as the more twisted the ionone ring the less the overlap between the ionone ring π system and that of the polyene. **Figure 16** shows the results of the calculation showing effect of twisting the ionone ring on the absorption maximum.

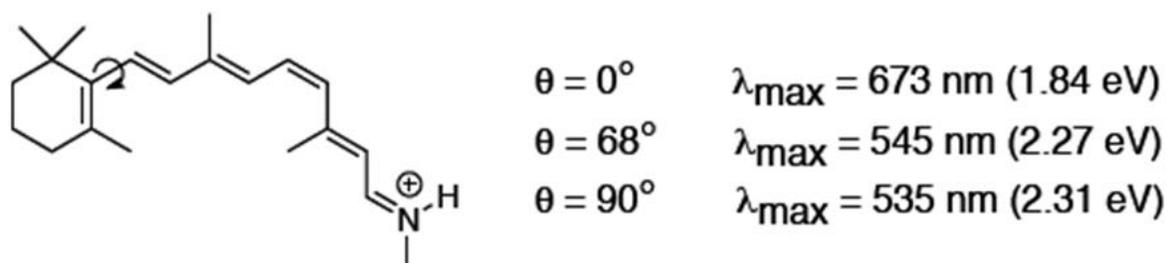


Figure 16: Results of the computational analysis of 11-*cis*-retinal-PSB absorption with different torsion of the C6 – C7 bond highlighted.⁷⁸

The value obtained for the 90° torsion (535 nm) was in good agreement with the value obtained for 5,6-dihydroretinal in the gas phase.⁷⁸ This is not surprising as in the case of the 90° torsion there is no overlap of the ionone ring π system with the polyene which should be similar to not having a ionone ring π system as is the case in 5,6-dihydroretinal. These results would therefore suggest that in the case of red and green rhodopsin the chromophore is more planar in comparison to blue rhodopsin. Trabanino and coworkers reiterated this hypothesis in 2006.⁸¹

The most red shifted retinal-PSB calculated in the gas phase has a absorption of 610 nm. This provided a new platform for further computational studies. Volker Buss and coworkers carried out calculations of rhodopsin absorptions starting from the most red shifted calculated retinal-PSB (610 nm) and proceeded to build in a counterion, followed by other protein residues.⁸² The results showed that the introduction of the counterion caused the most blue shifting from 610 nm to 486 nm. The introduction of the additional residues resulted in a small red shift through the dipolar interaction with the chromophore. In these calculations, the conformation of the chromophore was set to play a minimal role in wavelength tuning in rhodopsin. Therefore, the conclusion was that it is the counterion and not the residues in the binding pocket that dictates the absorption maxima of the retinal-PSB. This statement proved to be faulty as the removal of the counterion in bovine rhodopsin resulted in only a 30 nm red shift.⁶⁹

Olivucci and Ryde performed calculations at the CASPT2//CASSCF level (second-order multiconfigurational perturbation theory) for WT rhodopsin and two mutants and isorhodopsin in QM/MM structures based on two crystal structures.⁷¹ The results showed that the dipoles or polarizable residues do affect the electronic characteristics of the chromophore. The limitation is that different calculations could predict different results. Notwithstanding, calculation of retinal-PSB absorption in the protein is not trivial as there are many variables; steric, electronic and otherwise, that could greatly affect the electronic profile of the retinal-PSB.

In conclusion, performing computational analysis on the retinal-PSB within the protein environment is a difficult task. There are many factors that need to be considered such as; dynamics and polarizability of the residues in the binding pocket. Crystallographic data on the

rhodopsins are limited. Development of computational studies on proteins that differ in only a few amino acids and exhibit distinctive absorption profile would serve as a good platform for probing factors for quantifying the electronic excitation energies.

I.6 STRATEGIES FOR SPECTRAL TUNING IN A RHODOPSIN MIMIC

The factors that are important for spectral tuning are still not completely discerned. This is mainly due to a lack of crystallographic data for the color rhodopsins and the mutant designed for the study of wavelength regulation. Rhodopsin being a transmembrane protein, expression and crystallization is rather challenging. Rhodopsin is also very light sensitive, which makes the preparation more wearisome. Therefore, the design of a rhodopsin mimic that binds retinal as a protonated Schiff base would provide an excellent platform for probing the factors governing wavelength regulation. The mimic has to be amenable to mutagenesis and routine crystallization conditions.

Previous studies on wavelength regulation showed that two important factors were being considered; the difference in charge distribution between the ground state and the excited state, and the electrostatic interactions between the residues and the retinal-PSB in both the ground state and the excited state.

The electronic profile of the carbons in all-*trans*-retinylidene in the ground state showed that C15, the iminium carbon, has the largest distribution of positive charge. As shown in **Figure 17**, all the odd numbered carbons exhibit a higher distribution of positive charge than the even numbered carbons. This is due to the resonance structures of all-*trans*-retinylidene to delocalize the positive charge along the polyene. In the excited state, the distribution of positive charge is very different (**Figure 18c**). The positive charge has moved away from the iminium and towards

the ionone ring (C5 to C9). C11, C13 and the iminium nitrogen are now negatively charged. Increasing the negative dipole in the ionone ring region can stabilize the excited state thus reducing the energy gap between the excited state and the ground state.

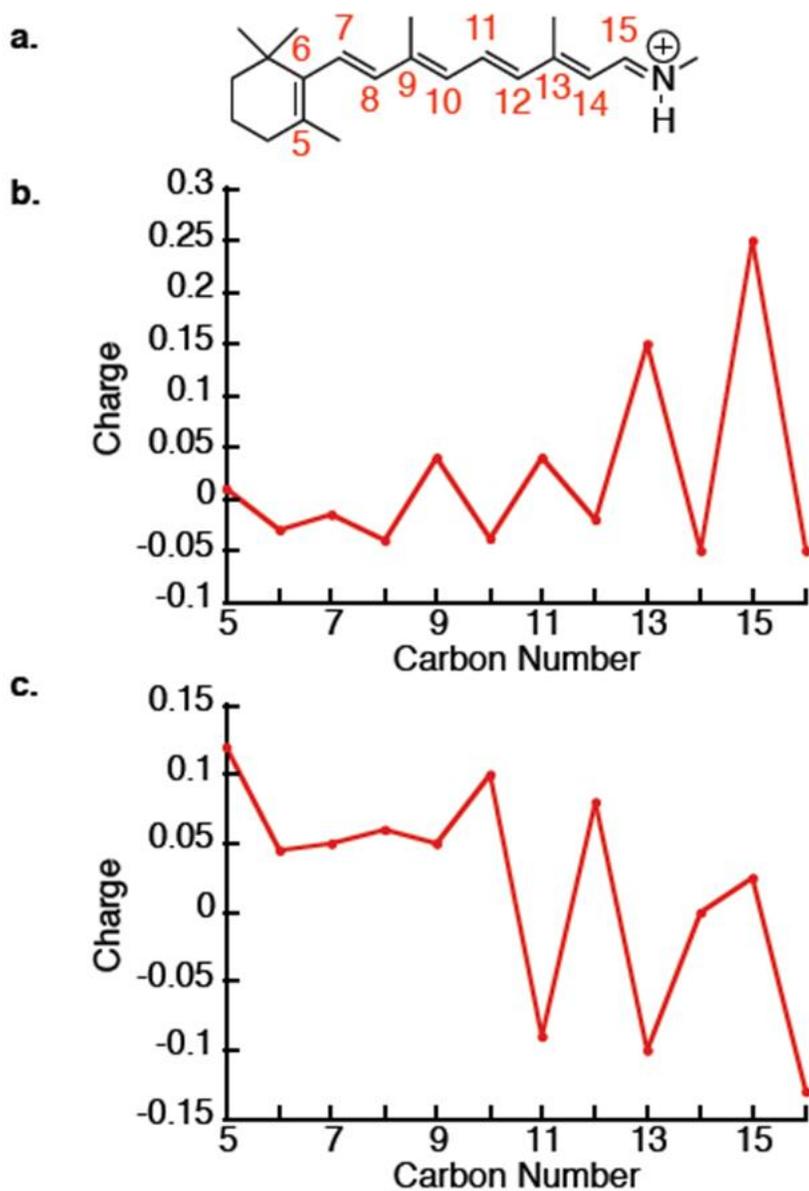


Figure 17: Charge distribution of retinal-PSB. **a.** All-*trans*-retinal-PSB with numbered polyene carbons. Calculated charged state of the carbon atoms in the **b.** ground state and **c.** excited state.⁸³

The alternate single bond /double bond character varies between the Schiff base and protonated Schiff base and the ground state and excited state. Buss and coworkers performed calculations on the bond distance of all-*trans*-retinal SB and PSB in solution as well as the rhodopsin retinal-PSB. **Figure 18a** shows that the Schiff base is more blue shifted and has the most alternate single bond–double bond character. However, when the Schiff base gets protonated it red shifts 80 nm from 360 nm to 440 nm. The single bond-double bond character becomes less significant resulting in a red shift looking at the bond distances from C9–C10. This phenomenon is a result of the delocalization of the positive charge along the polyene. **Figure 18b** shows that the more stable the resonance structures, the more delocalized the positive charge, which will result in a slighter degree of single bond-double bond alteration.

The cyanine dyes are different in that there is no difference in the stability of the resonance structures (**Figure 18b**). As a result of this very little single bond-double bond alteration is observed. The red shift observed for these cyanine dyes is due to the fact that the cyanine dyes are inherently more red shifted than retinal with the same number of double bonds.⁸⁴ This suggests that stabilizing the alternate resonance structure would be another way to cause red shift. Introduction of favorable electrostatic interactions, such as negative dipoles in the ionone ring region, could facilitate this.

Previous studies and recent computational studies on the rhodopsin models seem to corroborate this finding that the overall electrostatic interaction with the chromophore is essential in spectral tuning. This phenomenon could be due to electrostatic interactions with the ground state, the excited state, or electrostatic interactions that stabilize the alternate resonance structures.

Ullmann and coworkers carried out calculations to obtain the overall electrostatic potential projected onto the Van der Waals surface of the chromophore.⁸⁵ For this study, sensory rhodopsin II (500 nm), bacteriorhodopsin (560 nm) and halorhodopsin (570 nm). This would describe the electrostatic interaction between the protein and the chromophore. Since the conformation of the side chains of the residues is essential to determine the electrostatic potential projected on the chromophore, a high resolution crystal structure of the retinal-bound protein is required.

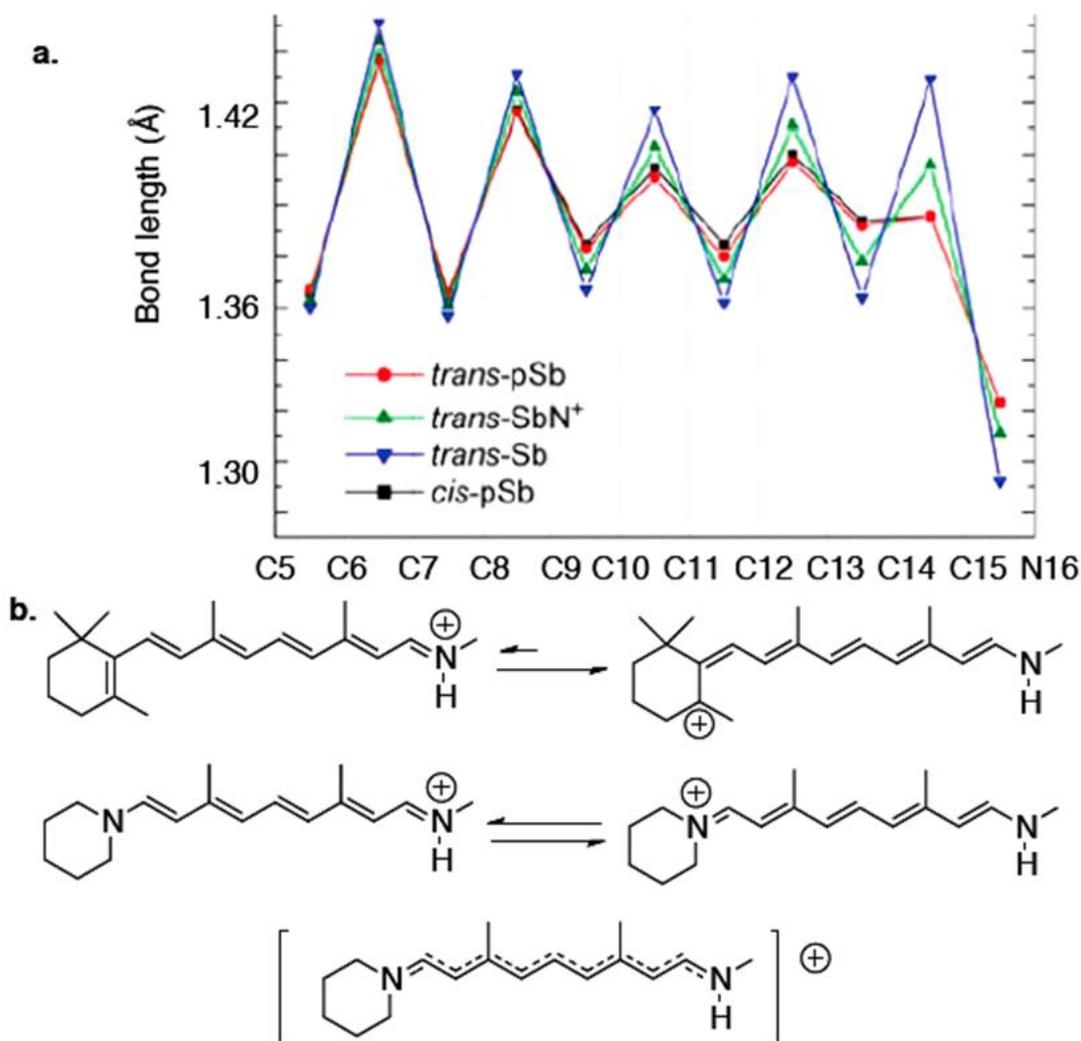


Figure 18: **a.** The calculated bond distances for different retinal compounds.⁸³ **b.** Resonance structures of cyanine dyes and retinal-PSB.

Similar calculations were done in our lab using the three color rhodopsin models and bovine rhodopsin crystal structures.⁸⁶ The results obtained were consistent with reported data for the color rhodopsins (**Figure 19**).

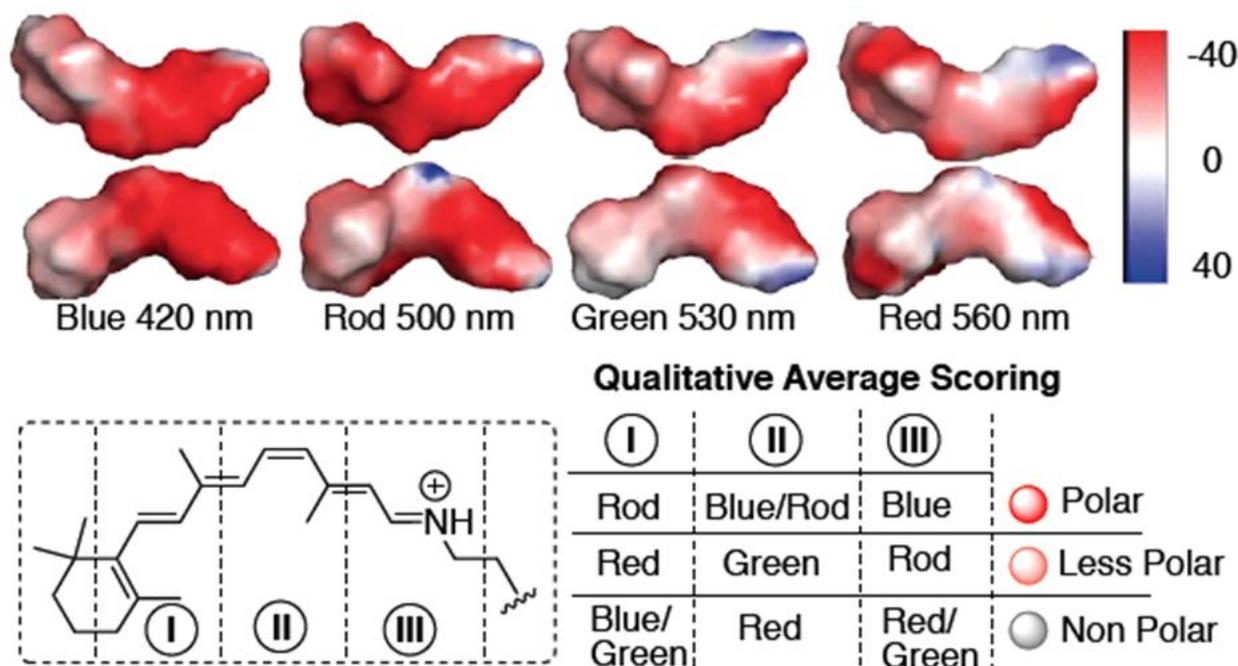


Figure 19: Electrostatic potential calculation. Electrostatic potential calculation (APBS suite) of blue, rod, green, and red opsin (calculations of blue, green and red are based on available homology models) projected on the Van der Waals surface of retinal (the electrostatic potential of the retinylidene chromophore was set to zero as to only illustrate the electrostatic contribution of the protein).⁸⁶ The chromophore is divided into the three segments; the qualitative average score for each segment represents the overall electrostatic potentials that lead to wavelength regulation of each opsin.

The results obtained showed that for blue rhodopsin, the most negative potential is seen around the Schiff base region, as expected. The negative potential in the Schiff base region would stabilize the positive charge on the iminium thus reducing delocalization. On the other hand, the negative potential for the red and green opsins is localized more toward the ionone ring region that would lead to more delocalization and results in a red shift. These results show that

subtle changes in the electronic environment of the protein could greatly affect the absorption profile of the opsins.

These electrostatic calculations will be used as a guide for analysis of the spectral tuning in future studies on wavelength regulation.

I.7 AN OVERVIEW OF PROTEIN X-RAY CRYSTALLOGRAPHY

X-ray crystallography can be described as a form of microscopy with very high resolution. It enables the visualization of a detailed model of a large molecule (macromolecules) at the atomic level. These models provide a platform for enhancing the understanding of protein function.⁸⁹

This data is obtained from a single crystal that consists of identical molecules in an ordered array. The amount of detail or resolution is governed by the wavelength of the electromagnetic radiation used. For example, in the case of light microscopy the shortest wavelength utilized is 300 nm. In the case of electron microscopy, the wavelength could be below 10 nm. However in order to visualize proteins at the atomic level the wavelength required would be around 1 Å, which would be X-rays (**Figure 20**).⁸⁸

When an object is seen, light rays are diffracted by the object and enter the eye through the lens, which then reconstructs an image on the retina. In the simple microscope, the object is placed just beyond the focal point of the lens; this is called the objective lens. The lens collects the light that is reflected from the object and forms an image beyond the focal point on the opposite side of the lens. This phenomenon is quite different where X-rays are concerned. Even though individual atoms diffract X-rays, it is not possible to form an image of a single molecule because lenses cannot focus X-rays. Measuring the directions and intensities of the diffracted rays and then using a computer to simulate an image-forming lens circumvent this problem.

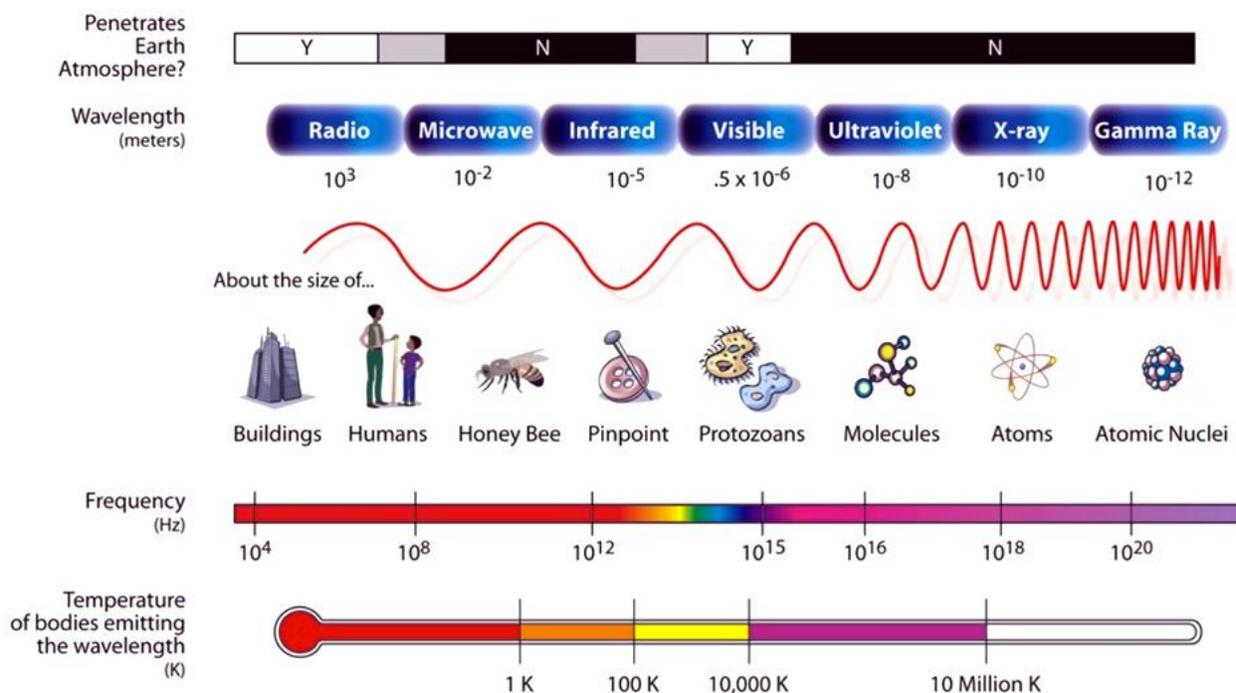


Figure 20: The Electromagnetic Spectrum.⁸⁸

The computer in this case is behaving as the lens generating the image. Secondly, a single molecule is a very weak scatterer of X-rays and so most of the X-rays will pass through a single molecule without being diffracted. Therefore, the diffracted beams would be too weak for detection. Analyzing the diffraction from crystals because crystals are an organized array of similar molecules with identical orientations solves this problem. Within the crystals are asymmetric units, which are the smallest unit volume that contains all of the structural information and by application of the symmetry operations can reproduce the unit cell (**Figure 21**).⁹⁰

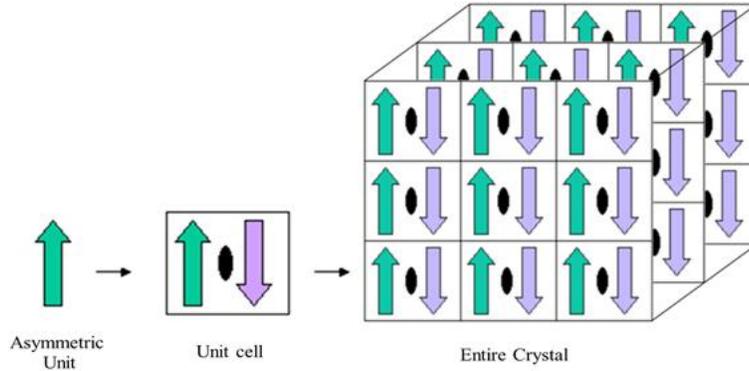


Figure 21: Diagram illustrating the relationship between the asymmetric unit, the unit cell and the entire crystal.

Six numbers describe the dimensions of the unit cell. These numbers correspond to the lengths of the three unique edges **a**, **b**, and **c**; and the three unique angles α , β , and γ (**Figure I-22**). A cell which has $a \neq b \neq c$ and $\alpha \neq \beta \neq \gamma$, as in figure I-23 is called triclinic. A cell which has $a \neq b \neq c$ and $\alpha = \gamma = 90^\circ$, and $\beta > 90^\circ$ is monoclinic. If $a = b$, $\alpha = \beta = 90^\circ$ and $\gamma = 120^\circ$, the cell is hexagonal. If $a = b = c$ and all the angles, α , β , and $\gamma = 90^\circ$ then the cell is cubic. If $a = b \neq c$, the cell is tetragonal; and if $a \neq b \neq c$ then the cell is orthorhombic. These crystal systems serve as the platform for the thirteen unique lattice types.

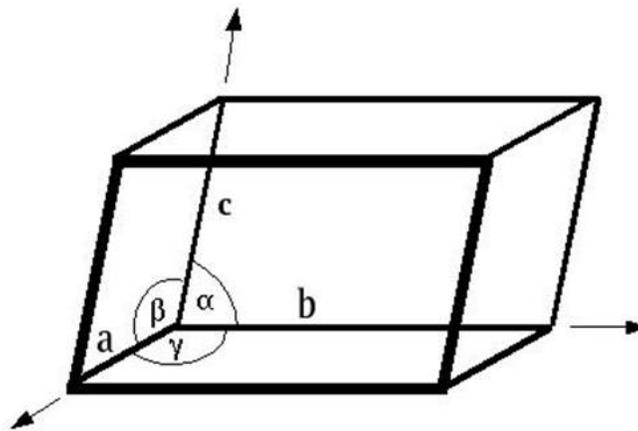


Figure 22: Triclinic unit cell, with edges **a**, **b**, and **c** and angles α , β , and γ .

When collecting the data, the crystal is placed in the path of a narrow beam of X-rays. The detector is an X-ray film that shows dark spots when developed. These dark spots are referred to as reflections because they emerge from the crystals as if reflected from the planes of the atoms. Crystals diffract the source beam into many discrete beams, each of which produces a distinct reflection on the film. The darker the reflection, the greater the intensity of the X-ray beams that reaches a particular position of the crystal. The positions and intensities of these reflections contain the information needed to determine molecular structures.

In the imaginary three-dimensional space of the diffraction pattern, each reflection can be assigned three indices (coordinates). This space is known as the reciprocal space. In crystallography h , k and l are used to designate each reflection in reciprocal space (**Figure 23**).^{89 91} The index h identifies the number of planes in the set per unit cell in the x direction, or, equivalently the number of sections that cut the \mathbf{a} edge of each unit cell. The indices k and l tell how many of such planes exist in the y and z directions per unit cell (**Figure 24**).⁸⁹

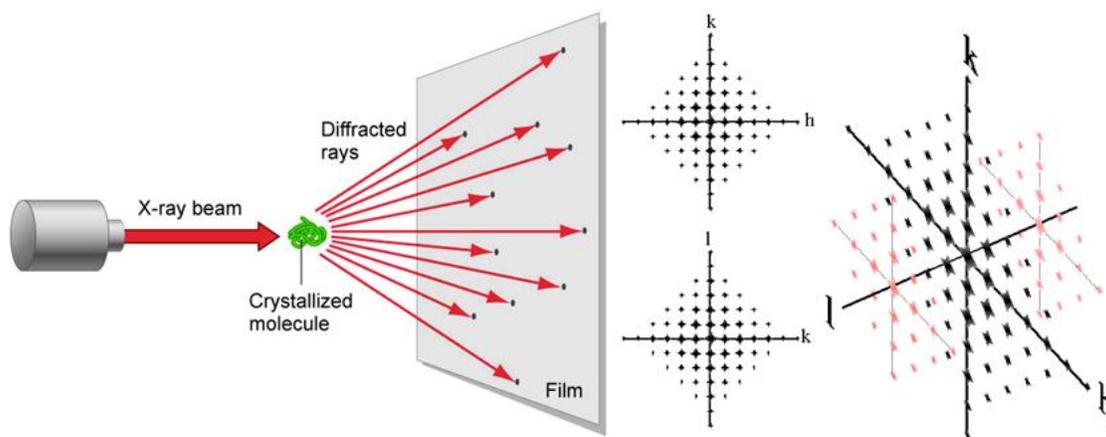


Figure 23: Diagrams illustrating a crystal diffracting the source beam into many discrete beams, each of which produces a distinct reflection on the film and the hkl designation of each reflection.

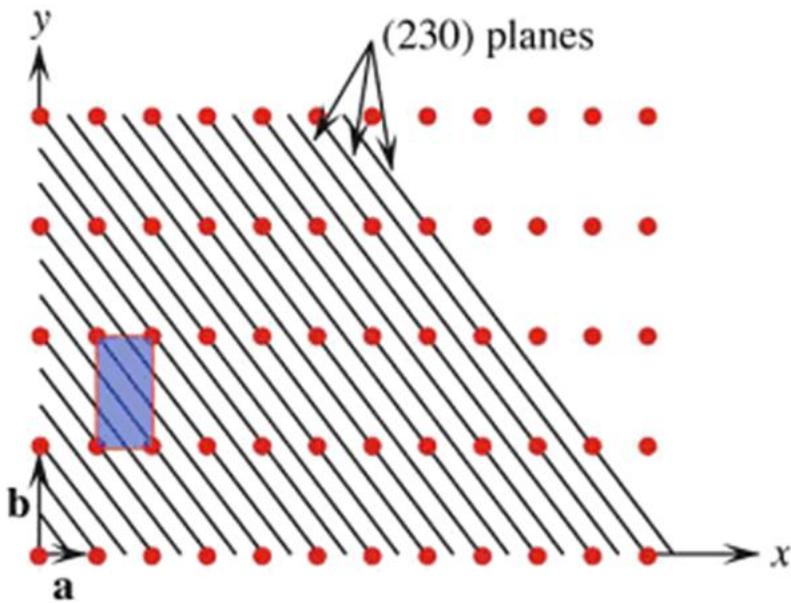
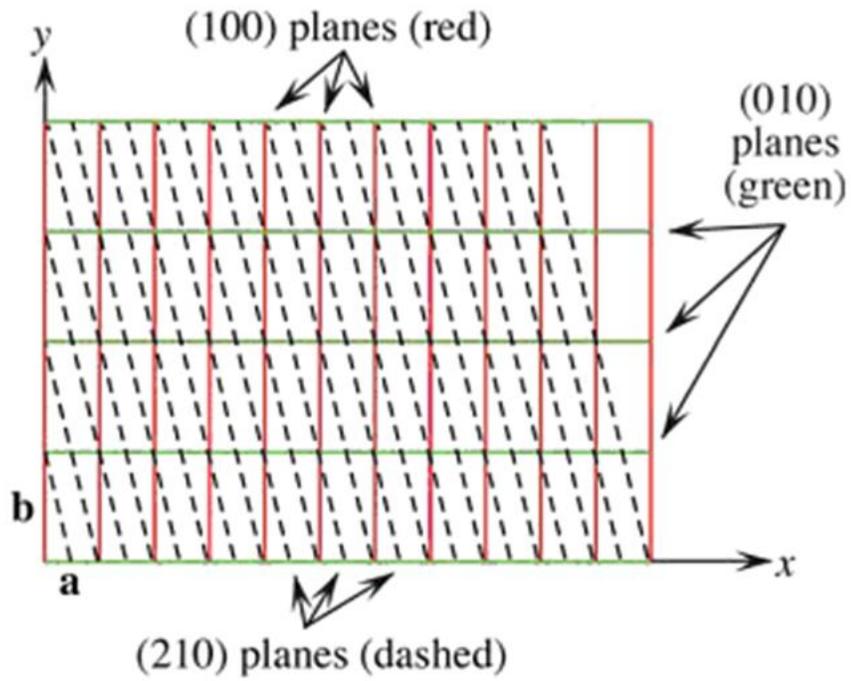


Figure 24: **a.** (210) planes in a two-dimensional section of the lattice. **b.** (230) planes in a two dimensional section of the lattice.

In the figures above, the different sets of equivalent parallel planes have different interplanar spacing. This spacing is denoted as d . The interplanar spacing decreases with an increase in any of the indexes; that is to say the more planes within the unit cell the closer the planes are to each other. W.L. Bragg showed that for a set of planes with index hkl and interplanar spacing d_{hkl} produces a diffracted beam when bombarded with X-rays of a wavelength λ at an angle of θ are reflected at the same angle if θ meets the following condition; where n is an integer.

$$2d_{hkl} \sin \theta = n\lambda$$

Bragg's conditions may be applied to reciprocal lattice which is the space occupied by the reflections. The higher the index of a reciprocal lattice point, the larger the diffraction angle. The larger the diffraction angles the higher the resolution.⁸⁹

The diffractions are characterized by the structure factor, \mathbf{F}_{hkl} , which is a mathematical function describing the phase and amplitude of the diffracted waves. The diffraction data is then taken through a series of computations (Fourier Sum), which generates the electron density map. The electron density can be described as the Fourier transform of the structure factors, \mathbf{F}_{hkl} , that can be used to convert the crystallographic data into an image of the unit cell (**Figure 25**).^{93 94}

Since \mathbf{F}_{hkl} is a function that possesses amplitude, frequency (which are correlated to the intensities of the reflections) as well as phase, the phase problem has to be addressed. The phase of \mathbf{F}_{hkl} cannot be obtained from the measurement of the intensity of a single reflection; therefore to obtain the electron density map, the phase of every diffracted ray must be determined.

The solution to the phase problem is the use of Multi- wavelength Anomalous Diffraction (MAD) method. In this method the electrons of the atoms absorb X-rays of particular

wavelengths. The electrons would then reemit X-rays after some delay. This delay would result in a phase shift of all the reflections, which is known as the anomalous dispersion. The solution for the phases can then be determined from the analysis of the phase shift.⁹³

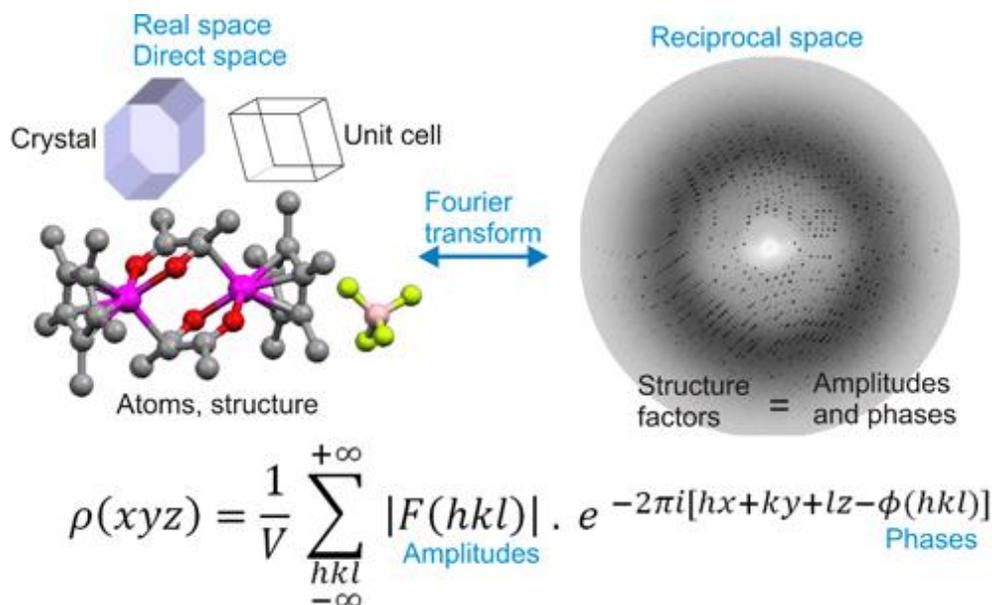


Figure 25: An illustration of the relationship between the diffraction data and the structure.⁹⁴

REFERENCES

REFERENCES

1. Wade, N. J. (1998) *A Natural History of Vision*, Bradford Books.
2. Woodbridge, F. J. E. (1983) *Aristotle's Vision of Nature*, Greenwood Press Reprint.
3. Kepler, J.; Donahue, W. H. (2000) *Optics* Green Lion Press.
4. Newton, I.; Cohen, B. J.; Einstein A.; Whittaker, E. (1952) *Opticks: Or a Treatise of the reflections, refractions inflections and colours of light-based on the Fourth Edition London, 1730*, Dover Publications.
5. Crescitelli, F. (1977) Freidrich William Kuhne. The Centennial of Rhodopsin. *Arch. Ophthalamol.* 95, 1766.
6. Wald, G.; Brown, P. K. Human Rhodopsin. *Science* 1958, 127, 222-226.
7. <http://www.familyconnect.org/images/parentsite/HumanEyeDiagram.jpg>
8. (a) Hubbard, R., Cis-Trans Isomers of Vitamin-a and Retinene in Rhodopsin System. *Fed Proc* 1952, 11(1), 233-233; (b) Hubbard, R.; Wald, G., Cis-Trans Isomers of Vitamin-A and Retinene in the Rhodopsin System. *J Gen Physiol* 1952, 36 (2), 269-31 5; (c) Hubbard, R.; Wald, G., Cis-Trans Isomers Vitamin-A and Retinene in Vision. *Science* 1952, 115 (2977), 60-63; (d) Dieterle, J. M.; Robeson, C. D., Crystalline Neoretinene-B. *Science* 1954, 120 (3110), 219-220.
9. Wang, J. K.; Mcdowell, J. H.; Hargrave, P. A., Site of Attachment of 11-Cis-Retinal in Bovine Rhodopsin. *Biochemistry* 1980, 19(22), 5111-5117.
10. Wald, G., The Biochemistry of Vision. *Annu Rev Biochem* 1953, 22, 497-526.
11. Shichida, Y.; Matsuyama, T., Evolution of opsins and phototransduction. *Philosophical Transactions of the Royal Society B-Biological Sciences* 2009, 364 (1531), 2881-2895.
12. <http://www.fz-juelich.de/inb/inb-1/Photoreception/>.
13. <http://webvision.med.utah.edu/photo1.html>
14. Pugh, E. N. Rods are rods and cones cones, and (never) the twian shall meet. *Neuron* 2001, 32, 375-376.
15. Hoffmann, R. *The Same and not the same.* 1995 Columbia University Press, New York, NY.

16. Waddell, W. H.; Crouch, R.; Nakanishi, K.; Turro, N. J., Quantitative Aspects of Photochemistry of Isomeric Retinals and Visual Pigments. *J. Am. Chem. Soc.* **1976**, *98* (14), 4189-4192.
17. Stryer, L.; Hurley, J. B.; Fung, B. K. K., Transducin - an Amplifier Protein in Vision. *Trends Biochem Sci* **1981**, *6* (9), 245-247.
18. Hofmann, K. P.; Scheerer, P.; Hildebrand, P. W.; Choe, H. W.; Park, J. H.; Heck, M.; Ernst, O. P., A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem Sci* **2009**, *34* (11), 540-552.
19. (a) Vuong, T. M.; Chabre, M.; Stryer, L., Millisecond Activation of Transducin in the Cyclic-Nucleotide Cascade of Vision. *Nature* **1984**, *311* (5987), 659-661; (b) Guy, P. M.; Koland, J. G.; Cerione, R. A., Rhodopsin-Stimulated Activation Deactivation Cycle of Transducin - Kinetics of the Intrinsic Fluorescence Response of the Alpha-Subunit. *Biochemistry* **1990**, *29* (30), 6954-6964.
20. Salon, J. A.; Lodowski, D. T.; Palczewski, K., The Significance of G Protein-Coupled Receptor Crystallography for Drug Discovery. *Pharmacol Rev* **2011**, *63* (4), 901-937.
21. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M., Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289* (5480), 739-745.
22. Bennet, N.; Michelvillaz, M.; Kuhn, H. Light-Induced interaction between rhodopsin and the Gtp-binding Protein Metarhodopsin-II is the major photoproduct involved. *Eur. J. Biochem* **1982**, *127*, 97-103.
23. Borhan, B.; Souto, M. L.; Imai, H.; Shichida, Y.; Nakanishi, K. Movement of retinal along the visual transduction path. *Science* **2000** *288*, 2209-2212.
24. Emeis, D.; Kuhn, H.; Reichert, J.; Hofmann, K.P. Complex-formation between Metarhodopsin II and GTP binding protein in bovine photoreceptor membranes leads to a shift of the photoproduct equilibrium. *Febs. Lett.* **1982** *143*, 29-34.
25. Kliger, D. S.; Lewis, J.W. Spectral and kinetic characterization of visual pigment photointermediates. *Israel J. Chem.* **1995**, *35*, 289-307.
26. Kliger, D. S.; Lewis, J.W.; Szundi, I. Time resolved spectroscopy of visual pigment intermediates. *Abstr. Pap. Am. Chem.* **1995** *S 217*, U323-U323.
27. Souto, M. L.; Borhan, B.; Nakanishi, K. Low-temperature photoaffinity labeling of rhodopsin and intermediates along transduction path. *Vertebrate Phototransduction and the Visual Cycle* **2000**, *Pt B 316*, 425-435.
28. Collins, F. D. Rhodopsin and the yellow indicator. *Nature* **1953**, *171*, 469-472.

29. Morton, R. A.; Pitt, G.A.J. Rhodopsin. IX. pH and the hydrolysis of indicator yellow. *Biochem. J.* **1955**, *59*, 128-134.
30. Rando, R. R. Polyenes and Vision. *Chem. Biol.* **1996**, *3*, 255-262.
31. Shicida, Y.; Matuoka, S.; Yoshizawa, T. Formation of photorhodopsin, a precursor of bathorhodopsin, detected by picosecond laser photolysis at room temperature. *Photobioch Photobiop.* **1984**, *7*, 221-228.
32. Shieh, T.; Han, M.; Sakmar, T.P.; Smith, S.O. The steric trigger in rhodopsin activation. *J Mol. Biol.* **1997**, *269*, 373-384.
33. Yan, B.; Nakanishi, K.; Spudich, J.L. Mechanism of activation of sensory rhodopsin I- Evidence of steric trigger. *P Natl Acad Sci USA* **1991**, *88*, 9412-9416.
34. Nakamichi, H.; Okada, T., Crystallographic analysis of primary visual photochemistry. *Angewandte Chemie-International Edition* **2006**, *45* (26), 4270-4273.
35. Choe, H.-W.; Kim, Y. J.; Park, J. H.; Morizumi, T.; Pai, E. F.; Krauss, N.; Hofmann, K. P.; Scheerer, P.; Ernst, O. P., Crystal structure of metarhodopsin II. *Nature* **2011**, *471* (7340), 651 - U137.
36. Standfuss, J.; Edwards, P. C.; D'Antona, A.; Fransen, M.; Xie, G.; Oprian, D. D.; Schertler, G. F. X., The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* **2011**, *471* (7340), 656-660.
37. Imai, H.; Terakita, A.; Tachibanaki, S.; Iriamoto, Y.; Yoshizawa, T.; Shichida, V., Photochemical and biochemical properties of chicken blue-sensitive cone visual pigment. *Biochemistry* **1997**, *36* (42), 12773-12779.
38. Travis, G. H.; Golczak, M.; Moise, A. R.; Palczewski, K., Diseases caused by defects in the visual cycle: Retinoids as potential therapeutic agents. In *Annual Review of Pharmacology and Toxicology*, 2007; Vol. 47, pp 469-512.
39. Kiser, P. D.; Golczak, M.; Lodowski, D. T.; Chance, M. R.; Palczewski, K., Crystal structure of native RPE65, the retinoid isomerase of the visual cycle. *P Natl Acad Sci USA* **2009**, *106* (41), 17325-17330.
40. Wald, G., Human Vision and the Spectrum. *Science* **1945**, *101* (2635), 653-658.
41. Brown, P. K.; Wald, G., Visual Pigments in Single Rods + Cones of Human Retina - Direct Measurements Reveal Mechanisms of Human Night +Color Vision. *Science* 1964, *144* (361), 45.
42. Kleinschmidt, J.; Harosi, F. I., ANION SENSITIVITY AND SPECTRAL TUNING OF CONE VISUAL PIGMENTS INSITU. *P Natl Acad Sci USA* **1992**, *89* (19), 9181-9185.

43. Suzuki, T.; Makinotasaka, M., Analysis of Retinal and 3-Dehydroretinal in the Retina by High-Pressure Liquid-Chromatography. *Anal Biochem* **1983**, *129* (1), 111-119.
44. Crescitelli, F., THE NATURAL-HISTORY OF VISUAL PIGMENTS -1990. *Progress in Retinal Research* **1991**, *11*, 1-32.
45. Tsin, A. T. C.; Alvarez, R. A.; Fong, S. L.; Bridges, C. D. B., Conversion of Retinol to 3,4-Didehydroretinol in the Tadpole. *Comp Biochem PhysB* **1985**, *81* (2), 415-421.
46. Motto, M. G.; Sheves, M.; Tsujimoto, K.; Baloghnaier, V.; Nakanishi, K., Opsin Shifts in Bovine Rhodopsin and Bacteriorhodopsin - Comparison of 2 External Point-Charge Models. *J Am Chem Soc* **1980**, *102* (27), 7947-7949.
47. Blatz, P. E.; Baumgart, N.; Balasubbr.V; Balasubbr.P; Stedman, E., WAVELENGTH REGULATION IN VISUAL PIGMENT CHROMOPHORE-LARGE INDUCED BATHOCROMIC SHIFTS IN RETINOL AND RELATED POLYENES. *Photochemistry and Photobiology* **1971**, *14* (4), 531.
48. Blatz, P. E.; Mohler, J. H.; Navangul, H. V., ANION-INDUCED WAVELENGTH REGULATION OF ABSORPTION MAXIMA OF SCHIFF-BASES OF RETINAL. *Biochemistry* **1972**, *11* (5), 848.
49. Honig, B.; Greenberg, A. D.; Dinur, U.; Ebrey, T. G., VISUAL-PIGMENT SPECTRA - IMPLICATIONS OF PROTONATION OF RETINALSCHIFF-BASE. *Biochemistry* **1976**, *15* (21), 4593-4599.
50. Andersen, L. H.; Nielsen, I. B.; Kristensen, M. B.; El Ghazaly, M. O. A.; Haacke, S.; Nielsen, M. B.; Petersen, M. A., Absorption of Schiff-base retinal chromophores in vacuo. *J Am Chem Soc* **2005**, *127* (35), 12347-12350.
51. Rajput, J.; Rahbek, D. B.; Andersen, L. H.; Hirshfeld, A.; Sheves, M.; Altoe, P.; Orlandi, G.; Garavell, M., Probing and Modeling the Absorption of Retinal Protein Chromophores in Vacuo. *Angewandte Chemie-International Edition* **2010**, *49* (10), 1790-1793.
52. Livnah, N.; Sheves, M., Model Compounds Can Mimic Spectroscopic Properties of Bovine Rhodopsin. *J Am Chem Soc* **1993**, *115* (1) 351-353.
53. Rosenber. B; Krigas, T. M., SPECTRAL SHIFTS IN RETINAL SCHIFF BASE COMPLEXES. *Photochemistry and Photobiology* **1967**, *6* (10), 769.
54. Baasov, T.; Sheves, M., ALTERATION OF PKA OF THE BACTERIORHODOPSIN PROTONATED SCHIFF-BASE - A STUDY WITH MODEL COMPOUNDS. *Biochemistry* **1986**, *25* (18), 5249-5258.

55. Honig, B.; Dinur, U.; Nakanishi, K.; Baloghnaïr, V.; Gawinowicz, M.A.; Arnaboldi, M.; Motto, M. G., External Point-Charge Model for Wavelength Regulation in Visual Pigments. *J Am Chem Soc* **1979**, *101* (23), 7084-7086.
56. Nakanishi, K.; Baloghnaïr, V.; Arnaboldi, M.; Tsujimoto, K.; Honig, B., An External Point-Charge Model for Bacteriorhodopsin to Account for Its Purple Color. *J Am Chem Soc* **1980**, *102* (27), 7945-7947.
57. Beppu, Y.; Kakitani, T., Theoretical-Study of Color Control Mechanism in Retinal Proteins. 1. Role of the Tryptophan Residue, Tyrosine Residue and Water Molecule. *Photochemistry and Photobiology* **1994**, *59* (6), 660-669.
58. Luecke, H.; Schobert, B.; Lanyl, J. K.; Spudich, E. N.; Spudich, J. L., Crystal structure of sensory rhodopsin II at 2.4 angstroms: Insights into color tuning and transducer interaction. *Science* **2001**, *293* (5534), 1499-1503.
49. Luecke, H.; Schobert, B.; Richter, H. T.; Cartailier, J. P.; Lanyl, J.K., Structure of bacteriorhodopsin at 1.55 angstrom resolution. *Journal of Molecular Biology* **1999**, *291* (4), 899-911.
60. (a) Irving, C. S.; Byers, G. W.; Leermake, Pa, Effect of Solvent Polarizability on Absorption Spectrum of All-Trans-Retinylypyrrolidiniminium Perchlorate. *J Am Chem Soc* **1969**, *91* (8), 2141; (b) Irving, C. S.; Byers, G. W.; Leermake, Pa, Spectroscopic Model for Visual Pigments - Influence of Microenvironmental Polarizability. *Biochemistry* **1970**, *9* (4), 8584.
61. Mathies, R.; Stryer, L., RETINAL HAS A HIGHLY DIPOLAR VERTICALLY EXCITED SINGLET-STATE - IMPLICATIONS FOR VISION. *P Natl Acad Sci USA* **1976**, *73* (7), 2169-2173.
62. Baasov, T.; Sheves, M., Model Compounds for the Study of Spectroscopic Properties of Visual Pigments and Bacteriorhodopsin. *J Am Chem Soc* **1985**, *107* (25), 7524-7533.
63. Nathans, J.; Thomas, D.; Hogness, D. S., MOLECULAR-GENETICS OF HUMAN COLOR-VISION - THE GENES ENCODING BLUE, GREEN, AND RED PIGMENTS. *Science* **1986**, *232* (4747), 193-202.
64. Neitz, M.; Neitz, J.; Jacobs, G. H., SPECTRAL TUNING OF PIGMENTS UNDERLYING RED-GREEN COLOR-VISION. *Science* **1991**, *252* (2008), 971-974.
65. Asenjo, A. B.; Rim, J.; Oprian, D. D., MOLECULAR DETERMINANTS OF HUMAN RED/GREEN COLOR DISCRIMINATION. *Neuron* **1994**, *12* (5), 1131-1138.
66. Lin, S. W.; Kochendoerfer, G. G.; Carroll, H. S.; Wang, D.; Mathies, R. A.; Sakmar, T. P., Mechanisms of spectral tuning in blue cone visual pigments- Visible and raman spectroscopy of blue-shifted rhodopsin mutants. *Journal of Biological Chemistry* **1998**, *273* (38), 24583-24591.

67. Nakayama, T. A.; Zhang, W. M.; Cowan, A.; Kung, M., Mutagenesis studies of human red opsin: Trp-281 is essential for proper folding and protein- retinal interactions. *Biochemistry* **1998**, *37* (50), 17487-17494.
68. Zhukovsky, E. A.; Oprian, D. D., EFFECT OF CARBOXYLIC-ACID SIDE-CHAINS ON THE ABSORPTION MAXIMUM OF VISUAL PIGMENTS. *Science* **1989**, *246* (4932), 928-930.
69. Sakmar, T. P.; Franke, R. R.; Khorana, H. G., THE ROLE OF THE RETINYLIDENE SCHIFF-BASE COUNTERION IN RHODOPSIN IN DETERMINING WAVELENGTH ABSORBENCY AND SCHIFF-BASE PKA. *P Natl Acad Sci USA* **1991**, *88* (8), 3079-3083.
70. (a) Lin, J. Y., A user's guide to channelrhodopsin variants: features, limitations and future developments. *Exp Physiol* **2011**, *96* (1), 19-25; (b) Suzuki, D.; Irieda, H.; Homma, M.; Kawagishi, I.; Sudo, Y., Phototactic and Chemotactic Signal Transduction by Transmembrane Receptors and Transducers in Microorganisms. *Sensors-Basel* **2010**, *10* (4), 4010-4039; (c) Purcell, E. B.; Crosson, S., Photoregulation in prokaryotes. *Curr Opin Microbiol* **2008**, *11* (2), 168-178; (d) Sasaki, J.; Spudich, J. L., Signal transfer in haloarchaeal sensory rhodopsin-transducer complexes. *Photochemistry and Photobiology* **2008**, *84* (4), 863-868.
71. Fu, H.-Y.; Lin, Y.-C.; Chang, Y.-N.; Tseng, H.; Huang, C.-C.; Liu, K.-C.; Huang, C.-S.; Su, C.-W.; Weng, R. R.; Lee, Y.-Y.; Ng, W. V.; Yang, C.-S., A novel six-rhodopsin system in a single archaeon. *J. Bacteriol.* **2010**, *192* (22), 5866-5873.
72. Kolbe, M.; Besir, H.; Essen, L. O.; Oesterhelt, D., Structure of the light-driven chloride pump halorhodopsin at 1.8 angstrom resolution. *Science* **2000**, *288* (5470), 1390-1396.
73. (a) Vogeley, L.; Sineshchekov, O. A.; Trivedi, V. O.; Sasaki, J.; Spudich, J. L.; Luecke, H., Anabaena sensory rhodopsin: A photochromic color sensor at 2.0 angstrom. *Science* **2004**, *306* (5700), 1390-1393; (b) PebayPeyroula, E.; Rummel, G.; Rosenbusch, J. P.; Landau, E. M., X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science* **1997**, *277* (5332), 1676-1681.
74. Mogi, T.; Marti, T.; Khorana, H. G., STRUCTURE-FUNCTION STUDIES ON BACTERIORHODOPSIN 9 SUBSTITUTIONS OF TRYPTOPHAN RESIDUES AFFECT PROTEIN-RETINAL INTERACTIONS IN BACTERIORHODOPSIN. *Journal of Biological Chemistry* **1989**, *264* (24), 14197-14201.
75. Greenhalgh, D. A.; Farrens, D. L.; Subramaniam, S.; Khorana, H. G., HYDROPHOBIC AMINO-ACIDS IN THE RETINAL-BINDING POCKET OF BACTERIORHODOPSIN. *Journal of Biological Chemistry* **1993**, *268* (27), 20305-20311.
76. Subramaniam, S.; Marti, T.; Khorana, H. G., PROTONATION STATE OF ASP (GLU)-85 REGULATES THE PURPLE-TO-BLUE TRANSITION IN BACTERIORHODOPSIN MUTANTS ARG-82- ALA AND ASP-85- GLU – THE BLUE FORM IS INACTIVE IN PROTON TRANSLOCATION. *P Natl Acad Sd USA* **1990**, *87* (3), 1013-1017.

77. Shimono, K.; Iwamoto, M.; Sumi, M.; Kamo, N., Effects of three characteristic amino acid residues of Pharaonis phoborhodopsin on the absorption maximum. *Photochemistry and Photobiology* **2000**, *72* (1), 141-145.
78. Wada, A.; Tsutsumi, M.; Inatomi, Y.; Imai, H.; Shichida, Y.; Ito, M., Retinoids and related compounds. Part 26. Synthesis of (11 Z)-8,18-propano- and methano-retinals and conformational study of the rhodopsin chromophore. *Journal of the Chemical Society-Perkin Transactions 1* **2001**, (19), 2430-2439.
79. Sakurai, M.; Sakata, K.; Saito, S.; Nakajima, S.; Inoue, Y., Decisive role of electronic polarization of the protein environment in determining the absorption maximum of halorhodopsin. *J Am Chem Soc* **2003**, *125* (10), 3108-3112.
80. Cembran, A.; Gonzalez-Luque, R.; Altoe, P.; Merchan, M.; Bernardi, F.; Olivucci, M.; Garavelli, M., Structure, spectroscopy, and spectral tuning of the gas-phase retinal chromophore: The beta-Ionone "handle" and alkyl group effect. *J Phys Chem A* **2005**, *109* (29), 6597-6605.
81. Trabanino, R. J.; Vaidehi, N.; Goddard, W. A., III, Exploring the molecular mechanism for color distinction in humans. *Journal of Physical Chemistry B* **2006**, *110* (34), 17230-17239.
82. Sekharan, S.; Sugihara, M.; Buss, V., Origin of spectral tuning in rhodopsin - It is not the binding pocket. *Angewandte Chemie-International Edition* **2007**, *46* (1-2), 269-271.
83. Sekharan, S.; Weingart, O.; Buss, V., Ground and excited states of retinal Schiff base chromophores by multiconfigurational perturbation theory. *Biophysical Journal* **2006**, *91*(1), L7-L9.
84. Derguini, F.; Caldwell, C. G.; Motto, M. G.; Baloghnaïr, V.; Nakanishi, K., Bacteriorhodopsins Containing Cyanine Dye Chromophores - Support for the External Point-Charge Model. *J Am Chem Soc* **1983**, *105* (3), 646-648.
85. Kloppmann, E.; Becker, T.; Ullmann, G. M., Electrostatic potential at the retinal of three archaeal rhodopsins: Implications for their different absorption spectra. *Proteins* **2005**, *61*(4), 953-965.
86. Lee, K. S. S. Michigan State University, East Lansing, 2011
87. Wang, W. Michigan State University, East Lansing, 2012
88. http://en.wikipedia.org/wiki/Electromagnetic_spectrum
89. Rhodes, G., (2006) *Crystallography made Crystal Clear: A guide for users of macromolecular models. 3rd Edition*. Academic Press.
90. http://www.rcsb.org/pdb/education_discussion/look-at-structures/images/biol_unit/unit.png

91. <http://www.roywmacdonald.com/siteimages/Crystal%20System%20Diagram.jpg>
92. <http://www.amanzi.org/craig/diagrams/recip.gif>
93. Giacovazzo, C., *Phasing in Crystallography: A Modern Perspective*. International Union of Crystallography. *Oxford Science Publications*.
94. www.xtal.iqfr.csic.es

CHAPTER II. ENGINEERING AND CRYSTALLIZATION OF THE 2ND

GENERATION RHODOPSIN MIMIC, CRBPII

II.1 AN INTRODUCTION: THE 1ST GENERATION RHODOPSIN MIMIC

Interest in the phenomenon of protein-engendered wavelength regulation was inspired by the fact that a single chromophore, 11-*cis*-retinal, is responsible for the observation of a wide range of absorptions from 420 nm to 570 nm.^{1 2 3} Studying these visual pigments provides a good platform for probing protein-chromophore interactions that could lead to the different absorption maxima observed. Several different approaches were taken to try to dissect the factors that govern wavelength regulation. As discussed in the previous chapter, site directed mutagenesis studies of the rhodopsins were done in order to determine the factors that are responsible for the opsin shift. Irrational screening methods have been used to select for rhodopsin mutants with different absorptions.⁴ The conclusions from these studies indicate that the residues surrounding the chromophore as well as the residues outside the binding pocket could affect the spectral profile of the visual pigments. However, difficulties associated with handling membrane bound proteins have resulted in the lack of crystallographic data of these mutants.

Therefore using rhodopsin, a membrane bound protein, to study the intricacies of wavelength regulation is not sensible as it is not amenable to routine mutagenesis or crystallization conditions and the expression is rather poor. Even though several studies have been done using different rhodopsin mutants, the lack of crystallographic data makes it very difficult to substantiate the hypotheses put forward. As a result of this, our lab used an appropriate model system to elucidate the factors responsible for wavelength regulation.

The goal was to develop a model protein system that binds retinal as a protonated Schiff base, thus mimicking rhodopsin. To this end, Cellular Retinoic Acid Binding Protein II (CRABP II) was chosen as the rhodopsin substitute. CRABP II is a small cellular protein that belongs to the family of intracellular lipid binding proteins (iLBP's). It consists of 137 amino acid residues. It forms a 10 β -sheet barrel and has a large binding pocket for its ligand all-*trans*-retinoic acid (**Figure 26**).⁵

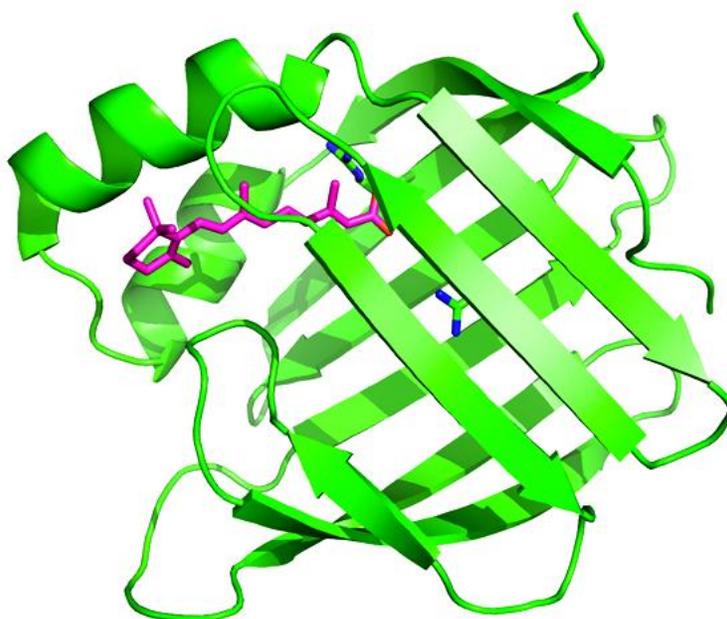


Figure 26: Crystal structure of WT-CRABP II with bound retinoic acid (PDB code: 1CBS).

CRABP II is soluble, easily expressed and purified in large quantities. CRABP II is also amenable to routine mutagenesis and so provides a platform for protein design. Since 11-*cis* retinal is very light sensitive, all-*trans*-retinal was used instead. This should not pose a problem since both isomers share a similar mechanism in relation to spectral tuning.

All-*trans*-retinoic acid and all-*trans*-retinal are structurally similar, but their affinities for CRABP II are very different. The dissociation constant for all-*trans*-retinal is 6 μ M while the

dissociation constant of all-*trans*-retinoic acid is 1nM.⁶ The crystal structure of WT-CRABPII bound to all-*trans*-retinoic acid shows that the key interaction is a salt bridge with Arg 132 and also Arg 111 with the aid of a highly ordered water network illustrated in **Figure 27**.

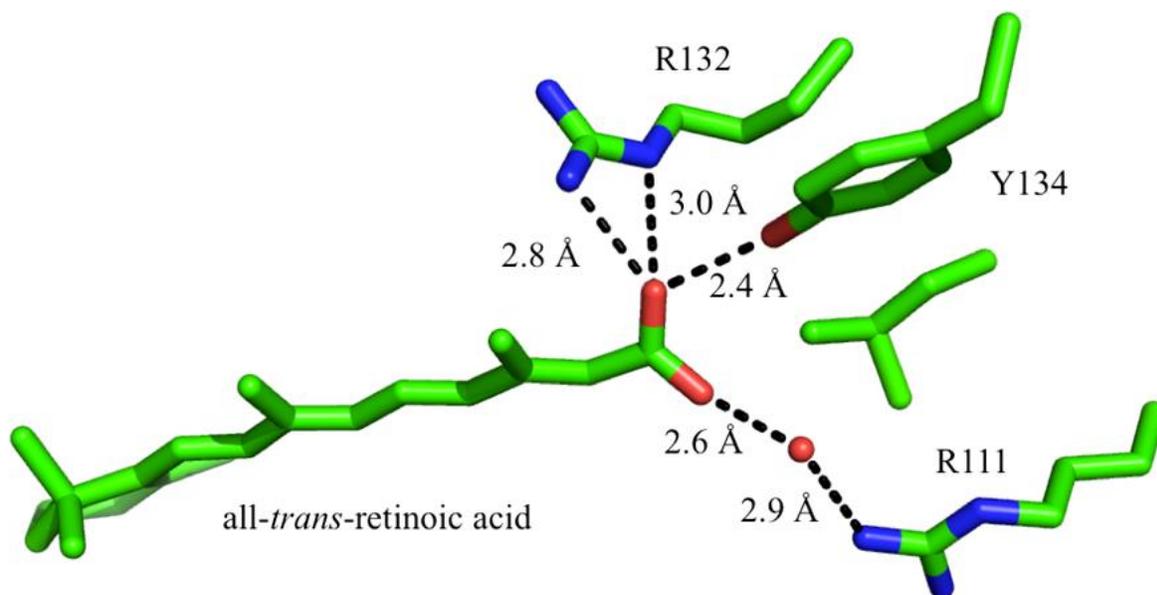


Figure 27: Key interactions of the chromophore, all-*trans*-retinoic acid, in the binding pocket.

The fact that retinal lacks the carboxylate that is essential for binding, explains the low binding affinity observed.

The first step towards the engineering of the rhodopsin mimic was to introduce a lysine residue that would react retinal to form a Schiff base. Initially it was thought that increasing the hydrophobicity in the vicinity of the lysine would ensure Schiff base formation. It was later shown however, that this was not true. The key element was the trajectory of the nucleophilic lysine to the aldehyde of the all-*trans*-retinal. Based on crystallographic studies, Bürgi and Dunitz showed that for a successful nucleophilic attack on a carbonyl, the nucleophile must approach the carbonyl in a trajectory that bisects the plane of the aldehyde. The optimal angle

was determined to be 107° .⁷ This angle represents the maximum overlap of the π^* of the carbonyl and the lone pair of the nucleophile, in this case the amine group of the lysine while minimizing repulsion with the π bonding electron cloud. To this end, Arg 132 was chosen as the site for introducing the lysine residue (**Figure 28**).

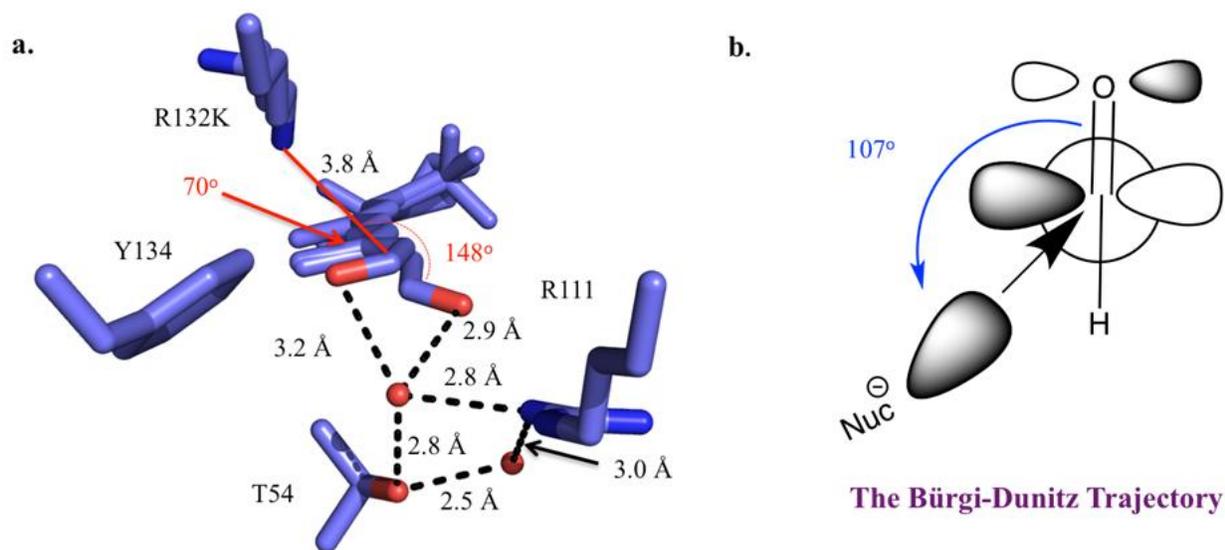


Figure 28: **a.** Hydrogen bonding interactions between R132K:Y134F-CRABPII and all-*trans*-retinal. **b.** The Bürgi-Dunitz trajectory.

The lysine residue was introduced at position 132 and the tyrosine at position 134 was mutated to phenylalanine. The introduction of phenylalanine at position 134 was to make the environment close to the 132 position more hydrophobic as well as to nullify possible interactions with the aldehyde or the lysine. It was observed that the double mutant, R132K:Y134F, did not form a Schiff base. This was further confirmed by crystallographic data. The crystal structure suggested that Arg 111 and Thr 54 are interacting with the aldehyde through a water mediated hydrogen bonding network (**Figure 28a**). These interactions resulted

in the aldehyde adopting an unfavorable conformation thus preventing the nucleophilic attack of the amine group of the lysine.

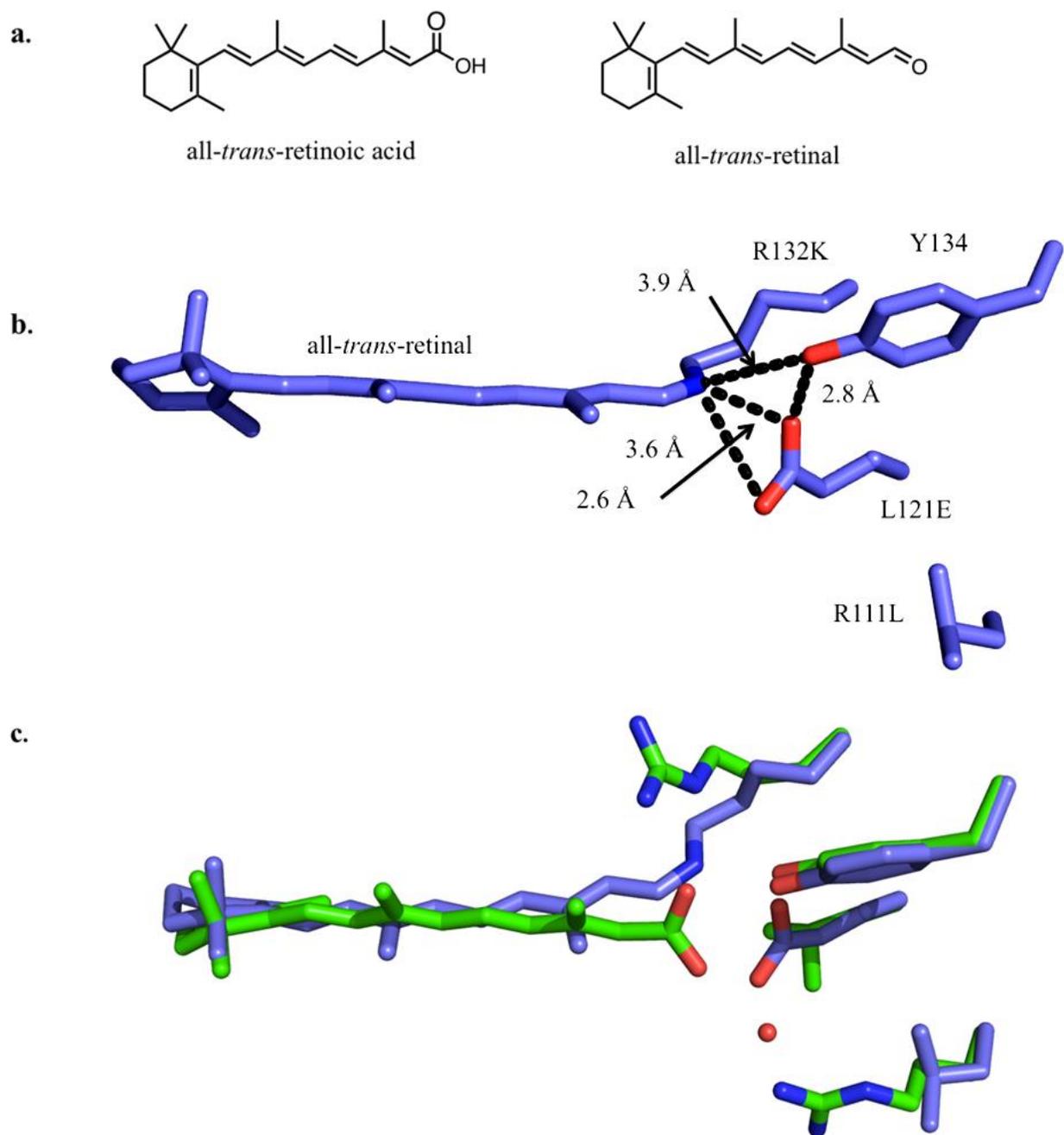


Figure 29: **a.** Structures of all-*trans*-retinoic acid and all-*trans*-retinal. **b.** Crystal structure of R132K:R111L:L121E. **c.** Overlay of the structures of WT-CRABP II-retinoic acid (green, PDB code: 1CBS) and R132K:R111L:L121E – bound to the retinylidene PSB (cyan, PDB code: 2G7B).

These observations suggested that removal of Arg 111 and Thr 54, would allow for the aldehyde to then adopt the favorable conformation for nucleophilic attack. The mutant R132K:R111L was made to disrupt the water mediated hydrogen bonding interaction that resulted in the aldehyde adopting an unfavorable conformation. Even though the water network was disrupted there was still no evidence of PSB formation. This could be due to the lack of stabilization of the PSB is reflected in its low pK_a value. In the rhodopsins, a counterion is present to stabilize the resultant PSB. The crystal structure revealed that the leucine at position 121 would be the ideal position to introduce the counterion, glutamate. These findings led to the design of the triple mutant R132K:R111L:L121E (KLE), which formed a stable PSB upon binding retinal (**Figure 29**). The pK_a for this triple mutant was 8.7, which means that the PSB is stable under physiological conditions. Substitution of the glutamate with polar residues such as glutamine gave similar results to what was observed for the double mutant (R132K:R111L). It can therefore be concluded that the glutamate is indeed the counterion.^{6 8}

II.2 STUDY USING C15 AS THE CHROMOPHORE FOR WAVELENGTH TUNING

With this triple mutant in hand, further mutagenesis studies were done to probe the factors governing wavelength regulation. However, the bound chromophore in the R132K:R111L:L121E (KLE) based mutants did not respond to changes in polarity as shown in **Table 8**.^{6 9} Examination of the crystal structures revealed position 59 to be the optimal position for modulating the polarity around the ionone ring region. The introduction of different residues at this position should result in wavelength modulation. Residues with negative polarity should result in red shift as a result of delocalization of the positive charge along the polyene while the

residues with a positive polarity would result in blue shift. Disappointedly, the incorporation of different residues did not have an effect on the wavelength.

Table 8: CRABP II mutants with all-*trans*-retinal and all-*trans* C15-retinal.

| CRABP II Mutants | λ_{max} with all- <i>trans</i> -retinal | λ_{max} with all- <i>trans</i> C15-retinal |
|------------------------|--|---|
| R132K:R111L:L121E:R59 | 449 nm | n.d |
| R132K:R111L:L121E:R59E | 450 nm | 424 nm |
| R132K:R111L:L121E:R59Q | 444 nm | 413 nm |
| R132K:R111L:L121E:R59L | 443 nm | 391 nm |
| R132K:R111L:L121E:R59W | 442 nm | 404 nm |

The wavelength maxima observed for KLE was 449 nm, while the series of KLE:R59X mutants had wavelength maximum ranging from 442 – 450 nm. Reexamination of the crystal structure of KLE revealed that the ionone ring of the retinal was exposed to the aqueous environment (**Figure 30**). This would nullify any effect of the residues introduced. This phenomenon is in contrast to the typical rhodopsin system where the chromophore is fully embedded inside the binding pocket.

Two solutions were considered to address this issue; the first is to use a shorter retinal analogue as chromophore and the other is to engineer another protein that is able to completely encapsulate the retinal in the binding pocket away from the aqueous environment. Dr. Lee carried out the studies using the C15 retinal analogue and Dr. Jia was able to obtain a crystal structure of the mutant R132K:R111L:L121E:R59W with C15 retinal (**Figure 30 c,d**). The crystal structure shows the C-15 retinal fully embedded in the binding pocket.¹⁰

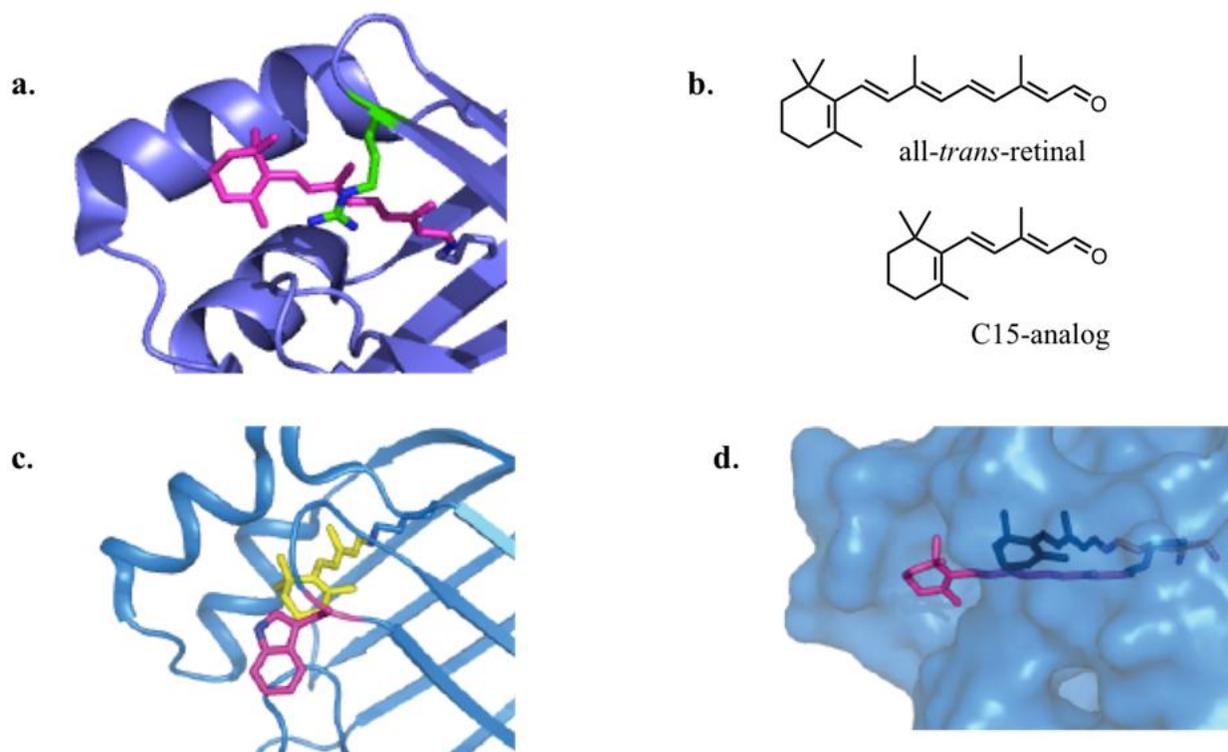


Figure 30: Crystal structure of CRABPII with all-*trans* C15-retinal analogue. **a.** Crystal structure of R132K:R111L:L121E (KLE) with arginine 59 highlighted in green. **b.** The chemical structure of all-*trans*-retinal and all-*trans*-C15-retinal analogue. **c.** Crystal structure of R132K:R111L:L121E:R59W with all-*trans*-C15-retinal analogue with the Tryptophan at position 59 highlighted in yellow. **d.** The overlaid structures of R132K:R111L:L121E (KLE) with all-*trans*-retinal and R132K:R111L:L121E:R59W with all-*trans*-C15-retinal analogue.²¹

The mutants previously discussed (**Table 8**) were bound to the all-*trans*-C15-retinal analogue. The result was that sensitivity of the chromophore to polarity changes was obtained by sequestering the chromophore in the binding pocket. The introduction of residues with negative polarity resulted in red shift. The introduction of tryptophan resulted in red shift apparently due to its polarizable π -electron cloud.¹¹

The all-*trans*-C15-retinal analogue was bound to several other mutants of R59 to prove that sequestering the chromophore is essential for wavelength regulation.

II.3 DEVELOPMENT OF THE 2ND GENERATION RHODOPSIN MIMIC, CRBP II

The studies using CRABP II showed that the chromophore has to be completely buried on the binding pocket, but in order to fully understand the mechanism of wavelength regulation in the visual pigments retinal has to be the chromophore used. This led to the consideration of the protein Cellular Retinol Binding Protein II (CRBP II) as an alternative candidate for a rhodopsin mimic. Human CRBP II is found in the small intestine epithelium. It is thought to participate in the uptake and/or intracellular metabolism of vitamin A. Vitamin A is a fat-soluble vitamin necessary for growth, reproduction, differentiation of epithelial tissues, and vision. CRBP II may also modulate the supply of retinoic acid to the nuclei of endometrial cells during the menstrual cycle.^{12 13} CRBP II binds both retinol and retinal with dissociation constants of 10 nM and 90 nM respectively. The crystal structure of the WT-CRBP II showed that the chromophore is bound deeper in the binding pocket. CRBP II also belongs to the iLBP family. The sequence identity between CRBP II and CRABP II is only 35% yet their backbone structure is very similar. CRBP II has a 10 β -sheet barrel with 2 helices at the ‘mouth’ of the binding pocket, similar to CRABP II. An overlay of the crystal structures of CRABP II and CRBP II revealed that the binding pocket of CRBP II is 5Å deeper than that of CRABP II (**Figure 31**).

The difference in the binding pockets of these two proteins could be linked to their function. CRABP II is responsible for transferring retinoic acid to the protein Retinoic Acid Receptor (RAR), which is a gene regulation protein. In order for the efficient transfer, one would expect that the retinoic acid should not be completely buried in the binding pocket (**Figure 31 b**). Therefore the shallowness and openness of the binding pocket of CRABP II would be essential for function.^{14 15}

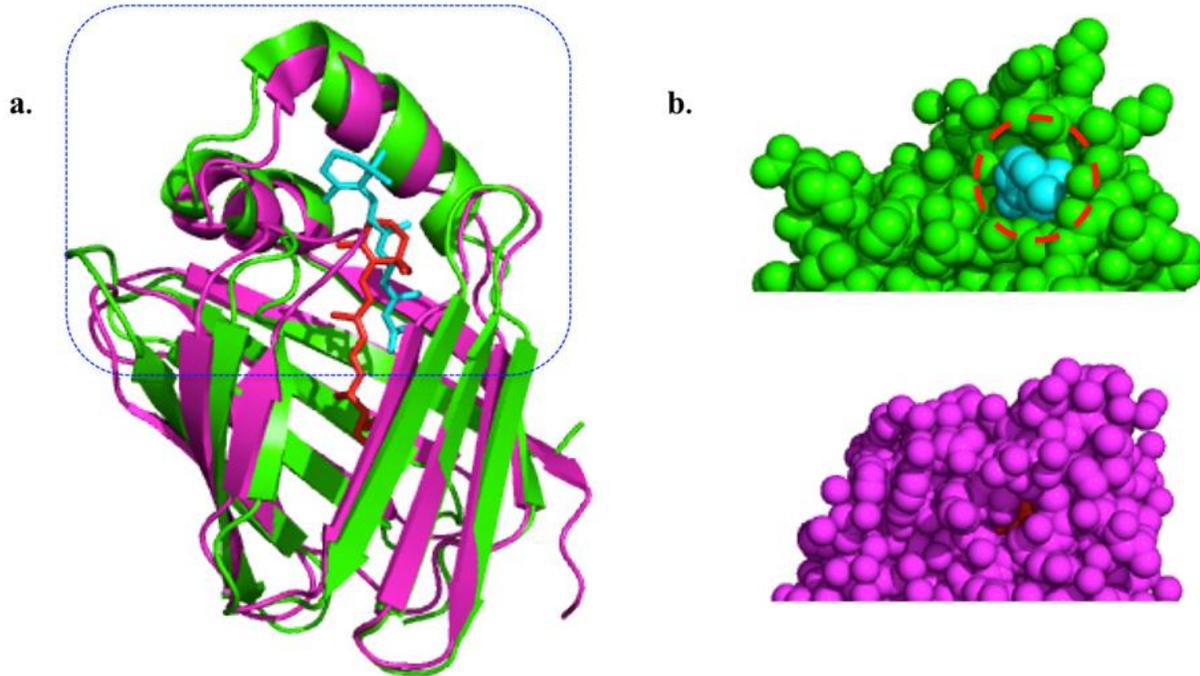


Figure 31: Comparison of WT-CRBPII and WT-CRABPII. **a.** Overlaid crystal structures of WT-CRBPII (purple, 2RCT) bound to all-*trans*-retinol (red) and WT-CRABPII (green, 1CBS) bound to all-*trans*-retinoic acid (cyan). **b.** The space filling model of the region highlighted by the rectangle of CRABPII (top) and CRBPII (bottom) showing the ionone region.

For CRBPII, however, the retinal/retinol has to be completely embedded in the binding pocket. This is to protect the retinal/retinol from oxidation, degradation and isomerization during transfer (**Figure 31b**).

The complete encapsulation of the retinol/retinal in the binding pocket of CRBPII makes it the perfect candidate for reengineering into a rhodopsin mimic.

II.4 ENGINEERING OF THE 2ND GENERATION RHODOPSIN MIMIC, CRBPII

The expression of the WT-CRBPII was done in the *E.coli* pET17b vector system. Dr. Wenjing Wang did the original optimization and characterization of the protein using this expression system. Both tagged and tagless constructs were made, however the tagless pET17b

construct was the construct used. The purification of the protein was done using ion exchange chromatography. The pI of CRBP_{II} is ~5.1 and so an anion exchange resin at pH 8.0 was used. The protein's negative charged at this pH will cause it stick to positively charged column. The expression yields of the WT and most of the mutants are very high (up to 100 mg/L).

The crystal structure of WT-CRBP_{II} bound to all-*trans*-retinol was already available.¹⁶ Dr. Zahra Nossoni repeated the crystallization and was able to get crystal structures of CRBP_{II} bound to all-*trans*-retinol and all-*trans*-retinal (not reported) at a resolution of 1.2Å. These crystal structures were very essential in the reengineering of CRBP_{II} into a rhodopsin mimic.

Overlay of both crystal structures revealed that all-*trans*-retinol and all-*trans*-retinal adopt very similar conformations in the binding pocket (**Figure 32**). The only difference was the orientation of the oxygen atom, which is sp³ hybridized in all-*trans*-retinol and sp² hybridized in all-*trans*-retinal. All the hydrogen bonding interactions observed for the all-*trans*-retinol structure were maintained in the all-*trans*-retinal structure.

As with CRABP_{II} a nucleophilic lysine residue must be introduced to facilitate the formation of a PSB with retinal. The crystal structures revealed three possible sights for the introduction of the lysine; these are positions 51, 106 and 108. Similar in silico studies that were used for CRABP_{II}, were carried out on these three residues to determine which would adopt the correct geometry for formation of the PSB. The results showed that glutamine at position 108 would adopt the correct Bürgi-Dunitz trajectory to attack the aldehyde.

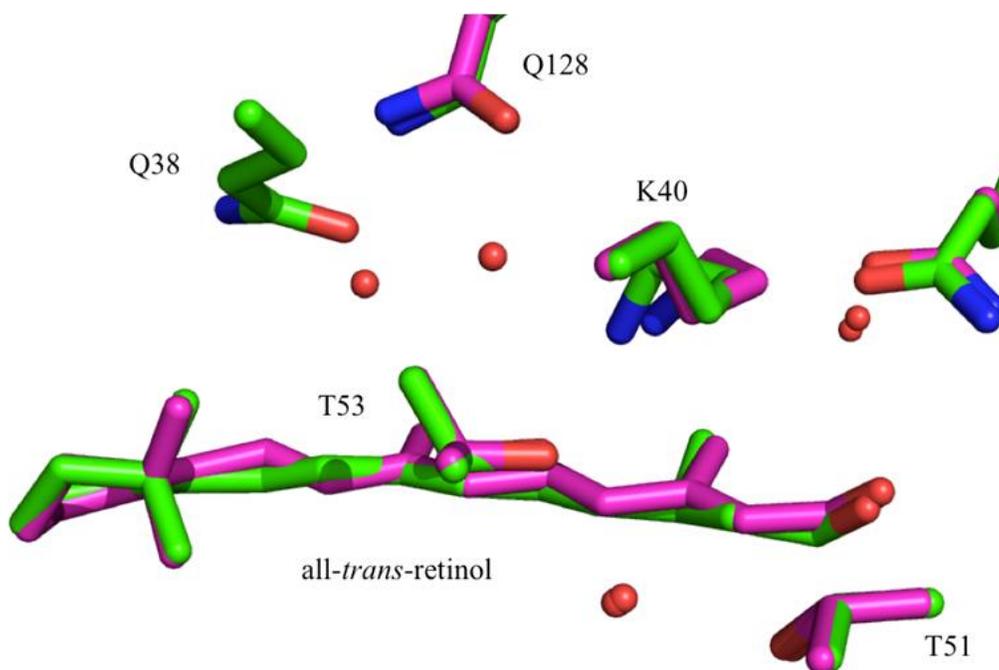


Figure 32: Overlaid crystal structures of WT-CRBPII with all-*trans*-retinol (green) and all-*trans*-retinal (magenta). The hydrogen bonding network is highlighted in both systems.

Although the results showed that the most favorable position was position 108. Dr. Wang did mutations to all three sites T51K, W106K and Q108K. The results showed that mutating T51 and W106 was unfavorable to the stability and solubility of the protein. W106 is a highly conserved residue in the lipid binding protein family and mutation of this position resulted in an unstable protein while with the T51 mutation several expression issues were encountered.¹⁷ The Q108K mutant gave the best results. The UV-vis binding studies of this mutant proved that the protein binds to retinal but the amount of PSB formed (506 nm) was low, most of the bound protein was Schiff base (365 nm). Examination of the crystal structure revealed that lysine at position 40 (K40) could be interfering with the stability of the PSB (**Figure 33**).

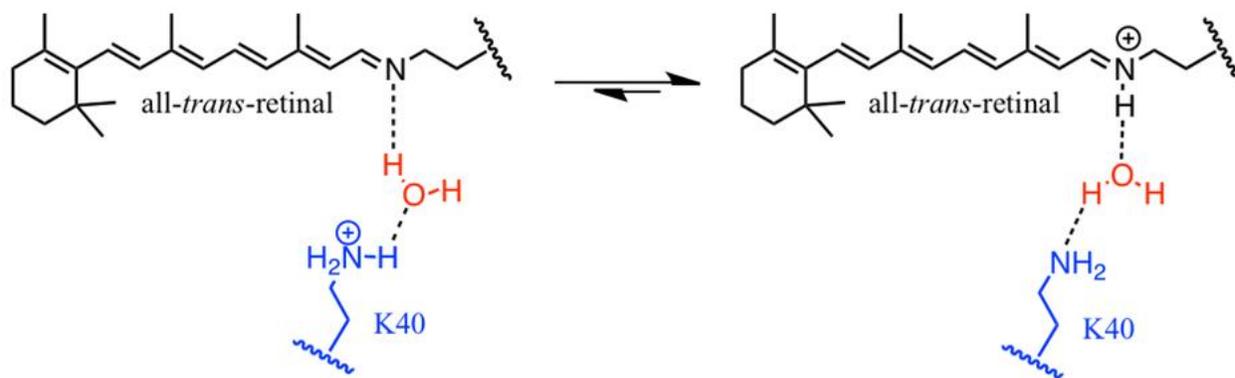


Figure 33: Proposed mechanism for proton transfer in single mutant Q108K.

The two different charged species are in equilibrium with each other with the equilibrium lying towards the right. The pK_a of protonated lysine is ~ 3 pK_a units lower than that of the PSB. This could explain the major absorption peak observed at 365 nm and the minor absorption peak at 506 nm. Therefore, K40 was mutated to leucine to give the double mutant Q108K:K40L (KL) in order to prevent this proton transfer. This mutation rectified the problem and the retinal-bound mutant Q108K:K40L gave a stable PSB with a $pK_a = 8.3$ and a maximum absorption at 508 nm. This double mutant was used as the foundation to begin mutagenesis studies to probe the factors that govern wavelength regulation. Over 200 mutants have been made in our lab and over 30 crystallographic data sets have been obtained.

II.5 CRYSTALLIZATION OF THE 2ND GENERATION RHODOPSIN MIMIC, CRBP_{II}

In order to probe the factors governing wavelength regulation, crystal structures are essential. Encouraged by the ease of crystallization of the WT CRBP_{II} with all-*trans*-retinal, we were able to formulate a series of crystallization conditions that worked for many of the mutants

of CRBP_{II} and some of the 2nd generation CRABP_{II} mutants. These crystal structures provided a wealth of information about some factors that are key to wavelength regulation.

Gratifyingly we were able to obtain the crystal structure of the double mutant Q108K:K40L (KL) (obtained by Dr. Zahra Nossoni) which had electron density confirming the formation of the PSB. Comparison of the double mutant with the WT revealed that conformation of the chromophore is slightly different **Figure 36**. The retinal was a little deeper in the binding pocket than in the WT. This could be a result of the fact that in the case of KL the retinal is covalently bound while in the case of the WT the retinal is held in the binding pocket by hydrogen bonding interactions. The chromophore in both cases adopted a *6-s-trans* conformation. Calculations have shown that the *6-s-trans* conformation is 0.6 kcal higher in energy compared to the *6-s-cis* conformation.¹⁸ The *6-s-trans* conformation observed for CRBP_{II} is as a result of the tighter packing between the chromophore and the protein.

Closer examination of the KL structure revealed the different sites that are in close proximity to the chromophore that can be mutated to determine the effect on the wavelength. The sites that are mutated are highlighted in **Figure 32**. The approach is very similar to that which was taken with CRABP_{II}.

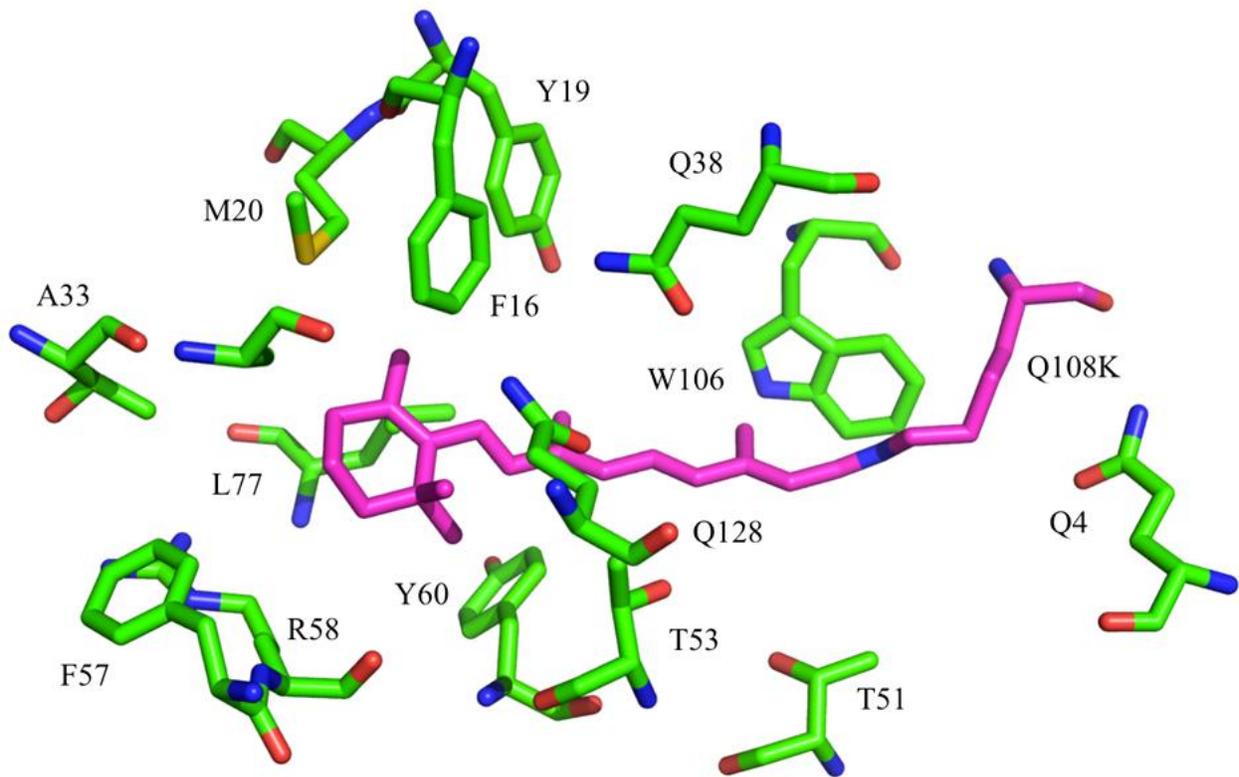


Figure 34: Crystal structure of Q108K:K40L bound to all-*trans*-retinal with surrounding residues highlighted.

As discussed before, the other visual pigments and CRABPII required a counterion for the stabilization of the PSB. However in the case of CRBPII, specifically KL, we were able to obtain a stable PSB in the absence of a counterion. Closer examination of the crystal structure revealed that a possible π -cation interaction with Trp 106 and a water mediated hydrogen bonding network between the PSB and Gln 4 could be stabilizing the PSB. With KL in hand, mutations were done that targeted different zones along the chromophore. These mutations would change the electrostatic environment around the chromophore **Figure 35**.

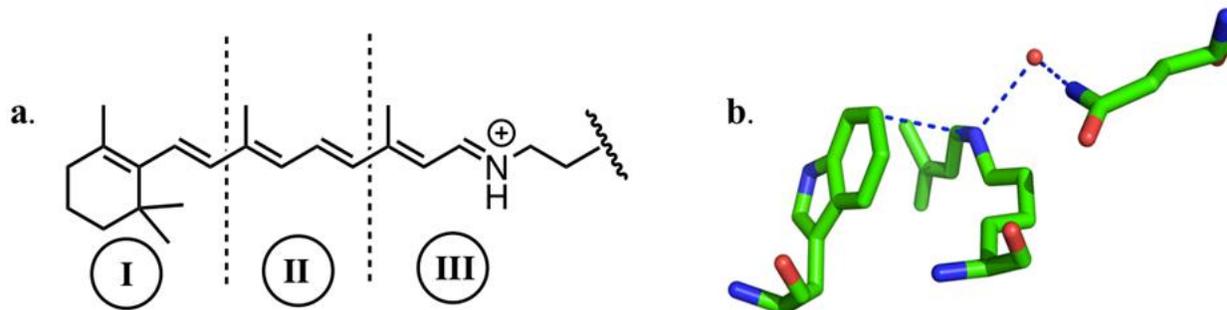


Figure 35: **a.** Protonated Schiff Base of all-*trans*-retinal. The chromophore is divided into three zones designated **I**, **II** and **III** to clarify the discussion of the different mutations. **b.** The possible stabilization of the protonated Schiff Base by Trp 106 and water mediated hydrogen bonding with Gln 4 is illustrated.

Table 9 illustrates the modifications made to zones II and III of the bound chromophore. The residues targeted in this case were Thr 51 (zone III), Thr 53 and Tyr 19 (zone II). Changes to zone II did not result in the formation of red shifted complexes, as the maximum change in the absorption was only 5 nm (entries 3 and 5). However, the mutation of Thr51 (zone III) to a hydrophobic amino acid (Val) resulted in a 25 nm red shift. This is a result of the removal of the polar Thr51, which due to its proximity to the PSB prevent the delocalization of the positive charge along the polyene. Changes to zone I, the ionone ring region, did not correlate well with the postulate of charge delocalization. The introduction of polar or acidic residues in this region resulted in a blue shift. This is contrary to the generally accepted postulate that the introduction of polar or acidic residues in the ionone ring region should result in red shifting.¹⁹ This will be discussed in detail later in the chapter.

Table 9: Effects of mutations on the different zones of the bound chromophore.

| Zone | CRBP II Mutant | λ_{\max} (nm) | Protein Shift (nm) |
|----------|-------------------|-----------------------|--------------------|
| - | KL | 508 | 0 |
| III | KL:T51V | 533 | 25 |
| II | KL:T53C | 513 | 5 |
| II + III | KL:T51V:T53C | 539 | 31 |
| II | KL:Y19W | 513 | 5 |
| II + III | KL:T51V:Y19W | 537 | 29 |
| II + III | KL:T15V:T53C:Y19W | 538 | 30 |

Protein shift is with reference to Q108K:K40L

Examination of the crystal structures (Dr. Zahra Nossoni) proved to be very essential to optimizing the binding pocket. This led to the introduction of several other mutations that resulted in the heptamutant KL:T51V:T53C:Y19W:T29L:R58W which had a maximum absorbance of 591 nm. The question then was could we get more red shifted? The disruption of the water mediated hydrogen bonding network between Thr 51 and Gln 4 that stabilizes the charge on the iminium could result in the bathochromic shift desired. In order to explore this Gln 4 was mutated to Phe, Trp, Ala, Leu and Thr. All these mutations led to a substantial bathochromic shift (591 nm to 610 – 613 nm). The introduction of the positively charged Arg resulted in even more red shifting to give a system with an absorption maximum of 622 nm. A final surge towards the most red shifted pigment was to further encapsulate the chromophore in the binding pocket. The crystal structures revealed that the Alanine at position 33 was the optimum site to mutate. This residue resides on the helix which acts as the ‘mouth’ of the

binding pocket. The incorporation of A33W to give the nonamutant, KL:T51V:T53C:Y19W:T29L:R58W:Q4R:A33W led to the formation of the most red shifted pigment ever to be reported with a maximum absorption of 644 nm.

II.6 EFFECTS OF THE THR51 MUTATION ON WAVELENGTH REGULATION

As previously discussed there are several postulates governing wavelength regulation, one of which focuses on the planarity of the chromophore. The more planar the chromophore the more red shift would be observed due to increased delocalization of the charge. The alternating double and single bonds in the chromophore allows for the different degrees of twisting along the polyene. The planarity is dependent on the steric and electronic characteristics of the binding pocket.

As discussed previously the most drastic red shift was observed when Thr 51 was mutated to the hydrophobic valine (**Table 9**). The removal of the polar residue in close proximity to the iminium allowed for the delocalization of charge along the polyene resulting in the bathochromic shift. Based on these observations, several other mutations were made at position 51 to determine if it is possible to obtain a more red shifted pigment. Residues of varying sizes and electronic properties were considered these were: Tryptophan, Phenylalanine, Isoleucine, Methionine, Alanine and Glycine. **Table 10** shows the results obtained for two sets of CRBP II mutants.

The mutation of T51 to glycine gave poor expression. With respect to alanine the mutation was successful only for the KL:R58F series. The results showed that only when T51 was changed to isoleucine that further red shifting was observed (an additional 12 to 14 nm with

respect to T51V). This bathochromic shift could be a result of hydrophobic packing against the chromophore, specifically the PSB region, which resulted in rigidifying the system even more.

Table 10: Effects of T51 mutations.

| CRBP _{II} Mutant | λ_{\max} (nm) | Protein Shift ^a (nm) | CRBP _{II} Mutant | λ_{\max} (nm) | Protein Shift ^b (nm) |
|---------------------------|-----------------------|---------------------------------|---------------------------|-----------------------|---------------------------------|
| KL | 508 | - | KL:R58F | 527 | - |
| KL:T51V | 533 | 25 | KL:R58F:T51V | 561 | 34 |
| KL:T51M | 534 | 26 | KL:R58F:T51M | 562 | 35 |
| KL:T51F | 531 | 23 | KL:R58F:T51F | 565 | 37 |
| KL:T51I | 545 | 37 | KL:R58F:T51I | 575 | 48 |

^a Protein Shift is with reference to Q108K:K40L (KL) ^b Protein Shift is with reference to Q108K:K40L:R58F

Attempts made to obtain crystal structures to prove this hypothesis, however, were unsuccessful.

Based on these observations, attempts were made to incorporate the T51I mutation to the most red shifted CRBP_{II} mutant to date, KLVCWLW:Q4R:A33W, which has a maximum absorbance of 644nm to see if we can get further red shifting. Disappointedly attempts to change the T51V to T51I in several of the longer series mutants were futile.

II.7 EXPLORING THE EFFECTS OF MUTATIONS ON THE 6-*s*-cis AND 6-*s*-trans CONFORMATION OF THE CHROMOPHORE, ALL-*trans*-RETINAL

In the early stages of the development of the most red shifted CRBP_{II} complex, the first high resolution (1.2 Å) structure was that of the octamutant KLVCWLW:Q4H. This mutant has a

maximum absorbance of 594 nm. With this mutant in hand, attempts were made to introduce polar residues in the ionone ring region in order to facilitate more red shifting. Examination of the crystal structure revealed position 77 as a possible place to begin this analysis. The mutation L77T was introduced to give the mutant KLVCWLW:Q4H:L77T. Based on the postulates put forward in regards to factors governing wavelength regulation, the introduction of polar residues in the ionone ring region should aid the delocalization of the charge resulting in red shifting. However, when the polar residue was introduced in this case a hypsochromic shift was observed. This was contrary to what was expected. This phenomenon leads to the question of what is happening within the protein with the introduction of the polar residue in the ionone region.

Gratifyingly we were able to obtain the crystal structure of the KLVCWLW:Q4H:L77T mutant. Examination of the crystal structure revealed that the chromophore does not adopt the same conformation in both mutants **Figure 36**. The chromophore in KLVCWLW:Q4H adopted the 6-*s*-trans (all-*trans*) conformation. The protonated Schiff base was also trans. The chromophore in KLVCWLW:Q4H:L77T, however, adopted the 6-*s*-cis conformation and the protonated Schiff base was cis. The change in conformation from 6-*s*-trans to 6-*s*-cis explains the blue shift observed with the introduction of L77T. The 6-*s*-trans conformation allows for a more efficient overlap of the p-orbitals in the pi system of the polyene. This allows for more delocalization that results in red shifting. (The torsion angle between the 6 – 7 bond for the L77 is 16° and the torsion angle between the 6 – 7 bond for the L77T is -28°).

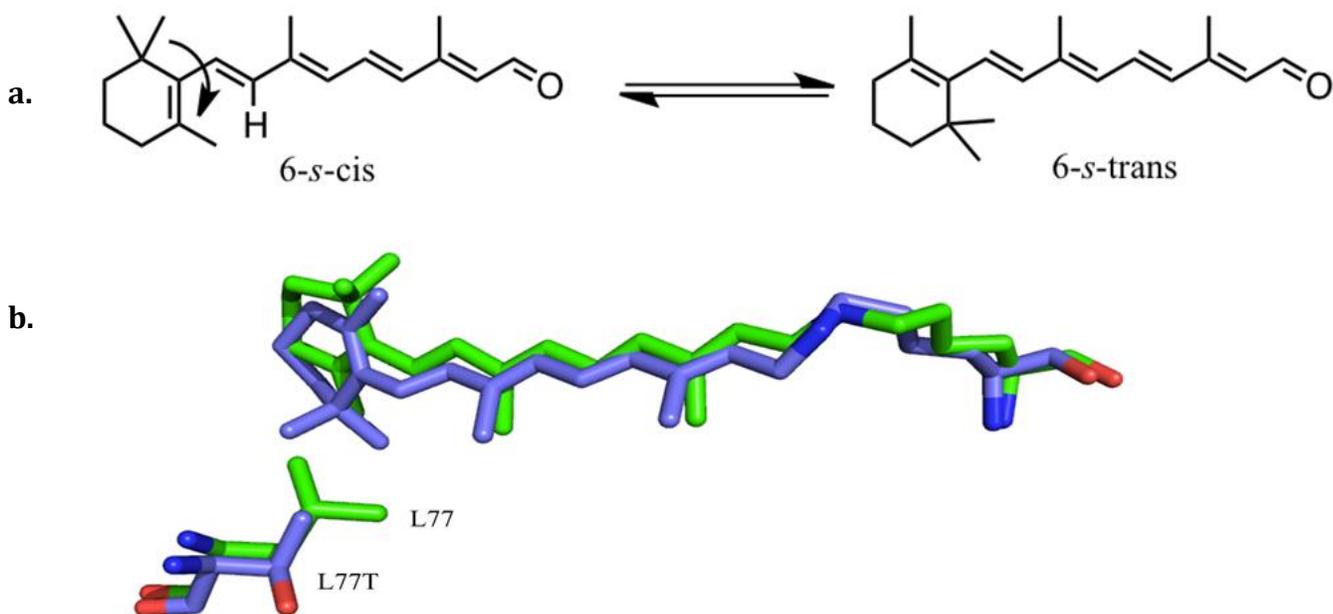


Figure 36: a. Chemical structure of the 6-*s*-cis and 6-*s*-trans conformations of retinal.^{20 18} b. Overlay of the crystal structures of KLVCWLW-Q4H (green) and KLVCWLW-Q4H-L77T (magenta). Position 77 is highlighted in both cases. Distance between L77 and the gem dimethyl of the L77T structure is 1.8 Å which would result in unfavorable steric interactions.

The mutation of the original leucine to threonine seemed to have created a ‘hole’ which can be occupied by the geminal methyl groups resulting in the change in conformation. In light of this observation, we decided to explore this phenomenon even further to see if we would be able to control the conformation of the chromophore through the introduction of residues with the desired characteristics to favor one conformation over the other. We refer to this hypothesis as the ‘bump and hole’ strategy **Figure 39**. Several other mutant pairs were used to explore this phenomenon. This explains what was discussed earlier in the chapter about the effect of the introduction of polar residues in the ionone ring region (Zone I).

The converse was then explored. The L77T mutation was introduced to the mutant with the highest pK_a, KLVYW, for which the chromophore has the 6-*s*-cis conformation (these

mutants were made by Dr. Wenjing Wang). This was done initially to see if the introduction of the polar residue would lead to an increase in the pK_a.

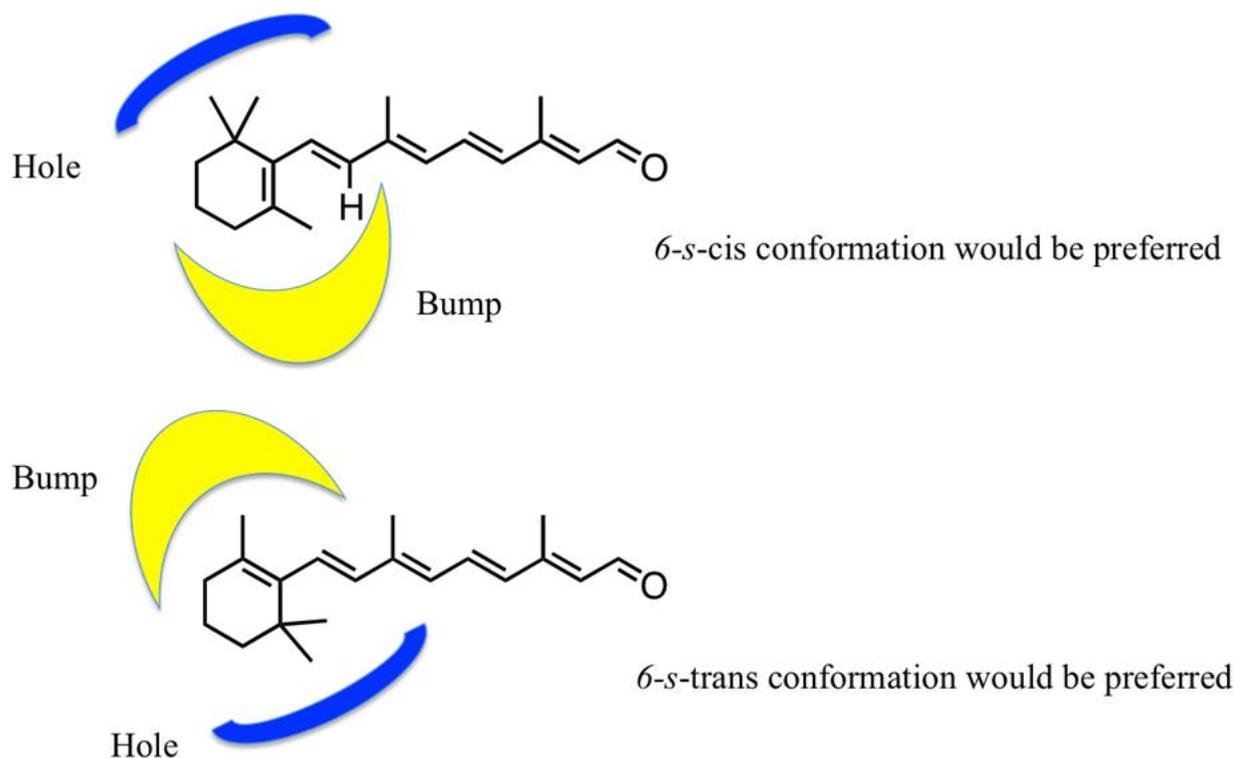


Figure 37: Diagram illustrating the 'bump and hole' strategy that would favor the 6-s-cis and 6-s-trans conformations.

However, based on the observation for the KLVCWLW-Q4H and KLVCWLW-Q4H:L77T pair it would be interesting to see if there were any changes in the conformation of the chromophore. We were able to get crystal structures of both KLVYW (Dr. Zahra Nossoni) and KLVYW:L77T. The results showed that there were no significant changes in the conformation of the chromophore. The only difference was that there was a slight difference in the trajectory of the ionone ring. The introduction of the bulkier leucine residue caused the ionone ring to move about 2 Å from its original trajectory. It is interesting to note that contrary to what was

observed for the mutants KLVCWLW-Q4H and KLVCWLW-Q4H :L77T, the iminium in both cases (KLVYW and KLVYW:L77T) was cis. Therefore it can be concluded that the 6-*s*-trans to 6-*s*-cis isomerization observed with the introduction of the L77T mutation is not solely based on the ‘bump and hole’ hypothesis. The conformation of the iminium seems to play a role. To explore this phenomenon the L77T mutation should be introduced to mutants that when bound to all-*trans*-retinal adopts the 6-*s*-trans conformation with a trans iminium.

Examination of the crystal structure revealed other residues that could be considered for this study. These residues are: T53, Q38 and S55. We were able to crystallize several of the mutants made (Ipek Yapici) but unfortunately the crystal quality of mutants of Q38 and S55 was poor. We were able to obtain crystals for the mutant pair KLVWW and KLVWW:T53L. The results showed that there was no change in the conformation of the chromophore. Based on the data in hand, it seems the conformation of the chromophore is more sensitive to the L77 position. Ipek Yapici in the Borhan lab is continuing to look into the effects of other residues on chromophore conformation.

II.8 MATERIALS AND METHODS

UV-vis spectra were recorded using a Cary 300 Bio WinUV, Varian spectrophotometer.

Fluorescence was recorded with a Fluorolog-3 (Instruments S.A., Inc.) fluorometer.

II.8.1 Mutagenesis

All mutations were made using *Stratagene's QuickChange Site-Directed Mutagenesis Protocol*.

Table 11: PCR reaction solutions.

| DNA template | Primer forward | Primer Reverse | Turbo DNA polymerase | dNTP (10mM of each) | Buffer (10x) | Water (dd) | Total volume |
|--------------|----------------|----------------|----------------------|---------------------|--------------|------------|--------------|
| 70 ng | 20 pmol | 20 pmol | 1 unit | 1 μ L | 5 μ L | / | 50 μ L |

Table 12: PCR temperature control cycles.

| | | |
|------|------------------------|--------|
| 1 x | 95 °C | 3 min |
| 20 x | 95 °C | 30 sec |
| | (T _m -4) °C | 1 min |
| | 72 °C | 10 min |
| 1 x | 72 °C | 10 min |
| 1 x | 25 °C | 5 min |

Note: T_m represents the melting temperature of the primer, which was calculated using the website shown below.

<http://www.promega.com/techserv/tools/biomath/calc11.htm>

II.8.2 WT-CRBPII gene sequence

ATG CCA AAC TTC TCT GGC AAC TGG AAA ATC ATC CGA TCG GAA AAC TTC
GAG GAA TTG CTC AAA GTG CTG GGG GTG AAT GTG ATG CTG AGG AAG ATT
GCT GTG GCT GCA GCG TCC AAG CCA GCA GTG GAG ATC AAA CAG GAG GGA
GAC ACT TTC TAC ATC AAA ACC TCC ACC ACC GTG CGC ACC ACA GAG ATT AAC
TTC AAG GTT GGG GAG GAG TTT GAG GAG CAG ACT GTG GAT GGG AGG CCC
TGT AAG AGC CTG GTG AAA TGG GAG AGT GAG AAT AAA ATG GTC TGT GAG
CAG AAG CTC CTG AAG GGA GAG GGC CCC AAG ACC TCG TGG ACC AGA GAA
CTG ACC AAC GAT GGG GAA CTG ATC CTG ACC ATG ACG GCG GAT GAC GTT
GTG TGC ACC AGG GTC TAC GTC CGA GAG TGA

II.8.3 WT-CRBPII amino acid sequence

MPNFSGNWKIIRSENFEE LLKVLGVNV **M**LRKIAVAAASKPA
VEIKQEGDTFYIKTSTTVRTTEINFKVGEEFEEQTV DGRPCK
SLVKWESENK **M**VCEQKLLKGEGPKTSWTRELTNDGELILT
MTADDVVCTRVYVRE Stop

II.8.4 Primers used for mutagenesis

T51D

Forward: 5'-GGTGATAACTTCAAGGATAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTATCCTTGAAGTTATCACC-3'

T51I

Forward: 5'-GGTGATAACTTCAAGATTA AAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTAATCTTGAAGTTATCACC-3'

T51V

Forward: 5'-GGTGATAACTTCAAGGTAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTACCTTGAAGTTATCACC-3'

T51C

Forward: 5'-GGTGATAACTTCAAGTGTA AAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTACACTTGAAGTTATCACC-3'

T51N

Forward: 5'-GGTGATAACTTCAAGAACA AAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTGTTCTTGAAGTTATCACC-3'

T51M

Forward: 5'-GGTGATAACTTCAAGATGAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTACCTTGAAGTTATCACC-3'

T51A

Forward: 5'-GGTGATAACTTCAAGGCAAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTCGTCTTGAAGTTATCACC-3'

T51G

Forward: 5'-GGTGATAACTTCAAGGGCAAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTCCGCTTGAAGTTATCACC-3'

T51F

Forward: 5'-GGTGATAACTTCAAGTTTAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTAAACTTGAAGTTATCACC-3'

T51W

Forward: 5'-GGTGATAACTTCAAGTGGAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTCCACTTGAAGTTATCACC-3'

T53V

Forward: 5'-CTTCAAGACAAAAGTCACTAGCACATTCCG-3'

Reverse: 5'-CGGAATGTGCTAGTGACTTTTGTCTTGAAG-3'

T53L

Forward: 5'-CTTCAAGACAAAAGTCACTAGCACATTCCG-3'

Reverse: 5'-CGGAATGTGCTAGTCAGTTTTGTCTTGAAG-3'

L77T

Forward: 5'-CACAAAGAGCACCGATAACCGGC-3'

Reverse: 5'-GCCGGTTATCGGTGCTCTTTGTG-3'

L77V

Forward: 5'-CACAAAGAGCGTAGATAACCGGC-3'

Reverse: 5'-GCCGGTTATCTACGCTCTTTGTG-3'

II.8.5 XL1-Blue competent cells preparation

XL1-Blue *E. coli* cells were streaked on agar plate with tetracycline as antibiotic (12.5 mg/L) and incubated overnight at 37 °C.

The following day a colony was picked from the plate using a sterile wooden applicator. The colony was inoculated in sterile LB solution (10 mL), with tetracycline (12.5 mg/L), in a shaker at 220 RPM overnight. The overnight culture (1 ml) was transferred to sterile LB solution

(200 mL) with tetracycline (12.5 mg/L). the solution was kept shaking at 220 RPM at 37 °C for about 2 h or until the OD was between 0.4 and 0.8. At the same time, a 500 mL centrifuge bottle was sterilized with 20% bleach for 2 h and rinsed with sterile water 6 to 7 times before use.

The cells were harvested by centrifugation at 3500 RPM for 5 min. at 4 °C. The cells were resuspended with sterile sodium chloride solution (0.9 %, 100 mL). The cells were harvested by centrifugation at 3000 RPM for 4 min at 4 °C.

The following steps were then followed:

A: The cell pellet was resuspended with calcium chloride solution (100 mM, 15% glycerol v/v, 10 mL) on ice and was incubated for at least 30 min. The cells were fractioned in sterile eppendorf tubes in 100 µL portions and flash frozen with liquid nitrogen.

B: The cell pellet was resuspended with calcium chloride solution (100 mM, 50 mL) and incubated on ice for 30 min. The cells were spun down by centrifugation at 3000 RPM for 4 min and then resuspended with calcium chloride solution (100 mM, 15% glycerol v/v, 10 mL) on ice. The cells were fractioned in sterile eppendorf tubes in 100 µL portions and flash frozen with liquid nitrogen.

Note: The cells prepared using protocol **A** uses one step less than protocol **B**. This results in cells that are less competent. The cells that are used for the transformation of the PCR product was prepared using protocol **B** as the cells need to be highly competent. This is due to the fact that the DNA from the PCR reaction is nicked and does not go through the cell membrane with the same efficiency as circular DNA during the heat shock process. Therefore protocol **B** is a better protocol for competent cells preparation. This protocol was used to prepare the BL21 competent cells the only difference is chloramphenicol was used instead of tetracycline.

II.8.6 Heat shock transformation

The DpnI digested PCR solution (5 μ L) was added to competent cells (50 or 100 μ L). The cells were then incubated on ice for 10 to 30 min, followed by heat shock at 42 °C for 50 sec. The cells were put back on ice immediately. Sterile LB solution (500 μ L) was added to the cells and shaken at 37 °C for 1 h. The cells were then harvested by centrifugation at 5000 RPM for 1 min. The cells were resuspended in LB (50 μ L) and then plated on an agar plate with the appropriate antibiotics.

II.8.7 Protein expression and purification in the pET-17b system

CRBP_{II} DNA was transformed into BL21 (DE3) pLysS competent cells. A single colony was picked from the transformation plate and inoculated in 15 or 50 mL of LB with concentration of ampicillin at 100 mg/L and chloramphenicol at 27 mg/L at 37 °C overnight. The next day the growth was transferred to 1 L of sterile LB media with the same concentration of ampicillin and chloramphenicol. The media was incubated at 37 °C until the OD₆₀₀ reached 0.4 to 0.8. Isopropyl-1-thio-D-galactopyranoside (IPTG, Gold Biotechnology) was added to a final concentration of 1mM to induce protein expression, and the cell culture was shaken at RT or 16 °C, depending on the mutant, overnight.

Cells were harvested by centrifugation at 5000 RPM for 10 min in Beckmann J2-21M/E centrifuge and resuspended with 50 ml Tris binding buffer (10 mM Tris, pH=8.0), followed by 3 min sonication (3 time, 1 min per pulse, pulse=60%, power 60%) to lyse the cells. DNase (20 μ L, Roche, recombinant, 25 unit/ μ L) was added to the lysed cells and incubated for 30 min. The

lysed cell solution was then spun down for 30 min at 5000 RPM. The supernatant was then applied to the fast Q ion exchange column, which was pre-equilibrated (washed with 50 mL 2 M NaCl, and equilibrated with 2 x 50 mL washes of 10 mM Tris, pH=8.0, buffer). The protein was then washed with 10 mM Tris, pH=8.0 and eluted using the elution buffer, 10 mM Tris, 100 mM NaCl, pH=8.0. The eluted protein was desalted using an Amicon filter membrane (10 kDa cut off) and then applied to a source Q Fast Protein Liquid Chromatography (FPLC). The protein was subjected to a gradient sodium chloride solution from 0 mM to 1 mM at pH=8.0 (25 mM Tris and Tris-HCl). The protein was eluted at ~40 mM sodium chloride.

II.8.8 Crystallization Methods

The purified protein was then concentrated to concentrations ranging from 10 – 20 mg/ml. Two equivalents of the ligand (*all-trans*-retinal) were then added and the resulting mixture incubated at 4 °C overnight or 6-8 h at room temperature (depends on the mutant). The UV-vis of the sample was taken to ensure that all the protein is bound. Two 24 well plates were set up using the hanging drop method. The condition in each well is shown in **Table 13**. The drop is comprised of 1µl of the well solution and 1µl of the protein sample.

Table 13: Crystallization Conditions.**a) Box 5**

| Reservoir | PEG (4K) 40% | Salt NH₄OAc (4M) | Buffer NaOAc (1M) pH= 4.0 | Buffer NaOAc (1M) pH= 4.5 | Buffer NaOAc (1M) pH= 4.8 | H₂O |
|------------------|-------------------------|--|--|--|--|-----------------------|
| A1 | 625 | 25 | 100 | | | 250 |
| A2 | 750 | 25 | 100 | | | 125 |
| A3 | 750 | 25 | 100 | | | 125 |
| A4 | 800 | 25 | 100 | | | 75 |
| A5 | 875 | 25 | 100 | | | 0 |
| A6 | 950 | 25 | 100 | | | 0 |
| B1 | 625 | 25 | | 100 | | 250 |
| B2 | 750 | 25 | | 100 | | 125 |
| B3 | 750 | 25 | | 100 | | 125 |
| B4 | 800 | 25 | | 100 | | 75 |
| B5 | 875 | 25 | | 100 | | 0 |
| B6 | 950 | 25 | | 100 | | 0 |
| C1 | 625 | 25 | | | 100 | 250 |
| C2 | 750 | 25 | | | 100 | 125 |
| C3 | 750 | 25 | | | 100 | 125 |
| C4 | 800 | 25 | | | 100 | 75 |
| C5 | 875 | 25 | | | 100 | 0 |
| C6 | 950 | 25 | | | 100 | 0 |
| D1 | 625 | 37.5 | 100 | | | 212.5 |
| D2 | 750 | 37.5 | 100 | | | 87.5 |
| D3 | 750 | 37.5 | 100 | | | 87.5 |
| D4 | 800 | 37.5 | 100 | | | 37.5 |
| D5 | 875 | 37.5 | 100 | | | 0 |
| D6 | 950 | 37.5 | 100 | | | 0 |

Table 13 (cont'd)

b) Box 6

| Reservoir | PEG (4K) 40% | Salt NH₄OAc (4M) | Buffer NaOAc (1M) pH= 4.0 | Buffer NaOAc (1M) pH= 4.5 | Buffer NaOAc (1M) pH= 4.8 | H₂O |
|------------------|-------------------------|--|--|--|--|-----------------------|
| A1 | 625 | 0 | 100 | | | 275 |
| A2 | 750 | 0 | 100 | | | 150 |
| A3 | 750 | 0 | 100 | | | 150 |
| A4 | 800 | 0 | 100 | | | 100 |
| A5 | 875 | 0 | 100 | | | 25 |
| A6 | 950 | 0 | 100 | | | 0 |
| B1 | 625 | 25 | 100 | | | 268.75 |
| B2 | 750 | 25 | 100 | | | 143.75 |
| B3 | 750 | 25 | 100 | | | 143.75 |
| B4 | 800 | 25 | 100 | | | 93.75 |
| B5 | 875 | 25 | 100 | | | 18.75 |
| B6 | 950 | 25 | 100 | | | 0 |
| C1 | 625 | 25 | 100 | | | 262.5 |
| C2 | 750 | 25 | 100 | | | 137.5 |
| C3 | 750 | 25 | 100 | | | 137.5 |
| C4 | 800 | 25 | 100 | | | 87.5 |
| C5 | 875 | 25 | 100 | | | 12.5 |
| C6 | 950 | 25 | 100 | | | 0 |
| D1 | 625 | 37.5 | | | 100 | 257.5 |
| D2 | 750 | 37.5 | | | 100 | 132.5 |
| D3 | 750 | 37.5 | | | 100 | 132.5 |
| D4 | 800 | 37.5 | | | 100 | 82.5 |
| D5 | 875 | 37.5 | | | 100 | 7.5 |
| D6 | 950 | 37.5 | | | 100 | 0 |

***Total volume in each well = 1ml**

***The volumes indicated in the table are in μ L.**

Table 14: Data Collection and refinement statistics. The ligand is all-*trans*-retinal in each case.

| | KL-T51V- R58Y-Y19W- L77T | KL-T51V- T53C-R58W- T29L-Y19W- Q4H-L77T | KL-R58E | KL-T51V- T53V-R58Y- Y19W | KL-T51V- T53L-R58W- Y19W | KL-T51V- T53V-R58W- Y19W |
|---|---|--|-----------------|---|---|---|
| Data collection | | | | | | |
| Space group | C121 | C121 | P1 | P1 | P1 | P1 |
| Cell Dimensions | | | | | | |
| <i>a</i> (Å) | 29.85 | 29.98 | 29.62 | 29.53 | 30.02 | 29.53 |
| <i>b</i> (Å) | 64.96 | 64.99 | 36.11 | 35.67 | 35.84 | 35.67 |
| <i>c</i> (Å) | 128.91 | 129.19 | 63.86 | 64.48 | 64.51 | 63.48 |
| Wavelength (Å) | 1.0 50 – 1.3 | 1.0 50 – 1.3 | 1.0 50 – 1.3 | 1.0 50 – 1.5 | 1.0 50 – 1.77 | 1.0 50 – 1.9 |
| Resolution (Å) | 90% | 91% | 80% | 97% | 97% | 95% |
| Completeness | | | | | | |
| Refinement | | | | | | |
| Resolution (Å) | 1.6 | 1.7 | 1.8 | 2.0 | 1.85 | 2.0 |
| No. of Reflections | 28290 | 23577 | 17103 | 14954 | 19420 | 15386 |
| R _{factor} / R _{free} | 20/26 | 20/30 | 21/32 | 20/26 | 19/25 | 20/27 |

Table 14 (cont'd)

| | KL-T51V- R58Y-Y19W- L77T | KL-T51V- T53C-R58W- T29L-Y19W- Q4H-L77T | KL-R58E | KL-T51V- T53V-R58Y- Y19W | KL-T51V- T53L-R58W- Y19W | KL-T51V- T53V-R58W- Y19W |
|-----------------------------|---|--|----------------|---|---|---|
| No. of atoms | | | | | | |
| Protein | 2190 | 2190 | 2190 | 2190 | 2190 | 2190 |
| Ligand | 26 | 26 | 26 | 20 | 40 | 40 |
| Water | 143 | 519 | 297 | 95 | 111 | 71 |
| Acetate | 0 | 0 | 0 | 0 | 0 | 2 |
| B-factors | | | | | | |
| Overall | 34.947 | 26.810 | 42.290 | 34.975 | 28.163 | 34.542 |
| Main (Chain A) | 31.785 | 22.401 | 37.957 | 30.756 | 23.488 | 36.481 |
| Main (Chain B) | 28.736 | 20.162 | 41.525 | 31.232 | 23.847 | 30.955 |
| Ligand | 65.171 | 43.865 | 62.409 | 57.147 | 80.336 | 50.193/67.522 |
| R.M.S deviations | | | | | | |
| Bond lengths (Å) | | | | | | |
| Overall | 0.022 | 0.019 | 0.013 | 0.016 | 0.020 | 0.016 |
| Bond Angles (°) | | | | | | |
| Overall | 2.097 | 2.006 | 1.783 | 1.890 | 2.006 | 1.819 |

REFERENCES

REFERENCES

1. Wald, G., Mechanism of human color vision. *Am. J. Ophthalmol*, **1965**, 60 (6), 1132.
2. Wald, G., Receptors of human color vision. *Science* **1964**, 145 (363), 10074.
3. Brown, P. K. Wald, G. Visual pigments in single rods + cones of human retina - direct measurements reveal mechanisms of human night + color vision. *Science* **1964**, 144 (361), 45.
4. Kim, S. Y.; Waschuk, S. A. Brown, L. S.; Jung, K. H., Screening and characterization of proteorhodopsin color-tuning mutations in *Escherichia coli* with endogenous retinal synthesis. *Bba-Bioenergetics* **2008**, 1777 (6), 504-513.
5. (a) Vaezeslami, S.; Mathes, E.; Vasileiou, C. , B.; Geiger, J.H., The structure of apo wild-type cellular retinoic acid binding protein II at 1.4 Å and its relationship to ligand binding and nuclear translocation. *J. Mol. Biol.* **2006**, 363 (3), 687-701; (b) Kleywegt, G. J.; Bergfors, T.; Senn, H.; Lemotte, P.; Gsell, B.; Shudo, K.; Jones, T. A., Crystal-structures of cellular retinoic acid binding protein-I and protein-II in complex with all-trans-retinoic acid and a synthetic retinoid. *Structure* **1994**, 2 (12), 1241-1258.
6. Vasileiou, C.; Vaezeslami, S.; Crist, R. M.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B., Protein design: Reengineering Cellular Retinoic Acid Binding Protein II into a rhodopsin protein mimic. *J. Am. Chem. Soc.* **2007**, 129 (19). 6140-6148.
7. Burgi, H. B.; Dunitz, J. D.. From crystal statics to chemical dynamics. *Accounts Chem. Res.* **1983**, 16 (5). 153-161.
8. (a) Crist, R. M.; Vasileiou, C.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B.. Engineering a rhodopsin protein mimic. *J. Am. Chem. Soc.* **2006**, 128 (14), 4522-4523; (b) Vasileiou, C.; Wang, W. J.; Jia, X. F.; Lee, K. S. S.; Watson, C. T.; Geiger, J. H.; Borhan, B., Elucidating the exact role of engineered CRABPII residues for the formation of a retinal protonated Schiff base. *Proteins* **2009**, 77 (4), 812-822.
9. Lee, K. S. S., PhD. Dissertation, Michigan State University, East Lansing, 2011.
10. Lee, K. S. S.; Jia, X.; Wang, W.; Berbasova, T.; Choi, Y.; Nossoni, F.; Geiger, J.H.; Borhan, B. *ChemPlusChem*, **2012**, 4, 273 – 276.
11. Beppu, Y.; Kakitani, T., Theoretical-study of color control mechanism in retinal proteins role of the tryptophan residue, tyrosine residue and water molecule. *Photochem. Photobiol.* **1994**, 59 (6), 660-669.
12. Wolf, G., The intracellular vitamin-a-binding proteins - an overview of their functions. *Nutr. Rev.* **1991**, 49 (1), 1-12.13.

13. Li, E.; Norris, A. W., Structure/function of cytoplasmic vitamin A binding proteins. *Annu. Rev. Nutr.* **1996**, *16*, 205-234.
14. Wolf, G., Cellular retinoic acid-binding protein II: A coactivator of the transactivation by the retinoic acid receptor complex RAR.RXR. *Nutr. Rev.* **2000**, *58* (5), 151-153.
15. Durand, B.; Saunders, M.; Leroy, P.; Leid, M.; Chambon, P., Alltrans and 9-cis retinoic acid induction of CRABP II transcription is mediated by rar-rxr heterodimers bound to dr1 and dr2 repeated motifs. *Cell* 1992, *71*(1), 73-85.
16. (a) Winter, N. S.; Bratt, J. M.; Banaszak, L. J., Crystal-structures of holo and apo Cellular Retinol-Binding Protein-II. *J. Mol. Biol.* **1993**, *230* (4), 1247-1259; (b) Tarter, M.; Capaldi, S.; Carnizo, M. E.; Ambrosi, E.; Perduca, M.; Monaco, H. L., Crystal structure of human cellular retinol-binding protein II to 1.2 angstrom resolution. *Proteins* **2008**, *70* (4), 1626-1630.
17. Tarter, M.; Capaldi, S.; Carnizo, M. E.; Ambrosi, E.; Perduca, M.; Monaco, H. L., Crystal structure of human cellular retinol-binding protein II to 1.2 Å resolution. *Proteins* **2008**, *70* (4), 1626-1630.
18. Rajput, J.; Rahbek, D. B.; Andersen, L. H.; Hirshfeld, A.; Sheves, M.; Altoe, P.; Orlandi, G.; Garavelli, M., Probing and modeling the absorption of retinal protein chromophores in vacuo. *Angew. Chem.Int. Ed. Engl.* **2010**, *49* (10), 1790-1793.
19. Blatz, P. E.; Mohler, J. H.; Navangul, H. V., Anion-induced wavelength regulation of absorption maxima of Schiff bases of retinal. *Biochemistry* **1972**, *11* (5), 8484.
20. (a) Hamanaka, T.; Kakudo, M.; Ashida, T.; Mitsui, T., Crystal structure of all-trans retinal. *Acta. Crystall. B-Stru.* **1972**, *B 28* (Jan15), 2144; (b) Gilarch, R.; Kane, I. L.; Karte, J.; Sperling, W., Crystal structure of visual chromophores, 11 -cis and all-trans retinal. *Nature* **1971**, *232* (5307), 187.
21. Wang, W. Michigan State University, East Lansing, 2012

CHAPTER III. DEVELOPMENT AND CRYSTALLIZATION OF CRBPII DERIVATIVES AS FLUORESCENT PROTEIN TAGS

III.1 GREEN FLUORESCENT PROTEIN (GFP) AND ITS DERIVATIVES

The development of fluorescent protein tags has allowed for the exploration of cellular processes on the molecular level. This has resulted in a significant increase in knowledge of cellular processes. The fusion of fluorescent protein tags to a protein of interest allows for the visualization of the protein in the cellular environment. Fluorescent protein tags provide a non-invasive approach to study the expression level, localization and movement of the protein of interest throughout the cell. The use of fluorescent protein tag has also been used to elucidate molecular function and key protein-protein interactions of the protein of interest.¹

Green fluorescent protein (GFP) is the most widely used fluorescent protein for imaging in live cells. The development and utility of GFP has made it a benchmark for the development of other fluorescent protein tags. The discovery of GFP and its impact on the field of cellular biology has led to the scientists who were very instrumental to this end, to be awarded the Nobel Prize in 2008.

GFP was first discovered in the species *Aequorea victoria* by Osamu Shimomura while he was a student at Princeton University in the 1960s.² In 1979, the chemical structure of the chromophore formed in GFP was proposed to be 4-(*p*-hydroxybenzylidene) imidazolidin-5-one (HBI).³ This proposal was later verified and the structure was confirmed. In 1992, the GFP gene was sequenced and cloned by Douglas Prasher. However, he was unable to observe autofluorescence when the GFP was expressed in *E. coli*. This was found to be a result of the

extra amino acid residues at the N-terminus.⁴ It was at this point that Martin Chalfie began to work with the GFP gene and was able to successfully express GFP that was able to generate the fluorophore autocatalytically. Roger Tsien's group also worked on GFP and was able to successfully generate a brighter GFP monomer. They also developed variants of this monomer, which gave access to different fluorescent tags that exhibit different emission wavelengths, thus different colors are observed.

The GFP protein consists of 11-stranded β -barrel with an α -helix running down the middle of the barrel (**Figure 38a**).⁵ The helix is very important in this structure as the residues responsible for the formation of the fluorophore are located on this helix. These residues are located at positions 65, 66 and 67. The formation of the fluorophore begins with the cyclization of the amide backbone followed by dehydration to form the resultant imidazolidin-5-one ring. The oxidation of this imidazolidin-5-one intermediate leads to the formation of the mature GFP fluorophore.⁶ The proposed mechanism for the formation of the fluorophore is illustrated in **Figure 38b**.

It was observed that the wild type GFP was not very bright. The absolute brightness is defined as the radiant intensity of the emission to the illumination intensity of the excitation. This observation was attributed to the high pKa of the tyrosine at position 66. In the highly fluorescent form (quantum efficiency 65%) of the fluorophore, Tyr 66 is deprotonated and in the case of the wild type GFP the protonated form is the most dominant. This was attributed to the hydrogen bonding interaction between serine at position 65 and the tyrosine at position 66. The replacement of Ser 65 with Ala, Cys, Leu or Thr, resulted in the deprotonated state of the

fluorophore to be formed exclusively.^{7 8} The mutations disrupted the hydrogen bonding network which resulted in a lowering of the pKa of Tyr 66.

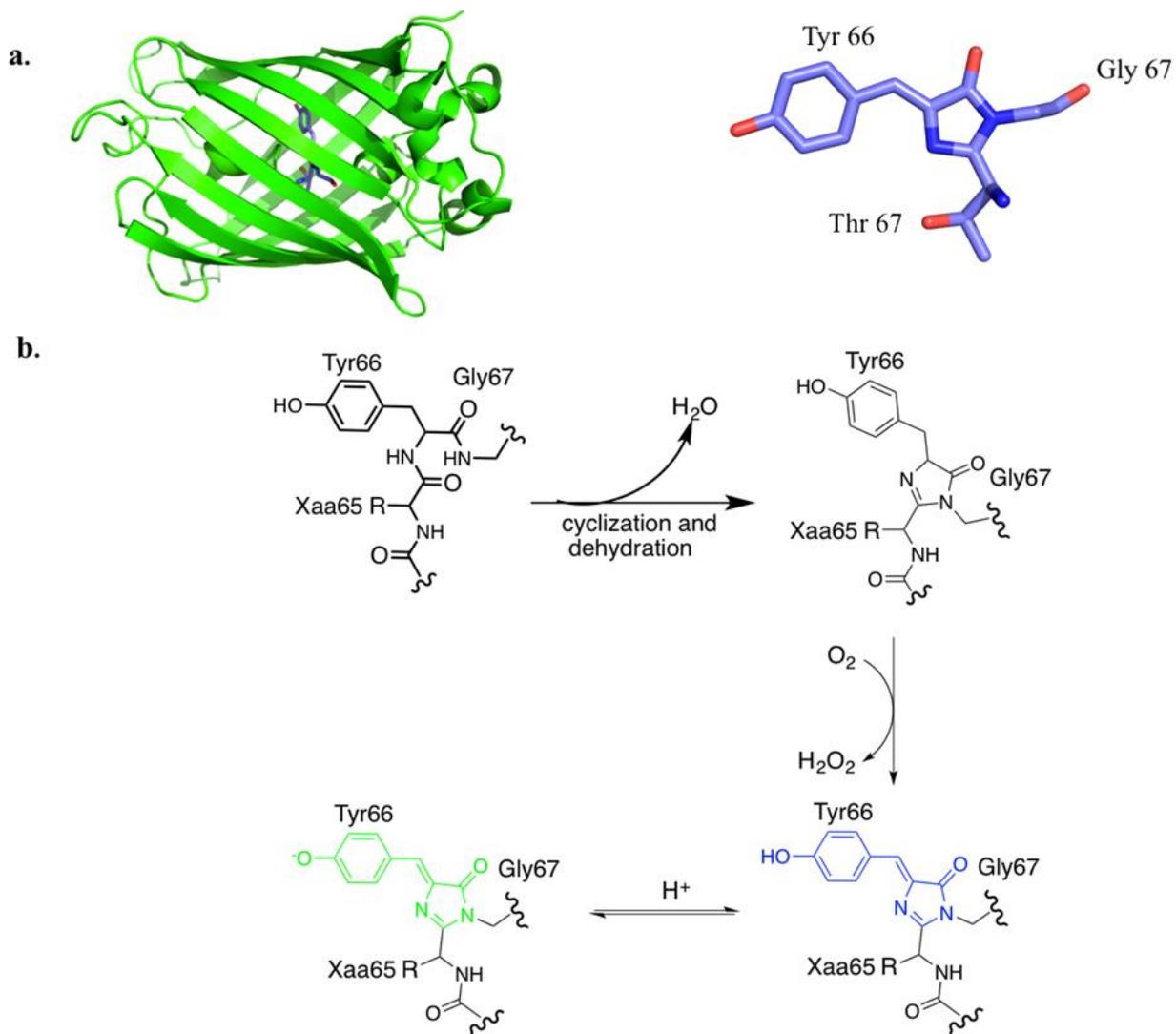


Figure 38: a. Crystal structure (left) and fluorophore (right) of EGFP (PDB code: 2Y0G). **b.** Molecular mechanism illustrating the maturation of the fluorophore in green fluorescent proteins.

It was found that the mutation of Ser 65 to Thr accelerated the maturation of the fluorophore four fold. This result greatly increased the potential use of GFP as a fluorescent tag.

This GFP variant served as a platform for the development of other fluorescent proteins such as blue, cyan and orange fluorescent proteins. This was done by either changing the residue

at position 66 or by modifying the environment surrounding the fluorophore in the binding pocket.⁹

With the availability of these different GFP variants, one can only suspect that the development of a red variant would be next in order to enhance the color palette. However attempts to engineer a red fluorescent protein from the native GFP were futile until 2008.¹⁰ The initial lack of success in the development of red fluorescent protein from the native GFP prompted the search for another source. In 2001, Remington and coworkers isolated and crystallized *Discosoma* red fluorescent protein (DsRed); a red fluorescent protein found in coral. The crystal structure revealed that DsRed shared similar structural features to GFP even though there is only 23% sequence homology.¹¹ The fluorophore in DsRed is more conjugated than the fluorophore of GFP. This extra conjugation is a result of another oxidation step (**Figure 39**).

DsRed forms a tetramer with a maturation time of 10 hours at 37°C. In order to make this system as effective as GFP, the generation of monomeric DsRed that maintains the ability to fluoresce is essential. Glick and coworkers carried out mutagenesis on DsRed in order to generate a monomeric, functioning DsRed with a faster maturation time.¹² It required a total of 33 mutations to achieve this goal. In 2004, Tsien and coworkers used this engineered monomeric DsRed as the platform for the development of improved red, orange and yellow fluorescent proteins. These different derivatives exhibit different spectral characteristics, maturation time and stability.¹³ The development of these systems allows for the visualization of more than one process by using the appropriate tags.¹⁴

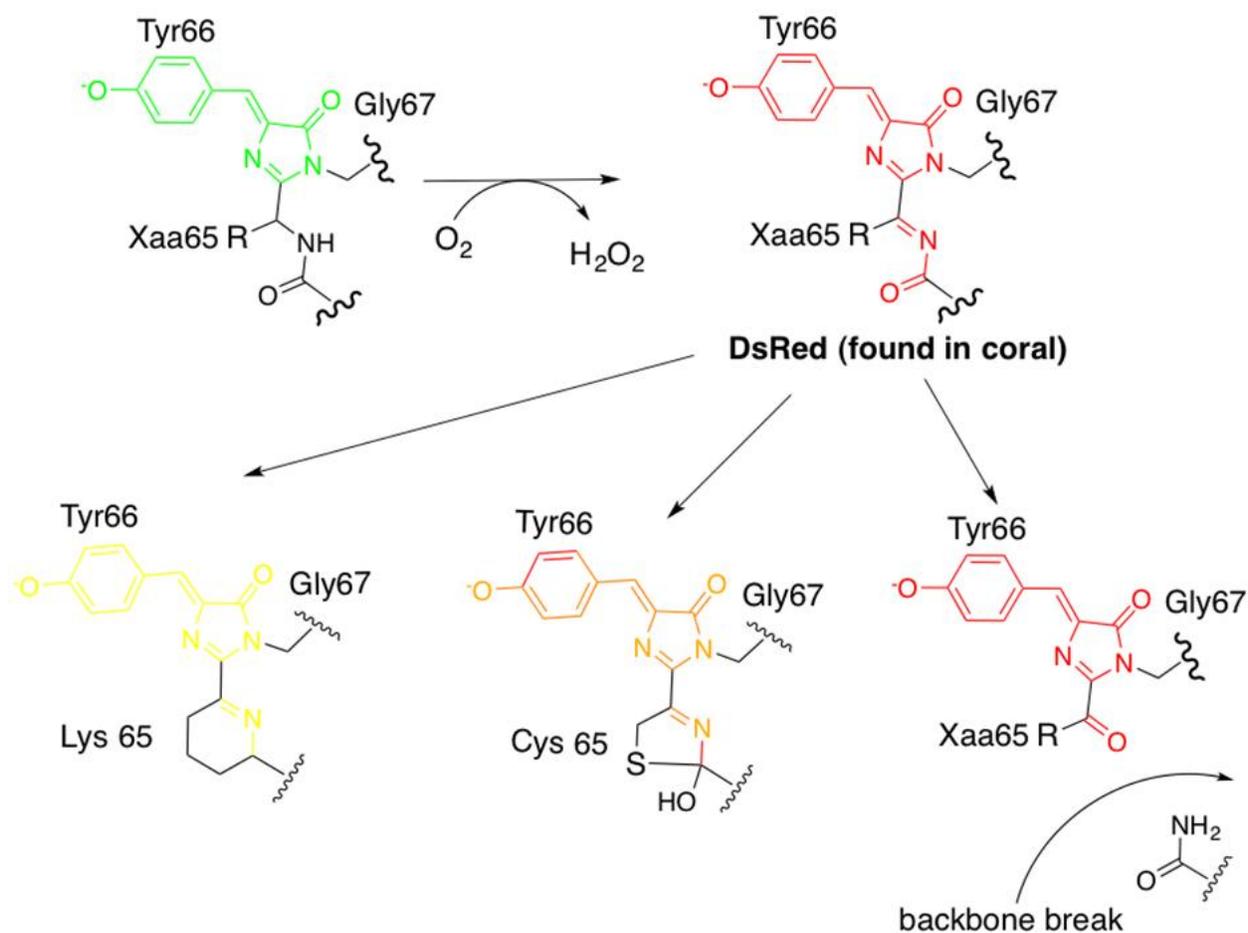


Figure 39: Mechanism for the formation of the fluorophore and molecular structure of the different variants developed from DsRed.

Fluorescent proteins have been used extensively in the field of cellular biology. They have been used to monitor movement of the protein through the cell, as a signal read out of transcription activation, and also to study protein-protein interactions and conformational changes using FRET.^{15 16 17} A more recently developed use of fluorescent protein is for super-resolution microscopic studies. These studies will allow for the visualization of a single protein within the cell.^{18 19}

There are two major classes of fluorescent proteins. These are: reversible and irreversible photoactivatable fluorescent proteins. For the reversible fluorescent proteins, fluorescence can be

turned on and off at will or the fluorescence can be switched by irradiation at different wavelengths. The latter phenomenon is usually due to *cis-trans* isomerization of a double bond that could lead to conformational changes which could then lead to spectral changes.²⁰ Irreversible photoactivation involves chemical structural changes of the chromophore; cleavage of the chromophore from the backbone of the protein.²¹

III.2 OTHER FLUORESCENT TAGS

III.2.1 SNAP tag

The SNAP tag is a commercially available system that allows each protein to be specifically tagged with a suitable fluorescent dye. Methods such as these are used in biochemistry and more often, in order to investigate the function of proteins and enzymes in living cells. SNAP protein is a derivative of the ubiquitous mammalian cellular enzyme O-6-Alkylguaninealkyltransferase (hAGT), which has the task of DNA repair at guanine sites. It irreversibly transfers the alkyl group to one of its cysteine groups in the active site from the alkylated guanine, in order to carry out its function as a DNA repair protein (**Figure 40a**).

Tainer and coworkers discovered that substitutions did not affect the activity of the alkyl transferase enzyme.²² This discovery led Johnsson and coworkers to covalently label the hAGT protein with any molecule at the para position of the benzyl group (**Figure 40b**).²³

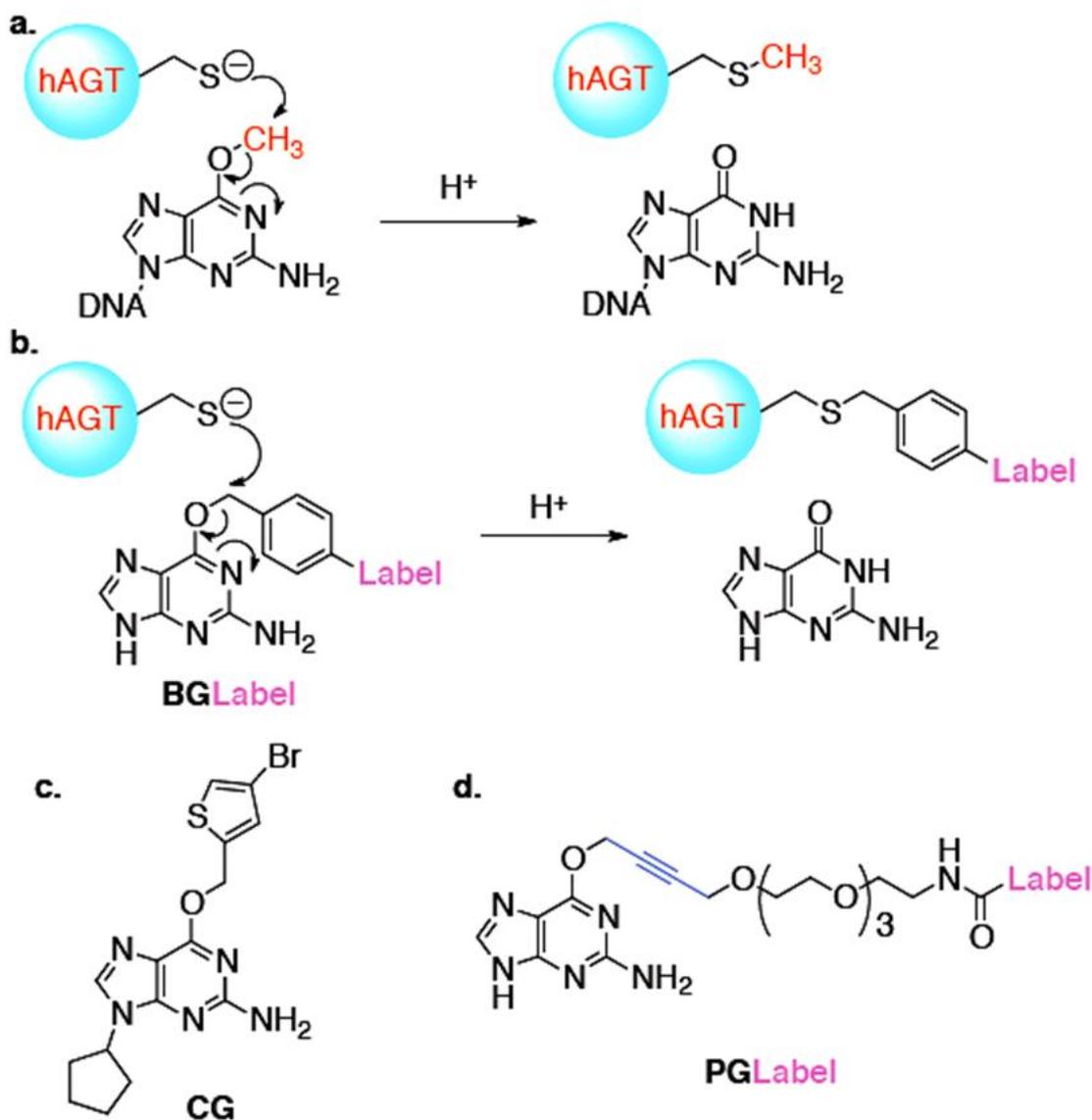


Figure 40: Illustration of SNAP protein tag mechanisms.²³ **a.** The mechanism of the DNA repair enzyme hAGT. **b.** General ways of labeling hAGT with a labeled benzyl guanine. **c.** Molecular Structure of WT-hAGT inhibitor CG. **d.** Molecular structure of labeled propargyl guanine.

The development of the hAGT protein tag involved the removal of 30 amino acid residues from the C-terminus. These 30 amino acids were shown not to affect the activity of the enzyme. The truncated hAGT was used as the platform for developing a more efficient SNAP tag system. Initial results showed that there was a lot of background fluorescence due to the activity of the endogenous hAGT protein. In order to solve this problem, the mutation G106W

was made to increase the enzyme's activity towards the unnatural substrate benzyl guanine (BG), **(Figure 40b)**.

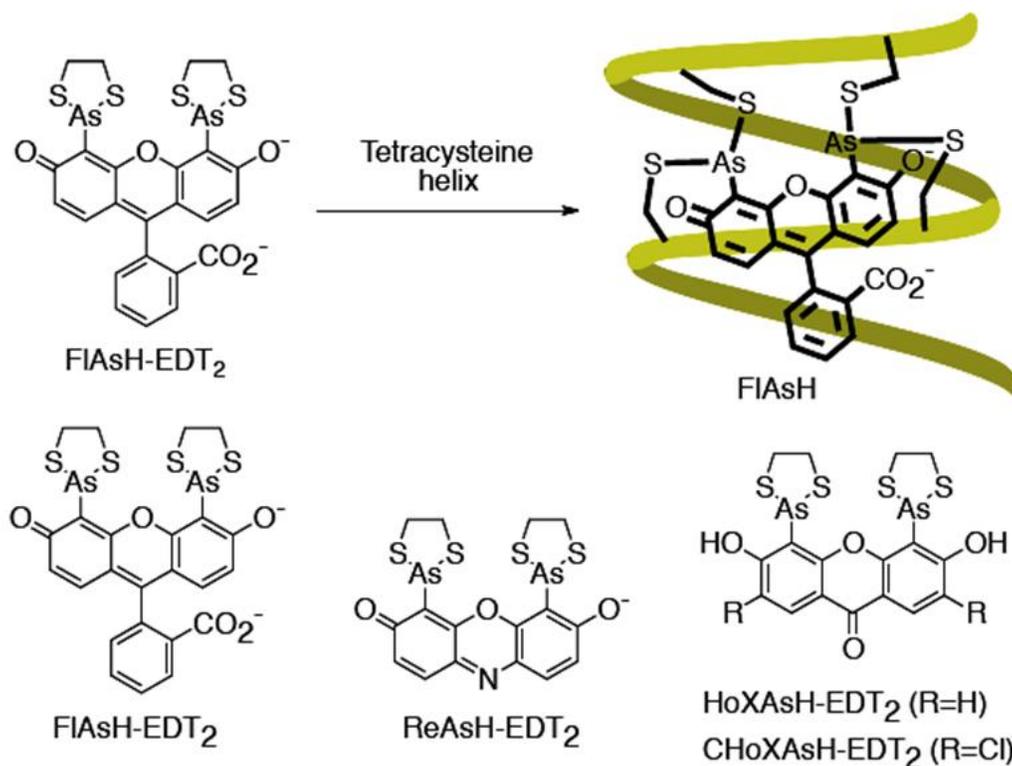
In order to reduce the background fluorescence issue, an inhibitor cyclopentyl guanine (CG), was added **(Figure 40c)**. CG is specific for WT-hAGT protein and it has no effect on the activity of BG with the mutant hAGT developed through directed evolution.²⁴

The SNAP tag is widely used in the field of cellular biology. The versatility it provides comes from the fact that different fluorophores with different spectral properties can be bound to the benzyl guanine moiety. By applying cell permeable or cell impermeable fluorophores, it is possible to visualize targets on the outer membrane as well as intracellular proteins.

The SNAP tag system has limitations. This technique is not selective with different labels and so multiple tags cannot be used at the same time. Engineering mutants that are very substrate specific solved this problem. For example the derivative ^LAGT which is specific for O⁶ – propargyl guanine **(Figure 40d)**.²⁵

III.2.2 FAsH tag

Fluorescein Arsenical Helix binder (FAsH) was developed by Tsien and coworkers.²⁶ The design of this tag involved the use of a small peptide that can fold into an α -helix with four cysteine residues strategically placed at positions i , $i + 1$, $i + 4$, $i + 5$. The cysteine residues allow for the chelation to the bis-arsenic fluorescein, which enhances the fluorescence by a thousand fold. The enhancement has been attributed to the restriction of the rotation upon binding to the peptide. This restriction in rotation causes a decrease in the non-radiative decay. The drawback of the FAsH tag is the toxicity of arsenic.



FIAsh - $\lambda_{Ex} = 500$ nm, $\lambda_{Em} = 529$ nm; ReAsH - $\lambda_{Ex} = 593$ nm, $\lambda_{Em} = 608$ nm; CHoXAsH - $\lambda_{Ex} = 380$ nm, $\lambda_{Em} = 430$ nm.

Figure 41: Illustration of FIAsh mechanism and three different organoarsenic compounds used for the different color FIAsh system.²⁶

Therefore in order to reduce the toxicity to live cells 1,2-ethanediol (EDT) has to be added to prevent the chelation of the tag to other cysteine residues within the cells. The FIAsh tag has association and dissociation rate constants ranging from $10 \text{ M}^{-1}\text{s}^{-1}$ to $1 \text{ M}^{-1}\text{s}^{-1}$. This allows for the addition for micromolar quantities of the arsenic compounds and quick removal to reduce toxicity.

Several variants of the FIAsh tag have been developed that span the entire visible spectrum. For each of these variants a different arsenic compound is used (**Figure 41**). Even though there is access to the different colored FIAsh systems, the major limitation is the inability of the different

organoarsenic compounds to distinguish between the helices. This prevents the use of two of the organoarsenic compounds at the same time in the same cell.

III.3 ENGINEERING CRBP_{II} INTO A FLUORESCENT PROTEIN TAG

With the development of the different fluorescent protein tags, the development of red and near infra-red fluorescent tags is still lacking. These systems would be quiet promising as the absorption maxima would be out of the range of the absorption of molecules within the cell. This would result in a significant reduction in the occurrence of background fluorescence. Even though there are a few red and near infra-red fluorescent protein tags, they lack good photostability. Therefore a more photostable fluorescent protein tag is needed to complete the existing fluorescent protein palette.

It was to this end that CRBP_{II} was considered. As mentioned before CRBP_{II} was engineered into a rhodopsin mimic to probe the factors governing wavelength regulation. CRBP_{II} was engineered to bind to all-*trans*-retinal as a protonated Schiff base (PSB) and was shown to regulate the wavelength of the bound chromophore greater than 170 nm. The high expression levels of CRBP_{II} in *E. coli* and the fact that this protein is amenable to mutagenesis and routine crystallization, led to consideration of developing this protein into a fluorescent tag.

The color spectrum that is observed for CRBP_{II} bound to all-*trans*-retinal through wavelength regulation could be applied to CRBP_{II} as a fluorescent protein tag. This would require the use of the appropriate fluorophore and strategic mutagenesis studies.

The strategy to make CRBP_{II} into a useful fluorescence tag is to utilize the information obtained from using all-*trans*-retinal as the chromophore. The results show that the Schiff base

formed upon the binding of all-*trans*-retinal red shifts 60 nm upon protonation. This makes the conjugated polyene aldehyde a suitable target as a fluorophore precursor for CRBP_{II}. If the absorption of the protonated Schiff base formed by the fluorophore is in the red and near infra-red region and the Schiff base is much more blue-shifted, then the background fluorescence would be drastically reduced. The large binding cavity of CRBP_{II} allowed for the exploration of several ligands as potential fluorophores.

Figure 42 shows the general structure of the potential fluorophores. The idea is to have a polyene system just as all-*trans*-retinal and an electron donating group, which is in conjugation to the polyene system. This would create a push pull system that increases the conjugation along the polyene. The increase in conjugation increases the double bond character of the alternate single bonds thus rigidifying the system and reducing the non-radiative relaxation. This would result in an increase in the quantum efficiency of the system. With CRBP_{II} being amenable to mutagenesis, the binding pocket can be modified to provide a more rigid environment for the fluorophore. This would allow for a further increase in the quantum efficiency

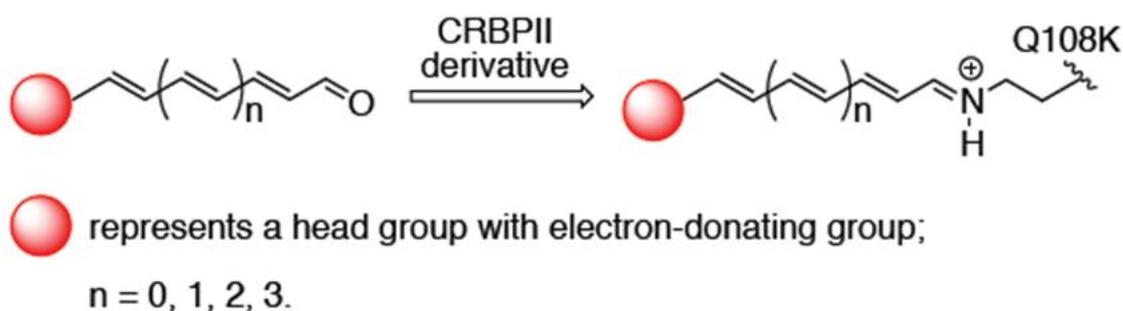


Figure 42: General structure of the fluorophores to be in the development of CRBP_{II} into a fluorescent tag.

III.4 CRYSTALLIZATION OF CRBP2 AS A FLUORESCENT PROTEIN TAG

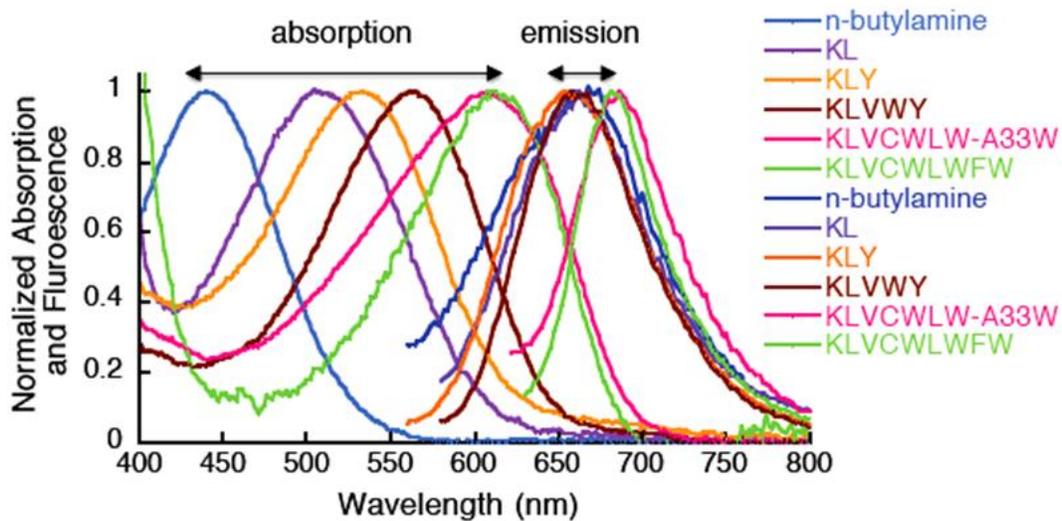
III.4.1 CHARACTERIZATION OF DIFFERENT POLYENE FLUOROPHORE BOUND TO CRBP2 MUTANTS

As discussed in previous chapters, wavelength regulation of retinal-PSB bound to CRBP2 mutants can be observed as a result of different protein-chromophore interactions, both steric and electrostatic. Retinal and retinal-PSB have been reported to have 0.02% quantum efficiency in solution. This would be ideal in the development of the retinal bound CRBP2 into a fluorescent tag. The idea would be that since the quantum efficiency of the chromophore outside of the binding pocket is low then background would not be much of an issue. However, once the chromophore is bound as the PSB it would be rigidified and could result in an increase in the quantum efficiency. In 2011, a rhodopsin based fluorescent probe was developed to probe the electrostatic potential across the membrane of *E. coli*.²⁷ The quantum efficiency of this system is 0.2%. This showed a ten fold increase when compared to the chromophore in solution. This system could serve as a platform to test the retinal-PSB CRBP2 system. Dr. Wenjing Wang then proceeded to determine the quantum efficiency of several CRBP2 mutants bound to retinal (**Table 15**). The results showed that the quantum efficiencies were comparable to what was observed for the chromophore in solution as well as the rhodopsin fluorescent tag. In addition, even though the different mutants had different . excitation wavelengths, the emission wavelengths did not vary as much (**Figure 43**).

To this end, several of the CRBP2 mutants were characterized using different fluorophores (**Figure 44**). The fluorophores are similar in structure to retinal in that they possess a polyene chain with an aldehyde group. This would allow for this system to have the same binding mode as retinal, a protonated Schiff base.

Table 15: Fluorescent characterization of retinal-PSB.

| Retinal PSB | Excitation (nm) | Emission (nm) | Quantum Yield (%) |
|-------------------|-----------------|---------------|-------------------|
| n-Butylamine | 440 | 660 | 0.033 |
| KL | 506 | 614/660 | 0.029 |
| KL:R58Y | 533 | 654 | 0.12 |
| KL:T51V:R58Y | 561 | 660 | 0.14 |
| KL:T51V:R58Y:Y19W | 563 | 656 | 0.14 |
| KL:T51V:R58Y:A33W | 566 | 662 | 0.13 |
| KL:T51V:R58F | 561 | 664 | 0.12 |
| KLVCWLW:A33W | 605 | 684 | 0.079 |
| KLVCFLWF:A33W | 610 | 676 | 0.18 |

**Figure 43:** Normalized UV-VIS and fluorescence spectra of retinal-PSB with *n*-butylamine and different CRBP II mutants.³⁰

The fluorophores used have electron-donating groups resulting in the formation of a push-pull system. The head group of the merocyanine polyene is more electron donating as compared to the azulene polyene, and so forms a stronger push pull system. The syntheses of these fluorophores were reported by Dr. Lee.²⁸

As shown in **figure 44c** the formation of the protonated Schiff base of the azulene polyene and n-butylamine, resulted in a 100 nm red shift. The incubation of several CRBP2 mutants with the azulene polyene gave a wide absorption range, 622 nm to 841 nm.³⁰ This was very interesting as the azulene polyene could be an excellent target for near-IR fluorescence. However, the fluorescence of the azulene polyene-CRBP2 complex had lower quantum efficiencies than retinal.

This phenomenon has dampened the prospects for using the azulene polyene as a fluorescent tag. The other fluorophore of choice was the merocyanine polyene aldehyde, which forms a stronger push pull system. Cyanine compounds are well known for their fluorescent properties.²⁹ **Figure 44c** shows the absorption of the merocyanine polyene to be 467 nm in ethanol and it red shifts upon PSB formation with n-butylamine in PBS buffer to 576 nm. Further red shifting is observed when this fluorophore is bound to the protein (602 nm).

The absorbance of the three species is distinct and the extinction coefficient increases as you go from the free aldehyde to the merocyanine polyene complexed with CRBP2, making this an excellent target for designing a fluorescent tag. The disadvantage, however, is the fact that there is only a 30 nm difference in wavelength between the PSB in solution and the PSB in the binding pocket of the protein. This will lead to a significant amount of background fluorescence.

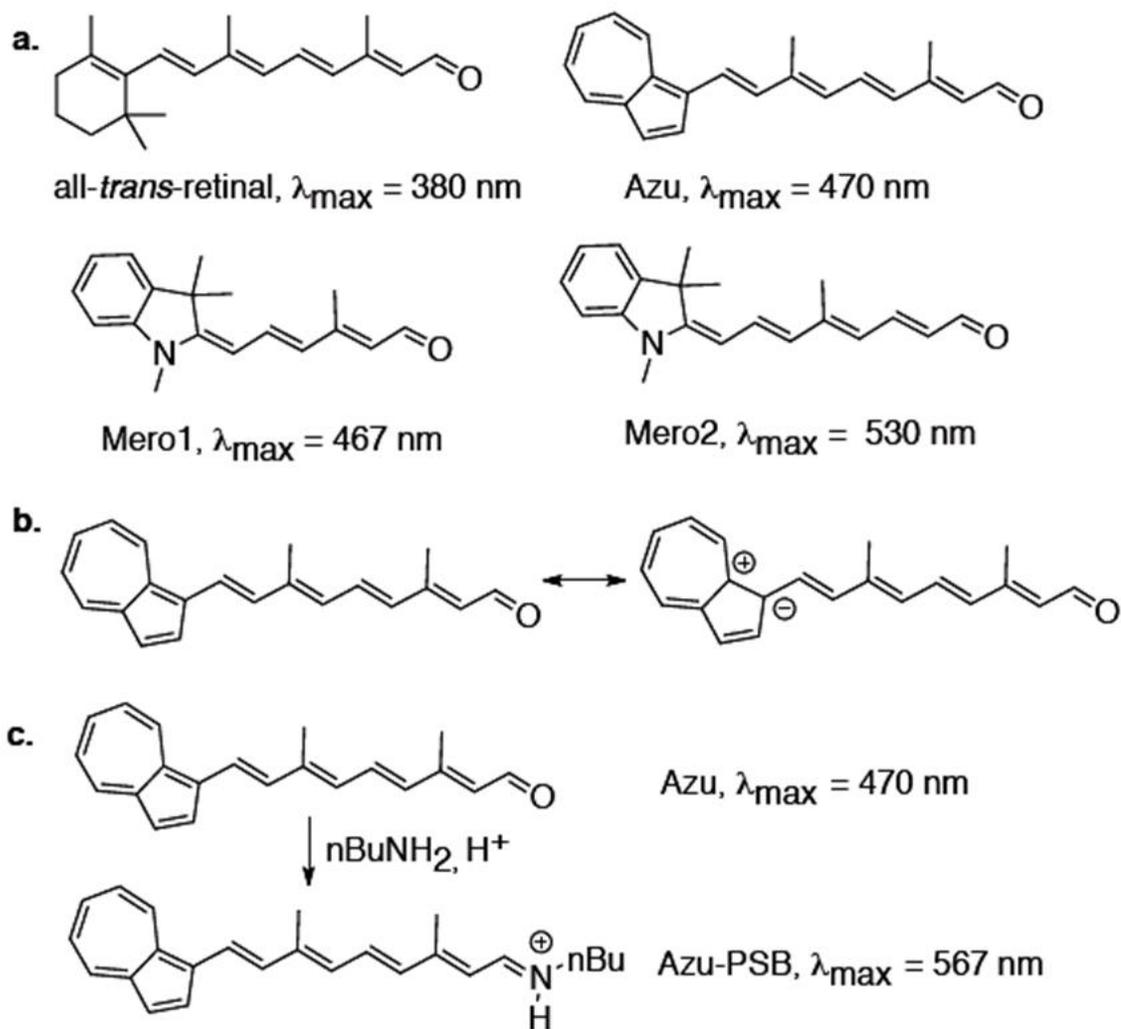


Figure 44: **a.** Structures of the different fluorophores.^{28 30} **b.** Resonance of the azulene polyene head group. **c.** Formation of a protonated Schiff base of azulene polyene with n-butylamine

The incubation of merocyanine polyene with different CRBP_{II} mutants did not result in wavelength regulation. This is due to the fact that upon formation of the PSB, the positive charge is delocalized towards the two nitrogen ends and as a result increases the population of the planar form due to less single bond/double bond alteration (**Figure 45**). The restricted rotation about the bonds also limits the number of vibrational energy levels, which increases the extinction coefficient and makes the peak sharper.

The ability to modulate wavelength in the retinal-PSB system is due to the fact that the positive charge can be localized towards the Schiff base or the ionone ring. This is not the case with the merocyanine polyene.

Even though we were not able to observe wavelength regulation with the merocyanine polyene, the quantum efficiencies of the CRBP_{II} mutants bound to the merocyanine polyene were much higher than those reported for the retinal bound counterparts. The double mutant Q108K:K40L (KL) when bound to retinal had a quantum efficiency of 0.029% while the same mutant bound to merocyanine polyene is 17%.

III.4.2 EFFECT OF THE T51V MUTATION ON THE QUANTUM EFFICIENCY OF CRBP_{II} MUTANTS BOUND TO MEROCYANINE POLYENE

Intrigued by the results for the KL mutant other CRBP_{II} mutants were bound to the merocyanine aldehyde and the quantum efficiencies were determined. An interesting finding was that the introduction of the mutation T51V (Q108K:K40L:T51V) resulted in the doubling of the quantum efficiency with respect to KL, 17% to 32%. This phenomenon begged the question of how the mutation of a single site from a polar threonine to a hydrophobic valine could result in such a drastic increase in the quantum efficiency. Delightfully we were able to grow crystals of several CRBP_{II} mutants bound to the merocyanine polyene in the same crystal screen used when growing the CRBP_{II} mutants bound to retinal. The data obtained for both the mutants KL and KL:T51V revealed that the fluorophore has two different trajectories (**Figure 46**).

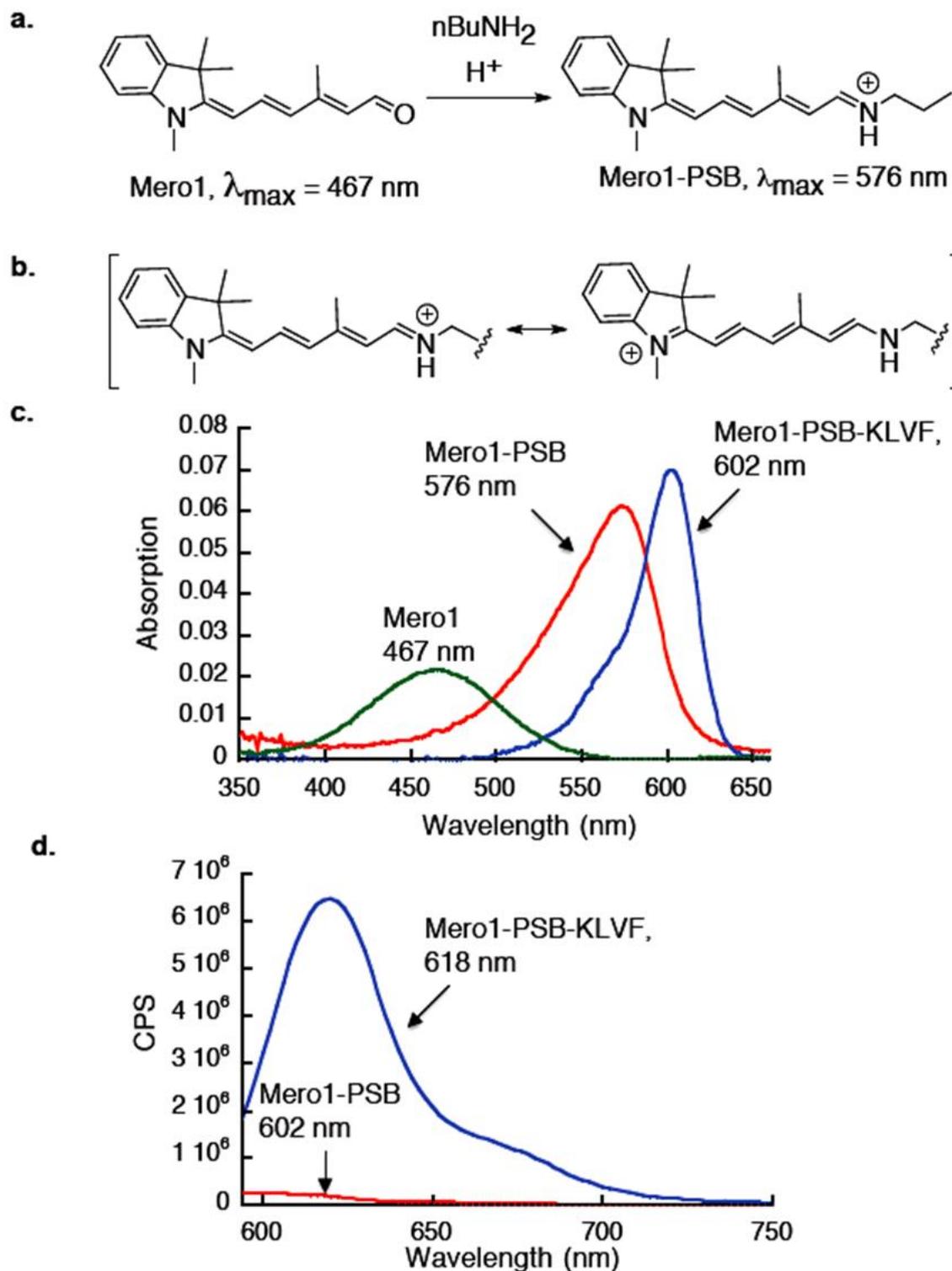


Figure 45: Characterization of merocyanine polyene.³⁰ **a.** Formation of merocyanine polyene-PSB. **b.** Resonance structure of merocyanine polyene-PSB. **c.** UV-Vis spectra of merocyanine polyene, PSB and PSB formed with CRBP II mutant KLVF (Q108K:K40L:T51V:R58F). **d.** Fluorescence spectra of merocyanine polyene.

In the case of the mutant KL, the head group of the fluorophore is in close proximity to three aromatic residues, Tyr 60, Phe 57 and Phe 16 as shown in **Figure 46b**. When the T51V mutation was introduced the data shows that the head group of the fluorophore becomes sandwiched between Tyr 60 and Phe 16 and has also move closer to Phe 57 as shown in **Figure 46c**. This observation then led to the conclusion that the fluorophore is more rigidified in the binding pocket of the mutant KL:T51V leading to the drastic increase in the quantum efficiencies.

As shown in **Figure 46a**, very little change is observed in the protein backbone except for the residue at the 51 position. The mutation of threonine to valine resulted in the inward movement of the residue towards the Schiff base region of the fluorophore. Further examination of the crystal structures revealed that it is not just a simple movement of the residue at position 51. The mutation of threonine to valine resulted in the disruption of the water-mediated hydrogen bonding network between Thr 51, Thr 53 and the main chain carbonyl of Val 62 (**Figure 47**).

The crystal structure shows that the disruption of the water mediated hydrogen bonding network resulted in the outward movement of the β -sheet bearing the Val 62 residue by 1 Å which led to the inward movement of valine at position 51 towards the Schiff base region of the fluorophore. This movement was attributed to the hydrophobic interactions that are enhanced by the introduction of valine. Based on these observations, the mutation T53V was done to give the tetramutant KL:T51V:T53V.

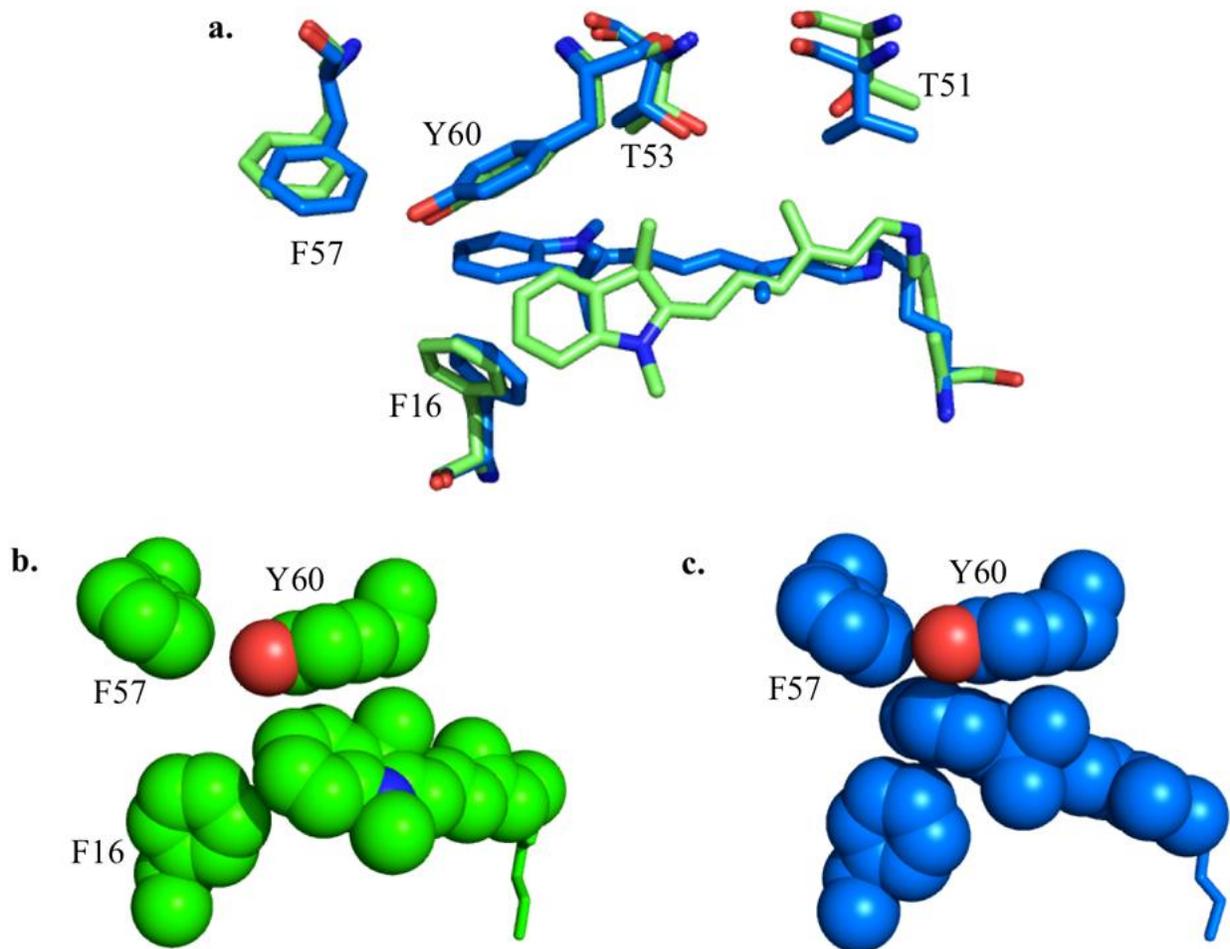


Figure 46: Comparison of Q108K:K40L (KL) and Q108K:K40L:T51V (KL:T51V) bound to merocyanine polyene. **a.** Overlaid crystal structures of KL (green) and KL:T51V (blue). Space filling models of **b.** KL and **c.** KL:T51V showing the surrounding residues.

Table 16: Showing the distances of the aromatic residues in the mutants KL versus KL:T51V bound to merocyanine.

| Residue | KL | KL:T51V |
|---------|-------|---------|
| Phe 16 | 4.2 Å | 3.6 Å |
| Phe 57 | 6.7 Å | 4.1 Å |
| Tyr 60 | 4.0 Å | 3.6 Å |

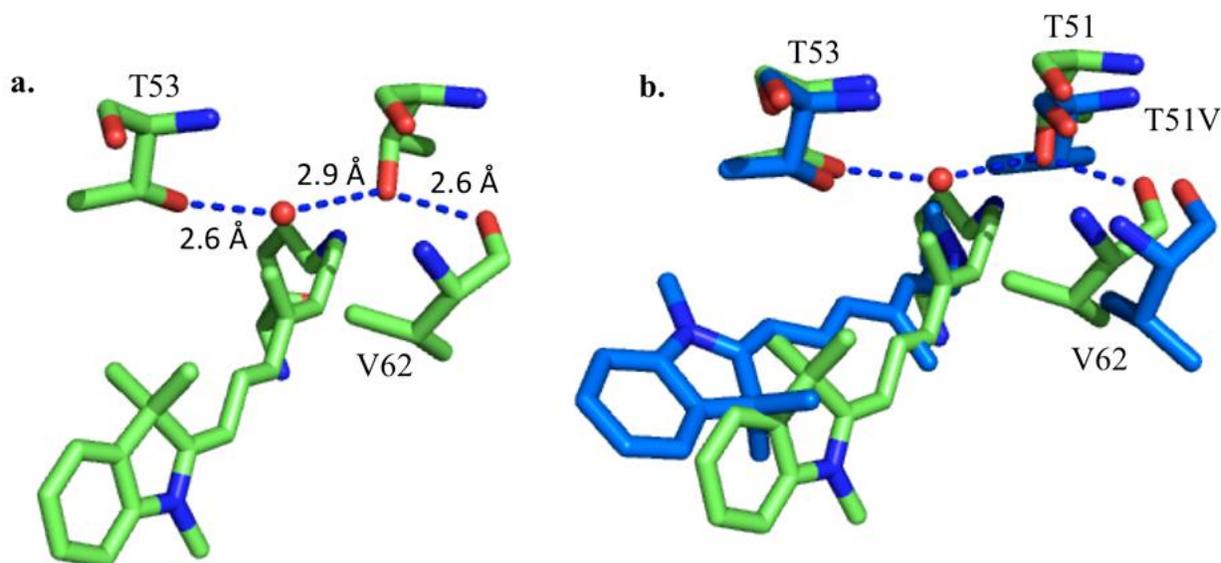


Figure 47: a. Water mediated hydrogen bonding network in Q108K:K40L (KL) between T51, T53 and main chain carbonyl of V62. **b.** Overlaid crystal structures of KL (green) and KL:T51V (blue).

However, the quantum efficiencies of both KL:T51V (32%) and KL:T51V:T53V (30%) were the same. This may mean that increasing the hydrophobicity is not the way to approach increasing the quantum efficiency of the system. Examination of the crystal structure revealed that the orientation of the merocyanine polyene in the binding pocket of the tetramutant, KL:T51V:T53V, is the same as that observed for KL:T51V (**Figure 48c**). From these observations, it seems that mutation at position 51 greatly affects the quantum efficiency of this system by tightening the packing around the fluorophore.

Initially it was thought that the change in the trajectory of the fluorophore in the mutant KL:T51V was due to possible steric interactions that resulted from the inward movement of T51V. However, modeling of the crystal structures revealed that both conformations of the fluorophore fit quite well in the binding pocket of KL without any unfavorable steric interactions. However, the KL merocyanine conformation does not fit well in the KL:T51V structure. The data revealed unfavorable steric interactions, as there is only a 2.6 Å distance

between the merocyanine C20 methyl and the V51 carbonyl group. So the conformation observed for the mutant KL is lower energy for some reason, and the introduction of the T51V mutation enforces the higher energy and higher quantum efficiency KL: T51V conformation.

In order to confirm the conformation of the fluorophore in KL:T51V, another set of crystals were grown in the same conditions as mentioned before. However, in this case crystals were taken from the wells corresponding to three different pHs (4.0, 4.5 and 4.8). Refinement of the data sets obtained showed that the conformation of the fluorophore is the same. The data obtained was for pH 4.0 and 4.5, the crystals from pH 4.8 did not give a good data set. Therefore, the reason for the fluorophore to adopt this conformation in KL:T51V could be that this conformation is the most stable of all the possible conformations within the binding pocket.

III.4.3 EXPLORING THE EFFECTS OF DIFFERENT T51 MUTATIONS ON THE QUANTUM EFFICIENCY OF CRBPII MUTANTS BOUND TO MEROCYANINE POLYENE

The results from trying to increase the quantum efficiency of the system by increasing the hydrophobicity of the binding pocket proved that position 51 is key as other mutations had little effect. Therefore mutations at position 51 were screened to determine the effect it would have on the quantum efficiency (quantum yield).

The residues used for this screening were larger, non-polar and polarizable. The idea was that since the introduction of valine at position 51 resulted in the movement of the beta sheet towards the fluorophore, then the introduction of larger residues would possibly rigidify the system even more resulting in an increase in the quantum yield.

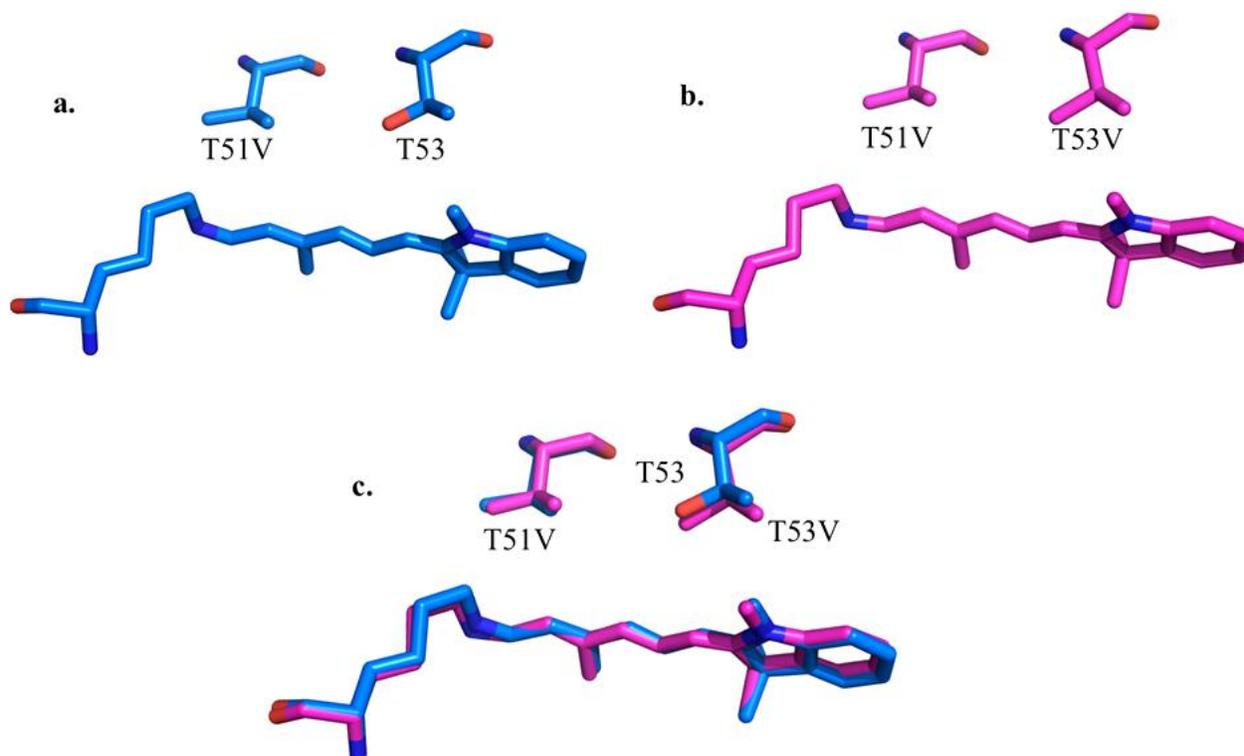


Figure 48: **a.** Crystal structure of KL:T51V. **b.** Crystal Structure of the tetra-mutant KL:T51V:T53V. **c.** Overlaid crystal structures of KL:T51V (blue) and KL:T51V:T53V (magenta).

Mutagenesis was done on two groups of CRBP_{II} mutants, the KL:T51V series and the KL:T51V:R58F series. The R58F mutation was chosen because the results from the series of mutants made to explore wavelength regulation using retinal as ligand proved the R58F mutation was one of the mutations that resulted in more red shifting. The crystallographic data of a few mutants with this mutation (Dr. Zahra Nossoni) showed that the R58F goes deeper into the binding pocket and stacks closer to the ionone region of retinal. Therefore if the same phenomenon occurs for the merocyanine polyene then the introduction of this mutation could result in an increase in the quantum efficiency. This proved not to be the case when the quantum yields of KL:T51V and KL:T51V:R58F were determined. The quantum efficiencies were 32% and 26% respectively.

Table 17 shows the quantum efficiencies of the different mutants of the KL:T51X series. The results proved to be rather disappointing as the T51V mutation proved to be the most effective mutation with quantum yield 32%.

Table 17: Quantum efficiency of the different variants of the KL:T51X series.

| Mutant | Maximum absorption (nm) | Quantum yield |
|---------------|--------------------------------|----------------------|
| KL | 600 | 17% |
| KL:T51V | 601 | 32% |
| KL:T51I | 600 | 21% |
| KL:T51M | 598 | 28% |
| KL:T51F | 603 | 32% |

Table 18: Quantum efficiency of the different variants of the KL:T51X:R58F series.

| Mutant | Maximum absorption (nm) | Quantum yield |
|---------------|--------------------------------|----------------------|
| KL:T51A:R58F | 601 | 17% |
| KL:T51V:R58F | 596 | 26% |
| KL:T51I:R58F | 600 | 37% |
| KL:T51M:R58F | 598 | 36% |
| KL:T51F:R58F | 604 | 37% |

The introduction of isoleucine and methionine actually had quantum yields less than that observed for the T51V mutation, while phenylalanine was the same.

Based on the results obtained for the KL:T51X series, the KL:T51X:R58F series was considered. The effects of the mutations appeared to be affected by the incorporation of the R58F mutation. **Table 17** shows the quantum efficiencies of the mutant of the KL:T51X:R58F series.

Even though the introduction of the R58F mutation to the KL:T51V (**Table 17**) resulted in a slight decrease in the quantum efficiency, when the same mutations as the KL:T51X series, were made at position 51, the results showed an enhancement of the quantum efficiencies. It is interesting to note that the introduction of smaller residues such as alanine resulted in drastic decreases in the quantum yield. KL:T51A:R58F (**Table 18**) and KL (**Table 17**) have the same quantum yield of 17%. This phenomenon shows the importance of the size of the residue at position 51. Attempts to incorporate smaller residues (glycine) and larger residues (tryptophan) were futile. As shown in **Table 18**, the introduction of alanine at position 51 resulted in the same quantum yield as the mutant KL. Structural data of the mutants of the KL:T51X:R58F series would be quite insightful. Crystals were obtained for the mutants KL:T51A:R58F and KL:T51I:R58F. However, the data for KL:T51I:R58F:L117E did not show the presence of the fluorophore in the binding pocket. This mutant (KL:T51I:R58F:L117E) has a quantum yield of 38% which is the same as KL:T51I:R58F which has a quantum yield of 37%. Therefore the data for this mutant should be a good representation of what is observed for KL:T51I:R58F:L117E.

Examination of the data revealed that the introduction of the R58F mutation resulted in a different orientation of the fluorophore (**Figure 49**). This change in the trajectory of the fluorophore is a result of the inward movement of the phenylalanine at position 58 which led to

the movement of tyrosine 60 towards the middle of the polyene region of the fluorophore. As mentioned before, the movement of the phenylalanine at position 58 was observed in the case of mutants bound to retinal. However, while this phenomenon resulted in an increase in the wavelength in the case of retinal; it did not affect the fluorescence of the merocyanine polyene-CRBPII complex. KL:T51V:R58F:L117E mutant was used as the base of comparison for the KL:T51X:R58F series as we were unable to obtain usable data for the KL:T51V:R58F mutant. The quantum yields of KL:T51V:R58F and KL:T51V:R58F:L117E being the same, 26%, shows that the L117E mutation does not affect the fluorescence. The role of L117E will be discussed later.

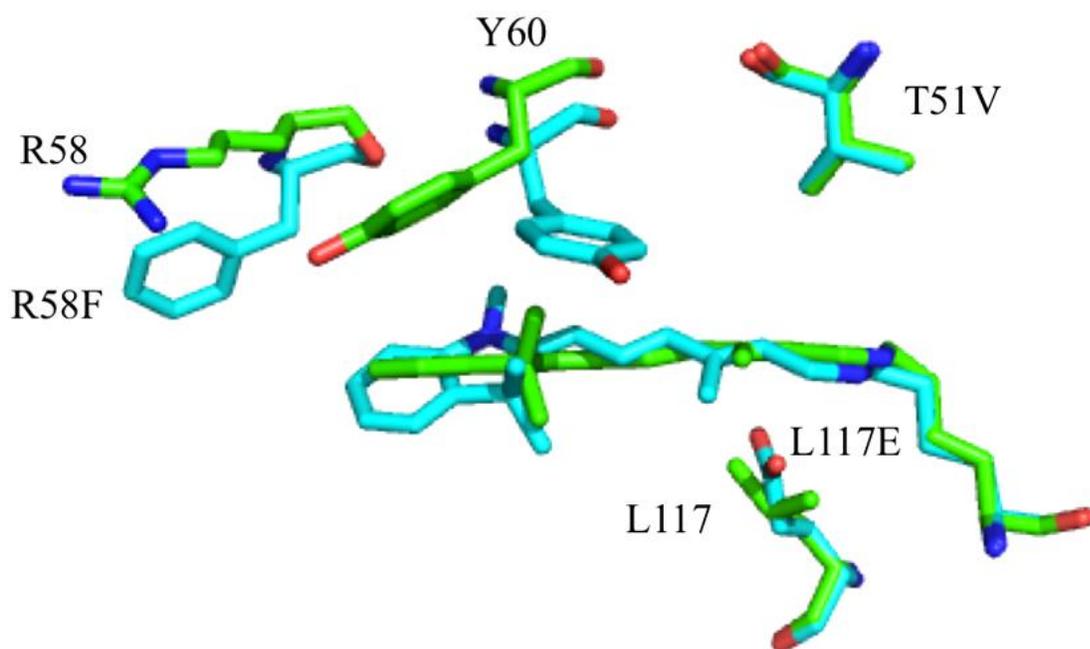


Figure 49: Overlaid structures of KL:T51V (green) and KL:T51V:R58F:L117E (cyan) highlighting the residues in the binding pocket.

The comparison of the data obtained for the mutants, KL:T51V:R58F:L117E and KL:T51I:R58F:L117E revealed that the trajectory of the fluorophore did not change and there were insignificant changes in the protein backbone (**Figure 50**). It can be concluded then that the

changes in the protein backbone observed with the introduction of valine is not changed by the introduction of the larger residues. Due to the increase in the size of the residues the fluorophore is even more rigidified. This increase in the hydrophobic packing against the fluorophore resulted in an 11% increase in the quantum yield.

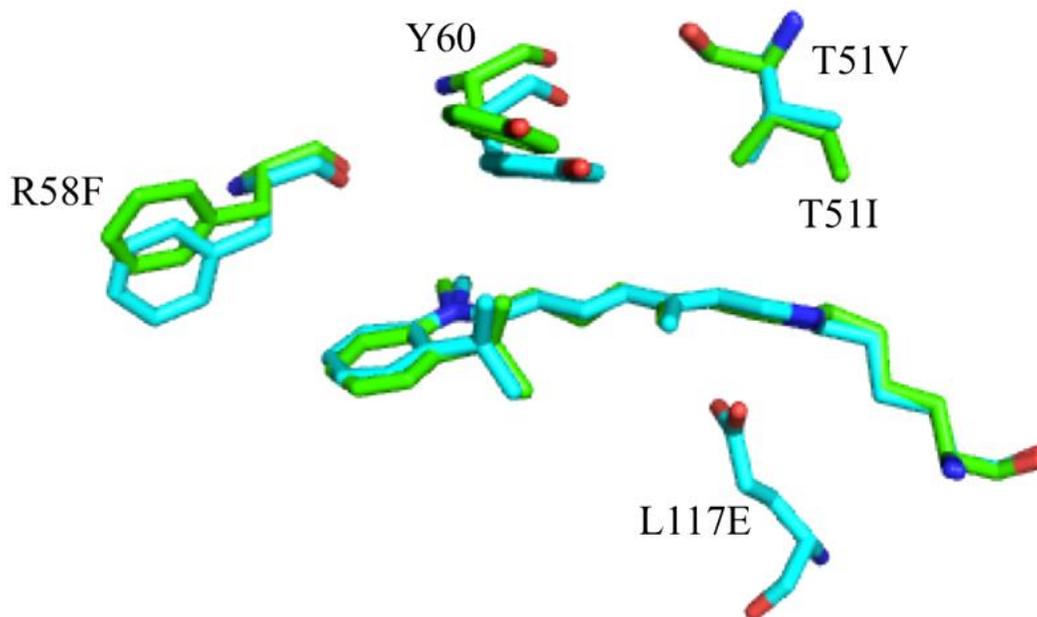


Figure 50: Overlaid structures of KL:T51V:R58F:L117E (blue) and KL:T51I:R58F:L117E (green) highlighting the residues in the binding pocket.

This effect however was nullified by the introduction of alanine, KL:T51A:R58F. Though a hydrophobic residue, alanine is small, even smaller than the wild type residue threonine. As a result, the fluorophore adopted the open conformation observed for the double mutant KL, which is reflected in the quantum yield of both mutants, 17% (**Figure 51**).

It is interesting to note that the same movement of the protein backbone observed for KL:T51V was observed for KL:T51A:R58F as the introduction of alanine will disrupt the water mediated hydrogen bonding network just the same (**Figure 47**). The draw back however is that

alanine is small and is about 5Å away from the fluorophore therefore, the interaction between the fluorophore and the alanine is minimal. This further emphasizes the effects of the residue at position 51 on the quantum yield.

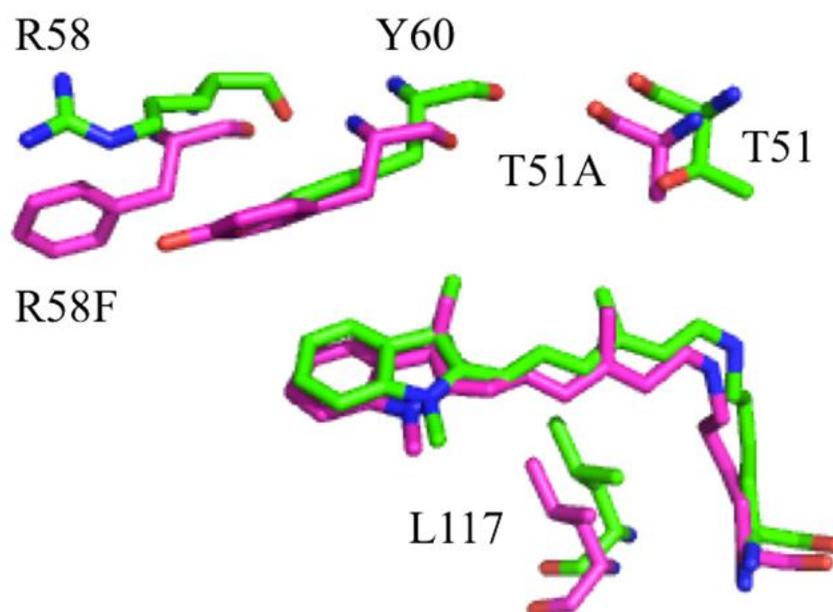


Figure 51: Overlaid structures of KL:T51A:R58F (magenta) and KL (green) highlighting the residues in the binding pocket.

III.4.4 EFFECTS OF L117E MUTATION ON THE UTILITY OF CRBP II MUTANTS BOUND TO MEROCYANINE POLYENE AS A FLUORESCENT TAG

The CRBP II mutant KL:T51V:R58F was used as the model for designing this system as a fluorescent tag. The brightness of KL:T51V:R58F bound to merocyanine is about 25 fold higher than that of the merocyanine polyene PSB in solution. The comparison of the merocyanine polyene bound to KL:T51V:R58F with a red fluorescent protein, mRaspberry, which has similar excitation (594 nm) and emission, revealed that our system is three fold brighter. This

phenomenon makes the CRBP_{II} system a formidable alternative to currently available red fluorescent protein tags.

In order to develop this system into a fluorescent tag that can be used *in vivo*, another factor to consider is the kinetics of binding of the merocyanine polyene. UV-VIS- monitored kinetics studies were conducted on three CRBP_{II} mutants (KL, KL:T51V, KL:T51V:R58F) at room temperature. The results obtained were in agreement with the results obtained by Dr. Wenjing Wang. The results show that the KL mutant has the shortest half-life and the half-lives increase dramatically once the T51 mutation is made. This means that PSB formation is hindered by the mutation of T51.

The kinetics of PSB formation could be the result of the binding of the merocyanine polyene, or the SB bond formation, protonation events are rarely rate limiting. The incubation of the CRBP_{II} mutants with merocyanine indicates that binding occurs almost immediately. A red shift is observed as soon as the merocyanine polyene is added. It can then be concluded that the rate determining step is the PSB formation. The increase in the maturation time as a result of the T51 mutation suggests that T51 plays a role in the PSB formation. T51 could be possibly activating and/or orienting the aldehyde in the Burgi-Dunitz trajectory through hydrogen bonding for nucleophilic attack of the amine group of lysine 108.

To investigate the likelihood of the use of KL:T51V:R58F mutant as a fluorescent protein tag, initial *in vivo* assays were done using *E. coli* (Dr. Wenjing Wang). The experiment was carried out using wide-field fluorescent microscopic studies. The cells were grown under typical cell growth conditions for *E. coli*, then induced for expression after which the merocyanine polyene was added. The cells were then incubated at 37^oC with vigorous shaking. The results showed that the merocyanine polyene was able to cross the membrane barrier.

To prove that this system can be used as a general fluorescent tag, a mammalian expression system for CRBP was constructed by Dr. Wang. Constructs were made that contained a CRBP II KL:T51V:R58F, EGFP and RB (RetinoBlastoma) triple fusion starting with the pEGFP-RB plasmid (**Figure 52**).

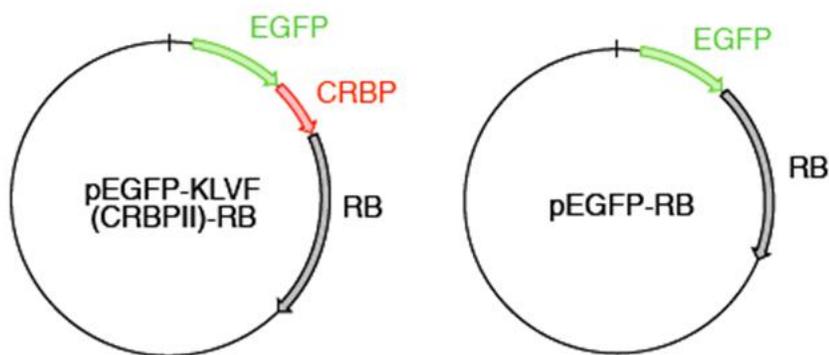


Figure 52: Constructs of pEGFP-KLVF(CRBP II)-RB and pEGFP-RB.

The EGFP provides a positive control for the expression and localization the triple fusion. It also provides a direct comparison of the relative brightness of CRBP-merocyanine fluorescence with EGFP. Retinoblastoma (RB) is a tumor suppressor protein that plays an important role by suppressing tumor cell division.³⁷ This protein is localized in the nucleus. The pEGFP-RB and pEGFP-CRBP II constructs were used as controls. Gratifyingly, red fluorescence was observed only when CRBP II was induced in the construct and merocyanine is added (**Figure 53**).

The drawback of this system is the long incubation and maturation times of KL:T51V:R58F after merocyanine addition. Another factor is the overnight incubation to get rid of the background fluorescence. Strategies need to be developed to avoid the background fluorescence. This will allow for the visualization shortly after the fluorophore is added, instead of waiting for 12 h after the removal of the excess fluorophore. The use of CRBP II mutants with faster kinetics and equivalent quantum yields could be a possible solution. Another solution

could be to use a fluorophore with different spectral characteristics that would avoid the overlap with the background fluorescence.

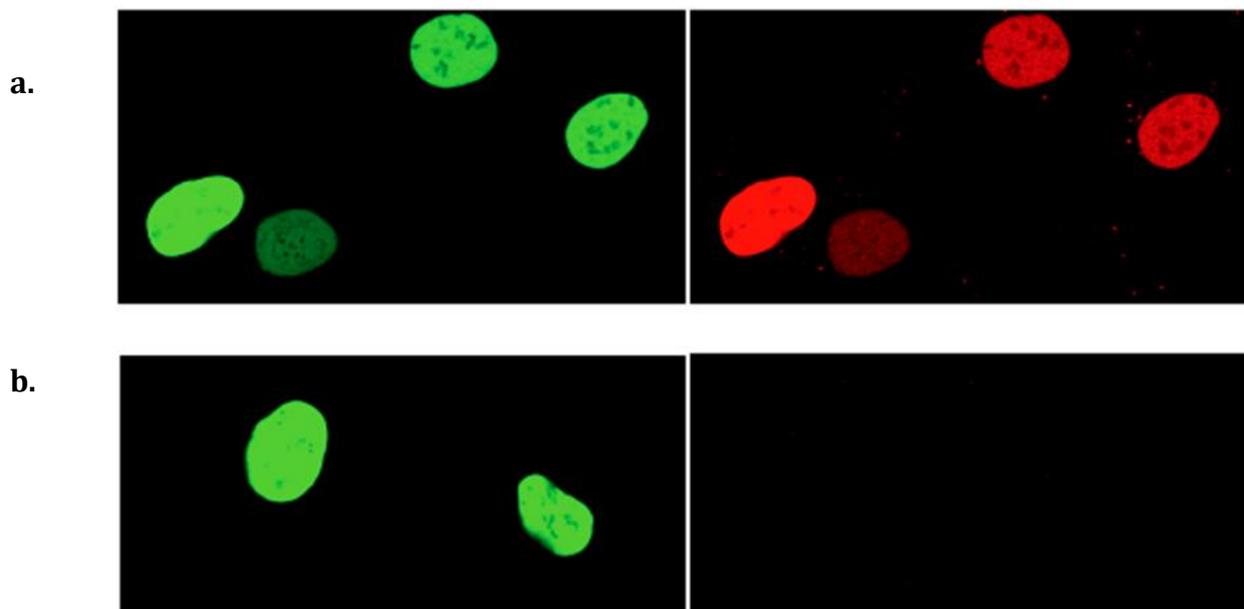


Figure 53: Confocal pictures of **a.** pEGFP-KLVF(CRBPII)-RB construct and **b.** pEGFP-RB construct. EGFP was excited at 488 nm and KLVF-merocyanine polyene was excited at 594 nm.³⁰

Since the mutation of T51 resulted in slow PSB formation, the introduction of an alternative ‘activator residue’ that is in close proximity to the Schiff base region could allow for the decrease in the maturation time. Examination of the crystal structure revealed that the leucine at position 117 would be an ideal place to start (**Figure 54a**). Introduction of an acidic residue at position 51 would nullify the quantum yield increase observed when threonine is replaced with a large hydrophobic residue. Though this mutation resulted in a decrease in the maturation time of the PSB, this system would not be favorable as the system should have similar quantum yields to the KL:T51V:R58F mutant. Therefore the other position to consider is position 117. By mutating

this position it may be possible to retain the increased quantum efficiency of the T51V mutation while still increasing kinetics of the PSB formation.

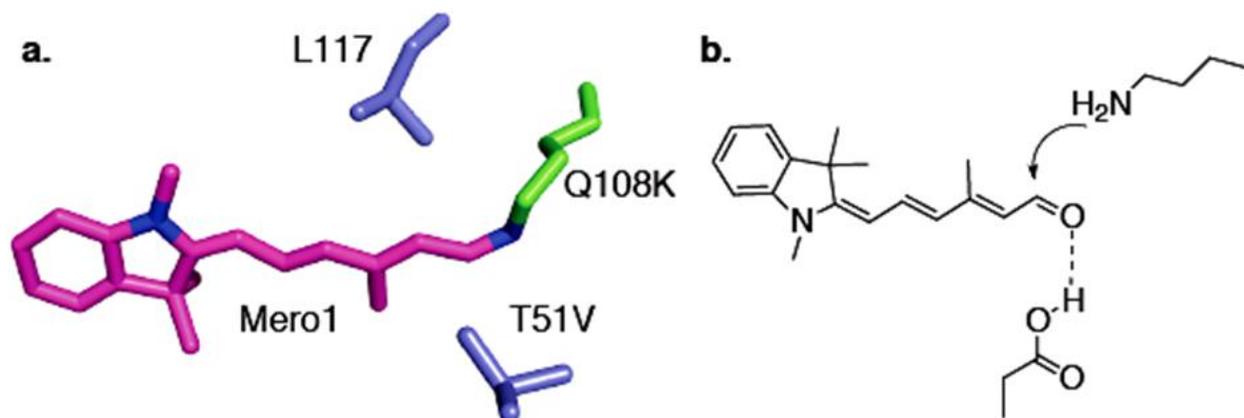


Figure 54: **a.** Crystal structure of KL:T51V with residues L117 and T51V highlighted. **b.** Proposal for the mechanism of activation of the aldehyde by the introduction of an acidic residue at position 51 or 117.

The role of the T51 residue in the mutant KL is assumed to be a hydrogen bond formation between the carbonyl group and the hydroxyl group of the T51. This hydrogen bond orients the aldehyde in a trajectory favorable for PSB formation. This phenomenon could explain the decrease in the rate of PSB formation when T51 is mutated to valine. The absence of this hydrogen bonding element in the T51V mutant results in the fluorophore taking more time to adopt the correct trajectory for PSB formation. This phenomenon could also explain why the merocyanine in the mutant KL:T51V adopts a different conformation than that observed in KL. Attempts were then made to obtain an energy minimized model of the merocyanine aldehyde in the binding pocket of the KL mutant and the results showed that T51 could interact with the aldehyde through a hydrogen bond (**Figure 55**).

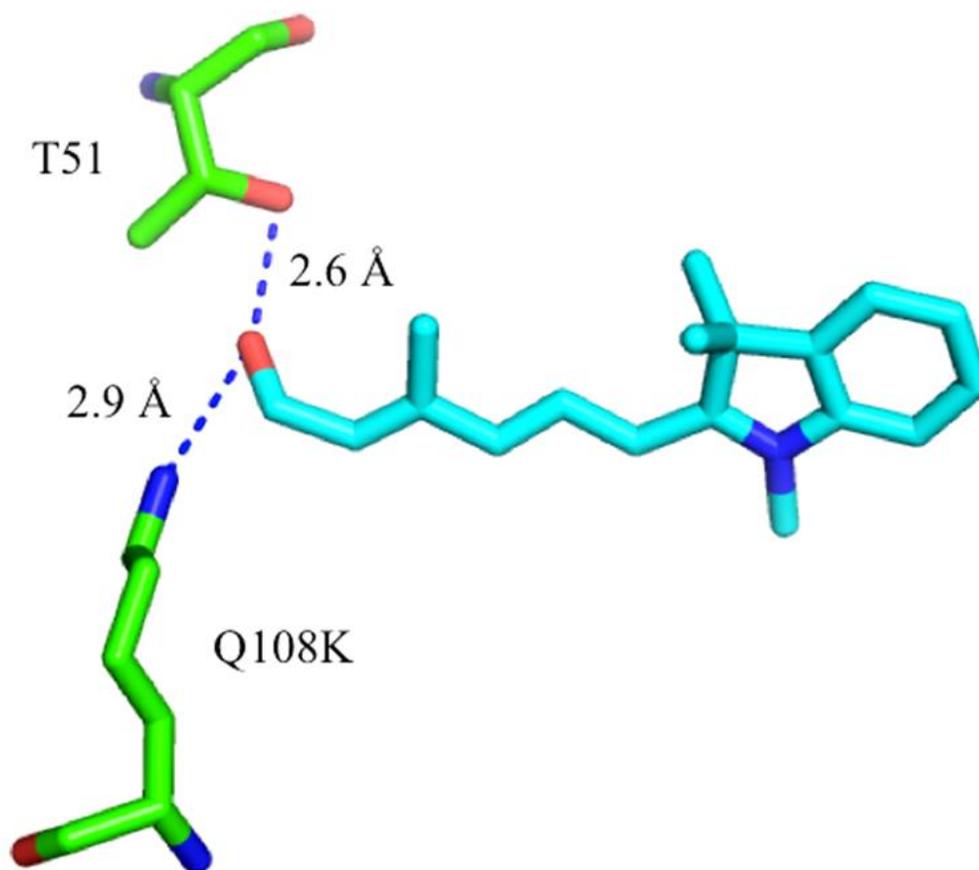


Figure 55: Energy minimized model of KL with merocyanine aldehyde (cyan). The interactions between residues Q108K and T51 are highlighted.

Examination of the structure revealed that the introduction of a polar residue at position 117 could possibly restore the rate. The mutation L117E was then introduced to the mutant KL:T51V:R58F and several other CRBP_{II} mutants and the results showed a dramatic decrease in the maturation time from the order of hours to the order of minutes (**Figure 56**). The conclusion drawn from these results is that the glutamate at position 117 restored and enhanced the role that T51 played in the PSB formation kinetics (**Figure 57**). It is interesting to note that the mutants that possess the L117E were not able to form a PSB with retinal. From UV-VIS studies it was observed that the retinal enters the binding pocket but there was still no evidence of PSB formation even after 24 hours of incubation. This phenomenon would suggest that the

mechanism of how the PSB is formed in the case of the fluorophore, merocyanine is different than that of retinal. Upon observation of an overlay of a CRBP II mutant (KLVCWLW:Q4R) bound to all-*trans*-retinal and KLVF:L117E bound to merocyanine, it was revealed there are unfavorable interactions between the L117E and all-*trans*-retinal (**Figure 57**).

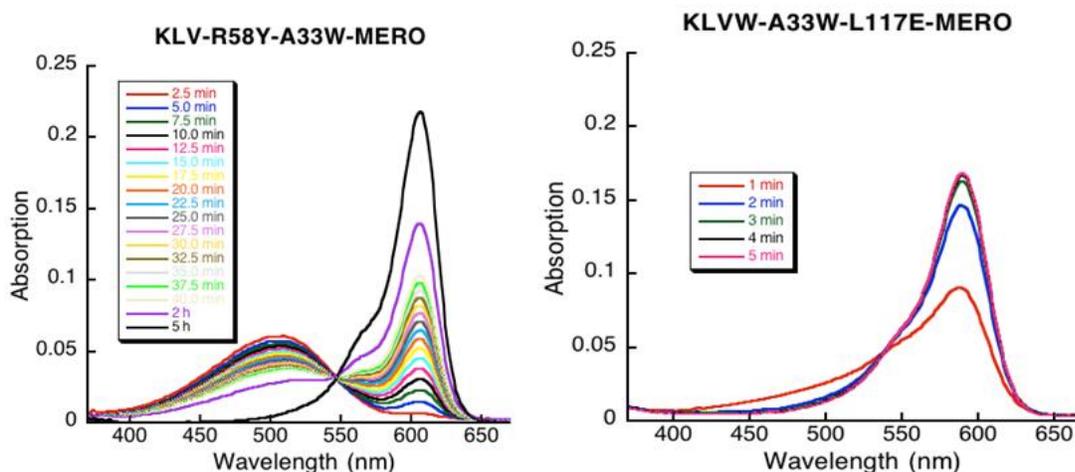


Figure 56: UV-VIS spectra showing the effect of the L117E mutation on the kinetics of PSB formation with merocyanine polyene.³⁰ Mutants are KL:T51V:R58Y:A33W and KL:T51V:R58W:A33W:L117E.

The crystal structures also revealed that the lysine (Q108K) could form a salt bridge with L117E as the distance between the two residues is only 3.0 Å. This would mean that the amine is not free for nucleophilic attack on the aldehyde. The question then is why do the mutants that possess the L117E mutation form a PSB with merocyanine aldehyde? The answer could lie in the fact that the merocyanine polyene is a push pull system; the protonated amine could serve to activate the aldehyde and free up the amine of the lysine (Q108K) for nucleophilic attack (**Figure 61**). This cannot happen with retinal and so the amine remains trapped.

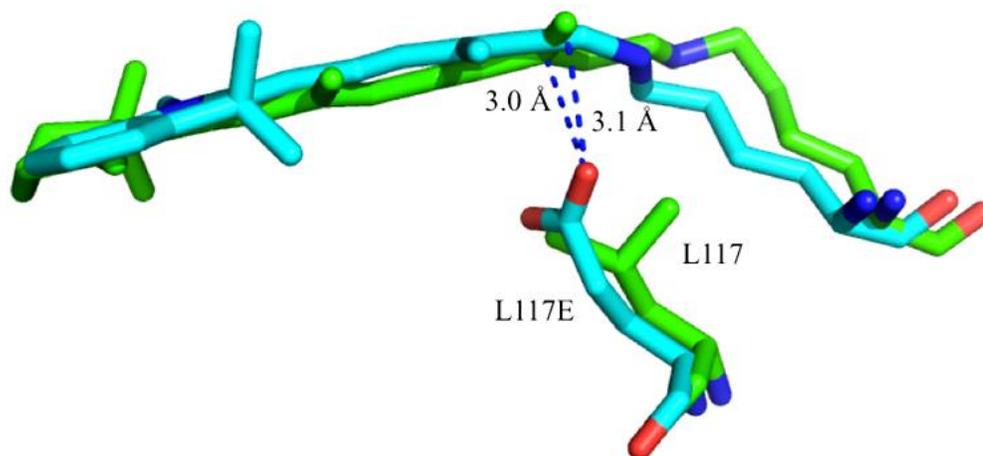


Figure 57: Overlaid structures of KL:T51V:R58F:L117E bound to merocyanine (cyan) and KLVCWLW:Q4R (PDB ID: 4EEJ) bound to all-*trans*-retinal (green) highlighting the residues L117 and L117E. The distances between the polyene and the oxygen of L117E is 3.0 Å. The distance between the C20 methyl group and the oxygen of L117E is 3.1 Å.

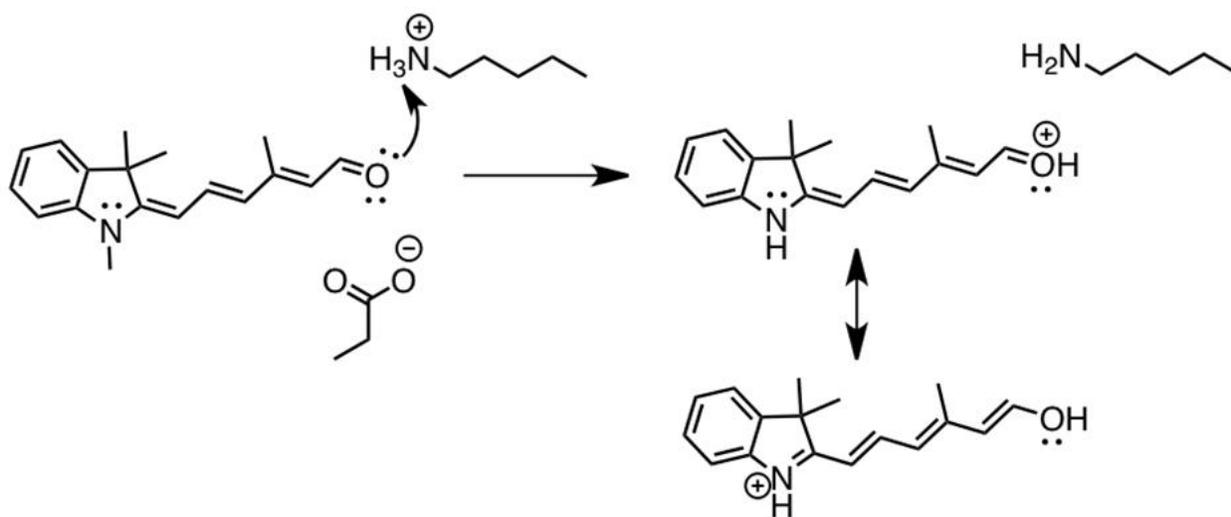


Figure 58: Proposed mechanism for activation of merocyanine polyene via a proton transfer in the presence of L117E mutation.

In order to explore this hypothesis, mutants that have the L117E mutation but do not possess the lysine (Q108K) mutation were made. This would allow for assessment of the interactions before PSB formation. The Q108K mutation was changed to Q108L in the mutant Q108K:K40L:T51V:R58F:L117E. Disappointedly, attempts to obtain crystals of the mutant Q108L:K40L:T51V:R58F:L117E with retinal or merocyanine polyene were unsuccessful.

III.4.5 CHARACTERIZATION OF JULOLIDINE BOUND TO CRBP_{II} MUTANTS AS FLUOROPHORIC pH SENSORS

As previously discussed, the CRBP_{II} mutants have been characterized using retinal, azulene and merocyanine. Julolidine is another fluorophore that was used. Initial results were very similar to those obtained for azulene, even though the absorption maximum was in the near IR region, the quantum yields were low. With the positive results for the kinetics of the L117E mutants with merocyanine the decision was made to revisit the other fluorophores. It was then discovered by Tetyana Berbasova in the Borhan lab that when julolidine is bound to KL:T51V:R58F:L117E the absorption maximum changes with pH up to about 40 nm (595 nm to 634 nm). Several other mutants were tried and it was determined that for all the other mutants two mutations were crucial for this property. These mutations are L117E and Q4X. This result has provided the platform for developing CRBP_{II} mutant as a fluorophoric pH sensors.

UV-vis studies were done for several mutants and the results showed that the mutation at the Q4 position plays an important role. The Q4 mutations were initially done for the wavelength regulation studies. It was discovered that this mutation played a role in the geometry of the PSB formed. In the case of julolidine the Q4 mutation proved to affect the ability to shift the absorption maxima with changing pH, except in the case of the mutant KL:T51V:R58F:L117E. Only when Q4 is mutated to hydrophobic residues is this change in absorption maximum observed (**Figure 59**). It was also determined that if Q4 is mutated but L117 is not this effect is completely nullified. This further reiterates the point that L117E is essential for this process. Attempts to obtain crystals of the mutants for which this phenomenon is observed were unsuccessful.

Further studies are currently being undertaken by Tetyana Berbasova and Elizabeth Santos from the Borhan lab to determine other residues that could lead to a more drastic change in wavelength.

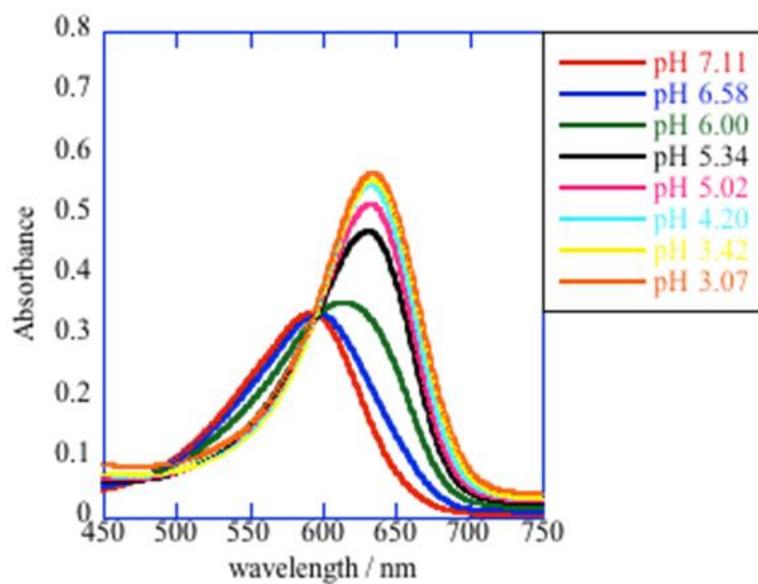


Figure 59: UV-Vis spectra showing the pH titration of the mutant Q108K:K40L:T51V:R58F:L117E:Q4F (KLVF-L117E-Q4F).

III.5 MATERIALS AND METHODS

UV-vis spectra were recorded using a Cary 300 Bio WinUV, Varian spectrophotometer.

Fluorescence was recorded with a Fluorolog-3 (Instruments S.A., Inc.) fluorometer.

III.5.1 Mutagenesis

All mutations were made using *Stratagene's QuickChange Site-Directed Mutagenesis Protocol*.

Table 19: PCR Reaction Solutions.

| DNA template | Primer forward | Primer Reverse | Turbo DNA polymerase | dNTP (10mM of each) | Buffer (10x) | Water (dd) | Total volume |
|--------------|----------------|----------------|----------------------|---------------------|--------------|------------|--------------|
| 70 ng | 20 pmol | 20 pmol | 1 unit | 1 μ L | 5 μ L | / | 50 μ L |

Table 20: PCR temperature control cycles.

| | | |
|------|------------------------|--------|
| 1 x | 95 °C | 3 min |
| 20 x | 95 °C | 30 sec |
| | (T _m -4) °C | 1 min |
| | 72 °C | 10 min |
| 1 x | 72 °C | 10 min |
| 1 x | 25 °C | 5 min |

Note: T_m represents the melting temperature of the primer, which was calculated using the website shown below.

<http://www.promega.com/techserv/tools/biomath/calc11.htm>

III.5.2 WT-CRBPII gene sequence

ATG CCA AAC TTC TCT GGC AAC TGG AAA ATC(10) ATC CGA TCG GAA AAC TTC
GAG GAA TTG CTC AAA(20) GTG CTG GGG GTG AAT GTG ATG CTG AGG AAG ATT
GCT GTG GCT GCA GCG TCC AAG CCA GCA(40) GTG GAG ATC AAA CAG GAG GGA
GAC ACT TTC(50) TAC ATC AAA ACC TCC ACC ACC GTG CGC ACC ACA GAG ATT
AAC TTC AAG GTT GGG GAG GAG TTT GAG GAG CAG ACT GTG GAT GGG AGG
CCC TGT AAG AGC CTG GTG AAA TGG GAG AGT GAG AAT AAA ATG GTC TGT
GAG CAG AAG CTC CTG AAG GGA GAG GGC CCC AAG ACC TCG TGG ACC AGA
GAA CTG ACC AAC GAT GGG GAA CTG ATC CTG ACC ATG ACG GCG GAT GAC
GTT GTG TGC ACC AGG GTC TAC GTC CGA GAG TGA

III.5.3 WT-CRBPII amino acid sequence

MPNFSGNWKIIRSENFEE LLKVLGVNV MLRKIAVAAASKPA
VEIKQEGDTFYIKTSTTVRTTEINFKVGEEFEEQTV DGRPCK
SLVKWESENK MVCEQKLLKGE GPKTSWTRELTNDGELILT
MTADDVVCTRVYVRE Stop

III.5.4 Primers used for mutagenesis

T51D

Forward: 5'-GGTGATAACTTCAAGGATAAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTATCCTTGAAGTTATCACC-3'

T51I

Forward: 5'-GGTGATAACTTCAAGATTA AAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTAATCTTGAAGTTATCACC-3'

T51V

Forward: 5'-GGTGATAACTTCAAGGTAAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTACCTTGAAGTTATCACC-3'

T51C

Forward: 5'-GGTGATAACTTCAAGTGTA AAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTACACTTGAAGTTATCACC-3'

T51N

Forward: 5'-GGTGATAACTTCAAGAACA AAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTGTTCTTGAAGTTATCACC-3'

T51M

Forward: 5'-GGTGATAACTTCAAGATGAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTTACCTTGAAGTTATCACC-3'

T51A

Forward: 5'-GGTGATAACTTCAAGGCAAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTCGTCTTGAAGTTATCACC-3'

T51G

Forward: 5'-GGTGATAACTTCAAGGGCAAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTCCGCTTGAAGTTATCACC-3'

T51F

Forward: 5'-GGTGATAACTTCAAGTTTAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTAAACTTGAAGTTATCACC-3'

T51W

Forward: 5'-GGTGATAACTTCAAGTGGAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTCCACTTGAAGTTATCAC-3'

T53V

Forward: 5'-CTTCAAGACAAAAGTCACTAGCACATTCCG-3'

Reverse: 5'-CGGAATGTGCTAGTGACTTTTGTCTTGAAG-3'

III.5.5 XL1-Blue competent cells preparation

XL1-Blue *E. coli* cells were streaked on agar plate with tetracycline as antibiotic (12.5 mg/L) and incubated overnight at 37 °C.

The following day a colony was picked from the plate using a sterile wooden applicator. The colony was inoculated in sterile LB solution (10 mL), with tetracycline (12.5 mg/L), in a shaker at 220 RPM overnight. The overnight culture (1 mL) was transferred to sterile LB solution (200 mL) with tetracycline (12.5 mg/L). The solution was kept shaking at 220 RPM at 37 °C for about 2 h or until the OD was between 0.4 and 0.8. At the same time, a 500 mL centrifuge bottle was sterilized with 20% bleach for 2 h and rinsed with sterile water 6 to 7 times before use.

The cells were harvested by centrifugation at 3500 RPM for 5 min. at 4 °C. The cells were resuspended with sterile sodium chloride solution (0.9 %, 100 mL). The cells were harvested by centrifugation at 3000 RPM for 4 min at 4 °C.

The following steps were then followed:

A: The cell pellet was resuspended with calcium chloride solution (100 mM, 15% glycerol v/v, 10 mL) on ice and was incubated for at least 30 min. The cells were fractioned in sterile eppendorf tubes in 100 μ L portions and flash frozen with liquid nitrogen.

B: The cell pellet was resuspended with calcium chloride solution (100 mM, 50 mL) and incubated on ice for 30 min. The cells were spun down by centrifugation at 3000 RPM for 4 min and then resuspended with calcium chloride solution (100 mM, 15% glycerol v/v, 10 mL) on ice. The cells were fractioned in sterile eppendorf tubes in 100 μ L portions and flash frozen with liquid nitrogen.

Note: The cells prepared using protocol **A** uses one step less than protocol **B**. This results in cells that are less competent. The cells that are used for the transformation of the PCR product was prepared using protocol **B** as the cells need to be highly competent. This is due to the fact that the DNA from the PCR reaction is nicked and does not go through the cell membrane with the same efficiency as circular DNA during the heat shock process. Therefore protocol **B** is a better protocol for competent cells preparation. This protocol was used to prepare the BL21 competent cells the only difference is chloramphenicol was used instead of tetracycline.

III.5.6 Heat shock transformation

The DpnI digested PCR solution (5 μ L) was added to competent cells (50 or 100 μ L). The cells were then incubated on ice for 10 to 30 min, followed by heat shock at 42 °C for 50 sec. The cells were put back on ice immediately. Sterile LB solution (500 μ L) was added to the cells and shaken at 37 °C for 1 h. The cells were then harvested by centrifugation at 5000 RPM

for 1 min. The cells were resuspended in LB (50 μ L) and then plated on an agar plate with the appropriate antibiotics.

III.5.7 Protein expression and purification in pET-17b system

CRBP^{II} DNA was transformed into BL21 (DE3) pLysS competent cells. A single colony was picked from the transformation plate and inoculated in 15 or 50 mL of LB with concentration of ampicillin at 100 mg/L and chloramphenicol at 27 mg/L at 37 °C overnight. The next day the growth was transferred to 1 L of sterile LB media with the same concentration of ampicillin and chloramphenicol. The media was incubated at 37 °C until the OD₆₀₀ reached 0.4 to 0.8.

Isopropyl-1-thio-D-galactopyranoside (IPTG, Gold Biotechnology) was added to a final concentration of 1mM to induce protein expression, and the cell culture was shaken at RT or 16 °C, depending on the mutant, overnight.

Cells were harvested by centrifugation at 5000 RPM for 10 min in Beckmann J2-21M/E centrifuge and resuspended with 50 mL Tris binding buffer (10 mM Tris, pH=8.0), followed by 3 min sonication (3 time, 1 min per pulse, pulse=60%, power 60%) to lyse the cells. DNase (20 μ L, Roche, recombinant, 25 unit/ μ L) was added to the lysed cells and incubated for 30 min. The lysed cell solution was then spun down for 30 min at 5000 RPM. The supernatant was then applied to the fast Q ion exchange column, which was pre-equilibrated (washed with 50 mL 2 M NaCl, and equilibrated with 2 x 50 mL washes of 10 mM Tris, pH=8.0, buffer). The protein was then washed with 10 mM Tris, pH=8.0 and eluted using the elution buffer, 10 mM Tris, 100 mM NaCl, pH=8.0. The eluted protein was desalted using an Amicon filter membrane (10 kDa cut off) and then applied to a source Q Fast Protein Liquid Chromatography (FPLC). The protein

was subjected to a gradient sodium chloride solution from 0 mM to 1 mM at pH=8.0 (25 mM Tris and Tris-HCl). The protein was eluted at ~40 mM sodium chloride.

III.5.8 Crystallization Methods

The purified protein was then concentrated to concentrations ranging from 10 – 20 mg/ml. Two equivalents of the ligand (merocyanine) were then added and the resulting mixture incubated at 4 °C overnight or 6-8 h at room temperature (depends on the mutant). The UV-vis of the sample was taken to ensure that all the protein is bound. Two 24 well plates were set up using the hanging drop method. The condition in each well is shown in **Table 21**. The drop is comprised of 1µL of the well solution and 1µL of the protein sample.

Table 21: Crystallization Conditions.**a) Box 5**

| Reservoir | PEG (4K) 40% | Salt NH₄OAc (4M) | Buffer NaOAc (1M) pH= 4.0 | Buffer NaOAc (1M) pH= 4.5 | Buffer NaOAc (1M) pH= 4.8 | H₂O |
|------------------|-------------------------|--|--|--|--|-----------------------|
| A1 | 625 | 25 | 100 | | | 250 |
| A2 | 750 | 25 | 100 | | | 125 |
| A3 | 750 | 25 | 100 | | | 125 |
| A4 | 800 | 25 | 100 | | | 75 |
| A5 | 875 | 25 | 100 | | | 0 |
| A6 | 950 | 25 | 100 | | | 0 |
| B1 | 625 | 25 | | 100 | | 250 |
| B2 | 750 | 25 | | 100 | | 125 |
| B3 | 750 | 25 | | 100 | | 125 |
| B4 | 800 | 25 | | 100 | | 75 |
| B5 | 875 | 25 | | 100 | | 0 |
| B6 | 950 | 25 | | 100 | | 0 |
| C1 | 625 | 25 | | | 100 | 250 |
| C2 | 750 | 25 | | | 100 | 125 |
| C3 | 750 | 25 | | | 100 | 125 |
| C4 | 800 | 25 | | | 100 | 75 |
| C5 | 875 | 25 | | | 100 | 0 |
| C6 | 950 | 25 | | | 100 | 0 |
| D1 | 625 | 37.5 | 100 | | | 212.5 |
| D2 | 750 | 37.5 | 100 | | | 87.5 |
| D3 | 750 | 37.5 | 100 | | | 87.5 |
| D4 | 800 | 37.5 | 100 | | | 37.5 |
| D5 | 875 | 37.5 | 100 | | | 0 |
| D6 | 950 | 37.5 | 100 | | | 0 |

Table 21 (cont'd)

b) Box 6

| Reservoir | PEG (4K) 40% | Salt NH₄OAc (4M) | Buffer NaOAc (1M) pH= 4.0 | Buffer NaOAc (1M) pH= 4.5 | Buffer NaOAc (1M) pH= 4.8 | H₂O |
|------------------|-------------------------|--|--|--|--|-----------------------|
| A1 | 625 | 0 | 100 | | | 275 |
| A2 | 750 | 0 | 100 | | | 150 |
| A3 | 750 | 0 | 100 | | | 150 |
| A4 | 800 | 0 | 100 | | | 100 |
| A5 | 875 | 0 | 100 | | | 25 |
| A6 | 950 | 0 | 100 | | | 0 |
| B1 | 625 | 25 | 100 | | | 268.75 |
| B2 | 750 | 25 | 100 | | | 143.75 |
| B3 | 750 | 25 | 100 | | | 143.75 |
| B4 | 800 | 25 | 100 | | | 93.75 |
| B5 | 875 | 25 | 100 | | | 18.75 |
| B6 | 950 | 25 | 100 | | | 0 |
| C1 | 625 | 25 | 100 | | | 262.5 |
| C2 | 750 | 25 | 100 | | | 137.5 |
| C3 | 750 | 25 | 100 | | | 137.5 |
| C4 | 800 | 25 | 100 | | | 87.5 |
| C5 | 875 | 25 | 100 | | | 12.5 |
| C6 | 950 | 25 | 100 | | | 0 |
| D1 | 625 | 37.5 | | | 100 | 257.5 |
| D2 | 750 | 37.5 | | | 100 | 132.5 |
| D3 | 750 | 37.5 | | | 100 | 132.5 |
| D4 | 800 | 37.5 | | | 100 | 82.5 |
| D5 | 875 | 37.5 | | | 100 | 7.5 |
| D6 | 950 | 37.5 | | | 100 | 0 |

***Total volume in each well = 1ml**

***The volumes indicated in the table are in μ L.**

Table 22: Data Collection and refinement statistics. The ligand is merocyanine in each case.

| | KL | KL-T51V | KL-T53V-T51V | KL-R58F-T51I | KL-R58F-T51V-L117E | KL-R58F-T51A | KL-T53C-T51V | KL-T51V-R58W-Y19W | KL-R58F-T51V | KL-R58F-T51I-L117E |
|---|-----------|----------------|---------------------|---------------------|---------------------------|---------------------|---------------------|--------------------------|---------------------|---------------------------|
| Data collection | | | | | | | | | | |
| Space group | P1 | P1 | P1 | P1 | P1 | P1 | P1 | P1 | P1 | P1 |
| Cell Dimensions | | | | | | | | | | |
| <i>a</i> (Å) | 29.63 | 29.62 | 29.47 | 29.77 | 29.87 | 30.23 | 29.64 | 29.68 | 36.05 | 29.77 |
| <i>b</i> (Å) | 35.95 | 35.91 | 35.64 | 36.03 | 35.97 | 36.40 | 35.80 | 36.21 | 54.54 | 36.03 |
| <i>c</i> (Å) | 63.69 | 64.07 | 64.20 | 63.81 | 63.87 | 65.27 | 63.87 | 64.08 | 69.59 | 63.81 |
| Wavelength (Å) | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| Resolution (Å) | 50 – 1.34 | 50 – 1.32 | 50 – 1.37 | 50 – 2.1 | 50 – 1.6 | 50 – 1.5 | 50 – 1.4 | 50 – 1.4 | 50 – 1.6 | 50 – 1.5 |
| Completeness | 96% | 96% | 94% | 87% | 92% | 95.4% | 93% | 94% | 91% | 82% |
| Refinement | | | | | | | | | | |
| Resolution (Å) | 1.5 | 1.5 | 1.5 | 2.1 | 1.7 | 1.5 | 1.5 | 1.5 | 1.7 | 2.0 |
| No. of Reflections | 35556 | 34941 | 36613 | 11631 | 27654 | 35366 | 41795 | 37685 | 45139 | 12820 |
| R _{factor} / R _{free} | 19.7/24.3 | 20.2/24.5 | 20.0/23.4 | 23.1/33.8 | 21.5/26.1 | 22/26.7 | 18/22 | 18.5/22.1 | 29/37 | 20/26 |

Table 22 (Cont'd)

| | KL | KL-T51V | KL-T53V-T51V | KL-R58F-T51I | KL-R58F-T51V-L117E | KL-R58F-T51A | KL-T53C-T51V | KL-T51V-R58W-Y19W | KL-R58F-T51V | KL-R58F-T51I-L117E |
|-------------------------|-----------|----------------|---------------------|---------------------|---------------------------|---------------------|---------------------|--------------------------|---------------------|---------------------------|
| No. of atoms | | | | | | | | | | |
| Protein | 2190 | 2190 | 2190 | 2190 | 2190 | 2190 | 2190 | 2190 | 4380 | 2190 |
| Ligand | 48 | 48 | 48 | 20 | 20 | 20 | 40 | 40 | 0 | 0 |
| Water | 143 | 198 | 174 | 63 | 137 | 139 | 217 | 187 | 163 | 113 |
| Acetate | 1 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| B-factors | | | | | | | | | | |
| Overall | 23.242 | 22.106 | 29.135 | 48.308 | 21.607 | 25.934 | 21.741 | 19.265 | 34.471 | 24.304 |
| Main (Chain A) | 19.765 | 17.159 | 24.112 | 46.899 | 19.102 | 22.801 | 17.511 | 16.611 | 32.206 | 22.011 |
| Main (Chain B) | 20.705 | 19.400 | 26.800 | 44.453 | 18.035 | 21.119 | 19.169 | 16.072 | 32.501 | 21.482 |
| Main (Chain C) | | | | | | | | | 34.905 | |
| Main (Chain D) | | | | | | | | | 31.920 | |
| Ligand | 25.4 | 26.0 | 29.6 | 63.9 | 57.187 | 64.981 | 26.973 | 22.509 | | |
| R.M.S deviations | | | | | | | | | | |
| Bond lengths (Å) | 0.024 | 0.021 | 0.021 | 0.012 | 0.020 | 0.022 | 0.023 | 0.024 | 0.021 | 0.016 |
| Bond Angles (°) | 2.344 | 2.346 | 2.325 | 1.540 | 2.268 | 2.304 | 2.370 | 2.713 | 2.146 | 1.855 |

III.5.9 Quantum efficiency determination of CRBP_{II} mutants bound to merocyanine polyene aldehyde

The quantum efficiency of CRBP_{II} mutants bound to merocyanine polyene aldehyde were determined by comparing to Oxazine 170 (Ox-170) and Oxazine 1 (Ox-1), which have quantum efficiencies of 58.5% and 13.9% respectively, when excited at 565 nm.

Three samples of Oxazine 170 (Ox-170) and Oxazine 1 (Ox-1) with absorption at 565 nm ranging from 0.01 to 0.1 were excited at 565 nm, fluorescence emission spectra were collected from 575 nm to 800 nm. The total photons emitted, which are the integration of the emission spectrum, were plotted against the absorption at 565 nm. A linear function was fitted to the points to get a slope for CRBP_{II} mutants bound to merocyanine polyene aldehyde.^{31 32}

The quantum efficiency can be efficiency can be derived from the following equation:

*Quantum Efficiency = Slope of the merocyanine polyene-PSB / slope of the standard (Oxazine 170 (Ox-170) and Oxazine 1 (Ox-1)) * quantum yields of the standards.*

REFERENCES

REFERENCES

1. (a) Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R.Y., Review - The fluorescent toolbox for assessing protein location and function. *Science* **2006**, *312* (5771), 217-224; (b) Cho, W. H.; Stahelin, A. V., Membrane-protein interactions in cell signaling and membrane trafficking. *Annu. R. Biom.* **2005**, *34*, 119-151.
2. Zimmer, M., GFP: from jellyfish to the Nobel prize and beyond. *Chem. Soc. Rev.* **2009**, *38* (10), 2823-2832.
3. Cody, C. W.; Prasher, D. C.; Westler, W. M.; Prendergast, F.G.; Ward, W. W., Chemical Structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry* **1993**, *32*(5), 1212-1218.
4. Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F.G.; Cormier, M. J., Primary structure of the *aerquores- victoria* green-fluorescent protein. *Gene* **1992**, *111* (2), 229-233.
5. Ormo, M.; Cubitt, A. B.; Kaillo, K.; Gross, L. A.; Tsien, R.Y.; Remington, S. J., Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **1996**, *273* (5280), 1392-1395.
6. Jackson, S. E.; D Craggs, T.; Huang, J. R., Understanding the folding of GFP using biophysical techniques. *Expert Rev. Proteomic* **2006**, *3* (5), 545-559.
7. Jung, G.; Zumbusch, A., Improving autofluorescent proteins: Comparative studies of the effective brightness of green fluorescent protein (GFP) mutants. *Microsc. Res. Techniq.* **2006**, *69* (3), 175-185.
8. Heim, R.; Cubitt, A. B.; Tsien, A. V., Improved green fluorescence. *Nature* **1995**, *373* (6516), 663-664.
9. (a) Helm, R.; Prasher, D. C.; Tsien, R. Y., Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *P. Natl. Acad. Sci. USA* **1994**, *91* (26), 12501-12504; (b) Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y., Understanding, Improving and Using Green Fluorescent Proteins. *Trends Biochem. Sci.* **1995**, *20* (11), 448-455; (c) Hem, R.; Tsien, R. V., Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* **1996**, *6* (2), 178-182.
10. Mishin, A. S.; Subach, F. V.; Yampolsky, I. V.; King, W.; Lukyanov, K. A.; Verkhusha, V. V., The first mutant of the *Aequorea victoria* green fluorescent protein that forms a red chromophore. *Biochemistry* **2008**, *47* (16), 4666-4673.

11. Yarbrough, D.; Wachter, R. M.; Kallio, K.; Matz, M. V.; Remington, S. J., Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-angstrom resolution. *P. Natl. Acad. Sci. USA* **2001**, *98* (2), 462-467.
12. Bevis, B. J.; Glick, B. S., Rapidly maturing variants of the *Discosoma* red fluorescent protein (DsRed). *Nature Biotechnology* **2002**, *20*(1), 83-87.
13. Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y., Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp red fluorescent protein. *Nat. Biotechnol.* **2004**, *22* (12), 1567-1572.
14. Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A., Fluorescent Proteins and their applications in Imaging Living Cells and Tissues. *Physiol. Rev.* **2010**, *90* (3), 1103-1163.
15. (a) Rayan, C.; Guet, J. E.; Taulier, N.; Pincet, F.; Urbach, W., Recent Applications of Fluorescence Recovery after Photobleaching (FRAP) to Membrane Bio-Macromolecules. *Sensors-Basel* **2010**, *10* (6), 5927-5948; (b) McNally, J. G., Quantitative FRAP in analysis of molecular binding dynamics in vivo. *Method. Cell Biol.* **2008**, *85*, 329-351.
16. Anderson, M. T.; Tjioe, I. M.; Lorincz, M. C.; Parks, D. R.; Herzenberg, L. A.; Nolan, G. P.; Herzenberg, L. A., Simultaneous fluorescence-activated cell sorter analysis of two distinct transcriptional elements within a single cell using engineered green fluorescent proteins. *P. Natl. Acad. Sci. USA* **1996**, *93* (16), 8508-8511.
17. (a) Susperreguy, S.; Prendes, L. P.; Desbats, M. A.; Charo, N. L.; Brown, K.; MacDougald, . A.; Kerppola, T.; Schwartz, J.; Piwien-Pilipuk, C., Visualization by BiFC of different C/EBP beta dimers and their interaction with HP1 alpha reveals a differential subnuclear distribution of complexes in living cells. *Exp Cell Res* **2011**, *317* (6), 706-723; (b) Kerppola, T. K., Biomolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu Rev Biophys* **2008**, *37*, 465-487; (c) Kerppola, T. K., Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. *Nat. Protoc.* **2006**, *1* (3), 1278-1286; (d) Hu, C. D.; Chinenov, Y.; Ranez-Carrozzi, V. R.; Grinberg, A. V.; Kerppola, T. K., geIFRET and BiFC: Visualization of protein interactions in vitro and in living cells. *Biophysical Journal* **2002**, *82* (1), 2a-2a; (e) Hoffmann, C.; Gaietta, G.; Bunemann, M.; Adams, S. R.; Oberdorff-Maass, S.; Behr, B.; Vilardaga, J. P.; Tsien, R. Y.; Eisman, M. H.; Lohse, M. J., A FAsH-based FRET approach to determine G protein - coupled receptor activation in living cells. *Nat. Methods* **2005**, *2* (3), 171-176.
18. Flors, C., Photoswitching of monomeric and dimeric DNA-intercalating cyanine dyes for super-resolution microscopy applications. *Photoch. Photobio. Sci* **2010**, *9* (5), 643-648.
19. (a) Schermelleh, L.; Heintzmann, R.; Leonhardt, H., A guide to super-resolution fluorescence microscopy. *J. Cell Biol* **2010**, *190* (2), 165-175; (b) Huang, B., Super-resolution optical microscopy: multiple choices. *Curr. Opin. Chem. Biol.* **2010**, *14* (1), 10-14; (c) Huang, B.; Bates,

M.; Zhuang, X. W., Super-Resolution Fluorescence Microscopy. *Annu. Rev. Biochem.* **2009**, *78*, 993-1016.

20. (a) Faro, A. A.; Carpentier, P.; Jonasson, G.; Pompidor, G.; Arcizett, D.; Demachy, I.; Bourgeois, D., Low-Temperature Chromophore Isomerization Reveals the Photoswitching Mechanism of the Fluorescent Protein Padron. *J. Am. Chem. Soc.* **2011**, *133* (41), 16362-16365; (b) Subach, F. V.; Patterson, G. H.; Renz, M.; Lippincott-Schwartz, J.; Verkhusha, V. V., Bright Monomeric Photoactivatable Red Fluorescent Protein for Two-Color Super-Resolution sptPALM of Live Cells. *J. Am. Chem. Soc.* **2010**, *132* (18), 6481-6491; (c) Nienhaus, K.; Nar, H.; Heilker, R.; Wiedenmann, J.; Nienhaus, G. U., Trans-cis isomerization is responsible for the red-shifted fluorescence in variants of the red fluorescent protein eqFP611. *J. Am. Chem. Soc.* **2008**, *130* (38), 12578.

21. (a) Henderson, J. N.; Gepshtein, R.; Heenan, J. R.; Kallio, K.; Huppert, D.; Remington, S. J., Structure and Mechanism of the Photoactivatable Green Fluorescent Protein. *J. Am. Chem. Soc.* **2009**, *131* (12), 4176; (b) Subach, F. V.; Malashkevich, V. N.; Zencheck, W. D.; Xiao, H.; Filonov, G. S.; Alma, S. C.; Verkhusha, V. V., Photoactivation mechanism of PAmCherry based on crystal structures of the protein in the dark and fluorescent states. *P Natl. Acad. Sci. USA* **2009**, *106* (50), 21097-21102; (c) Brakemann, I.; Stiel, A. C.; Weber, G.; Andresen, M.; Testa, I.; Grotjohann, T.; Leutenegger, M.; Plessmann, U.; Urlaub, H.; Eggeling, C.; Wahl, M. C.; Hell, S. W.; Jakobs, S., A reversibly photoswitchable GFP-like protein with fluorescence excitation decoupled from switching. *Nat. Biotechnol.* **2011**, *29* (10), 942-U132.

22. Daniels, D. S.; Woo, T. T.; Luu, K. X.; Noll, D. M.; Clarke, N. D.; Pegg, A. E.; Tam, J. A., DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat. Struct. Mol. Biol* **2004**, *11* (8), 714-720.

23. (a) George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K., Specific labeling of cell surface proteins with chemically diverse compounds. *J. Am. Chem. Soc.* **2004**, *126* (29), 8896-8897; (b) Juillerat, A.; Gronemeyer, I.; Keppler, A.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K., Directed evolution of O-6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. *Chem. Biol.* **2003**, *10* (4), 313-317; (c) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K., A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **2003**, *21* (1), 86-89; (d) Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K., Labeling of fusion proteins with synthetic fluorophores in live cells. *P. Natl. Acad. Sci. USA* **2004**, *101* (27), 9955-9959.

24. Juillerat, A.; Heinis, C.; Sielaff, I.; Barnikow, J.; Jaccard, H.; Kunz, B.; Terskikh, A.; Johnsson, K., Engineering substrate specificity of O-6-alkylguanine-DNA alkyltransferase for specific protein labeling in living cells. *Chembiochem* **2005**, *6* (7), 1263-1269.

25. Heinis, C.; Schmitt, S.; Kindermann, M.; Godin, G.; Johnsson, K., Evolving the substrate specificity of O-6-alkylguanine-DNA alkyltransferase through loop insertion for applications in molecular imaging. *ACS Chem. Biol.* **2006**, *1* (9), 575-584.

26. Griffin, B. A.; Adams, S. R.; Tsien, R. Y., Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **1998**, *281* (5374), 269-272.
27. Kralj, J. M.; Hochbaum, D. R.; Douglass, A. D.; Cohen, A. E., Electrical Spiking in *Escherichia coli* Probed with a Fluorescent Voltage-Indicating Protein. *Science* **2011**, *333* (6040), 345-348.
28. Lee, K. S. S. Michigan State University, East Lansing, 2011.
29. Sameiro, M.; Goncalves, T., Fluorescent Labeling of Biomolecules with Organic Probes. *Chem. Rev.* **2009**, *109* (1), 190-212.
30. Wang, W. Michigan State University, East Lansing, 2012.
31. <http://www.chem.ufl.edu/~kschanze/manuals/quantumyieldguide.pdf>
32. Rurack, K.; Spieles, M., Fluorescence quantum yields of a series of red and near-infrared dyes emitting at 600-1000 nm. *Anal. Chem.* **2011** *83* (4) 1232-1242.

**CHAPTER IV: (I) TOTAL SYNTHESIS OF GAMMA-
CARBOXYGLUTAMIC ACID AND APPLICATIONS TOWARDS
PEPTIDE SYNTHESIS. (II) CRYSTALLIZATION OF NR2a SUBUNIT
WITH CONANTOKIN T (Con-T)**

IV.1 AN INTRODUCTION

Carboxyglutamic acid (or the conjugate base, carboxyglutamate), is an uncommon amino acid introduced into proteins by a post translational modification of glutamic acid residues. This modification is found, for example, in clotting factors and other proteins of the coagulation cascade. This modification introduces an additional carboxyl group on the gamma carbon of glutamic acid that results in an increase in the affinity for calcium ions. In the blood coagulation cascade, vitamin K is required to introduce gamma-carboxylation of clotting factors II, VII, IX, X and protein Z. The mechanism by which this process occurs is still being debated, however the general scheme of the role of vitamin K is illustrated in **Figure 60** below.¹

The coagulation system overlaps with the immune system. The formation of blood clots can result in physically trapping invading microbes. Some of the products of the coagulation cascade were found to contribute to the inherent immune system by their ability to increase vascular permeability. In addition, some products of the coagulation system also exhibit anti microbial activity. An example, beta lysine, which is a protein made by the platelets can act as a cationic detergent which results in the lysis of many gram-positive bacteria.

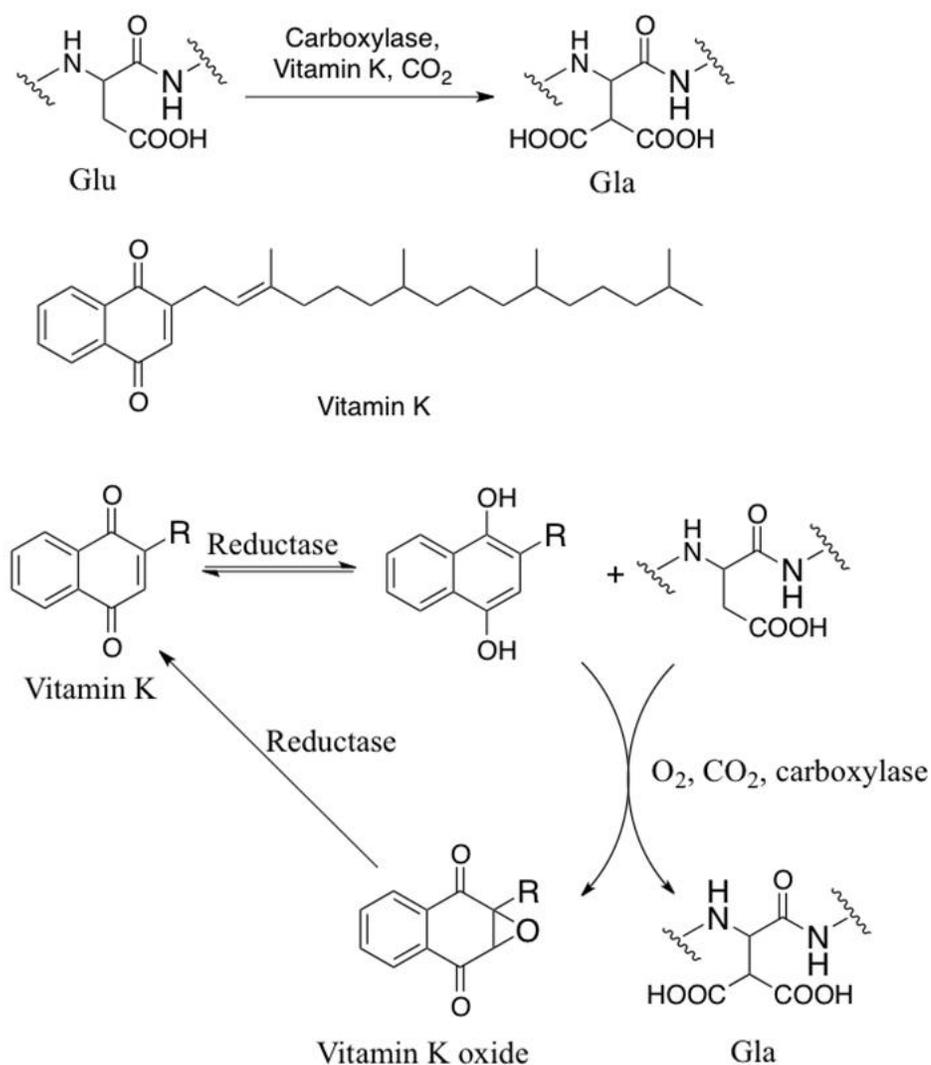


Figure 60: Proposed mechanism for the carboxylation of glutamic acid

Several different syntheses of gamma-carboxyglutamic acid have been reported. Two of the reported syntheses were racemic while the other three were chiral (these syntheses were of the L or S enantiomer). For the purpose of this research only the chiral syntheses were considered. The chiral syntheses of gamma-carboxyglutamic acid all involved the use of a chiral template. The details of each synthesis will be discussed later.

IV.2 PREVIOUSLY REPORTED SYNTHESSES OF L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla)

The earliest reported chiral synthesis involved the use of DL-serine followed by chiral separation of the enantiomers. Even though in the end both enantiomers were obtained, this would not be described as a true chiral synthesis **Figure 61**.² The DL-serine benzyl ester (**I**) was used to give benzoic 2-((benzyloxy)carbonyl)amino)-3-chloropropanoic anhydride (**II**), which was then treated with the enolate of malonic acid. The resultant product (**III**) was then separated using chiral HPLC yielding ~50% yield of each enantiomer with a > 95% ee.

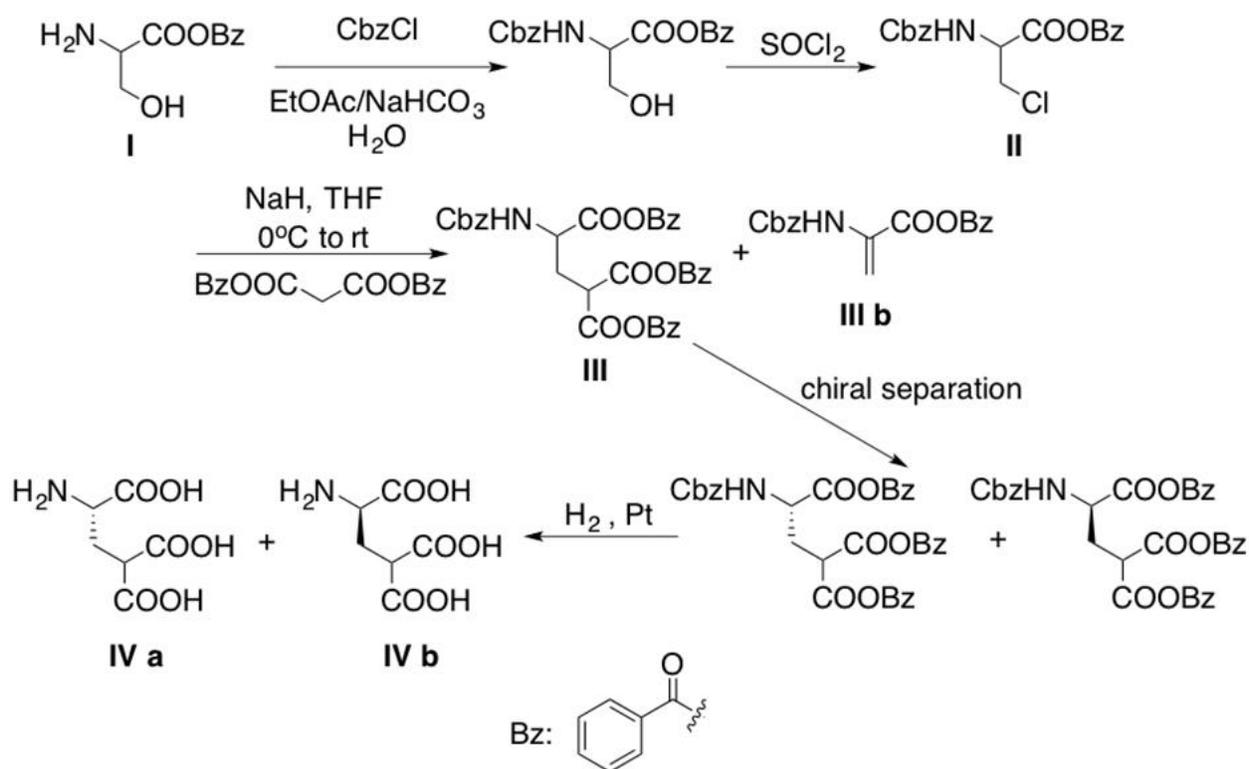


Figure 61: Earliest reported synthesis of DL gamma-carboxyglutamic acid (Gla).

Subsequent syntheses utilize different chiral templates that would lead to a single enantiomer (**Figure 62** and **Figure 63**). Scheme 2 makes use of an oxazinone template to access the chiral product while scheme 3 utilizes the Garner's aldehyde to access the chiral product.^{3,4}

a.

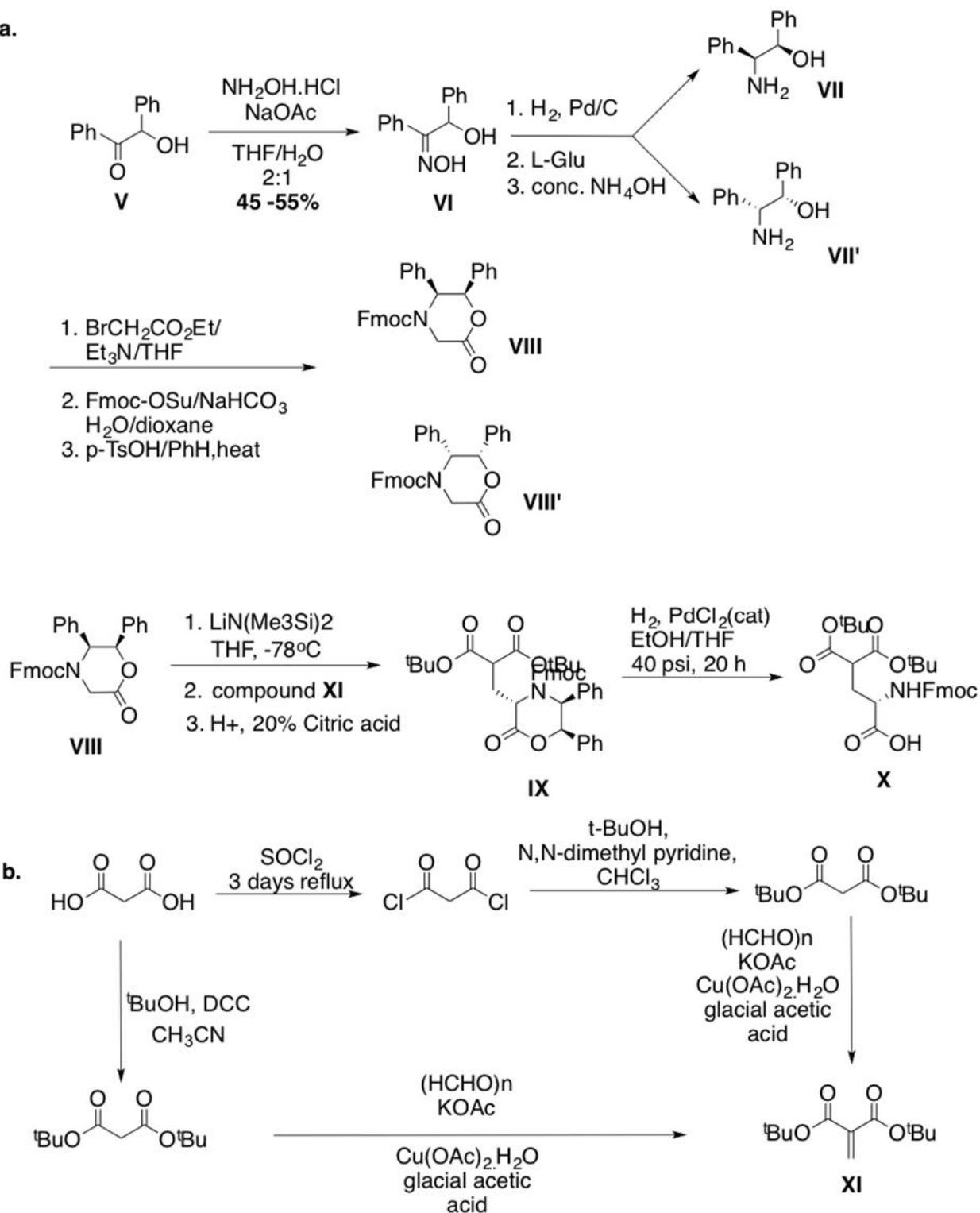


Figure 62: a. Synthesis of L-gamma-carboxyglutamic acid using the oxazinone template. b. Synthesis of compound **XI**.

The synthesis of the oxazinone template though racemic initially gave access to a single enantiomer using chiral resolution. The conditions employed in the chiral resolution allowed for the crystallization of the desired enantiomer of the oxazinone. Therefore the subsequent steps of the synthesis can be considered as chiral.

The synthesis involving the Garner's aldehyde (**Figure 63**) is the most efficient reported to date. This route gave access to the 9-fluorenylmethoxycarbonyl (Fmoc) protected L-Gla with an overall yield of 60% and > 99% ee.⁴

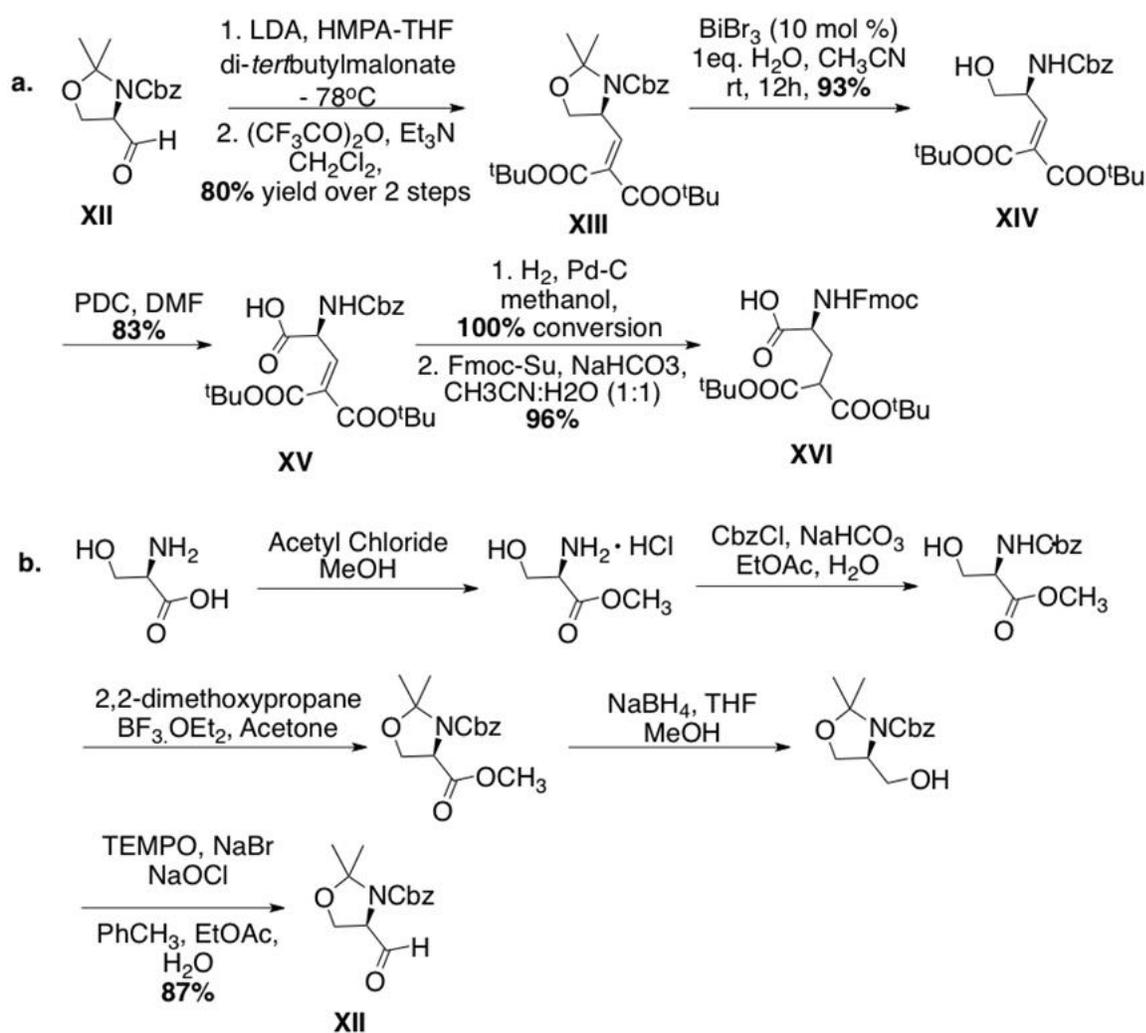


Figure 63: **a.** Synthesis of L-gamma-carboxyglutamic acid using the Garner's aldehyde template. **b.** Synthesis of Garner's aldehyde, compound **XII**.

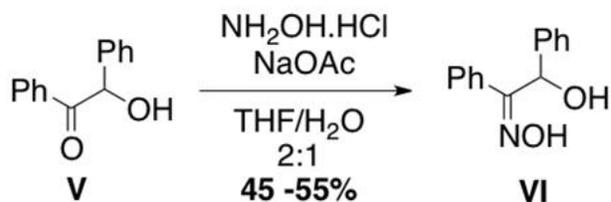
The schemes illustrated above were pursued in order to synthesize L-Gla efficiently and with high ee. The L-Gla synthesized is protected and ready for peptide synthesis (will be discussed further later). The details of the experiments are discussed below.

IV.3 SYNTHESIS OF L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla) VIA THE CHIRAL OXAZINONE TEMPLATE

The L- gamma-carboxyglutamic acid (Gla) was to be synthesized via Scheme 2. The rationale behind this decision was that this route would give access to both D- and L- enantiomers of this amino acid analogue. Of the various syntheses reported, the synthesis outlined in the scheme above is the shortest. The transformations employ the use of readily available and reasonably cheap reagents. This makes this route very efficient way, in my opinion, to access both enantiomers of Gla.

The racemic synthesis of the oxazinone template was undertaken skipping the resolution step just to become familiar with the reactions and to determine how well the transformations work.

Benzoin oxime was prepared from benzoin and hydroxylamine hydrogen chloride.⁵



When 2.5 g or less of the benzoin was used the recrystallization worked but the yield was less than 5%. When the mass was increased to 20 g the recrystallization failed as the crude product

did not dissolve in ether. As a result, the product was purified via triturating with chloroform. The H-NMR obtained showed minimum impurities and was taken into the next transformation.

The oxime **VI** prepared above was then reduced using Pd/C and hydrogen to give the amino alcohol. The resolution of the racemic amino alcohol was done using L- glutamic acid.⁶

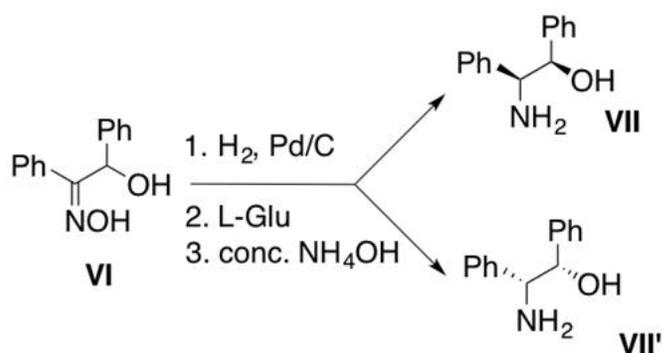


Table 23: α_D Readings for **VII'** at different concentrations. Reported $\alpha_D = -50.3^\circ$

| Concentration (c) (c=1, 10 mg/ml) | α_D |
|-----------------------------------|------------------|
| 3.0 | -39.2° to -39.4° |
| 1.5 | -40.5° to -40.7° |
| 0.65 | -44.9° to -45.1° |
| 0.25 | -40° to -55° |

The yield of the reduction however was very poor (30 – 40%). Efforts to improve the yields of this reaction were undertaken. These strategies included an increase in catalyst loading, and increasing the pressure to 60 psi. This resulted in an increase dramatic increase in the yield from ~40% to 83%. The yield of the crystallization for the resolution was also very poor as only milligram quantities (80-100 mg) of crystallized product was obtained when 15 – 20 g of the racemic amino alcohol was used. Smaller amounts (1.0 g – 5.0 g) were used in subsequent

attempts which resulted in the isolation of **2b** as the L-glutamate salt with ~80% ee (**Table 23 entry 3**).

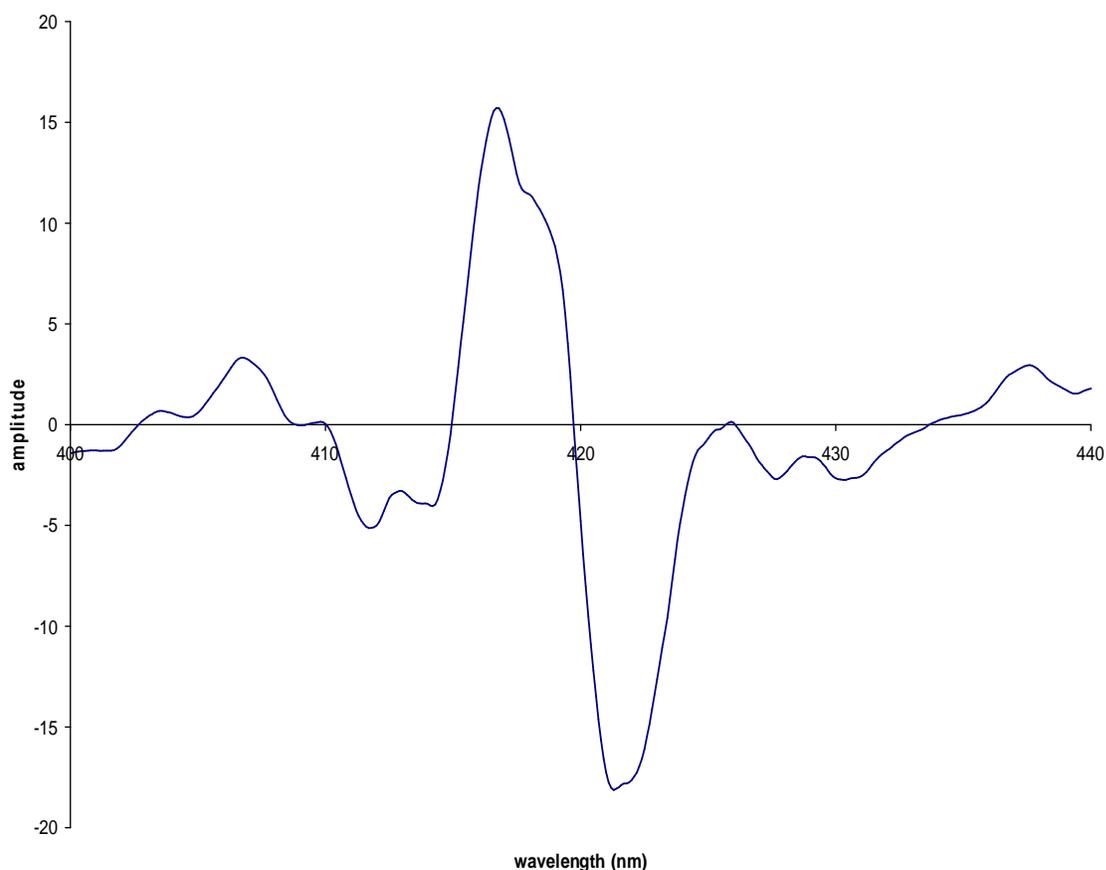
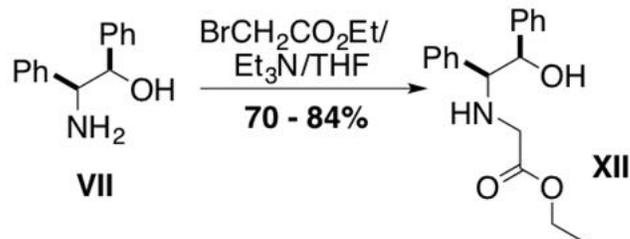
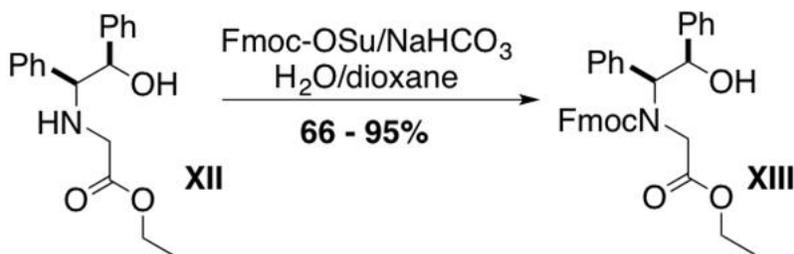


Figure 64: ECCD of VII'.

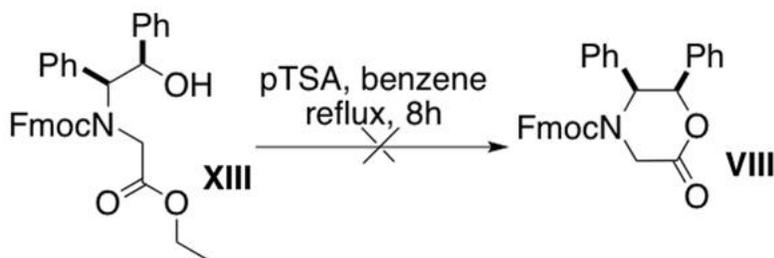
ECCD (exciton coupling circular dichroism) was used to determine the chirality of the amino alcohol. The result shown above (**Figure 64**) shows a negative ECCD, which corresponds to the signal expected for **2b**. This expectation was preceded by the results obtained for a similar substrate with the same stereochemistry; the difference is that a methyl group replaces one of the phenyl groups. Efforts to crystallize the other isomer were futile. The subsequent steps of the synthesis were carried out using the racemic amino alcohol.



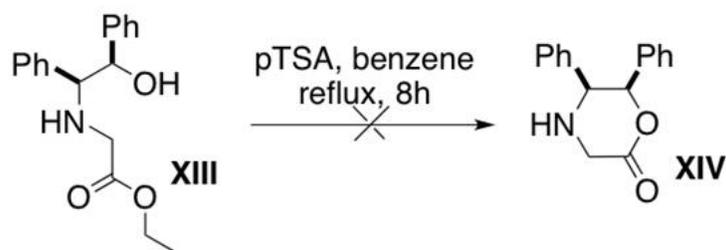
The glycinate **XII** was prepared by treatment of **VII** with bromoethylacetate.⁷ The reaction proceeded well affording the desired product in decent yields. The initial attempts of the recrystallization steps afforded the product in lower yield (~50 %). Concentrating the mother liquor after the first crop of crystals was obtained and cooling in ice rectified this. This afforded more crystals and so improving the yields.



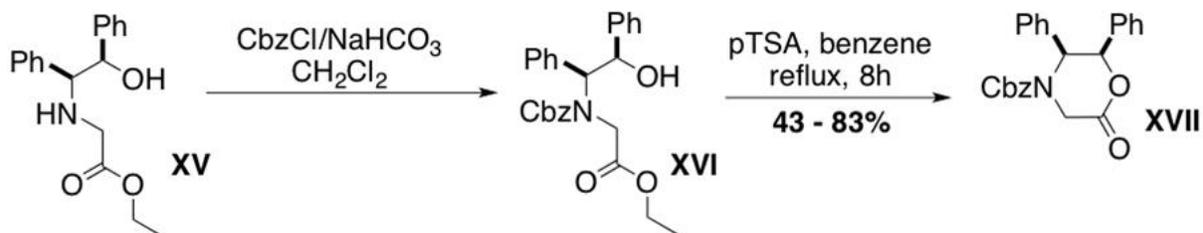
The N-Fmoc-glycinate was prepared by treatment of glycinate **XII** with Fmoc-OSu.⁸ The reaction proceeded well providing the desired product in moderate to excellent yields. No purification was necessary as product showed presence of very little impurities by H-NMR, and was taken to the subsequent step.



Attempts to synthesize **VIII** from **XIII** were unsuccessful. After refluxing for 16 -24 h only starting material was detected by TLC. It was found that if the reaction mixture was heated for longer than 24h degradation of the starting material was observed, as the peaks observed for the Fmoc protecting group seem to have disappeared from H-NMR.

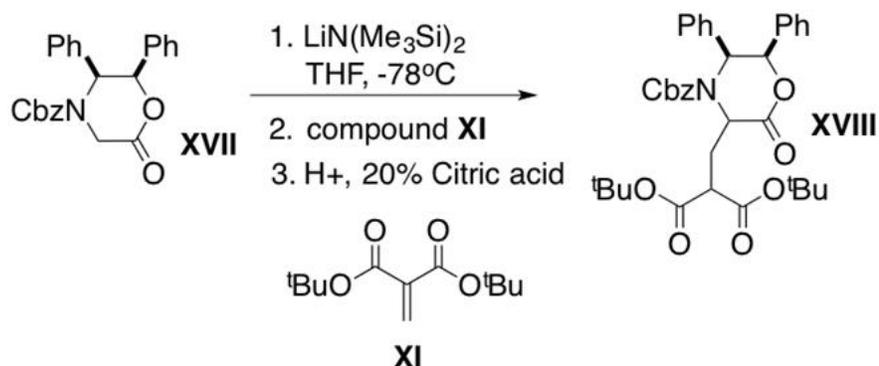


The unsuccessful attempts to obtain the cyclized product N-Fmoc- oxazinone led to the decision to attempt the cyclization of the unsubstituted N-oxazinone **XIII**. However, these attempts were unsuccessful as refluxing for 6 – 8 hours resulted in decomposition of the starting material. This was confirmed by H-NMR.



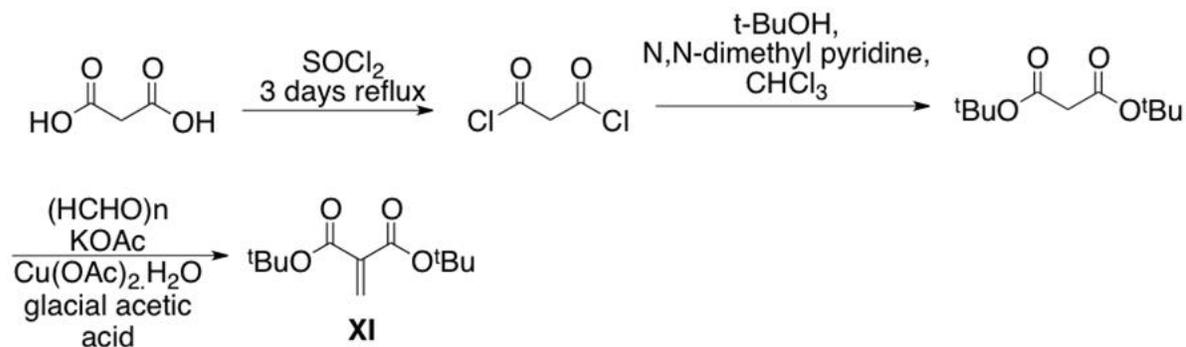
In light of the fact that the previous attempts to obtain the N-Fmoc oxazinone **VIII** and the unsubstituted N-oxazinone **XIV** were unsuccessful; the decision was then made to carry out this sequence of reactions as reported using benzyloxycarbonyl (Cbz) as the protecting group on nitrogen.⁹ This would lead to the formation of N-Cbz-oxazinone **XVII**. The N-Cbz glycinate **XVI** was obtained with quantitative yields, TLC shows complete conversion. The H-NMR

confirmed the presence of the product. No further purification was done and the product was then taken to the cyclization step to give the N-Cbz-oxazinone.

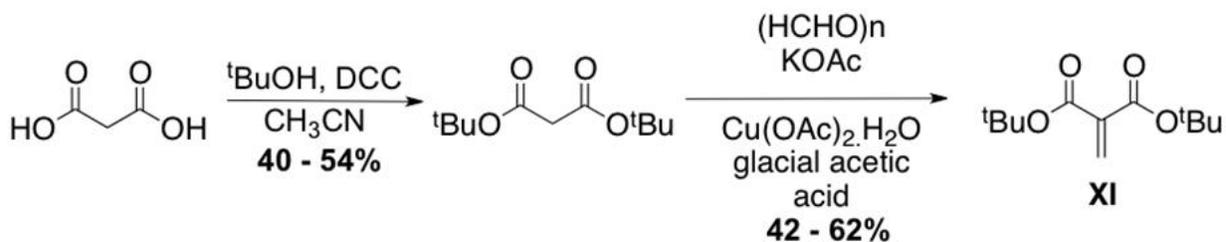


The N-Cbz-oxazinone **XVII** was then treated with di-tert-butylmethylene malonate in an attempt to obtain Gla-oxazinone **XVIII**. Attempts to synthesize **XVIII** using LiHMDS were unsuccessful, the H-NMR and mass of the product did not correspond to the reported data for the desired product.²³ NaHMDS was tried instead and the results were the same, the desired product was not obtained. This was further confirmed by elemental analysis that showed that the formula of the product obtained is $\text{C}_{27}\text{H}_{25}\text{NO}_6$ while the formula for the desired product is $\text{C}_{36}\text{H}_{41}\text{NO}_8$. It was then decided to try this reaction using lithium diisopropylamine (LDA) another bulky base. TLC showed complete conversion and only one spot with $R_f = 0.80$. However, when the crude was placed on the column two colored bands were observed; a dark brown band ($R_f = 0.80$) and a yellow band ($R_f = 0.41$). In both cases the same solvent systems were employed. The products from this reaction were not solid as expected for the desired product and both products were soluble in chloroform, which is not characteristic of the desired product. The H-NMR of both products did not correspond to the reported data for the Gla oxazinone **XVIII**.

The original approach was to synthesize **XI** via the route shown below⁷:



After heating the mixture for 2 days a black tar-like solid could be seen accumulating in the reaction flask. After day 3, the vacuum distillation to isolate the malonyl dichloride was unsuccessful. The H-NMR of the black tar-like solid showed a large multiplet between 1.0 – 3.0 ppm. The reaction was repeated several times with a similar outcome. Therefore it was decided to use an alternative route using DCC coupling to synthesize the ester directly from the acid.⁸



The yields of the DCC coupling reaction ranged from 40 – 54%, which is much lower than the reported yield of 92%. There was complete conversion based on TLC and so the low yields could be attributed to the purification step that involved the vacuum distillation of the product. The yields for the Knoevenagel condensation ranged from 42 – 62%, which was better than the reported yield of 42%.

The inability to obtain the coupled substrate, compound **XVIII**, led to the decision to attempt the synthesis using Garner's aldehyde as the chiral template. Details of the synthesis are discussed below.

IV.4 SYNTHESIS OF L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla) VIA THE GARNER'S ALDEHYDE TEMPLATE

The Garner's aldehyde was synthesized from L-serine which would lead to D-gamma-carboxyglutamic acid.^{10 11}

The synthesis of **XII** from L-serine proceeded very well with excellent yields and high optical purity. The α_D for N-Cbz-L-serine methyl ester ranged from -14.3 to -14.0 which is in very good agreement to the literature value of -14.5; and for (R)-3-benzyl 4-methyl 2,2-dimethyloxazolidine-3,4-dicarboxylate, α_D ranged from -47.5 to -48.6 which also corresponds well to the reported value of -48.3. The subsequent reduction of (R)-3-benzyl 4-methyl 2,2-dimethyloxazolidine-3,4-dicarboxylate to the alcohol was successful. However, the yields were not as great as the reported yield of 88%. This could be attributed to potential loss of the product during purification via column chromatography. Attempts to improve the yield to match what is reported were unsuccessful.

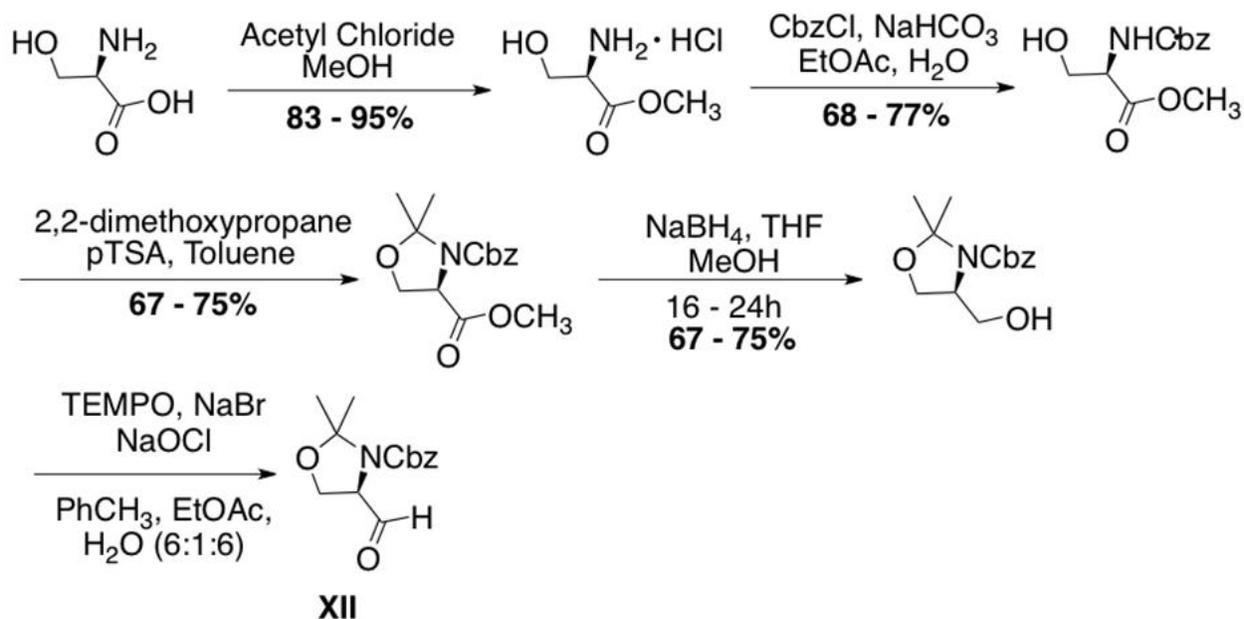
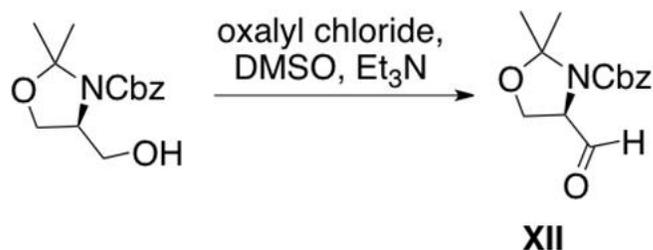


Figure 65: Synthesis of Garner's Aldehyde **XII**.

There were some issues with the conversion of the alcohol to aldehyde **XII** using TEMPO as the oxidant.¹¹ For this reaction analysis of the crude product obtained showed the presence of the desired aldehyde product. However, when the crude was taken through purification via column chromatography using silica gel, the product was lost. This phenomenon suggests that degradation of the product might be taking place due to the acidity of the silica gel and so for subsequent trials a higher percentage (3% instead of 1%) of triethylamine was used to prepare the basified silica gel was used. This did not modify the outcome and so it was then decided to try the Swern oxidation to get to the target, aldehyde **XII**. This reaction was very successful affording quantitative yields of the product, Garner's aldehyde. The purification was done using a basified silica gel plug instead of a column.¹²



Aldehyde **XII** was then taken through several transformations illustrated in **Figure 66** to obtain the desired product.¹³

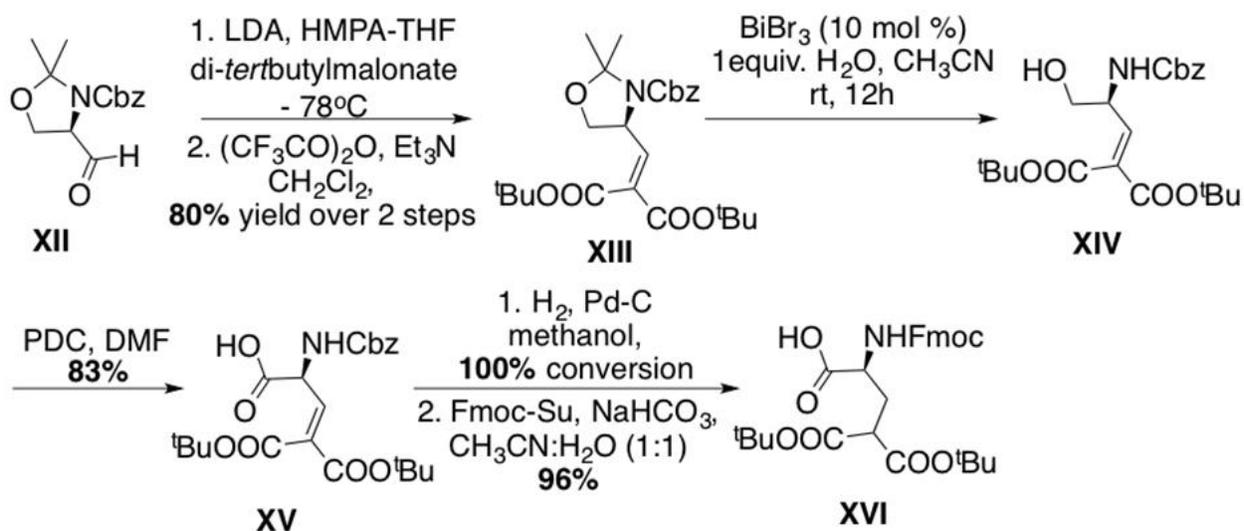
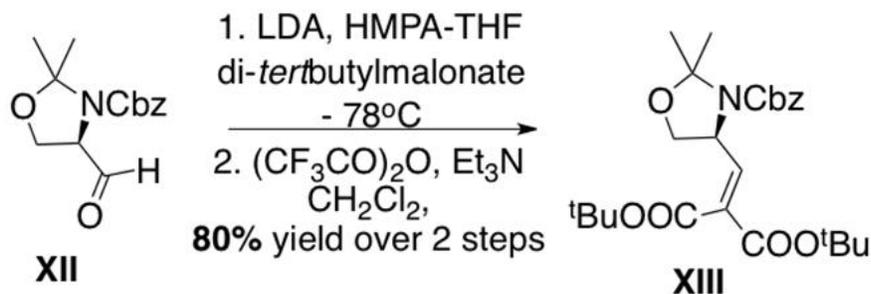
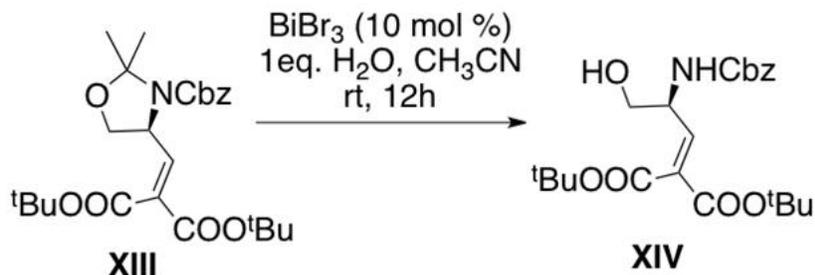


Figure 66: Synthesis of L- γ -carboxyglutamic acid (Gla) from **XII**.



The Knoevenagel condensation of **XII** with di-*tert*-butylmethylene malonate proceeded very well. The yield was as reported in the literature.¹³ Attempts were then made to remove the acetonide protection in compound **XIII**.



These attempts were unsuccessful as in some cases the starting material was recovered and in other cases the mass of the product was equal to the mass of the starting material but the

H-NMR did not match which suggested that some rearrangement might have occurred. Several other Lewis acids were used in an attempt to facilitate the acetonide deprotection. The conditions tried resulted in either degradation or recovery of the starting material. **Figure 67** shows the conditions employed in attempts to deprotect the acetonide.^{14 15 16 17}

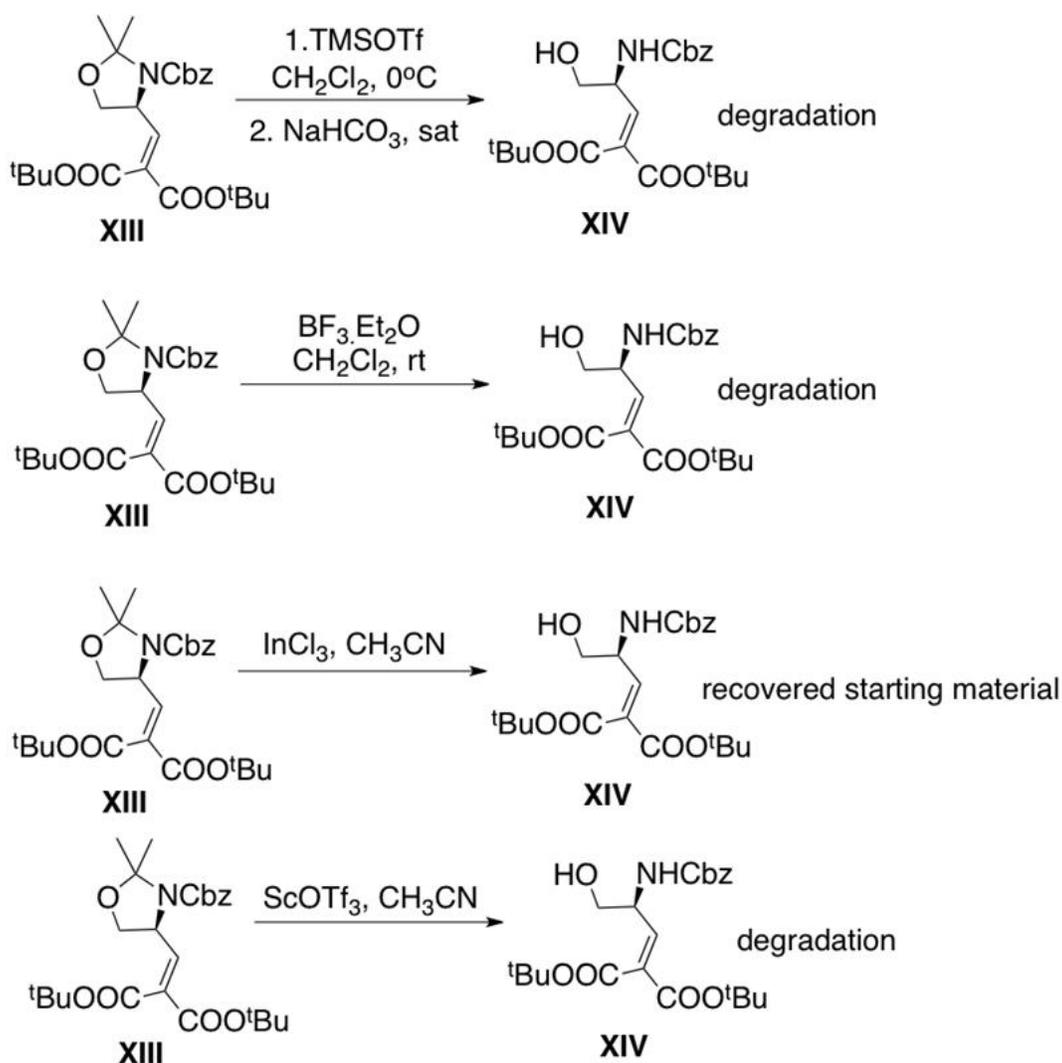


Figure 67: Conditions used in an attempt to deprotect the acetonide.

The disappointing outcome led to the decision to revert to the original conditions but with one minor change. The BiBr₃ was dried in a desiccator under an inert atmosphere for 24h. The reaction gave the desired product with a 15% yield. With this less than satisfactory yield other

alternatives were explored. Browsing the literature revealed a report that explored the use copper (II) salts as mild Lewis acids for acetonide deprotection in the presence sensitive functional groups. This led to the exploration of the different copper (II) salts to determine the appropriate salt that would work for this transformation. To our delight copper (II) chloride was the salt that worked best to give the desired product. The reaction was repeated several times giving a range of 60 – 72% yield.

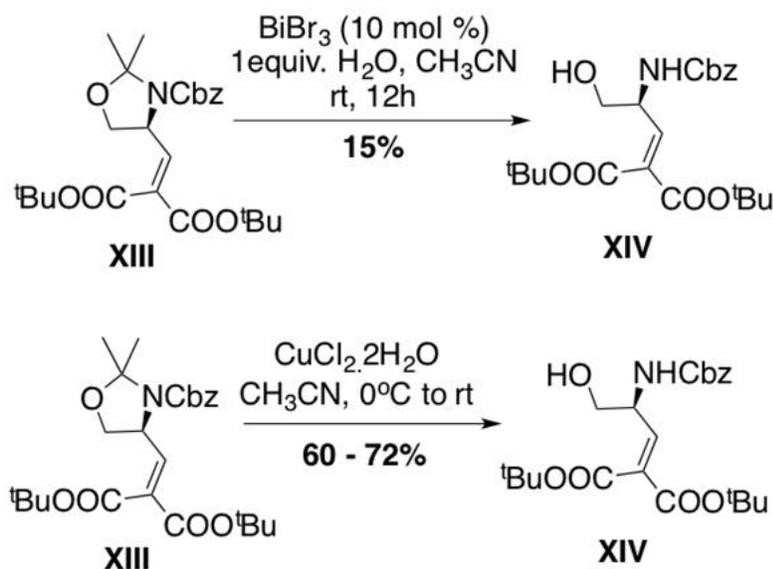


Figure 68: Conditions used that successfully deprotected the acetonide.

Compound XIV was taken through the next series of transformations without further purification until the final step (**Figure 69**).¹³ The oxidation with 4.0 equivalents of pyridinium dichlorochromate (PDC) proceeded with full conversion (the reaction was monitored by the disappearance of the starting material). The product was extracted using diethyl ether (ether). The subsequent steps proceeded with full conversion. However there were some issues with the purification of the product. The final product (white solid) was discolored (green oil) due to the presence of residual chromium (III). The ¹H-NMR revealed that the product was masked with a lot of impurities as only a few of the peaks corresponding to the product were seen. The analysis

of the crude product using electrospray mass spectrometry revealed the fragmentation pattern innate to that of the product. Several markers such as the loss of the *tert*-butyl groups as well as the sequential loss of a mass of 44 assisted in the confirmation of the presence of the product. The loss of a mass of 44 corresponds to the loss of the geminal CO₂ groups.

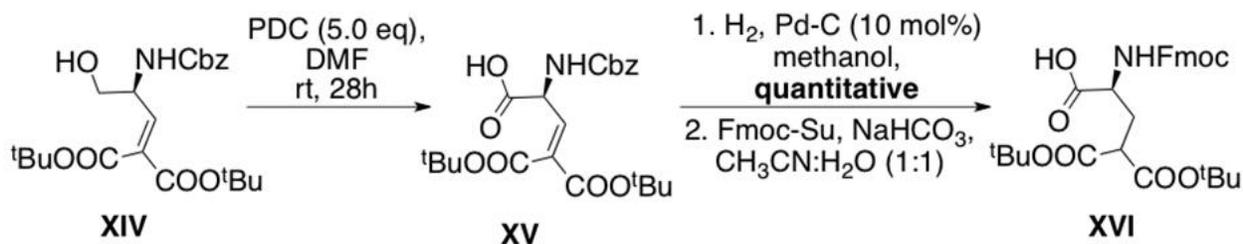


Figure 69: Steps leading to the final product compound **XVI**.

Attempts to isolate the product were futile. Several runs through the silica gel column did not completely remove the impurities and resulted in a substantial loss of the product with each run. This phenomenon would suggest that in order to get a substantial amount of the product, large amounts of the starting materials would be required. This is not very economical and thus led to the decision to attempt to circumvent the oxidation step.

IV.5 SYNTHESIS OF L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla) USING L-SERINE AS A CHIRAL TEMPLATE

The scheme was then reviewed and the decision was made to use the starting L-serine as the chiral template (**Figure 63**). This resulted in a completely different approach that utilizes steps from different schemes and the incorporation of some new transformations (**Figure 70**). The idea was to retain the carboxyl group of the serine as the α -carboxyl group and so avoid the oxidation step.

This scheme incorporated reactions described in the earliest reported synthesis of L-Gla using a racemic serine as the starting material (**Figure 64**). The first two steps of the transformation proceeded with yields comparable to the reported yields. With the N-Cbz serine methyl ester in hand the following steps were a little different than what was reported.

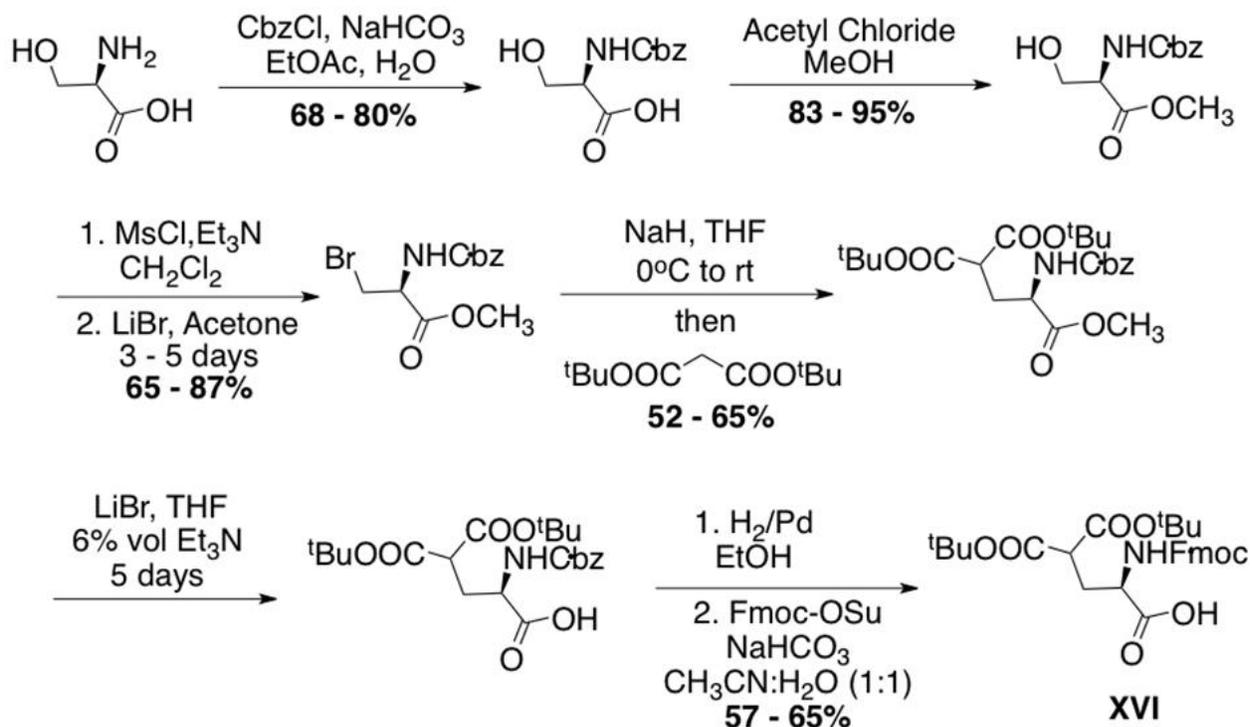


Figure 70: Synthesis of L-Gla from L-serine as chiral template.

The N-Cbz serine methyl ester was then converted to the chloride, bromide, mesylate and tosylate, each of which was then treated with the enolate of di-ethyl malonate to see which condition was ideal for this transformation. In all cases, except for the bromide, the starting material was recovered. The bromide was then treated with the enolate of di-*tert*-butyl malonate and the desired product was obtained with moderate yields. It was also found that if the mixture of the di-*tert*-butyl malonate and NaH was allowed to stir for 2 – 3h under an inert atmosphere

the yield increases. The subsequent steps were carried out without any purification until the last step.

The next step of hydrolyzing the methyl ester proved to be less trivial than anticipated. The aim was to find a protocol that selectively hydrolyzes the methyl ester in the presence of the *tert*-butyl esters.¹⁸ Browsing the literature revealed a report that explored the use lithium salts to selectively hydrolyze methyl esters in the presence of other esters. **Table 24** shows the different conditions used.

The mechanism of this reaction explains the selectivity of this reaction for methyl esters (**Figure 71**). The initial step in the reaction involves the coordination of the lithium to the alkoxy oxygen of the ester. The bulkier alkoxy groups would hinder the coordination and would explain why the esters with isopropyl and *tert*-butyl groups are stable under these conditions. The drawback of this reaction however, is that for straight-chained alkyl groups such as ethyl and propyl groups the selectivity diminishes.

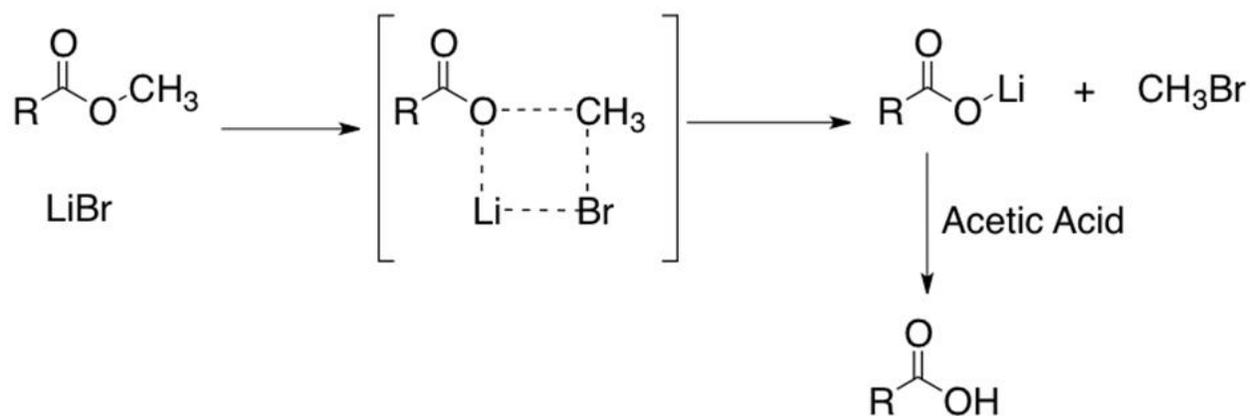
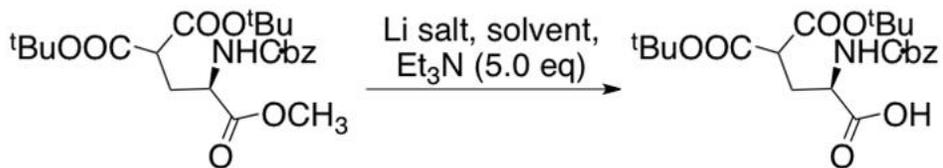


Figure 71: Mechanism of hydrolysis of the methyl ester.

Table 24: Showing the conditions used using different Lithium salts.

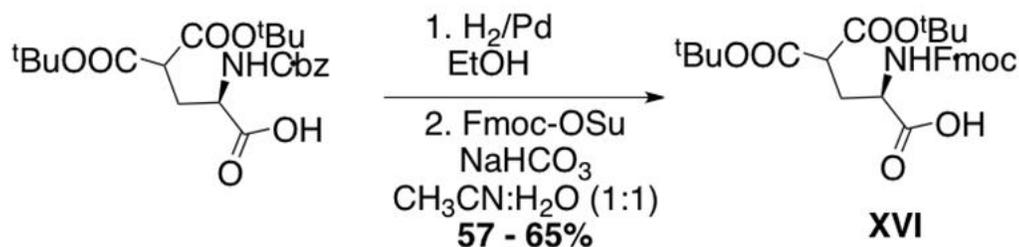


| Lithium salt | Equivalents | Solvent | Vol % of H ₂ O |
|-------------------|-------------|--------------------|------------------------------|
| LiCl | 2.0 | THF | 3.0 |
| LiCl | 5.0 | THF | 6.0 |
| LiCl | 2.0 | CH ₃ CN | 3.0 |
| LiCl | 5.0 | CH ₃ CN | 6.0 |
| LiBr | 2.0 | THF | 3.0 |
| LiBr | 5.0 | THF | 6.0 |
| LiBr | 2.0 | CH ₃ CN | 3.0 |
| LiBr | 5.0 | CH ₃ CN | 6.0 |
| LiBF ₃ | 2.0 | CH ₃ CN | 3.0 |
| LiBF ₃ | 5.0 | CH ₃ CN | 6.0 |

*The highlighted entry was the condition that gave the best result.

This reaction was monitored by the disappearance of the starting material. The clear oil obtained was then taken through the final two steps of the scheme to give the desired product, L-Gla.

These reactions were carried out following the reported protocol.



The purification of the final product, L-Gla proved to be more tedious than expected. The reported protocol reports purification using column chromatography using 20% ethyl acetate/hexanes, however this did not allow for an efficient removal of the impurities. There were four distinct products each of which exhibit a very different R_f value in solvents with different percentages of ethyl acetate. Therefore to address the issues with purification a gradient column was run using solvents ranging from 0% - 100% ethyl acetate. Some of the product was lost as reflected in the moderate yields.

With the desired product in hand, attempts were then made to synthesize peptides, conopeptides, which have this unnatural amino acid in their sequence. These Gla residues are important in the structure and function of these peptides. This phenomenon will be discussed in more detail in the next section.

IV.6 SYNTHESIS OF CONOPEPTIDES USING L-GAMMA-CARBOXYGLUTAMIC ACID (L-Gla)

Knowledge of the interactions that results in the formation of stabilized secondary protein structures is essential for understanding how proteins fold. A lot of time and effort has been expended in order to elucidate the chemical and physical principles that govern the formation of α -helices, the most common secondary structure of proteins. It has been discovered that the helical structure of most proteins is dependent on hydrophobic interactions, which are very difficult to

control when considering hydrophobic repacking, disulfide cross-linking and metal binding. In order to explore such interactions it would be logical to work with short peptides (~20 residues) which can be easily manipulated and can then provide a wealth of information about how sequence affects the structure of peptides / proteins.

The exploration of these systems is however hampered by the fact that peptides (~ 20 residues) rarely adopt any type of secondary conformation in an aqueous environment. This spurred scientists' interest to design systems in which these peptides are able to form helices via metal bridging using transition metal complexes.^{19 20 21} These helices are able to organize into higher order systems such as coiled coils and helix bundles. The success of these systems led to the design of short peptides bearing unnatural amino acids.

In the design of these higher order systems it was found that the metal ion is essential for initiating helix formation and helix – helix association. However, the stability of these complexes is dependent on the interstrand hydrophobic interactions. Therefore this process cannot be considered as being totally dependent on metal ion chelation.

Thankfully there are natural occurring systems in which helix formation is totally and in some cases mostly dependent on the chelation of metal ions.²² One example of such systems is conotoxins, which is a member of the family of conopeptides. The resultant association of the helices formed after metal chelation was originally discovered in conopeptides and is referred to as the 'metallo-zipper' motif.

The cone snails, *Conus* genus, are venomous marine mollusks that use small structured toxic peptides (conotoxins) as a means of prey capture and defense against predators. There are over 500 species within the *Conus* genus and each member is capable of expressing

approximately 100 different conotoxins. This variety is essential as several of these peptides work together in order to effectively impose the desired physiological effect on the prey.

Conantokins have been explored in order to understand the relationship between these peptides and hypervalent metal ions, which allows for such a dramatic change in structure.

The peptide amino acid sequences of the four known members of the conantokin family are shown in **Figure 72**.

Con G: GE γ γ LQ γ NQ γ LIR γ KSN
Con T: GE γ γ YQKML γ NLR γ AEVKKNA
Con T (K γ 7): GE γ γ YQ γ ML γ NLR γ AEVKKNA
Con R: GE γ γ VAKMAA γ LAR γ NIAKGCKVNCYP

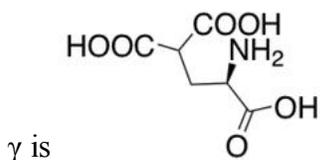


Figure 72: Amino acid sequence of a few of the known Conantokins.

The presence of the gamma-carboxyglutamic acid (Gla) residues confers the ability of these peptides to bind to hypervalent metal ions. This comes as no surprise since there are two carboxyl groups on the side chain of this molecule.

Two of the members of conantokins, Con-G and Con-T, have been examined in detail by Castellino and co-workers.^{23 24} Con-G and Con-T contain 17 and 21 amino acids respectively. Con-G was isolated from the species *Conus geographus* and Con-T from *Conus tulipa*. A common feature of conotoxins is the high abundance of gamma-carboxyglutamic acid residues. There are namely four residues per molecule in Con-T and five residues per molecule in Con-G.

It was found that in the presence of metal ions such as Ca^{2+} , Mg^{2+} , Zn^{2+} , and Cu^{2+} , the structure of both Con-G and Con-T is changed from a linear structure to a helical structure which further associates to give an anti-parallel dimer.^{23 25 26 27}

The growing interest in these peptides is due to: 1) they are small peptides (12-30 residues), which means that they would be amenable to peptide synthesis conditions (Solid Phase Peptide Synthesis) and mutagenic manipulations; 2) their ability to adapt a helical conformation in the presence of hypervalent metal ions; 3) their impeccable specificity for the NMDA (N-methyl-D-Aspartate) receptor, which make these peptides plausible candidates for use in drug therapy.

The N-methyl-D-aspartate receptor (NMDAR) is an ionotropic receptor for glutamate, which is its selective specific agonist (**Figures 73**). Activation of NMDA receptors results in the opening of an ion channel that is nonselective to cations. This allows flow of Na^+ and small amounts of Ca^{2+} ions into the cell and K^+ out of the cell. Calcium flux through NMDARs is thought to play a critical role in synaptic plasticity, a cellular mechanism for learning and memory. The NMDA receptor is distinct in that it is both ligand-gated and voltage-dependent. This makes the conotoxin peptides a possible cure for mental diseases such as Alzheimer's and Parkinson's disease.^{28 29 30}

Con-G and Con-T are members of the ω -conantokins which are specific for Ca^{2+} channels. A lot of work has been done by Castellino and coworkers to elucidate the residues that confer this antagonistic activity to these peptides.^{31 32 33 34} It is interesting to note that Con-G and Con-T are not the only members of Conus genus that exhibit antagonistic activity on the NMDA receptor, Con R from the species *Conus radiatus*²⁸ and Con-Pr1, Pr2 and Pr3 from the species *Conus parius* are other examples.³⁵

As mentioned before, another interesting feature of the conantokins is their ability to associate to form dimers and in some cases trimers when treated with hypervalent metal ions. When Con-G or Con-T (K7 γ), variant of Con-T (which does not form a dimer upon metal chelation) is bound to Ca²⁺ the association to form anti-parallel dimers was observed. The phenomenon was verified by the crystal structure obtained by Dr. Sara Cnudde from our lab in collaboration with Castellino.²⁷

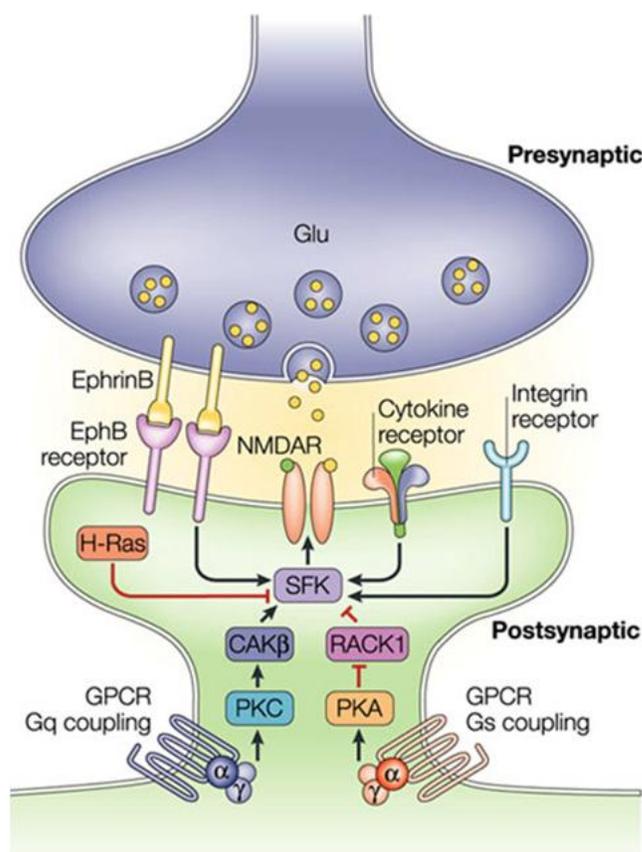


Figure 73: Diagram of the Synapse with the N-methyl-D-Aspartate Receptor (NMDAR).^{36 37}

These crystal structures allowed for keen examination of these systems and provided a wealth of information about the interface of interaction between the two peptides, and highlighted the similarities and surprising differences between these structures. If successful with

obtaining these peptides, routine crystallization would be undertaken. Systems in which the peptides are bound to different metals will also be explored, as it is crucial for understanding the effects of the metal ion on the structure and the interactions between peptides that facilitate helix association upon metal chelation.

Upon fully understanding the ‘language of the peptides’, manipulations that allow for the application of the ‘metallo-zipper’ motif for building novel structures and explore interactions of these peptides with macromolecular systems can be considered. These systems have the potential to be utilized in chemical reactions. The specificity exhibited by these peptides one can envision the utility of these peptides as a scaffold to engineer systems that will be able to bind various biological targets.

In order to engineer the different peptides in an attempt to build these novel macromolecular systems, the key residue for metal ion binding, Gla, had to be synthesized. With Gla in hand we would be able to access many variants of the conotoxins.

The synthesis of these peptides proved to be a little challenging. Reports from the Castellino and the Geiger laboratories shows that several residues are conserved across the family of conopeptides and the residues L-Gla 4, 7,10 and Arg 13 are essential for the formation of the helix upon metal binding. Therefore, the decision was made to synthesize a few peptides to test the role of other residues of the peptides (see **Figure 74**).

Considering the essential residues of Con-G, the decision was made to synthesize a truncated version of the peptide. The syntheses were carried out using the Solid Phase Peptide Synthesis (SPPS) protocol.

Solid-phase peptide synthesis (SPPS), pioneered by Robert Bruce Merrifield resulted in a paradigm shift within the peptide synthesis community.

It is now the accepted method for creating peptides and proteins in the lab in a synthetic manner.

Con G: **GE $\gamma\gamma$ LQ γ NQ γ LIR γ KSN**

Con G (13-mer): **$\gamma\gamma$ LQ γ NQ γ LIR γ K**

Ala-Con G (13-mer): **$\gamma\gamma$ AA γ AA γ AAA γ K**

Con G (14-mer): **W $\gamma\gamma$ LQ γ NQ γ LIR γ K**

Ala-Con G (14-mer): **W $\gamma\gamma$ AA γ AA γ AAA γ K**

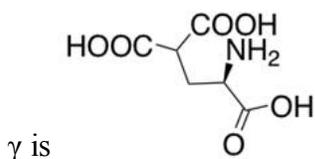


Figure 74: Sequence of four modified members of the Conantokin family that were synthesized and the original sequence of Con-G.

SPPS allows the synthesis of natural peptides that are difficult to express in bacteria, the incorporation of unnatural amino acids peptide/protein backbone modification, and the synthesis of D-proteins, which consist of D-amino acids.

SPPS involves the use of a solid supports, which are usually small insoluble and porous solid beads that are treated with functional units ('linkers') on which peptide chains can be built. The peptide will remain covalently attached to the bead until cleaved from it by a reagent such as anhydrous hydrogen fluoride or trifluoroacetic acid. The peptide is thus 'immobilized' on the solid-phase and can be retained during a filtration process, whereas liquid-phase reagents and by-products of synthesis are flushed away.

The general principle of SPPS is one of repeated cycles of coupling-wash-deprotection-wash. The free N-terminal amine of a solid-phase attached peptide is coupled to a single N-protected amino acid unit. This unit is then deprotected, revealing a new N-terminal amine to which a further amino acid may be attached. The superiority of this technique partially lies in the

ability to perform wash cycles after each reaction, removing excess reagent with the growing peptide of interest remaining covalently attached to the insoluble resin.

The overwhelmingly important consideration is to generate extremely high yield in each step. For example, if each coupling step were to have 99% yield, a 26-amino acid peptide would be synthesized in 77% final yield (assuming 100% yield in each deprotection); if each step were 95%, it would be synthesized in 25% yield. Thus each amino acid is added in major excess (2~10x) and coupling amino acids together is highly optimized by a series of well-characterized agents.

There are two majorly used forms of SPPS – 9-fluorenylmethyloxycarbonyl (Fmoc) and *tert*-butyloxycarbonyl (Boc). Unlike ribosome protein synthesis, solid-phase peptide synthesis proceeds in a C-terminal to N-terminal fashion. The N-termini of amino acid monomers is protected by either of these two groups and added onto a deprotected amino acid chain. Automated synthesizers are available for both techniques and they can also be performed manually.

The protocol employed for the syntheses of the peptides mentioned above utilizes the Fmoc protecting group. To test the efficiency of the synthesis the first five amino acid residues of the Con-G (13-mer) were coupled manually via the Fmoc protocol. Gratifyingly, the coupling proceeded with very good efficiency. The electrospray mass spectrometry analysis showed that the majority of the product was indeed the desired tetrapeptide (see experimental **Figure 79**). However, when the entire sequence of the Con-G (13-mer) synthesized via this method, electrospray mass spectrometry analysis showed a significant decrease in the yield of the desired peptide. There were quite a number of by products and the intensity of the UV signal at 210 nm (absorption of peptide linkage) was very low. This posed a problem in the purification process

via High Performance Liquid Chromatography (HPLC). This led to the decision to add a tryptophan residue at the beginning of the sequence. The introduction of the tryptophan residue 'should not' affect the structure of the helix since it is at a position that is not embedded in the structure. The disadvantage however, of carrying out the synthesis manually is the amount of reagents required for each step and the low yield in the end was not very encouraging. Therefore, the decision was made to proceed using an automated system.

The use of the automated system did not improve the results. The results showed a very low concentration of the desired peptide or no peptide. This phenomenon led to us to question the coupling efficiency of the modified amino acid, L- Gla. Therefore, the decision was made to incorporate glutamic acid, which has a good coupling efficiency, and proline, which is known to pose problems in coupling, into the sequence where L-Gla would be. This experiment used the truncated (8-mer). The sequences attempted are shown below:

| | |
|---------------------------|----------|
| Con G (8-mer): | NQELIREK |
| Ala-Con G (8-mer): | AAEAAAEK |
| Con G (8-mer): | NQPLIRPK |
| Ala-Con G (8-mer): | AAPAAAPK |

Figure 75: Sequence of the shortened modified members of the Conantokin family that were synthesized.

The result showed that both substitutions worked very well and so it turns out that the problem was either with L-Gla or the conditions that were currently being used. (See experimental **Figures 81 - 85**)

A recent report revealed that L-Gla has to be given a longer coupling time, up to one and a half times the time compared to the natural amino acids. This led to a revision of the protocol being used. The coupling times for L-Gla were changed and different coupling reagents (HATU, PyBOP and HBTU) were used in an attempt to improve the yield of the desired peptide.

The syntheses were carried out only using the 8-mer sequences in order to conserve on reagents. Disappointedly, the results were not as exciting as expected. The yields of the 8-mers were very low which would in turn explain the low yield or non-existence of the full peptide.

IV.7 EXPRESSION AND CRYSTALLIZATION OF NR2A WITH CONANTOKIN T (Con-T)

Another interesting feature of conantokins is their affinity for the N-methyl-D-Aspartate Receptor (NMDAR). The NMDAR is a glutamate receptor responsible for controlling synaptic plasticity and memory function. The NMDAR is a specific type of ionotropic glutamate receptor. NMDA (N-methyl-D-aspartate) is the name of a selective agonist that binds to NMDA receptors but not to other 'glutamate' receptors. Activation of NMDA receptors results in the opening of an ion channel that is nonselective to cations with an equilibrium potential near 0 mV. A property of the NMDA receptor is its voltage dependent activation, as a result of ion channel block by extracellular Mg^{2+} ions. This allows the flow of Na^+ and small amounts of Ca^{2+} ions into the cell and K^+ out of the cell.

The NMDA receptor forms a heterotetramer between two GluN1 and two GluN2 subunits (the subunits were previously denoted as NR1 and NR2), two obligatory NR1 subunits and two regionally localized NR2 subunits **Figure 76**. Two ligands are required for the activation of these receptors, Glutamate and glycine or D- serine. The conantokins have been shown to be potent antagonists of the NMDAR. Con-G shows potential as a neuroprotective agent in ischemic and excitotoxic brain injury, neuronal apoptosis, pain, epilepsy, and as a research tool in drug addiction and Alzheimer's disease. Con-T acts by inhibiting NMDAR-mediated Ca^{2+} influx in neurons in the central nervous system. Though extensive research has

been done in this field, the interactions of the conantokins with the NMDAR is still not fully understood due to the lack of crystal structures. Therefore our goal is to crystallize the NMDAR bound to Con-G, which binds NR1A and Con-T, which binds NR2A. These crystal structures would be pivotal in the design therapeutics for the neurological diseases that are linked to these receptors.

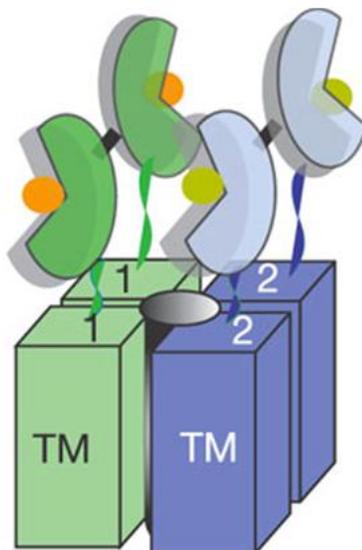


Figure 76: Diagram of the N-methyl-D-Aspartate Receptor (NMDAR). **TM** (Transmembrane portion of the receptor); **1** (subunit 1, NR1); **2** (subunit 2, NR2).^{36 37}

Eric Gouaux and coworkers reported the crystal structures of both NR1A and NR2A with the native ligands, glycine and glutamate respectively.^{36 37} A crystal structure of the NR1-NR2A heterodimer was also obtained to probe the subunit arrangement and function in the NMDA receptors. It should be noted however, that the crystal structures obtained were not of the full receptor but of the ligand-binding core, S1S2 **Figure 77**. The amine terminal domain (ATD) and the carboxyl terminal domain (CTD) were removed as well as the transmembrane portion of each receptor. The corresponding S1S2 domains were connected using a glycine-threonine (GT) linker.

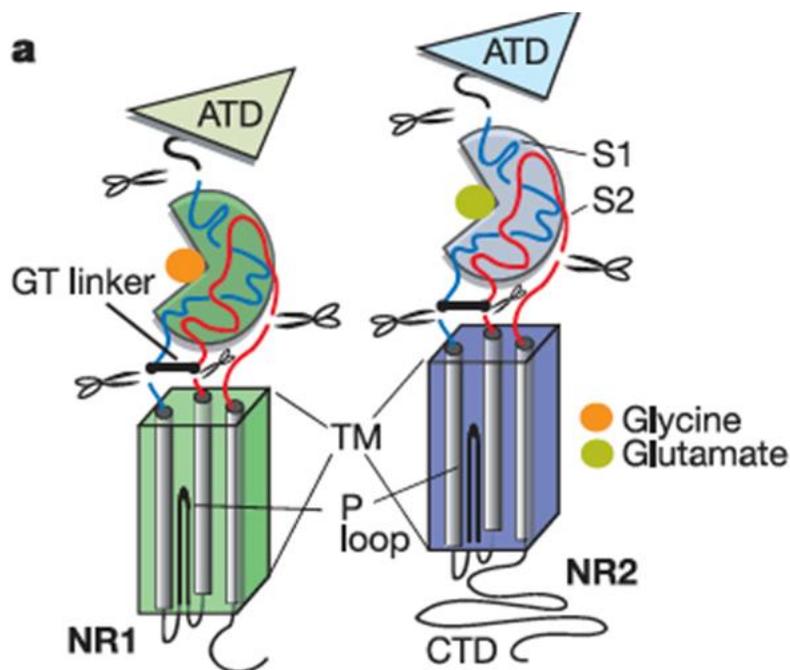


Figure 77: Diagram of the construct of the N-methyl-D-Aspartate Receptor (NMDAR) that was crystallized. Diagram shows the amine terminal domain (ATD), carboxyl terminal domain (CTD) and the transmembrane domain (TM).^{36 37}

The crystal structure of the NR2A ligand-binding core revealed recognition elements for glutamate and suggests a mechanism by which the NR2A subunit binds to glutamate. The major difference between NR2A and the other glutamate receptors is the negatively charged residue that plays a role in the binding of the positive amine group of the agonist. For the NR2A receptor the residue is an aspartate residue at position 731, while for the non-NMDA receptors this residue is a glutamate.

The availability of these crystal structures has provided a platform to explore the possibility of obtaining a crystal structure of the NR2A receptor bound to Con-T. This crystal structure would provide a wealth of information that would prove to be quite valuable in the assessment of the conopeptides as therapeutics. The decision was made to follow the protocol published by Eric Gouaux and coworkers to explore this possibility.

The gene of the truncated NR2A gene was synthesized (by Origene). The gene was then inserted into the pET22b(+) plasmid between the BamHI and HINDIII sites. (**Figure 78**).



Figure 78: pET22b(+) cloning /expression region

The pelB leader region is essential for the expression of proteins that possess disulfide bonds. This sequence is responsible for taking the mature protein into the periplasmic space of the bacterial cell away from the reducing environment of the cytoplasm. This prevents the reduction of the disulfide bonds, which are integral to the integrity of the protein structure. The sequence between the pelB leader and the BamHI site is for the signal peptidase. This signal peptidase cleaves the leader peptide once the protein is safe in the periplasmic space.

The gene was successfully cloned into the pET22b(+) plasmid and test expression was done using the reported protocol. The expression conditions were 0.5mM isopropyl- β -thio-D-thiogalctosidase (IPTG), followed by incubation for 24 hours at 20°C using the Origami B (DE3) cell line. The desired protein (NR2A) will have a His-tag at the C-terminus. Disappointedly when these conditions were used there was no evidence of expression of the desired protein via His-tag purification protocol.

Several different expression conditions were tried using different cell lines, different temperatures, different expression times and different IPTG concentrations.

Table 25: Showing the conditions used using different cell lines, expression conditions, IPTG concentrations and expression times using pET22b(+).

| Cell Lines | Expression Temperatures/ °C | IPTG Concentrations/mM | Expression Times /hours |
|-------------------|--------------------------------|---------------------------|----------------------------|
| BL21(DE3)pLysS | 4, 16, 20 and 37 | 0.25, 0.5, 0.75, 1.0 | 3, 6, 8, 16 and 24 |
| Origami(DE3)pLysS | 4, 16, 20 and 37 | 0.25, 0.5, 0.75, 1.0 | 3, 6, 8, 16 and 24 |
| Tuner (DE3) | 4, 16, 20 and 37 | 0.25, 0.5, 0.75, 1.0 | 3, 6, 8, 16 and 24 |

The results from each of these expression showed that the optimum for all three cell lines was actually 16°C for 16 h with 0.5 mM of IPTG. The problem however, was that even though the protein was expressed, it was in the inclusion bodies in each case. Attempts to denature and refold the protein to obtain a pure sample were futile. The sodium dodecyl sulfate (SDS) gel of the elutions of the urea His-tag purification showed significant amounts of impurities. The impurities seem to either have His-Tag or are closely associated with the protein of interest. The SDS gel shows two significant bands, one of which runs at a larger mass of about 75kD and the other runs at around the 37kD marker. An anti-His western blot of the SDS gel showed that the band at around 37kD gave a positive result, while the band at 75 kD did not. It can be therefore concluded that the impurities are indeed associated with the protein of interest.

Reexamination of the construct led to decision that the pelB leader sequence is unnecessary for NR2A, which has no disulfide bonds. Therefore attempts were made to remove this sequence from the construct. Two routes were considered: 1) change the BamHI cut site at the beginning of the gene to NdeI and vice versa. If the two cut sites are the same then there is

the possibility of removing the pelB leader sequence and ligate to the beginning of the gene sequence right after the ribosome binding site (RBS). This route was aborted as attempts to design suitable primers were futile. For every possibility the primers formed hairpins.

2) The other route was to use different compatible plasmids. In this case, another construct was made using a similar protocol as before using a modified pET28b that would have a small ubiquitin like modifier protein (SUMO) sequence inserted between the gene and the His-tag. The unmodified pET28b plasmid was also used. The differences between the pET22b(+) and pET28b, both modified and unmodified, are, the His-tag is on the N-terminus (pET28b) instead of the C-terminus (pET22b(+)) and there is no pelB leader sequence.

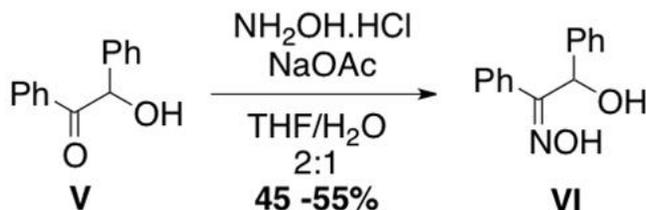
The same conditions were applied as before (see **Table 25**) and the optimum conditions were similar to what was discovered before. The only difference is that the Origami(DE3)pLysS is not compatible with the pET28b, both modified and unmodified. The SDS gels revealed that the protein was in the soluble fraction and not in the inclusion bodies, however, the yield was low. Attempts to improve the yield by modifying the optimum conditions proved to be futile. The best yield obtained to date is 1 mg/L.

Work is currently underway (Ipek Yapici and Janice Chiou) to rectify this problem. A new construct has been designed to incorporate the gene of a fluorescent protein before the gene of the desired protein. This would allow for a possible visual confirmation of the expression of the protein and could also result in an increase in the expression yield.

IV.8 MATERIALS AND METHOD

IV.8.1 Synthesis of L-gamma-carboxyglutamic acid (L-Gla)

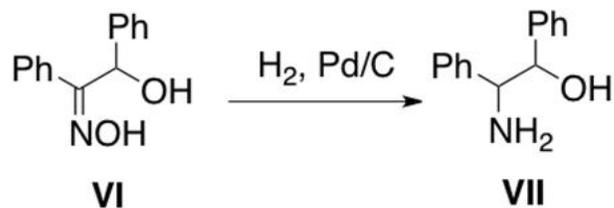
Synthesis of Benzoin oxime (**VI**) was carried out as follows:



Benzoin (2.12g, 10mmol) and hydroxylamine hydrogen chloride (0.70g, 10mmol) were dissolved THF: H₂O (2:1) (10mL). Sodium acetate (0.82g, 10mmol) was added and the resulting mixture left to stir overnight. The THF was removed in vacuo, the resulting product was removed by filtration from aqueous solution. The residue was then dissolved in ether and cooled to 0 °C to precipitate the benzoin oxime **VI**. The oxime was recrystallized from benzene to give a white solid. (0.101g, 4.4%) melting point range 149 – 151 °C; reported range 147 – 148 °C.

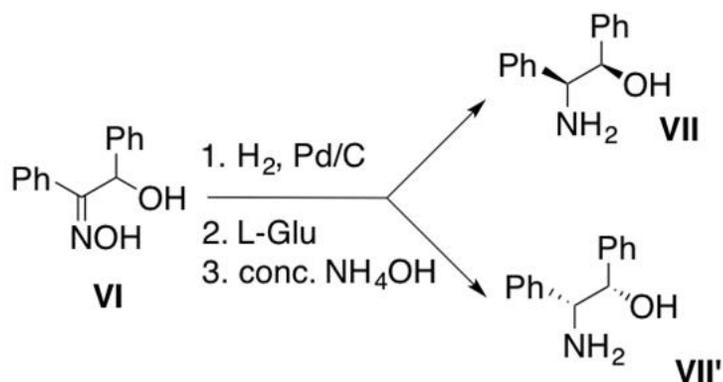
¹H-NMR (300 MHz, DMSO) δ 10.30 (1H, s), δ 7.08– 6.80 (5H, m), δ 5.10 (1H, d, J=4.5 Hz), δ 4.80 (1H, d, J=4.0Hz) ppm. ¹³C NMR (100 MHz, CDCl₃, δ): 128.31, 128.22, 128.31, 127.83, 127.67, 127.57, 127.08, 126.94, 78.41, 61.88 ppm.

Synthesis of 1,2-Diphenyl-2-aminoethanol (**VII**) was carried out as follows:



A mixture of **VI** (1.36g, 6.0 mmol), ethanol (11.2 mL), hydrogen chloride (0.24g, 0.20mL) and 5% palladium- charcoal (0.16g) was hydrogenated at 40 psi. The hydrogenation was complete in 3 hours (may take longer, reaction was followed by TLC). H₂O (10mL) was then added to the mixture. The catalyst was removed via filtration. The filtrate was diluted with H₂O (40 mL) after which concentrated ammonium hydroxide was added in excess. The resulting precipitate was filtered and washed with H₂O. (0.379g, 30%)

¹H-NMR (300 MHz, CDCl₃) δ 7.40- 7.00 (10H, m), δ 4.70 (1H, d, J=6.3Hz), δ 4.20 (1 H, d, J=6.3Hz), δ 1.80 (3H, broad singlet) ppm. ¹³C NMR (100 MHz, CDCl₃, δ): 128.31, 128.22, 128.31, 127.83, 127.67, 127.57, 127.08, 126.94, 78.41, 61.88 ppm.



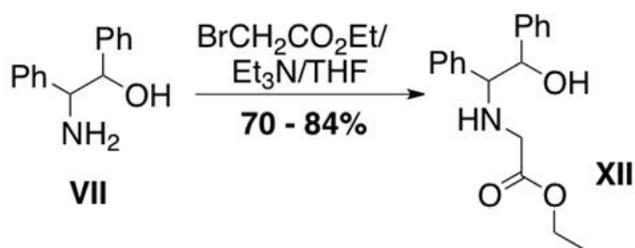
For the resolution of **VII**, a mixture the amino alcohol (2.00g, 9.0mmol) and of L-glutamic acid (1.39g, 9.0 mmol) was dissolved in hot boiling 50% ethanol (60 mL). The resulting mixture was stored overnight at RT. The fine needles were collected and washed with ice cold 50% ethanol and dried. The free base was obtained by dissolving the crystals in warm water then precipitating with excess ammonia. The precipitate was washed and dried with water and dried. No crystals

were observed. To obtain the other isomer, the mother liquor was concentrated to a quarter of its volume after which 95% ethanol was added. The flask was left undisturbed overnight. No crystals were observed.

The procedure was repeated using 1.0g of the amino alcohol. 80mg (16 %) of crystals were obtained from the first crystallization.

$^1\text{H-NMR}$ (300 MHz, CDCl_3) δ 7.40- 7.00 (10H, m), δ 4.70 (1H, d, $J=6.3\text{Hz}$), δ 4.20 (1 H, d, $J=6.3\text{Hz}$), δ 1.80 (3H, broad singlet) ppm. $^{13}\text{C NMR}$ (100 MHz, CDCl_3 , δ): 128.31, 128.22, 128.31, 127.83, 127.67, 127.57, 127.08, 126.94, 78.41, 61.88 ppm.

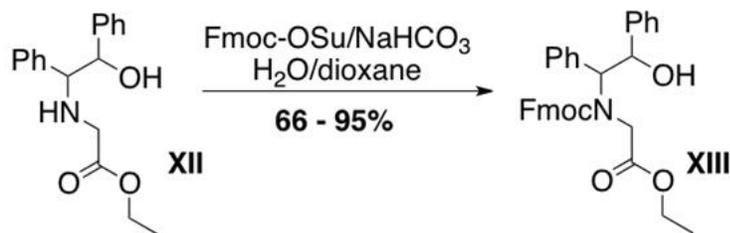
Synthesis of glycinate (**XII**) was carried out as follows:



To a well stirred suspension of **VII** (500mg, 2.25mmol, 1 equiv.) in dry THF was added ethyl bromoacetate (600mg, 3.50mmol, 1.55 equiv) followed by addition of triethylamine (0.65 ml, 2.0 equiv). After being stirred vigorously for 18 h, the mixture was filtered to remove $\text{Et}_3\text{N-HBr}$. The filtrate was evaporated under vacuum. The solid residue was washed with cold water which was then recrystallized from hot absolute ethanol.

$^1\text{H-NMR}$ (500 MHz, CDCl_3) δ 1.20 (3H, t, $J=7.5\text{Hz}$), δ 3.15 (1H, $\frac{1}{2}$ AB q, $J=17.5$ Hz), δ 3.29 (1H, $\frac{1}{2}$ AB q, $J=17.5$ Hz), δ 3.95 (1H, d, 6.0 Hz), δ 4.11 (2H, q, $J=7.1$ Hz), δ 4.80 (1H, d, $J=6.0$ Hz), δ 7.17-7.32 (10H, m) ppm. $^{13}\text{C NMR}$ (100 MHz, CDCl_3 , δ): 172.2, 140.2, 138.5, 128.45, 128.33, 128.13, 127.88, 126.95, 68.29, 60.76, 48.40 ppm.

Synthesis of N-Fmoc-glycinate (**XIII**) was carried out as follows:

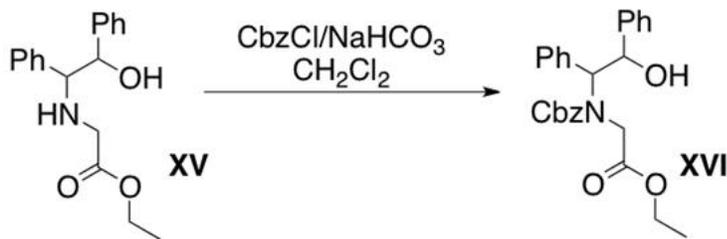


A mixture of **4** (50mg, 0.015mmol), NaHCO_3 (65mg, 0.78mmol), and H_2O (1mL) was cooled in an ice bath. To the vigorously stirring solution was added dropwise Fmoc-Su (107mg, 0.32mmol) in dioxane. After the addition the ice bath was removed and the suspension stirred for an additional 1h. H_2O (1mL) was added and the solution extracted with EtOAc (3 x 5mL). the combined organic layers were then washed with 0.2 M HCl (5mL) and brine (5mL) and dried with anhydrous Na_2SO_4 . The organic layers were then concentrated to a clear oil. Toluene (1 mL) was added and the solution concentrated to a solid. The crude product was taken to the next step.

$^1\text{H-NMR}$ -crude (300 MHz, CDCl_3): δ 1.22 (3H, t, $J=7.5\text{Hz}$), δ 3.33 (2H, d, $J=10\text{ Hz}$), δ 4.08-4.56(4H, dd, $J=7.5\text{ Hz}$, $J=14\text{Hz}$) δ 4.35 (1H, t, 7.5 Hz), δ 4.55 (1H, d, 7.5 Hz) δ 7.20-7.45 (16H, m), δ 7.61(1H, d, $J=7.5\text{ Hz}$), δ 7.76 (1H, d, 7.5 Hz) ppm.

$^{13}\text{C NMR}$ (75 MHz, CDCl_3 , δ): 168.5, 150.61, 142.4, 141.28, 130.03, 128.55, 128.19, 127.95, 127.41, 125.2, 120.15, 72.92, 68.91, 46.32, 25.43 ppm.

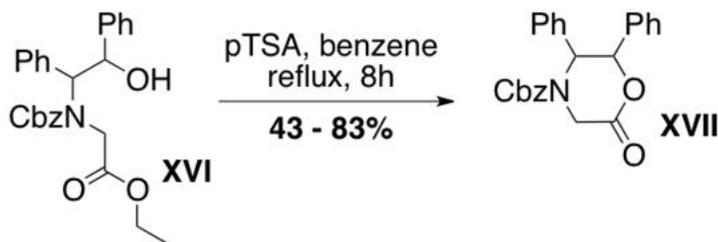
Synthesis of N-Cbz-glycinate (**XVI**) was carried out as follows:



To a vigorously stirred mixture of the ethyl ester **XV** (500 mg, 1.67 mmol, 1.0 equiv) in CH₂Cl₂ (8.0mL) and saturated aqueous NaHCO₃ (8.0mL) was added benzylchloroformate (316 mg, 1.83 mmol, 1.1 equiv.), the mixture was then stirred for 12 h. The aqueous layer was separated and extracted with 3 x CH₂Cl₂ and the combined organic layers were combined and washed with water. The combined organic extracts were dried over anhydrous sodium sulfate, filtered, and concentrated to give colorless oil, which was carried on crude.

¹H NMR (300 MHz, CDCl₃) δ 1.02 (3H, t, J=7.1Hz), δ 3.7-4.0 (5H, m), δ 4.98-5.18 (2H, m,) δ 5.42 – 5.53 (2H, m), δ 7.1-7.5 (15H, m) ppm. ¹³C NMR (125 MHz, CDCl₃, δ): 171.65, 156, 140.94, 136.14, 135.3, 130.32, 129.04, 128.71, 128.55, 128.42, 128.37, 128.26, 127.73, 127.16, 126.95, 126.21, 74.53, 61.43, 53.39 ppm.

Synthesis of N-Cbz- oxazinone (**XVII**) was carried out as follows:

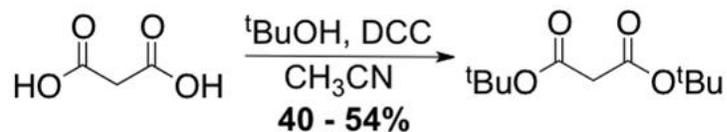


To a stirred solution of crude **XVI** (800mg, 1.67 mmol, 1.0 equiv) in benzene (20mls) in a 25ml one-necked round bottom flask equipped with a Soxhlet extractor packed with CaCl₂ was added p-toluenesulfonic acid monohydrate (33.3mg, 0.167 mmol, 0.1equiv.). The mixture was brought to reflux for 8 h. The mixture was allowed to cool, and the resultant precipitate was collected washed with water, and recrystallized from hot absolute ethanol.

¹H NMR (300 MHz, CDCl₃) δ 4.60 (2H, AB q, J=17.6 Hz), δ 5.06 (2H, AB q, J= 12.6 Hz), δ 5.29 (1H, d, J=3Hz), δ 6.24 (1H, d, J=3Hz), δ 6.66 (2H, s), δ 6.86 (1H, s), δ 7.0-7.4 (12H, m)

ppm. ^{13}C NMR (125 MHz, CDCl_3 , δ): 167.30, 158.52, 128.65, 128.49, 128.32, 128.25, 128.20, 128.11, 127.99, 127.73, 127.47, 126.38, 126.17, 81.02, 80.41, 68.06, 67.74, 60.33, 59.13 ppm.

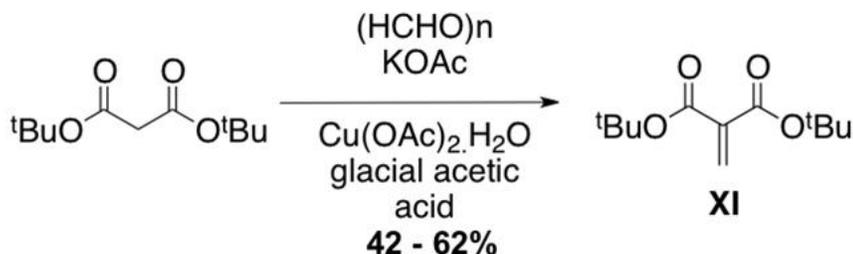
Synthesis of Di-tert-butyl malonate was carried out as follows:



To a solution of malonic acid (1.06g, 10mmol, 1.0 equiv.) and tert butyl alcohol (1.44g, 20mmol, 2.0 equiv.) in dry acetonitrile (30 mL) was added a solution of DCC (4.21g, 20mmol, 2.0 equiv.) in acetonitrile (20 mL). The reaction mixture was then allowed to stir at rt for 20 minutes, filtered and evaporated. The residue was bulb-to-bulb distilled using the Kugelrohr apparatus to give the pure product, di- tert- butyl malonate.

$^1\text{H-NMR}$ (300 MHz, CDCl_3) δ 1.44 (18H, s), δ 3.15 (2H, s) ppm. ^{13}C NMR (125 MHz, CDCl_3 , δ): 166.12, 81.54, 44.28, 27.87 ppm.

Synthesis of Di-tert-butyl methylene malonate was carried out as follows:

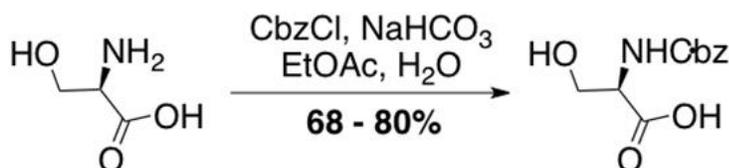


In a 25 mL round bottom flask fitted with a condenser, and under N_2 atmosphere, paraformaldehyde, di-tert-butyl malonate, cupric acetate monohydrate, potassium acetate and glacial acetic acid (1.3mL) were mixed at room temperature. The resulting mixture was then heated to 100°C while being stirred for 2 h. The acetic acid was then removed under reduced

pressure first at room temperature then at 100 °C. The residue was distilled in the Kugelrohr apparatus at a temperature range of 110 – 140 °C. The distilled product was dissolved in ether (1.3 mL) and washed with a 5% aqueous solution NaHCO₃ (2 x 1 mL). The aqueous layer was extracted with ether (2 x 1.3 mL). The organic layer were combined and dried with Na₂SO₄. The solvent was removed in vacuo, and the residue distilled using the Kugelrohr apparatus. The product was collected in a temperature range of 110 – 120 °C.

¹H-NMR (500 MHz, CDCl₃) δ 1.49(18H, s), 6.22 (2H, s) . ¹³C NMR (125 MHz, CDCl₃, δ): 163.73, 138.34, 130.78, 81.85, 28.07, 28.02, 27.96 ppm.

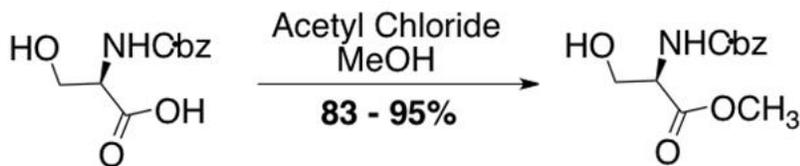
Synthesis of N-Cbz-L-serine was carried out as follows:



To a well stirred suspension of L-serine in ethyl acetate (8.3 mL) and H₂O (8.3 mL) sodium bicarbonate (1.58g, 19 mmol) was added. Benzylchloroformate (1.62g, 9.5 mmol) was dissolved in ethyl acetate (8.3 mL) and then added dropwise over 30 minutes at 0 °C. The mixture was left to stir overnight. The organic layer was extracted using ethyl acetate. The combined organic layers were combined and dried with anhydrous sodium sulfate. All the ethyl acetate was removed in vacuo to give a white solid.

¹H-NMR (300 MHz, CDCl₃) δ 7.40-7.20 (5H, m), δ 4.99 (1H, s), δ 4.18 (1H, m), δ 3.83- 3.11 (2H, dd J=7.8Hz, 24 Hz) ppm.

Synthesis of N-Cbz-L-methyl serinate was carried out as follows:

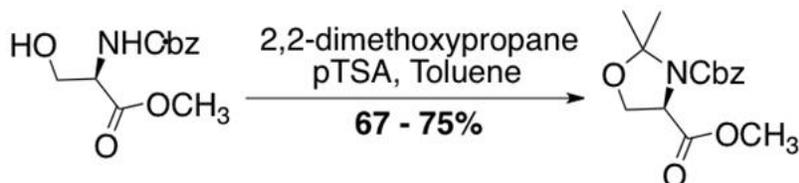


Acetyl chloride (1.55g, 19.7 mmol) was added dropwise over 15 min to MeOH (9.8 mL) at 0 °C.

The solution was then stirred at 0 °C for another 15 min. The white solid product from above, N-Cbz-L-serine (1.80g, 7.1 mmol) was then added in small portions and the resulting mixture was stirred for an extra 30 minutes. The mixture was then heated slowly and maintained at reflux for 2 h. The resulting solution was cooled and the solvent removed in vacuo. The residue was dissolved in chloroform (16.3 mL) and washed with 10 % sodium bicarbonate solution (2 x 6.5 mL), and dried with anhydrous sodium sulfate. The mixture was concentrated in vacuo to give product as a colorless oil.

¹H-NMR (300 MHz, CDCl₃) δ 7.36-7.23 (5H, m), δ 5.86 (1H, s), δ 5.01 (2H, s), δ 4.40 (1H, m), δ 4.00- 3.73(2H, dd) J=3.3Hz, 14.1Hz, δ 3.73 (3H, s) ppm. ¹³C NMR (125 MHz, CDCl₃, δ): 170.96, 156.22, 136.04, 128.70, 128.53, 128.36, 128.24, 128.10 67.19, 63.24, 56.02, 52.70 ppm.

Synthesis of acetonide was carried out as follows:



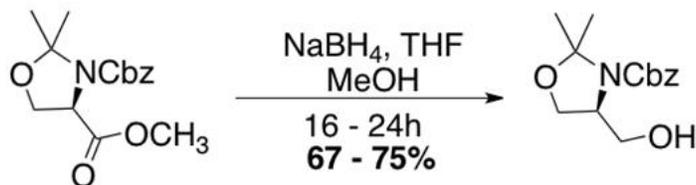
To a solution of the ester (1.45g, 5.7 mmol) in dry toluene (19.8 mL) was added 2,2-dimethoxypropane (1.43 mL) and pTSA•H₂O (27 mg, 0.14mmol). The solution was refluxed for 30 minutes then 12 mL of the methanol/toluene mixture was distilled off. To the residue 2,2-dimethoxypropane (0.92 mL, 5.7 mmol) was added and the mixture refluxed for another 30

minutes. 9.0 mL of methanol/toluene was distilled off, (reaction was followed by TLC). The cooled solution was poured into a saturated sodium bicarbonate solution (6 mL). The organic layers were extracted with ether (2 x 6 mL). The organic layers were combined and washed with brine and dried with anhydrous sodium sulfate. The solvent was removed in vacuo to give the crude product as amber oil. The crude was purified via flash chromatography (ethyl acetate: hexanes, 1:1).

$^1\text{H-NMR}$ α -form (300 MHz, CDCl_3) δ 7.34-7.25 (5H,m), δ 5.17 (2H, d) $J=1.8$ Hz, δ 4.44 (1H, m) δ 4.17- 4.06 (2H, m) δ 3.71(3H, s), δ 1.86 (3H, s), δ 1.54 (3H, s) ppm. ^{13}C NMR (125 MHz, CDCl_3 , δ): 171.22, 151.75, 136.32, 128.41, 127.74, 95.48, 66.70, 66.57, 58.86, 52.37, 24.92 , 24.10 ppm.

$^1\text{H-NMR}$ β -form (300 MHz, CDCl_3) δ 7.34-7.25 (5H,m), δ 5.00 (2H, d) $J=1.8$ Hz, δ 4.51 (1H, m) δ 4.17- 4.06 (2H, m) δ 4.06 (3H, s), δ 1.62 (3H, s), δ 1.47 (3H, s) ppm. ^{13}C NMR (125 MHz, CDCl_3 , δ): 171.09, 151.75, 136.01, 128.34, 127.99, 94.81, 66.15, 60.34, 59.55, 52.50, 26.02, 25.16 ppm.

Synthesis of reduced acetonide was carried out as follows:



To a well stirred and cooled (-10°C , ice +NaCl) solution of acetonide ester (0.99g, 3.3mmol) in THF (12 mL), solid NaBH_4 (0.51g, 13.4 mmol) was added. The resulting mixture was stirred for 30 minutes at the same temperature after which MeOH was added dropwise. The mixture was then stirred for 16 h at room temp. After 16 h H_2O was added carefully to the reaction mixture

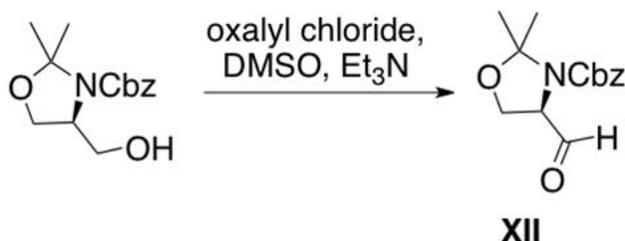
and stirred for another 30 mins. Most of the organic solvent was removed in vacuo after which brine was added. The mixture was extracted using ethyl acetate (3 x 7.5mL). The combined extracts were then washed with brine (5mL) and dried with anhydrous sodium sulfate. The ethyl acetate was removed under vacuum and the residue purified with 1:1 ethyl acetate/hexanes by flash chromatography to give a colorless oil as product.

$^1\text{H-NMR}$ α -form, 70% (300 MHz, C_6D_6) δ 7.28-7.06 (5H, m), δ 5.10 (2H, s), δ 4.00- 3.79 (2H, m), δ 3.85 (2H, m), δ 3.42-3.36 (1H, m), δ 1.78 (3H, s), δ 1.57(3H, s) ppm.

$^1\text{H-NMR}$ β -form, 30% (300 MHz, C_6D_6) δ 7.28-7.06 (5H, m), δ 4.98 (2H, d) $J=12.3$ Hz, δ 4.00- 3.79 (2H, m), δ 3.85 (2H, m), δ 3.42-3.36 (1H, m), δ 1.49 (3H, s), δ 1.32 (3H, s).

$^{13}\text{C NMR}$ α and β -forms (125 MHz, CDCl_3 , δ): 152.16, 154.00, 135.70, 137.30, 126.75, 127.25, 127.47, 127.49, 127.57, 127.62, 127.74, 127.92, 128.10, 128.19, 94.10, 94.16, 67.44, 66.64, 65.25, 65.10, 60.25, 59.52, 26.97, 26.28, 24.45 ppm.

Synthesis of acetonide protected Aldehyde (**XII**) was carried out as follows:

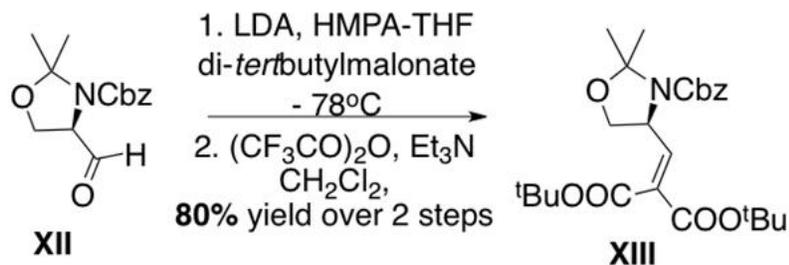


Oxalyl chloride (0.6 mL, 5.93 mmol) was added at -78°C to a solution of DMSO (0.9 mL, 13.0 mmol) in dry DCM. The mixture was then stirred for 10 mins. After stirring a solution of the acetonide alcohol (1.57g, 5.93 mmol) in dry DCM was carefully added and the mixture stirred for another 15 mins. Et_3N (4.20 mL, 30mmol) was then added at once and after 10 mins the reaction mixture was allowed to warm to room temp. and quenched with H_2O (30 mL). The

organic layer was separated and the aqueous phase extracted with ether (3 x 25mL). The combined organic layers were dried with anhydrous sodium sulfate and the solvent removed under reduced pressure to give 5.

¹H-NMR α and β -forms (300 MHz, CDCl₃) δ 9.55 (1H, s), δ 9.60 (1H, s), δ 7.35- 7.25 (5H, m), δ 5.10 & 5.18 (2H, s), δ 4.39 (1H, dd) J=7.5 Hz, J=24.3 Hz, δ 4.28-3.94 (2H, m), δ 1.65(3H, s), 1.58(3H,s), 1.56 (3H, s), 1.49 (3H, s) ppm. ¹³C NMR α and β -forms (125 MHz, CDCl₃, δ): 169.41, 149.86, 136.07, 134.74, 128.63, 128.49, 128.47, 128.19, 128.09, 128.02, 127.98, 127.91, 127.72, 127.56, 96.94, 95.44, 68.32, 66.76, 65.63, 25.74, 24.06 ppm.

Synthesis of 2S-(3-benzyloxycarbonyl-2,2-dimethyloxazolidin-4-ylmethylene)-malonic acid di-tert-butyl ester_(**XIII**) was carried out as follows:



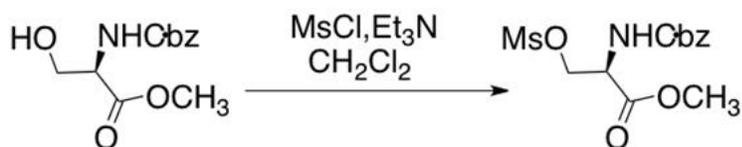
To a solution of diisopropylamine (0.5 mL, 4.2 mmol) in anhydrous THF (5.3 mL) was added n-BuLi (1.42 mL, 1.6 M in hexane, 2.3 mmol) at 0 °C under N₂, and the mixture were stirred for 30 min. After stirring for an additional 30 min at -78 °C, anhydrous HMPA (0.76 mL, 0.23 mmol) was added, and the mixture was stirred for 30 min. A solution of di-*tert*-butyl malonate (0.5 g, 2.3 mmol) in THF (1.8 mL) was injected into the above mixture. After 30 min, a solution of D-Garner's aldehyde (0.5g, 2.0 mmol) in THF (3.5 mL) was introduced, and the reaction

mixture was stirred for 2 h at -78°C . The mixture was quenched with saturated aqueous NH_4Cl and extracted with ether.

The organic layer was washed with brine and dried (Na_2SO_4). Removal of the solvents afforded a crude oil. To the mixture of the above oil and Et_3N (1.3 mL, 9.2 mmol) in CH_2Cl_2 at 0°C was added $(\text{CF}_3\text{CO})_2\text{O}$ (0.64 mL, 45.5 mmol). The reaction was stirred at 0°C for 12 h and at rt for 6 h, quenched with saturated NH_4Cl , and extracted with CH_2Cl_2 . After drying (Na_2SO_4), the extracts were filtered and concentrated under reduced pressure. The crude product was purified by flash column chromatography on silica gel to afford pure **XIII** (0.70 g, 80%).

^1H NMR (300 MHz, CDCl_3): 7.26–7.36 (5H, m), 6.73 (1H, d, $J = 8.8$ Hz), 5.13 (2H, m), 4.80 (1H, m), 4.22 (1H, m), 3.86 (1H, dd, $J = 3.2, 9.2$ Hz), 1.43–1.56 (24H, m) ppm. ^{13}C NMR (125 MHz, CDCl_3 , δ): 163.90, 163.06, 152.09, 146.75, 136.32, 128.52, 128.33, 128.11, 128.03, 127.76, 127.60, 95.11, 82.17, 81.95, 68.39, 66.64, 55.73, 27.94, 26.12, 23.87 ppm.

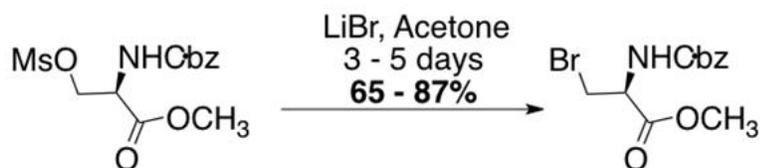
Synthesis of the Mesylate was carried out as follows:



The alcohol (200 mg, 0.80 mmol) was dissolved in dichloromethane (5 ml) and cooled to 0°C . Triethylamine (0.2 ml) was then added followed by methanesulfonylchloride (MsCl) (0.1 ml, 90 mg, 0.80 mmol). The resulting mixture was stirred at room temperature for 20 mins. The reaction was then quenched with water and extracted using dichloromethane. The combined organic layers were combined and dried over Na_2SO_4 and concentrated to give a solid. The solid was

dissolved in a little dichloromethane and passed through a silica gel plug to remove the excess ammonium salt byproduct. The crude solid was taken to the next step.

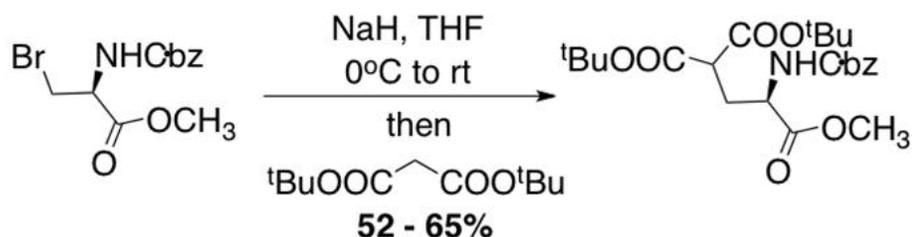
Synthesis of the Bromide was carried out as follows:



The mesylate (170 mg, 0.51 mmol) was dissolved in acetone (10 ml). Lithium bromide (44.6 mg, 0.51 mmol) was then added and the resulting mixture was vigorously stirred at room temperature for 5 days. The reaction was quenched with water and then extracted using ethyl acetate. The combined organic layers were dried over Na₂SO₄ and concentrated. The crude was then purified using column chromatography using 30% ethyl acetate/hexanes to give a white solid (130 mg).

¹H-NMR (300 MHz, CDCl₃) δ 7.36-7.30 (5H, m), δ 5.63 (1H, s), δ 5.13 (2H, s), δ 4.80- 4.72 (1H, m), δ 3.98-3.69 (2H, m), δ 3.80 (3H, s) ppm. ¹³C NMR (125 MHz, CDCl₃, δ): 169.30, 155.53, 135.93, 128.57, 128.52, 128.30, 128.21, 128.12, 67.31, 54.29, 53.08, 33.67 ppm.

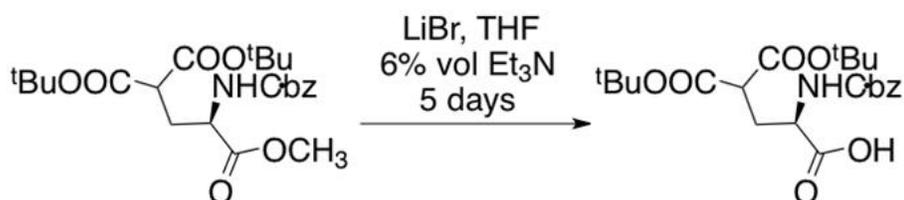
Synthesis of the N-Cbz-L-gammacarboxyglutamic acid di-*tert*-butyl methyl ester was carried out as follows:



To a suspension of NaH in anhydrous tetrahydrofuran (THF) (10 ml) at 0°C under N₂ atmosphere, di-*tert*-butyl malonate (218 mg, 1.01 mmol) was added. The resulting mixture was then stirred for 2h at 0°C after which the bromide (397 mg, 1.01 mmol) was added. The reaction mixture was allowed to warm up to room temperature and stirred vigorously overnight under the inert N₂ atmosphere. The reaction was then quenched with water and then extracted using ethyl

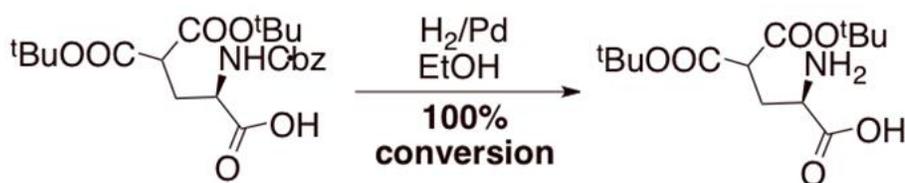
acetate. The combined organic layers were dried over Na_2SO_4 and concentrated. Attempts to purify the product resulted in loss of the product and so the crude product was taken to the next step.

Hydrolysis of the N-Cbz-L-gammacarboxyglutamic acid di-*tert*-butyl methyl ester was carried out as follows:



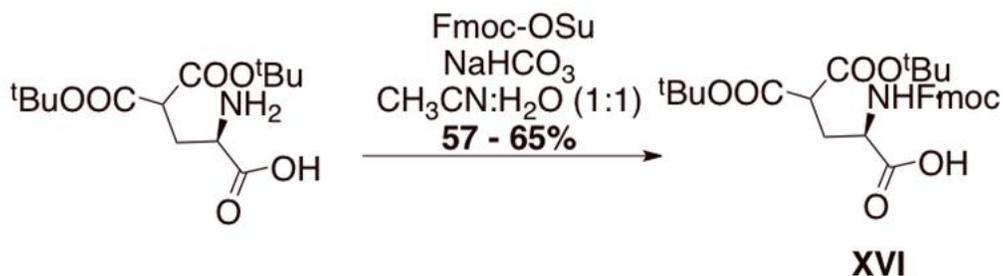
The ester (10 mg, 0.02 mmol) was dissolved in anhydrous tetrahydrofuran (THF) (0.1 mL) then 6 vol% of water (6.0 μl) was added. The resulting solution was then charged with triethylamine (13.0 μl) followed by lithium bromide (8 mg, 0.10 mmol). The reaction mixture was then stirred vigorously at room temperature for 5 days. The reaction was then quenched with the addition of acetic acid.

Synthesis of L-gammacarboxyglutamic acid di-*tert*-butyl ester was carried out as follows:



To the crude carboxylic acid (200 mg), ethanol (2.6 ml) and 10% Pd/C (40 mg) were added. The resulting mixture was stirred vigorously under a H_2 atmosphere until the hydrogen is absorbed to saturation. The mixture was then filtered through a celite plug and concentrated to give a crude oil, which was taken directly through to the Fmoc protection step.

Synthesis of N-Fmoc-L-gammarcoxyglutamic acid di-*tert*-butyl ester was carried out as follows:



A mixture of the above oil (145 mg), *N*-(9-fluorenylmethoxycarbonyloxy) succinimide (Fmoc-OSu) (203 mg, 0.43 mmol) and NaHCO₃ (203 mg, 1.72 mmol) in water and acetonitrile (1:1, 3 ml) was stirred at room temperature overnight. The acetonitrile was then evaporated under pressure and the pH of the resulting solution was adjusted to pH = 4 using 10 % citric acid. The mixture was then extracted with ethyl acetate. The combined organic layers were washed with brine, dried over Na₂SO₄ and concentrated. The crude product was purified by chromatography to give a solid (152 mg, 60%).

¹H-NMR (300 MHz, CDCl₃) δ 7.80 (2H, d, J = 7.5 Hz), δ 7.45 (2H, dd, J = 3.0, 7.2 Hz) δ δ 7.57 (2H, dd, J = 7.1 Hz), δ 7.40 (2H, dd, J = 7.2 Hz), δ 7.40 (2H, ddd, J = 1.3, 7.5 Hz), δ 5.56 (1H, d, J = 7.8 Hz), δ 4.45 (2H, m), δ 4.33 (1H, m), δ 4.22 (1H, t, J = 7.0 Hz), δ 3.41 (1H, t, J = 7.1 Hz), 2.43 (1H, m), 2.21 (1H, m), 1.45 – 1.49 (18H, m) ppm. ¹³C NMR (125 MHz, CDCl₃, δ): 176.0, 168.50, 168.30, 155.89, 143.93, 143.60, 141.41, 128.30, 127.93, 128.30, 125.21, 119.80, .82.5, 67.81, 67.28, 52.29, 50.58, 47.0, 30.4, 26.78 ppm.

IV.8.2 Peptide Synthesis

IV.8.2.1 Method used for the manual synthesis

The Wang resin was pre-swelled in a disposable 5 ml disposable syringe with frit by rotating with dichloromethane (3.3ml for 52.2 mg of resin) for one hour. The resin was then washed 5 x 4 ml portions of dimethylformamide (DMF). The *N*-9-fluorenylmethoxycarbonyloxy group was removed using by treatment with 5 bed volumes of 20% piperidine in DMF for 20 minutes. Meanwhile the first amino acid (5.0 equivalents) in the sequence (C-terminus) was dissolved in DMF along with HATU. Diisopropylcarbodiimide (DIC) (5.0 equivalents) then added and the resulting mixture was stirred at room temperature for 30 minutes. After 20 minutes the resin was washed with DMF 5 x 4 ml portions. The amino acid, HATU and DIC mixture was then added to the resin and incubated with rotation for 30 minutes. Other coupling reagents such as PyBOP and HBTU were used.

Treating the resin with 20% piperidine in DMF for 3 minutes deprotected the Fmoc terminal amino acid of the growing chain. The resin was then washed with 5 x 4 ml portions of DMF. During this the next amino acid is prepared using the protocol described above for the next coupling. This whole process was repeated until the sequence is complete. The peptide was cleaved using trifluoroacetic acid (TFA).

IV.8.2.2 Method used to synthesize the peptides via Peptide Synthesizer

method-listing: 04-05-12- 10 microdouble_mc. M PM
path: X:\Li\Research\Peptide Synthesizer-MSU\04-05-1 2 Camille\
** 04.04.2012 - 12:22:20
** Software: MultiPep
Prepare
1 Memo ResPep SL version, 10 pmol peptide synthesis
2 RinseNeedle 1000 /2000 ul
3 WashColumns 300 p1, Reservoir->Peptides, 4x
4 Extract 30s
Cycle: 1 ->2
5 Deprotection 200
pI,
Piperidine->Peptides
6 Deprotection 180 p1, Piperidine->Peptides
7 RinseNeedle 1000 / 2000 ul
8 WashColumns 300 p1, Reservoir->Peptides, 2x
9 WashColumns 250 p1, Reservoir->Peptides, 6x
10 Extract 30s
11 Coupling 180+11 +2+160->Peptides /... I...
12 Coupling 180+1 1+2+160->Peptides /... I...
13 Capping 200 p1, CapMixture->Peptides
14 RinseNeedle 1000 / 2000 ul
15 WashColumns 300 p1, Reservoir->Peptides, 2x
16 WashColumns 250 p1, Reservoir->Peptides, 5x
17 Extract 30s
Cycle: 3->5
18 Deprotection
19 Deprotection
20 RinseNeedle
21 WashColumns
22 WashColumns
23 Extract
24 Coupling
25 Coupling
26 Capping
27 RinseNeedle
28 WashColumns
29 WashColumns
30 Extract
200 p1, Piperidine->Peptides
180 p1, Piperidine->Peptides
1000 / 2000 ul
300 p1, Reservoir->Peptides, 2x
250 p1, Reservoir->Peptides, 6x

30s
180+1 1+2-4-68->Peptides /... /...
180+11 +2+68->Peptides L. /...
200 p1, CapMixture->Peptides
1000 / 2000 ul
300 p1, Reservoir->Peptides, 2x
250 p1, Reservoir->Peptides, 5x
30s
Cycle: 6 -> 6
31 Deprotection 200 p1, Piperidine->Peptides
32 Deprotection 180 p1, Piperidine->Peptides
33 RinseNeedle 1000 / 2000 ul
34 WashColumns 300 p1, Reservoir->Peptides, 2x
35 WashColumns 250 p1, Reservoir->Peptides, 6x
36 Extract 30s
37 Coupling 180+1 1+2+160->Peptides I... /...
38 Coupling 180+1 1+2+160->Peptides I... /...
39 Capping 200 p1, CapMixture->Peptides
40 RinseNeedle 1000 / 2000 ul
41 WashColumns 300 p1, Reservoir->Peptides, 2x
42 WashColumns 250 p1, Reservoir->Peptides, 5x
43 Extract 30s
Cycle: 7->8
44 Deprotection 200 p1, Piperidine->Peptides
45 Deprotection 180 p1, Piperidine->Peptides
46 RinseNeedle 1000 / 2000 ul
47 WashColumns 300 p1, Reservoir->Peptides, 2x
48 WashColumns 250
p1,
Reservoir->Peptides, 6x
49 Extract 30s
50 Coupling 180+1 1+2+68->Peptides I... I...
51 Coupling 180+11 +2+68->Peptides I... I...
52 Capping 200 p1, CapMixture->Peptides
53 RinseNeedle 1000 / 2000 ul
54 WashColumns 300 p1, Reservoir->Peptides, 2x
55 WashColumns 250 p1, Reservoir->Peptides, 5x
56 Extract 30s
Cycle: 9 -> 9
57 Deprotection 200 p1, Piperidine->Peptides
58 Deprotection 180 p1, Piperidine->Peptides
59 RinseNeedle 1000 / 2000 ul
60 WashColumns 300 p1, Reservoir->Peptides, 2x
61 WashColumns 250 p1, Reservoir->Peptides, 6x
62 Extract 30s
63 Coupling 180+1 1+2+160->Peptides I... I...

64 Coupling 180+11+2+1 60->Peptides I.. I...
65 Capping 200 p1, CapMixture->Peptides
66 RinseNeedle 1000 / 2000 ul
67 WashColumns 300 p1, Reservoir->Peptides, 2x
68 WashColumns 250 p1, Reservoir->Peptides, 5x
69 Extract 30s
Cycle: 10->11
70 Deprotection
71 Deprotection
72 RinseNeedle
73 WashColumns
74 WashColumns
75 Extract
76 Coupling
77 Coupling
78 Capping
79 RinseNeedle
80 WashColumns
81 WashColumns
82 Extract
200 p1, Piperidine->Peptides
180 p1, Piperidine->Peptides
1000/2000 ul
300 p1, Reservoir->Peptides, 2x
250 p1, Reservoir->Peptides, 6x
30s
180+1 1+2+68->Peptides /... /...
180+1 1+2+68->Peptides /... /...
200 p1, CapMixture->Peptides
1000/2000 ul
300 i.il, Reservoir->Peptides, 2x
250 p1, Reservoir->Peptides, 5x
30s
Cycle: 12->13
83 Deprotection 200 p1, Piperidine->Peptides
84 Deprotection 180 p1, Piperidine->Peptides
85 RinseNeedle 1000 / 2000 ul
86 WashColumns 300 p1, Reservoir->Peptides, 2x
87 WashColumns 250 p1, Reservoir->Peptides, 6x
88 Extract 30s
89 Coupling 180+11+2+160->Peptides/... /...
90 Coupling 180+1 1+2+160->Peptides I... /...
91 Capping 200 p1, CapMixture->Peptides
92 RinseNeedle 1000 / 2000 ul
93 WashColumns 300 p1, Reservoir->Peptides, 2x
94 WashColumns 250 p1, Reservoir->Peptides, 5x

95 Extract 30s
Cycle: 14 -> max
(->
14)
96 Deprotection 200 p1, Piperidine->Peptides
97 Deprotection 180 p1, Piperidine->Peptides
98 RinseNeedle 1000 / 2000 ul
99 WashColumns 300 p1, Reservoir->Peptides, 2x
100 WashColumns 250 p1, Reservoir->Peptides, 6x
101 Extract 30s
102 Coupling
103 Coupling
104 Capping
105 RinseNeedle
106 WashColumns
107 WashColumns
108 Extract
180+1 1+2+68->Peptides I... I...
180+11+2+68->Peptides I... I...
200 pI,
CapMixture->Peptides
1000/2000 ul
300 p1, Reservoir->Peptides, 2x
250 p1, Reservoir->Peptides, 5x
30s
Final
109 Deprotection 200 p1, Piperidine->Peptides
110 Deprotection 180 p1, Piperidine->Peptides
111 Deprotection 180 p1, Piperidine->Peptides
112 RinseNeedle 1000 / 2000 ul
113 WashColumns 300 p1, Reservoir->Peptides, 2x
114 WashColumns 250 p1, Reservoir->Peptides, 6x
115 WashColumns 250 p1, DCM->Peptides, 3x
116 Extract 300s
117 RinseNeedle 1000 / 3000 ul
end of method ***

IV.8.3 HPLC run protocol

Buffer A: 0.06% trifluoroacetic acid (TFA) in H₂O.

Buffer B: 0.052% trifluoroacetic acid (TFA) in 80% acetonitrile in H₂O.

Concentration of peptides used in each run is 5 mg/ml.

Table 26: HPLC run for Con-G.

| Time | %A | %B |
|-------|-----|-----|
| 0:00 | 100 | 0 |
| 5:00 | 70 | 30 |
| 25:00 | 70 | 30 |
| 65:00 | 70 | 30 |
| 88:00 | 0 | 100 |

**Detection wavelength was 210 nm

Table 27: HPLC run for Ala-Con-G.

| Time | %A | %B |
|-------|----|-----|
| 0:00 | 0 | 10 |
| 5:00 | 0 | 10 |
| 25:00 | 0 | 40 |
| 65:00 | 0 | 40 |
| 88:00 | 0 | 100 |

**Detection wavelength was 210 nm

IV.8.4 Mass Spectra Data of the peptides synthesized

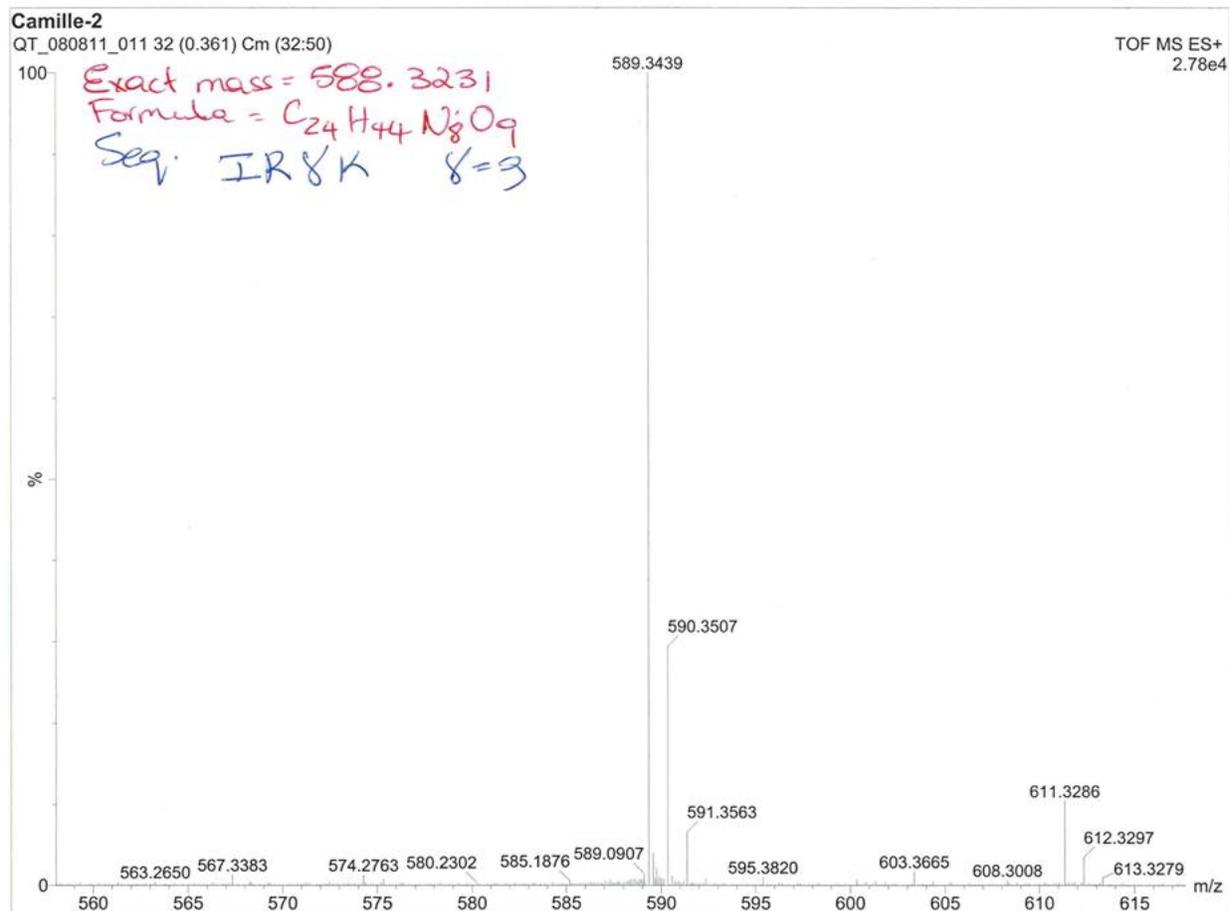


Figure 79: Mass Spectra of the Con-G-tetramer (sequence IRγK). Major peak corresponds to the $[M + H]^+$; Observed Mass = 589.3439; Expected Mass = 588.3231.

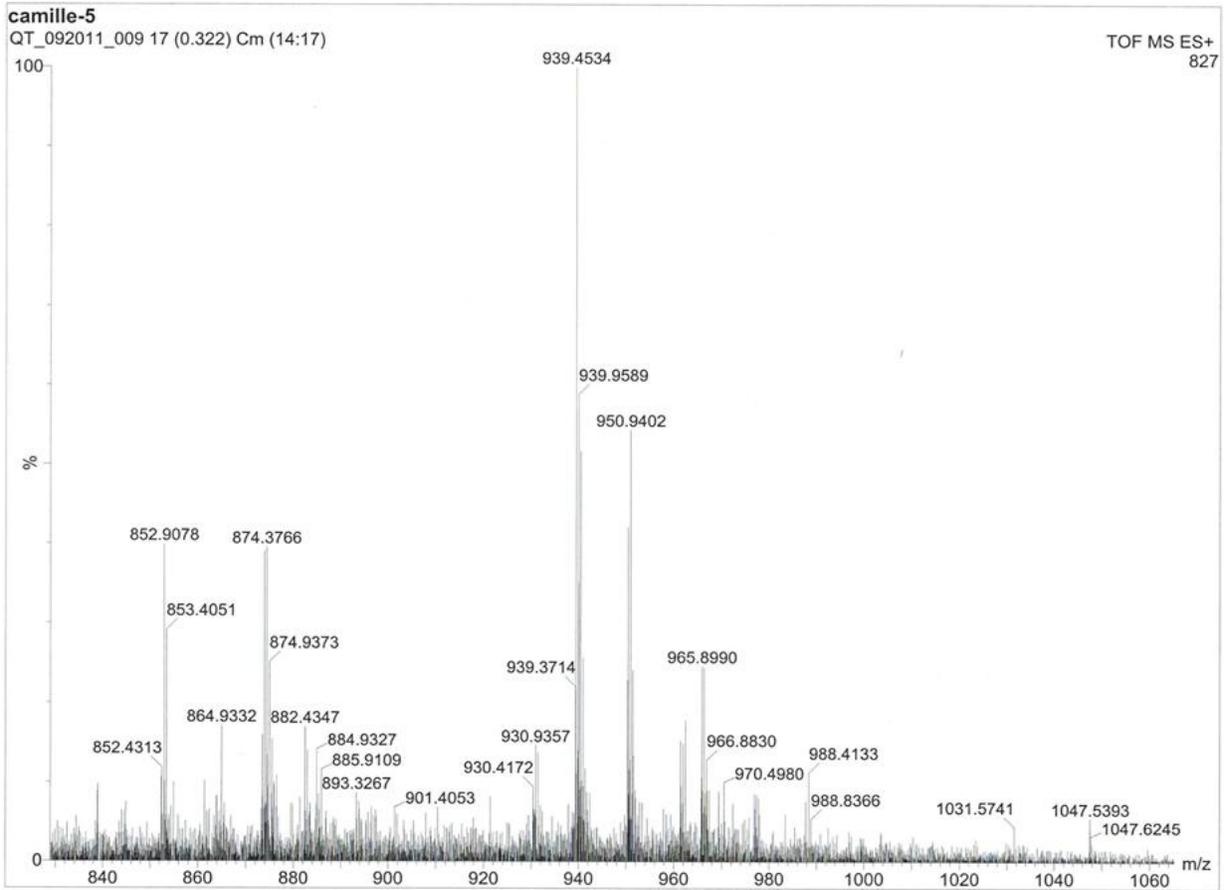


Figure 80: Mass Spectra of Con-G-13 mer (sequence $\gamma\gamma\text{LQ}\gamma\text{NQ}\gamma\text{LIR}\gamma\text{K}$). Observed Mass = 939.4534 which corresponds to $[\text{M} + 2\text{H}]^{2+}$; Expected Mass = 1876.7810.

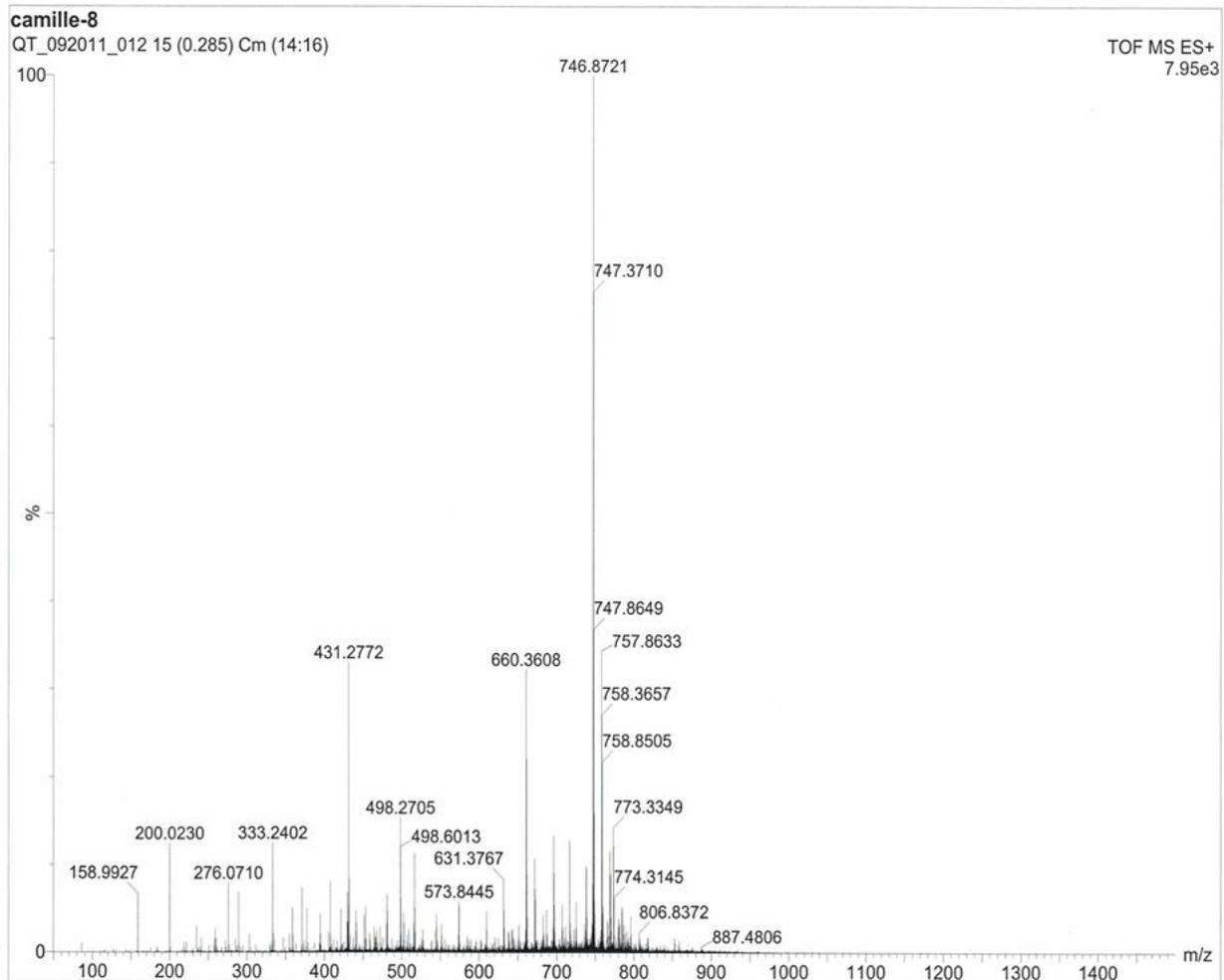


Figure 81: Mass Spectra of Ala Con-G-13 mer (sequence $\gamma\gamma\text{AA}\gamma\text{AA}\gamma\text{AAR}\gamma\text{K}$). Observed mass = 746.8721 which corresponds to the $[\text{M} + 2\text{H}]^{2+}$; Expected Mass = 1491.5961.

LIC_MALDI_02_27_12_02_D2 #1-55 RT: 0.00-0.52 AV: 55 NL: 1.01E6
T: ITMS + p MALDI Full ms [350.00-1200.00]

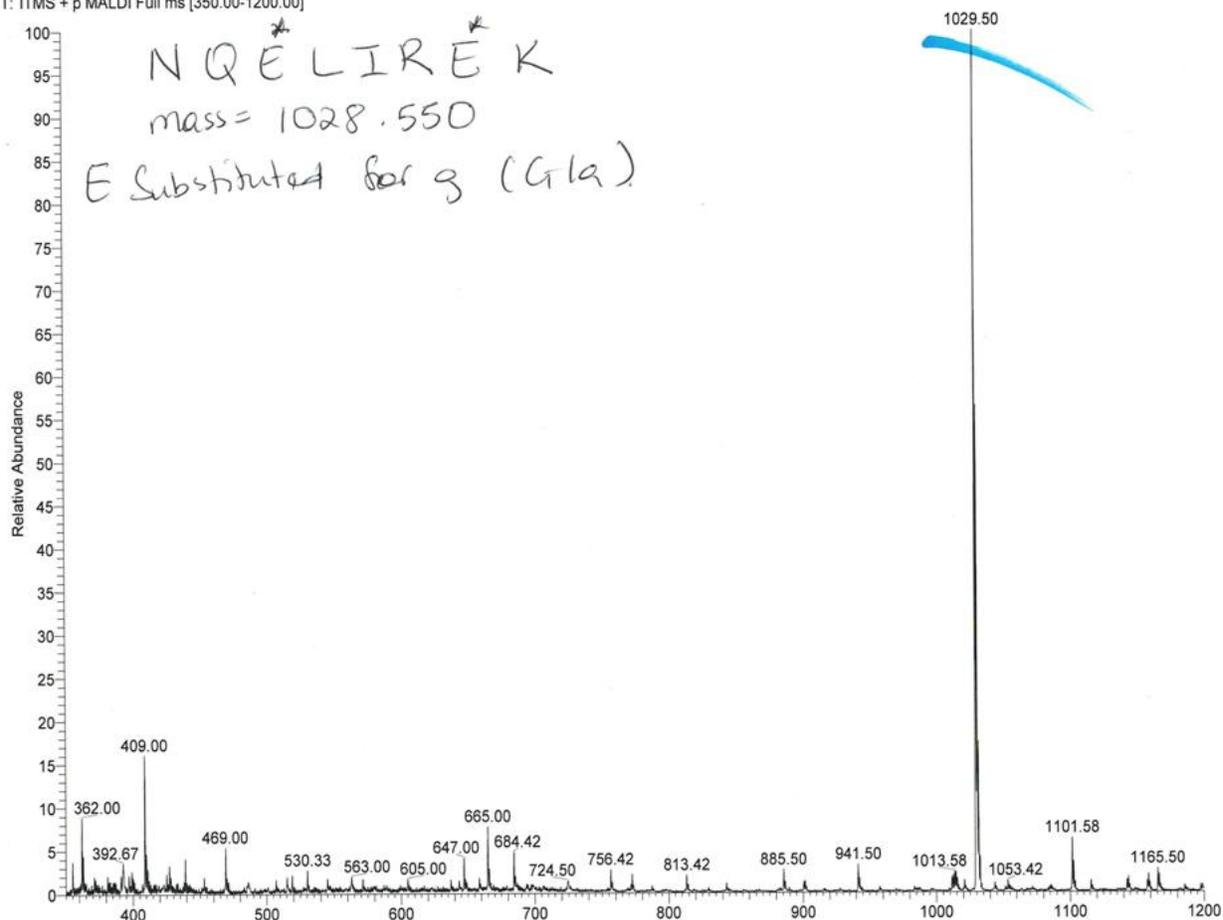


Figure 82: Mass Spectra of Con-G-8mer with Glutamate substituted for L-Gla (sequence NQELIREK). Observed mass $[M + H]^+ = 1029.50$; Expected Mass = 1028.550.

LiC_MALDI_02_27_12_04_D4 #1-53 RT: 0.00-0.51 AV: 53 NL: 1.54E5
T: ITMS + p MALDI Full ms [350.00-1200.00]

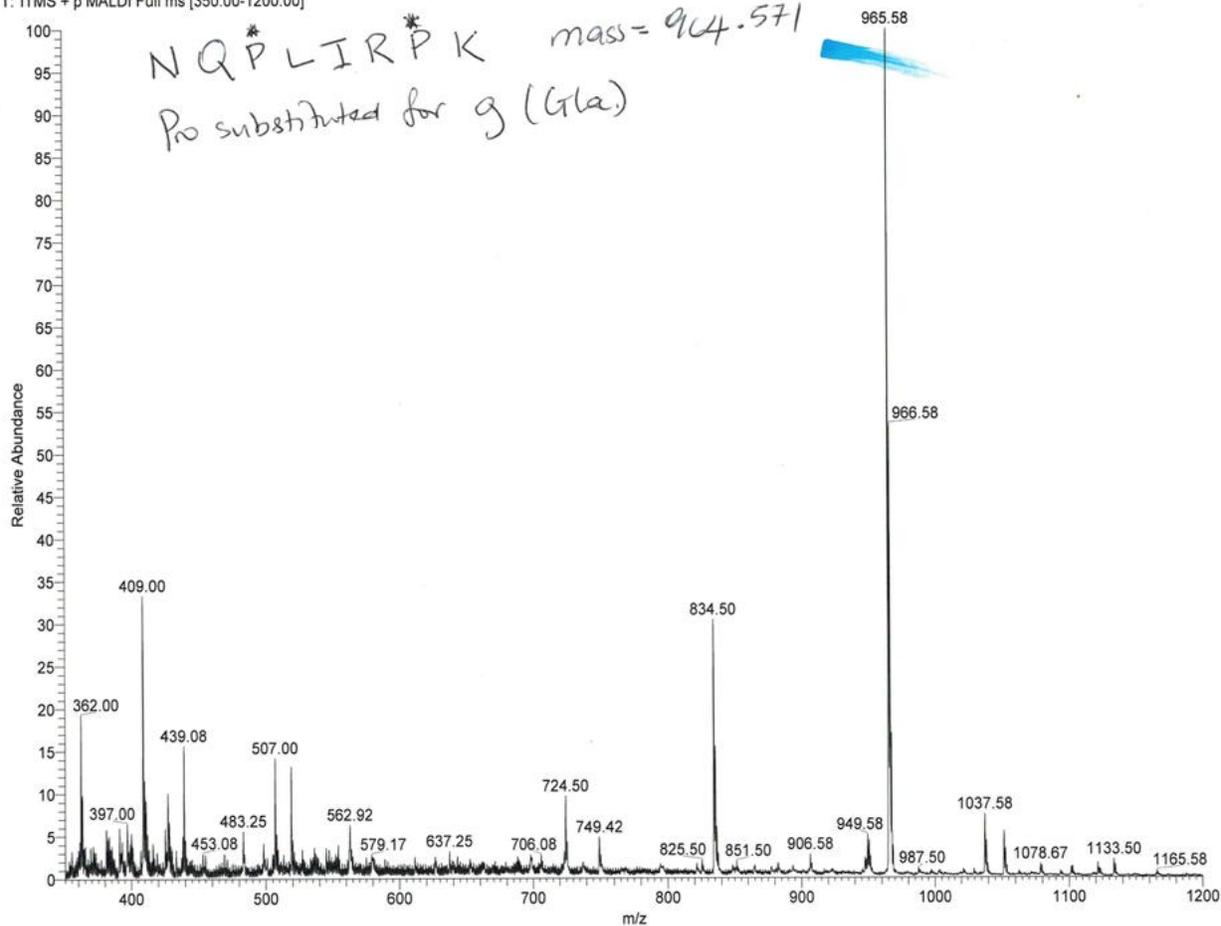


Figure 83: Mass Spectra of Con-G-8mer with Proline substituted for L-Gla (sequence NQPLIRPK). Observed Mass $[M + H]^+ = 965.58$; Expected Mass = 964.571.

LIC_MALDI_02_27_12_01_D1 #1-60 RT: 0.00-0.64 AV: 60 NL: 6.81E5
T: ITMS + p MALDI Full ms [350.00-1200.00]

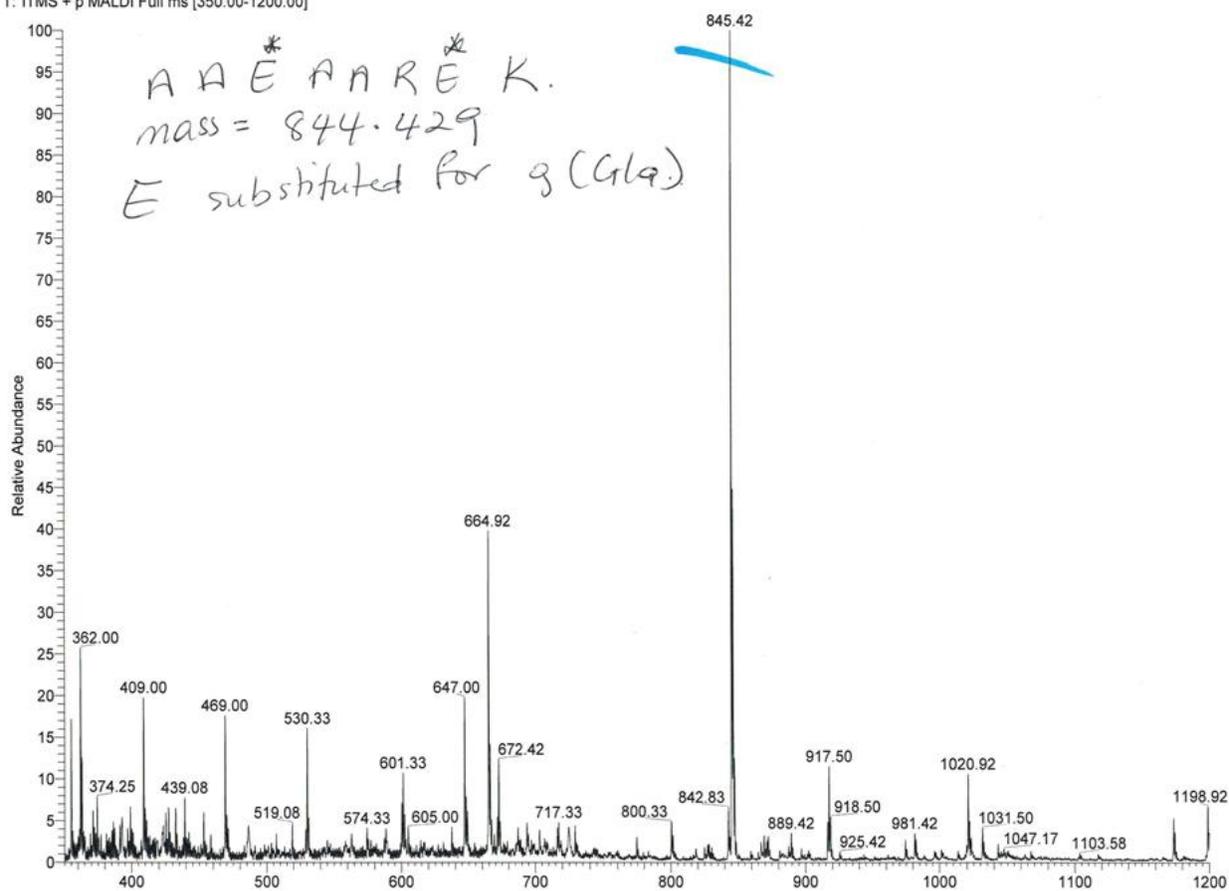


Figure 84: Mass Spectra of Ala-Con-G-8mer with Glutamate substituted for L-Gla (sequence AAEAAREK). Observed Mass $[M + H]^+ = 845.42$; Expected Mass = 844.429.

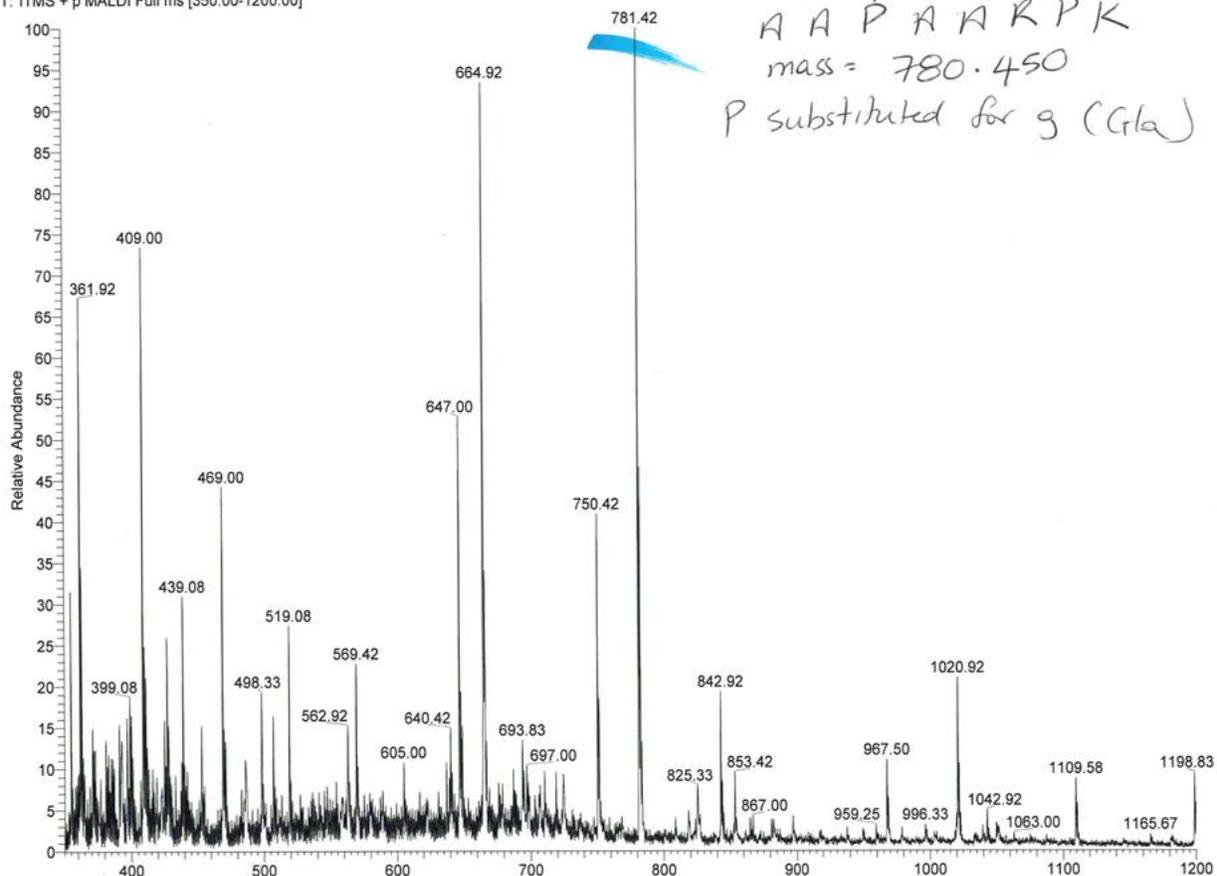
LIC_MALDI_02_27_12_03_D3 #1-55 RT: 0.00-0.64 AV: 55 NL: 9.60E4
T: ITMS + p MALDI Full ms [350.00-1200.00]

Figure 85: Mass Spectra of Ala-Con-G-8mer with Proline substituted for L-Gla (sequence AAPAARPK). Observed mass $[M+H]^+ = 781.420$; Expected Mass = 780.450.

IV.8.5 Expression of NR2A with Conantokin T (Con-T).

UV-vis spectra were recorded using a Cary 300 Bio WinUV, Varian spectrophotometer.

IV.8.6 Cloning

IV.8.6.1 Plasmid DNA Double Digestion Protocol

The samples of the plasmid with the gene (PUCMinus) and the plasmid into which the gene is to be inserted (pET22b(+), pET28(b)) were prepared using the amounts specified in the table below:

Table 28: Double Digestion Reaction solution.

| | |
|-------------------|------------|
| DNA (Plasmid) | Up to 1 µg |
| Buffer (10x) | 2 µl |
| Enzyme 1 (10U/µl) | 1 µl |
| Enzyme 2 (10U/µl) | 1 µl |
| H ₂ O | - |
| Total Volume | 20 µl |

http://www.methods.info/Methods/RNA_DNA/restr_analysis.html

The samples were then incubated for 3 hours at 37^oC. A DNA agarose gel of the samples was run after which the DNA was purified using the Wizard SV Gel and PCR Clean-Up System protocol.

IV.8.6.2 Plasmid DNA Ligation Protocol

The vector (pET22b(+)) (50 ng) was combined with three times the amount of insert (gene cut out from the PUCMinus vector). The volume was adjusted to 10 µl using dd H₂O. 10 µl of 2X Quick ligation buffer was added after which the mixture was mixed thoroughly. 1 µl of Quick ligase was then added and mixed. The mixture was then allowed to sit at room temperature for 30 mins. The DNA was then chilled on ice after which it was transformed into competent cells of XL1Blue for amplification.

IV.8.7 NR2A gene sequence

GGATCCGAATTTATGGGCCCGGATGATAACCATCTGAGCATTGTGACCCTGGAAGA
AGCGCCGTTTGTGATTGTGGAAGATATTGATCCGCTGACCGAAACCTGCGTGCGCAA
CACCGTGCCGTGCCGCAAATTTGTGAAAATTAACAACAGCACCAACGAAGGCATGA
ACGTGAAAAAATGCTGCAAAGGCTTTTGCATTGATATTCTGAAAAAACTGAGCCGC
ACCGTGAAATTTACCTATGATCTGTATCTGGTGACCAACGGCAAACATGGCAAAAA
AGTGAACAACGTGTGGAACGGCATGATTGGCGAAGTGGTGTATCAGCGCGCGGTGA
TGGCGGTGGGCAGCCTGACCATTAACGAAGAACGCAGCGAAGTGGTGGATTTTAGC
GTGCCGTTTGTGGAACCGGCATTAGCGTGATGGTGAGCCGCGGCACCCAGGTGAC
CGGCCTGAGCGATAAAAAATTCAGCGCCCGCATGATTATAGCCCGCCGTTTCGCTT
TGGCACCGTGCCGAACGGCAGCACCGAACGCAACATTCGCAACAACACTATCCGTATA
TGCATCAGTATATGACCCGCTTTAACCAGCGCGGCGTGGAAGATGCGCTGGTGAGCC
TGAAAACCGGCAAACCTGGAGCGTTTATTTATGATGCGGCGGTGCTGAACTATAAAG
CGGGCCGCGATGAAGGCTGCAAACCTGGTGACCATTGGCAGCGGCTATATTTTTGCGA
CCACCGGCTATGGCATTGCGCTGCAGAAAGGCAGCCCGTGGAACGCCAGATTGAT
CTGGCGCTGCTGCAGTTTGTGGGCGATGGCGAAATGGAAGAACTGGA AACCTGTG
GCTGACCGGCATTTGCCATAACTAAAAGCTT

IV.8.8 NR2A amino acid sequence

MGPDDN HLSIVTLEEA PFVIVEDIDP LTETCVRNTV PCRKFVKINN
STNEG M N V K K CCKGFCIDIL KKL SRTVKFT YDLYLVTNGK HGKKVNNVWN
GMIGE V V Y Q R A V M A V G S L T I N E E R S E V V D F S V P F V E T G I S V M V S R G T Q V T
GLSDKKFQRP HDYSPPFRFGTVPNGSTERNIRNNYPY M H Q Y M T R F N Q R G V
EDALVSLKTG KLDAFIYDAA VLNYKAGRDEGCKL VTIGSG YIFATTGYGI
ALQKGS PWKR QIDLALLQFV GDGEM E E L E T L W L T G I C H N * K L

IV.8.9 XL1-Blue competent cells preparation

XL1-Blue *E. coli* cells were streaked on agar plate with tetracycline as antibiotic (12.5 mg/L) and incubated overnight at 37 °C.

The following day a colony was picked from the plate using a sterile wooden applicator. The colony was inoculated in sterile LB solution (10 mL), with tetracycline (12.5 mg/L), in a shaker at 220 RPM overnight. The overnight culture (1 mL) was transferred to sterile LB solution (200 mL) with tetracycline (12.5 mg/L). the solution was kept shaking at 220 RPM at 37 °C for about 2 h or until the OD was between 0.4 and 0.8. At the same time, a 500 mL centrifuge bottle was sterilized with 20% bleach for 2 h and rinsed with sterile water 6 to 7 times before use.

The cells were harvested by centrifugation at 3500 RPM for 5 min. at 4 °C. The cells were resuspended with sterile sodium chloride solution (0.9 %, 100 mL). The cells were harvested by centrifugation at 3000 RPM for 4 min at 4 °C.

The following steps were then followed:

A: The cell pellet was resuspended with calcium chloride solution (100 mM, 15% glycerol v/v, 10 mL) on ice and was incubated for at least 30 min. The cells were fractioned in sterile eppendorf tubes in 100 µL portions and flash frozen with liquid nitrogen.

B: The cell pellet was resuspended with calcium chloride solution (100 mM, 50 mL) and incubated on ice for 30 min. The cells were spun down by centrifugation at 3000 RPM for 4 min and then resuspended with calcium chloride solution (100 mM, 15% glycerol v/v, 10 mL) on ice. The cells were fractioned in sterile eppendorf tubes in 100 µL portions and flash frozen with liquid nitrogen.

Note: The cells prepared using protocol A uses one step less than protocol B. This results in cells that are less competent. The cells that are used for the transformation of the PCR product was prepared using protocol B as the cells need to be highly competent. This is due to the fact that the DNA from the PCR reaction is nicked and does not go through the cell membrane with the same efficiency as circular DNA during the heat shock process. Therefore protocol B is a better protocol for competent cells preparation. This protocol was used to prepare the BL21 competent cells the only difference is chloramphenicol was used instead of tetracycline.

IV.8.10 Heat shock transformation

The DpnI digested PCR or ligation solution (5 μ L) was added to competent cells (50 or 100 μ L). The cells were then incubated on ice for 10 to 30 min, followed by heat shock at 42 °C for 50 sec. The cells were put back on ice immediately. Sterile LB solution (500 μ L) was added to the cells and shaken at 37 °C for 1 h. The cells were then harvested by centrifugation at 5000 RPM for 1 min. The cells were resuspended in LB (50 μ L) and then plated on an agar plate with the appropriate antibiotics.

IV.8.11 Mutagenesis

All mutations were made using *Stratagene's QuickChange Site-Directed Mutagenesis Protocol*.

Table 29: PCR Reaction solution.

| DNA template | Primer forward | Primer Reverse | Turbo DNA polymerase | dNTP (10mM of each) | Buffer (10x) | Water (dd) | Total volume |
|--------------|----------------|----------------|----------------------|---------------------|--------------|------------|--------------|
| 70 ng | 20 pmol | 20 pmol | 1 unit | 1 μ L | 5 μ L | / | 50 μ L |

Table 30: PCR temperature control cycles.

| | | |
|------|------------------------|--------|
| 1 x | 95 °C | 3 min |
| 20 x | 95 °C | 30 sec |
| | (T _m -4) °C | 1 min |
| | 72 °C | 10 min |
| 1 x | 72 °C | 10 min |
| 1 x | 25 °C | 5 min |

Note: T_m represents the melting temperature of the primer, which was calculated using the website shown below.

<http://www.promega.com/techserv/tools/biomath/calc11.htm>

IV.8.12 Primers used for mutagenesis:

Primers used for frame shift correction in pET22b (+)

Forward: 5'-CGGATCCGGAATTCATGGGC-3'

Reverse: 5'-GCCCATGAATTCCGGATCCG-3'

Primers used for pelB Leader sequence loop out in pET22b (+)

Forward: 5'-TAAATGATTGTCGTCAGGGCCCATGTATATCTCCTTC-3'

Reverse: 5'-GAAGGAGATATACATGGGCCCTGACGACAATCATTTA-3'

IV.8.13 Protein expression and purification in pET-22b(+) and in pET-28b systems

NR2A DNA was transformed into competent cells (BL21 (DE3) pLysS, Origami (DE3) pLysS and Tuner (DE3)) . A single colony was picked from the transformation plate and inoculated in

15 or 50 mL of LB with appropriate antibiotics at 37 °C overnight. The next day the growth was

transferred to 1 L of sterile LB media with the same concentration of antibiotics. The media was incubated at 37 °C until the OD₆₀₀ reached 0.8. Isopropyl-1-thio-D-galactopyranoside (IPTG, Gold Biotechnology) was added to final concentrations of 0.25, 0.5, 0.75 or 1.0mM to induce protein expression, and the cell culture was the shaken at 4 °C, RT ,16 °C, 20°C or 37 °C depending on the plasmid and cell line. The times of incubation varied from 3, 6, 8, 16 and 24 h.

Cells were harvested by centrifugation at 5000 RPM for 10 min in Beckmann J2-21M/E centrifuge and resuspended with 50 mL 20 mM imidazole HEPES buffer pH=8.0 with 0.5 % CHAPS and 250 mM NaCl), followed by 3 min sonication (3 time, 1 min per pulse, pulse = 60%, power 60%) to lyse the cells. DNase (20 µL, Roche, recombinant, 25 unit/µL) was added to the lysed cells and incubated for 30 min. The lysed cell solution was then spun down for 30 min at 5000 RPM. The supernatant was then purified via Ni-NTA purification protocol. He supernatant was allowed to bind to the Ni-NTA resin, which was cleaned and equilibrated with the resuspension buffer, for 30 min. The resin was then washed with the resuspension buffer and the protein was eluted using 250 mM imidazole HEPES buffer pH=8.0 with 0.5 % CHAPS and 250 mM NaCl as the elution buffer. UV-vis and SDS-PAGE was used to detect the fractions containing the expected protein. The fractions were then dialyzed overnight against HEPES buffer pH = 8.0 with 0.5 % CHAPS.

IV.8.14 Inclusion bodies purification

Pellet washing buffer: 20 mM imidazole, HEPES buffer pH=8.0 with 0.5%CHAPS and 250 mM NaCl)

Solubilizing buffer: 20 mM imidazole, 6M Urea HEPES buffer pH=8.0 with 0.5 % CHAPS and 250 mM NaCl)

Denaturing washing buffer: 20 mM imidazole, 6M Urea HEPES buffer pH=8.0 with 0.5 % CHAPS and 250 mM NaCl)

Denaturing elution buffer: 250 mM imidazole, 6M Urea HEPES buffer pH=8.0 with 0.5 % CHAPS and 250 mM NaCl)

Refolding buffer: 10 mM imidazole, HEPES buffer pH=8.0 with 0.5 % CHAPS and 250 mM NaCl)

Cells were harvested by centrifugation at 5000 RPM for 10 min in Beckmann J2-21M/E centrifuge and resuspended with 50 mL 20 mM imidazole HEPES buffer pH=8.0 with 0.5%CHAPS and 250 mM NaCl), followed by 3 min sonication (3 time, 1 min per pulse, pulse = 60%, power 60%) to lyse the cells. DNase (20 μ L, Roche, recombinant, 25 unit/ μ L) was added to the lysed cells and incubated for 30 min. The lysed cell solution was then spun down for 30 min at 5000 RPM. The pellet was then washed 2 or 3 times using the washing buffer. The pellet was then dissolved in the solubilizing buffer. The solubilized pellet was then spun down for 30 min at 8000 RPM and the supernatant was then applied to the denatured Ni-NTA resin. The protein was eluted using 250 mM imidazole, 6M Urea, HEPES buffer pH=8.0 with 0.5%CHAPS and 250 mM NaCl as the elution buffer. UV-vis and SDS-PAGE was used to detect the fractions containing the expected protein. The fractions were then dialyzed overnight against HEPES buffer pH=8.0 with 0.5%CHAPS.

REFERENCES

REFERENCES

1. Dowd, P.; Wook Ham, S.; Geib, S.J., Mechanism of Action of Vitamin K. *J.Am.Chem.Soc.*, **1991**, *113*, 7734 - 7743.
2. *Helvetica Chimica Acta*, **1977**, *60*, 804.
3. Schuerman, M.A.; Keverline, K. I.; Hiskey, R.G., Chiral Synthesis of L- γ -Carboxyglutamic Acid (L-Gla). *Tetrahedron Letters*, **1995**, *36*, 825 – 828.
4. Jiang, S.; Lai, C. C.; Kelley, J.A.; Roller, P.P., A Practical Synthesis of fully Protected L- γ -Carboxyglutamic Acid. *Tetrahedron Letters*, **2006**, *47*, 23 – 25.
5. Kolasa, T.; Miller, M. J., Reactions of Alpha-Hydroxy Carbonyl Compounds with Azodicarboxylates and Triphenylphosphine: Synthesis of Alpha-N-Hydroxy Amino Acid Derivatives. *J. Org. Chem.* **1987**, *52*, 4984.
6. Weilard, J., Pfister, K., Swanzey, E.F., Preparation of the Stereoisomeric α , β -Diphenyl- β -hydroxyethylamines. *J. Am. Chem. Soc.* **1951**, *73*, 1216.
7. John Wiley and Sons, Organic Synthesis, Collective Volume 4. p. 263-264.
8. Shelkov, R.; Nahmany, M.; Melman, A., Acylation through Ketene Intermediates. *J. Org. Chem.* **2002**, *67*, 8982.
9. Reyes, A.; Regla, I.; Fragnaso, M.C.; Vallejo, L.A.; Demare, P.; Jiménez-Vázquez, H.A.; Ramírez Y.; Juaristi, E.; Tamariz, J., Stereoselective Tandem Michael-Intramolecular Cyclization Approach to Functionalized Pyrroloisindolones. *Tetrahedron*, **1999**, *55*, 11187 – 11202.
10. Monache, G.D.; Di Giovanni, M. C.; Maggio, F.; Misiti, D.; Zappia, G., Convenient Preparation of *O*-Benzyl-*N*-Cbz-*D*-Serinol, An Intermediate in β -Hydroxy- α -amino Amino Acid Synthesis. *Synthesis*, **1995**, *09*, 1155.
11. Konas, D. W.; Pankuch, J. J.; Coward, J. K., The Synthesis of (2*S*)-4,4-Difluoroglutamyl γ -Peptides Based on Garner's Aldehyde and Fluoro-Reformatsky Chemistry. *Synthesis*, **2002**, *17*, 2616.
12. Enders, D.; Vicario, A.; Wolberg, M.; Müller, M., Asymmetric Total Synthesis of (-)-Callystatin A and (-)-20-*epi*-Callystatin A Employing Chemical and Biological Methods. *Chem. Eur. J.* **2002**, *08*, 4272 - 4284.
13. Jiang, S.; Lai, C. C.; Kelley, J.A.; Roller, P.P., A Practical Synthesis of fully Protected L- γ -Carboxyglutamic Acid. *Tetrahedron Letters*, **2006**, *47*, 23 – 2

14. Poon, K.W. C.; Lovell, K.M.; Dresner, K.N.; Datta, A., Trimethylsilyl Trifluoromethanesulfonate (TMSOTf) Assisted Facile Deprotection of *N*, *O*- Acetonides. *J. Org. Chem.* **2008**, *73*, 752 – 755.
15. Brovotto, M.; Seoane, G.; Stereoselective Synthesis of 3-Oxygenated-*cis*-dialkyl-2,5-substituted Tetrahydrofurans from Cyclohexadienediols. *J. Org. Chem.* **2008**, *73*, 5776 – 5785.
16. Pfrengle, F.; Dekaris, V.; Schefzig, L.; Zimmer, R.; Reissig, H-U., Indium Trichloride Mediated Cleavage of Acetonides in the Presence of Acid Labile Functional Groups – Enhancing the Synthetic Utility of 1,3-Dioxolanyl-Substituted 1,2-Oxazines. *Synlett*, **2008**, *19*, 2965 – 2968.
17. Avenozo, A.; Cativiela, C., Corzana, F.; Peregrina, J.M.; Zurbano, M.M., Preparation and Synthetic Applications of (*S*)- and (*R*)-*N*-Boc-*N*,*O*-isopropylidene- α -methylserinals: Asymmetric Synthesis of (*S*)- and (*R*)-2-Amino-2-methylbutanoic Acids (Iva). *J. Org. Chem.* **1999**, *64*, 8220 – 8225.
18. Mattsson, S.; Dahlstrom, M.; Karlsson, S., Mild Hydrolysis of Esters Mediated by Lithium Salts. *Tetrahedron Letters* **2007**, *48*, 2497- 2499.
19. Brown, J.E.; Klee, W.A., Helix-Coil Transition of the Isolated Amino Terminus of Ribonuclease. *Biochemistry*, **1971**, *10*, 470-476.
20. Ghadiri, M.Reza; Choi C., Secondary Structure Nucleation in Peptides. Transition Metal ion stabilized Alpha Helices. *J.Am.Chem.Soc.* **1990**, *112*, 1630-1632.
21. Reza Ghadiri, M., Katharina Fernholz, A., Peptide Architecture: Design of Stable Alpha Helical Metallopeptides via a Novel Exchange-Inert Ruthenium(III) Complex. *J.Am.Chem.Soc.* **1990**, *112*, 9633-9635.
22. Prorok, M.; Wander, S. E.; Blandl, T.; Castellino, F.J., Calcium Binding Properties of Synthetic γ -Carboxyglutamic Acid-Containing Marine Cone Snail “Sleeper” Peptides, Conantokin-G and Conantokin-T. *Biochemistry*, **1996**, *35*, 16528-16534.
23. Prorok, M.; Zajicek, J.; Blandl, T.; Castellino, F. J., Metal-Ion-Binding Properties of Synthetic Conantokin-G. *Biochemical Journal*, **1997**, *328*, 777-783.
24. Chen, Z.; Blandl, T.; Prorok, M.; Warder, S.E.; Li, L.; Zhu, Y.; Pederson, L. G.; Ni, F.; Castellino, F.J., Conformational Changes in Conantokin –G Induced upon Binding of Calcium and Magnesium as Revealed by NMR Structural Analysis. *Journal of Biological Chemistry*, **1998**, *273*, 16248-16258.

25. Prorok, M.; Dai, Q.; Castellino, F.J., A Single Amino Acid Replacement Results in the Ca^{2+} - Induced Self-Assembly of a Helical Conantokin-Based Peptide. *Biochemistry*, **2004**, *43*, 13225-13232.
26. Dai, Q.; Sheng, Z.; Geiger, J. H.; Castellino, F.J., Helix-Helix Interactions between Homo- and Heterodimeric γ -Carboxyglutamate-containing Conantokin Peptides and Their Derivatives. *Journal of biological Chemistry*, **2007**, *282*, 12641-12649.
27. Cnudde, S.E., Prorok, M., Dai, Q., Geiger, J.H., Castellino, F.J. The Crystal Structures of the Calcium-Bound Con-G and Con-T[K7 γ] Dimeric Peptides Demonstrate a Metal Dependent Helix-Forming Motif. *J. Am. Chem.Soc.* **2007**, *129*, 1586-1593.
28. Prorok, M.; Blandl, T.; Castellino, F.J., Structure-Function Relationships of the NMDA Receptor Antagonist Peptide, Conantokin-R. *FEBS Letters*, **2000**, *470*, 139-146.
29. Skolnick, P.; Boje, K.; Miller, R.; Pennington, M.; Maccacchini, M-L., Non-competitive Inhibition of *N*- methyl-*D*-Aspartate by Conantokin-G: Evidence for an Allosteric Interaction at Polyamine Sites. *Journal of Neurochemistry*, **1992**, *59*, 1516-1521.
30. Kendrick, S.J.; Lynch, D.R.; Pritchett, D.B., Characterization of Glutamate Binding Sites in Receptors Assembled from Transfected NMDA Receptor Subunits. *Journal of Neurochemistry*, **1996**, *66*, 608-616.
31. Prorok, M.; Wander, S. E. ; Blandl, T.; Castellino, F.J., Amino Acid Determinants for NMDA Receptor inhibition by Conantokin-T. *Journal of Neurochemistry*, **2001**, *77*, 812-822.
32. Klein, R.C.; Prorok, M.; Galdzicki, Z.; Castellino, F.J., The Amino Acid Residue at Sequence Position 5 in the Conantokin Peptides Partially Governs Subunit-Selective Antagonism of Recombinant *N*- methyl-*D*-Aspartate Receptors. *Journal of Biological Chemistry*, **2001**, *276*, 26860-26876.
33. Prorok, M.; Blandl, T.; Castellino, F.J., NMDA-receptor Antagonist Requirements in Conantokin-G. *FEBS Letters*, **1998**, *435*, 257-262.
34. Warder, S. E.; Prorok, M.; Chen, Z.; Li, L.; Zhu, Y.; Pedersen, L.G. Ni, F.; Castellino, F.J., The Roles of Individual γ -Carboxyglutamate residues in the Solution Structure and Cation-Dependent Properties of Conantokin-T. *Journal of Biological Chemistry*, **1998**, *273*, 7512-7522.
35. Teichert, R.W.; Jimenez, E.C.; Twede, V.; Watkins, M.; Hollmann, M.; Bulaj, G.; Olivera, B.M., Novel Conantokins from *Conus parius* Venom Are Specific Antagonists of *N*- methyl-*D*-Aspartate Receptors. *Journal of Biological Chemistry*, **2007**, *282*, 36905-36913.

36. Furukawa, H.; Gouaux, E., Mechanisms of Activation, Inhibition and Specificity: Crystal Structures of the NMDA Receptor NR1 Ligand-Binding Core. *The EMBO Journal*, **2003**, *22*, 2873 – 2885.
37. Furukawa, H.; Singh, S.K.; Mancusso, R.;Gouaux, E., Subunit Arrangement and Function in NMDA Receptors. *Nature*, **2005**, *438*, 185 – 192.