

STUDY OF PHOTO REACTION AT ATOMIC RESOLUTION IN A RHODOPSIN MIMIC
AND
STUDIES OF DOMAIN SWAPPING IN ILBP FAMILY MEMBERS

By

Nona Ehyaei

A DISSERTATION

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

Chemistry - Doctor of Philosophy

2020

ABSTRACT

STUDY OF PHOTO REACTION AT ATOMIC RESOLUTION IN A RHODOPSIN MIMIC AND STUDIES OF DOMAIN SWAPPING IN ILBP FAMILY MEMBERS

By

Nona Ehyaei

Signal transduction usually involves the binding of signaling molecules, ligands to the receptors. Herein, we will explore signal transduction in vision, a family of GPCR, and, more specifically, its rhodopsin subfamily, which plays a crucial role in color vision and sensing. There have been many mutagenesis studies done using Raman spectroscopy, crystallography, and NMR to understand the mechanism of the wavelength regulation and the photoisomerization of rhodopsin proteins. Moreover, recent studies have successfully indicated the intermediates of rhodopsin's photocycle, using time-resolved experiments, femtosecond x-ray laser (x-ray free-electron laser, XFEL), and cryokinetic data. All of these studies demonstrate the different important biophysical characteristics of rhodopsin systems; however, they have some limitations. Rhodopsins are membrane proteins, and their expression, purification, mutagenesis, and crystallization are very challenging. Also, these proteins evolve a lot during evolution. As a result of environmental changes and developments during evolution, many amino acid residues in rhodopsins become conserved residues; therefore, it is hard to illustrate every single residue's effect on wavelength tuning through mutagenesis studies. Therefore, we use Cellular Retinoic Acid Binding Protein II (CRABP II) and Cellular Retinol Binding Protein II (CRBP II) as mimics to study rhodopsin systems. Their solubility, small size, substantial binding pocket and ease of crystallization makes them a great candidate for our purpose. By using high resolution X-ray crystallography and spectroscopy, we were successful in mimicking wavelength tuning as well as photoisomerization

cycle in rhodopsin mimic templates. Another application of these templates is designing new fluorescent dyes. Many fluorophores have been designed based on hCRBP2 template in collaboration with Prof. Borhan's group to reach the FarRed-NearIR emission. We postulate mechanisms of these new fluorophores using our structural analysis.

During our studies on hCRBP2, my lab-mates characterized the domain swapped dimer as a folding product for this protein. In domain swapping, two or more monomers exchange an identical part of their structures to form a dimer or higher-order oligomer. Almost all of the studies on DSD hCRBP2 have been done through bacterial expression. To find out the physiological relevancy of this phenomenon, we tried to investigate the existence of the DSD form in mammalian expression. Also, Since the existence of domain swapping for hCRBP2 is likely to have physiological importance, we investigate the mechanism of domain swap dimerization in the other members of iLBP family. We characterized the Domain swapped dimer for WT-human fatty acid binding protein 5 (hFABP5) bound to palmitic acid as a natural product during the E. coli expression. The existence of Domain swapping in FABP5 as another member of the iLBP family is another reason that indicates the formation of DSD as a natural kinetic product during the folding process, which may indicate a common folding pathway for these two proteins.

Dedicated to my beloved family and James Geiger for his support.

ACKNOWLEDGEMENTS

First, I'd like to thank my advisor and my mentor, Prof. James H. Geiger. He showed me how to have the confidence to initiate new projects and work independently. He helped me to find my path through barriers and hurdles and to achieve my goals. He guided me like a great coach with my major decisions particularly with finding my future career. He is one of the nicest and kindest people I have ever seen in my whole life and I am forever grateful that I was lucky enough to know him as my advisor.

I also would like to thank my committee members for their great support. They truly guided me during my Ph.D. program. I am glad to have Prof. Heedeok Hong as my second reader. He helped me a lot during my graduate studies at Michigan State University. He was the first person who taught me about protein folding and piqued my interest in biophysical science. I am very thankful and honored to have the support of Prof. David Weliky. I was fortunate to attend his courses which were the most advantageous courses that I have had during my educational experience. Prof. Weliky, thank you so much for showing me how to think and participate in scientific discussions. I would like to thank Prof. Jian Hu for his supports during my seminars and for helping me to find my future position. I would also like to express my gratitude to all of my committee and my advisors for reading this dissertation.

I am also very thankful for knowing Prof. Babak Borhan. He always helped me to improve my research, and to find my way from the first time that I joined MSU. I want to thank Prof. Borhan and Dr. Chrysoula Vasileiou for their helpful ideas during our weekly meetings. I enjoyed our collaboration with Borhan's group especially, Soham Maity, Dr. Wei Sheng, Dr. Elizabeth Santos for letting me broaden my knowledge.

I also would like to thank Prof. William Henry and Dr. Stacy Hovde for their support and generosity. I have used many of the instruments and facilities in their laboratory. Furthermore, I have been learned many new things from Dr. Stacy Hovde and really enjoy working with her. I am very glad to know Prof. William Henry and have his support in better understanding of my research.

During my graduate studies, I have learned many new techniques from my lab mates. I have learned many new techniques from Dr. Hadi Nayebi Gavgani, and Dr. Zahra Assar-Nossoni. They always helped me to improve in my research and I am glad to know them. I also thankful for being able to work with Hadi Nayebi in the same lab. He helped me to have a better understanding in biochemistry and he was available to support me. Thank you Hadi for your everything. I also want to thank Dr. Alireza Ghanbarpour for his supports. I truly enjoyed having scientific discussions with him. I also want to thank our new member of our laboratory, Courtney Bingham, for being so nice and helping me during past two years. I would like to thank Dr. Rafida Nossoni for helping me in crystallography and being very understanding during hard times. Also, a big thank you to Matison Pawlowski, Joelle Eaves, Mustapha Akhdar, and Carlos Ferran-Heredia for helping me during the experiments in the past five years.

I am grateful for having the support of my dear friends especially, Shams Vahedi at MSU. Shams, thank you for supporting me throughout the difficult times in the past year.

Lastly, I would like to thank my loving parents Minoos and Reza, and my bright brothers, Danial and Dana. I am extremely appreciative to have their care and encouragement during each and every stage of my life. I have learned from them to be strong, rational, and respectful to myself and others. I could never ask for a better family.

TABLE OF CONTENTS

| | |
|---|-----------|
| LIST OF TABLES..... | x |
| LIST OF FIGURES..... | xii |
| KEY TO SYMBOLS AND ABBREVIATIONS..... | xxii |
| CHAPTER I: ROLE OF WATER MOLECULES IN WAVELENGTH TUNING OF HCRBP II | 1 |
| I-1 INTRODUCTION | 1 |
| I-1-1 Signal Transduction | 1 |
| I-1-2 Color Vision..... | 1 |
| I-1-3 Rods and Cones..... | 2 |
| I-2 RHODOPSIN | 4 |
| I-3 VISUAL PHOTOTRANSDUCTION CASCADE | 7 |
| I-4 WAVELENGTH TUNING IN RHODOPSINS | 10 |
| I-5 WHY RHODOPSIN MIMIC SYSTEMS | 21 |
| I-5-1 Using Intracellular Lipid Binding Proteins (ILBPs) as Rhodopsin mimics | 22 |
| I-5-2 Human cellular retinoic acid binding protein II (hCRABP II) | 24 |
| I-5-3 Human cellular retinol binding protein II (hCRBP II)..... | 25 |
| I-6 WAVELENGTH REGULATION IN HCRBP II | 26 |
| I-7 ROLE OF WATER MOLECULES IN WAVELENGTH REGULATION OF HCRBP II | 31 |
| I-7-1 Q108K; K40L as a reference | 31 |
| I-7-2 Introducing the negative counterions to the PSB region (Blue shifted spectra) | 33 |
| I-7-3 Removing water molecule 2 | 34 |
| I-7-4 Introducing the bulky hydrophobic residues in the ionone region | 36 |
| I-7-5 Mutating Gln4 (removing water molecule 1) | 38 |
| I-7-6 Removing the water network between Q38 and Q128 (Water 3 and 4)..... | 40 |
| I-8 CONCLUSION | 42 |
| I-9 EXPERIMENTAL | 42 |
| I-9-1 Site-Directed Mutagenesis | 42 |
| I-9-2 Material and Methods: Protein Expression and Purification (bacterial expression).... | 42 |
| I-9-3 Crystallization and refinement | 44 |
| I-9-4 Tips for growing crystals from saturated solution | 45 |
| REFERENCES | 53 |
| CHAPTER II: ISOMERIZATION IN HCRBP II | 60 |
| II-1 PHOTOISOMERIZATION CYCLE IN ANIMAL RHODOPSIN | 60 |
| II-2 PHOTOISOMERIZATION CYCLE IN MICROBIAL RHODOPSIN | 62 |
| II-3 PHOTOISOMERIZATION OF ALL TRANS TO 15-CIS IN HCRABP II | 64 |
| II-4 PHOTOISOMERIZATION OF THE 13-CIS TO ALL TRANS RETINAL IN HCRABP II (SIMILAR TO BACTERIAL RHODOPSIN) | 70 |
| II-5 ISOMERIZATION IS DETECTED IN HCRBP II | 72 |

| | |
|---|------------|
| II-5-1 Clues for isomerization in hCRBP _{II} | 72 |
| II-5-2 Photoisomerization on hCRBP _{II} system was detected in solution..... | 74 |
| II-5-3 pK _a measurements for M1:Q4A..... | 80 |
| II-5-4 Photoisomerization on hCRBP _{II} system was detected in crystals | 81 |
| II-5-5 Detecting the potential third intermediate in the isomerization process | 86 |
| II-5-6 Investigate Photoisomerization in another mutant of hCRBP _{II} : Q108K:K40L:T51V:T53C:R58W:Y19W:T29L:Q4R (M1;Q4R)..... | 91 |
| II-5-7 Investigate Photoisomerization in another mutant of hCRBP _{II} : (K40L:T51V:T53C:R58W:Y19W:T29L:Q4A:Q38L) | 93 |
| II-5-8 Photo switching experiment in the presence of Q4 in some different mutant of hCRBP _{II} | 96 |
| II-6 CONCLUSION | 96 |
| II-7 EXPERIMENTAL..... | 96 |
| II-7-1 Site directed mutagenesis, protein expression and purification | 96 |
| II-7-2 UV-vis Measurements | 97 |
| II-7-3 pK _a determinations | 97 |
| II-7-4 Crystallization..... | 97 |
| II-7-4-1 For growing crystals in the dark (Trans PSB)..... | 97 |
| II-7-4-2 For SB products..... | 97 |
| II-7-4-3 UV irradiated crystals (PSB product)..... | 98 |
| REFERENCES..... | 104 |
| CHAPTER III: STRUCTURE INSIGHTS FOR NEW FLUORESCENT PROTEIN TAGS..... | |
| III-1 FLUORESCENT PROTEINS | 108 |
| III-2 APPLICATION OF FLUORESCENT IMAGING..... | 110 |
| III-3 SYNTHETIC FLUOROCHROMES AND DYES..... | 114 |
| III-4 FR-1V DYE AND PHOTO SWITCHING MECHANISM | 119 |
| III-5 FR1V-α- CYANO | 130 |
| III-6 CONCLUSION | 133 |
| III-7 EXPERIMENTAL..... | 133 |
| III-7-1 Site directed mutagenesis and protein expression | 133 |
| III-7-2 Crystallization..... | 134 |
| REFERENCES..... | 142 |
| CHAPTER IV : DOMAIN SWAPPING IN HCRBP_{II} (IN VIVO STUDIES) | |
| IV-1 DOMAIN SWAPPING | 150 |
| IV-2 DOMAIN SWAPPING IN HCRBP_{II} | 151 |
| IV-3 INVESTIGATE THE FOLDING INTERMEDIATES OF HCRBP_{II} (USING CD AND FLUORESCENCE SPECTROSCOPY)..... | 154 |
| IV-4 IN VITRO REFOLDING EXPERIMENT TO INVESTIGATE THE EFFECT OF LIGAND BINDING ON THE FOLDING PATHWAY OF HCRBP_{II}..... | 156 |
| IV-5 STUDY THE OCCURRENCE OF DOMAIN SWAPPING BY MUTATIONAL ANALYSIS OF HCRBP_{II} (PHASE RELATIONSHIP)..... | 157 |
| IV-6 EXPRESSION OF DSD HCRBP_{II} IN MAMMALIAN CELLS | 162 |
| IV-6-1 Expression of hCRBP _{II} in Hela cells and detection using western blotting..... | 162 |

| | |
|---|------------|
| IV-6-2 Investigate the size of hCRBP II in mammalian expression..... | 167 |
| IV-7 DOMAIN SWAPPING IN RETINOID BINDING PROTEIN 7 AND FABP2 | 170 |
| IV-8 CONCLUSION..... | 171 |
| IV-9 EXPERIMENTAL | 172 |
| IV-9-1 Site directed mutagenesis..... | 172 |
| IV-9-2 Primers | 173 |
| IV-9-3 In vitro refolding..... | 174 |
| IV-9-4 Site directed mutagenesis (pCMV-Vector)..... | 176 |
| IV-9-5 Amino acid sequence of hCRBP II with flag tag in N terminus..... | 176 |
| IV-9-6 Mammalian expression | 176 |
| REFERENCES..... | 178 |

CHAPTER V: DOMAIN SWAPPING OF FATTY ACID-BINDING PROTEIN 5

| | |
|--|------------|
| (FABP5) | 182 |
| V-1 FATTY ACID-BINDING PROTEINS (FABP)..... | 182 |
| V-2 APO MONOMER STRUCTURE OF FABP5..... | 186 |
| V-3 DOMAIN SWAPPED DIMER STRUCTURES FOR FABP5 BOUND TO PALMITIC ACID..... | 188 |
| V-4 EFFORTS FOR GROWING APO DSD HFABP5 CRYSTALS..... | 189 |
| V-5 INTERACTION OF PALMITIC ACID WITH HFABP5..... | 191 |
| V-6 COMPARISON BETWEEN MONOMER AND DSD HFABP5 IN COMPLEX WITH PALMITIC ACID..... | 194 |
| V-7 CONFORMATIONAL CHANGE BETWEEN HFABP5 DSD BOUND TO DIFFERENT LIGANDS | 196 |
| V-8 DETERMINATION OF MELTING POINT OF FABP5 BY USING THERMAL SHIFT ASSAY (TSA) | 197 |
| V-9 STUDY THE OCCURRENCE OF DOMAIN SWAPPING BY MUTATIONAL ANALYSIS OF HUMAN FABP5 | 199 |
| V-10 INVESTIGATE OF THE DETERMINANTS OF DIMER/MONOMER RATIO DURING THE EXPRESSION OF WT HFABP5..... | 202 |
| V-11 HFABP5 AS A NEW RHODOPSIN MIMIC SYSTEM | 203 |
| V-12 CONCLUSION | 205 |
| V-13 EXPERIMENTAL..... | 206 |
| V-13-1 Site-Directed Mutagenesis | 206 |
| V-13-2 Amino acid sequence of FABP5:..... | 206 |
| V-13-3 Primers..... | 206 |
| V-13-4 Protein Expression and Purification of FABP5..... | 208 |
| V-13-5 Thermal shift assay (TSA) | 209 |
| V-13-6 Crystallization and refinement | 211 |
| V-13-6-1 Crystallization of Apo FABP5..... | 211 |
| V-13-6-2 Crystallization of FABP5 domain swapped dimer bound to palmitic acid (first structure)..... | 212 |
| V-13-6-3 Crystallization of FABP5 domain swapped dimer bound to palmitic acid (second structure)..... | 212 |
| REFERENCES..... | 216 |

LIST OF TABLES

| | |
|--|-----|
| Table I-1: Comparison between the absorption of KL, Q10K;T51D, and Q108K;K40L;L117E | 34 |
| Table I-2: Anion Exchange purification protocol for hCRBP _{II} , adjusted with 50 mM Tris, pH 8 buffer. The proteins elute between 4% -8 % 2M NaCl..... | 44 |
| Table I-3: X-ray crystallographic data and refinement statistics for KL; T53C hCRBP _{II} bound to retinal | 47 |
| Table I-4: X-ray crystallographic data and refinement statistics for KL; R58E hCRBP _{II} bound to retinal | 48 |
| Table I-5: X-ray crystallographic data and refinement statistics for KL; T51V hCRBP _{II} bound to retinal | 49 |
| Table I-6: X-ray crystallographic data and refinement statistics for KL; T51V;53C; R58W; T29L; Y19W; Q4A hCRBP _{II} bound to all trans retinal incubation in the dark | 50 |
| Table I-7: X-ray crystallographic data and refinement statistics for KL; T51V;53C; R58W; T29L; Y19W; Q4A hCRBP _{II} bound to cis-retinal after UV irradiation. | 51 |
| Table II-1: X-ray crystallographic data and refinement statistics for M1: Q4A mutant bound to all trans retinal in absence of visible light. | 99 |
| Table II- 2: X-ray crystallographic data and refinement statistics for M1;Q4A bound to all trans retinal in presence of visible light. | 100 |
| Table II-3: X-ray crystallographic data and refinement statistics for M1:Q4A bound to 15 cis retinal after UV exposure..... | 101 |
| Table II-4: X-ray crystallographic data and refinement statistics for M1;Q4A;Q38L bound to all trans retinal..... | 102 |
| Table II-5: X-ray crystallographic data and refinement statistics for M1;Q4A;Q38L bound to all trans retinal after UV exposure. | 103 |
| Table III-1: Several mutants of hCRBP _{II} that photo switch with this FR-1v by UV exposure. | 121 |
| Table III-2: X-ray crystallographic data and refinement statistics for m1;q4a;q38l bound to fr-1v | 135 |
| Table III-3: X-ray crystallographic data and refinement statistics for M1:Q4A:Q38L bound to fr-1v after UV exposure. | 136 |

| | |
|---|-----|
| Table III-4: X-ray crystallographic data and refinement statistics for KLVS bound to FR-1v.. | 137 |
| Table III-5: X-ray crystallographic data and refinement statistics for KLVS bound to FR-1v.. | 138 |
| Table III-6: X-ray crystallographic data and refinement statistics for Q108K:K40E:T51C:T53S:R58L:Q38F:Q4F bound to FR-1cyno..... | 139 |
| Table III-7: X-ray crystallographic data and refinement statistics for Q108K:K40L:T51V:T53S:R58H:Y19W:A33Y bound to Thiophenol. | 140 |
| Table III-8: X-ray crystallographic data and refinement statistics for Q108K:K40L:T51V:T53S:R58H:F16Y:Y19W:A33H:L117C bound to thiophenol. | 141 |
| Table IV-1: PCR protocol for mutagenesis | 172 |
| Table V-1: Crystallographic data of FABP5 (pseudo monomer). | 213 |
| Table V-2: Crystallographic data of FABP5 domain swapped dimer binds to palmitic acid (first structure). | 214 |
| Table V- 3: Crystallographic data of FABP5 domain swapped dimer binds to palmitic acid.... | 215 |

LIST OF FIGURES

| | |
|---|----|
| Figure I-1: Structure of human eye..... | 2 |
| Figure I-2: Schematic structure of Rod and Cone cells..... | 3 |
| Figure I-3: A) all-trans retinal B) retinal bound to bacteriorhodopsin..... | 4 |
| Figure I-4: Biological functions of microbial rhodopsins as ion pumps, Ion channels and sensors. ¹¹ | 4 |
| Figure I-5: Retinal forms by the Cleavage of β -carotene. | 6 |
| Figure I-6: Retinal makes Schiff base with lysine residue in rhodopsin. | 6 |
| Figure I-7: Isomerization of retinal in A) bacteriorhodopsin B) Microbial rhodopsin..... | 7 |
| Figure I-8: Shows the three subunits of transducin and the dissociation..... | 8 |
| Figure I-9: Structure of the Rhodopsin-GT complex. ²³ | 9 |
| Figure I-10: Schematic structure of Rhodopsin-GT complex ²³ | 9 |
| Figure I-11: Absorption of the protein depends on the position of the negative charge or negative dipole moments..... | 11 |
| Figure I-12: Electrostatic potential calculations for sensory rhodopsin II (SRII) 486 nm, bacteriorhodopsin (BR) 552 nm, and halorhodopsin (HR) 576 nm..... | 12 |
| Figure I-13: Electrostatic potential calculations for Rhodopsin, blue, green and red opsin..... | 13 |
| Figure I-14: Overlaid structures of the BR (CYAN - PDB ID: 1C3W) and SRII (Pink- PDB ID: 1JGJ) | 14 |
| Figure I-15: The water-mediated hydrogen networks in rhodopsin A) Ground state B) activated state. ^{42 4041} | 14 |
| Figure I-16: Wat 2a and Wat2b in the vicinity of the retinal in rhodopsin (PDB: 1L9H)..... | 15 |
| Figure I-17: Glu181 plays as counterion in metaII rhodopsin (3PQR). | 16 |
| Figure I-18: Difference in electron density maps A) 16ns and B)290ns after activation, obtained by time- resolved room temperature crystallography. (positive density Blue and negative density Gold). Even in 16ns, there is gold negative density for water ⁴⁰²⁴⁰ | 17 |

| | |
|--|----|
| Figure I-19: Water network in binding site of jumping spider rhodopsin 1. | 18 |
| Figure I-20: Comparisons of the retinal binding pockets between mono bovine rhodopsin (1GZM) and bistable Jumping spider rhodopsin 1 (6I9K) and squid rhodopsins(2Z93). | 18 |
| Figure I-21: The gas phase absorption of retinal n-butyl iminium PSB | 19 |
| Figure I-22: β -ionone compares with the α -ionone and β -damascon structures..... | 20 |
| Figure I-23: Structures and UV-vis absorption spectra of ATR and all-trans-retinal (3,4-didehydro-retinal), and their analogue. | 20 |
| Figure I-24: The evolution of rhodopsins. ⁶⁷ | 22 |
| Figure I-25: Structure of A-FABP5 a member of iLBP (PDB:3FR4). 10 beta strand with two alpha helixes as a lid. | 23 |
| Figure I-26: Structure of hCRABPII bound to retinal (PDB:2FR3)..... | 24 |
| Figure I-27: Reengineered hCRABPII protein retinal pigments ranging from 482nm to 630nm. ⁸⁰ | 25 |
| Figure I-28: Q108K; K40L (KL) hCRBPII structure bound to retinal through lysine residue (PDB: 4EXZ). ⁸⁷ | 26 |
| Figure I-29: Reengineered hCRBPII protein retinal pigments ranging from 425nm to 644nm... .. | 27 |
| Figure I-30: The three regions of the retinal in wavelength tuning study. ⁸⁷ | 27 |
| Figure I-31: Shows important residues for wavelength tuning in the hCRBPII system..... | 28 |
| Figure I-32: The overlaid structure of KL:T51V:T53C:R58W:Y19W:T29L (cyan, 591 nm) and of KL:T51V:T53C:R58W:Y19W:T29L:Q4R (green, 622 nm)..... | 29 |
| Figure I-33: Demonstrates electrostatic charge distribution on the surface of the retinal for Q108K:K40L mutant of hCRBPII on the left and Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R on the right. ⁸⁷ | 30 |
| Figure I-34: Absorption spectra for hCRBPII mutants bound to retinal. ^{87 88} | 30 |
| Figure I-35: Water molecules and residues around the retinal (5 angstrom) that have effect on wavelengthtuning (PDB:4RUU). ⁸⁰ | 31 |
| Figure I-36: KL structure (PDB:4RUU) A) mol A B)mol B. The relative orientation of Schiff base is different in the two molecules in the asymmetric unit..... | 32 |

| | |
|--|----|
| Figure I-37: Schiff base region | 33 |
| Figure I-38: Overlaid structures of KL (4RUU) with KL; T51D (6E5S). Most of the structures are similar, except the T51D is 1 angstrom closer to the PSB region. | 34 |
| Figure I-39: A) water network between Thr53 and Thr51 in KL (PDB:4RUU). B) KLT51V (PDB:5FAZ), removal of water molecule. C) KLT53C, there is no change in water molecule .. | 35 |
| Figure I-40: Ionone region | 36 |
| Figure I-41: KL;R58E. Glu58 is immersed through the solvent. | 36 |
| Figure I-42: KLT53C. There are two water molecules near to the ionone ring. | 37 |
| Figure I-43: Overlaid structure of KLT53C and KL. Removal of water molecules near to the ionone ring by addition of R58W mutation. | 37 |
| Figure I-44: Shows hydrophobic interactions with the nitrogen in all trans retinal bound to KLVCWLWQ4A hCRBP II. | 39 |
| Figure I-45: 15-cis retinal bound to KLVCWLWQ4A hCRBP II. Positive charge on cis iminium stabilized through pi-cation interaction with Trp106..... | 40 |
| Figure I-46: Hydrogen bonding between Gln38 and Gln128 and two water molecules in KL structure (PDB:4RUU)..... | 41 |
| Figure I-47: Gln38-Gln128 Water molecules in the two chains of crystal structure of KL:T51V:R58Y:Y19W:Q38L (PDB:..... | 41 |
| Figure I-48: hCRBP II bound to retinal A) KL; T51V;53C; R58W; T29L; Y19W; Q4A B)KL; T51V; T53C. hCRBP II bound to FR-1 C) KL; T51V: T53S D) KL;T51V; T53S; R58W; T29L;Q38L;Q128L;Y19W..... | 46 |
| Figure II-1: The photocycle in Bovine rhodopsin. The absorption wavelengths and lifetimes of the intermediates are given. ¹³ | 61 |
| Figure II-2: Trp265 changes its position from ground state (Green) and excited state (Cyan) as retinal isomerizes in rhodopsin. | 62 |
| Figure II-3: Photocycle in microbial rhodopsin..... | 63 |
| Figure II-4:Time-resolved serial crystallography resolves ultrafast atomic motions of retinal and shows the rotation of the all trans conformation to 13-cis retinal over the course of a few hundred femtoseconds in bacteriorhodopsin. ¹⁵ | 64 |
| Figure II-5: Conversion of the PSB to SB overtime in M1 hCRABP II. ^{7 25} | 66 |

| | |
|--|----|
| Figure II-6: Overlaid structures of 15-cis PSB and all-trans SB in R111K:R134F:T54V:R132Q:P39Y: R59Y (M2) hCRABPII. ^{7 25} | 67 |
| Figure II-7: Overlaid structures of 15-cis PSB (pink) and all-trans SB (green) in R111K:R134F:T54V:R132Q:P39Y: R59Y (M2) hCRABPII. The trans isomer stabilized through the hydrophobic interactions. Cis iminium stabilized through the π -cation interaction with Trp108. ⁷ | 68 |
| Figure II-8: Reversible photocycle for R111K:R134F:T54V:R132Q:P39Y: R59Y (M2) hCRABPII..... | 69 |
| Figure II-9: Isomerization of hCRABPII system in crystal. ⁷ | 69 |
| Figure II-10: Isomerization of 15cis retinal to all trans retinal in hCRABPII system..... | 70 |
| Figure II-11: Photoisomerization of 13-cis retinylidene to all trans retinylidene in microbial rhodopsin..... | 71 |
| Figure II-12: Photoisomerization of 13-cis retinylidene to all trans retinylidene in hCRABPII system. | 72 |
| Figure II-13: Positive charge on cis iminium stabilized through the water network with Gln4 and π -cation interaction with Trp106 (PDB:4RUU). | 73 |
| Figure II-14: The overlaid structure of KL:T51V:T53C:R58W:Y19W:T29L (cyan, 591 nm) and of KL:T51V:T53C:R58W:Y19W:T29L:Q4R (green, 622 nm)..... | 74 |
| Figure II-15: UV measurement for M1;Q4A after incubation with retinal for 30min. The measurement have been done at pH 7.4. Maximum PSB in this pH (gray line). PSB is not stable and it converts back to SB (360nm) over time in presence of the visible light (or it can happen thermally) (green line). | 75 |
| Figure II-16: Formation of PSB after bringing the solution from pH8 down to pH4 in the absence of visible light. | 76 |
| Figure II-17: Thermal conversion of the PSB to SB overtime in pH4. | 77 |
| Figure II-18: Conversion of the PSB to SB in presence of the visible light. Most of the PSB converts to SB after around 2 hours..... | 78 |
| Figure II-19: Conversion of the SB to PSB by shining 30s UV exposure..... | 79 |
| Figure II-20: Reversible photoswitching system for M1; Q4A hCRBPII system. By UV exposure the SB converts to PSB (orange line). Visible light for 20min converts the PSB back to SB..... | 79 |

| | |
|--|----|
| Figure II-21: Longer exposure of the solution to the UV light lead to quenching the chromophore. | 80 |
| Figure II-22: A) pK _a measurement for the chromophore after 30min incubation of the protein. pK _a is around 7. B) pK _a measurement after 24 hours incubation of the solution in pH 4.5. pK _a is measured around 3.3. | 81 |
| Figure II-23: UV exposure for 2min on M1;Q4A crystal. Crystal changed the color from the colorless to blue. | 82 |
| Figure II-24: Conversion of the all trans retinal to 15-cis retinal in a single crystal of the M1;Q4A. the overlaid structure of the trans and cis isoform demonstrate the change in the position of the Gln38. | 83 |
| Figure II-25: A) Hydrophobic interaction of the nitrogen with surrounding residues in all trans retinal bound to M1;Q4A B) π-cation interaction of the cis iminium in M1;Q4A. | 84 |
| Figure II-26: Isomerization of the all-trans retinal to 15 cis iminium in M1;Q4A hCRBP II. | 85 |
| Figure II-27: Photocycle in M1;Q4A crystals. | 85 |
| Figure II-28: Detecting the two intermediates in the isomerization process of M1;Q4A. | 86 |
| Figure II-29: Crystal of M1;Q4A changes the color from blue to colorless by visible light. | 86 |
| Figure II-30: Removing the water molecules between Gln38 and chromophore by visible light exposure. | 87 |
| Figure II-31: Formation of the blue crystal in the dark might be related to trans iminium intermediate. | 88 |
| Figure II-32: Overlaid structures of the M1;Q4A (green), with M1;Q4A;Q38L (blue). The only difference is removing the two water molecules near to Gln38. | 89 |
| Figure II-33: pK _a measurement for chromophore in M1;Q4A;Q38. solution were bring down to pH 1.4. Lowering the pH lead to the aggregation of the protein. From the change in the absorption vs pH graph, it is clear that the pK _a is lower than 3.3. | 89 |
| Figure II-34: Three intermediates detected for isomerization of the 15cis to all trans retinal in M1;Q4A construct. | 90 |
| Figure II-35: Formation of PSB after bringing the pH down to 4 for M1;Q4R construct. Maximum PSB formed after 46min (dark green line) pH4. | 91 |

| | |
|---|-----|
| Figure II-36: Shinning UV light (orange, gray and yellow line) after reaching the maximum PSB in the pH4 (Blue line) in M1;Q4R hCRBPPII construct. The SB converts to PSB by UV exposure.(pH4)..... | 92 |
| Figure II-37: Visible light exposure to the solution converts the PSB to SB in M1;Q4R hCRBPPII (pH4)..... | 92 |
| Figure II-38: UV exposure to the solution of M1;Q4A;Q38L convert the SB to PSB. Measurements are in pH4. | 93 |
| Figure II-39: Converting the PSB back to SB by visible light. (pH4)..... | 94 |
| Figure II-40: UV exposure on the M1;Q4A;Q38L crystal. | 94 |
| Figure II-41: Crystal structure of M1;Q4A;Q38L A) before UV exposure B)after UV exposure. Both structures have the all trans retinal conformation. | 95 |
| Figure III-1: GFP structure (pdb:1ema)..... | 109 |
| Figure III-2: Shows some of the fluorescent proteins with different emission wavelengths. ²⁶ . | 110 |
| Figure III-3: a)intermolecular FRET b)intramolecular FRET for detecting conformational changes c) cameleon intramolecular FRET probe d) intramolecular phosphorylation-sensitive fret probes copied from ref. ⁴⁸ | 112 |
| Figure III-4 a) single SLN identification by intraoperative fluorescence imaging b) detection of lymph node by multispectral FMI system. ⁵⁰ | 113 |
| Figure III-5: crystal structure of BphPs bound to BY (pdb:3c2w). ⁶¹ | 115 |
| Figure III-6: Demonstrates the reaction of snap-tag, and Clip-tag.poi (protein of interest). ⁷⁰ .. | 116 |
| Figure III-7: Malachite green (mg)..... | 117 |
| Figure III-8: Photoactive yellow protein tags (PYPs) bound to 4- hydroxycinnamic acid (pdb:2phy). ⁷⁹ | 118 |
| Figure III-9: Merocyanine retinal aldehyde (MCRA) bound to hCRABPII. ⁸⁴ | 119 |
| Figure III-10: FR-1v bound to Lys108 | 120 |
| Figure III-11: Absorption and emission for PS4 before (left) and after UV exposure (right). Absorption in dash line and emission solid line. ^{70 93} | 121 |
| Figure III-12: Postulated mechanism for photoswitching in FR-1V bound to hCRBPPII. ⁷⁰ | 122 |

| | |
|---|-----|
| Figure III-13: UV exposure on Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v crystals lead to change in color of crystals. | 123 |
| Figure III-14: Crystals of Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v before (left) and after UV exposure (left)..... | 123 |
| Figure III-15: Overlaid crystals of Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v before (blue) and after UV exposure (yellow). The only difference is changing the conformation of Trp58. | 124 |
| Figure III-16: Absorption spectra for Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1V in pH7.4. SB converts to PSB via UV exposure (left). PSB converts back to SB overtime thermally (right)..... | 125 |
| Figure III-17: Acidifying solution of Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v leads to conversion to the SB (λ_{MAX} 380nm) to 320-340nm peak. The pka for this form measured to be around 3.5..... | 126 |
| Figure III-18: Absorption spectra for Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v in pH4. SB and 320nm-340nm peaks converts to PSB via UV exposure (left). PSB converts back to mostly to 340-320nm peak overtime thermally (right). | 127 |
| Figure III-19: Absorption spectra for Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v in pH3.2. 320nm-340nm peaks converts to PSB via UV exposure (left). PSB converts back to 340-320nm peak in around 5min thermally (right). | 128 |
| Figure III-20: Conversion of SB to PSB via UV irradiation. A) for Q108K;K40L;T51V;T51C B) Q108K;K40L;T51V;T51S bound to FR-1V in pH4. | 129 |
| Figure III-21: Crystals of Q108K;K40L;T51V;T53C before and after UV exposure. Density for the chromophore is not sufficient to predict the mechanism of photo switching. | 130 |
| Figure III-22: FR1V- α - cyano structure. | 130 |
| Figure III-23: UV elimination breaks the sulfide bond between the Cys51 and the chromophore. | 131 |
| Figure III-24: UV exposure to the 330nm peak converts it to 445nm (left). The 445nm peaks converts back to 330nm in the presence of visible light (right)..... | 131 |
| Figure III-25: Electron density for chromophore in Q108K;K40E;T51C;T53S;R58L;Q38F;Q4F construct (green). | 132 |
| Figure III-26: Electron density for FR1V- α - cyano in Q108K;K40E;T51C;T53S;R58L;Q38F;Q4F construct. | 133 |

| | |
|---|-----|
| Figure IV-1: Demonstrates the concept of domain swapping..... | 151 |
| Figure IV-2: Structure of domain swapped dimer for hCRBP II..... | 152 |
| Figure IV-3: Proposed folding pathway for hCRBP II. Open monomer is as an intermediate in the folding (PDB:4ZH6). Domain swapped dimer and monomer shows as folding products. | 153 |
| Figure IV-4:Unfolding spectra of the WT hCRBP II in presence of different concentration of Gu-HCl. a and b) monitoring by CD measurements. c and d) Tryptophan fluorescent spectroscopy to study the intermediates in folding pathway of hCRBP II. Purple line in b and d demonstrates the third degree polynomial fit..... | 155 |
| Figure IV-5: Chromatogram of size exclusion chromatography after conducting the refolding experiment in high concentration. The left figure is the chromatogram for refolding in the presence of ligand and the right one is the result of refolding in the absence of ligand. | 157 |
| Figure IV-6: The "phase problem" in domain swap dimerization. Top figure is for monomer hCRBP II. In the bottom of figure, the ideal and actual β strand in DSD form elucidated. | 158 |
| Figure IV-7: Residues 56-63 in the B subunit of Y60W-hCRBP II..... | 159 |
| Figure IV-8: Demonstrates the conformation of residues in different hCRBP II construct. O: out of the binding pocket, I: inside the binding pocket, S: toward the solvent..... | 160 |
| Figure IV-9: a) SEC chromatogram for hCRBP II N59L. b) SDS PAGE from fraction 14-22 of SEC. | 161 |
| Figure IV-10:Chromatogram of Source Q Size Exclusion chromatography of a) R58L hCRBP II b) R58Q hCRBP II..... | 161 |
| Figure IV-11: Multiphoton fluorescence image of HeLa cells with cytoskeletal microtubules (magenta) and DNA (cyan)..... | 163 |
| Figure IV-12: Schematic illustration of steps in western blotting..... | 164 |
| Figure IV-13: 6-well culture plate used for expression of hCRBP II. There is the same amount of HeLa cells in each well. | 165 |
| Figure IV-14: Exposed x-ray film from the western blotting. Row 1, 2 and 3 extract from controls (cells without changing media, cells with changing the media in transfection). Row 4, 5, 6 extract from transfection with hCRBP II plasmid (transfection with 350 ng of extract). | 166 |
| Figure IV-15: Exposed x-ray after western blotting. Sample 1 and 2 are from 8 ng extract of control and 8ng extract of empty pCMV-flag tag vector. Sample 3,4 and 5 are from 8 ng,4 ng and 2 ng extract of infection with 750ng plasmid, respectively. | 167 |

| | |
|---|-----|
| Figure IV-16: Chromatogram of size exclusion chromatography. 1.5mg of domain swapped dimer hCRBP2 (Q108K:T51D mutant) ran through the superdex 75. | 168 |
| Figure IV-17: Chromatogram of size exclusion chromatography (superdex 75) for 0.5 mammalian expression extract..... | 169 |
| Figure IV-18: X-ray film from western blotting for extract before running the SEC column and fraction 14- 20 after SEC column. | 170 |
| Figure IV-19: Amino acid sequence alignment for hCRBP2 and FABP2. The red box shows the amino acids in the hinge loop and its neighbor residues important in domain swapping. | 171 |
| Figure IV-20: Dialysis cassettes uses for refolding experiments..... | 175 |
| Figure V-1: Crystal structure of FABP4 (PDB;4NNS). | 182 |
| Figure V-2: Structure of human hFABP5 in complex with Linoleic acid (PDB: 4LKT)..... | 183 |
| Figure V-3: Structure of human hFABP5 in complex with palmitic acid (PDB: 1b56)..... | 183 |
| Figure V-4: Structure of domain swapped dimer for FABP5 bound to 2-AG(4AZM).Two chains are in pink and green color..... | 184 |
| Figure V-5: Molecular structure of AEA and 2-AG..... | 185 |
| Figure V-6: Overlaid structures of FABP5 bound to 2-AG(green), and FABP5 bound to AEA(orange)..... | 185 |
| Figure V-7: a) chromatogram of size exclusion chromatography. b) SDS PAGE of fractions 15 to 22. Fractions 18, 19, 20 were mixed for putting the crystallization box. | 186 |
| Figure V-8: The overlaid structure of our apo monomer FABP5 in green (chain A green and chain B pink) vs reported apo monomer FABP5 in gray(PDB:4LKP). | 187 |
| Figure V-9: a) SEC for FABP5 b) crystal structure of FABP5 domain swapped dimer with AEA as a ligand in extra electron density..... | 188 |
| Figure V-10: Structure of domain swapped dimer hFABP5 bound to palmitic acid..... | 189 |
| Figure V-11: Crystal structure of FABP5 with palmitic acid in the binding pocket. | 190 |
| Figure V-12: Overlay of two FABP5 domain swapped structures bound to palmitic acid (first structure is in cyan and the most recent one is in orange). The arrows point to the hinge loops. | 191 |
| Figure V-13: Interaction of palmitic acid in the binding pocket of hFABP5. | 192 |

| | |
|--|-----|
| Figure V-14: Interaction of hFABP5 with the carboxylic group of palmitic acid..... | 192 |
| Figure V-15: SEC chromatogram for WT hFABP5 before (Orange graph) and after refolding (Blue graph). Dimer/monomer ratio remain constant..... | 193 |
| Figure V-16: overlaid structure of WT monomer hFABP5 (pink) and DSD hFABP5 (cyan) bound to palmitic acid..... | 194 |
| Figure V-17: Palmitic acids in overlaid structure of WT monomer hFABP5 (pink) and DSD hFABP5 (cyan). | 195 |
| Figure V-18: Overlay between a) structures of DSD FABP5 binds to 2-AG (green) and FABP5 binds to palmitic acid (blue) b) structures of DSD FABP5 binds to AEA (pink) and binds to palmitic acid (blue). | 196 |
| Figure V-19: Thermal stability assay for monomer FABP5 (fraction 21 in blue and 22 in red).198 | |
| Figure V-20: Thermal stability assay for dimer FABP5 (fraction 18 in blue and 19 in red)..... | 198 |
| Figure V-21: a) Domain swapped dimer of FABP5 (PDB 4AZR), there is a hydrogen bond between Q65 and T57. b) monomer structure (purple, PDB ID 4LKP) Gln65 is pointing out of toward the solvent and T64 makes a hydrogen bond to Thr57 via water network..... | 200 |
| Figure V-22: Shows SEC after expression of Q65A FABP5. | 201 |
| Figure V-23: a) Size exclusion chromatogram of Gln65 mutants of FABP5. b. SEC of T64 mutants of FABP5 (Thr63E pink, T63L green)..... | 201 |
| Figure V-24: Chromatogram of SEC column after a) expression of FABP5 in 25°C. b) expression in 20°C. | 203 |
| Figure V-25: Sequence alignment of WT-hFABP5 and WT-hCRBP II. | 204 |
| Figure V-26: Overlaid structures of monomer Q108K:K40L:hCRBP II bound to retinal (PDB: 4RUU) (Orange) with WT hFABP5 (Blue). Lys60 might interfere with binding the retinal in hFABP5..... | 205 |
| Figure V-27: TSA data for fractions 16-22 of size exclusion chromatography for WT FABP5. Data demonstrates the gradual switch from 40 °C to 65 °C..... | 210 |

KEY TO SYMBOLS AND ABBREVIATIONS

| | |
|---------|---|
| Å | Angstrom |
| PDB ID | Protein Data Bank Identifier |
| PBS | Phosphate-buffered saline |
| bR | Bacteriorhodopsin |
| TM | Transmembrane |
| 7TM | seven transmembrane |
| cGMP | cyclic guanosine monophosphate |
| iLBP | intracellular Lipid Binding Proteins |
| hCRBPII | human Cellular Retinol Binding Protein II |
| FABP5 | Fatty acid binding protein 5 |
| SB | Schiff base |
| PSB | Protonated Schiff Base |
| ESPT | Excited-state proton transfer |
| FP | Florescence protein |
| GFP | Green florescence protein |
| EGFP | Enhanced green florescence protein |
| YFP | Yellow florescence protein |
| RS-FP | Reversibly photoswitchable fluorescence protein |
| PYP | Photoactive yellow protein |
| FLIP | Fluorescence loss in photobleaching |
| FLINC | Fluorescence fluctuation increase by contact |

| | |
|---------------|---|
| FLAP | fluorescence localization after photobleaching |
| FLIM | Florescent lifetime imaging microscopy |
| PALM | photoactivated localization microscopy |
| PAM | photoacoustic microscopy |
| CALI | chromophore-assisted light inactivation |
| FALI | fluorophore-assisted light inactivation |
| FMI | fluorescence molecular imaging |
| <i>p</i> -HBI | 4-(<i>p</i> -Hydroxy-benzylidene)-5-imidazolinone |
| NP | nanoparticle |
| SNP | Semiconductor nanoparticle |
| POI | protein of interest |
| ROI | region of interest |
| PPI | protein-protein interaction |
| σ | Sigma |
| WT | Wild type |
| IPTG | Isopropyl β -D-1-thiogalactopyranoside Polymerase |
| PCR | Polymerase Chain Reaction PEG Polyethylene glycol |
| SDS-PAGE | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| Kd | Kilo Dalton |
| E. Coli's | Escherichia coli |
| min | Minute |
| S | Second |
| h | Hour |

| | |
|--------------|---|
| UV | Ultraviolet ight |
| Vis | Visible light |
| IR | infrared |
| NIR | near infrared |
| Abs | Absorbance |
| Em | Emission |
| SS | Stock shift |
| LSS | Large stock shift |
| mM | Milimolar |
| μ M | Micromolar |
| nM | Nanomolar |
| mol | mole |
| mmol | Milimole |
| mg | Miligram |
| ml | Mililiter |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxynucleotide triphosphates rpm Deoxynucleotide triphosphates |
| FPLC | Fast protein liquid chromatography |
| FTIR | Fourier transform infrared spectroscopy |
| $^{\circ}$ C | Degrees of centigrade |
| K | Degrees of kelvin |
| RT | Room temprature |
| pH | Logarithmic scale of hydrogen ion activity |

| | |
|--------|------------------------------------|
| NaOH | Sodium hydroxide |
| RMSD | Root mean square deviation |
| Amp | Ampicillin |
| Kan | Kanamycin |
| LB | Luria Broth |
| FBS | fetal bovine serum |
| DMEM | Dulbecco's Modified Eagle's Medium |
| Ala, A | Alanine Arg, |
| R | Argannine |
| Asn, N | Asparagine |
| Asp, D | Aspartate |
| Cys, C | Cysteine |
| Gln, Q | Glutamine |
| Glu, E | Glutamate |
| His, H | Histidine |
| Ile, I | Isoleucine |
| Leu, L | Leucine |
| Lys, K | Lysine |
| Met, M | Methionine |
| Phe, F | Phenylalanine |
| Pro, P | Proline |
| Ser, S | Serine |
| Thr, T | Threonine |

Trp, W Tryptophan

Tyr, Y Tyrosine

Val, V Valine

CHAPTER I: ROLE OF WATER MOLECULES IN WAVELENGTH TUNING OF HCRBP II

I-1 INTRODUCTION

I-1-1 Signal Transduction

Signal transduction is a process in which chemical or physical signals lead to a following molecular event in cells, which is known as a signaling pathway. Proteins that are responsible for detecting these signals are receptors.¹ Signal transduction usually involves the binding of signaling molecules, ligands to the receptors. Most of these ligands, such as growth factors, cytokines, and neurotransmitters, are soluble molecules that bind to cell surface receptors. However, some other ligands are soluble lipid molecules such as steroid hormones, since they need to cross the membrane and reach to nuclear receptors. Herein, we will explore more into signal transduction in vision, a family of GPCR, and, more specifically, its rhodopsin subfamily, which plays a crucial role in color vision and sensing.^{2 3 4 5}

I-1-2 Color Vision

Color vision is the ability to recognize different wavelengths of light, and it is a multistep process. When light hits an object, some of it absorbed, and some of it reflected by the object. Light passes through the cornea, the other layer of the eye. Then the cornea bends the light through the pupil and lens. The lens focuses the light through the retina, which is a layer of nerve cells located in the back of the eye (**Figure1-1**). Finally, Photons of the light make electrical signals which travel from retina to the brain to form images that we see. As mentioned, the retina contains visual neural

cells.⁶ These photoreceptors are located in the Macula of the retina and categorized into two categories of Rods and Cones cells. Scientist estimates that we have 6 million Cones and 110 million Rods in our retina.⁷

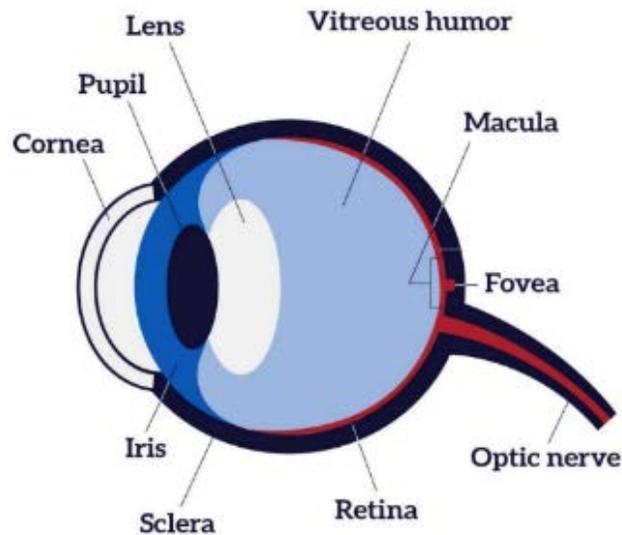


Figure I-1: Structure of human eye.

Cones are activated in a bright environment and Rods are more sensitive and stimulated in darker environments. Cons contain three photopigments (Red, green, and Blue) activated in a different wavelength of visible light. Through activation of the Cons by visible light, Cons send a signal to the visual cortex; Our brain processed different colors depending on the number of Cons that were activated, the strengths of their signal, and our past visual experience with the objects. In the darker environments, if only the Rods are activated, we do not recognize the colors, and we usually see them as gray. When one or more of the Cons types are not functioning, colorblindness can happen.⁸

I-1-3 Rods and Cones

Up to now, we gained a general view of the color vision. How exactly photons of light converts to electrical signals in the retina. Let's delve more deeply into the structure of Rods and Cones and

how the signaling happens through them. Rods and Cones are contained an outer and inner segments (**Figure 1-2**).

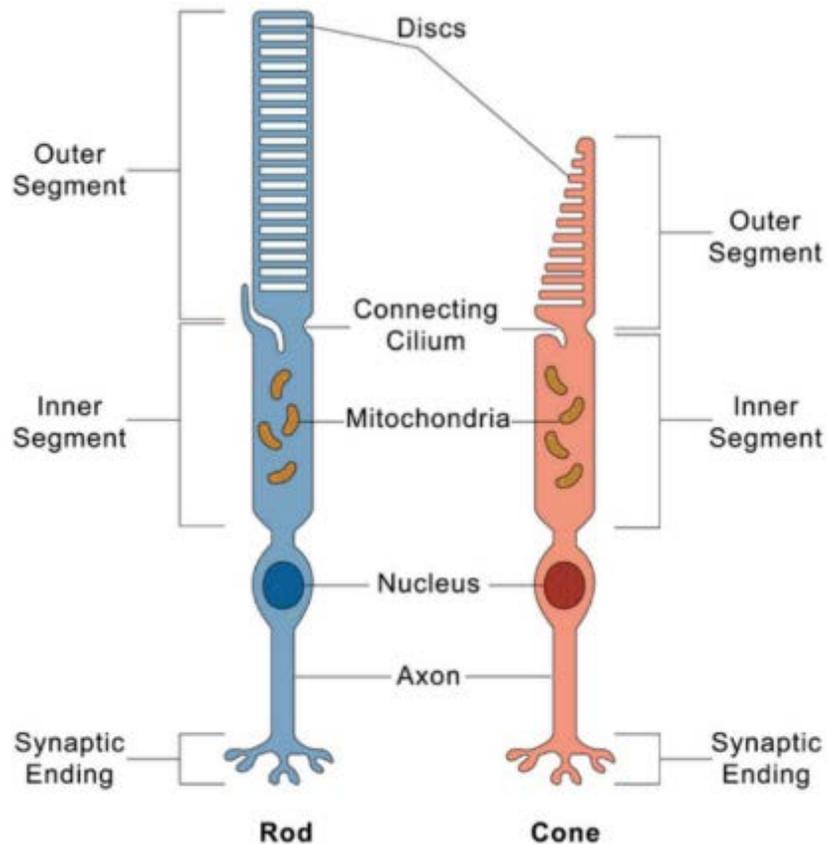


Figure I-2: Schematic structure of Rod and Cone cells.

The outer layer of these visual cells contains membrane proteins called opsins. The combination of these proteins with their chromophore, retinal, is called rhodopsin (**Figure I-3**). The N terminus of rhodopsins are exposed to the extracellular surface, and the C terminus is located in the cytoplasmic region. The only pigment in Rod cells is Rod rhodopsin with an absorption 500nm. Cone cells contain three different rhodopsins: Blue (10%), green (30%), and Red (60%) Rhodopsin. The wavelength absorptions for these pigments in humans are 425nm, 530, and 560nm, respectively.^{9 4}

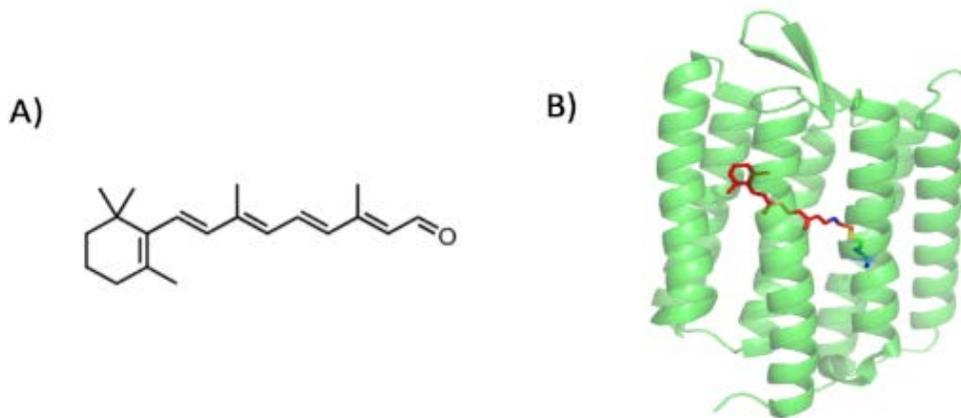


Figure I-3: A) all-trans retinal B) retinal bound to bacteriorhodopsin.

I-2 RHODOPSIN

Rhodopsins are a large family of proteins, and they play a crucial role as ion channels, ion pumps, signaling, sensing, and vision. The two major categories of this family are microbial rhodopsin and animal rhodopsins.¹⁰ Microbial rhodopsins usually function as energy production and signaling. Biological functions of the microbial rhodopsins are illustrated (**Figure I-4**).

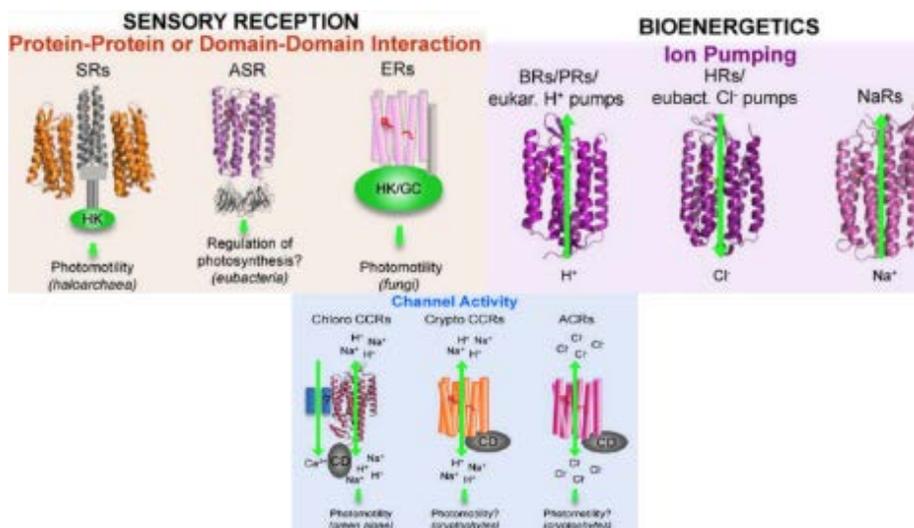


Figure I-4: Biological functions of microbial rhodopsins as ion pumps, Ion channels and sensors.¹¹

Animal rhodopsins are part of a larger family of Guanine nucleotide-binding protein-coupled receptors (GPCR) employed in color vision, sensing, and non-visual transduction.^{12 13} Structures of rhodopsins are similar and have a seven transmembrane helix fold (**Figure 1-3**). Although the structures of microbial and animal rhodopsins are similar, they have little sequence similarity.¹⁴ The first structure of visual rhodopsin was solved in 2000 by Palczewski at 2.8 Å resolution (PDB: 1F88). In the structure, the chromophore, retinal, is bound through a covalent bond, Schiff Base, to Lys296 in the binding pocket of the protein. The chromophore is buried inside the binding pocket, and is parallel to the surface of the membrane, helping to absorb the maximum amount of light. Retinal has interaction with the residues of the binding pocket, and it is locked in that position. The visual phototransduction, conversion of the photon of the light to an electrical signal in the retina, occurs via rhodopsin proteins.¹⁵ George Wald elucidated this process, and he received a Noble Prize in 1967 for his great work in this area. The protein's ability to absorb light comes from the chromophore, retinal, inside the binding pocket of the protein, opsin. Retinal forms by the Cleavage of β -carotene (**Figure I-5**).^{16 17 18} Retinal bound via a covalent bond to the binding pocket of the protein. Retinal aldehyde makes a Schiff base to the Lys residue inside the binding pocket of the protein (Lys296), forming the retinylidene protein. The Schiff base can be protonated; delocalization of the charge along the polyene region leads to different absorption of the protein in the visible region (**Figure 1-6**).

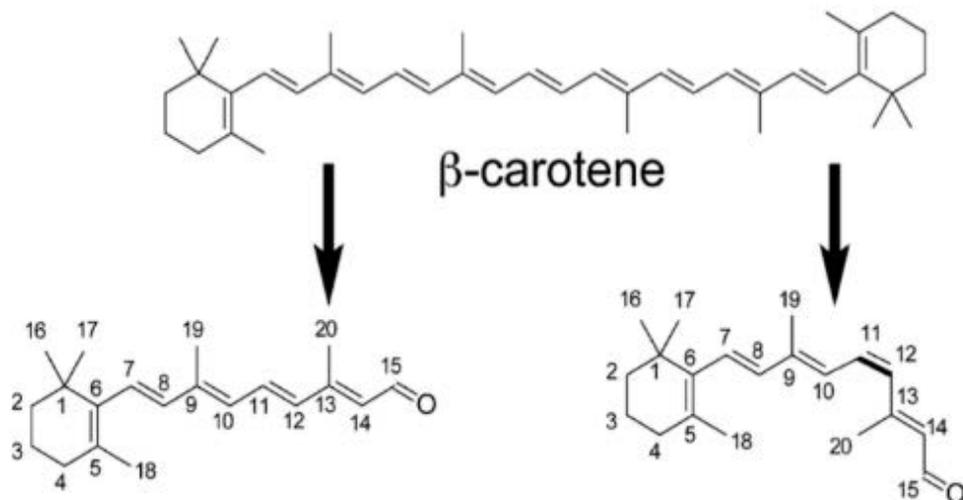


Figure I-5: Retinal forms by the Cleavage of β -carotene.

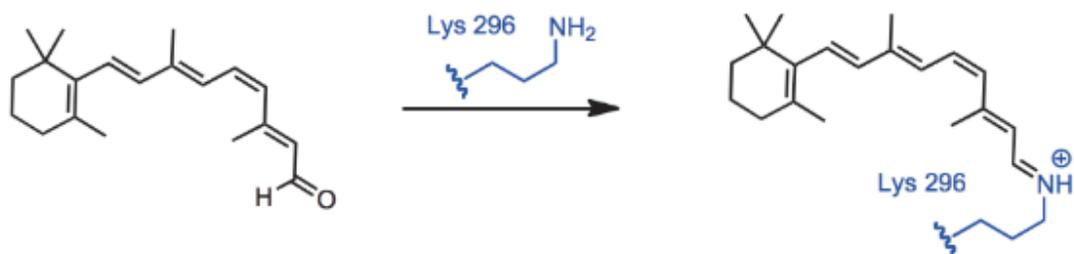


Figure I-6: Retinal makes Schiff base with lysine residue in rhodopsin.

Exposure to visible light leads to the photoisomerization process in rhodopsin, which is the critical event in the sensing and vision process. In Animal rhodopsin, the chromophore that binds to the opsin is 11-*cis* retinal, and upon light absorption, it isomerizes to the most stable thermodynamic product, all *trans*-retinal. The thermodynamic product then hydrolyzes, leaves the binding pocket, and another 11-*cis* retinal will be substituted in the system. In microbial rhodopsin, the isomerization will be different. Retinal first bound through the all-*trans*-retinal to the lysine; after

light exposure, the chromophore isomerizes to the 13-*cis* isomeric form, and it will convert back to all-trans form through dark adaptation (**Figure 1-7**).

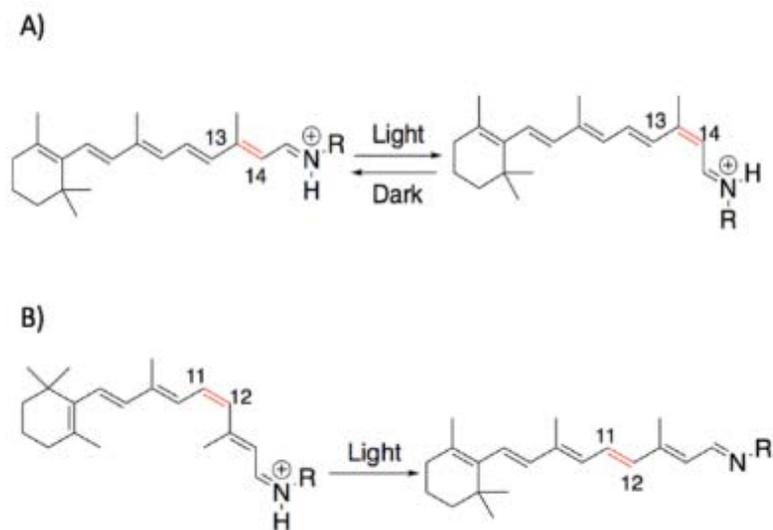


Figure 1-7: Isomerization of retinal in A) bacteriorhodopsin B) Microbial rhodopsin.

This photoisomerization process happens in picoseconds. The conformational change of the chromophore causes the change in the transmembrane helices and leads to a phototransduction cascade. To elucidate intermediates of photoisomerization in rhodopsin proteins, and driving forces behind this process, many studies using structural biology and time-resolved spectroscopy have been done so far.¹⁸ I will explain the photoisomerization in Rhodopsin proteins and photoisomerization in our engineered proteins, in more detail in the next chapter.

I-3 VISUAL PHOTOTRANSDUCTION CASCADE

Charged ions passage through ion channels inside the cell. These channels are activated by the cyclic guanosine monophosphate (cGMP) in rods and cones.¹⁹ In the presence of cGMP, this influx of ions leads to the release of the neurotransmitter glutamate. In the phototransduction process, active rhodopsin state first binds to the GDP-bound transducin heterodimer and exchanges the

GDP for GTP transducin, which has three subunit α , β , γ .^{20 21} Subunit α dissociate from the other two subunits and binds to phosphodiesterase (PDE). This binding cause removing the γ inhibitory subunit of PDE to dissociate, making it activated. The activated PDE converts the cGMP to GMP, which leads to closing the ion channels. Each activated rhodopsin will convert 103 cGMP to GMP. Closing the ion channels and hyperpolarization of the membrane leads to lowering the release of glutamate. Decreasing the number of glutamates that are neurotransmitters leads to a signal for the presence of light. When activated rhodopsin becomes deactivated, cells return to their normal states. This phototransduction cascade is illustrated (**Figure 1-8**).²²

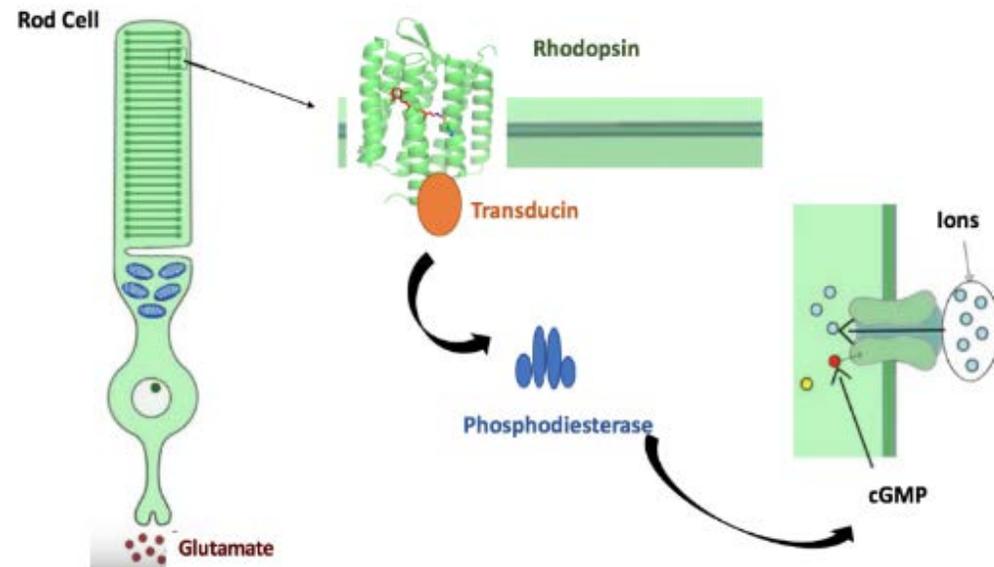


Figure I-8: Shows the three subunits of transducin and the dissociation.

With the development of high-resolution Cryo-EM microscopy, the structure of the light-activated Rhodopsin-GT complex in the presence and absence of a G-protein-stabilizing nanobody has been obtained (**Figure I-9**) (**Figure I-10**).^{17 23 24}

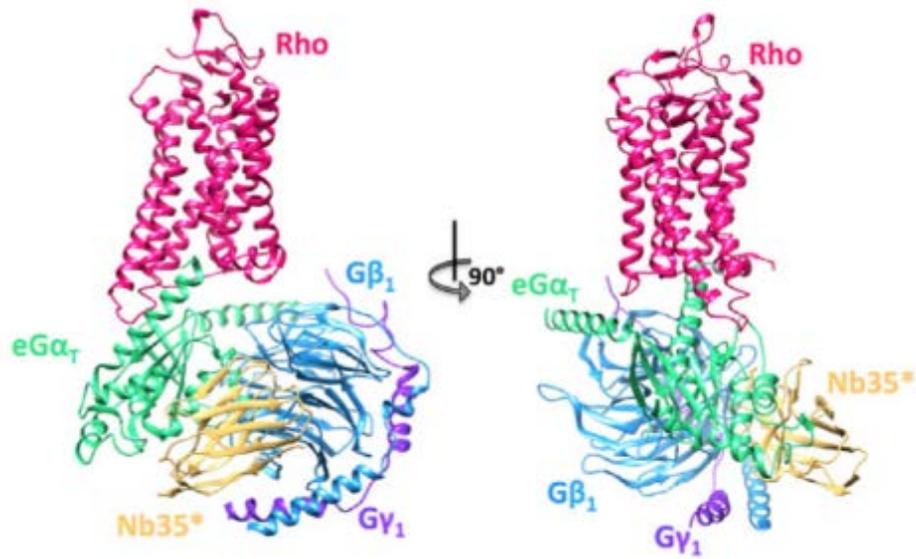


Figure I-9: Structure of the Rhodopsin-GT complex.²³

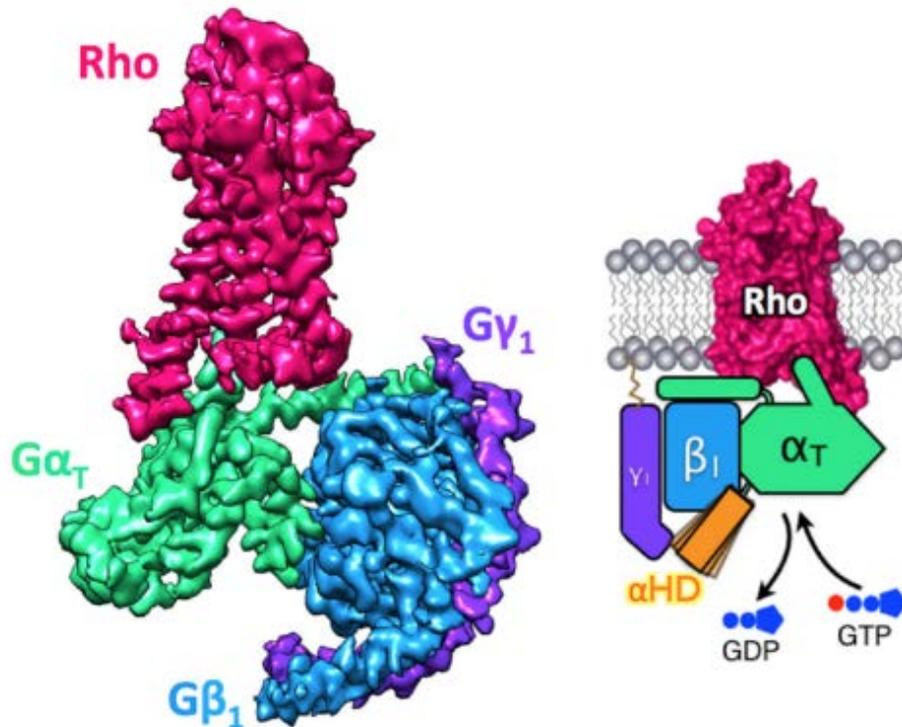


Figure I-10: Schematic structure of Rhodopsin-GT complex²³.

We now understand the color vision process and the critical role of neural vision cells and rhodopsin proteins. But, how exactly wavelength regulation happens in a color pigment like retinal? Understanding the wavelength regulation is essential to discover the mechanism and photophysics of rhodopsin proteins in vision. Furthermore, it helps us in designing new chromophores and fluorophores that I explain more in chapter III.

I-4 WAVELENGTH TUNING IN RHODOPSINS

Retinal binds through the covalent bond via the formation of an iminium Protonated Schiff Base (PSB). The absorption of free retinal is 380nm. The Absorption of the Schiff base when it is not protonated (SB) is 365nm, and when it is protonated (PSB) is 440nm in ethanol. However, different rhodopsins have different PSB absorptions depending on their protein environments. The difference between absorption of PSB in protein compared to ethanol is called “opsin shift,” and it is used as a starting factor for the elucidation of wavelength tuning in rhodopsin proteins.^{25 26 27}

Wavelength regulation on rhodopsins has been studied for more than 50 years. Theories about wavelength regulation are mostly based on conformational changes and electrostatic interactions. The point charge theory is based on electrostatic interactions and explains the stabilizing of the positive charge of the iminium with negative charges or negative dipole moments.^{28 29} Point charge theory focused on the fact that in the conjugated system, the delocalization of charge leads to more red-shifted absorption. Therefore, a higher number of negative charges near the iminium, lead to stabilizing the positive charge and more blue-shifted spectra. In contrast, introducing negative charges near the end of the polyene region of the chromophore leads to delocalizing of the positive charge and more red-shifted spectra (**Figure I-11**).^{30 31 32 33}

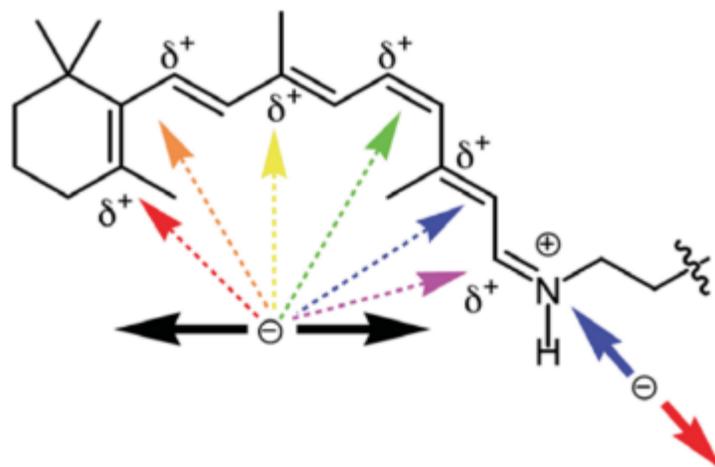


Figure I-11: Absorption of the protein depends on the position of the negative charge or negative dipole moments.

In 1993, interactions of counterions with Protonated Schiff Base (PSB) were studied by Sheves and his coworkers. In their model for rhodopsin, they showed a significant red shift absorption by removing the counterions from the PSB region.³⁴ They have demonstrated that changing the angles of counterions does not affect absorption; however, it can lead to different pKa environments for the chromophore. Since the counterion is the only negatively charged residue present in the chromophore's vicinity in both bacteriorhodopsin and bovine rhodopsin, the polarizability of the residues may play a role in the delocalization of the charge. In 1976, Mathies and Stryer demonstrated that permanent dipoles of the residues in the chromophore's vicinity could generate inducible dipole.³⁵ Nakanishi and Honig proved the negative charges' presence as the reason for stabilizing the positive charge on the iminium. Further studies have been done by mutating the Glu and Asp residues buried in rhodopsin. Mutating the Glu111 to Gln 3.45 Å away from the PSB nitrogen leads to a blue-shifted absorption (380nm) and lowering the pKa by 6 units. Mutating this counterion to other residues leads to 30nm red-shifted absorption as a result of destabilizing the

positive charge. Other studies tried to study the difference between colored opsin.³⁶ Nakanishi and Honig suggest that in bacteriorhodopsin, the point charge is on the ion ring; however, in bovine rhodopsin, the point charge is located in the middle of the polyene region. They explain the difference between the absorption of bacteriorhodopsin (560nm) and bovine rhodopsin (500nm) with their hypothesis. Electrostatic potential calculations have also been done to calculate the distribution of charge along the polyene for sensory rhodopsin II (SRII) 486 nm, bacteriorhodopsin (BR) 552 nm, and halorhodopsin (HR) 576 nm. These calculations demonstrate the negative dipole moments along the polyene (**Figure I-12**).^{37 38}

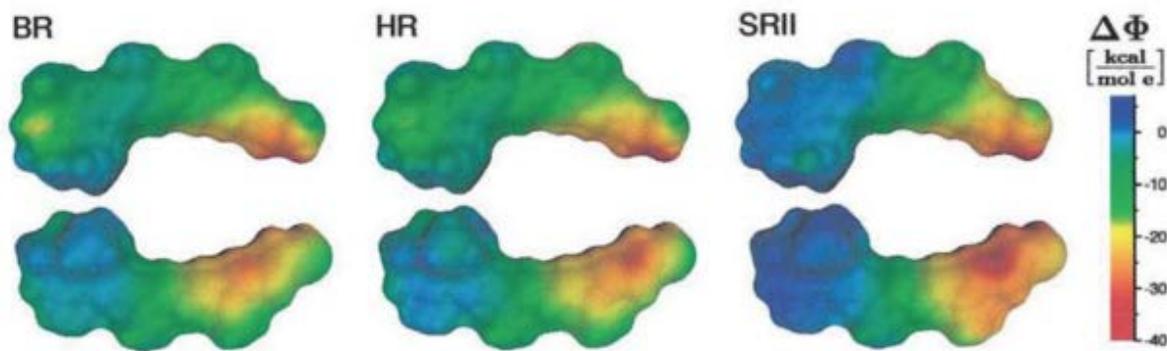


Figure I-12: Electrostatic potential calculations for sensory rhodopsin II (SRII) 486 nm, bacteriorhodopsin (BR) 552 nm, and halorhodopsin (HR) 576 nm.

Dr. Lee from Prof. Borhan's group has done the same electrostatic potential calculations for Rhodopsin and three other colored opsins. These results also are consistent with the wavelength regulation concluded from before (**Figure I-13**).³³ Further electrostatic potential calculations have been done for our designed rhodopsin template system which will be discussed in this chapter.

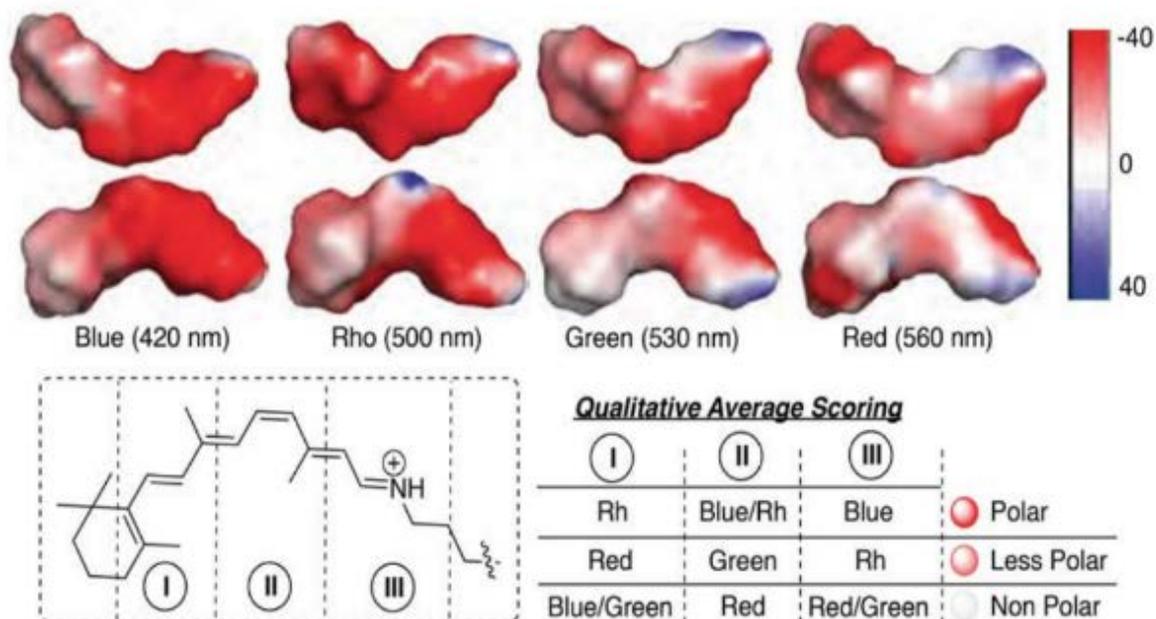


Figure I-13: Electrostatic potential calculations for Rhodopsin, blue, green and red opsin.

Among bacterial rhodopsins, SRII is 60nm blue shifted (500nm) compare to others, 560-590nm. Comparison between these rhodopsins followed by mutational studies revealed some critical residues like Val108, Gly130 and Thr 204 (the equivalent residues in BR are Methionine, Serine and Ala) in the binding pocket, which are essential for these differences. Another critical factor affecting the blue-shifted spectra is conformational changes of some residues in the binding pocket of the protein. Removal of hydroxyl groups of Ser141 and Thr142 near the B-ionone ring, changing the conformation of Arg 72, and change in the retinal position are other factors reported to affect the different absorption of SRII (**Figure I-14**).³⁹

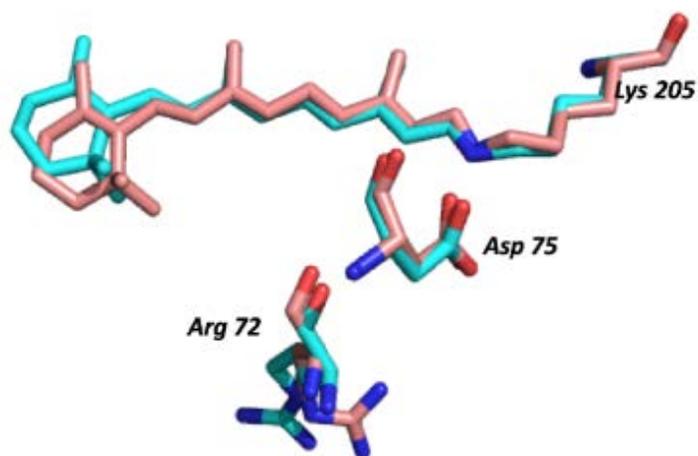


Figure I-14: Overlaid structures of the BR (CYAN - PDB ID: 1C3W) and SRII (Pink- PDB ID: 1JGJ).

Water molecules are another factor that play an essential role in wavelength tuning. Water molecules can stabilize the inactive and active states. The crystal structures of vertebrate active and inactivate state of rhodopsin have been obtained. Hydrogen bonds and water networks seem to stabilize the structures (**Figure I-15**). Water molecules are usually found in helices with irregularities with not enough hydrogen bonds.^{40 41 42}

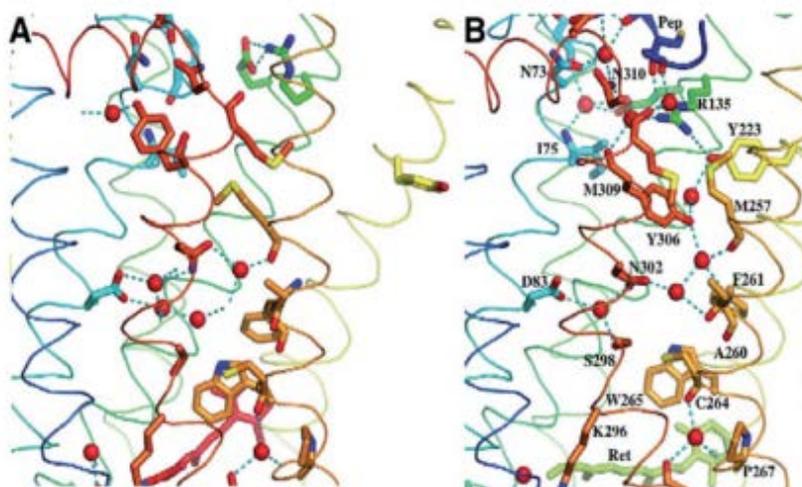


Figure I-15: The water-mediated hydrogen networks in rhodopsin A) Ground state B) activated state.^{42 40 41}

In the ground state and excited states of rhodopsin and bacteriorhodopsin, there are several highly ordered carboxyl side chains and water molecules that help to stabilize the positive charge on PSB. Upon the photoactivation process, these water molecules and polar residues change their arrangements, resulting in altering the speed of the isomerization process. The high-resolution structure of the ground state in bovine rhodopsin shows two water molecules near the PSB.^{43 44} The calculations demonstrate that each of them can contribute to a 30nm shift in the absorption. Wat2b has a hydrogen bond with Glu113 and stabilizes the negative charge of the counterion. Water 2a has hydrogen bonds with Glu181 and Ser186, and this network is mediated by Cys187 to Glu113 (**Figure I-16**).⁴⁴In Green and Red opsin, His 181 leads the chloride binding in the binding pocket.^{45 43 46} Wat2a and hydrogen binding with Glu181 probably substitute the role of histidine in bovine rhodopsin, and it affects wavelength tuning. Further mutation studies on Glu181 show the outcome in the spectra of rhodopsin.⁴⁵

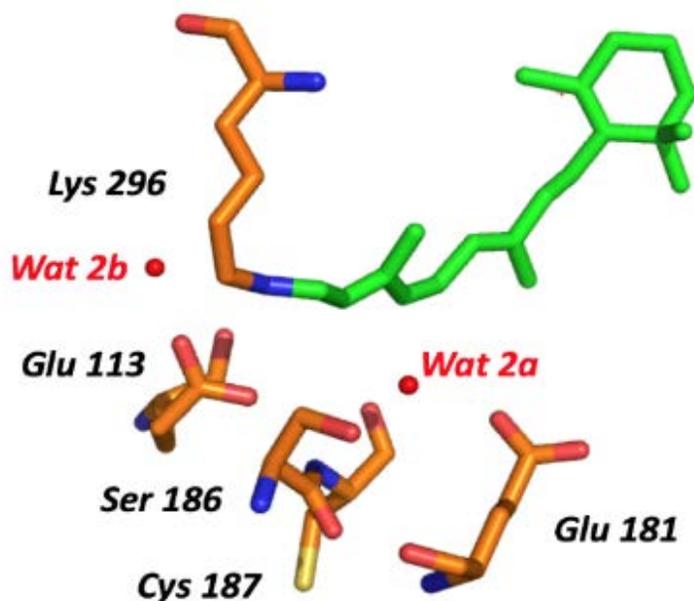


Figure I-16: Wat 2a and Wat2b in the vicinity of the retinal in rhodopsin (PDB: 1L9H).

From the crystal structure of the activated state of rhodopsin, it can detect the change in the arrangement of the water molecules and polar residues. Wat2a also helps the switch of the counter anion from Glu113 to Glu181 (**Figure I-17**).⁴⁷

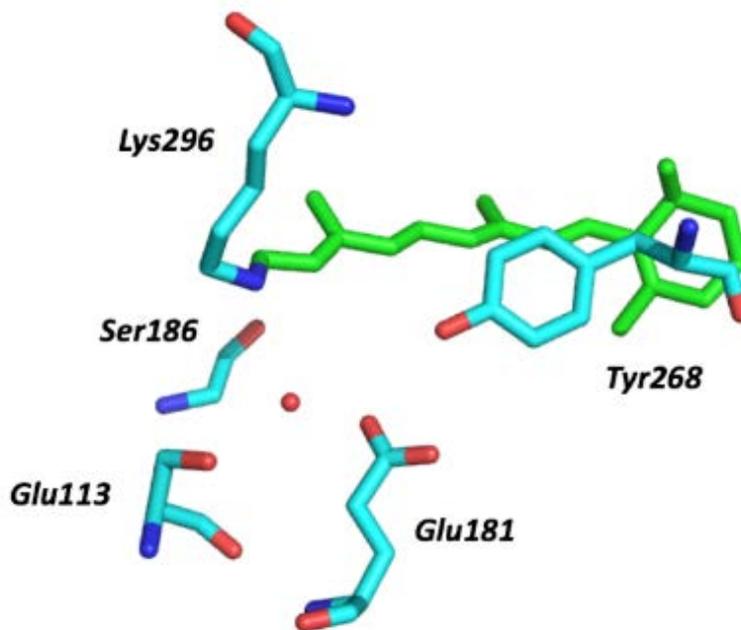


Figure I-17: *Glu181 plays as counterion in metaII rhodopsin (3PQR).*

In 2012, the role of the water networks further clarified by obtaining the structure of the active state of the rhodopsin. In the crystal structure, water molecules span from the binding pocket to the transmembrane side. Further, by developments in X-ray free-electron lasers (FELs) for serial femtosecond crystallography (SFX), and high brilliance synchrotrons for serial microsecond crystallography (SMX) help to better investigate the change in the water networks.^{48 49} Using these method helped to record the change in structure through photoactivation in short time shots. Differences between the ground state and activated stated maps could facilitate for getting insight upon the positions of water molecules. In the ground state of bacteriorhodopsin, there are three water molecules making hydrogen bonds with hydroxyl groups. 16ns snapshot structure after

activation of bacteriorhodopsin revealed the change in position of water molecules. The difference increases overtime among activation (**Figure I-18**)^{40 48}.

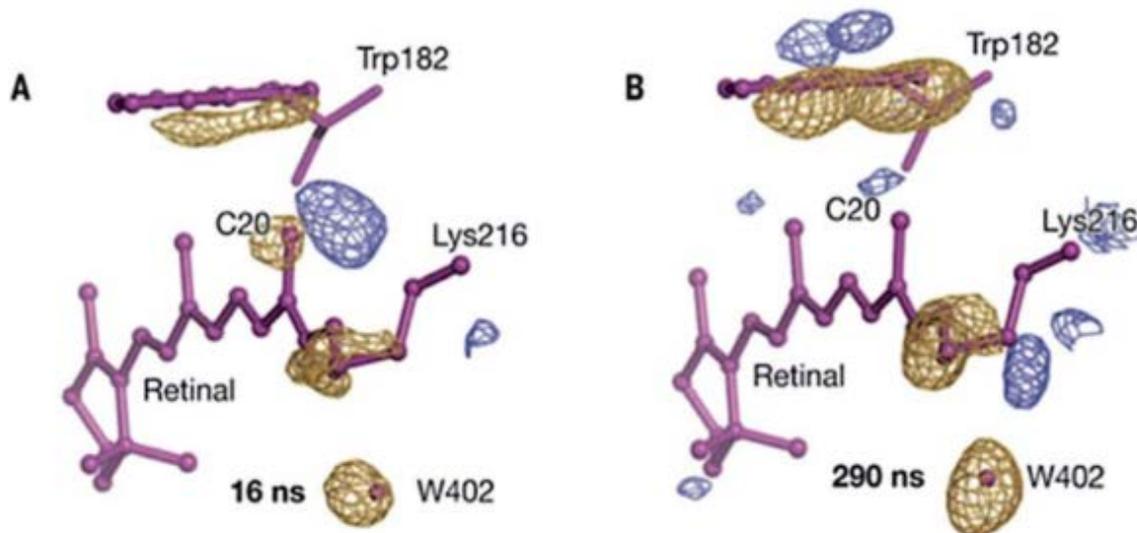


Figure I-18: Difference in electron density maps A) 16ns and B) 290ns after activation, obtained by time-resolved room temperature crystallography. (positive density Blue and negative density Gold). Even in 16ns, there is gold negative density for water402⁴⁰.

Rhodopsins also are categorized by the stability of their active states to mono and bistable groups. For many years, squid rhodopsin structures as bistable rhodopsin and bovine rhodopsin as a stable mono group have been compared.⁵⁰ Recently, the structure of another member of bistable rhodopsin, the jumping spider rhodopsin-1, bound to the inverse agonist 9-*cis* retinal have been found. This structure also revealed the highly ordered water networks in jumping spider rhodopsin-1 from binding pocket to the transmembrane (**Figure I-19**).⁵¹

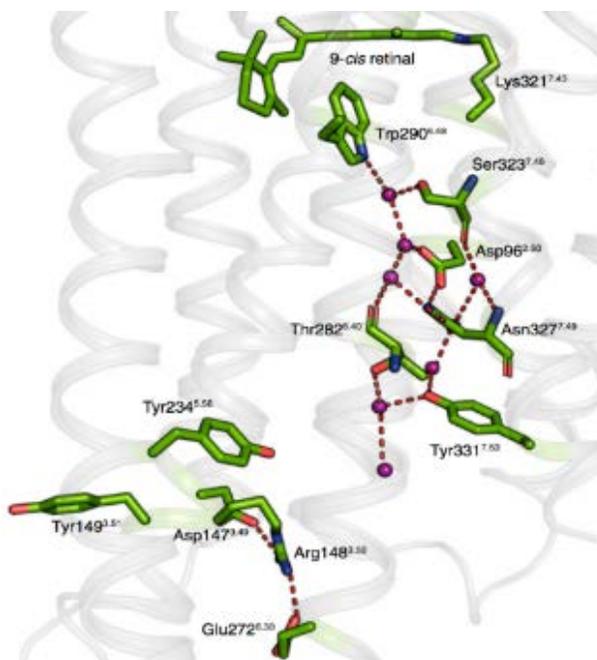


Figure I-19: Water network in binding site of jumping spider rhodopsin 1.

Comparison between Bovine rhodopsin, squid rhodopsin, and jumping spider rhodopsin-1 demonstrate the similarities between the water-mediated binding pocket of bovine rhodopsin and jumping spider rhodopsin-1 (**Figure I-20**).⁵¹

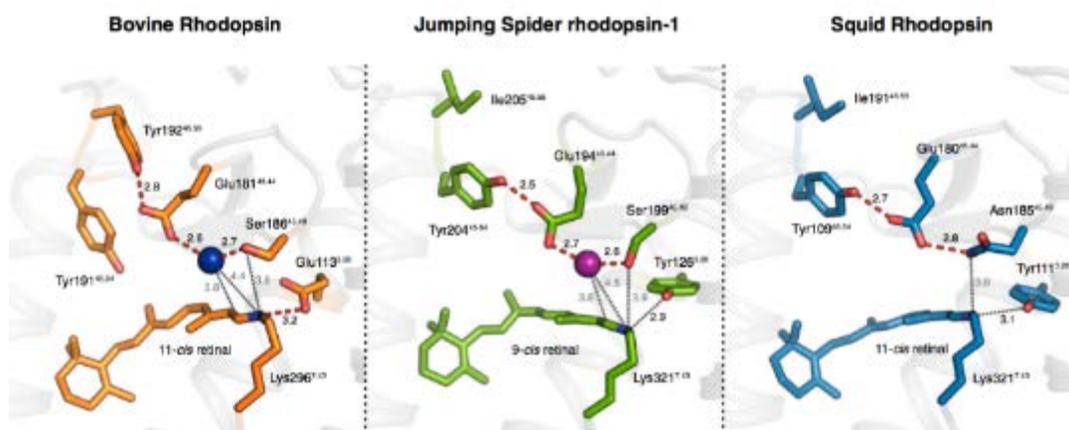


Figure I-20: Comparisons of the retinal binding pockets between mono bovine rhodopsin (1GZM) and bistable Jumping spider rhodopsin 1 (6I9K) and squid rhodopsins (2Z93).

The chromophore's conformation and planarity is another factor that can have a significant impact on the wavelength regulation by changing the level of the delocalization of the π -electrons. The conformational theory is based on the effect of the protein environment in the trajectory of the chromophore. The conjugated system between double bonds and traveling the positive charge of the PSB along the polyene depends on the planarity of the chromophore. The binding cavity can keep the polyene in a plane and keep the conjugated system, or it can make it twisted. When the chromophore is twisted, the PSB positive charge cannot be delocalized along the polyene, and it will lead to blue-shifted spectra. Steric repulsion of methyl groups and hydrogens, which causes the twisting of the chromophore, can be affected by the position and confirmation of residues of the protein. Gas-phase absorption of the retinal n-butyl iminium shows the broad absorption from 530nm to 610nm due to a change in B-ionone ring rotation, and the dihedral angle between the C5-C6 and C7-C8 double bonds of the chromophore. The difference between the absorption 6s-trans and twisted 6s-cis analog is form 610 to 630nm (**Figure I-21**).²⁹

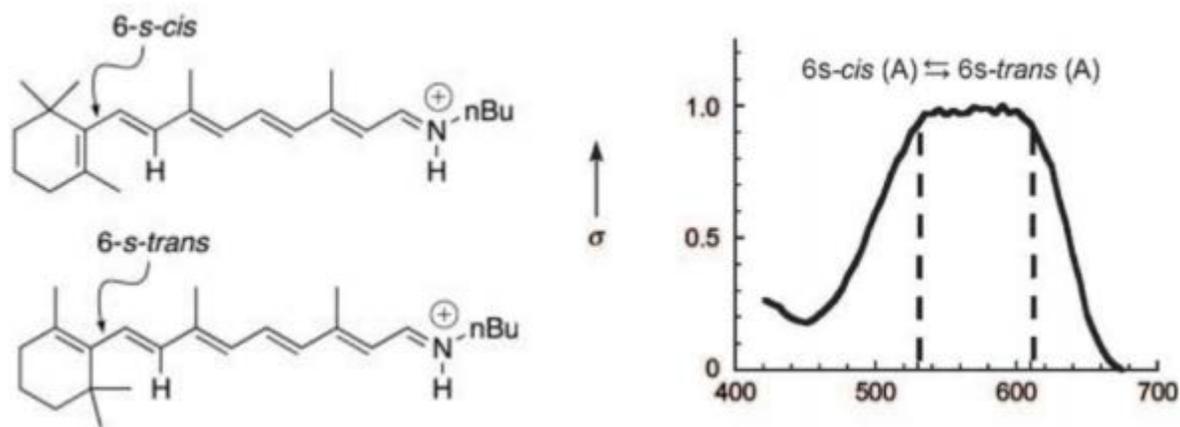


Figure I-21: The gas phase absorption of retinal n-butyl iminium PSB.

For making the blue-shifted pigments of Channelrhodopsins (ChRs), the 6s-trans to 6s-cis retinal isomerization has been used. The crystal structure of C1C2 reveals the 6s-trans retinal in the

binding pocket with 476nm.⁵² The double mutant of this protein demonstrates the 6s-*cis* retinal conformer with a 22nm blue shift compared to the wild type. β -ionone is another part of the chromophore that has been studied for many years. In one study, the conformation of the β -ionone compares with the α -ionone and β -damascon (**Figure I-22**)⁵³. The gas phase technique in combination with crystallography prove that the flatter potential energy surface of β -ionone ring makes it the most flexible structure for biological photoisomerization process.⁵⁴

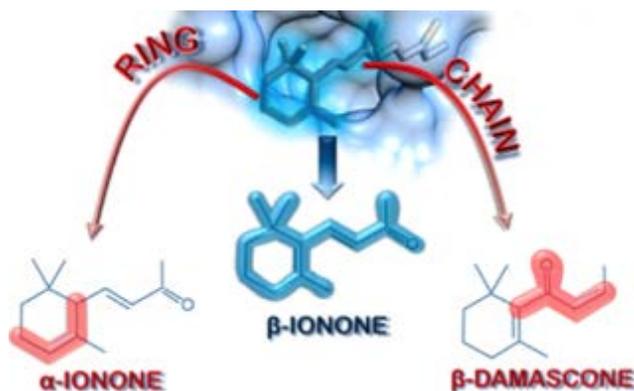


Figure I-22: β -ionone compares with the α -ionone and β -damascon structures.

Other attempts have been made to make more red-shifted channelrhodopsins to use in optogenetics. Inserting an extra C=C bond to all-trans-retinal (3,4-didehydro-retinal), lead to more red-shifted spectra in this system (**Figure I-23**).⁵⁵

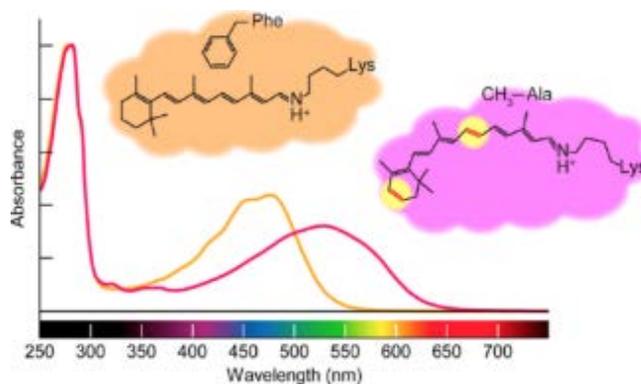


Figure I-23: Structures and UV-vis absorption spectra of ATR and all-trans-retinal (3,4-didehydro-retinal), and their analogue.

Other computational studies have been done to investigate the relationships between the amino acid sequence and conformational change of the chromophore with wavelength tuning.^{56 57 58 57} Poisson–Boltzmann/quantum chemical (PBQC) and charge density coupling (CDC), combined with the methodology of quantum mechanical/molecular mechanical (MM modeling) of the side chains, were used to explain the tuning of the retinal chromophore in green and red rhodopsins.⁵⁹ ^{60 61}The correlation between electrostatic interactions and wavelength tuning from the computational analysis in these cases were consistent with experimental data. Machine learning techniques also have been used to find the correlation between amino acid sequences and absorption wavelengths for around 700 microbial rhodopsins and their variants. Developing these techniques can help in predicting the absorptions and color tuning roles.⁶²

I-5 WHY RHODOPSIN MIMIC SYSTEMS

Rhodopsins are membrane proteins, and their expression, purification, mutagenesis, and crystallization are very challenging. As explained in the last part, there are many mutagenesis studies, physical studies like Raman spectroscopy, crystallography, and NMR have been done to understand the mechanism of the wavelength regulation and the photoisomerization of rhodopsin proteins.^{63 64} Moreover, recent studies successfully indicate the intermediates of rhodopsin's photocycle, using time-resolved experiments, femtosecond x-ray laser (x-ray free-electron laser, XFEL), and cryokinetic data. All of these studies demonstrate the different important biophysical characteristics of rhodopsin systems; however, they have some limitations.^{52 65 66} These proteins evolve a lot during evolution. Visual pigments and their proteins are known as one of the most convenient systems to study evolutionary changes. The evolutionary scheme of rhodopsins is illustrated in **(Figure 1-24)**.⁶⁷ Visual rhodopsins are one of the earliest proteins known in the animal kingdom.

| | |
|--|--|
| Cambrian Explosion (about 540 Ma) | Rhodopsins of the 2nd type (G-protein-coupled receptors) Visual pigment—rhodopsin |
| Bilateral animals (about 600 Ma) | Rhodopsins of the 2nd type (G-protein-coupled receptors) Visual pigment—rhodopsin |
| Metazoans (about 2 Ga) | Rhodopsins of the 2nd type (G-protein-coupled receptors) Visual pigment—rhodopsin |
| Eukaryotes (about 3.2 Ga) | Rhodopsins of the 1st type (ionic transportation) (algae, fungi) |
| Prokaryotes (about 3.8 Ga) | Rhodopsins of the 1st type (ionic transportation) Bacteriorhodopsin—photosynthesis |
| Beginnings of the Earth (4.567 Ga) (Khain, 2009) | |

Figure I-24: *The evolution of rhodopsins.*⁶⁷

As a result of environmental changes and developments during evolution, many amino acid residues in rhodopsins become conserved residues; therefore, it is hard to illustrate every single residue's effect on wavelength tuning through mutagenesis studies. Using evolutionarily naive photoactive systems can help us understand the evolution of these proteins and understand the impact of every single residue in the photocycle. Using an evolutionary naive soluble protein with high mutagenesis tolerance can provide us with a better template to study the photocycle. Our group, in collaboration with prof. Babak Borhan's group found two proper rhodopsin mimic systems: Cellular Retinoic Acid Binding Protein II (CRABPII) and Cellular Retinol Binding Protein II (CRBPII). These proteins are from the intracellular lipid-binding protein family.^{68 69 33}

I-5-1 Using Intracellular Lipid Binding Proteins (ILBPs) as Rhodopsin mimics

Intracellular lipid-binding proteins (iLBPs) are small (14-16kD) cytosolic proteins responsible for transporting various insoluble hydrophobic molecules, such as retinoids, retinal, fatty acids, cholesterol, and hemes, inside cells.⁷⁰ This family of proteins is in the cytosols of fish, birds,

amphibians, reptiles, and mammals.⁷¹ Many vital carrier proteins, such as retinol-binding proteins (RBP), cellular retinol-binding proteins (CRBP), cellular retinoic acid-binding proteins (RABP), and fatty acid-binding proteins (FABP) are part of this family. ILBPs are a subfamily of the lipocalin family.^{72 73} All of the family members have similar structures, including ten β strands and two α helices located at the mouth of the internal binding cavity. The β -barrel plays as a binding pocket for the ligand, and it is a relatively large binding pocket compared to the size of the protein. The α helices play as a lid for isolating the hydrophobic ligand inside the binding pocket (**Figure I-25**). Although crystal structures of most proteins in this family are similar, they have very low sequence homology.

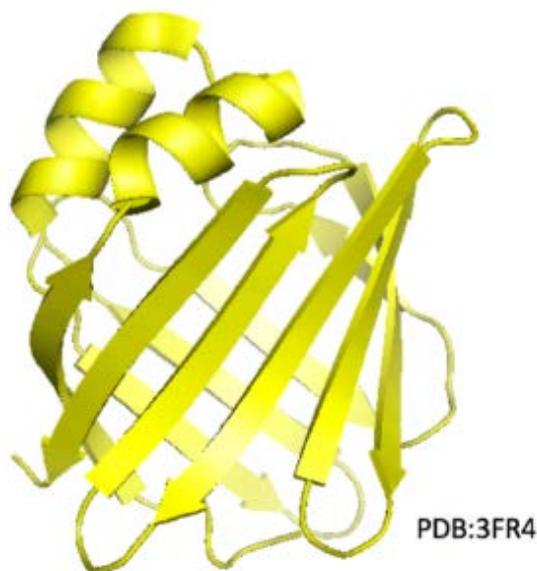


Figure I-25: Structure of A-FABP5 a member of iLBP (PDB:3FR4). 10 beta strand with two alpha helixes as a lid.

As discussed, we use Cellular Retinoic Acid Binding Protein II (CRABP II) and Cellular Retinol Binding Protein II (CRBP II) as templates for rhodopsin mimic systems.⁷⁴ Their solubility, small size, substantial binding pocket and ease of crystallization, makes them a great candidate for our purpose.^{75 76 77}

I-5-2 Human cellular retinoic acid binding protein II (hCRABPII)

The first template that has been studied is human cellular retinoic acid binding protein II (hCRABPII). This protein has 137 residues. It binds to all-trans-retinoic acid and delivers this hydrophobic molecule to retinoic acid receptors (RARs) in the nucleus.⁷⁸ Similar to rhodopsin, the chromophore designed to covalently attached to the Lys residue in the binding pocket. The first structure of this protein bound to all-trans-retinal was achieved by Dr. Soheila Vaezeslami, and also by Dr. Rafida Nossoni (**Figure 1-26**).

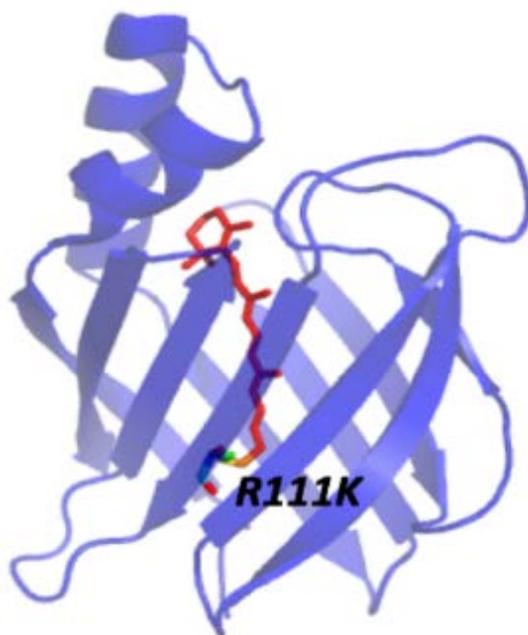


Figure I-26: Structure of hCRABPII bound to retinal (PDB:2FR3).

In collaboration with prof. Borhan's group, this protein has been used in many studies, such as wavelength regulation, pH sensing, and fluorescence studies. They were successful in making a wavelength-tunable system from 480nm to 630nm absorption by changing the binding pocket of hCRABPII protein (**Figure I-27**).⁷⁹



Figure I-27: Reengineered hCRABPII protein retinal pigments ranging from 482nm to 630nm.⁸⁰

Moreover, My former lab mates, Dr. Meisam Nosrati and Dr. Alireza Ghanbarpour, were successful in characterizing photoisomerization in the hCRABPII system.^{81 82} One of our final goals was using this system for cell-based assays and in vivo studies.⁸³ Developing new fluorophores can help us to use this system in many applications, such as imaging studies that I will explain more in chapter III.⁸⁴ However, trials for characterizing hCRABPII using fluorophores for in vivo studies were not successful. Therefore, our new studies mostly have been done in our second model, human cellular retinol-binding protein II. In the initiate CRABPII system, the ionone ring of the retinal was found to be water exposed. Therefore, we used our second template for wavelength regulation studies. In hCRBPII, the retinal buried in the binding pocket.

I-5-3 Human cellular retinol binding protein II (hCRBPII)

Human cellular retinol-binding protein II (hCRBPII) is another member of the iLBP family. hCRBPII binds to retinol and retinal with 10nM and 90nM dissociation constants, respectively. This protein is a small protein with 133 amino acids, and it is mostly found in the small intestine. It is responsible for the transport of retinol and retinal across enterocytes. Although the sequence identity between hCRBPII and hCRABPII is around 35%, their structures are similar, containing a ten strand β -barrel with two helices. The binding pocket of hCRABPII is more wide and open compare to hCRBPII. Also, the binding pocket of hCRBPII is around 5 angstroms deeper than

hCRABP_{II}. As a result, it helps the retinal to be wholly buried in the more enclosed binding pocket of hCRB_{II}. Since the retinal is fully embedded in the hydrophobic binding site of hCRB_{II} with low dielectric constant, change in the absorption is mostly affected by the interaction of protein and chromophore. Therefore, hCRB_{II} is a great template to investigate the wavelength tuning. Like the rhodopsin and hCRAB_{II} systems, the chromophore is bound covalently to a designed Lys residue in the binding pocket (**Figure 1-28**).^{85 86}

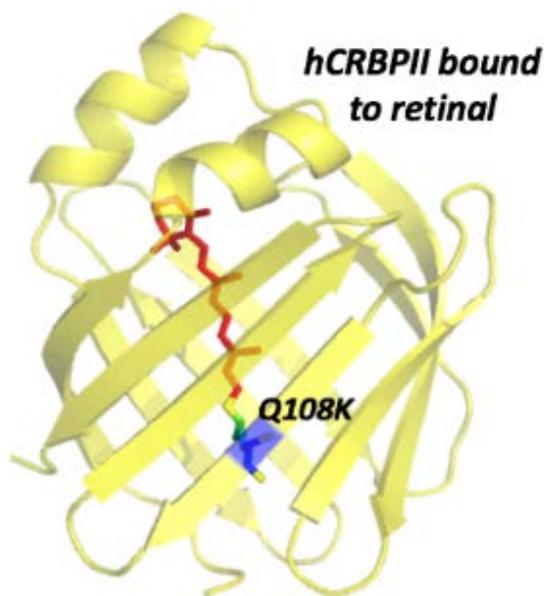


Figure I-28: Q108K; K40L (KL) hCRB_{II} structure bound to retinal through lysine residue (PDB: 4EXZ).⁸⁷

I-6 WAVELENGTH REGULATION IN HCRB_{II}

We have investigated the reasons behind wavelength tuning, using our templates, especially hCRB_{II}. Most of the initial experiments on wavelength regulation of hCRB_{II} have been done by Dr. Wenjing Wang from Prof. Borhan's group and Dr. Rafida Nossoni from our laboratory. Here, I will summarize some of their important findings in the wavelength tuning of hCRB_{II}. As

explained in their science paper, through a combination of mutations, they have successfully tuned the hCRBP_{II} template from 425nm to 644nm (**Figure I-29**).^{87 88 80}

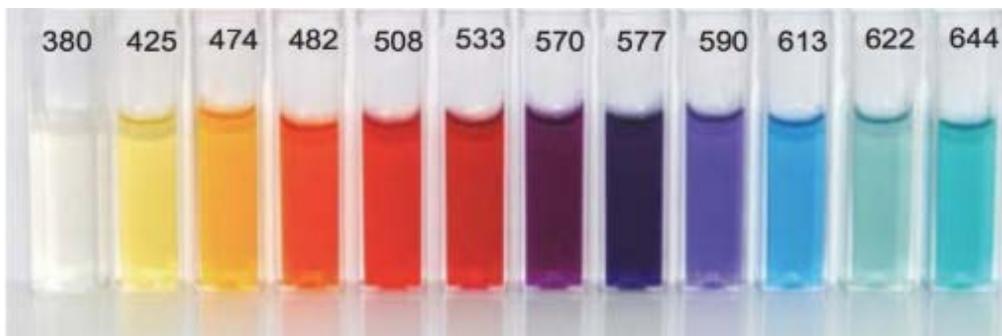


Figure I-29: Reengineered hCRBP_{II} protein retinal pigments ranging from 425nm to 644nm.

Unlike the absorption of SB, which is 360nm, the PSB absorption is not constant and depends on the delocalization of the positive charge on the polyene region. In this study, mutational analysis was explained based on the three regions on the chromophore (**Figure I-30**).⁸⁷

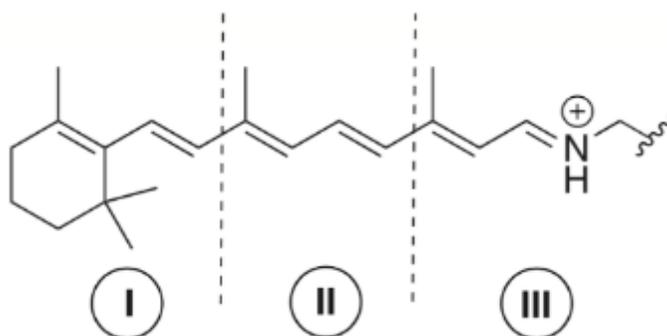


Figure I-30: The three regions of the retinal in wavelength tuning study.⁸⁷

In hCRBP_{II}, Q108K: K40L is the starting point since these mutations are essential for the proper binding of the retinal. By adding counterions next to the PSB (region III) and stabilizing the positive charge, the absorption blue shifted. Removing the counterions and water molecules in this region lead to red-shifted absorption. In the II and III regions, Thr51, Thr53, and Tyr19 were mutated. Combination of these mutations in Q108K:K40L:T51V:T53C: Y19W leads to a large bathochromic shift in the absorption. Unlike that expected, introducing of negative charge to the

ionone region (region I) did not result in red-shifted spectra. However, since the dielectric constant inside the binding pocket of the protein is lower than the aqueous environment, introducing the bulky hydrophobic residues in this region leads to the significant red-shifted spectra. In this region, mutating Arg58 located in the entrance of the binding pocket to an aromatic residue like Try inhibits the cavity's exposure to the bulk aqueous environment, resulting in the red-shifted spectra (**Figure I-31**). Introducing the R58W mutant lead to around a two-fold increase in red shifting absorption of internal mutants.

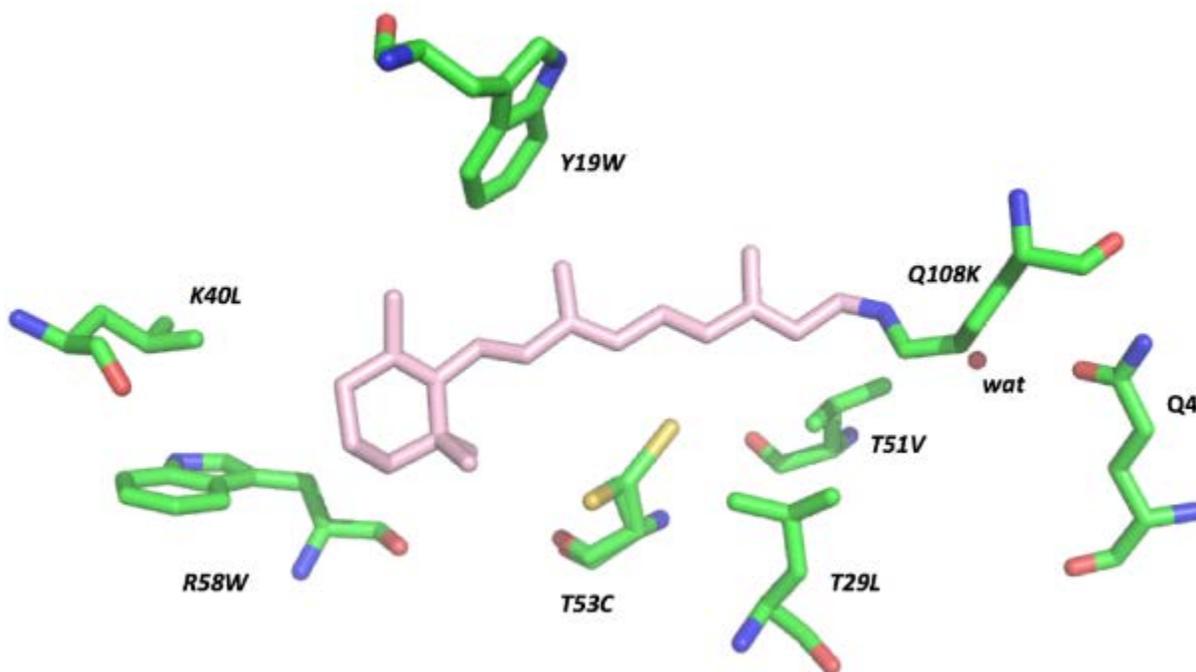


Figure I-31: Shows important residues for wavelength tuning in the hCRBP II system.

Another critical residue targeted for mutation was Gln4. As illustrated in (**Figure I-32**), Glutamine hydrogen bonds to an ordered water molecule proximal to the PSB. As a result of removing the water molecule by mutating the Glutamine, shifts the wavelength from 591nm in KL-T51V:T53C:Y19W:R58W:T29L to 613nm. As reported in the science paper, unlike other mutants,

structures of variants lacking Gln4 also lack the ordered water molecule, and have a *trans* conformation of the Schiff base (**Figure I-32**). The explanation for the *trans* conformation was not apparent when the original paper was published. One of the goals of this thesis was to determine the true conformation of the Schiff base in these mutants when the Schiff base is formed and the spectra were measured. Chapter II of the thesis describes these results. We hypothesized that the positive charge on the *cis* PSB stabilized with Try106 through Pi-cation interaction and the water molecule. By removing the water molecule, the positive charge is not stable, and it will convert to *trans* form over time.

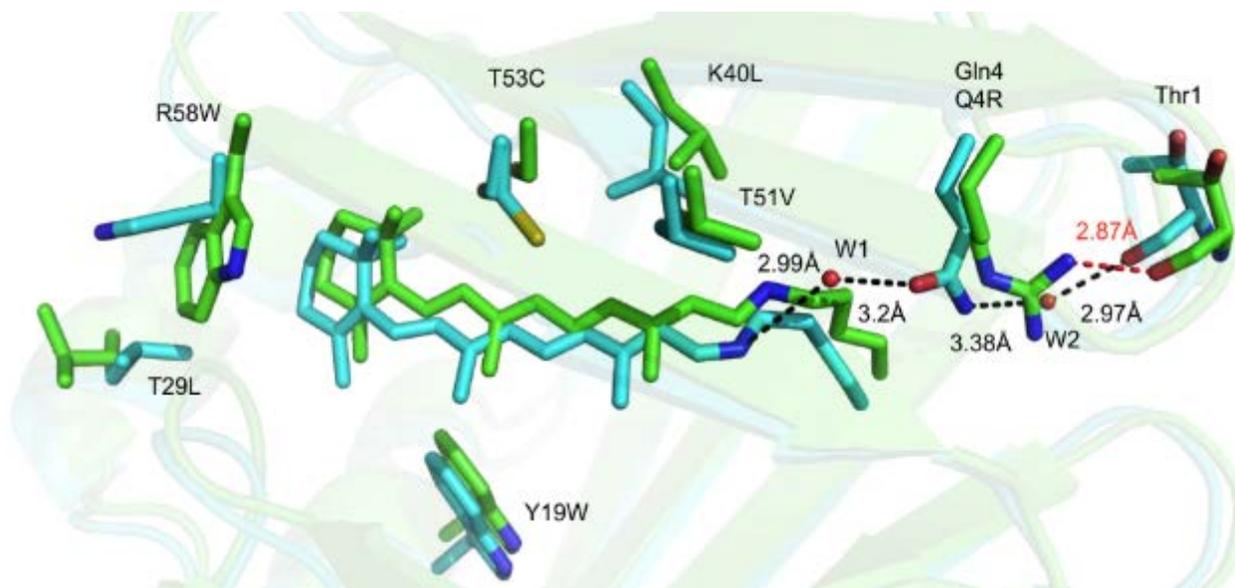


Figure I-32: The overlaid structure of KL:T51V:T53C:R58W:Y19W:T29L (cyan, 591 nm) and of KL:T51V:T53C:R58W:Y19W:T29L:Q4R (green, 622 nm).

Comparison of electrostatic calculations for KL and KL:T51V:T53C:R58W:Y19W:T29L:Q4R, the red-shifted hCRBP variant, suggests the red shifting is caused by an equal distribution of the charge along the polyene (**Figure I-33**).

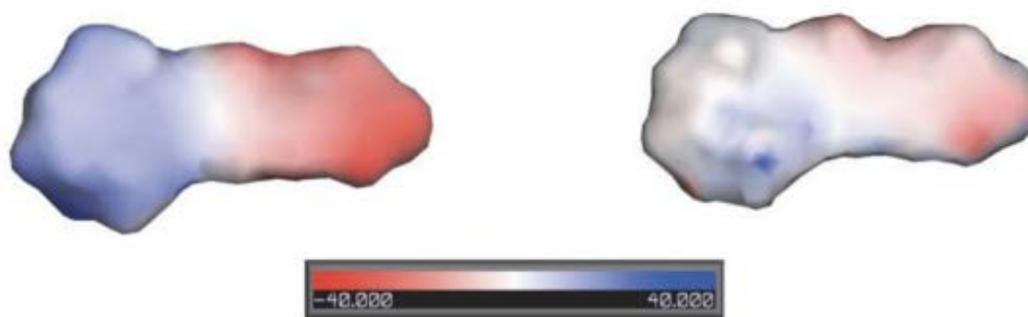


Figure I-33: Demonstrates electrostatic charge distribution on the surface of the retinal for *Q108K:K40L* mutant of *hCRBPII* on the left and *Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R* on the right.⁸⁷

To produce the most bathochromically shifted pigment, another bulky residue was introduced in zone I. Introducing A33W to KL-T51V:T53C:Y19W:R58W:T29L: Q4R lead to most red shifted variant so far discovered, with an absorption of 644nm. The summary of absorption for a spectrum of *hCRBPII* variants is illustrated in (Figure I-34).^{87 88}

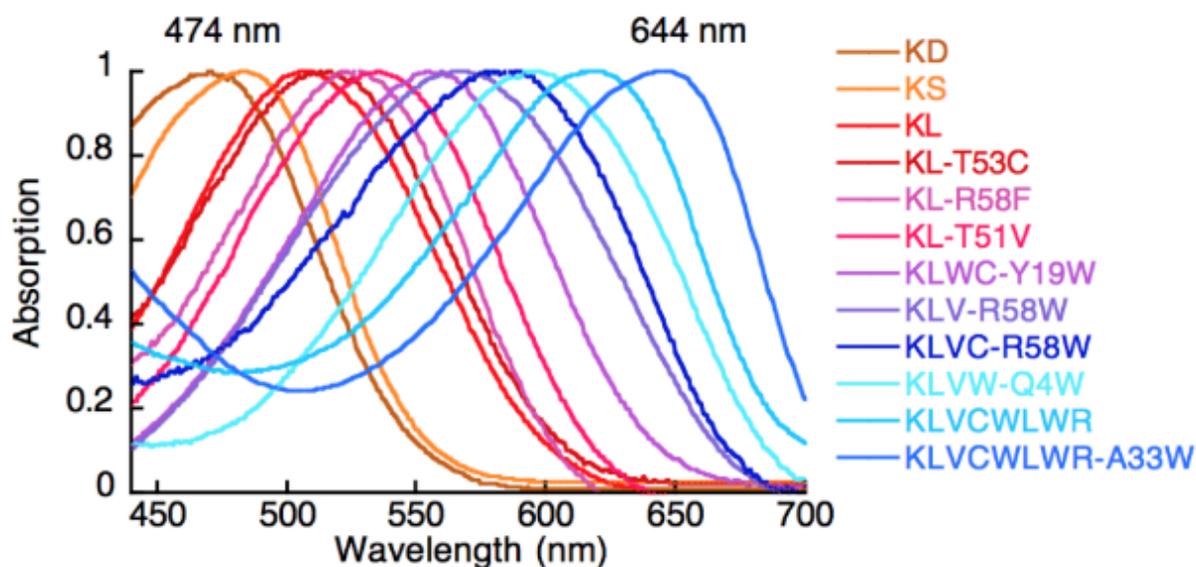


Figure I-34: Absorption spectra for *hCRBPII* mutants bound to retinal.^{87 88}

Up to this point, we reviewed the initial work on wavelength tuning of hCRBP. In the next part, I will discuss our recent discoveries of the key factors that affect the bathochromic shift of the system, especially the role of water molecules.

I-7 ROLE OF WATER MOLECULES IN WAVELENGTH REGULATION OF HCRBP

I-7-1 Q108K; K40L as a reference

As mentioned, we started with the KL template obtained by Dr. Meisam Nosrati and Dr. Zahra Nossoni. (Figure I-35) demonstrates the important residues and water molecules in KL structure that affect the wavelength tuning.

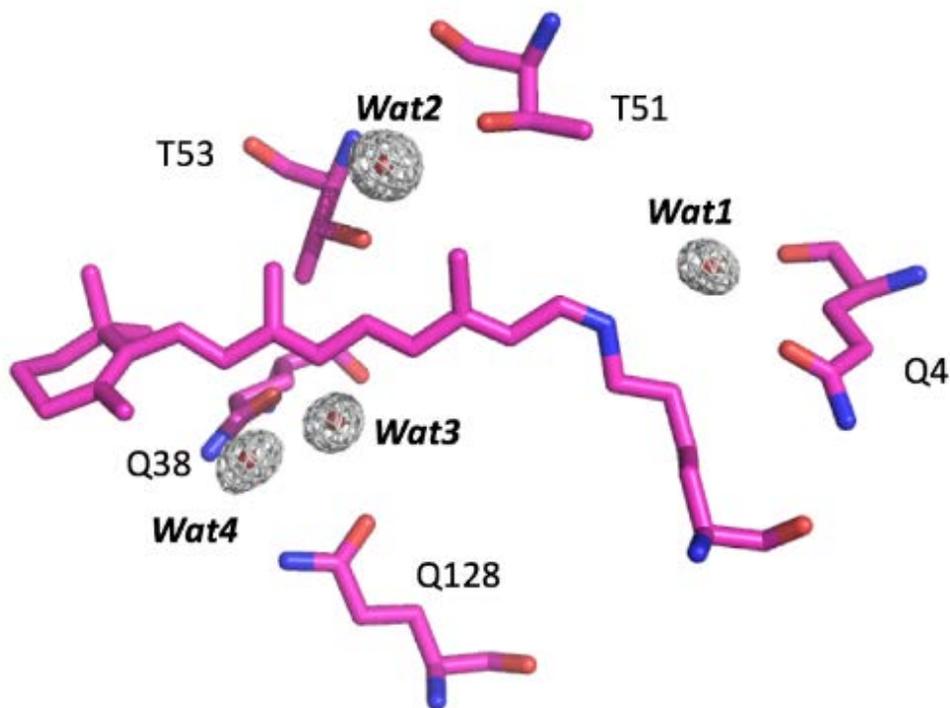


Figure I-35: Water molecules and residues around the retinal (5 angstrom) that have effect on wavelength tuning (PDB:4RUU).⁸⁰

KL crystalized in the P1 space group, and it has two molecules in the asymmetric unit. Conformation and orientation of the Schiff base are different in these two molecules. In Molecule A, the retinal is ordered with reasonable electron density and shows a pi-cation interaction between the iminium nitrogen and Trp106. In Molecule B, only first five atoms are ordered, and the PSB is pointed toward to protein environment (**Figure I-36**). Also, comparison between two KL structures (pdb:4RUU and pdb:4EXZ) shows that the whole chromophore is rotated perpendicular. Since, the protein does not hold the chromophore in the distinct conformation, wavelength tuning is not driven by confirmation of the chromophore, but by electrostatics of proteins.

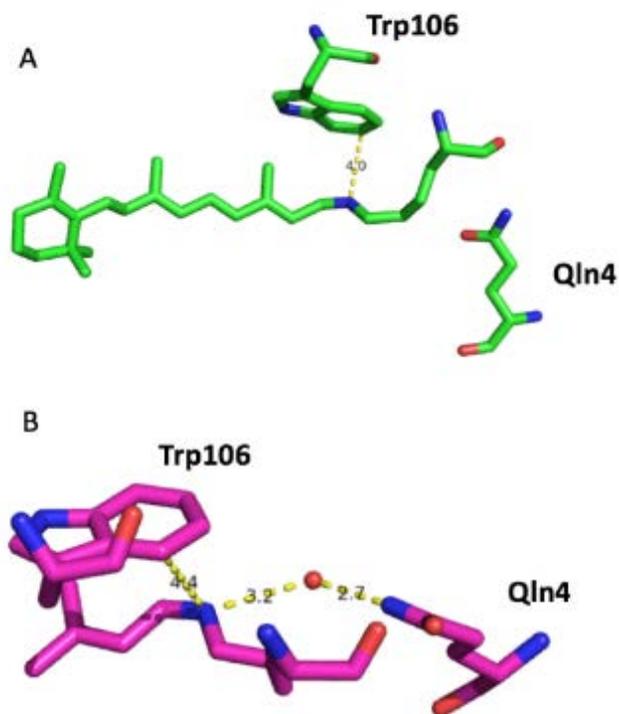


Figure I-36: KL structure (PDB:4RUU) A) mol A B)mol B. The relative orientation of Schiff base is different in the two molecules in the asymmetric unit..

I-7-2 Introducing the negative counterions to the PSB region (Blue shifted spectra)

To reach the blue-shifted spectra, negatively charged residues were added to the PSB region (**Figure I-37**). Similar to rhodopsin systems, by introducing new negative charges in the vicinity of the chromophore, the positive charge on the PSB becomes more localized, resulting in the more blue-shifted spectra.⁸⁷

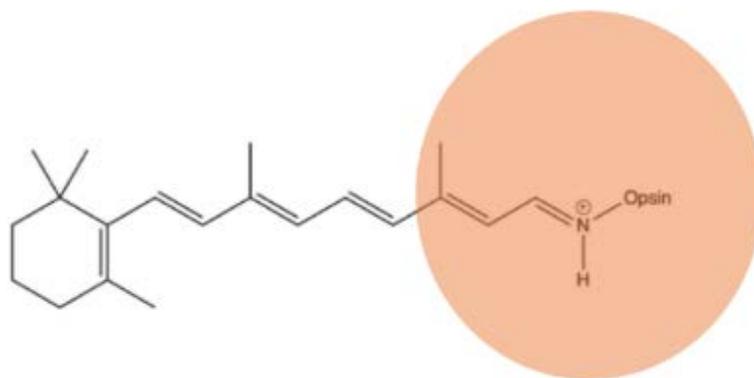


Figure I-37: Schiff base region.

The T51D mutation leads to lead to 32nm blue-shifted absorption spectra and. The domain swapped dimer structure of this mutant is obtained by Dr. Alireza Ghanbarpour and Dr. Meisam Nosrati. Glutamic acid is closer to the PSB of the chromophore compare to Thr51, and it leads to stabilizing the positive charge on PSB (**Figure I-38**).^{89 80} The conformation of the rest of the protein environment stays the same by T51D mutation. Introducing more negatively charged residue like L117E in this region also resulted in further blue shifting (**Table I-1**).

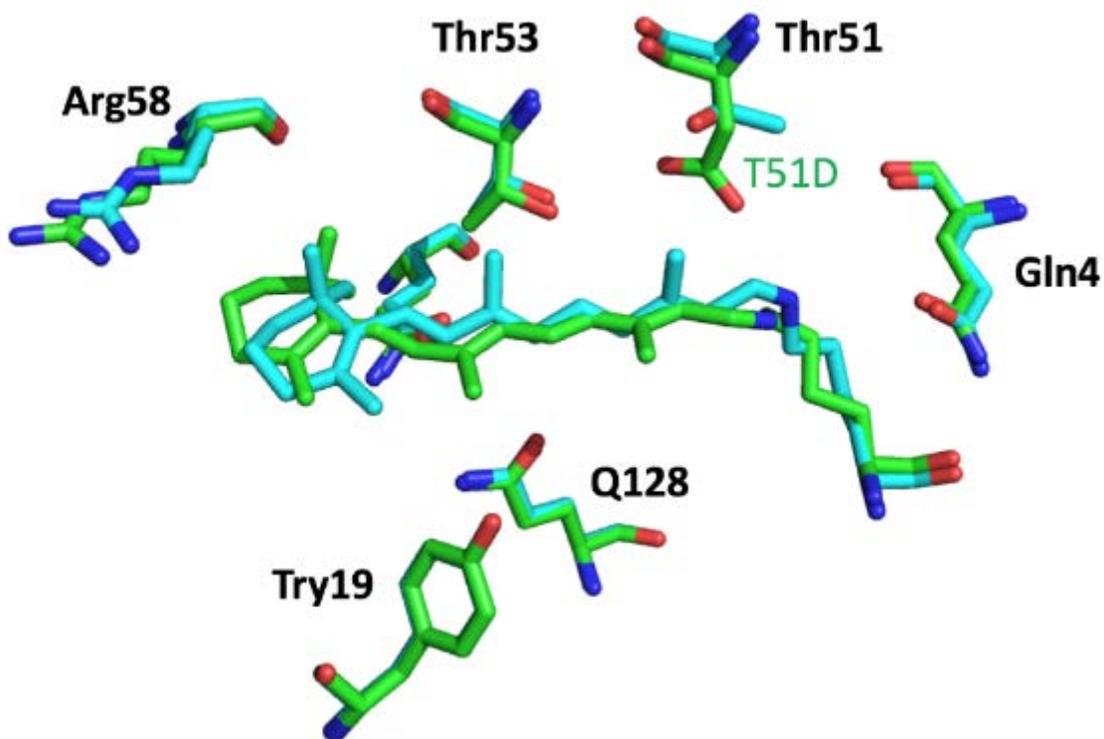


Figure I-38: Overlaid structures of KL (4RUU) with KL; T51D (6E5S). Most of the structures are similar, except the T51D is 1 angstrom closer to the PSB region.

Table I-1: Comparison between the absorption of KL, Q108K;T51D, and Q108K;K40L;L117E.

| Mutation | Absorption |
|------------------|------------|
| Q108K;K40L | 508nm |
| Q108K;T51D | 474nm |
| Q108K;T51D;L117E | 425nm |

I-7-3 Removing water molecule 2

As discussed for rhodopsin systems, change in the electrostatic interaction of the protein environment, as well as water network in the vicinity of the chromophore, affects the wavelength

tuning of the system.⁴³ In an effort to reach more red-shifted spectra in the hCRBP_{II} system, mutating both Thr51 and Thr53 resulted in a dramatic redshift by removing the water molecule between these two residues. Mutating Thr51 seems to have more effect in this absorption change. In KL: T53V structure, there is no significant change in the absorption and structure. However, KL: T51V resulted in removing the water molecule, and red-shifted spectra (**Figure I-39**).

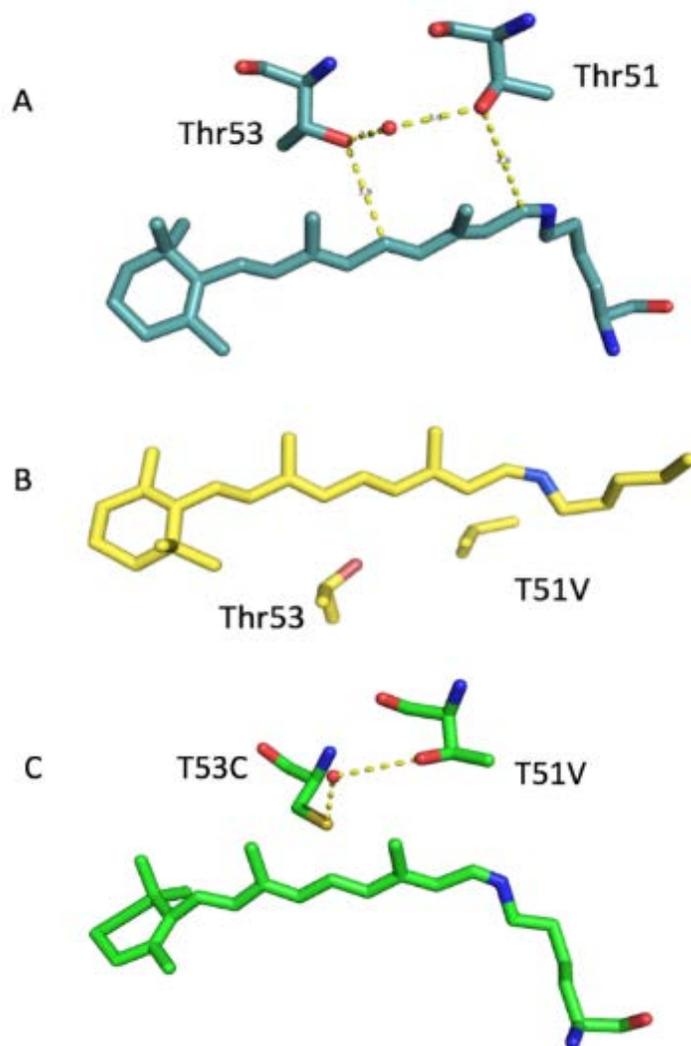


Figure I-39: A) water network between Thr53 and Thr51 in KL (PDB:4RUU). B) KLT51V (PDB:5FAZ), removal of water molecule. C) KLT53C, there is no change in water molecule.

I-7-4 Introducing the bulky hydrophobic residues in the ionone region

As discussed before, to reach further red-shifted absorption, more negative charges were introduced in the vicinity of the ionone region of the chromophore to delocalize the positive charge of the PSB (**Figure I-40**). Unlike what expected, the resulted absorption was blue shifted.

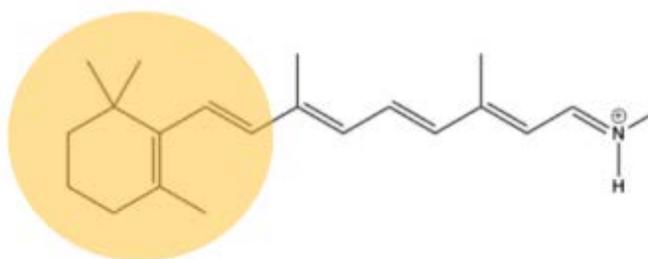


Figure I-40: Ionone region.

The KL:R58E structure is illustrated in (**Figure I-41**). The absorption of this construct blue shifted to 500nm. The reason behind this effect might be the change in position of the residue which is more exposed to the solvent.

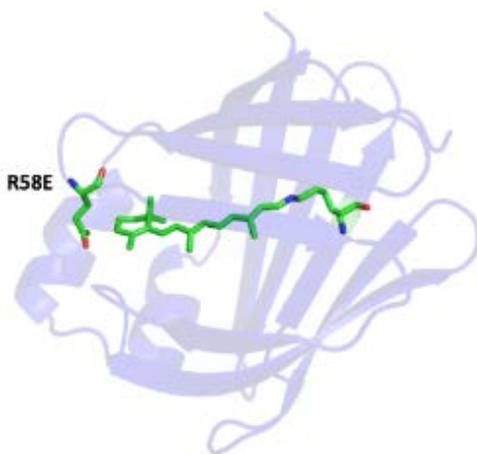


Figure I-41: KL;R58E. Glu58 is immersed through the solvent.

However, introducing bulky hydrophobic residues like Phe, Tyr, or Trp in this position lead to a large redshift. In most of the structures without having 58 mutated to a hydrophobic residue, there is at least one water molecule near the ionone ring of the chromophore. For example, two water molecules near the ionone are observed in the KLT53C structure (**Figure I-42**).⁸⁷

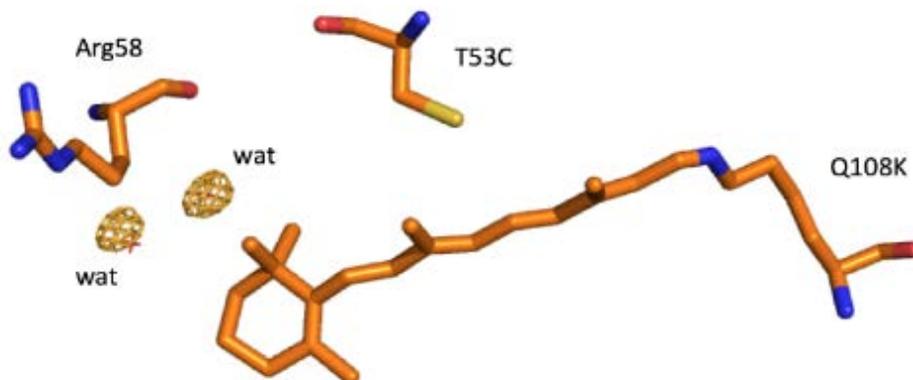


Figure I-42: KLT53C. There are two water molecules near to the ionone ring.

By adding bulky hydrophobic residues to 58, especially 58W, no ordered water molecules are observed near the ionone ring (**Figure I-43**).

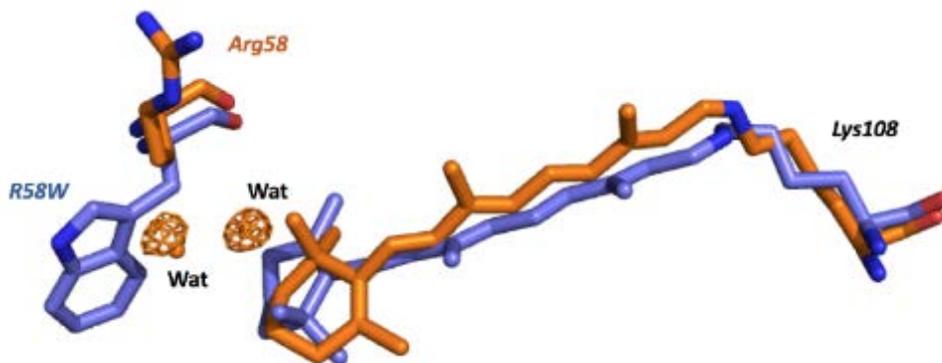


Figure I-43: Overlaid structure of KLT53C and KL. Removal of water molecules near to the ionone ring by addition of R58W mutation.

Hydrophobicity of this region inhibits the exposure of the binding cavity to the solvent. This can help to increase the effect of electrostatic interactions in the binding pocket. Another residue in this region, which has a significant effect on the absorption is Ala33. Mutating this residue to bulky hydrophobic residues, especially tryptophan, resulted in increasing the hydrophobicity of the entrance of the binding pocket and red-shifted absorption. In most of the cases, large red shifts in the absorption are observed for the constructs with the present of R58W mutation.

I-7-5 Mutating Gln4 (removing water molecule 1)

Previously, we explained that mutating Gln4 resulted in removing the water molecule 1 from the vicinity of the Schiff base and leading to a red-shifted absorption. However, since all of the structures with this mutant gave a *trans* conformation (different than other crystals which have *cis* conformation), we could not explain the resulting change in spectra by removal the of the water molecule.⁸⁷ The nitrogen of the Schiff base seems to have interaction with hydrophobic residues of the binding pocket. One of the examples is adding the Q4A mutant to KL; T51V;53C; R58W; T29L; Y19W lead to around 22nm red-shifted absorption. The color of the crystal, in this case, was light blue with *trans* conformation for the chromophore (**Figure I-44**). The hydrophobic residues around the Schiff base lead to lower pKa of the chromophore.

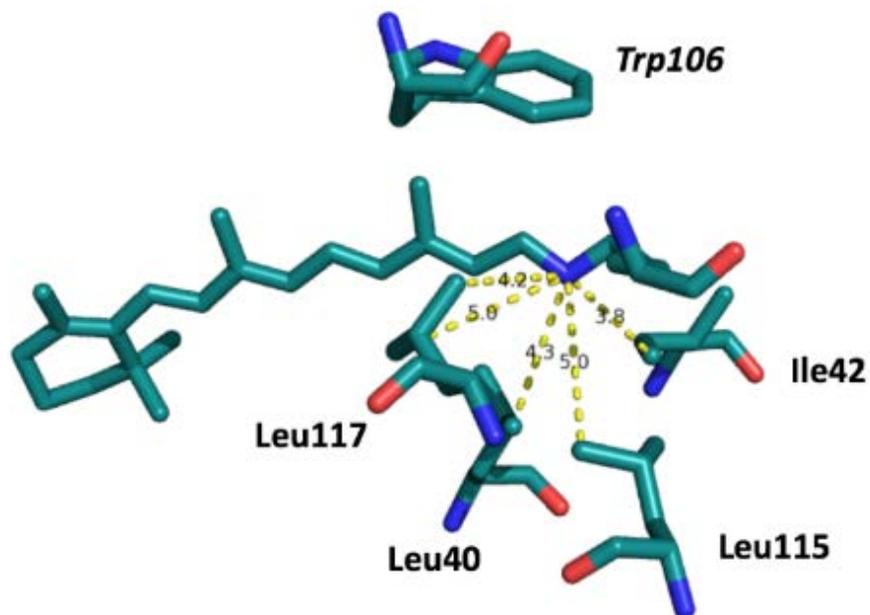


Figure I-44: Shows hydrophobic interactions with the nitrogen in all trans retinal bound to *KLVCWLWQ4A hCRBP II*.

Further trials for obtaining the *cis* structure will be discussed in the next chapter. UV exposure to the crystal isomerizes the trans conformation to *cis* resulting in the dark blue crystal consistent with 612nm absorption. In this structure, the positive charge on the nitrogen of PSB is having a pi-cation interaction with Trp106 leading to a higher pKa environment for the chromophore (**Figure I-45**).

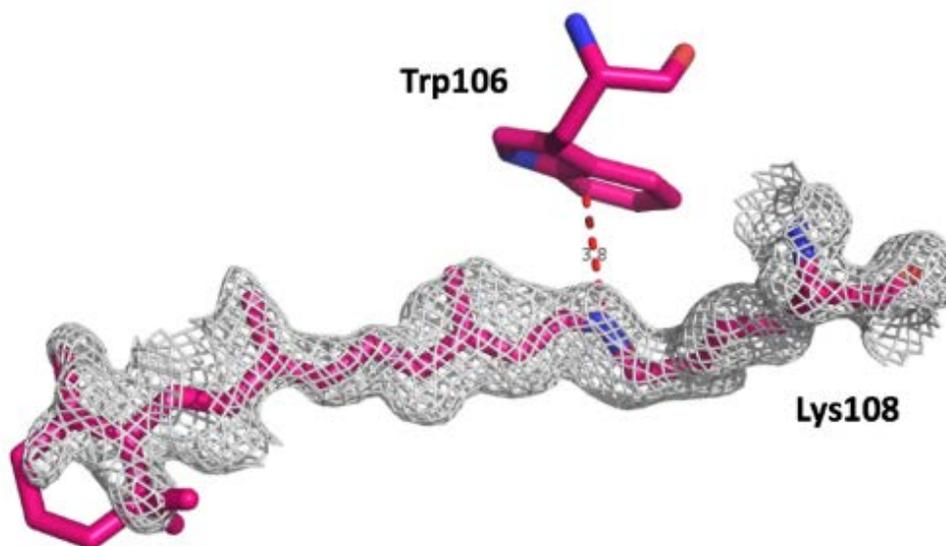


Figure I-45: 15-cis retinal bound to KLVCWLWQ4A hCRBPII. Positive charge on cis iminium stabilized through pi-cation interaction with Trp106.

I-7-6 Removing the water network between Q38 and Q128 (Water 3 and 4)

Gln38 and Gln128 are other residues that have an effect on the electrostatic interaction of the binding pocket. Gln38 and Gln128 have hydrogen bonds with two water molecules (**Figure 1-46**). The existence of these water networks near to the end of polyene and ionone ring helps with the polarity at the end of the chromophore and helping on the delocalization of the charge on the polyene.⁸⁷ Mutations of these residues lead to removing the water network leading to blue-shifted absorption spectra. As an example, the crystal structure of KL:T51V:R58Y:Y19W:Q38L construct demonstrates the removal of the water molecules leading to blue shifting compared to KL:T51V:R58Y:Y19W (**Figure I-47**).

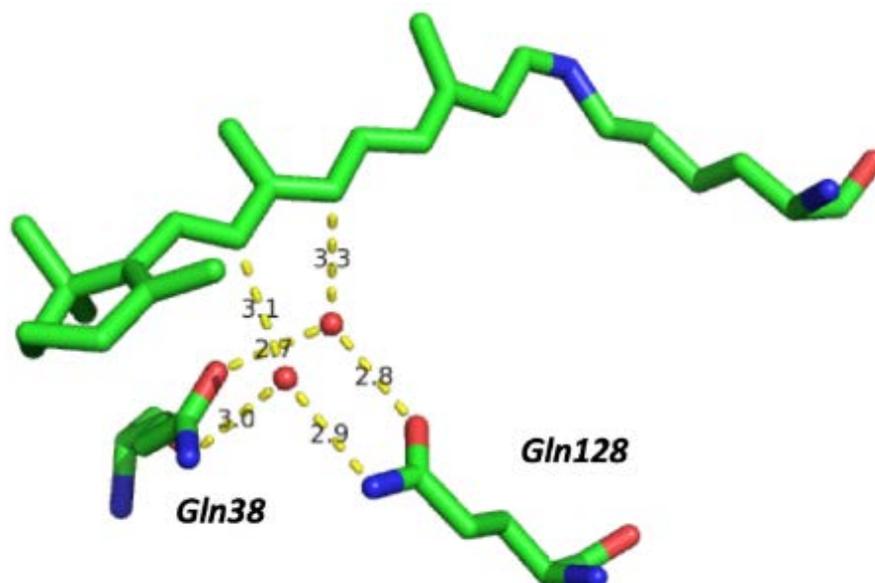


Figure I-46: Hydrogen bonding between Gln38 and Gln128 and two water molecules in KL structure (PDB:4RUU).

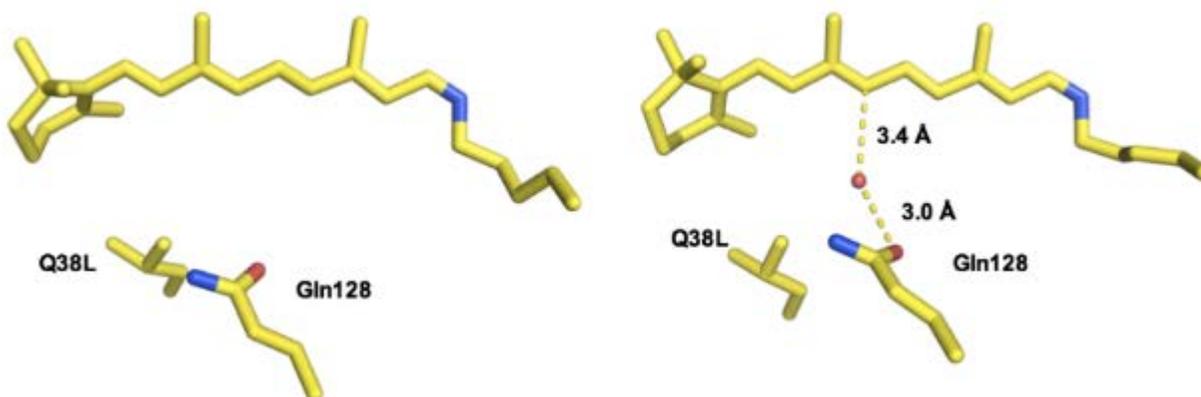


Figure I-47: Gln38-Gln128 Water molecules in the two chains of crystal structure of KL:T51V:R58Y:Y19W:Q38L (PDB:

I-8 CONCLUSION

Rhodopsins are photoreceptor systems and they have significant roles in living systems. These systems work based on the photo-isomerization of the retinal chromophore, concomitant changes in Schiff base pK_a and wavelength tuning. Our group was successful in mimicking wavelength tuning in a rhodopsin mimic template called hCRBP_{II}. In this chapter, we confirmed the effect of the residues and water molecules in the vicinity of the chromophore on wavelength tuning of this template.

I-9 EXPERIMENTAL

I-9-1 Site-Directed Mutagenesis

Please refer to Dr. Wenjing Wang's thesis.

I-9-2 Material and Methods: Protein Expression and Purification (bacterial expression)

The hCRBP_{II} gene was purchased from ATCC and cloned into the pET17b vector by Dr. Wenjing Wang. The NdeI and XhoI restriction enzyme used as a cutting site in the N- and C terminus of the vector. For the transformation step, 1 μ L of the plasmid was transformed in 50 μ L of *E. coli* DH5 α competent cells from Novagen® company. The cells were incubated for 30 min in ice and heat-shocked at 42°C for 45 seconds, then 400 μ L of Luria- Bertani broth was added, and the mixture was incubated at 37°C for two to three hours. After that, the cells were spread on an LB agar plate treated with ampicillin and incubated at 37°C for 16-20 hours. A single colony from the

plate was added to 10 mL of LB media contains 100 µg/mL ampicillin. The cell culture was grown for 12-16 hours at 37°C; then, the media were centrifuged at 14000 rpm for 1 min. DNA extraction and isolation from the cell pellet was done according to the manufacturer's instructions from Promega Wizard and SV Miniprep (A1330) DNA purification kit. For protein expression, the construct was transformed into BL2 *E. coli* competent cells from Invitrogen company and spread on the plate treated with ampicillin. A single colony was picked from the plate and transferred into 1L of LB media with ampicillin (100 mg/L) and incubated at 37°C until OD600 reached 0.7-1.0. 1 mM isopropyl β-D-1 thiogalactopyranoside (IPTG) from Gold Biotechnology was induced to the cells, and the cells incubate overnight at 25°C in the shaker. The transfected cells were harvested by centrifugation at 5000 rpm for 20 min. The cells were resuspended in lysate buffer containing 10 mM Tris, 10 mM NaCl pH 8.0, 50mL. The suspended cells were lysed by sonication, and the lysed cells were centrifuged at 4°C in 14,000 rpm for 20 min. For purification, the solution of the protein passed through Q Sepharose Fast Flow resin (GE Health Sciences), and the bound protein was eluted with 10 mM Tris, 150 mM NaCl, pH 8.0. The purity of the elution was checked with SDS PAGE and the pure fractions desalted by dialysis against 10 mM Tris pH=8.0 buffer. The desalted mixture was then loaded on a 15Q anion exchange column (GE Health Sciences), using the program described in (**Table I-2**).^{88 86}

Table I-2: Anion Exchange purification protocol for hCRBPII, adjusted with 50 mM Tris, pH 8 buffer. The proteins elute between 4% -8 % 2M NaCl.

| Description | | Parameters |
|-----------------|----------------------|-----------------------|
| Isocratic flow | pH=8.1, 0% 2M NaCl | 10.00 ml, 3.00 ml/min |
| Linear Gradient | pH=8.1, 0-4% 2M NaCl | 20.00 ml, 3.00 ml/min |
| Isocratic flow | pH=8.1, 4% 2M NaCl | 20.00 ml, 3.00 ml/min |
| Linear Gradient | pH=8.1, 4-8% 2M NaCl | 10.00 ml, 3.00 ml/min |
| Isocratic flow | pH=8.1, 8% 2M NaCl | 20.00 ml, 3.00 ml/min |

I-9-3 Crystallization and refinement

The concentrated, purified protein concentrated at 12-15 mg/ml and four equivalents of retinal added to the solution in the dark. The mixture sits at room temperature until maximum PSB formation, usually after 1-2 hours. Vapor diffusion crystallization was performed using a 24 well crystallization plate. The reservoir solution of 1 ml contains 25 % PEG 4000 and 0.1 M Ammonium Acetate and 0.1 M Sodium Acetate at pH range 4-4.8. Crystals usually appeared after 1-2 days at room temperature. Crystals were flash-frozen in a solution containing 18-20 % glycerol and the mother liquor in the reservoir. Data collection was through the Advanced Photon Source (APS) at (Argonne National Laboratory IL) LS-CAT (sector 21-ID-D, beamline D, F, or G) using a MAR300 detector for G and F beamline and 1.00Å wavelength radiation at 100K. For

refinement, the diffraction data were indexed, processed, and scaled using the HKL2000 program.¹⁵ The structure was obtained and solved by molecular replacement in PHASER in PHENIX and Q108K: K40L hCRBP_{II} (PDB: 4AZR) as a search model. The electron density map was produced by PHENIX. Refinement followed by model rebuilding and water and ligand replacements in COOT.^{90 91 92}

I-9-4 Tips for growing crystals from saturated solution

The higher concentration of proteins can lead to a saturated mixture of the protein with oiling out after one day. Typically, in these situations, if the proteins aggregated, the addition of salt (such as NaCl) to the protein solution may help. Lowering the temperature or lowering the amount of precipitant may help to prevent the formation of oily drops. Also, this saturated solution (oily drops) may use as seed for the next trials. In hCRBP_{II} crystals, I usually face this problem, forming oily drops instead of ordered crystals. In the hCRBP_{II} case, spreading this oily drops after one day helps in forming well-ordered crystals. I was successful in growing different mutants of this protein bound to retinal and FR1 (Fluorophore explained in Chapter III) with this method (**Figure I-48**) (**Table I-3, I-4, I-5, I-6, I-7**).

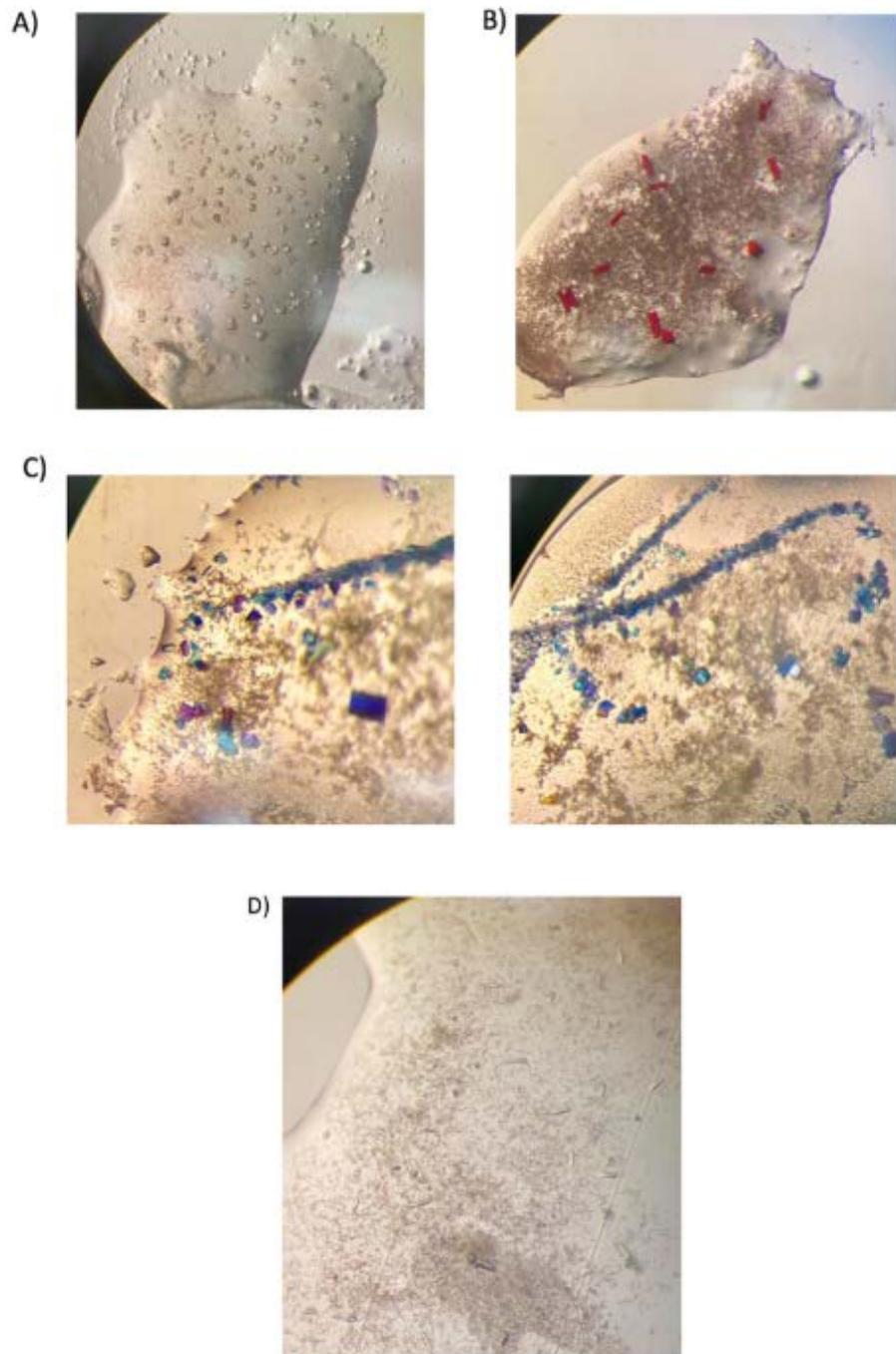


Figure I-48: *hCRBP II bound to retinal A) KL; T51V;53C; R58W; T29L; Y19W; Q4A B)KL; T51V; T53C. hCRBP II bound to FR-1 C) KL; T51V: T53S D) KL;T51V; T53S; R58W; T29L;Q38L; Q128L; Y19W.*

Table I-3: X-ray crystallographic data and refinement statistics for KL; T53C hCRBP II bound to retinal.

| | kIT53C |
|---------------------------------------|--|
| Wavelength | |
| Resolution range | 28.8 - 1.55 (1.605 - 1.55) |
| Space group | P 1 |
| Unit cell | 29.947 35.89 63.866 91.01 92.23 113.59 |
| Total reflections | |
| Unique reflections | 32793 (2523) |
| Multiplicity | |
| Completeness (%) | 92.79 (71.33) |
| Mean I/sigma(I) | |
| Wilson B-factor | 20.16 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 32791 (2523) |
| Reflections used for R-free | 1662 (130) |
| R-work | 0.1931 (0.2534) |
| R-free | 0.2346 (0.3125) |
| Number of non-hydrogen atoms | 2477 |
| macromolecules | 2204 |
| ligands | 48 |
| solvent | 225 |
| Protein residues | 266 |
| RMS(bonds) | 0.012 |
| RMS(angles) | 1.07 |

Table I-4: X-ray crystallographic data and refinement statistics for KL; R58E hCRBP II bound to retinal.

| | kl58E |
|---------------------------------------|---|
| Wavelength | |
| Resolution range | 31.89 - 1.8 (1.865 - 1.8) |
| Space group | P 1 |
| Unit cell | 29.624 36.107 63.864 90.836 87.593 66.833 |
| Total reflections | |
| Unique reflections | 18017 (796) |
| Multiplicity | |
| Completeness (%) | 79.99 (35.67) |
| Mean I/sigma(I) | |
| Wilson B-factor | 35.34 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 17989 (794) |
| Reflections used for R-free | 913 (39) |
| R-work | 0.2096 (0.3179) |
| R-free | 0.2577 (0.3898) |
| Number of non-hydrogen atoms | 2291 |
| macromolecules | 2184 |
| ligands | 8 |
| solvent | 99 |
| Protein residues | 266 |
| RMS(bonds) | 0.008 |
| RMS(angles) | 0.86 |

Table I-5: X-ray crystallographic data and refinement statistics for KL; T51V hCRBPII bound to retinal.

| | KL;T51V |
|---------------------------------------|---|
| Wavelength | |
| Resolution range | 32.83 - 1.4 (1.45 - 1.4) |
| Space group | P 1 |
| Unit cell | 29.898 36.323 63.987 87.166 92.478 115.28 |
| Total reflections | |
| Unique reflections | 45723 (4507) |
| Multiplicity | |
| Completeness (%) | 95.47 (93.72) |
| Mean I/sigma(I) | |
| Wilson B-factor | 18.59 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 45721 (4507) |
| Reflections used for R-free | 2324 (212) |
| R-work | 0.2120 (0.3101) |
| R-free | 0.2486 (0.3154) |
| Number of non-hydrogen atoms | 2542 |
| macromolecules | 2202 |
| ligands | 48 |
| solvent | 292 |
| Protein residues | 266 |
| RMS(bonds) | 0.009 |
| RMS(angles) | 1.28 |

Table I-6: X-ray crystallographic data and refinement statistics for KL; T51V;53C; R58W; T29L; Y19W; Q4A hCRBPII bound to all trans retinal incubation in the dark.

| | Q4ADRAKBLUETRANS(DARK) |
|---------------------------------------|--|
| Wavelength | |
| Resolution range | 27.66 - 1.4 (1.45 - 1.4) |
| Space group | P 1 |
| Unit cell | 30.842 35.918 64.062 85.95 86.434 65.334 |
| Total reflections | |
| Unique reflections | 41917 (4314) |
| Multiplicity | |
| Completeness (%) | 85.44 (87.77) |
| Mean I/sigma(I) | |
| Wilson B-factor | 12.34 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 41915 (4314) |
| Reflections used for R-free | 2059 (206) |
| R-work | 0.1859 (0.2141) |
| R-free | 0.2085 (0.2534) |
| Number of non-hydrogen atoms | 2535 |
| macromolecules | 2188 |
| ligands | 44 |
| solvent | 303 |
| Protein residues | 266 |
| RMS(bonds) | 0.006 |
| RMS(angles) | 0.96 |

Table I-7: X-ray crystallographic data and refinement statistics for KL;T51V;53C;R58W;T29L;Y19W;Q4A hCRBP II bound to cis-retinal after UV irradiation.

| | CISQ4A |
|---------------------------------------|---|
| Wavelength | |
| Resolution range | 21.28 - 1.261 (1.306 - 1.261) |
| Space group | P 1 |
| Unit cell | 30.923 36.038 64.026 86.203 86.555 65.059 |
| Total reflections | |
| Unique reflections | 58166 (5523) |
| Multiplicity | |
| Completeness (%) | 86.21 (81.77) |
| Mean I/sigma(I) | |
| Wilson B-factor | 13.07 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 58153 (5523) |
| Reflections used for R-free | 1959 (192) |
| R-work | 0.2042 (0.2444) |
| R-free | 0.2321 (0.2851) |
| Number of non-hydrogen atoms | 2513 |
| macromolecules | 2184 |
| ligands | 52 |
| solvent | 277 |
| Protein residues | 266 |
| RMS(bonds) | 0.006 |
| RMS(angles) | 0.97 |

REFERENCES

REFERENCES

1. Palczewski, K. Chemistry and biology of vision. *J. Biol. Chem.* **287**, 1612–1619 (2012).
2. Papin, J. A., Hunter, T., Palsson, B. O. & Subramaniam, S. Reconstruction of cellular signalling networks and analysis of their properties. *Nature Reviews Molecular Cell Biology* **6**, 99–111 (2005).
3. Berson, D. M. Phototransduction in ganglion-cell photoreceptors. *Pflugers Arch - Eur J Physiol* **454**, 849–855 (2007).
4. Shichida, Y. & Matsuyama, T. Evolution of opsins and phototransduction. *Philos Trans R Soc Lond B Biol Sci* **364**, 2881–2895 (2009).
5. Heck, M., Hofmann, K. P., Kraft, T. W. & Lamb, T. D. Phototransduction gain at the G-protein, transducin, and effector protein, phosphodiesterase-6, stages in retinal rods. *PNAS* **116**, 8653–8654 (2019).
6. Rando, R. R. Polyenes and vision. *Chemistry & Biology* **3**, 255–262 (1996).
7. Bowmaker, J. K. & Dartnall, H. J. Visual pigments of rods and cones in a human retina. *The Journal of Physiology* **298**, 501–511 (1980).
8. Lamb, T. D. Evolution of the genes mediating phototransduction in rod and cone photoreceptors. *Progress in Retinal and Eye Research* 100823 (2019)
doi:10.1016/j.preteyeres.2019.100823.
9. Molday, R. S. Photoreceptor membrane proteins, phototransduction, and retinal degenerative diseases. The Friedenwald Lecture. *Invest. Ophthalmol. Vis. Sci.* **39**, 2491–2513 (1998).
10. Kandori, H. Biophysics of rhodopsins and optogenetics. *Biophys Rev* (2020)
doi:10.1007/s12551-020-00645-0.
11. Govorunova, E. G., Sineshchekov, O. A., Li, H. & Spudich, J. L. Microbial Rhodopsins: Diversity, Mechanisms, and Optogenetic Applications. *Annu. Rev. Biochem.* **86**, 845–872 (2017).
12. Tsai, C.-J. *et al.* Crystal structure of rhodopsin in complex with a mini-Go sheds light on the principles of G protein selectivity. *Science Advances* **4**, eaat7052 (2018).
13. Filipek, S., Stenkamp, R. E., Teller, D. C. & Palczewski, K. G protein-coupled receptor rhodopsin: a prospectus. *Annu. Rev. Physiol.* **65**, 851–879 (2003).

14. Leung, N. Y. & Montell, C. Unconventional roles of opsins. *Annu Rev Cell Dev Biol* **33**, 241–264 (2017).
15. Palczewski, K. *et al.* Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science* **289**, 739–745 (2000).
16. Kim, Y.-K. *et al.* β -Carotene and its cleavage enzyme β -carotene-15,15'-oxygenase (CMOI) affect retinoid metabolism in developing tissues. *FASEB J* **25**, 1641–1652 (2011).
17. Wald, G. & Brown, P. K. Synthesis and Bleaching of Rhodopsin. *Nature* **177**, 174–176 (1956).
18. Ernst, O. P. *et al.* Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. *Chem. Rev.* **114**, 126–163 (2014).
19. Rhodopsin: A Structural Primer for G-Protein Coupled Receptors - Stenkamp - 2005 - Archiv der Pharmazie - Wiley Online Library.
<https://onlinelibrary.wiley.com/doi/abs/10.1002/ardp.200400995>.
20. Pittolo, S. *et al.* An allosteric modulator to control endogenous G protein-coupled receptors with light. *Nat. Chem. Biol.* **10**, 813–815 (2014).
21. Kong, Y. & Karplus, M. The Signaling Pathway of Rhodopsin. *Structure* **15**, 611–623 (2007).
22. Palczewski, K. G Protein–Coupled Receptor Rhodopsin. *Annu Rev Biochem* **75**, 743–767 (2006).
23. Gao, Y. *et al.* Structures of the Rhodopsin-Transducin Complex: Insights into G-Protein Activation. *Molecular Cell* **75**, 781-790.e3 (2019).
24. Tsai, C.-J. *et al.* Cryo-EM structure of the rhodopsin-G α i- β γ complex reveals binding of the rhodopsin C-terminal tail to the g β subunit. *eLife* **8**,.
25. Venkatakrisnan, A. J. *et al.* Diverse activation pathways in class A GPCRs converge near the G-protein-coupling region. *Nature* **536**, 484–487 (2016).
26. Venkatakrisnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185–194 (2013).
27. Kochendoerfer, G. G., Lin, S. W., Sakmar, T. P. & Mathies, R. A. How color visual pigments are tuned. *Trends in Biochemical Sciences* **24**, 300–305 (1999).
28. Kakitani, H., Kakitani, T., Rodman, H. & Honig, B. On the mechanism of wavelength regulation in visual pigments. *Photochem. Photobiol.* **41**, 471–479 (1985).

29. Probing and Modeling the Absorption of Retinal Protein Chromophores in Vacuo - Rajput - 2010 - Angewandte Chemie International Edition - Wiley Online Library. <https://onlinelibrary.wiley.com/doi/full/10.1002/anie.200905061>.
30. Zhou, X., Sundholm, D., Wesolowski, T. A. & Kaila, V. R. I. Spectral Tuning of Rhodopsin and Visual Cone Pigments. *J. Am. Chem. Soc.* **136**, 2723–2726 (2014).
31. Fitch, C. A. *et al.* Experimental pKa Values of Buried Residues: Analysis with Continuum Methods and Role of Water Penetration. *Biophysical Journal* **82**, 3289–3304 (2002).
32. Miyahara, T. & Nakatsuji, H. Light-Driven Proton, Sodium Ion, and Chloride Ion Transfer Mechanisms in Rhodopsins: SAC-CI Study. *J Phys Chem A* **123**, 1766–1784 (2019).
33. Lee, K. S. S. *et al.* Probing Wavelength Regulation with an Engineered Rhodopsin Mimic and a C15-Retinal Analogue. *ChemPlusChem* **77**, 273–276 (2012).
34. Han, M., DeDecker, B. S. & Smith, S. O. Localization of the retinal protonated Schiff base counterion in rhodopsin. *Biophys. J.* **65**, 899–906 (1993).
35. Mathies, R. & Stryer, L. Retinal has a highly dipolar vertically excited singlet state: implications for vision. *PNAS* **73**, 2169–2173 (1976).
36. Xie, P., Zhou, P., Alsaedi, A. & Zhang, Y. pH-dependent absorption spectra of rhodopsin mutant E113Q: On the role of counterions and protein. *Spectrochim Acta A Mol Biomol Spectrosc* **174**, 25–31 (2017).
37. Foster, K. W. *et al.* A rhodopsin is the functional photoreceptor for phototaxis in the unicellular eukaryote *Chlamydomonas*. *Nature* **311**, 756–759 (1984).
38. Kloppmann, E., Becker, T. & Ullmann, G. M. Electrostatic potential at the retinal of three archaeal rhodopsins: Implications for their different absorption spectra. *Proteins: Structure, Function, and Bioinformatics* **61**, 953–965 (2005).
39. Luecke, H., Schobert, B., Lanyi, 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* **293**, 1499–1503 (2001).
40. Lesca, E., Panneels, V. & Schertler, G. F. X. The role of water molecules in phototransduction of retinal proteins and G protein-coupled receptors. *Faraday Discuss.* **207**, 27–37 (2018).
41. Sun, X., Ågren, H. & Tu, Y. Functional water molecules in rhodopsin activation. *J Phys Chem B* **118**, 10863–10873 (2014).

42. Role of Bulk Water in Hydrolysis of the Rhodopsin Chromophore.
<https://www.jbc.org/content/286/21/18930.long>.
43. Okada, T. *et al.* Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5982–5987 (2002).
44. Okada, T. *et al.* The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **342**, 571–583 (2004).
45. Terakita, A., Yamashita, T. & Shichida, Y. Highly conserved glutamic acid in the extracellular IV-V loop in rhodopsins acts as the counterion in retinochrome, a member of the rhodopsin family. *Proceedings of the National Academy of Sciences* **97**, 14263–14267 (2000).
46. Sekharan, S. Water-mediated spectral shifts in rhodopsin and bathorhodopsin. *Photochem. Photobiol.* **85**, 517–520 (2009).
47. Choe, H.-W. *et al.* Crystal structure of metarhodopsin II. *Nature* **471**, 651–655 (2011).
48. Nango, E. *et al.* A three-dimensional movie of structural changes in bacteriorhodopsin. *Science* **354**, 1552–1557 (2016).
49. Schotte, F. *et al.* Watching a signaling protein function in real time via 100-ps time-resolved Laue crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19256–19261 (2012).
50. Murakami, M. & Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **453**, 363–367 (2008).
51. Crystal structure of jumping spider rhodopsin-1 as a light sensitive GPCR | PNAS.
<https://www.pnas.org/content/116/29/14547>.
52. Kato, H. E. *et al.* Crystal structure of the channelrhodopsin light-gated cation channel. *Nature* **482**, 369–374 (2012).
53. Shapes, Dynamics, and Stability of β -Ionone and Its Two Mutants Evidenced by High-Resolution Spectroscopy in the Gas Phase | The Journal of Physical Chemistry Letters.
<https://pubs.acs.org/doi/abs/10.1021/acs.jpcllett.8b00256>.
54. Toker, Y. *et al.* Counterion-controlled spectral tuning of the protonated Schiff-base retinal. *Phys. Rev. A* **98**, 043428 (2018).
55. Red-Tuning of the Channelrhodopsin Spectrum Using Long Conjugated Retinal Analogues | Biochemistry. <https://pubs.acs.org/doi/10.1021/acs.biochem.8b00583>.
56. Suomivuori, C.-M., Lang, L., Sundholm, D., Gamiz-Hernandez, A. P. & Kaila, V. R. I. Tuning the Protein-Induced Absorption Shifts of Retinal in Engineered Rhodopsin Mimics. *Chem. Eur. J.* **22**, 8254–8261 (2016).

57. An Average Solvent Electrostatic Configuration Protocol for QM/MM Free Energy Optimization: Implementation and Application to Rhodopsin Systems | Journal of Chemical Theory and Computation. <https://pubs.acs.org/doi/abs/10.1021/acs.jctc.7b00860>.
58. Hayashi, F. *et al.* Raftophilic rhodopsin-clusters offer stochastic platforms for G protein signalling in retinal discs. *Communications Biology* **2**, 1–12 (2019).
59. Hayashi, S. *et al.* QM/MM Geometry Optimization on Extensive Free-Energy Surfaces for Examination of Enzymatic Reactions and Design of Novel Functional Properties of Proteins. *Annu Rev Phys Chem* **68**, 135–154 (2017).
60. Collette, F., Renger, T., Müh, F. & Schmidt am Busch, M. Red/Green Color Tuning of Visual Rhodopsins: Electrostatic Theory Provides a Quantitative Explanation. *J. Phys. Chem. B* **122**, 4828–4837 (2018).
61. Lee, C., Sekharan, S. & Mertz, B. Theoretical Insights into the Mechanism of Wavelength Regulation in Blue-Absorbing Proteorhodopsin. *J. Phys. Chem. B* **123**, 10631–10641 (2019).
62. Inoue, K. *et al.* Exploration of natural red-shifted rhodopsins using a machine learning-based Bayesian experimental design. *bioRxiv* 2020.04.21.052548 (2020) doi:10.1101/2020.04.21.052548.
63. Duñach, M., Marti, T., Khorana, H. G. & Rothschild, K. J. Uv-visible spectroscopy of bacteriorhodopsin mutants: substitution of Arg-82, Asp-85, Tyr-185, and Asp-212 results in abnormal light-dark adaptation. *Proc Natl Acad Sci U S A* **87**, 9873–9877 (1990).
64. Kaur, J. *et al.* Solid-state NMR analysis of the sodium pump *Krokinobacter* rhodopsin 2 and its H30A mutant. *Journal of Structural Biology* **206**, 55–65 (2019).
65. Nogly, P. *et al.* Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. *Science* **361**, (2018).
66. Seddon, E. A. *et al.* Short-wavelength free-electron laser sources and science: a review. *Rep. Prog. Phys.* **80**, 115901 (2017).
67. Ostrovsky, M. A. Rhodopsin: Evolution and comparative physiology. *Paleontol. J.* **51**, 562–572 (2017).
68. Berbasova, T. *et al.* Light-Activated Reversible Imine Isomerization: Towards a Photochromic Protein Switch. *ChemBiochem* **17**, 407–414 (2016).
69. Wang, W., Geiger, J. H. & Borhan, B. The photochemical determinants of color vision: revealing how opsins tune their chromophore's absorption wavelength. *Bioessays* **36**, 65–74 (2014).

70. Banaszak, L. *et al.* Lipid-Binding Proteins: A Family of Fatty Acid and Retinoid Transport Proteins. in *Advances in Protein Chemistry* (eds. Anfinsen, C. B., Edsall, J. T., Richards, F. M. & Eisenberg, D. S.) vol. 45 89–151 (Academic Press, 1994).
71. Veerkamp, J. H., Kuppevelt, T. H. M. S. M. van, Maatman, R. G. H. J. & Prinsen, C. F. M. Structural and functional aspects of cytosolic fatty acid-binding proteins. *Prostaglandins, Leukotrienes and Essential Fatty Acids* **49**, 887–906 (1993).
72. Identification of intracellular carriers for the endocannabinoid anandamide | PNAS. <https://www.pnas.org/content/106/15/6375>.
73. Lücke, C., Huang, S., Rademacher, M. & Rüterjans, H. New insights into intracellular lipid binding proteins: The role of buried water. *Protein Sci* **11**, 2382–2392 (2002).
74. Vaezeslami, S., Mathes, E., Vasileiou, C., Borhan, 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.* **363**, 687–701 (2006).
75. Protein design: reengineering cellular retinoic acid binding protein II into a rhodopsin protein mimic. - Abstract - Europe PMC. <https://europepmc.org/article/med/17447762>.
76. Vaezeslami, S., Jia, X., Vasileiou, C., Borhan, B. & Geiger, J. H. Structural analysis of site-directed mutants of cellular retinoic acid-binding protein II addresses the relationship between structural integrity and ligand binding. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 1228–1239 (2008).
77. Gunasekaran, K., Hagler, A. T. & Gierasch, L. M. Sequence and structural analysis of cellular retinoic acid-binding proteins reveals a network of conserved hydrophobic interactions. *Proteins* **54**, 179–194 (2004).
78. Franzoni, L. *et al.* Structure and Backbone Dynamics of Apo- and Holo-cellular Retinol-binding Protein in Solution. *J. Biol. Chem.* **277**, 21983–21997 (2002).
79. Berbasova, T. *et al.* Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone. *J. Am. Chem. Soc.* **135**, 16111–16119 (2013).
80. Nosrati, M. Development of photoswitchable rhodopsin mimics: Spectroscopy, structural studies and wavelength regulation studies of these systems. (Michigan State University, 2015).
81. Ghanbarpour, A. *et al.* Mimicking Microbial Rhodopsin Isomerization in a Single Crystal. *J. Am. Chem. Soc.* **141**, 1735–1741 (2019).
82. Nosrati, M., Berbasova, T., Vasileiou, C., Borhan, B. & Geiger, J. H. A Photoisomerizing Rhodopsin Mimic Observed at Atomic Resolution. *J. Am. Chem. Soc.* **138**, 8802–8808 (2016).

83. Vasileiou, C. *et al.* Dissection of the Critical Binding Determinants of Cellular Retinoic Acid Binding Protein II by Mutagenesis and Fluorescence Binding Assay. *Proteins* **76**, 281–290 (2009).
84. Yapici, I. *et al.* “Turn-On” Protein Fluorescence: In Situ Formation of Cyanine Dyes. *J. Am. Chem. Soc.* **137**, 1073–1080 (2015).
85. Nossoni, Z. *et al.* Structures of holo wild-type human cellular retinol-binding protein II (hCRBP II) bound to retinol and retinal. *Acta Cryst D, Acta Cryst Sect D, Acta Crystallogr D, Acta Crystallogr Sect D, Acta Crystallogr D Biol Crystallogr, Acta Crystallogr Sect D Biol Crystallogr* **70**, 3226–3232 (2014).
86. Nossoni, Z. Structural insight in to the mechanism of wavelength tuning in a rhodopsin mimic and a single mutation resulted an extensive 3D domain swapped dimerization in hCRBP II. (Michigan State University, 2014).
87. Wang, W. *et al.* Tuning the electronic absorption of protein-embedded all-trans-retinal. *Science* **338**, 1340–1343 (2012).
88. Wang, W. Protein design: Reengineering of cellular retinol binding protein II (CRBP II) into a rhodopsin mimic, functionalization of CRBP II into a fluorescent protein tag and design of a photoswitchable protein tag. (Michigan State University, 2012).
89. Ghanbarpour, A. Functional Control of Soluble Rhodopsin Mimics Using High Resolution Structure-Based Design and Evaluation. (Michigan State University, 2019).
90. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25 (2010).
91. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
92. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D* **66**, 213–221 (2010).

CHAPTER II: ISOMERIZATION IN HCRBP II

In the previous chapter, we learned about our two protein design templates. hCRABP II and hCRBP II, to study the rhodopsin systems.^{1 2 3 4} Mutational studies in the hCRBP II system helps us to understand the reasons behind the wavelength tuning.⁵ As we discussed, another important phenomena in photo cascade of the rhodopsin proteins is isomerization of a double bond in the chromophore. My former lab mates Dr. Meisam Nosrati and Dr. Alireza Ghanbarpour were successful to mimic the isomerization cascade of rhodopsins in the hCRABP II system.^{6 7} Recently, I detected the isomerization in the hCRBP II template as well, which helps us to find more information about the isomerization in these templates. In this chapter, we first review the photoisomerization in both animal and bacteriorhodopsin, also findings by my group about photoisomerization in hCRABP II system. Moreover, I will explain about the photoisomerization in hCRBP II system which was recently detected.

II-1 PHOTOISOMERIZATION CYCLE IN ANIMAL RHODOPSIN

After a photon of light hits rhodopsin, the chromophore goes through the multistep conformational changes that is illustrated in **(Figure II-1)**.^{8 9} The isomerization from 11- *cis* (498nm) to all-*trans* happens in picoseconds in this photo cascade. Upon absorbing a photon of light, batho rhodopsin is the first stable intermediate (543nm), which was first determined in bovine rhodopsin. In this intermediate the conformation of the retinal changes from 11-*cis* to all *trans* retinal. The overall structure of the backbone of the opsin remain the same as the ground state. The subsequent intermediates are Lumi and Met I. Met I can further converts to Meta II, which can activate the G protein.^{10 11} Activation of the G-protein leads to the signaling cascade. Formation of the Meta II

state leads to conformational change in the alpha-helices in opsin. Meta II can generate free retinal or it can form the Meta III. Meta III is slower to form the free retinal compared to Meta II. Depending on the pH, it takes hours to minutes for Meta III to release the free retinal.¹² Rhodopsin is the first GPCR protein that crystallized in 2000 with 11-*cis* retinal in the binding pocket. 11 years later, one of the final photo cascade products, protein bound to all-*trans* retinal in Meta II crystallized.¹³

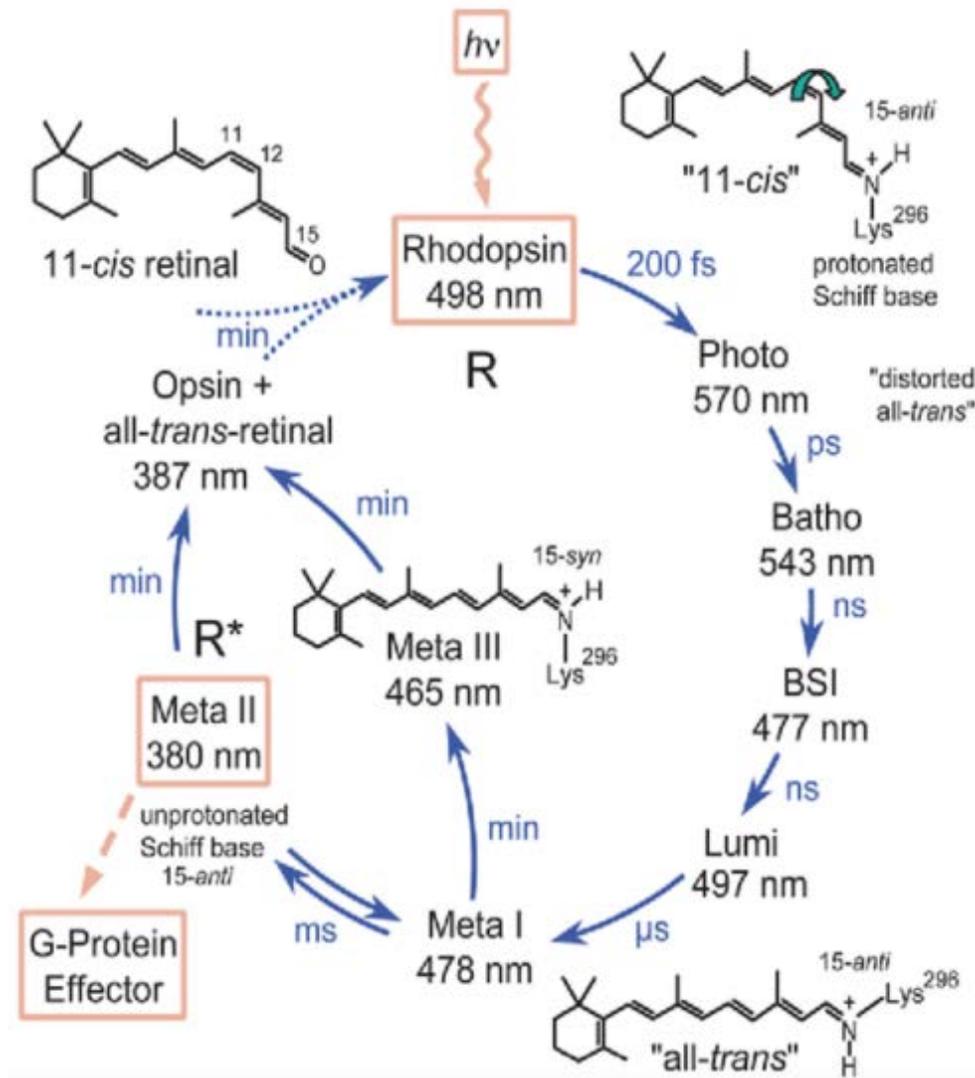


Figure II-1: The photocycle in Bovine rhodopsin. The absorption wavelengths and lifetimes of the intermediates are given.¹³

Formation of the Meta II lead to conformational change in the helical structure of the opsin protein. One of the most significant conformational changes is in TM6 containing Trp265 (**Figure 11-2**). Beside the vision, circadian rhythm is also regulated by the retinal photo isomerization in a rhodopsin family member in ganglion cells of the retina.

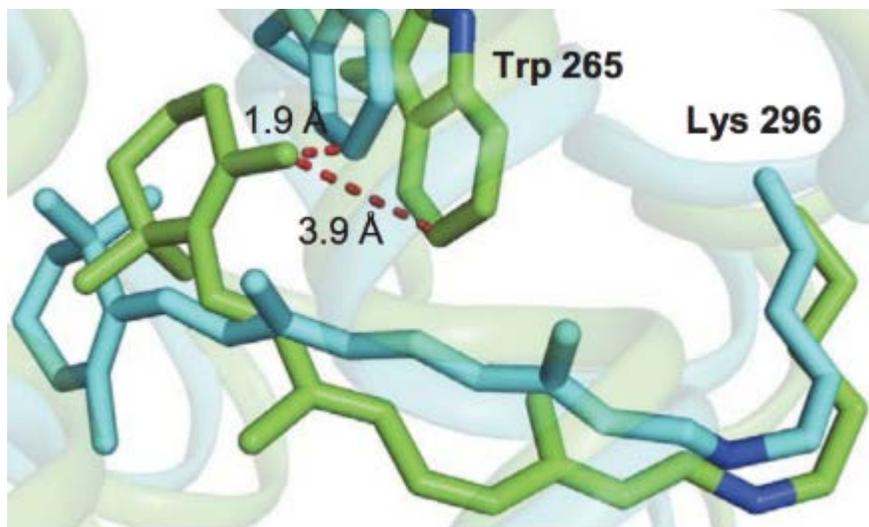


Figure II-2: Trp265 changes its position from ground state (Green) and excited state (Cyan) as retinal isomerizes in rhodopsin.

II-2 PHOTOISOMERIZATION CYCLE IN MICROBIAL RHODOPSIN

Photoisomerization in bacteriorhodopsin has been studied using NMR, Raman, FTIR, X-ray crystallography, cryo-EM, XFEL time-resolved crystallography, and theoretical calculations. The photocycle of bacteriorhodopsin, all-trans to 13-*cis* retinal isomerization, goes through multiple steps (**Figure 11-3**).^{9 14 15} Bacteriorhodopsin is one of the most well studied microbial rhodopsins. It acts as a unidirectional proton pump.^{16 17 18} This system is similar to animal rhodopsin because it passes through a many fast multi step process.¹⁹

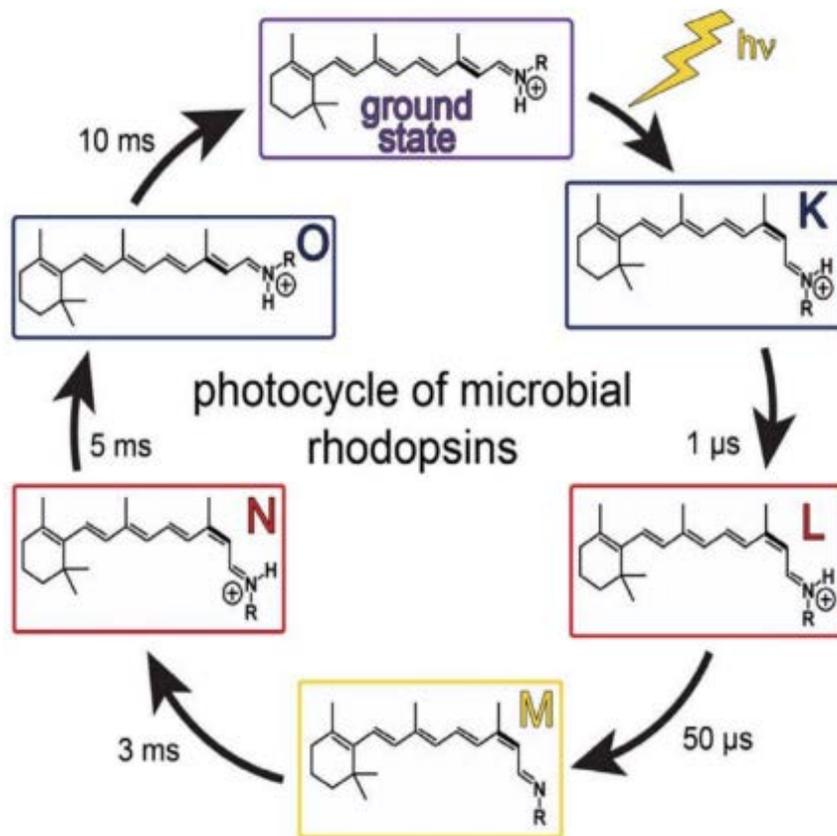


Figure II-3: Photocycle in microbial rhodopsin.

Two of these intermediates, O and K are red shifted, and the rest, L, M, and N are blue shifted. The isomerization first happens in the K intermediate, from all-trans retinal to 13-cis-15anti in 500fs.²⁰ After a microsecond, the retinal finds the planar conformation in the L intermediate to have hydrogen bonds with neighboring residues. Deprotonation happens in the M state which is significantly blue shifted. The M state first delivers the proton to the extracellular environment, and then faces to the intercellular space. Different studies confirm that the deprotonation happens through Asp96 and protonation happens through Asp85 in the M, N, and O states. Due to the fast kinetics of proton release from bacteriorhodopsin to the extracellular environment, there is uneven distribution of water network in bacteriorhodopsin system.^{21 22} There are 7 water molecules

through the extracellular side of the protein and two water molecules are toward the cytoplasmic side. The N intermediate generates the large conformational change of the backbone of the protein. The isomerization back to the all *trans* retinal happens in the O state.^{23 24 9} New studies through x-ray free electron lasers (XFEL) demonstrate the rotation of the all *trans* conformation to 13-*cis* retinal over the course of a few hundred femtoseconds in bacteriorhodopsin (**Figure II-4**).¹⁵

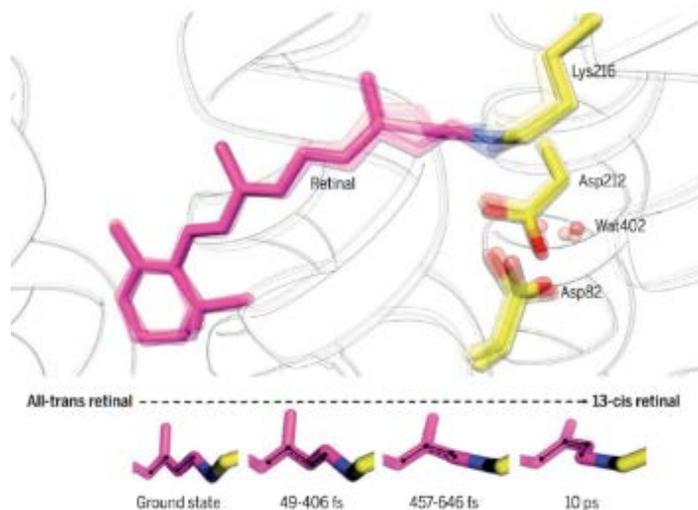


Figure II-4: Time-resolved serial crystallography resolves ultrafast atomic motions of retinal and shows the rotation of the all *trans* conformation to 13-*cis* retinal over the course of a few hundred femtoseconds in bacteriorhodopsin.¹⁵

II-3 PHOTOISOMERIZATION OF ALL TRANS TO 15-CIS IN HCRABPII

In the first chapter, I mentioned the two templates that we use in our laboratory to mimic rhodopsin proteins. Studies on the photoisomerization of hCRABPII system was initiated by Dr. Tetyana Barbasova.³ During studies for the design of a colorimetric proteinaceous pH sensor using hCRABPII template, they were successful in altering the pK_a of the iminium species from 2.4 to 8.1. During these studies, she observed the time-dependent conversion of the PSB ($\lambda_{max} > 450$ nm)

to the Schiff base (SB, $\lambda_{\max}=360$ nm).³ Follow up studies have been done when my former lab mates tried to increase the pKa of retinal in hCRABP_{II} by mutating residues of the binding pocket.³
^{25 7} Experiments show that hCRABP_{II} have lower pKas compared to hCRBP_{II} protein. To increase the pKa of the hCRABP_{II} protein, Dr. Meisam Nosrati compared the binding pocket of these two templates using the structures of hCRABP_{II} and hCRBP_{II}. He found that there are three Glu residues in the binding pocket of hCRBP_{II}, leading to a more hydrophilic binding cavity, which are not conserved in hCRABP_{II} protein. Another significant mutation in hCRBP_{II}, leading to higher pKa, is mutating Arg58 to Tyrosine. This resulted in R111K:Y134F:T54V:R132Q:P39Q:R59Y (M1) hCRABP_{II} mutant. Interestingly, acid-base titration of this mutant showed a broad peak from pH 4.3 to 10.5 for the chromophore. Dr. Meisam Nosrati showed that this indicates the change in the environment of the iminium nitrogen. Solution data shows that the maximum PSB is formed after around 3 hours. The PSB, in this case, is not stable and convert to SB after about 24 hours (**Figure II-5**). The crystal structure of this mutant after around 24-hour incubation with retinal is consistent with the *trans* structure, which is presumed to be the thermodynamic product (PDB:4YBP).²⁵

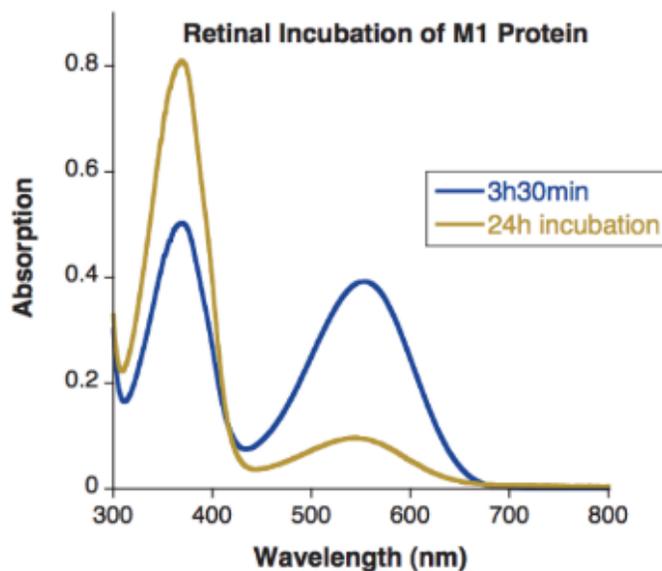


Figure II-5: Conversion of the PSB to SB overtime in M1 hCRABPII.^{7 25}

Crystallization immediately after the formation of the maximum PSB resulted in an ambiguous electron density, interpreted to be a mixture of *cis* and *trans* imine isomers. After a lot of efforts to accelerate the formation of PSB, Dr. Meisam Nosrati designed R111K:R134F:T54V:R132Q:P39Y: R59Y (M2) with the formation of the PSB in 20min. His crystallization trials after 20min and after 24-hour incubation with retinal lead to two different conformations, *cis*-PSB, and *trans*-SB, for the chromophore in the binding pocket (**Figure II-6**).

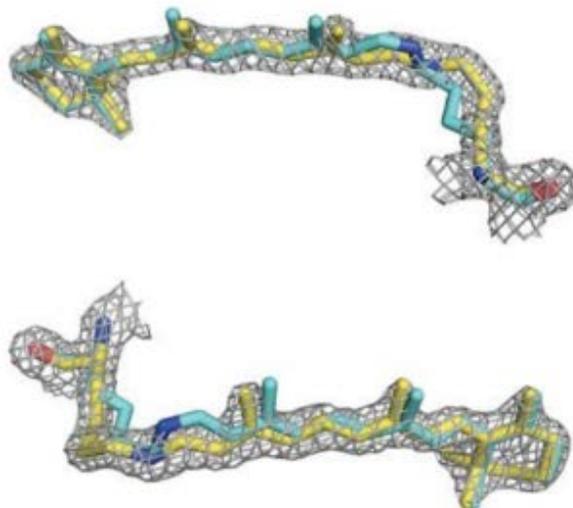


Figure II-6: Overlaid structures of 15-*cis* PSB and all-*trans* SB in
R111K:R134F:T54V:R132Q:P39Y: R59Y (M2) hCRABPII.^{7 25}

These results suggest that the *cis*-PSB forms as a kinetic product. This mutant of hCRABPII has the characterization of isomerization similar to the rhodopsin system. In hCRABPII crystals isomerization of 15-*cis*-PSB to all-*trans*-retinal is detected in a single crystal. Also, the pKa titration for this protein indicates two distinct pKa, 5.4 and 8.4, for the chromophore suggesting that the isomerization happens between two different pKa's similar to the rhodopsin system. The reasons behind the change in the pKa are explained by the hydrophobicity of the environment near the iminium region. In all-*trans*-SB retinal, the imine is pointed toward the hydrophobic residues in the binding pocket, lowering the pKa. However, in the *cis*-iminium isomer, the positive charge of the iminium is stabilized with the W108 through a π -cation interaction (**Figure 11-7**).

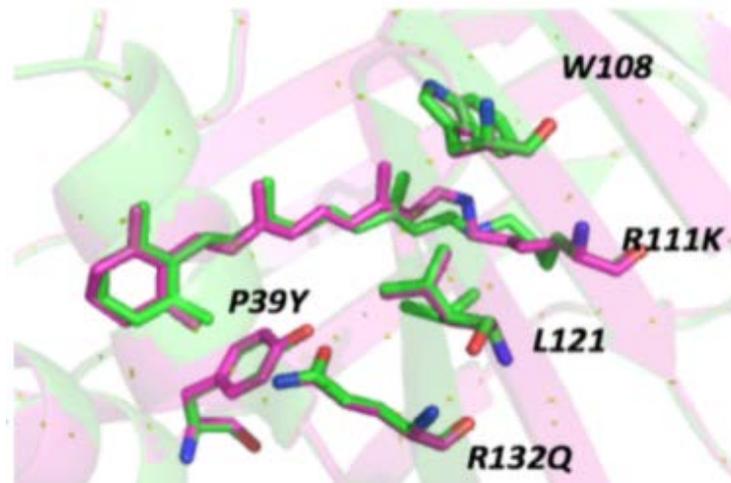


Figure II-7: Overlaid structures of 15-cis PSB (pink) and all-trans SB (green) in R111K:R134F:T54V:R132Q:P39Y: R59Y (M2) hCRABPII. The trans isomer stabilized through the hydrophobic interactions. Cis iminium stabilized through the π -cation interaction with Trp108.⁷

One of the most important characteristics of the rhodopsin system is their ability to photo-switch. In an effort to make the hCRABPII system photo switchable, the M2 mutant was incubated with retinal and after the formation of the maximum PSB, the complex was exposed to visible light. Exposure to visible light lead to conversion of PSB to SB relatively rapidly. After the conversion of the most of the PSB to SB, the complex was exposed to UV light, resulting in the reduction of the SB absorption and increasing the PSB peak. As measured by Dr. Nosrati, this conversion can be repeated for several cycles (**Figure II-8**).

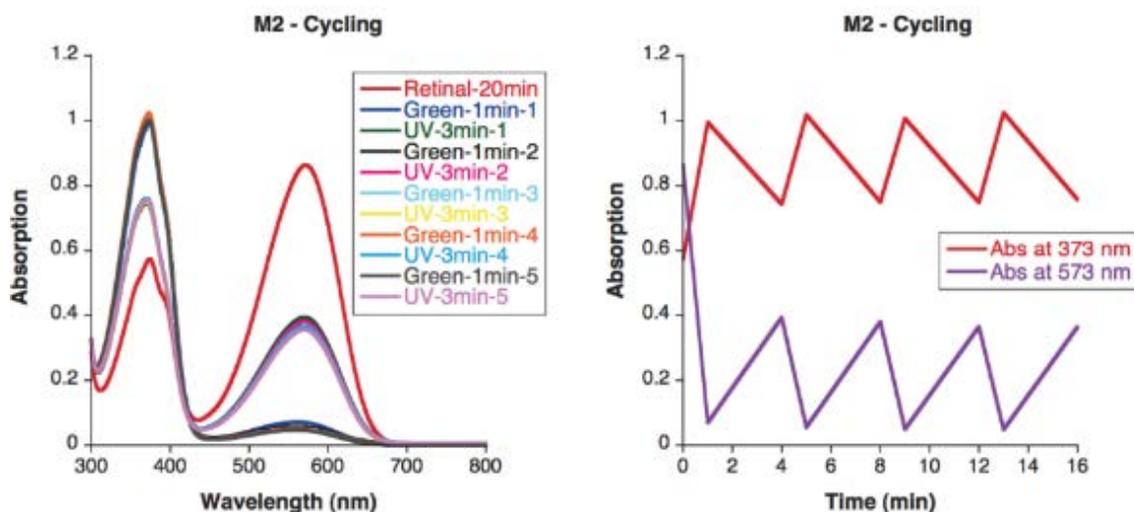


Figure II-8: Reversible photocycle for R111K:R134F:T54V:R132Q:P39Y: R59Y (M2) *hCRABPII*.

The same experiment has been done with a single crystal. Crystals were frozen before and after UV irradiation at pH 7.5, and they resulted in *trans*-SB and *cis*-PSB, respectively. The cycle is repeatable in crystals as well (**Figure II-9**).

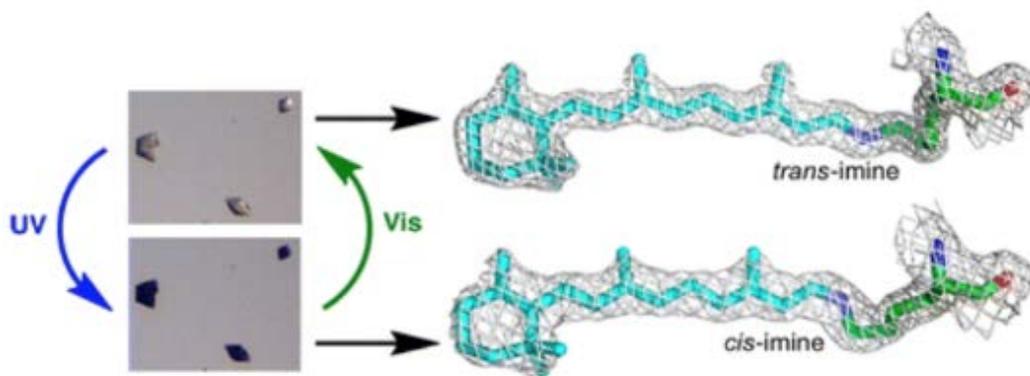


Figure II-9: Isomerization of *hCRABPII* system in crystal.⁷

Overall, Dr. Nosrati was successful to make the first photo switchable template for rhodopsin in our group, using the *hCRABPII* template. His proposed mechanism for isomerization in this

system is illustrate in (**Figure 11-10**). Isomerization happens between two distinct pKa regimes in this system.

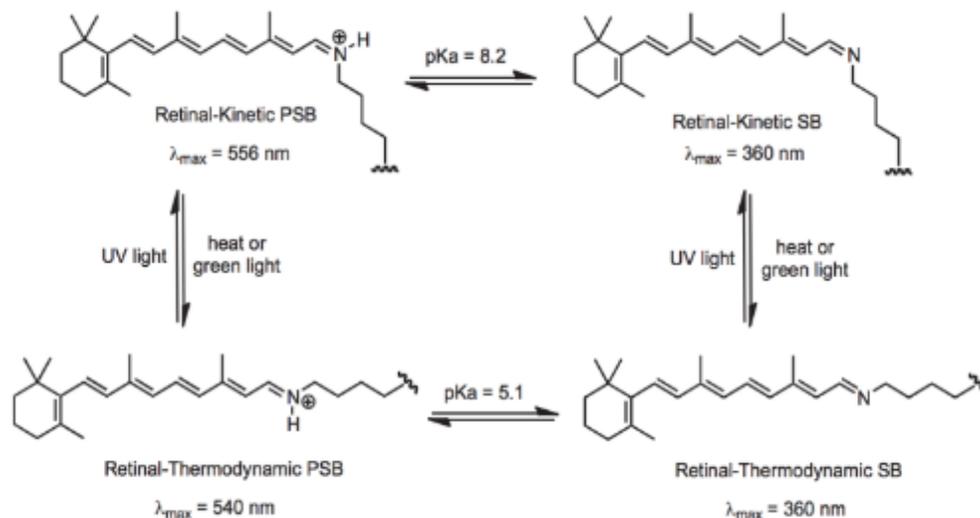


Figure II-10: Isomerization of 15cis retinal to all trans retinal in hCRABPII system.

Further computational studies have been done on this system using hybrid quantum mechanics/molecular mechanics (QM/MM). This studies suggest that the photocycle on M2 construct of hCRABPII passes through several intermediates. Green light exposure on M2 15C-PSB, resulted in 13- cis,15-cis product, following by isomerization into 13-cis,15- trans and, finally, an all-trans retinal.²⁶

II-4 PHOTOISOMERIZATION OF THE 13-CIS TO ALL TRANS RETINAL IN HCRABPII (SIMILAR TO BACTERIAL RHODOPSIN)

After the finding of the photo switchable property in the hCRABPII system, Dr. Alireza Ghanbarpour decided to mimic the photoisomerization of bacteriorhodopsin, 13-cis to all-trans

retinal, in this system (**Figure II-11**).^{27 6} His first trials using M1 construct to obtain the 13-*cis* retinal bound to hCRABPII were not successful. In order to make this system the same as bacteriorhodopsin, the 15-*cis* to all trans isomerization needed to be prevented. Also, the isomerization for C13-C14 double bond needs to be more favorable compare to the other double bond in the chromophore. By comparison to the bacteriorhodopsin protein, Dr. Ghanbarpour designed the L121 to aromatic residue to mimic the Trp86 steric interaction in the rhodopsin system. His efforts to crystalize M1-L121Y, M1-L121W, and M1-L121F was not successful; however the R111K:Y134F:T54V:R132Q:P39Q:R59Y:L121E (M1-L121E) variant, seems to prevent the imine isomerization. Structure of the M1-L121E bound to all trans retinal demonstrate the salt bridge between the trans iminium and Glu121.

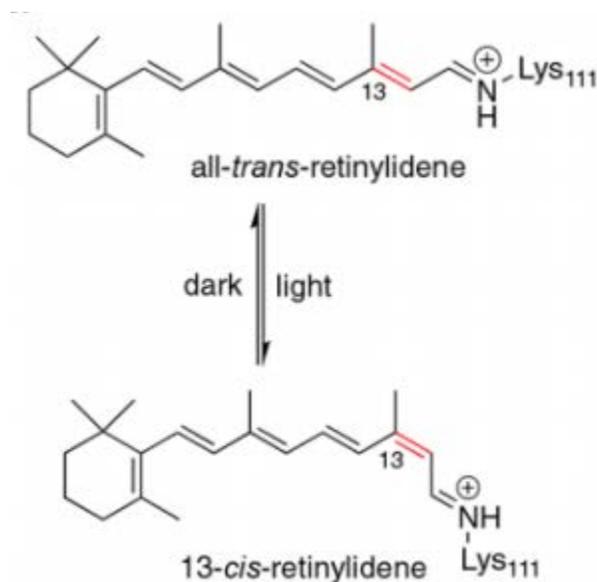


Figure II-11: Photoisomerization of 13-*cis* retinylidene to all trans retinylidene in microbial rhodopsin.

By adding L121E on M1 hCRABP_{II} construct and increasing the pK_a of chromophore, Dr. Ghanbarpour was successful to obtain the 13-*cis*-15-*syn* structure (PDB: 6MZQ, 6MOP) (**Figure II-12**).⁶

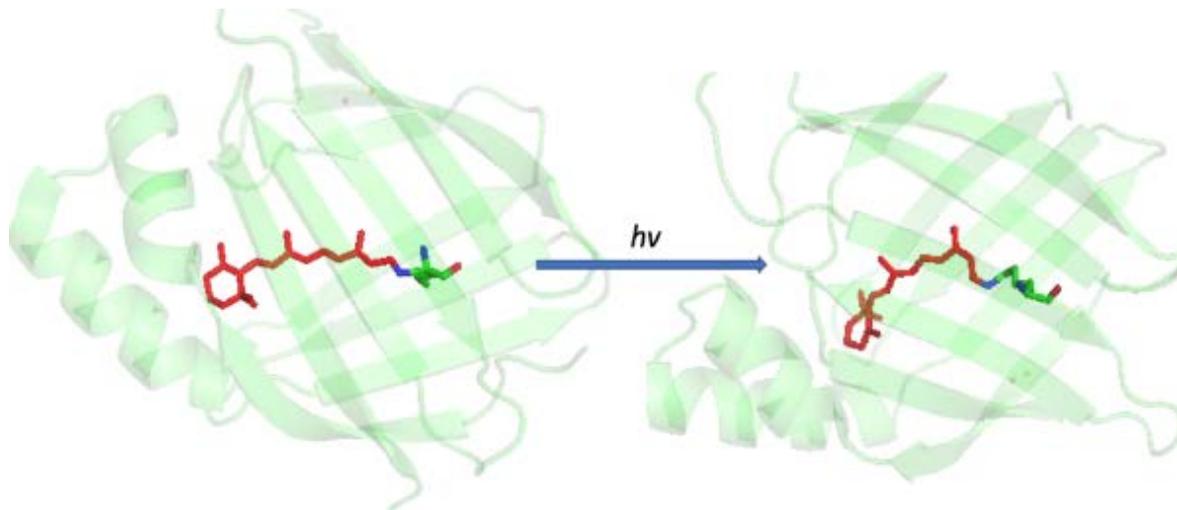


Figure II-12: Photoisomerization of 13-cis retinylidene to all-trans retinylidene in hCRABP_{II} system.

II-5 ISOMERIZATION IS DETECTED IN HCRBP_{II}

II-5-1 Clues for isomerization in hCRBP_{II}

As explained before in the first Chapter, hCRBP_{II} protein is bound to retinal via the Q108K: K40L (KL) double mutant by forming a Schiff base with Lys108.^{5 28} During our studies, this protein had been crystallized more than 40 times, and in most of these structures, retinal was bound through the *cis*-iminium to the protein. Since the crystallization condition for this protein is in pH 4, and this pH is much lower than pK_a of the chromophore in solution, we expect to have a positive charge on the Schiff base (PSB). The *cis* isomer seems to be stabilized and trapped via a π -cation interaction (iminium-W106) and a water network between Q4 and iminium (**Figure II-13**).^{29 30 5}

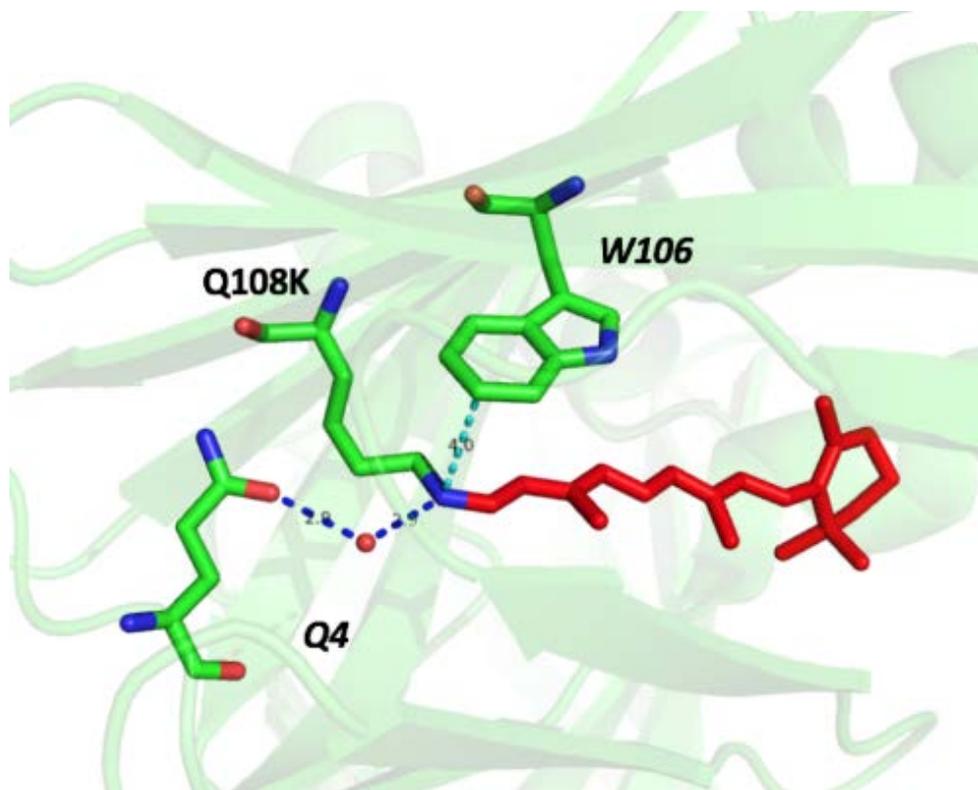


Figure II-13: Positive charge on *cis* iminium stabilized through the water network with Gln4 and π -cation interaction with Trp106 (PDB:4RUU).

Interestingly, in contrast to the other variants, all the structures that have mutation of Gln4 show the electron density consistent with the *trans*-isomer for the chromophore and no electron density for the water molecule. We hypothesized that by removing the water molecule, the *cis*-iminium is no longer favored over the *trans*-isomer. The overlaid structures of mutants with Gln4 with the rest of the structures indicates the removal of one water molecule near to the PSB region (**Figure II-14**). We predicted that removing this water molecule in our system may lead to ease of the conformational change to the *trans* isomer. We hypothesized that the formation of *trans* iminium structures in Gln4 mutants is due to the thermal isomerization of chromophore as it happens in the rhodopsin systems and also as we discovered for hCRABPII system before. To verify our

hypothesis, Q10K;K40L;T51V;T53C;R58W;T29L;Y19W;Q4A mutant (M1:Q4A), which is one of the most red-shifted mutant of hCRBP_{II} mutants, was investigated in this study.

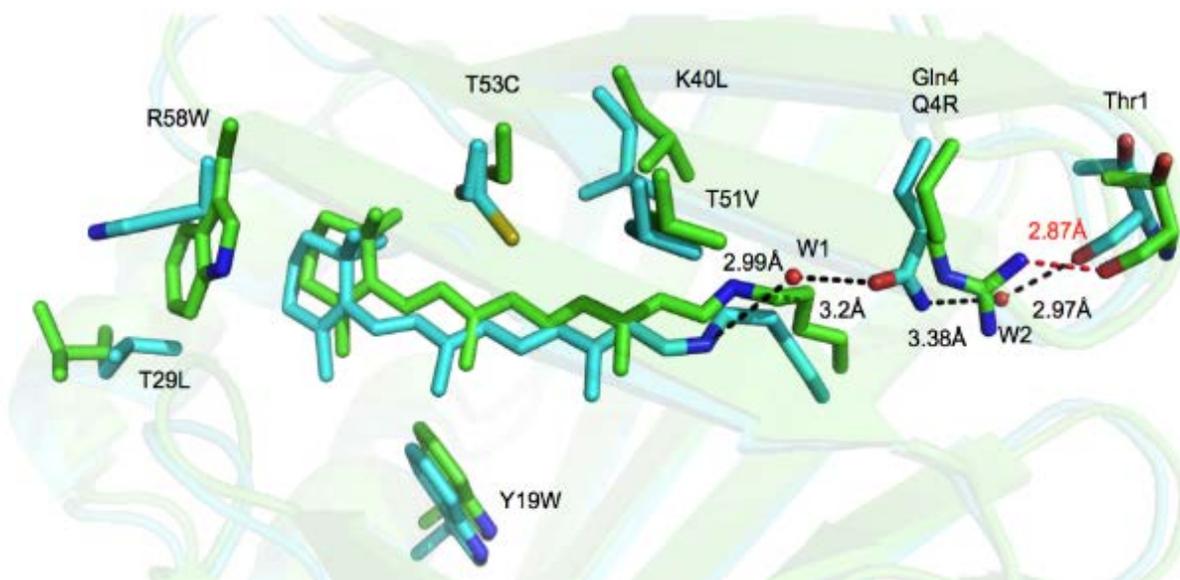


Figure II-14: The overlaid structure of KL:T51V:T53C:R58W:Y19W:T29L (cyan, 591 nm) and of KL:T51V:T53C:R58W:Y19W:T29L:Q4R (green, 622 nm).

II-5-2 Photoisomerization on hCRBP_{II} system was detected in solution

The UV-visible spectroscopy has been done to detect the changes in the absorption of the two isomeric state 15-cis/all-trans iminium. The M1: Q4A mutant solution incubated with retinal for 30min at pH 8 in the absence of visible light, and the UV measurement was done at physiological pH (**Figure II-15**). The 380nm absorption is related to the free retinal that shifted toward the 360nm, which is the Schiff base's absorption. By bringing the pH down from 8 to 7.4, some PSB (612nm) will initially form (maximum at around 45min). The PSB is not stable, and it will convert back to SB overtime.

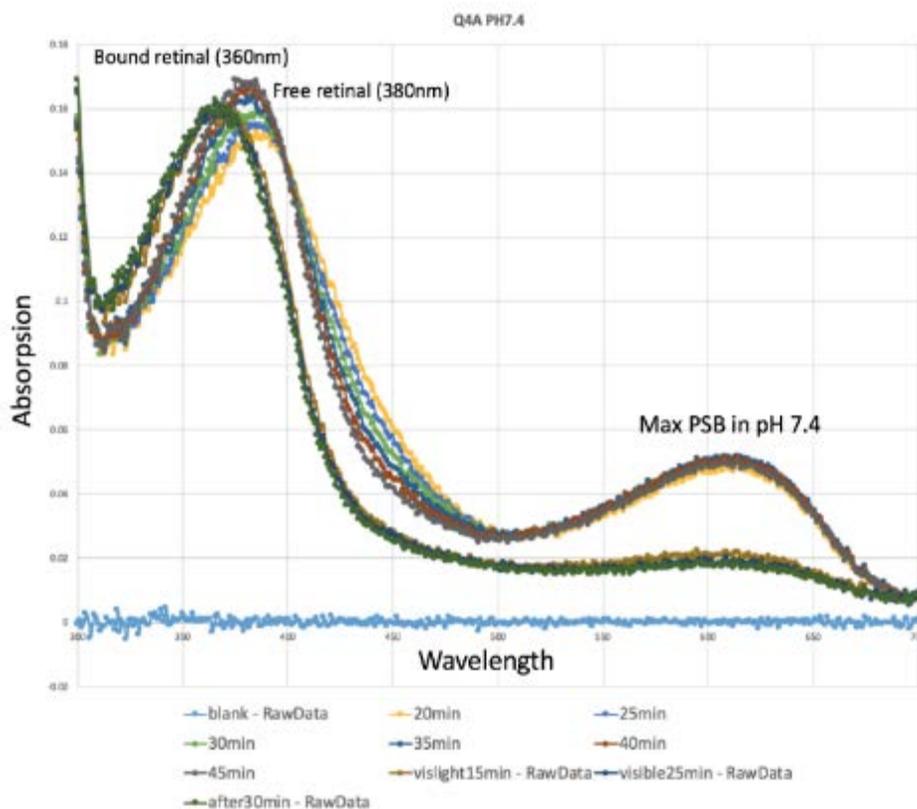


Figure II-15: UV measurement for M1;Q4A after incubation with retinal for 30min. The measurement has been done at pH 7.4. Maximum PSB in this pH (gray line). PSB is not stable and it converts back to SB (360nm) over time in presence of the visible light (or it can happen thermally) (green line).

Since all of the crystal structures of hCRBP_{II} were grown in pH 4, the UV spectroscopy has also been done at pH 4 to be consistent with structural analysis. Using pH4 also helped us to see the formation of PSB very clearly (Since the pK_a after 30-hour incubation of protein with ligand was measured to be 7, and pH 4 is below this pK_a). UV spectrum shows the maximum formation of PSB after 30min of bringing the solution down to pH4 in the absence of visible light (**Figure II-16**). The 360nm peak is related to the absorption of the imine bond, and the more red-shifted absorption (612nm) is associated with Protonated Schiff Base (PSB).

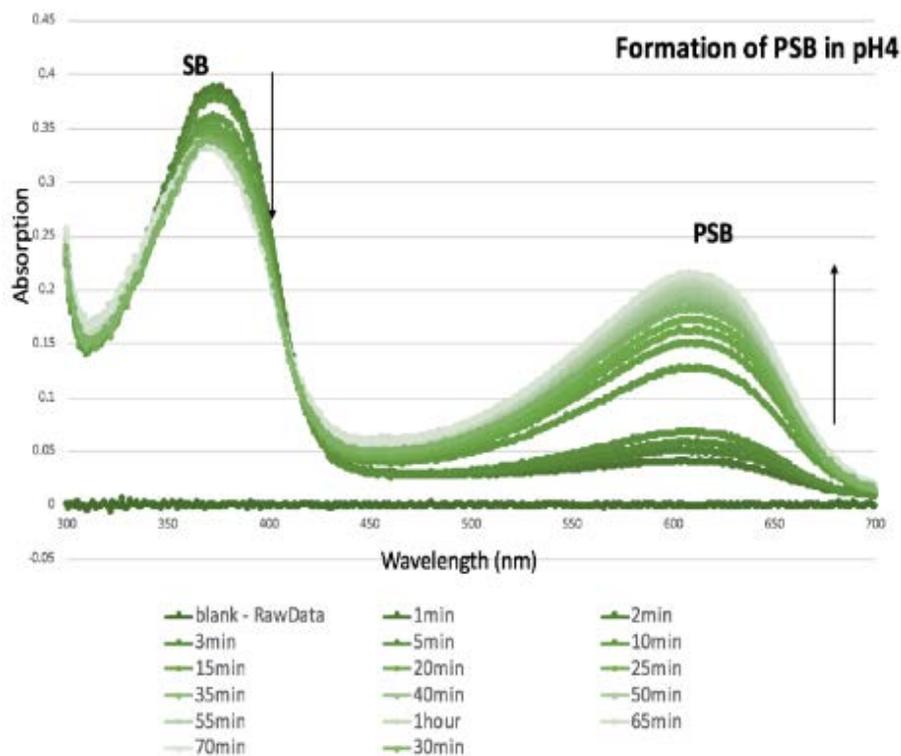


Figure II-16: Formation of PSB after bringing the solution from pH8 down to pH4 in the absence of visible light.

Interestingly, even at pH4, the PSB was not stable and converted to SB (λ_{max} 360) thermally over 24 hours in the absence of visible light (**Figure II-17**). As mentioned previously, by removing the Gln4 and removing the water molecule, we predict that the positive charge on the PSB is not stable because the retinal converts to the more stable thermodynamic product, all-*trans* imine.

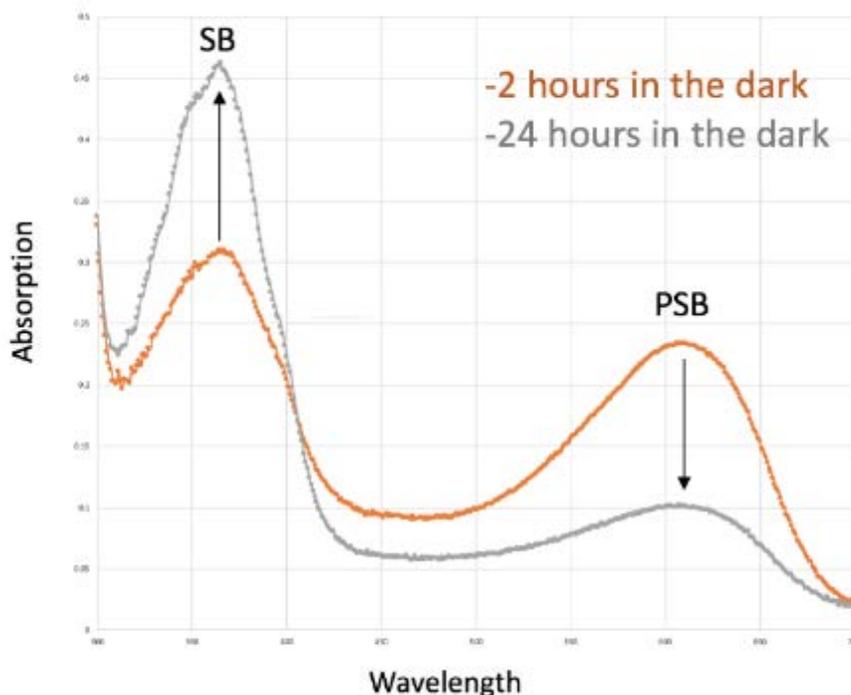


Figure II-17: Thermal conversion of the PSB to SB overtime in pH4.

To compare the rate of conversion of PSB to SB in visible light or in the absent of visible light (thermal conversion), we repeated this experiment by incubation of the solution in visible light. This leads to the loss of PSB and increasing the SB at a faster rate (**Figure II-18**). These results demonstrate that even in low pH (pH4), the PSB will convert to SB either thermally after around one day or by the presence of visible light after a couple of hours. As mentioned before, the first crystal obtained for this mutant (M1;Q4A) was consistent with all trans retinal conformation, which we believe is the thermodynamic product. Since, we incubated the solution and crystallization screen in the dark for a couple of days to get crystals, we predicted that this crystal is associated with the SB peak.

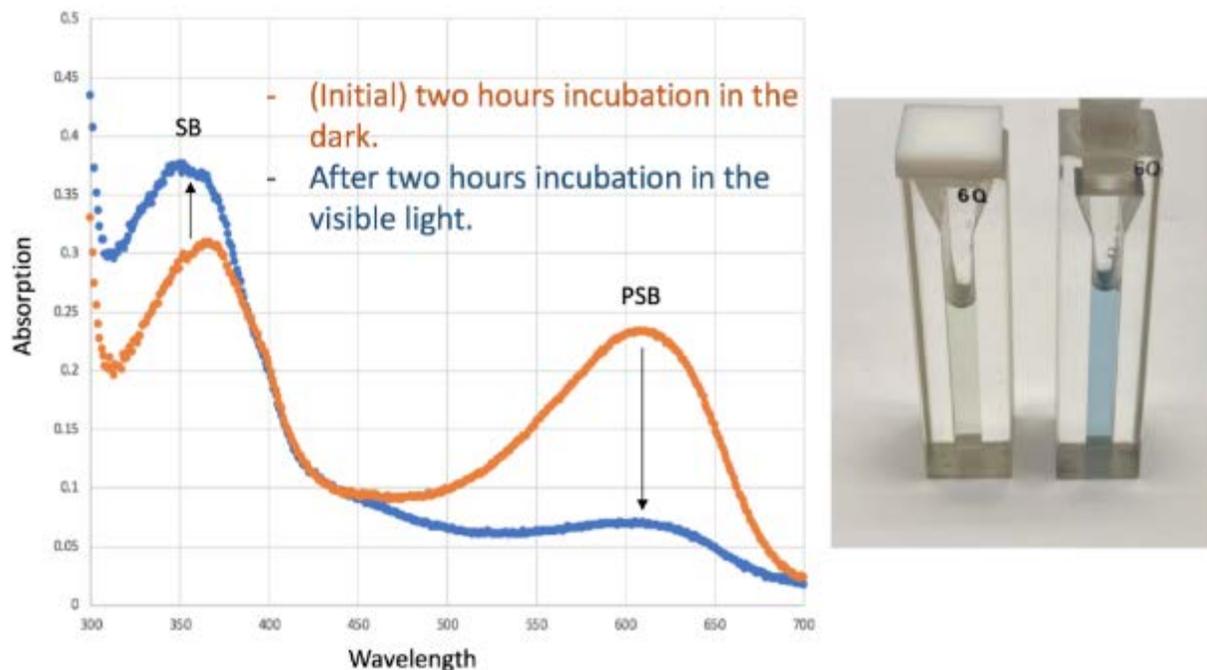


Figure II-18: Conversion of the PSB to SB in the presence of the visible light. Most of the PSB converts to SB after around 2 hours.

To investigate our hypothesis, we needed to obtain the crystal structure of the kinetic product. Therefore, we need to find a way to reproduce the PSB peak again. After removing all of the PSB thermally in the solution in pH4, by UV irradiation (340-360nm) on the solution, we were able to convert the SB (λ_{max} 360nm) to the PSB (λ_{max} 612nm). The intensity of the SB was much lower than before, which was due to the conversion to the PSB and also possibly quenching the chromophore. The maximum PSB intensity was gained by sinning UV light (360nm) for 30 seconds on the solution, and the intensity of the PSB, in this case, was even higher than it was measured before as the Maximum PSB in pH 4 (**Figure II-19**).

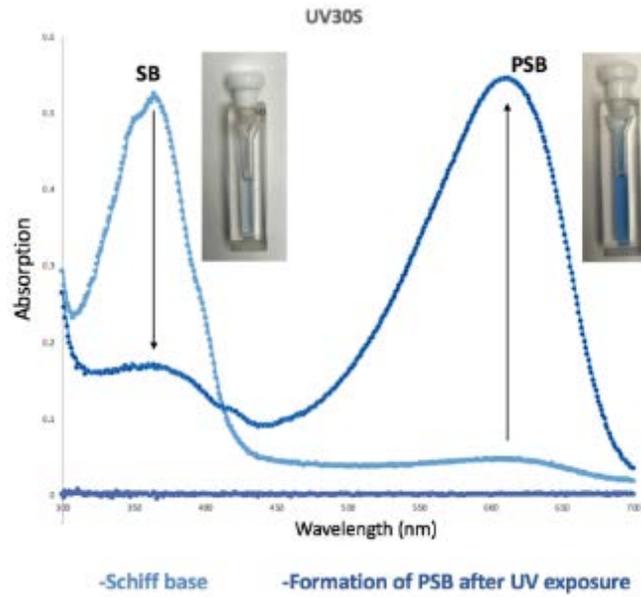


Figure II-19: Conversion of the SB to PSB by shining 30s UV exposure.

The PSB peak converts back to SB both thermally or by incubation in visible light, which demonstrates that we have a reversible photo switchable system (**Figure II-20**).

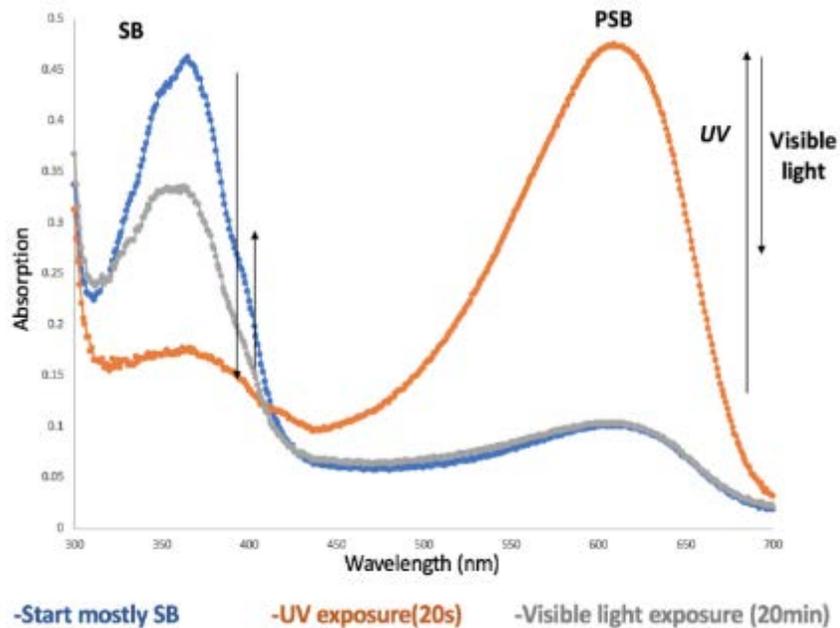


Figure II-20: Reversible photoswitching system for M1; Q4A hCRBP II system. By UV exposure the SB converts to PSB (orange line). Visible light for 20min converts the PSB back to SB.

Longer exposure of the solution to the UV irradiation led to a decrease in the intensity of both SB and PSB peaks, which demonstrates the quenching of the chromophore (**Figure II-21**).

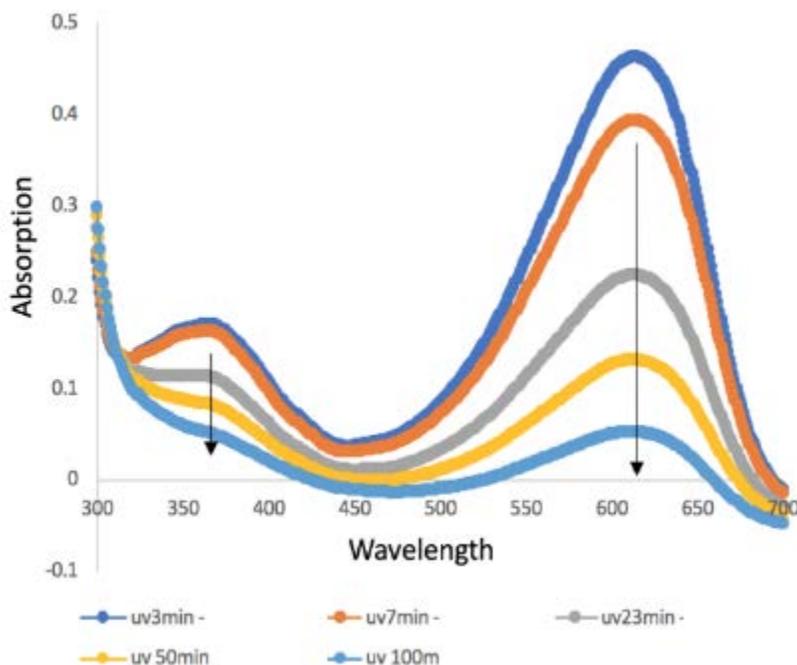


Figure II-21: Longer exposure of the solution to UV light lead to quenching the chromophore.

Based on the spectroscopic data, we hypothesized that the formation of trans iminium structures in M1; Q4 mutants is due to the thermal isomerization of chromophore as it happens in rhodopsin systems and also as we discovered for CRABPII system before.^{7 25} As we discussed previously, the reason behind this isomerization is changing in the pK_a environment for the chromophore.

II-5-3 pK_a measurements for M1:Q4A

We detect a huge change in the pK_a of these two forms of the complex. While the pK_a of the the chromophore was measured to be 7 after the binding of the retinal to the protein (around 30min incubation of the protein with retinal), measuring the pK_a after 24hours incubation of protein with the ligand, gives us the pK_a around 3.3. Similar to Rhodopsin and CRABPII systems,

photoisomerization results in a large change in the pK_a of iminium in these conformers. (pK_a 7 for PSB and $pK_a=3.3$ for the SB form) (**Figure II-22**). Based on these spectroscopic data, we assume that we have two distinct isomers that have been converted to each other, both thermally and photochemically.

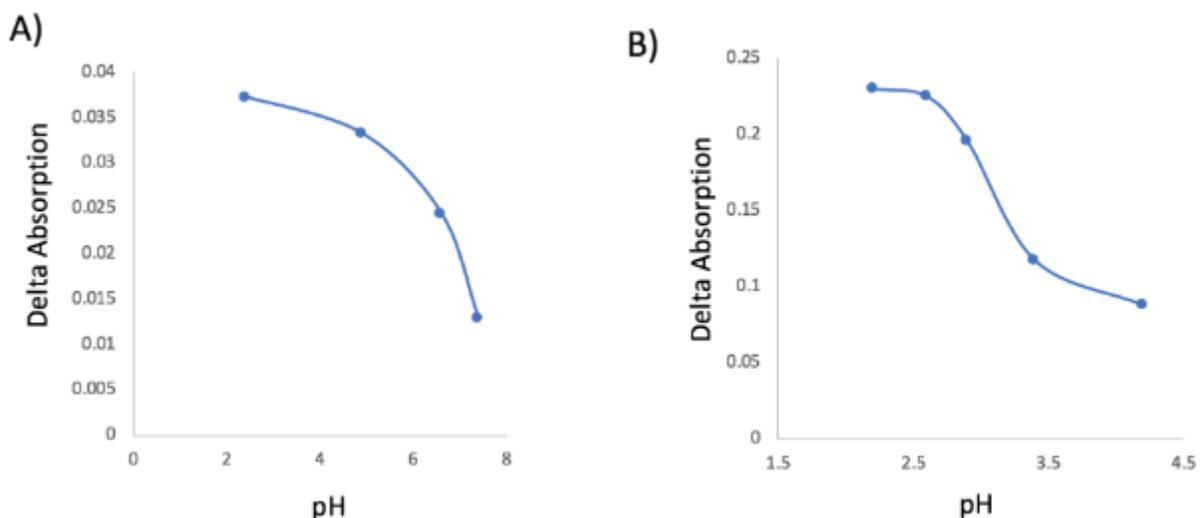


Figure II-22: A) pK_a measurement for the chromophore after 30min incubation of the protein. pK_a is around 7. B) pK_a measurement after 24 hours incubation of the solution in pH 4.5. pK_a is measured around 3.3.

II-5-4 Photoisomerization on hCRBP II system was detected in crystals

Based on the spectroscopic data, we hypothesized that the formation of trans iminium structures in M1;Q4 mutants is due to the photochemical isomerization of chromophore as it happens in rhodopsin systems and also as we discovered for CRABP II system before. In order to verify our hypothesis, change in the protonation state of the two photo isomers from high pK_a form (*cis* iminium (λ_{max} 612nm)) to low pK_a form (all-trans-retinal (λ_{max} 360)) was obtained by high-resolution X-ray crystallographic data for this mutant. As explained previously, growing the

crystals for this mutant lead to trans conformation for imine bond. We have tried to convert the trans to *cis* conformer by UV irradiation similar to what we expect from the solution data (**Figure II-23**). By exposing the crystal to the UV irradiation, we saw a huge change in the color of crystals.

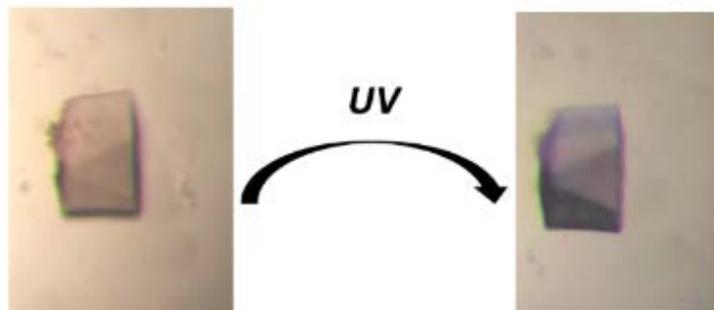


Figure II-23: UV exposure for 2min on M1;Q4A crystal. Crystal changed the color from the colorless to blue.

The first trials, crystals after a short time UV exposure (less than 1 min UV exposure), demonstrate the mixture of *cis* and trans isomer for the Schiff base (mostly *cis* conformation). Also, a comparison between all-trans-retinal and these structures from irradiated crystals shows floppy in the position of Q38 in the binding pocket. After many trials, I was successful in trapping the structure of the kinetic product (*Cis* isoform) after 2min UV exposure. Overlaid structure of trans imine and *cis* iminium is illustrated in (**Figure II-24**).

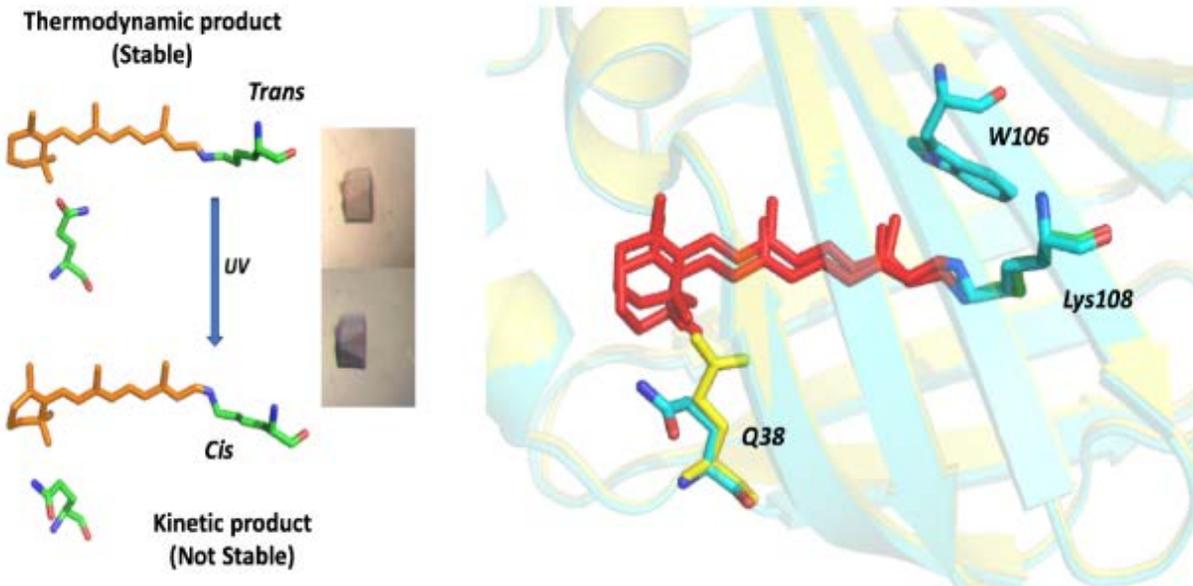


Figure II-24: Conversion of all *trans* retinal to 15-*cis* retinal in a single crystal of the M1;Q4A. the overlaid structure of the *trans* and *cis* isoform demonstrate the change in the position of the Gln38.

Change in the pKa environment can be explained by the crystal structures. In *trans* schiff base, the nitrogen is pointing toward the hydrophobic residues (low pKa). In *cis* protonated schiff base, the iminium is stabilized by the W106 through a π -cation interaction (**Figure II-25**).

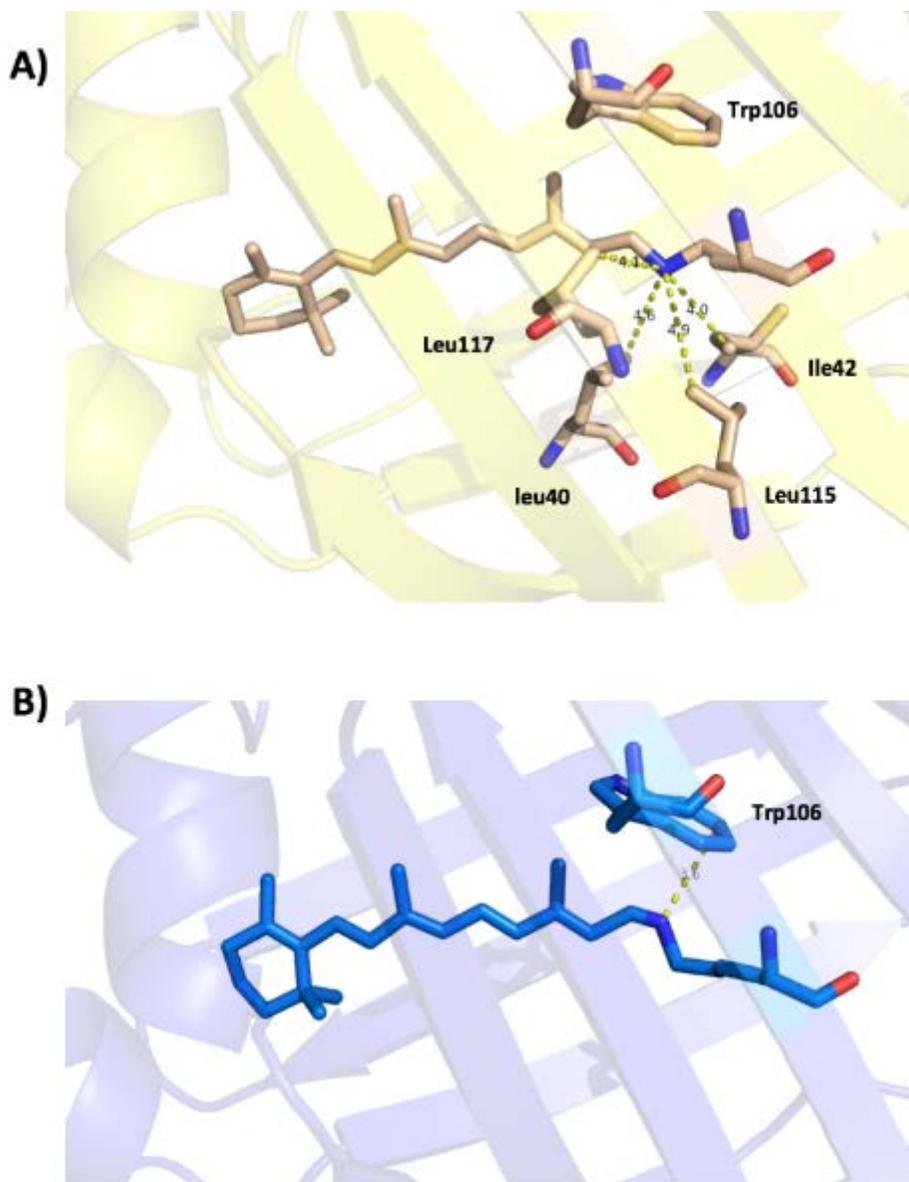


Figure II-25: A) Hydrophobic interaction of the nitrogen with surrounding residues in all trans retinal bound to M1;Q4A B) π -cation interaction of the cis iminium in M1;Q4A.

I was successful to detect photoisomerization of retinal in hCRBP II protein (**Figure II-26**).

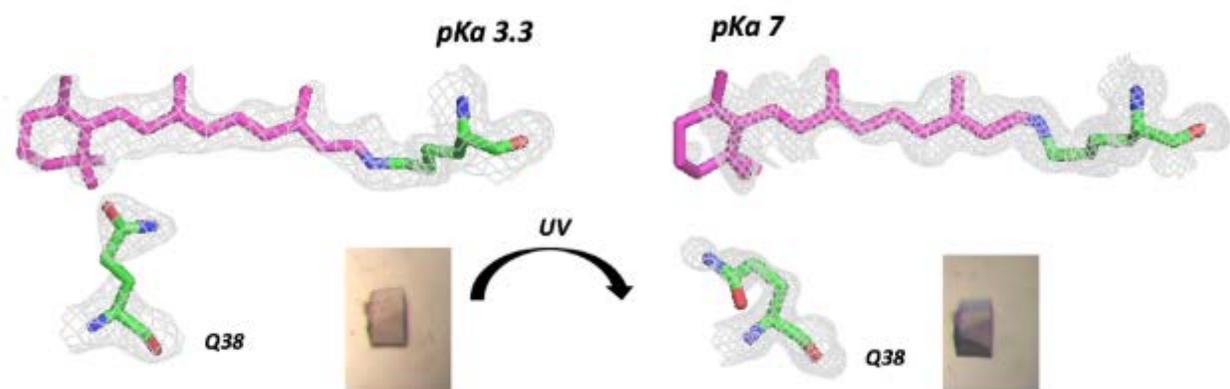


Figure II-26: Isomerization of all-trans retinal to 15 cis iminium in M1;Q4A hCRBP II.

To make a reversible photo switchable cycle in a solid-state, we incubate the crystals in the dark for a couple of hours after UV exposure. As we expect from the solution data, the chromophore should have the trans conformation again. I detect the change in the color of crystals; However, Since many chromophores quench by the UV exposure, we could not reproduce a decent electron density for the cycle (**Figure II-27**).

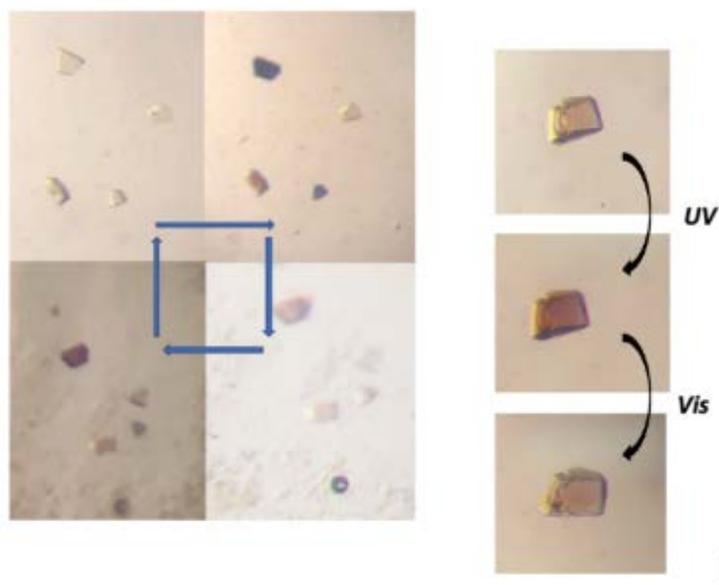


Figure II-27: Photocycle in M; Q4A crystals.

Up to this point we were able to catch two intermediates in the isomerization process of hCRBP_{II}, which was similar to what Dr. Meisam Nosrati was detected in hCRAB_{II} system, and we expected to happen (**Figure II-28**).

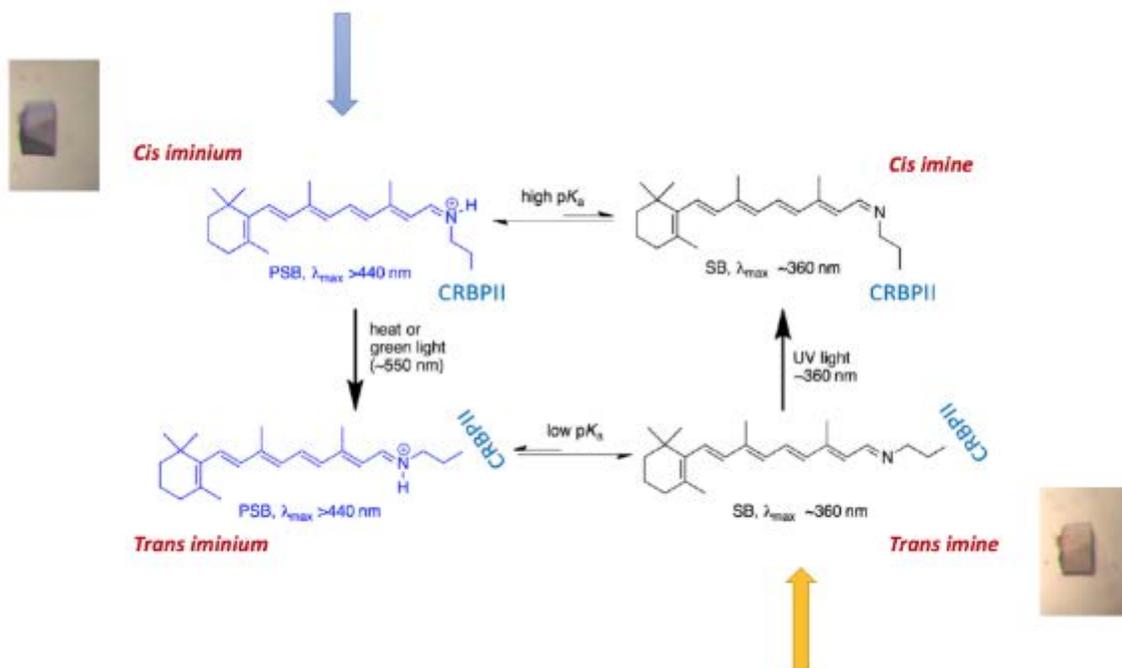


Figure II-28: Detecting the two intermediates in the isomerization process of M1;Q4A.

II-5-5 Detecting the potential third intermediate in the isomerization process

Another interesting observation that I had during this project was that the formation of crystals in the dark and the visible light showed different colors. Crystals formed in the dark were light blue color, and after exposing them to light, they change their color to colorless/yellow (**Figure II-29**).



Figure II-29: Crystal of M1;Q4A changes the color from blue to colorless by visible light.

Further, we show that the crystals which were grown in the light have the trans conformation, but with different water network surrounding the chromophore environment compare to the crystals grown in the dark. The only difference between these two structures is in 2 water molecules between Q38 and the chromophore (**Figure II-30**).

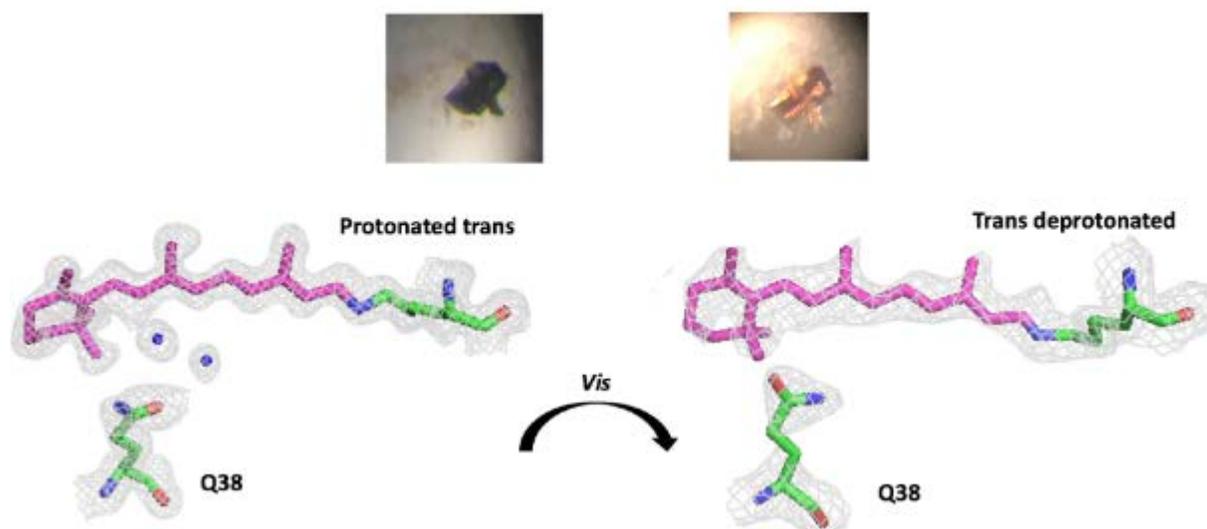


Figure II-30: Removing the water molecules between Gln38 and chromophore by visible light exposure.

Since the pK_a of trans iminium is 3.3, in pH4, some of the molecules have PSB, and getting blue crystals was expected (It is possible that PSB trans and PSB *cis* have the same absorption). However, after exposing these crystals to visible light, they change their colors to yellow/colorless. A possible explanation for this behavior is that exposure to visible light leads to an environment with an even lower pK_a (below 3.3), which leads to loss of all the PSB. Atomic resolution crystallography data demonstrate the same conformation for all the residues in the vicinity of the chromophore. The only difference is removal of two water molecules next to the Gln 38, which are far from the imine bond. The fact that changing the water network of the binding pocket can lead to two distinct pK_a 's is very interesting and it may shows the third intermediate in the

isomerization process (**Figure II-31**). Solution data also shows the removal of almost all of the PSB after around three hour exposure to the visible light.

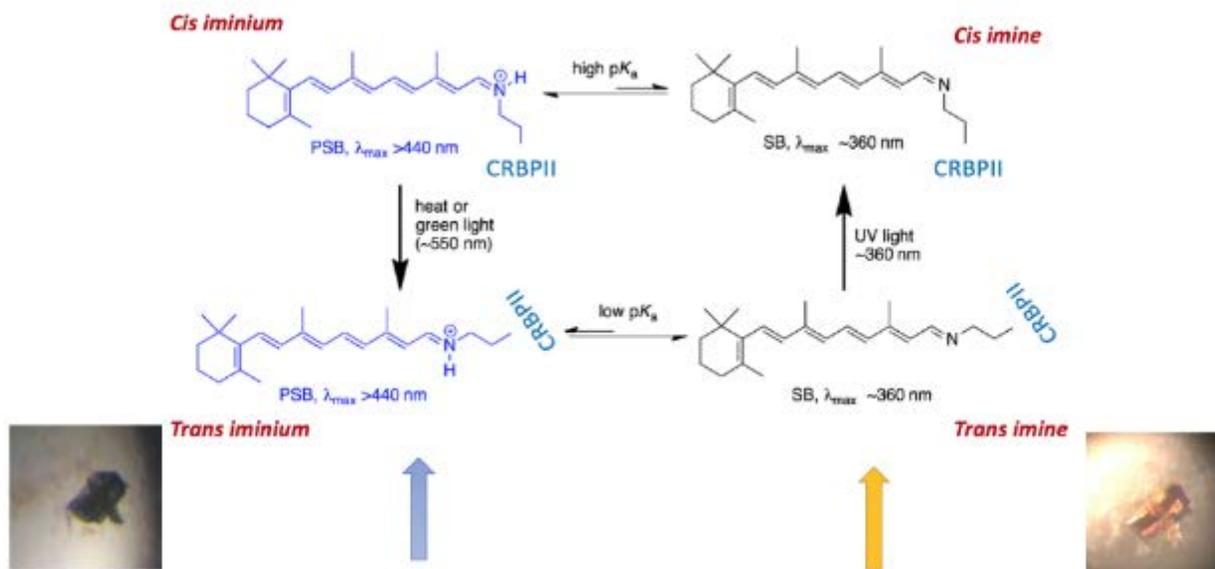


Figure II-31: Formation of the blue crystal in the dark might be related to *trans iminium* intermediate.

To investigate the effect of those two water molecules on the pK_a of the trans form, we mutate Glutamine to a nonpolar residue, Leucine. By mutating Glutamine to a nonpolar residue, we expect to remove these two water molecules from the vicinity of the chromophore, and we can measure the pK_a of the chromophore to see if it is lower than the pK_a of the *trans iminium* that was measured before (pK_a 3.3). I was successful in obtaining the crystal structure for this mutant, M1;Q4A;Q38L. As expected, the only difference between these two structures (M1;Q4A and M1;Q4A;Q38L) is these two water molecules and the rest of the structure is exactly the same (**Figure II-32**). From structural analysis, we predict that any change in pK_a of this construct compare to M1;Q4A is because of removing these two water molecules.

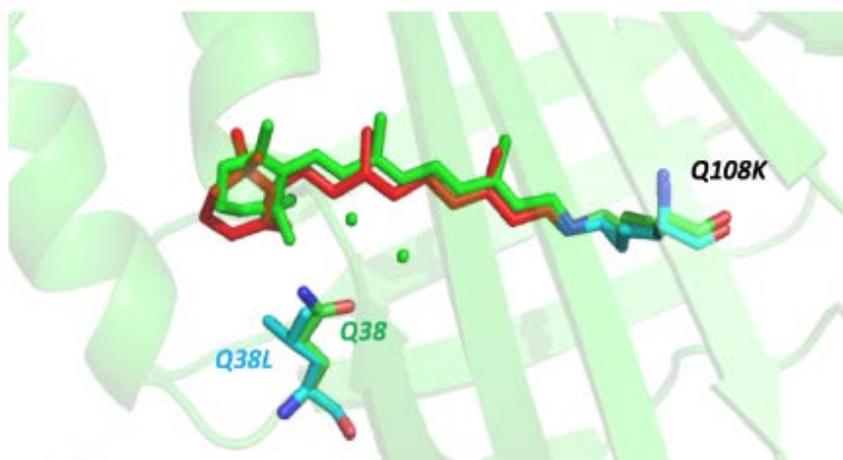


Figure II-32: Overlaid structures of the M1;Q4A (green), with M1;Q4A;Q38L (blue). The only difference is removing the two water molecules near Gln38.

At this point, we measured the pKa for the chromophore in this mutant (M1;Q4A;Q38L). The pKa of this mutant was below 3.3 as we predict and from the structural analysis (**Figure II-33**). This data supports our hypothesis that the pKa can be lowered by removing the water molecules.

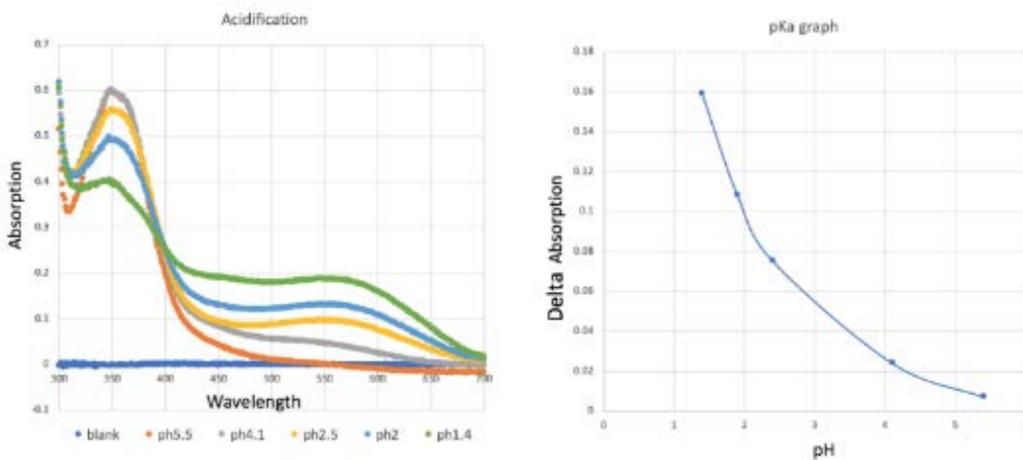


Figure II-33: pKa measurement for chromophore in M1;Q4A;Q38L. solution were bring down to pH 1.4. Lowering the pH lead to the aggregation of the protein. From the change in the absorption vs pH graph, it is clear that the pKa is lower than 3.3.

Overall, we have trapped three intermediates in this isomerization process (**Figure II-34**)

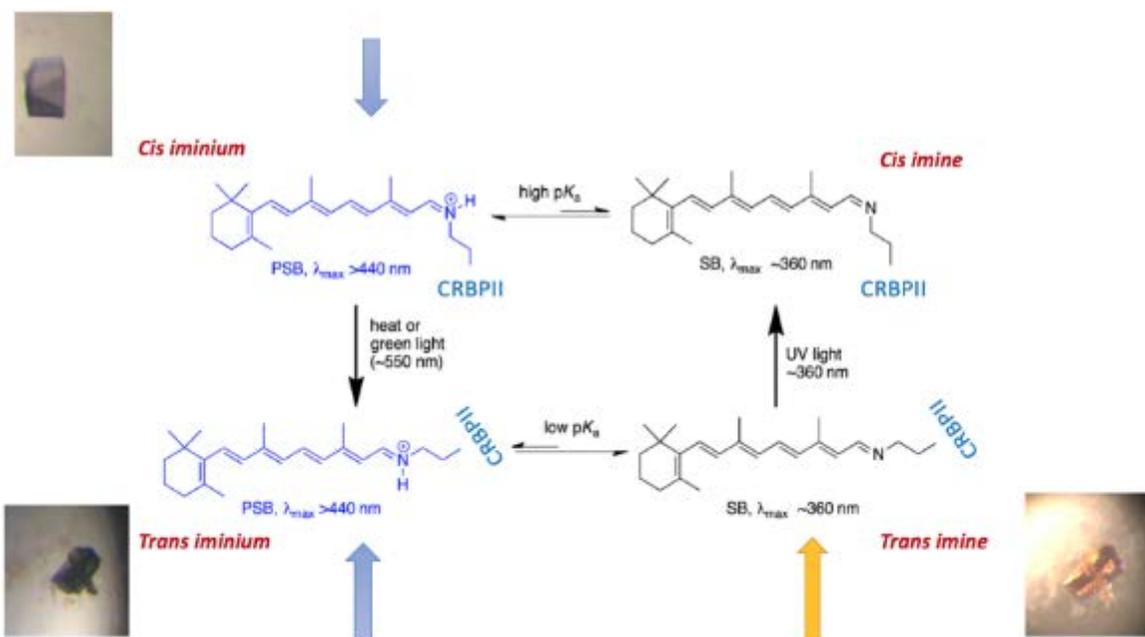
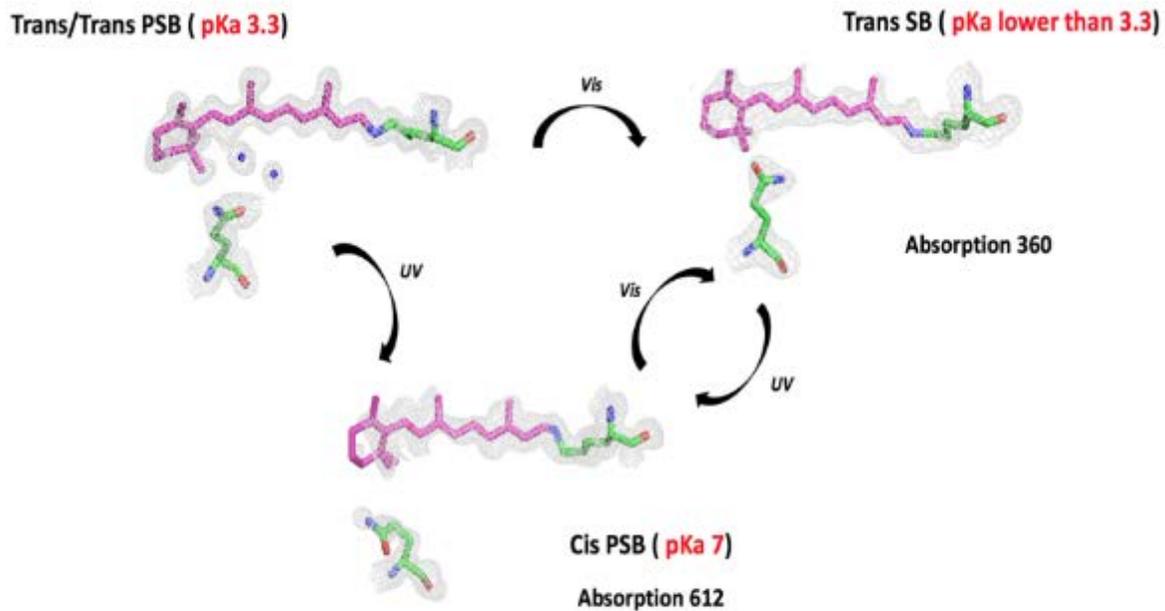


Figure II-34: Three intermediates detected for isomerization of the 15cis to all trans retinal in *MI;Q4A* construct.

II-5-6 Investigate Photoisomerization in another mutant of hCRBP II:

Q108K:K40L:T51V:T53C:R58W:Y19W:T29L:Q4R (M1;Q4R)

M1:Q4R with 622nm absorption is one of the most red shifted mutants of hCRBP II. The crystal structure of this mutant with all trans retinal was determined by Dr. Zahra Nossoni.^{29 30 5} Since the Q4 is mutated, the water molecule between the Q4 and imine bond is removed, and we expected to detect photo switching similar to M1:Q4A. The photoisomerization cycle was detected for this mutant in solution (**Figure II-35, II-36, II-37**). The pK_a is and it was measured by Dr. Wenjing Wang.

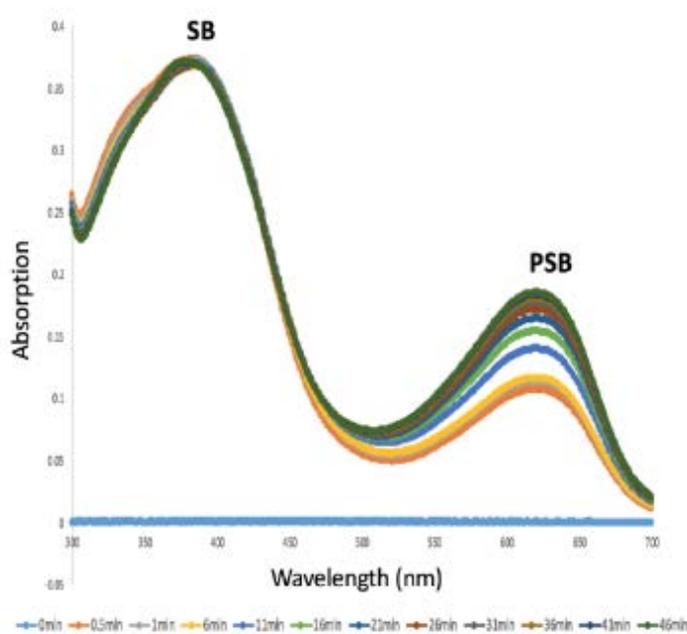


Figure II-35: Formation of PSB after bringing the pH down to 4 for M1;Q4R construct.

Maximum PSB formed after 46min (dark green line) pH4.

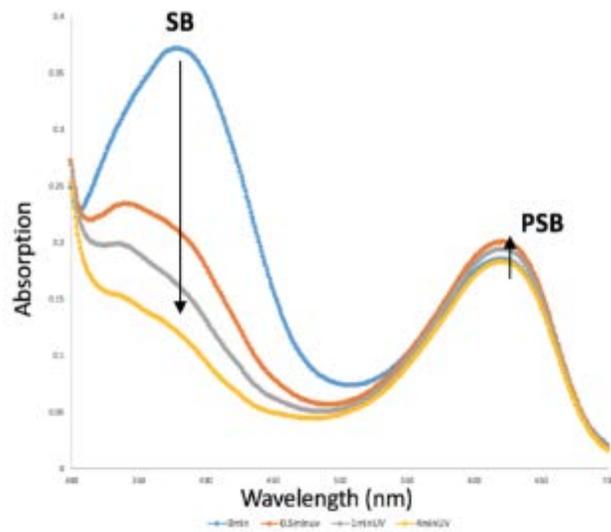


Figure II-36: *Shinning UV light (orange, gray and yellow line) after reaching the maximum PSB in the pH4 (Blue line) in M1;Q4R hCRBP II construct. The SB converts to PSB by UV exposure.(pH4).*

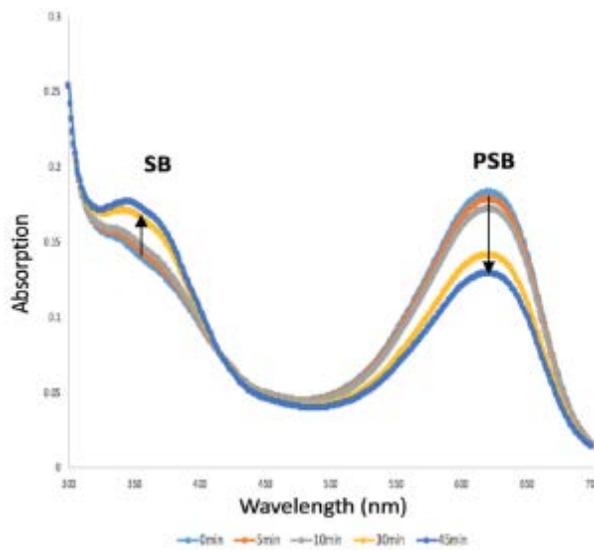


Figure II-37: *Visible light exposure to the solution converts the PSB to SB in M1;Q4R hCRBP II (pH4).*

**II-5-7 Investigate Photoisomerization in another mutant of hCRBP II:
(K40L:T51V:T53C:R58W:Y19W:T29L:Q4A:Q38L)**

After measuring the pka for the M1:Q4A:Q38L construct, we decided to perform photo switching experiments in this mutant as well. Binding the retinal to M1:Q4A:Q38L was very slow (around 5 hours). Conversion of SB to PSB (λ_{max} 533) detected also for M1:Q4A:Q38L mutant by UV exposure (Figure II-38 , Figure II-39).

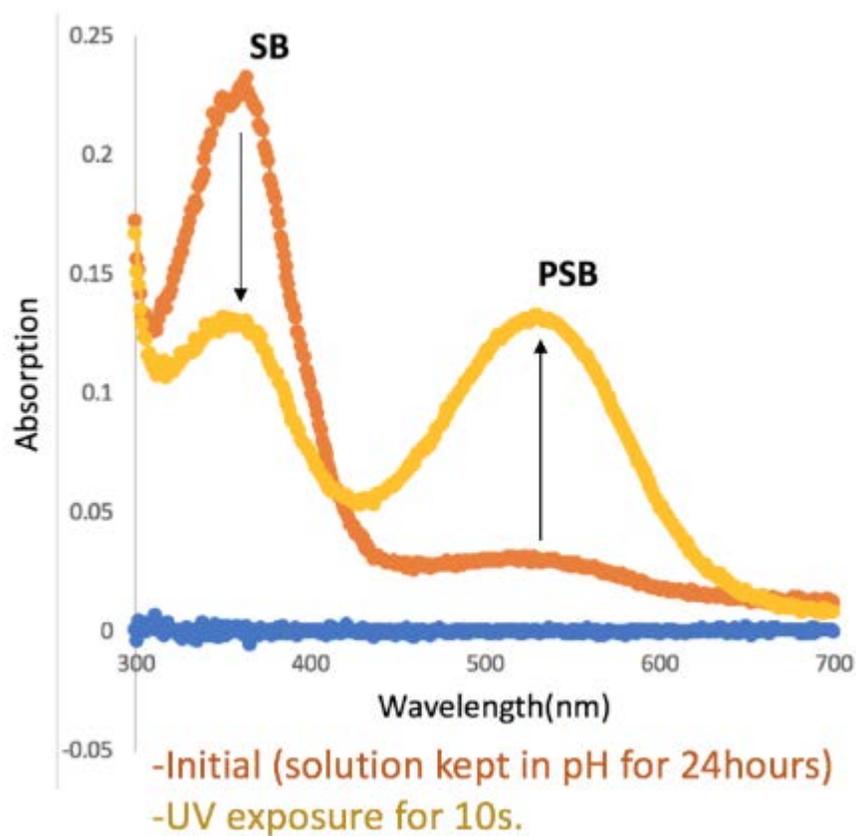


Figure II-38: UV exposure to the solution of M1;Q4A;Q38L convert the SB to PSB.

Measurements are in pH4.

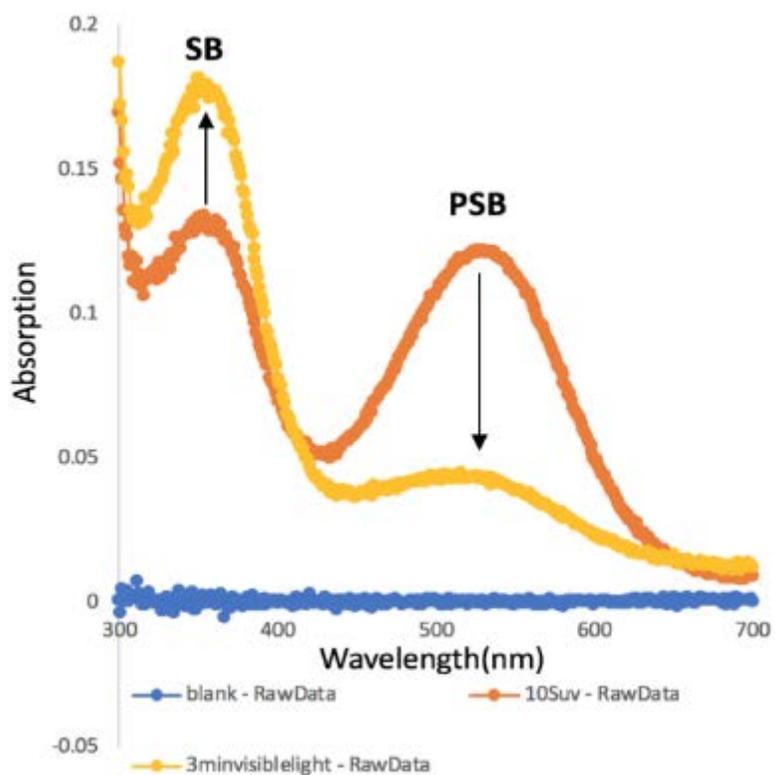


Figure II-39: Converting the PSB back to SB by visible light. (pH4).

Also change in the color of the crystal by UV exposure was detected similar to M1;Q4A (**Figure II-40**).

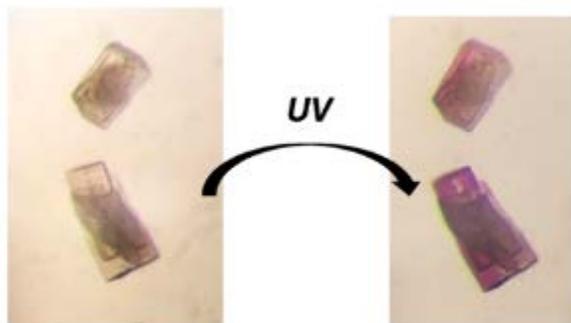


Figure II-40: UV exposure on the M1;Q4A;Q38L crystal.

Interestingly, structural analysis of M1:Q4A:Q38L mutant does not show the isomerization of chromophore. Both colorless and red crystal demonstrate the trans isoform for this mutant. (**Figure II-41**). The reason for changing the color in this crystal is not clear yet.

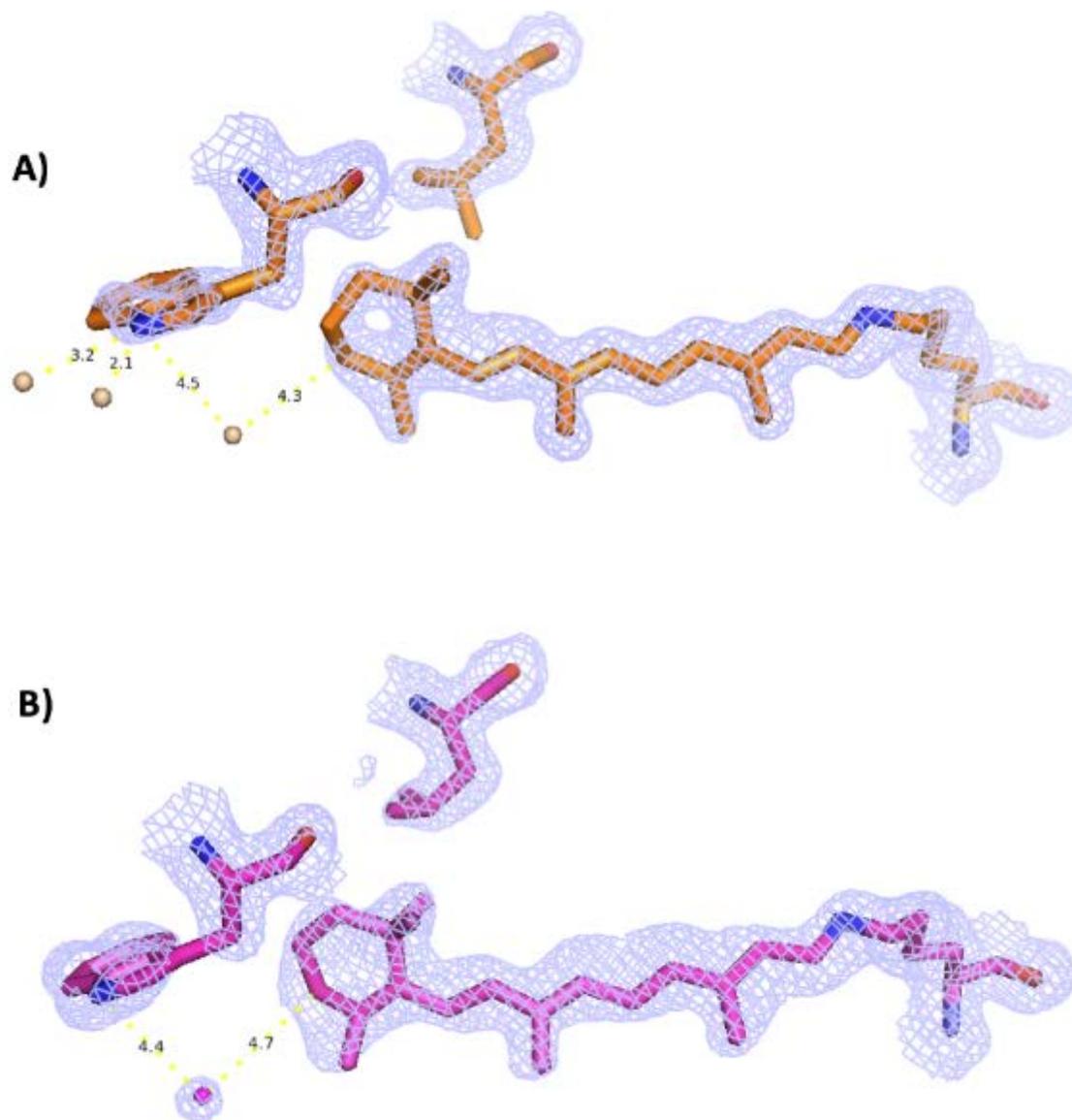


Figure II-41: Crystal structure of M1:Q4A:Q38L A) before UV exposure B) after UV exposure.

Both structures have the all trans retinal conformation.

II-5-8 Photo switching experiment in the presence of Q4 in some different mutant of hCRBP II

Photo switching experiments in solution have been done for retinal bound to some other constructs of hCRBP II such as Q108K:K40L and Q108L:K40L:T51V:T51S both having a Gln4. As we expected photo switching was not observed in the presence of the Gln4 and the water molecule in these constructs, demonstrating the stabilization of the high pK_a *cis*-imine by this interaction.

II-6 CONCLUSION

Based on the high-resolution X-ray crystallography and UV-vis spectroscopy, we demonstrated the photoisomerization of the retinal imine bond between two distinct pK_a regimes using the high-resolution X-ray crystallography in our engineered rhodopsin mimic system (hCRBP II). Interestingly, in this system, we also found out that we can change the pK_a of the *trans* isoform by visible light exposure, which is most probably caused by changing the water network in the vicinity of the chromophore. The fact that water molecules far from the imine region could have a huge effect on the pK_a of the imine is highly important in understanding the mechanism and nature of isomerization in the rhodopsin system. This designed reengineered hCRBP II as a photo switchable rhodopsin mimic can be used as a template for studying the photo physics of rhodopsin and designing photo switchable fluorescent protein tags.

II-7 EXPERIMENTAL

II-7-1 Site directed mutagenesis, protein expression and purification

Site directed mutagenesis was performed by Prof. Wenjing Wang. Refer to her thesis for details.³⁰ Protein expression and purification were performed identical as explained in chapter I.

II-7-2 UV-vis Measurements

All of the UV-vis measurements have been done at the designated pH 4 using PBS buffer. Deprotonated imine peaks (SB) appear at ~ 360 nm. Peaks with Abs max > 450 nm are considered PSB peaks.

II-7-3 pK_a determinations

pK_a values are based on the plot of the absorbance change vs pH by using the curve fit as explained before. Absorbance change of PSB during the acidification is plotted against pH.

$$\Delta A = \Delta A_0 / (1 + 10^{[pH - pK_a]})$$

Saturated citric acid was used for the acidification at first. For lower pH 6M HCl was used till complete SB to PSB conversion.

II-7-4 Crystallization

II-7-4-1 For growing crystals in the dark (Trans PSB)

Four equivalents of the retinal dissolved in Ethanol added to the concentrated protein at 15mg/ml for 2 hours in the dark. Vapor diffusion crystallization performed in the 24 well crystallization plates were performed as described in chapter I.

II-7-4-2 For SB products

Crystals grown in the presence of visible light. Crystals were frozen under the light of microscope as well.

II-7-4-3 UV irradiated crystals (PSB product)

Crystals were UV irradiated with a handheld TLC UV lamp for 2-3min in the dark. UV irradiated M1:Q4A crystals became Blue (622nm), and M1:Q4A:Q38L crystals became red (533nm) color.

The UV irradiated crystals were flash-frozen immediately under red microscope light in 12% PEG 3350, 10% NaCl at pH = 4 and 20 % glycerol. The irradiated crystals were in the same space group (P1) as SB crystals (**Table II-1, II-2, II-3, II-4, II-5**).

Table II-1: X-ray crystallographic data and refinement statistics for M1: Q4A mutant bound to all trans retinal in absence of visible light.

| | Q4ADRAKBLUETRANS(DARK) |
|---------------------------------------|--|
| Wavelength | |
| Resolution range | 27.66 - 1.4 (1.45 - 1.4) |
| Space group | P 1 |
| Unit cell | 30.842 35.918 64.062 85.95 86.434 65.334 |
| Total reflections | |
| Unique reflections | 41917 (4314) |
| Multiplicity | |
| Completeness (%) | 85.44 (87.77) |
| Mean I/sigma(I) | |
| Wilson B-factor | 12.34 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 41915 (4314) |
| Reflections used for R-free | 2059 (206) |
| R-work | 0.1859 (0.2141) |
| R-free | 0.2085 (0.2534) |
| Number of non-hydrogen atoms | 2535 |
| macromolecules | 2188 |
| ligands | 44 |
| solvent | 303 |
| Protein residues | 266 |
| RMS(bonds) | 0.006 |
| RMS(angles) | 0.96 |

Table II-2: X-ray crystallographic data and refinement statistics for M1;Q4A bound to all trans retinal in presence of visible light.

| | Q4AVISIBLELIGHTTRANS |
|---------------------------------------|--|
| Wavelength | |
| Resolution range | 29.48 - 2.11 (2.185 - 2.11) |
| Space group | P 1 |
| Unit cell | 30.755 35.773 64.153 86.18 86.478 65.224 |
| Total reflections | |
| Unique reflections | 12563 (1245) |
| Multiplicity | |
| Completeness (%) | 88.28 (88.93) |
| Mean I/sigma(I) | |
| Wilson B-factor | 21.28 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 12558 (1245) |
| Reflections used for R-free | 1265 (125) |
| R-work | 0.1825 (0.1974) |
| R-free | 0.2671 (0.2972) |
| Number of non-hydrogen atoms | 2390 |
| macromolecules | 2180 |
| ligands | 44 |
| solvent | 166 |
| Protein residues | 266 |
| RMS(bonds) | 0.009 |
| RMS(angles) | 1.19 |

Table II-3: X-ray crystallographic data and refinement statistics for M1:Q4A bound to 15 *cis* retinal after UV exposure.

| | CISQ4A |
|---------------------------------------|---|
| Wavelength | |
| Resolution range | 21.28 - 1.261 (1.306 - 1.261) |
| Space group | P 1 |
| Unit cell | 30.923 36.038 64.026 86.203 86.555 65.059 |
| Total reflections | |
| Unique reflections | 58166 (5523) |
| Multiplicity | |
| Completeness (%) | 86.21 (81.77) |
| Mean I/sigma(I) | |
| Wilson B-factor | 13.07 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 58153 (5523) |
| Reflections used for R-free | 1959 (192) |
| R-work | 0.2042 (0.2444) |
| R-free | 0.2321 (0.2851) |
| Number of non-hydrogen atoms | 2513 |
| macromolecules | 2184 |
| ligands | 52 |
| solvent | 277 |
| Protein residues | 266 |
| RMS(bonds) | 0.006 |
| RMS(angles) | 0.97 |

Table II-4: X-ray crystallographic data and refinement statistics for M1;Q4A;Q38L bound to all *trans* retinal.

| | M1:Q4A;Q38L YELLOW (BEFORE UV EXPOSURE) |
|---------------------------------------|---|
| Wavelength | |
| Resolution range | 28.61 - 1.401 (1.451 - 1.401) |
| Space group | P 1 |
| Unit cell | 30.991 36.091 64.047 86.107 86.561 65.368 |
| Total reflections | |
| Unique reflections | 47394 (4469) |
| Multiplicity | |
| Completeness (%) | 95.93 (91.28) |
| Mean I/sigma(I) | |
| Wilson B-factor | 17.34 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 47382 (4467) |
| Reflections used for R-free | 1957 (194) |
| R-work | 0.1962 (0.2761) |
| R-free | 0.2090 (0.2756) |
| Number of non-hydrogen atoms | 2519 |
| macromolecules | 2186 |
| ligands | 48 |
| solvent | 285 |
| Protein residues | 266 |
| RMS(bonds) | 0.007 |
| RMS(angles) | 1.01 |

Table II-5: X-ray crystallographic data and refinement statistics for M1;Q4A;Q38L bound to all trans retinal after UV exposure.

| | M1;Q4A;Q38L RETINAL PINK |
|---------------------------------------|--|
| Wavelength | |
| Resolution range | 31.86 - 1.5 (1.554 - 1.5) |
| Space group | P 1 |
| Unit cell | 30.979 36.479 63.896 90.812 93.401 115.905 |
| Total reflections | |
| Unique reflections | 38387 (3600) |
| Multiplicity | |
| Completeness (%) | 95.53 (89.06) |
| Mean I/sigma(I) | |
| Wilson B-factor | 22.37 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 38381 (3599) |
| Reflections used for R-free | 2001 (182) |
| R-work | 0.1980 (0.2625) |
| R-free | 0.2232 (0.2865) |
| Number of non-hydrogen atoms | 2419 |
| macromolecules | 2186 |
| ligands | 56 |
| solvent | 177 |
| Protein residues | 266 |
| RMS(bonds) | 0.008 |
| RMS(angles) | 1.08 |

REFERENCES

REFERENCES

1. Nossoni, Z. *et al.* Structures of holo wild-type human cellular retinol-binding protein II (hCRBP II) bound to retinol and retinal. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 3226–3232 (2014).
2. Vaezeslami, S., Mathes, E., Vasileiou, C., Borhan, 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.* **363**, 687–701 (2006).
3. Berbasova, T. *et al.* Light-Activated Reversible Imine Isomerization: Towards a Photochromic Protein Switch. *Chembiochem* **17**, 407–414 (2016).
4. Vasileiou, C. *et al.* Protein Design: Reengineering Cellular Retinoic Acid Binding Protein II into a Rhodopsin Protein Mimic. *J. Am. Chem. Soc.* **129**, 6140–6148 (2007).
5. Wang, W. *et al.* Tuning the electronic absorption of protein-embedded all-trans-retinal. *Science* **338**, 1340–1343 (2012).
6. Ghanbarpour, A. *et al.* Mimicking Microbial Rhodopsin Isomerization in a Single Crystal. *J. Am. Chem. Soc.* **141**, 1735–1741 (2019).
7. Nosrati, M., Berbasova, T., Vasileiou, C., Borhan, B. & Geiger, J. H. A Photoisomerizing Rhodopsin Mimic Observed at Atomic Resolution. *J. Am. Chem. Soc.* **138**, 8802–8808 (2016).
8. Bowmaker, J. K. & Dartnall, H. J. Visual pigments of rods and cones in a human retina. *The Journal of Physiology* **298**, 501–511 (1980).
9. Ernst, O. P. *et al.* Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. *Chem. Rev.* **114**, 126–163 (2014).
10. Movement of Retinal Along the Visual Transduction Path | Science. <https://science.sciencemag.org/content/288/5474/2209.abstract>.
11. Choe, H.-W. *et al.* Crystal structure of metarhodopsin II. *Nature* **471**, 651–655 (2011).
12. Zimmermann, K., Ritter, E., Bartl, F. J., Hofmann, K. P. & Heck, M. Interaction with transducin depletes metarhodopsin III: a regulated retinal storage in visual signal transduction? *J. Biol. Chem.* **279**, 48112–48119 (2004).
13. Palczewski, K. *et al.* Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science* **289**, 739–745 (2000).

14. Lanyi, J. K. Bacteriorhodopsin. *Annu. Rev. Physiol.* **66**, 665–688 (2004).
15. Nogly, P. *et al.* Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. *Science* **361**, (2018).
16. Rothschild, K. J. FTIR difference spectroscopy of bacteriorhodopsin: Toward a molecular model. *J Bioenerg Biomembr* **24**, 147–167 (1992).
17. Ion-pumping microbial rhodopsins.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4585134/>.
18. The Origin of Bond Selectivity and Excited-State Reactivity in Retinal Analogues | The Journal of Physical Chemistry A. <https://pubs.acs.org/doi/10.1021/acs.jpca.6b00701>.
19. Govorunova, E. G., Sineshchekov, O. A., Li, H. & Spudich, J. L. Microbial Rhodopsins: Diversity, Mechanisms, and Optogenetic Applications. *Annu. Rev. Biochem.* **86**, 845–872 (2017).
20. Schobert, B., Cupp-Vickery, J., Hornak, V., Smith, S. & Lanyi, J. Crystallographic structure of the K intermediate of bacteriorhodopsin: conservation of free energy after photoisomerization of the retinal. *J. Mol. Biol.* **321**, 715–726 (2002).
21. Bacteriorhodopsin's M412 intermediate contains a 13-cis, 14-s-trans, 15-anti-retinal Schiff base chromophore. - Abstract - Europe PMC. <https://europepmc.org/article/med/2751988>.
22. Sass, H. J. *et al.* Structural alterations for proton translocation in the M state of wild-type bacteriorhodopsin. *Nature* **406**, 649–653 (2000).
23. Crystal structure of the L intermediate of bacteriorhodopsin: evidence for vertical translocation of a water molecule during the proton pumping cycle. - Abstract - Europe PMC. <http://europepmc.org/article/med/14672661>.
24. Govorunova, E. G., Sineshchekov, O. A., Janz, R., Liu, X. & Spudich, J. L. Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics. *Science* **349**, 647–650 (2015).
25. Nosrati, M. Development of photoswitchable rhodopsin mimics: Spectroscopy, structural studies and wavelength regulation studies of these systems. (Michigan State University, 2015).
26. Manathunga, M. *et al.* Computational and Spectroscopic Characterization of the Photocycle of an Artificial Rhodopsin. *J. Phys. Chem. Lett.* **11**, 4245–4252 (2020).
27. Ghanbarpour, A. Functional Control of Soluble Rhodopsin Mimics Using High Resolution Structure-Based Design and Evaluation. (Michigan State University, 2019).

28. Wang, W., Geiger, J. H. & Borhan, B. The photochemical determinants of color vision. *Bioessays* **36**, 65–74 (2014).
29. Nossoni, Z. Structural insight in to the mechanism of wavelength tuning in a rhodopsin mimic and a single mutation resulted an extensive 3D domain swapped dimerization in hCRBP II. (Michigan State University, 2014).
30. Wang, W. Protein design: Reengineering of cellular retinol binding protein II (CRBP II) into a rhodopsin mimic, functionalization of CRBP II into a fluorescent protein tag and design of a photoswitchable protein tag. (Michigan State University, 2012).

CHAPTER III: STRUCTURE INSIGHTS FOR NEW FLUORESCENT PROTEIN TAGS

Fluorescent proteins can help to visualize biological processes that happen in the living organism. By using fluorescent dyes or fluorescent proteins, scientists can analyze the location, interaction, and dynamics of proteins at the molecular level.^{1 2 3} In this chapter, we first review different types of fluorescent proteins and fluorescent dyes, followed by a discussion of their applications in bioimaging and biosensing.^{4 5} We further discuss some of our designed fluorescent protein tags in collaboration with Prof. Borhan's group and our structural analysis.

III-1 FLUORESCENT PROTEINS

One of the most well-known tools for bioimaging is using the fluorescent protein. The most well-known intrinsic fluorescent protein, a green fluorescent protein (GFP), was discovered in the 1960s from the *Aequorea Victoria* jellyfish by Shimomura.^{6 7 8} Further, Jellyfish GFP has been engineered to produce, blue, cyan, yellow and other emitting colors. Also, fluorescent proteins and non-fluorescent chromoproteins from other species have been identified that emit in the orange, red, and far-red spectral regions.^{9 10} All of the fluorescent proteins are around 25kD, and they believed to have the same fold, and similar mechanism of fluorophore formation.^{11 12 13} The crystal structure of GFP has 11 β -strands with a central α -helix in the middle of the β -barrel. Along the α -helix, the 4-(p-hydroxy-benzylidene)-5-imidazolinone (p-HBI) chromophore forms via the cyclization of Ser65-Tyr66-Gly67 (**Figure III-1**).^{14 15} Based on the crystallographic data, through folding of the protein, the Gly67 amide nitrogen attacks the carbonyl carbon of the Ser65 and

forms the imidazoline-5-one intermediate by dehydration of the hemi-aminal. Further, the conjugated ring structure forms as a result of dehydrogenation of the C α -C β bond of Tyr66.^{15 17 18}

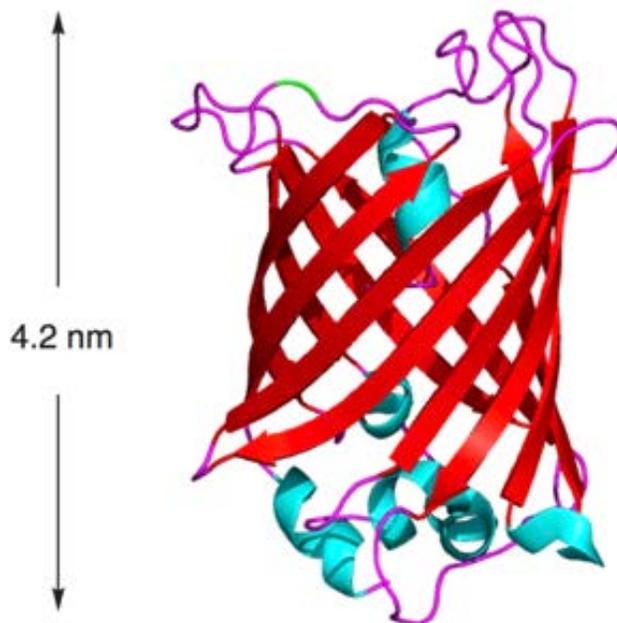
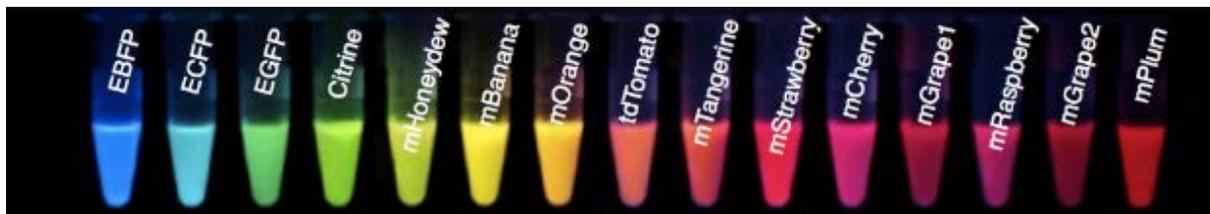


Figure III-1: GFP structure (pdb:1ema).

FP is so far exclusively Metazoan gene, and it has been identified in four different multicellular Cnidaria, Ctenophora, Arthropoda, and Chordata animals. GFP itself represents a small branch of Cnidaria.^{19 20} Most of the intrinsic FPs share a similar β -barrel structure containing the chromophore, and they use the same mechanism using the p-HBI scaffold to form the fluorophore. Some variations of these FPs have been developed via mutagenesis.²¹ Different fluorescence spectra depend on the molecular structure of the chromophore, and its interaction with the surrounding residues of the protein. Altering the hydrophobicity of side chains and the water network in the vicinity of the chromophore lead to formation of new artificial FPs with different spectroscopic behavior.^{12 22 23} Enhancement of the genetic variants of GFP leads to a broad range

of spectrum in the visible region. Different categories of fluorescent protein are blue FP (440-470nm), Cyan (471-500nm), Green (501-520nm), Yellow (521-550nm), Orange (551-571nm), Red (576-610nm), and far-red (611-660nm). Current FPs span emission wavelengths ranging from 442nm to 645nm (**Figure III-2**).⁹ Far-red and red fluorescent proteins are more favorable due to better tissue penetration and less damage to the cells.^{24 25}



*Figure III-2: Shows some of the fluorescent proteins with different emission wavelengths.*²⁶

Some of these FPs have a large Stokes shift such as Sapphire, mKeima, LSS-mKate2, LSSmOrange, CyOFP1, Sandercyanin. More than 100nm difference between excitation and emission of the fluorophores is considered to be a large Stokes shift (LSS).^{27 28 29} Another category of FP are photoactivatable FPs such as PA-GFP, PamKate. The fluorescence is turned on by an activation beam.^{30 31} In some FPs known as photoconvertible FPs, fluorescence wavelength can be irreversibly converted to another by a conversion beam such as mMaple, mEos2, Dendra2. Also, in some FPs like Padron, Dronpa, Dreiklang, and Kohinoor, fluorescence wavelengths can be reversibly switched between two distinct states.^{32 33} They are called as reversibly photo switchable FPs (rsFBs). Multicolor labeling for imaging can be achieved using different LSS and photoswitchable fluorescent proteins.^{34 35 36}

III-2 APPLICATION OF FLUORESCENT IMAGING

Fluorescent proteins have a variety of applications and they are mostly used in fluorescent bioimaging. They are used as an important tool to study cellular localization, protein trafficking,

protein-protein interaction, protein dynamics, gene expression, drug screening, signaling, etc for in vivo imaging.^{37 38} Maturation time, turnover rate, environmental sensitivity, blinking, and bleaching are important factors to choose the proper FP for the study.^{39 40 41} These photophysical properties bring variety of new applications in imaging studies. fluorescence recovery after photobleaching (FRAP), fluorescence loss in photobleaching (FLIP), and fluorescence localization after photobleaching (FLAP) are important methods that exploit photobleaching properties for FP migration in sample cells.^{42 43} Turnover is another photophysical property which is used to make Timer FPs.⁴⁴ In this type of FP, the FPs initially emits at a specific wavelength and emission wavelength changes after a certain time. Förster resonance energy transfer (FRET) is another technique which is widely used for protein-protein interaction. This method is based on the energy transfer between two light sensitive molecules. In this technique, after the donor chromophore goes to its excitation state, it can transfer energy to the acceptor nearby (less than 10 Å) chromophore by dipole-dipole coupling.^{45 43 46} Intramolecular FRET also can be used for detecting conformational changes as a result of protein-protein interaction. Another intramolecular FRET, like Cameleon for measuring intracellular Ca^{2+} and phosphorylation-sensitive FRET probes have been constructed as well (**Figure III-3**).^{47 48}

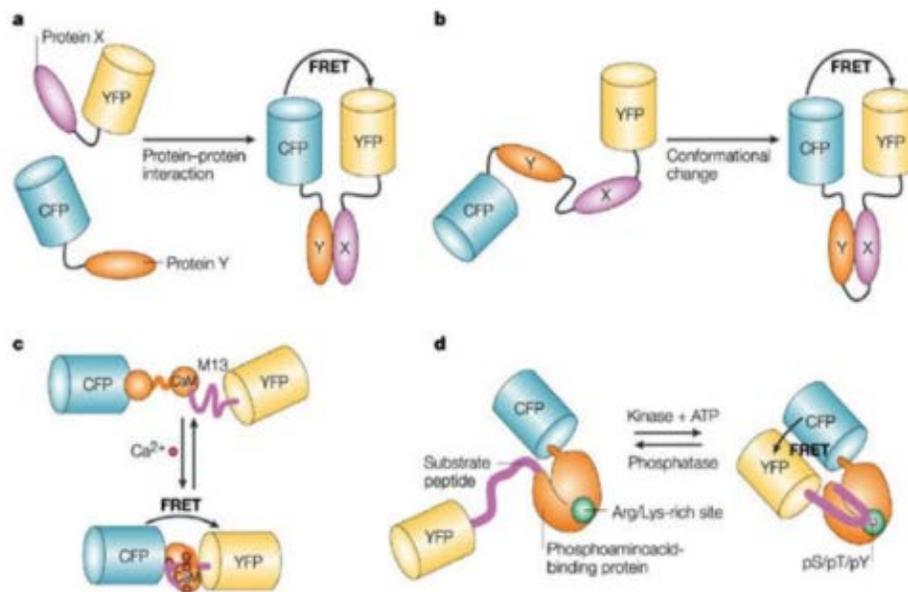


Figure III-3: a) intermolecular FRET b) intramolecular FRET for detecting conformational changes c) cameleon intramolecular FRET probe d) intramolecular phosphorylation-sensitive FRET probes copied from ref.⁴⁸

Protein-protein interaction can be also measured based on fluorescence intensity fluctuation, or blinking. In fluorescence cross-correlation spectroscopy (FCCS), the cross-correlation fluctuation of the two FP emissions can be measured to study the interaction of protein-protein interaction. Unlike FRET, this method is less dependant on the distance and orientations of the targets. An expanded version of this method, photochromic stochastic optical fluctuation imaging (pcSOFI) was used for detection of the activity of Kinase A.⁴⁹ Fluorescent-based biosensing is another application of this area. In this method, based on substrate-indicator recognition, fluorochromes are developed to respond to reaction based indicators, such as metals and redox and ions. Intra molecular charge transfer (ICT) is an example of these methods which interacts with analyte through electron donor acceptor part. Another example of biosensors, is cell membrane voltage sensing with fluorescent indicators. They are used mostly in neuron physiology for measuring ion

currents. Transmembrane adhesion receptors are another category of optical sensing for measuring the mechanical tension of the cell surface. Optical molecular imaging is a growing technique to use in tumor margin detection. Radiography, computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET) are common methods for cancer diagnosis. However, these methods cannot be used for an intraoperation. Surgeons use fluorescence molecular imaging (FMI) for navigating the surgical target, and to evaluate the level of a malignancy of tumor (**Figure III-4**).^{50 51 52} Combination of intraoperative fluorescence-guided surgery (FGS) for tumor removal and radio and chemotherapy helps the survival rate of patients. Two fluorescent dyes, indocyanine green (ICG), and fluorescein sodium have received FDA approval.⁵³ Most of the recent studies have focused on NIR-FGS. NIR-FGS are invisible and they do not interfere with appearance of target in surgical operations.⁵⁴ Also, they are less harmful to cells. photosensitizing-based chromophore/fluorophore-assisted light inactivation (CALI/FALI)⁵⁵ and photodynamic/photothermal therapy (PDT/PTT) are the two multimodality fluorescence bioimaging techniques that are currently being studied in this field.⁵⁶

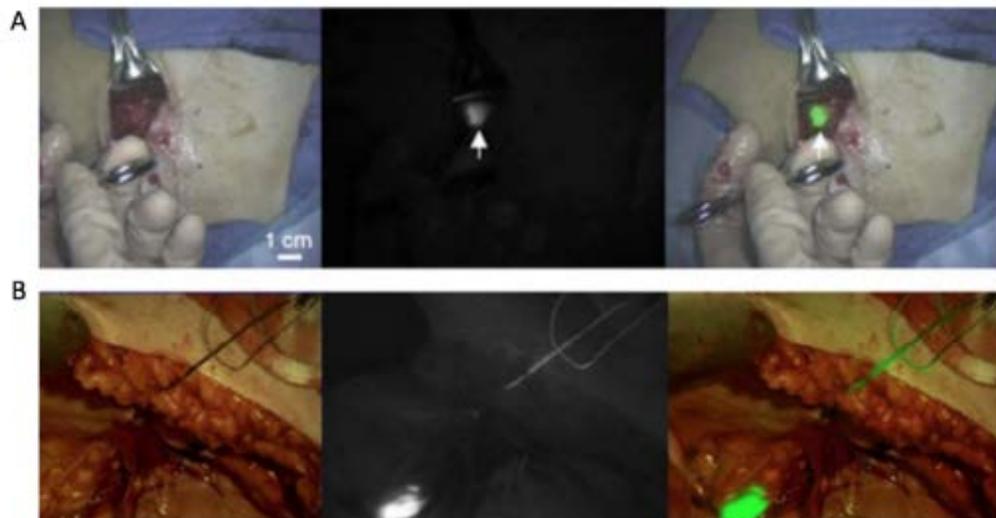


Figure III-4 a) single SLN identification by intraoperative fluorescence imaging b) detection of lymph node by multispectral FMI system.⁵⁰

Another new imaging agents are fluorescence polymer nano particles. They can enhance the brightness of imaging. The brightness of nanoparticles can be improved by increasing the number of particles. Due to the large size of nano particle dyes, they are usually more favorable in the whole animal labeling, coating, and sensing.^{57 58}

III-3 SYNTHETIC FLUOROCHROMES AND DYES

Although fluorescent proteins have been used in a variety of applications, they have some limitations. Unlike GFP, most of the FPs form dimer, or higher order oligomers when they are fused to other proteins.^{10 59} Also, fluorescent proteins needs oxygen to form their chromophore and they can be used only in obligate anaerobes. Due to limitation of FPs, synthetic fluorochromes and dyes have also been developed. In extrinsic fluorescent proteins, an exogenous ligand (fluorophore) binds to the endogenous biomolecule. UnaG for example does not have any intrinsic fluorescence, the fluorescence is triggered by binding to bilirubin.⁶⁰ Bacterial phytochromes photoreceptors (BphPs) binds to biliverdin IX α (BV) exhibit red-shifted NIR absorbance. BphPs bind BV in GAF domain followed by formation of a thioether bond with Cys12 in the PAS domain (**Figure III-5**).⁶¹ The engineered BphPs used for designing other fluorescent proteins.^{62 63}

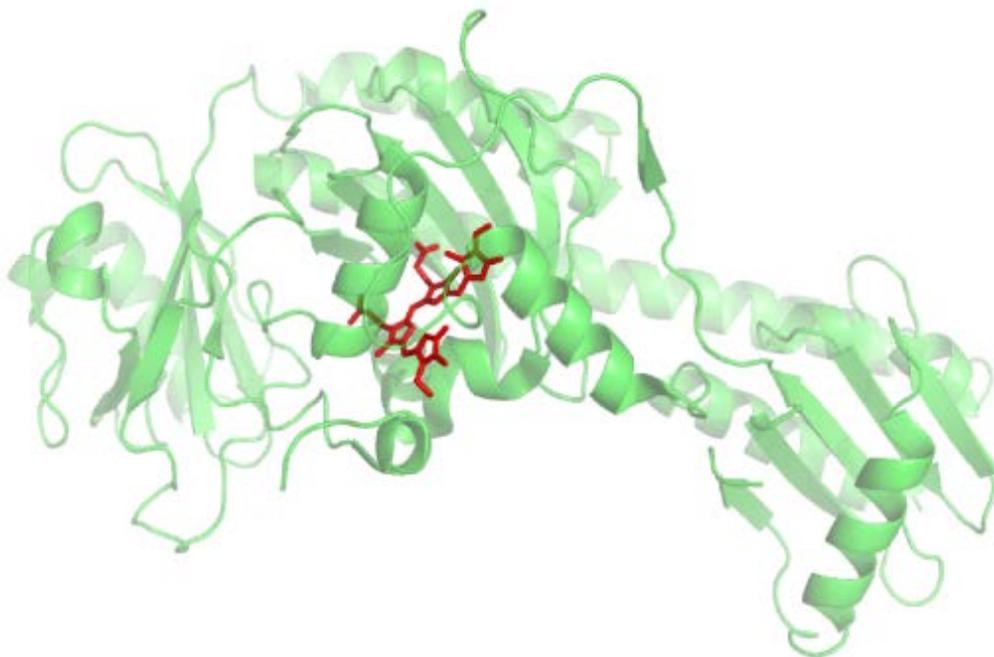


Figure III-5: crystal structure of BphPs bound to BY (pdb:3c2w).⁶¹

One of the most developed methods to make extrinsic fluorescent proteins is a chemical-labeling system. In this method, a tag fused with a target protein covalently interacts with a small molecule that plays as fluorophore after the binding. Variety of NIR fluorescent tags have been developed using this method. SNAP-tag, CLIP-tag, Halo-Tags are some of the examples of enzymes that used as a tag to a target protein in this method. One of the limitations of this method is that this method needs sequential two-substrate enzymatic reactions.^{64 65 66} SNAP-tag is an engineered O6-alkylguanine-DNA alkyl transferase that can catalyze the irreversible alkyl group transfer of O6-alkylguanine- or O6-benzylguanine-DNA. In this method, the dye is attached to the O6-benzylguanine substrate.^{67 68} CLIP-tag is engineered from the same enzyme, and uses O2-benzylcytosine as a substrate⁶⁹ (**Figure III-6**).

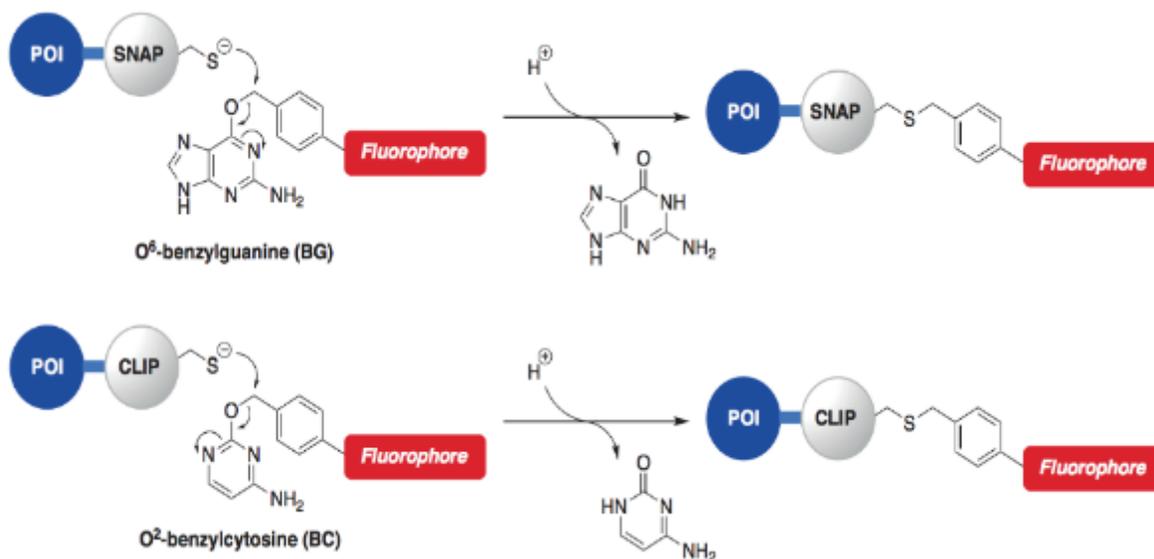


Figure III-6: Demonstrates the reaction of snap-tag, and Clip-tag.poi (protein of interest).⁷⁰

Fluorogenic TMP-tag also developed based affinity-based ligand-protein interactions. Unlike the aforementioned covalently self-modified enzymatic tags, this tag was developed based on noncovalent binding affinity of trimethoprim (TMP) with *E. coli* dihydrofolate reductase.^{71 72 73}

Similar to site specific chemical labeling, biorthogonal reactions have been used for designing new dyes. In specific chemical labeling, the protein of interest is labeled with a protein (usually around 200 amino acids) as tag. Due to the large size of the tag, the tag might interfere with the protein of interest. In contrast, instead of using a protein as a tag, the target protein is labeled with a functional groups. Developing the smart azide probes using Staundinger Ligation and Diels-Alder Cycloadditions are some of the examples of these probes. These methods usually needs many washing steps to remove the excess free fluorochromes. One of the recent developments in this area is using the quencher conjugated with the fluorescent probe. By doing so, the fluorophore remains in the dark state until the dye binds to the target. Small protein tags is one of the category of fluorogenic proteins which used as alternative for intrinsic fluorescent protein. Fluorogen or

dye remains in the dark state till they bound to their protein target. Fluorogen-activating proteins (FAPs) are a category of these fluorogenic proteins. They are single chain antibodies that reacts to a fluorogen. Thiazole orange and malachite green are two examples of these fluorogens (**Figure III-7**).⁷⁴ Due to the free rotation of their aromatic rings, they do not have emissions in the solution. After binding to the FAPs, the molecules become stable in their fluorescent states and they become activated.

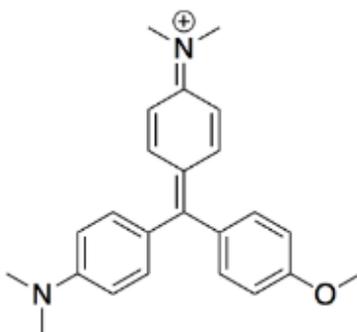


Figure III-7: Malachite green (mg).

Studies demonstrate great results for malachite green (MG) fluorogen. They are bright, wash-free labels with emission spectra in the far-red region. These properties make them favorable for imaging studies. FAP systems are used for multicolor imaging as well.^{74 75} Photoactive yellow protein tags (PYPs) are another category of small protein tags. PYP tags are engineered from photoactive yellow protein. PYP is a small, 125 amino acid, cytosolic protein (**Figure III-8**). The probe is based on the mechanism of static quenching, and it does not have emission in absent of PYP.⁷⁶ In the presence of PYP, the quencher (coumarin) dissociates from the fluorescent molecules. The coumarin based fluorophores usually have slow kinetics and they need wash steps. Therefore, some designed probes were based on 7-dimethylaminocoumarin thioester derivatives have been developed to improve the kinetics and removing the washing steps. PYP can be fused with nuclear localization signals (PYP-NLS) for cell imaging studies.^{77 78 79 80}

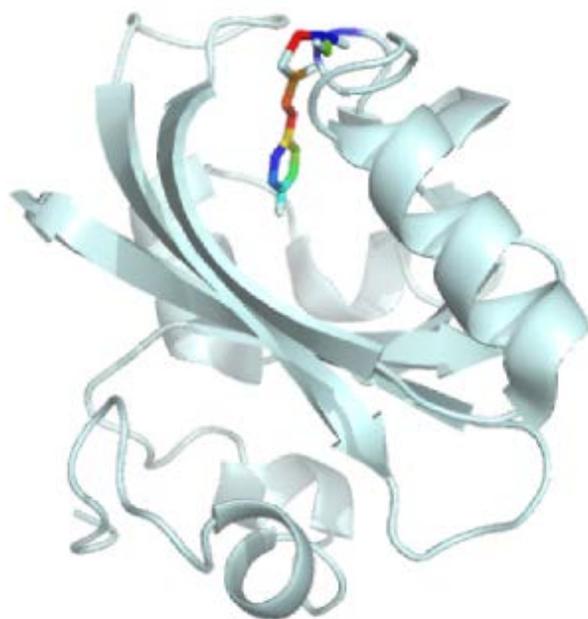


Figure III-8: Photoactive yellow protein tags (PYPs) bound to 4- hydroxycinnamic acid
(*pdb:2phy*).⁷⁹

Another small fluorescent protein is hCRABPII designed by collaboration between our lab and Prof. Borhan at Michigan State University. As explained in the previous chapter, we have used this system for many years to study wavelength tuning and isomerization similar to rhodopsin systems, also to design new fluorophores for imaging application.^{81 82} In CRABPII protein merocyanine binds to a lysine residue in the binding pocket (**Figure III-9**). Although this system has many advantageous such as high quantum yield, attempts to use this system for in vivo studies failed. Therefore, our new efforts for designing new dyes are based on the hCRBPII template. Many fluorophores have been designed based on hCRBPII by Prof. Borhan's group. Among them, I had the opportunity to work with FR1, FR0-Maleimide derivate, FR1-cyano, ThioFluor and Thio phenol chromophores.^{83 70} We will get into the details of each these chromophores in this chapter.

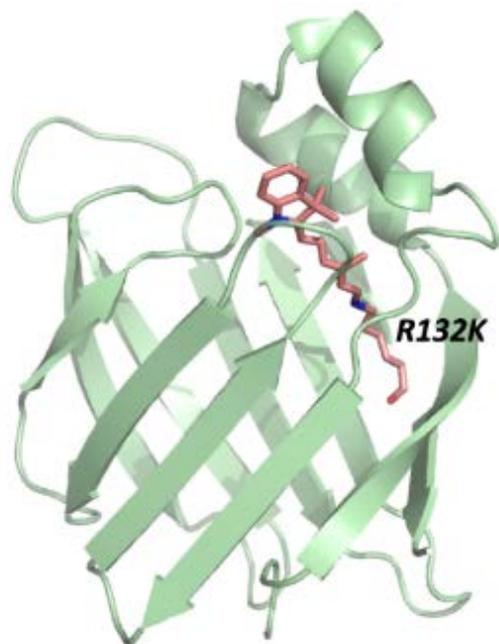


Figure III-9: Merocyanine retinal aldehyde (MCRA) bound to hCRABPII. ⁸⁴

Resolution and imaging depth are important factors in fluorescence imaging. By development of super-resolved fluorescence microscopy higher imaging resolutions are possible. NIR emissions still have the depth issue. Also, water and hemoglobin have the same absorptions in the visible and NIR regions. In the past decades there are many efforts in developing far-Red and NIR dyes with better depth and resolution. ^{85 86 87}

III-4 FR-1V DYE AND PHOTO SWITCHING MECHANISM

Another Chromophore that I had the opportunity to work on is FR-1V made by Dr. Wei Sheng. In an effort to make NIR emission and a large Stokes shift he developed this dye. FR-1V makes a Schiff base with Lys108 in hCRBPII similar to retinal chromophore (**Figure III-10**). ⁷⁰

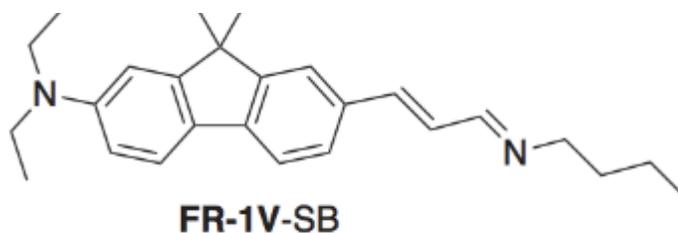


Figure III-10: FR-1v bound to Lys108.

To make this dyes useful for imaging studies, it needs to have low background for in vivo studies. Therefore, the emission of hCRBP2 complex should be different than FR-1V bound to other non-specific proteins. Since, emission of FR-1V-PSB (690nm) and FR-1V-SB (510nm) are relatively different in Ethanol, Dr. Wei Sheng tried to find a method to control the FR-1V complex exclusively as PSB with hCRBP2 and SB with other proteins. One of the mechanisms for controlling the amount of the PSB/SB is isomerization, which was explained in the last chapter. In case of retinal bound to hCRBP2 and hCRBP2 systems, we can increase the PSB/SB by UV exposure. By the time of this study, only isomerization in hCRBP2 system was discovered by Dr. Meisam Nosrati. Dr. Wei sheng tried the same procedure for isomerization of FR-1V bound to hCRBP2.⁹² He found several mutants that photo switch with this chromophore by UV exposure (**Table III-1**). Unfortunately, by the time of this study, none of these mutants were crystalized, and he published his results for two of these mutants, ps4 and ps9, by postulating the same mechanism as isomerization of hCRBP2 system.^{70 93}

Table III-1: Several mutants of hCRBP_{II} that photo switch with this FR-1v by UV exposure.

| Entry | hCRBP _{II} mutant | "OFF" ^a | | "ON" ^b | | "ON/OFF" PSB Abs ratio ^c |
|-------|-------------------------------------|---------------------|--------------------|----------------------|---------------------|---|
| | | $\lambda_{abs, SB}$ | $\lambda_{em, SB}$ | $\lambda_{abs, PSB}$ | $\lambda_{em, PSB}$ | |
| ps1 | Q108K:K40L:T51V:T53S | 370 | 437 | 637 | 710 | 1.8 |
| ps2 | Q108K:K40L:Y19W:R58Y | 395 | 455 | 568 | 674 | 1.2 |
| ps3 | Q108K:K40L:T51V:T53S:Y19W:R58L | 380 | 439 | 625 | 702 | 4.4 |
| ps4 | Q108K:K40L:T51V:T53S:Y19W:R58Y | 378 | 445 | 600 | 686 | 3.3 |
| ps5 | Q108K:K40L:T51V:T53S:Y19W:R58Y:S55Q | 376 | 440 | 616 | 696 | 5.1 |
| ps6 | Q108K:K40L:T51V:T53S:Y19W:R58W:L77W | 379 | 454 | 655 | 730 | 3.3 |
| ps7 | Q108K:K40L:T51V:T53C:Y19W:R58W:A33W | 377 | 429 | 619 | 677 | 2.0 |
| ps8 | Q108K:K40L:T51V:T53S:Y19W:R58W:V62S | 376 | 445 | 606 | 699 | 3.1 |
| ps9 | Q108K:K40L:T51V:T53S:Y19W:R58W:V62N | 378 | 451 | 605 | 611 | 2.6 |

Absorption and emission spectra for ps4 is illustrated in (Figure III-11).

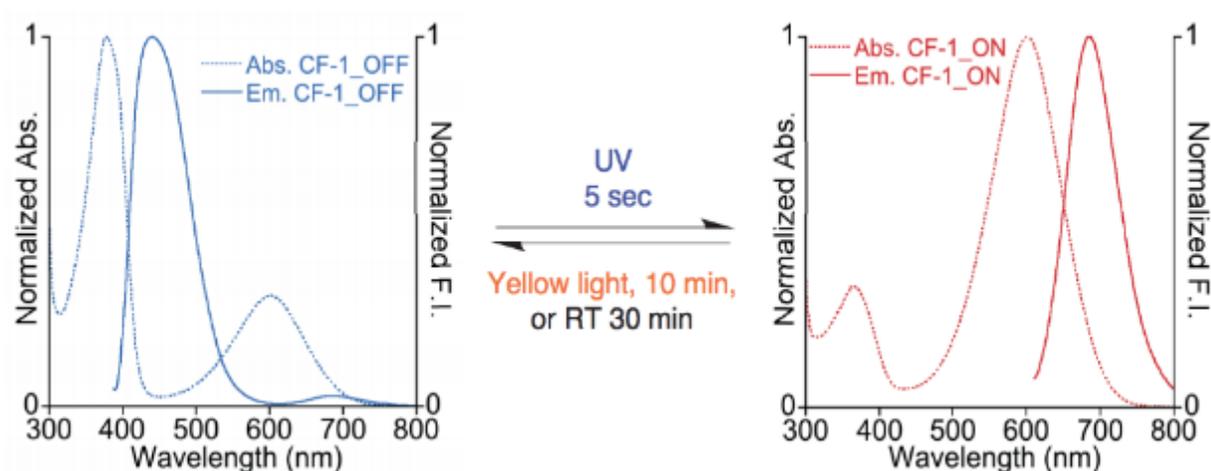


Figure III-11: Absorption and emission for PS4 before (left) and after UV exposure (right).

Absorption in dash line and emission solid line.^{70 93}

He explained his data and photo switching mechanism by referring to 15-*cis* to all-**trans** retinal isomerization in hCRAB_{II} that was discovered by Dr. Meisam Nosrati (Figure III-12).⁹³

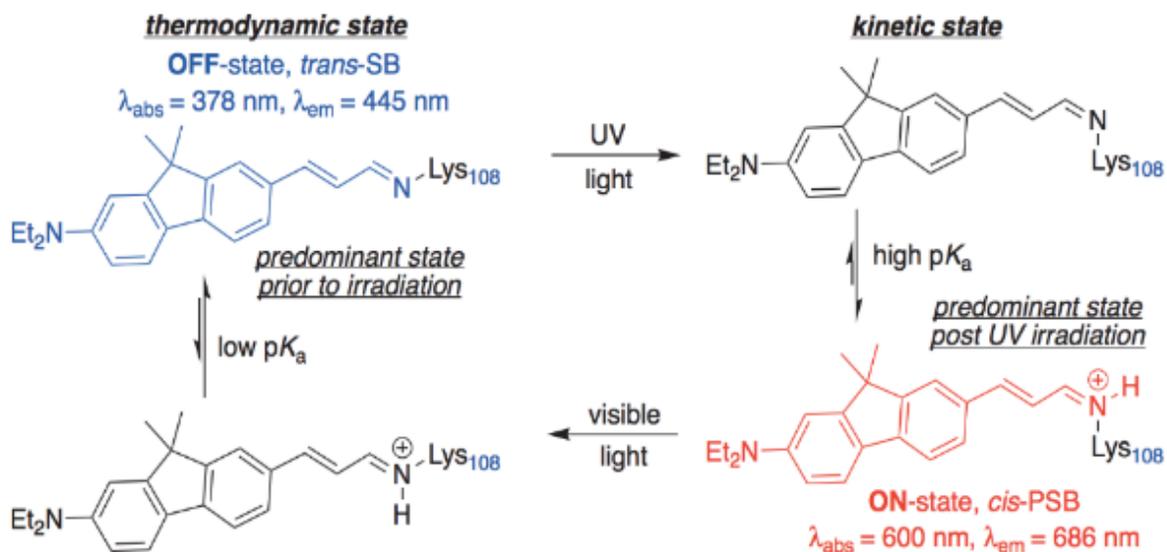


Figure III-12: Postulated mechanism for photoswitching in FR-1V bound to hCRBP1I.⁷⁰

As explained in the last chapter, I discovered the isomerization for 15-*cis* to all-*trans* retinal in hCRBP1I system in the absence of Gln4. The isomerization mechanism was discovered using a single crystal for Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A and following studies have been done also for Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L mutant. Since, I could crystallize these mutants with retinal, I decided to try them also for FR-1V fluorophore to find out the mechanism for photo switching in this dye. I was successful in crystallizing the Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1V. Crystal was colorless, and by UV exposure it changed to blue color, which shows a photo switching process (Figure III-13).

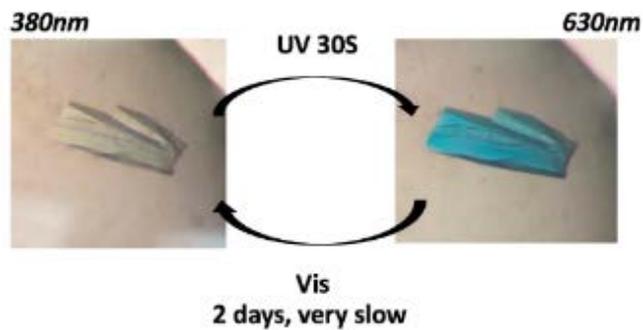


Figure III-13: UV exposure on Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v crystals lead to change in color of crystals.

Crystals were solved in high resolution around 1.4 Å. Interestingly, there is no evidence of the isomerization. Both crystals before and after UV exposure have *trans* conformation. The only difference between these two crystals is moving the Trp58 after shining the UV (**Figure III-14**), (**Figure III-15**).

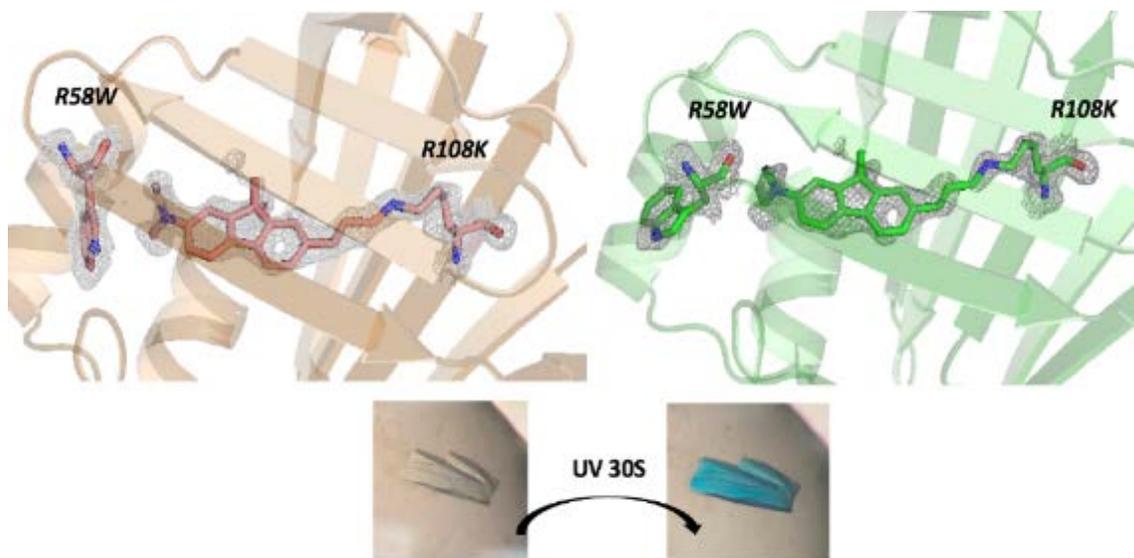


Figure III-14: Crystals of Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v before (left) and after UV exposure (left).

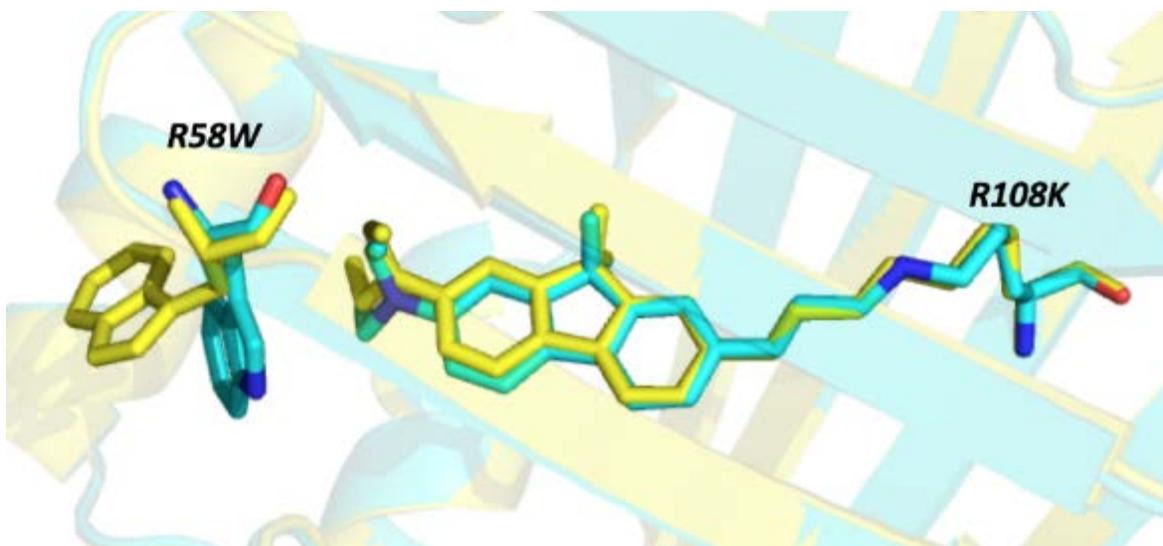


Figure III-15: Overlaid crystals of Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v before (blue) and after UV exposure (yellow). The only difference is changing the conformation of Trp58.

This observation lead us to think about another mechanism for photo switching other than isomerization, since the color change in crystals were clear. As mentioned, crystals of hCRBP II always grown and frozen in pH4. Therefore, the mechanism in pH 4 might be different than physiological pH. The absorption for this construct, Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L, bound to FR-1V in pH is illustrated (**Figure III-16**). At first, there is only SB peak observed with absorption around 380nm. UV exposure to this solution converts some of the SB to PSB (λ_{\max} 630nm). The PSB is not stable and it will convert back to SB thermally.

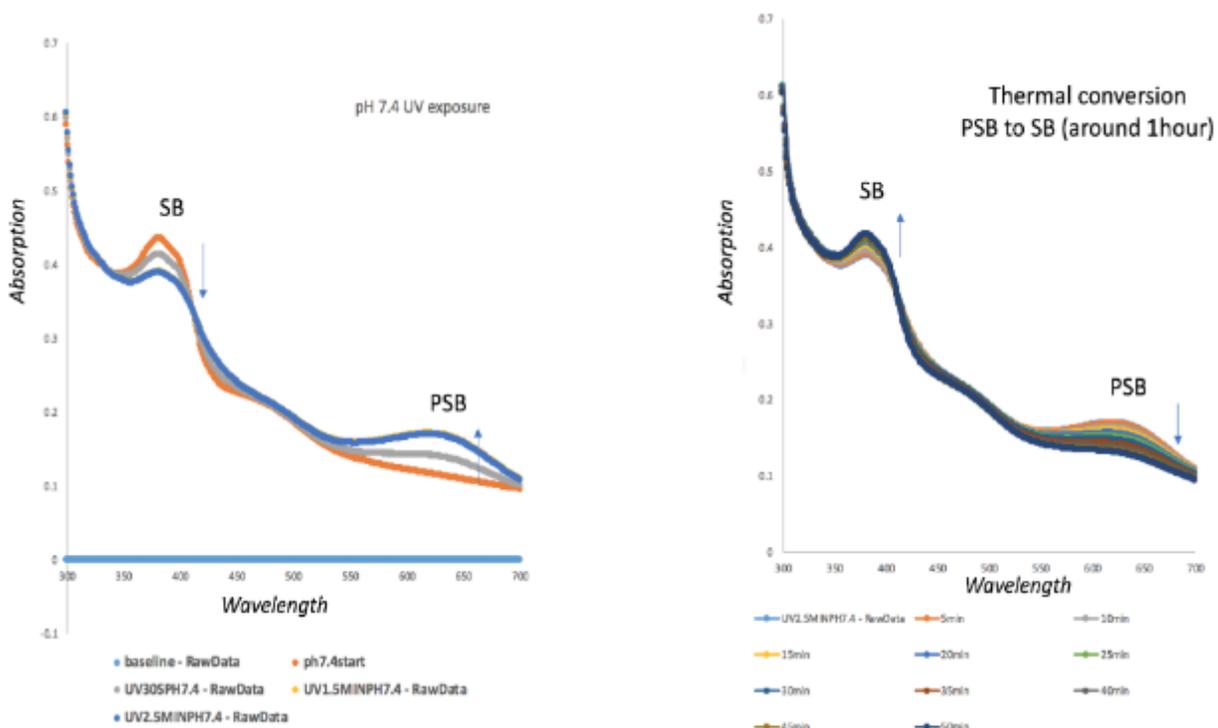


Figure III-16: Absorption spectra for Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-IV in pH7.4. SB converts to PSB via UV exposure (left). PSB converts back to SB overtime thermally (right).

As discussed, crystals were grown in pH 4. Therefore, I acidify the solution of the protein bound to chromophore. Interestingly, acidifying the solution leads to formation of two peaks around 320nm and 340nm, and does not show any formation of PSB (λ_{\max} 630nm) (**Figure III-17**). Our hypothesis is that the peaks around 340 are related to the double protonated form of the chromophore, when both nitrogens of the imine bond and Diethyl amine are protonated. The pK_a is around 3.5 for this form.

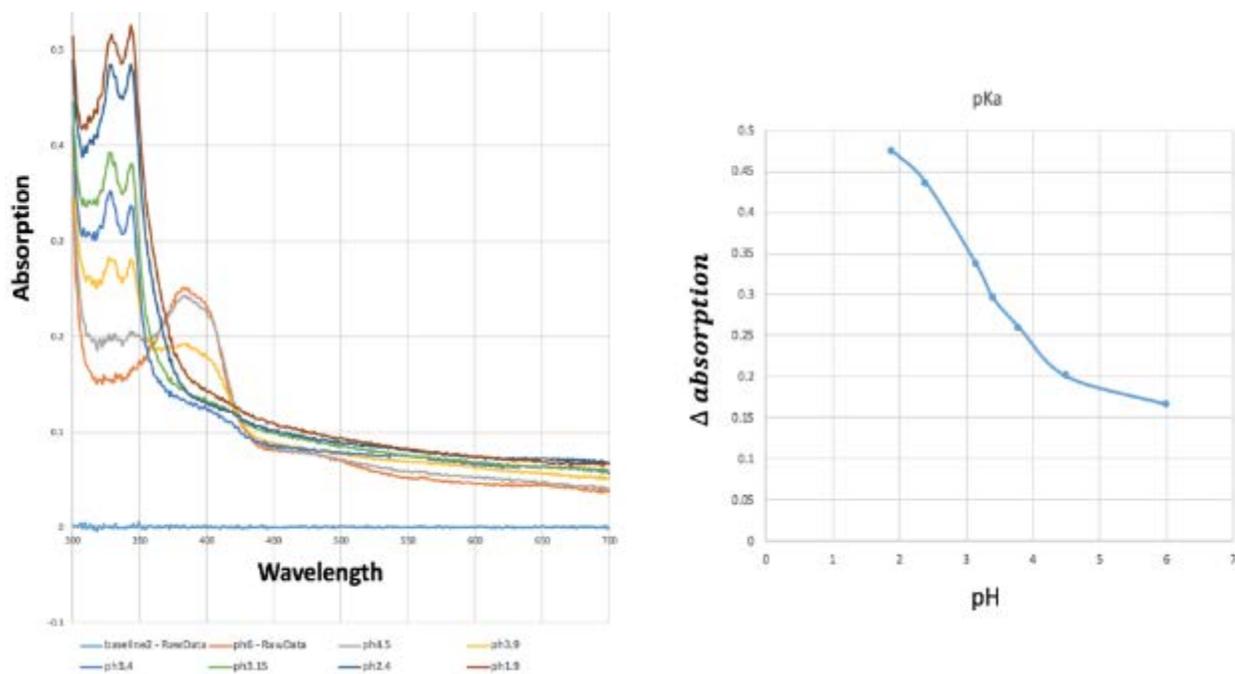


Figure III-17: Acidifying solution of Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v leads to conversion to the SB (λ_{MAX} 380nm) to 320-340nm peak. The pka for this form measured to be around 3.5.

Since the crystals are grown in pH4, I collected the absorption spectra in pH 4 as well. In pH4, there is mixture of the SB (λ_{max} 380nm) and 340nm and 320nm peaks. Shinning the UV light lead to form the PSB (λ_{max} 630nm), which is consistent with the color of crystal as blue. The PSB converts overtime to mostly 320nm and 340nm species (**Figure III-18**). The idea of the 340 peak to be related to double protonated form of the chromophore is consistent with our observation on the crystals. In the colorless crystal, Trp58 may have π -cation interaction with protonated diethyl amine. By UV exposure, and losing the proton, chromophore is in the PSB form, and no longer has the positive charge on the diethyl amine, and Trp58 no longer has the π -cation interaction and it rotate from the close interaction with diethyl amine (**Figure III-18**).

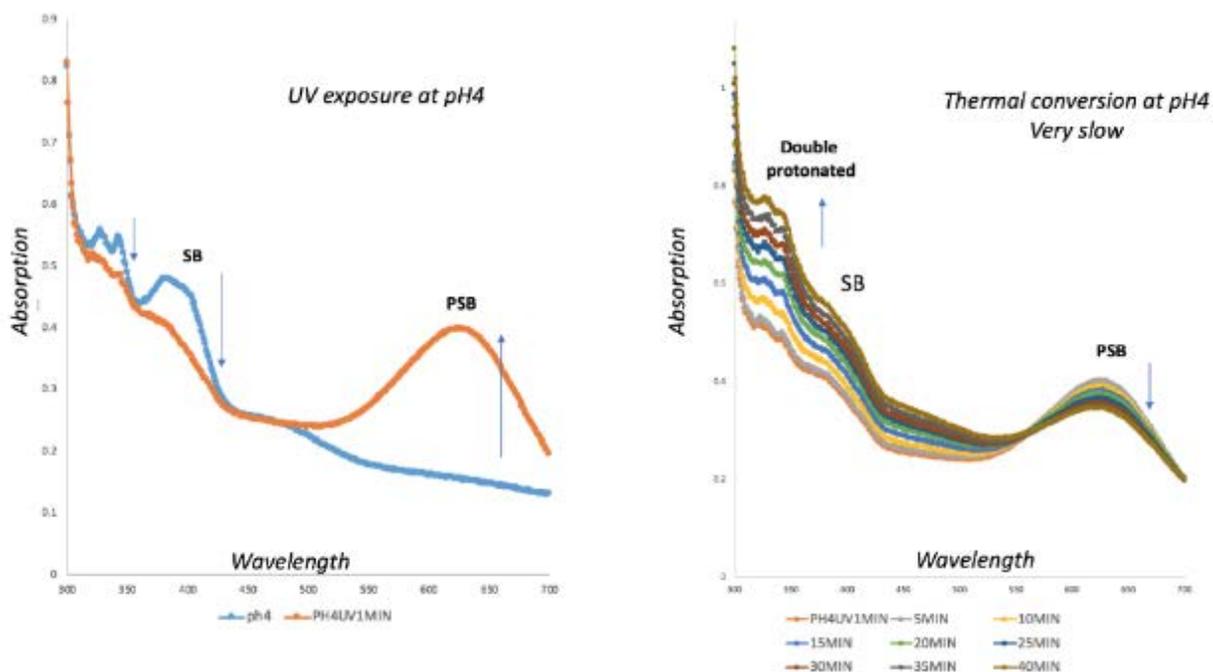


Figure III-18: Absorption spectra for Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v in pH4. SB and 320nm-340nm peaks converts to PSB via UV exposure (left). PSB converts back to mostly to 340-320nm peak overtime thermally (right).

To only have 420nm-440nm peaks, I brought the pH down to 3.2. There is no SB peak observed in this pH. By UV exposure, the PSB forms, but not as much as at pH4. The PSB convert back to 420nm-440nm peak very fast around 5 minutes (**Figure III-19**).

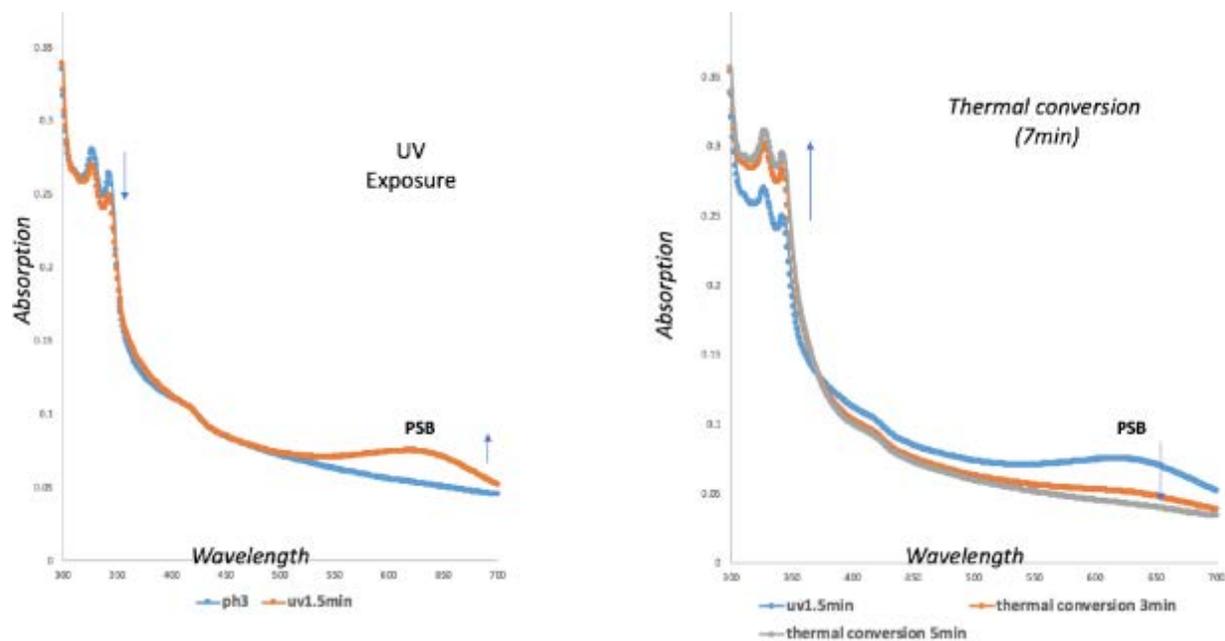


Figure III-19: Absorption spectra for Q108K;K40L;T51V;T51C;R58W;T28L;Y19W;Q4A;Q38L bound to FR-1v in pH3.2. 320nm-340nm peaks converts to PSB via UV exposure (left). PSB converts back to 340-320nm peak in around 5min thermally (right).

The idea of only changing the position of one residue, in our case Trp58, leading to photo switching of the chromophore is pretty interesting and it can be used in designing new photo switchable fluorophores. I have tried to crystallize some other mutants like Q108K;K40L;T51V;T53S and Q108K;K40L;T51V;T53C that are photo switchable with FR-1V (**Figure III-20**).

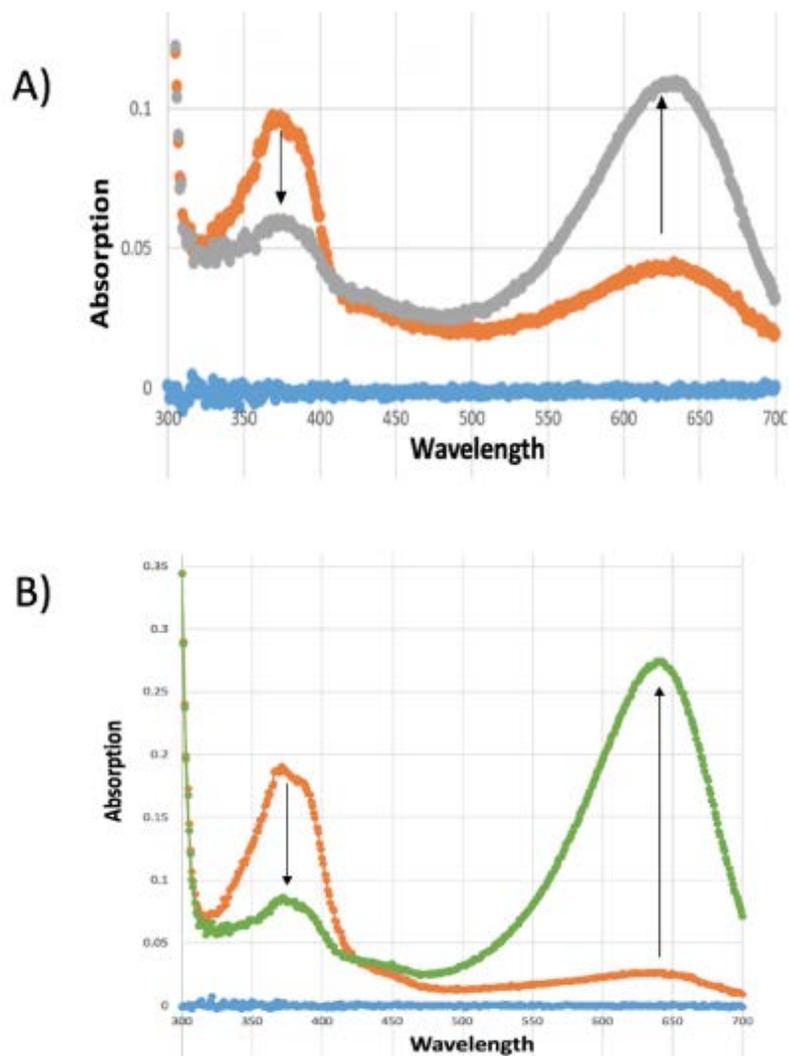


Figure III-20: Conversion of SB to PSB via UV irradiation. A) for Q108K;K40L;T51V;T51C B) Q108K;K40L;T51V;T51S bound to FR-1V in pH4.

I was successful in growing the crystals of Q108K;K40L;T51V;T53S bound to FR-1V. However, the color change in the crystals were not clear, via UV exposure, and there was not enough density for the chromophore in the binding pocket (**Figure III-21**).

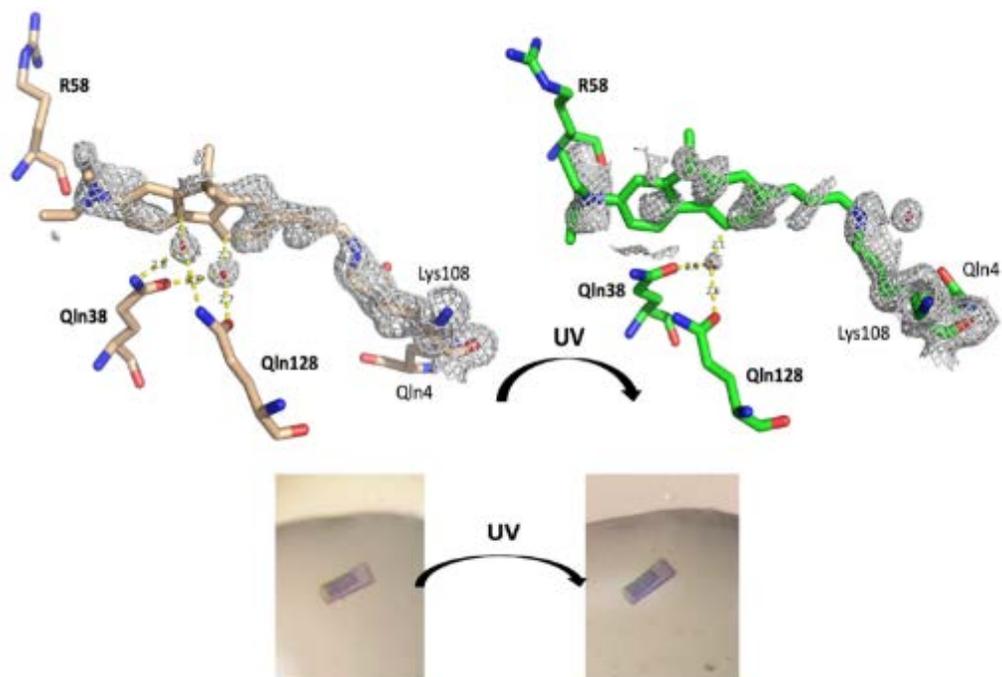


Figure III-21: Crystals of Q108K:K40L;T51V;T53C before and after UV exposure. Density for the chromophore is not sufficient to predict the mechanism of photo switching.

III-5 FR1V- α - CYANO

FR1V- α - cyano is another photo switchable dye synthesized by Soham Maity (**Figure III-22**). FR1V- α - cyano was engineered based on fluorine-based aldehydic dye (FR-1V), which is explained in the last part.

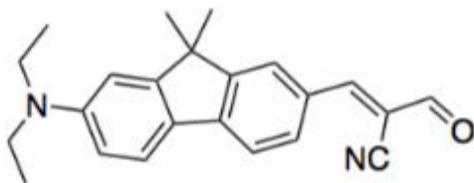


Figure III-22: FR1V- α - cyano structure.

The idea is that the chromophore binds to protein from two sides, making a Schiff base with Lys108 and sulfide bond with Cys 51. By UV exposure, the bond between Cys51 and Lys108 breaks (**Figure III-23**).

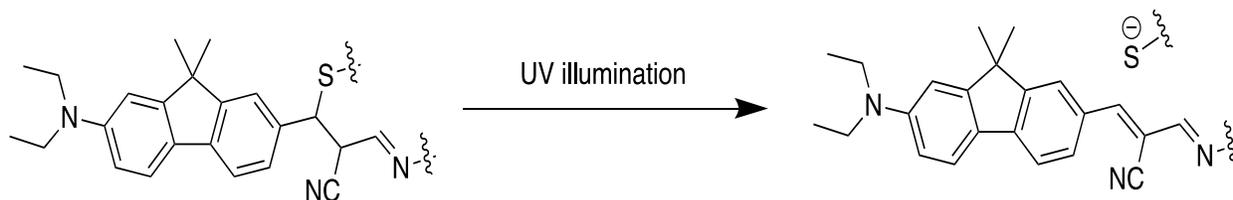


Figure III-23: UV elimination breaks the sulfide bond between the Cys51 and the chromophore.

Absorption spectra for Q108K:K40E:T51C:T53S:R58L:Q38F:Q4F mutant bound to FR1V- α -cyano is illustrated (**Figure III-24**). By UV exposure, the sulfide bond breaks and converts to 445nm which is the absorption of the imine bond. This process is reversible and the 445nm peaks convert to 330nm peaks in visible light, which shows the formation of the sulfide bond again.

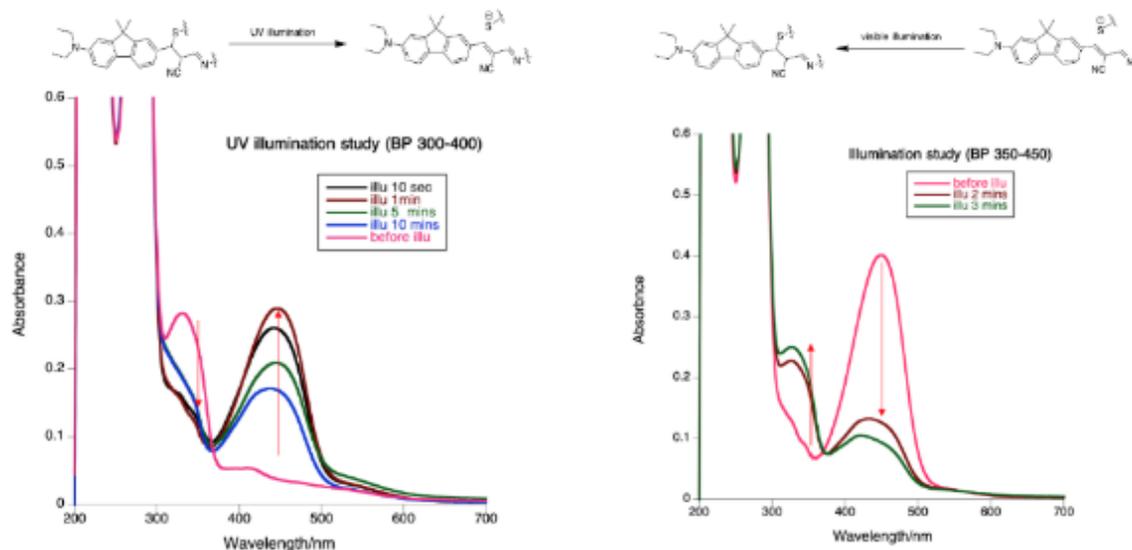


Figure III-24: UV exposure to the 330nm peak converts it to 445nm (left). The 445nm peak converts back to 330nm in the presence of visible light (right).

I have tried many times to crystalized this mutant bound to FR1V- α - cyano. Unfortunately, the electron density for chromophore it was not complete in this crystals as it shows in **(Figure III-25)**.

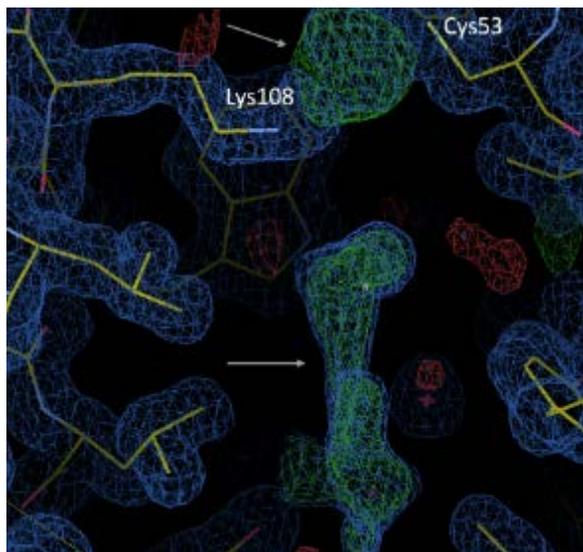


Figure III-25: Electron density for chromophore in Q108K:K40E:T51C:T53S:R58L:Q38F:Q4F construct (green).

Further trials to fit the chromophore in the electron density demonstrates different binding as expected to Cys51 **(Figure III-26)**. Overall, we cannot talk about binding of chromophore with certainty from this crystal data.

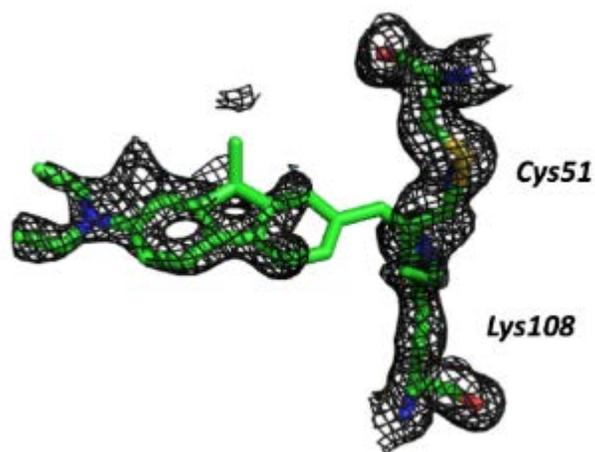


Figure III-26: Electron density for FRIV- α -cyano in Q108K:K40E:T51C:T53S:R58L:Q38F:Q4F construct.

III-6 CONCLUSION

Fluorescent protein tags have various applications specially in imaging studies. We used the hCRBP II templates for designing new fluorescent dyes. Our structural analysis and UV-fluorescent measurements help us to postulate the mechanism for these new designed dyes.

III-7 EXPERIMENTAL

III-7-1 Site directed mutagenesis and protein expression

For Site directed mutagenesis refer to Prof. Wenjing Wang's thesis. Protein expression and purification were performed similar as explained in chapter I.

III-7-2 Crystallization

Four equivalent of ligands (FR1, FR1-cyno, Thiophenol) in Ethanol added to the concentrated protein at 15mg/ml for 2hours in the dark. Vapor diffusion crystallization performed in the 24 well crystallization plates were performed as described in chapter I (**Table III-2, III-3, III-4, III-5, III-6, III-7, III-8**).

Table III-2: X-ray crystallographic data and refinement statistics for *m1:q4a:q38l* bound to fr-

Iv.

| | M1:Q4A:Q38L bound to FR1 (without UV exposure) |
|---------------------------------------|---|
| Wavelength | |
| Resolution range | 28.38 - 1.201 (1.244 - 1.201) |
| Space group | P 1 |
| Unit cell | 31.171 35.859 64.046 85.663 86.043 65.168 |
| Total reflections | |
| Unique reflections | 74191 (7109) |
| Multiplicity | |
| Completeness (%) | 94.77 (91.41) |
| Mean I/sigma(I) | |
| Wilson B-factor | 12.05 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 74189 (7109) |
| Reflections used for R-free | 1997 (195) |
| R-work | 0.1935 (0.2304) |
| R-free | 0.2045 (0.2511) |
| Number of non-hydrogen atoms | 2536 |
| macromolecules | 2182 |
| ligands | 54 |
| solvent | 300 |
| Protein residues | 266 |
| RMS(bonds) | 0.007 |
| RMS(angles) | 0.78 |

Table III-3: X-ray crystallographic data and refinement statistics for M1:Q4A:Q38L bound to *fr-1v* after UV exposure.

| | M1:Q4A:Q38L bound to FR1 (UV exposure) |
|---------------------------------------|--|
| Wavelength | |
| Resolution range | 32.62 - 1.301 (1.347 - 1.301) |
| Space group | P 1 |
| Unit cell | 31.1 35.997 64.013 86.111 86.5 65.102 |
| Total reflections | |
| Unique reflections | 58681 (5702) |
| Multiplicity | |
| Completeness (%) | 95.19 (92.50) |
| Mean I/sigma(I) | |
| Wilson B-factor | 13.50 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 58677 (5702) |
| Reflections used for R-free | 1985 (171) |
| R-work | 0.2081 (0.2680) |
| R-free | 0.2329 (0.2801) |
| Number of non-hydrogen atoms | 2435 |
| macromolecules | 2178 |
| ligands | 66 |
| solvent | 191 |
| Protein residues | 266 |
| RMS(bonds) | 0.007 |
| RMS(angles) | 0.99 |

Table III-4: X-ray crystallographic data and refinement statistics for KLVS bound to FR-1v.

| | KLVS bound to FR1 |
|---------------------------------------|--|
| Wavelength | |
| Resolution range | 29.06 - 1.45 (1.502 - 1.45) |
| Space group | P 1 |
| Unit cell | 29.593 36.087 63.587 90.813 92.292 112.905 |
| Total reflections | |
| Unique reflections | 41270 (4037) |
| Multiplicity | |
| Completeness (%) | 96.15 (94.06) |
| Mean I/sigma(I) | |
| Wilson B-factor | 15.25 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 41265 (4037) |
| Reflections used for R-free | 2005 (200) |
| R-work | 0.2057 (0.2415) |
| R-free | 0.2395 (0.2500) |
| Number of non-hydrogen atoms | 2463 |
| macromolecules | 2176 |
| ligands | 31 |
| solvent | 256 |
| Protein residues | 266 |
| RMS(bonds) | 0.007 |
| RMS(angles) | 1.01 |

Table III-5: X-ray crystallographic data and refinement statistics for KLVS bound to FR-1v.

| | KLVS bound to FR1 (UV exposure) |
|---------------------------------------|---|
| Wavelength | |
| Resolution range | 27.33 - 1.37 (1.419 - 1.37) |
| Space group | P 1 |
| Unit cell | 29.756 36.852 63.774 87.995 88.243 65.267 |
| Total reflections | |
| Unique reflections | 49131 (4848) |
| Multiplicity | |
| Completeness (%) | 94.99 (93.20) |
| Mean I/sigma(I) | |
| Wilson B-factor | 16.29 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 49094 (4841) |
| Reflections used for R-free | 2005 (191) |
| R-work | 0.2293 (0.3011) |
| R-free | 0.2539 (0.3248) |
| Number of non-hydrogen atoms | 2371 |
| macromolecules | 2176 |
| ligands | 27 |
| solvent | 168 |
| Protein residues | 266 |
| RMS(bonds) | 0.007 |
| RMS(angles) | 1.13 |

Table III-6: X-ray crystallographic data and refinement statistics for *Q108K:K40E:T51C:T53S:R58L:Q38F:Q4F* bound to *FR-1cyno*.

| | FR1cyno |
|---------------------------------------|----------------------------------|
| Wavelength | |
| Resolution range | 27.4 - 1.59 (1.647 - 1.59) |
| Space group | R 3 :H |
| Unit cell | 144.973 144.973 35.362 90 90 120 |
| Total reflections | |
| Unique reflections | 37132 (3731) |
| Multiplicity | |
| Completeness (%) | 99.59 (99.92) |
| Mean I/sigma(I) | |
| Wilson B-factor | 18.61 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 37128 (3729) |
| Reflections used for R-free | 1995 (203) |
| R-work | 0.1820 (0.2084) |
| R-free | 0.2231 (0.2542) |
| Number of non-hydrogen atoms | 2580 |
| macromolecules | 2185 |
| ligands | 41 |
| solvent | 354 |
| Protein residues | 267 |
| RMS(bonds) | 0.007 |
| RMS(angles) | 0.82 |

Table III-7: X-ray crystallographic data and refinement statistics for Q108K:K40L:T51V:T53S:R58H:Y19W:A33Y bound to Thiophenol.

| | Q108K:K40L:T51V:T53S:R58H:Y19W:A33Y bound to thiophenol |
|---------------------------------------|--|
| Wavelength | |
| Resolution range | 32.93 - 1.59 (1.647 - 1.59) |
| Space group | C 1 2 1 |
| Unit cell | 29.847 65.855 128.03 90 92.544 90 |
| Total reflections | |
| Unique reflections | 32054 (3199) |
| Multiplicity | |
| Completeness (%) | 95.75 (95.03) |
| Mean I/sigma(I) | |
| Wilson B-factor | 24.20 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 31934 (3173) |
| Reflections used for R-free | 1978 (190) |
| R-work | 0.2670 (0.3628) |
| R-free | 0.3009 (0.4082) |
| Number of non-hydrogen atoms | 2369 |
| macromolecules | 2186 |
| ligands | 40 |
| solvent | 143 |
| Protein residues | 266 |
| RMS(bonds) | 0.011 |
| RMS(angles) | 1.22 |

Table III-8: X-ray crystallographic data and refinement statistics for *Q108K:K40L:T51V:T53S:R58H:F16Y:Y19W:A33H:L117C bound to thiophenol*.

| | Q108K:K40L:T51V:T53S:R58H:F16Y:Y19W:A33H:L117C bound to thiophenol |
|---------------------------------------|---|
| Wavelength | |
| Resolution range | 29.57 - 1.531 (1.585 - 1.531) |
| Space group | P 1 |
| Unit cell | 29.899 36.014 64.011 90.439 92.585 114.121 |
| Total reflections | |
| Unique reflections | 34123 (2466) |
| Multiplicity | |
| Completeness (%) | 92.90 (66.50) |
| Mean I/sigma(I) | |
| Wilson B-factor | 24.86 |
| R-merge | |
| R-meas | |
| R-pim | |
| CC1/2 | |
| Reflections used in refinement | 34092 (2462) |
| Reflections used for R-free | 1997 (148) |
| R-work | 0.2137 (0.3224) |
| R-free | 0.2430 (0.3044) |
| Number of non-hydrogen atoms | 2389 |
| macromolecules | 2186 |
| ligands | 40 |
| solvent | 163 |
| Protein residues | 266 |
| RMS(bonds) | 0.008 |
| RMS(angles) | 1.42 |

REFERENCES

REFERENCES

1. Specht, E. A., Braselmann, E. & Palmer, A. E. A Critical and Comparative Review of Fluorescent Tools for Live-Cell Imaging. *Annu. Rev. Physiol.* **79**, 93–117 (2017).
2. Hilderbrand, S. A. & Weissleder, R. Near-infrared fluorescence: application to in vivo molecular imaging. *Curr Opin Chem Biol* **14**, 71–79 (2010).
3. Luo, S., Zhang, E., Su, Y., Cheng, T. & Shi, C. A review of NIR dyes in cancer targeting and imaging. *Biomaterials* **32**, 7127–7138 (2011).
4. Yuan, L., Lin, W., Zheng, K., He, L. & Huang, W. Far-red to near infrared analyte-responsive fluorescent probes based on organic fluorophore platforms for fluorescence imaging. *Chem. Soc. Rev.* **42**, 622–661 (2012).
5. Dean, K. M. & Palmer, A. E. Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat. Chem. Biol.* **10**, 512–523 (2014).
6. Primary structure of the *Aequorea victoria* green-fluorescent protein - ScienceDirect. <https://www.sciencedirect.com/science/article/pii/S096921999290691H>.
7. Green fluorescent protein as a marker for gene expression | Science. <https://science.sciencemag.org/content/263/5148/802>.
8. Zimmer, M. GFP: from jellyfish to the Nobel prize and beyond. *Chem. Soc. Rev.* **38**, 2823 (2009).
9. Olenych, S. G., Claxton, N. S., Ottenberg, G. K. & Davidson, M. W. The Fluorescent Protein Color Palette. *Current Protocols in Cell Biology* **36**, 21.5.1–21.5.34 (2007).
10. Day, R. N. & Davidson, M. W. The fluorescent protein palette: tools for cellular imaging. *Chem Soc Rev* **38**, 2887–2921 (2009).
11. Remington, S. J. Green fluorescent protein: A perspective. *Protein Sci* **20**, 1509–1519 (2011).
12. Stepanenko, O. V., Stepanenko, O. V., Kuznetsova, I. M., Verkhusha, V. V. & Turoverov, K. K. Beta-barrel scaffold of fluorescent proteins: folding, stability and role in chromophore formation. *Int Rev Cell Mol Biol* **302**, 221–278 (2013).
13. Nienhaus, K. & Nienhaus, G. U. Chromophore photophysics and dynamics in fluorescent proteins of the GFP family. *J Phys Condens Matter* **28**, 443001 (2016).

14. Crystal Structure of the *Aequorea victoria* Green Fluorescent Protein | Science. <https://science.sciencemag.org/content/273/5280/1392.long>.
15. The molecular structure of green fluorescent protein | Nature Biotechnology. <https://www.nature.com/articles/nbt1096-1246>.
16. Reaction Progress of Chromophore Biogenesis in Green Fluorescent Protein | Journal of the American Chemical Society. <https://pubs.acs.org/doi/10.1021/ja0580439>.
17. Sniegowski, J. A., Phail, M. E. & Wachter, R. M. Maturation efficiency, trypsin sensitivity, and optical properties of Arg96, Glu222, and Gly67 variants of green fluorescent protein. *Biochem. Biophys. Res. Commun.* **332**, 657–663 (2005).
18. Barondeau, D. P., Putnam, C. D., Kassmann, C. J., Tainer, J. A. & Getzoff, E. D. Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. *Proceedings of the National Academy of Sciences* **100**, 12111–12116 (2003).
19. Shagin, D. A. *et al.* GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. *Mol. Biol. Evol.* **21**, 841–850 (2004).
20. Labas, Y. A. *et al.* Diversity and evolution of the green fluorescent protein family. *Proceedings of the National Academy of Sciences* **99**, 4256–4261 (2002).
21. Zimmer, M. Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. *Chem. Rev.* **102**, 759–781 (2002).
22. Subach, F. V. & Verkhusha, V. V. Chromophore Transformations in Red Fluorescent Proteins. *Chem. Rev.* **112**, 4308–4327 (2012).
23. Sample, V., Newman, R. H. & Zhang, J. The structure and function of fluorescent proteins. *Chem. Soc. Rev.* **38**, 2852 (2009).
24. Piatkevich, K. D., Efremenko, E. N., Verkhusha, V. V. & Varfolomeev, S. D. Red fluorescent proteins and their properties. *RUSS. CHEM. REV.* **79**, 243–258 (2010).
25. Shcherbakova, D. M., Subach, O. M. & Verkhusha, V. V. Red Fluorescent Proteins: Advanced Imaging Applications and Future Design. *Angewandte Chemie International Edition* **51**, 10724–10738 (2012).
26. Wang, Y. & Wang, N. FRET and mechanobiology. *Integrative Biology* **1**, 565–573 (2009).
27. Zapata-Hommer, O. & Griesbeck, O. Efficiently folding and circularly permuted variants of the Sapphire mutant of GFP. *BMC Biotechnol.* **3**, 5 (2003).

28. Ghosh, S. *et al.* Blue protein with red fluorescence. *PNAS* **113**, 11513–11518 (2016).
29. Piatkevich, K. D. *et al.* Monomeric red fluorescent proteins with a large Stokes shift. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5369–5374 (2010).
30. Patterson, G. H. & Lippincott-Schwartz, J. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science* **297**, 1873–1877 (2002).
31. Grotjohann, T. *et al.* Diffraction-unlimited all-optical imaging and writing with a photochromic GFP. *Nature* **478**, 204–208 (2011).
32. Gurskaya, N. G. *et al.* Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. *Nat. Biotechnol.* **24**, 461–465 (2006).
33. mMaple: A Photoconvertible Fluorescent Protein for Use in Multiple Imaging Modalities. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0051314>.
34. Tiwari, D. K. *et al.* A fast- and positively photoswitchable fluorescent protein for ultralow-laser-power RESOLFT nanoscopy. *Nat. Methods* **12**, 515–518 (2015).
35. Ando, R., Mizuno, H. & Miyawaki, A. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science* **306**, 1370–1373 (2004).
36. Andresen, M. *et al.* Photoswitchable fluorescent proteins enable monochromatic multilabel imaging and dual color fluorescence nanoscopy. *Nat. Biotechnol.* **26**, 1035–1040 (2008).
37. Chudakov, D. M., Matz, M. V., Lukyanov, S. & Lukyanov, K. A. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiological Reviews* **90**, 1103–1163 (2010).
38. Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nat Methods* **2**, 905–909 (2005).
39. Megason, S. G. & Fraser, S. E. Imaging in systems biology. *Cell* **130**, 784–795 (2007).
40. Balla, T. Green Light to Illuminate Signal Transduction Events. *Trends Cell Biol* **19**, 575–586 (2009).
41. In Vivo and in Situ Tracking Cancer Chemotherapy by Highly Photostable NIR Fluorescent Theranostic Prodrug | Journal of the American Chemical Society. <https://pubs.acs.org/doi/10.1021/ja412380j>.
42. Morisaki, T. & McNally, J. G. Photoswitching-Free FRAP Analysis with a Genetically Encoded Fluorescent Tag. *PLoS One* **9**, (2014).

43. Molecules | Free Full-Text | Advanced Fluorescence Microscopy Techniques—FRAP, FLIP, FLAP, FRET and FLIM. <https://www.mdpi.com/1420-3049/17/4/4047>.
44. Monomeric Fluorescent Timers That Change Color From Blue to Red Report on Cellular Trafficking - PubMed. <https://pubmed.ncbi.nlm.nih.gov/19136976/>.
45. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik* **437**, 55–75 (1948).
46. Piston, D. W. & Kremers, G.-J. Fluorescent protein FRET: the good, the bad and the ugly. *Trends Biochem. Sci.* **32**, 407–414 (2007).
47. Whitaker, M. Genetically-encoded probes for measurement of intracellular calcium. *Methods Cell Biol* **99**, 153–182 (2010).
48. Markova, O. FLUORESCENT BIOSENSORS FOR LIVE CELLS ANALYSIS. (2008). doi:10.13140/RG.2.1.2841.5124.
49. Moeyaert, B., Vandenberg, W. & Dedecker, P. SOFIevaluator: a strategy for the quantitative quality assessment of SOFI data. *Biomed Opt Express* **11**, 636–648 (2020).
50. Chi, C. *et al.* Intraoperative Imaging-Guided Cancer Surgery: From Current Fluorescence Molecular Imaging Methods to Future Multi-Modality Imaging Technology. *Theranostics* **4**, 1072–1084 (2014).
51. Ntziachristos, V. Going deeper than microscopy: the optical imaging frontier in biology. *Nat. Methods* **7**, 603–614 (2010).
52. Fernández-Suárez, M. & Ting, A. Y. Fluorescent probes for super-resolution imaging in living cells. *Nat. Rev. Mol. Cell Biol.* **9**, 929–943 (2008).
53. A Review of Indocyanine Green Fluorescent Imaging in Surgery. <https://www.hindawi.com/journals/ijbi/2012/940585/>.
54. Bosma, S. E., van Driel, P. B., Hogendoorn, P. C., Dijkstra, P. S. & Sier, C. F. Introducing fluorescence guided surgery into orthopedic oncology: A systematic review of candidate protein targets for Ewing sarcoma. *J Surg Oncol* **118**, 906–914 (2018).
55. Hoffman-Kim, D., Diefenbach, T. J., Eustace, B. K. & Jay, D. G. Chromophore-assisted laser inactivation. *Methods Cell Biol.* **82**, 335–354 (2007).
56. Bharathiraja, S. *et al.* Photo-based PDT/PTT dual model killing and imaging of cancer cells using phycocyanin-polypyrrole nanoparticles. *European Journal of Pharmaceutics and Biopharmaceutics* **123**, 20–30 (2018).

57. Reisch, A. & Klymchenko, A. S. Fluorescent Polymer Nanoparticles Based on Dyes: Seeking Brighter Tools for Bioimaging. *Small* **12**, 1968–1992 (2016).
58. Wolfbeis, O. S. An overview of nanoparticles commonly used in fluorescent bioimaging. *Chem. Soc. Rev.* **44**, 4743–4768 (2015).
59. Advances in fluorescent protein technology | Journal of Cell Science. <https://jcs.biologists.org/content/120/24/4247>.
60. Kumagai, A. *et al.* A bilirubin-inducible fluorescent protein from eel muscle. *Cell* **153**, 1602–1611 (2013).
61. Yang, X., Kuk, J. & Moffat, K. Crystal structure of *Pseudomonas aeruginosa* bacteriophytochrome: Photoconversion and signal transduction. *PNAS* **105**, 14715–14720 (2008).
62. Shcherbakova, D. M., Baloban, M. & Verkhusha, V. V. Near-infrared fluorescent proteins engineered from bacterial phytochromes. *Curr Opin Chem Biol* **27**, 52–63 (2015).
63. Shcherbakova, D. M. & Verkhusha, V. V. Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat. Methods* **10**, 751–754 (2013).
64. Gautier, A. *et al.* An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chemistry & Biology* **15**, 128–136 (2008).
65. A general method for the covalent labeling of fusion proteins with small molecules in vivo | Nature Biotechnology. <https://www.nature.com/articles/nbt765?draft=marketing>.
66. Keppler, A., Pick, H., Arrivoli, C., Vogel, H. & Johnsson, K. Labeling of fusion proteins with synthetic fluorophores in live cells. *PNAS* **101**, 9955–9959 (2004).
67. Michnick, S. W., Ear, P. H., Manderson, E. N., Remy, I. & Stefan, E. Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nature Reviews Drug Discovery* **6**, 569–582 (2007).
68. Cole, N. B. Site-Specific Protein Labeling with SNAP-Tags. *Curr Protoc Protein Sci* **73**, 30.1.1-30.1.16 (2013).
69. Juillerat, A. *et al.* Directed evolution of O⁶-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. *Chem. Biol.* **10**, 313–317 (2003).
70. Sheng, W. Development of Novel Far-Red/Near-Infrared Dye-hCRBP^{II} Based Imaging Tags for Background-Free Live Cell Imaging. (Michigan State University, 2019).

71. Gallagher, S. S., Sable, J. E., Sheetz, M. P. & Cornish, V. W. An in vivo covalent TMP-tag based on proximity-induced reactivity. *ACS Chem. Biol.* **4**, 547–556 (2009).
72. Wombacher, R. *et al.* Live-cell super-resolution imaging with trimethoprim conjugates. *Nat. Methods* **7**, 717–719 (2010).
73. Hori, Y. & Kikuchi, K. Protein labeling with fluorogenic probes for no-wash live-cell imaging of proteins. *Curr Opin Chem Biol* **17**, 644–650 (2013).
74. Szent-Gyorgyi, C. *et al.* Fluorogen-activating single-chain antibodies for imaging cell surface proteins. *Nat. Biotechnol.* **26**, 235–240 (2008).
75. Telmer, C. A. *et al.* Rapid, specific, no-wash, far-red fluorogen activation in subcellular compartments by targeted fluorogen activating proteins. *ACS Chem. Biol.* **10**, 1239–1246 (2015).
76. Kyndt, J. A., Meyer, T. E., Cusanovich, M. A. & Van Beeumen, J. J. Characterization of a bacterial tyrosine ammonia lyase, a biosynthetic enzyme for the photoactive yellow protein. *FEBS Lett.* **512**, 240–244 (2002).
77. van der Horst, M. A., Arents, J. C., Kort, R. & Hellingwerf, K. J. Binding, tuning and mechanical function of the 4-hydroxy-cinnamic acid chromophore in photoactive yellow protein. *Photochem. Photobiol. Sci.* **6**, 571–579 (2007).
78. Development of Protein-Labeling Probes with a Redesigned Fluorogenic Switch Based on Intramolecular Association for No-Wash Live-Cell Imaging - Hori - 2012 - *Angewandte Chemie International Edition* - Wiley Online Library.
<https://onlinelibrary.wiley.com/doi/full/10.1002/anie.201200867>.
79. Borgstahl, G. E., Williams, D. R. & Getzoff, E. D. 1.4 Å structure of photoactive yellow protein, a cytosolic photoreceptor: unusual fold, active site, and chromophore. *Biochemistry* **34**, 6278–6287 (1995).
80. Kumar, N., Hori, Y., Nishiura, M. & Kikuchi, K. Rapid no-wash labeling of PYP-tag proteins with reactive fluorogenic ligands affords stable fluorescent protein conjugates for long-term cell imaging studies. *Chem. Sci.* **11**, 3694–3701 (2020).
81. Berbasova, T. *et al.* Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone. *J. Am. Chem. Soc.* **135**, 16111–16119 (2013).
82. Wang, W. *et al.* Tuning the electronic absorption of protein-embedded all-trans-retinal. *Science* **338**, 1340–1343 (2012).
83. Santos, E. M. Development of fluorescent protein tags for live-cell imaging. (Michigan State University, 2017).

84. Yapici, I. *et al.* “Turn-On” Protein Fluorescence: In Situ Formation of Cyanine Dyes. *J. Am. Chem. Soc.* **137**, 1073–1080 (2015).
85. Synthesis and Modification of Terrylenediimides as High-Performance Fluorescent Dyes - Nolde - 2005 - Chemistry – A European Journal - Wiley Online Library. <https://chemistry-europe.onlinelibrary.wiley.com/doi/full/10.1002/chem.200401177>.
86. Guo, Z., Park, S., Yoon, J. & Shin, I. Recent progress in the development of near-infrared fluorescent probes for bioimaging applications. *Chemical Society Reviews* **43**, 16–29 (2014).
87. Escobedo, J. O., Rusin, O., Lim, S. & Strongin, R. M. NIR Dyes for Bioimaging Applications. *Curr Opin Chem Biol* **14**, 64 (2010).
88. Weinberg, D. R. *et al.* Proton-coupled electron transfer. *Chem. Rev.* **112**, 4016–4093 (2012).
89. Stoner-Ma, D. *et al.* Observation of Excited-State Proton Transfer in Green Fluorescent Protein using Ultrafast Vibrational Spectroscopy. *J. Am. Chem. Soc.* **127**, 2864–2865 (2005).
90. Zhou, P. & Han, K. Unraveling the Detailed Mechanism of Excited-State Proton Transfer. *Acc. Chem. Res.* **51**, 1681–1690 (2018).
91. Ghanbarpour, A. Functional Control of Soluble Rhodopsin Mimics Using High Resolution Structure-Based Design and Evaluation. (Michigan State University, 2019).
92. Nosrati, M., Berbasova, T., Vasileiou, C., Borhan, B. & Geiger, J. H. A Photoisomerizing Rhodopsin Mimic Observed at Atomic Resolution. *J. Am. Chem. Soc.* **138**, 8802–8808 (2016).
93. Sheng, W. *et al.* A Near-Infrared Photoswitchable Protein–Fluorophore Tag for No-Wash Live Cell Imaging. *Angewandte Chemie International Edition* **57**, 16083–16087 (2018).

CHAPTER IV: DOMAIN SWAPPING IN HCRBP II (IN VIVO STUDIES)

IV-1 DOMAIN SWAPPING

During our studies on hCRBP II, my lab-mates characterized the domain swapped dimer as a folding product for this protein.^{1 2 3 4 5} In domain swapping two or more monomers exchange an identical part of their structures to form a dimer or higher-order oligomers (**Figure IV-1.**)⁶ The first evidence of domain swapping was discovered by Eisenberg and his group for the dimeric structure of diphtheria toxin. So far, domain swapping has been reported for more than 40 different cases. One of the first examples is Bovine pancreatic ribonuclease (RNase A), in which the N terminal fragments exchanged to form a domain swapped dimer (DSD).^{6 7 8} Domain swapping can be dynamic, where the barrier between monomer and DSD or oligomer is low, and it can be static, where the barrier is very high. The exchanged region can be very small, or it can be as large as half of a protein domain, even around 100 residues. In domain swapping, the exchanged loop between the swapped region in dimer is called the hinge region. The swapped region in most of the cases is either the N terminus or C terminus of the protein, but it can also be in the middle of the protein sequence like in blood coagulant factors IX/X-binding protein. In some DSD cases more than one region in a protein can swap.⁹ Domain swapping can exist in some proteins naturally as a folding product. In some DSD cases, domain swapping can be the consequence of unusual conditions like pH, temperature, and presence of denaturant. Domain swapping is responsible for the evolution of a number of large proteins. Also, it may lead to aggregation that might be reason for many diseases, such as Parkinson's disease, Alzheimer's disease, diabetes, and many others.¹⁰

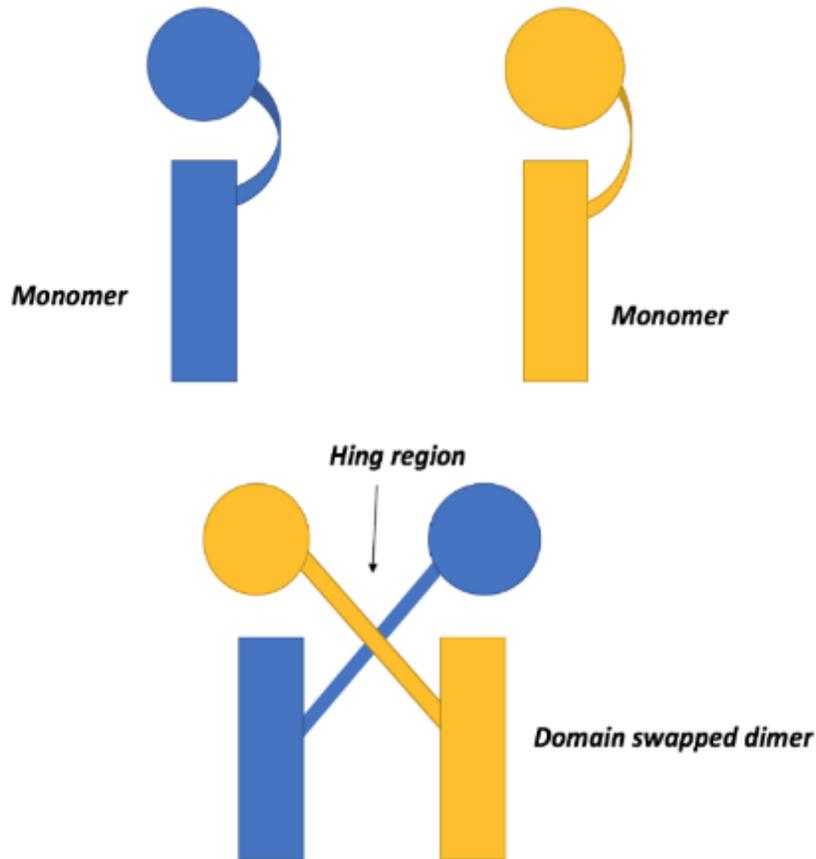


Figure IV-1: Demonstrates the concept of domain swapping.

IV-2 DOMAIN SWAPPING IN HCRBPII

As I explained before, Intracellular lipid binding proteins (iLBPs) are small cytosolic proteins responsible for transport of various insoluble hydrophobic molecules. All of the family members have similar structures including a ten stranded β barrel and two alpha helices located at the mouth of the internal binding cavity playing as the entrance to the binding pocket of the protein.¹⁵ Structures of almost all of the members of the iLBP family have been reported, which including fatty acid binding proteins and retinoic acid binding proteins.^{16 17} Most of the studies on the folding pathway of iLBPs have been carried out in the Gierasch lab. Their results which are mostly on human Cellular Retinoic acid binding protein I (hCRBPI), reports the early β barrel formation in

the folding pathway of this protein.^{18 19} The formation of the β -barrel as a meta stable intermediate during the folding process of iLBPs prevent this family from having aggregation and amyloid formation. During the efforts for using hCRBP2 as a rhodopsin mimic, my former lab mates reported the first DSD structure of this protein (**Figure IV-2**).¹

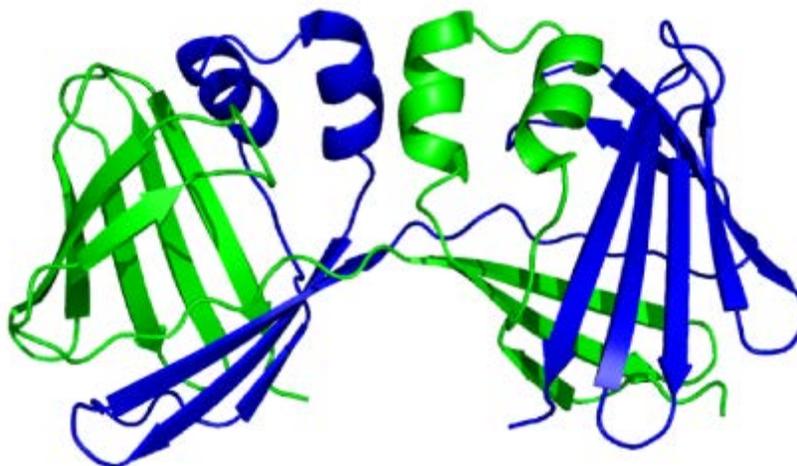


Figure IV-2: Structure of domain swapped dimer for hCRBP2.

The swapped region is pretty huge and it is around half of the protein. The exchanged region contains three β strands with two alpha helices. The existence of domain swapping for hCRBP2 can be seen as a new folding product for the iLBP family, and it may change the previous concepts of folding for this family, like early formation of β barrel. The existence of a domain swapped dimer for hCRBP2, has led our group to suggest a new folding pathway for this protein. Based on our structural work, we suggested at least one stable intermediate in the folding pathway of this protein. In our proposed mechanism, hCRBP2 folds via an “open monomer” and both monomer and domain swapped dimer are derived from this intermediate. Our group predicted that the N-terminal and C-terminal halves of hCRBP2 are capable of at least partially folding independently,

and make the “open monomer” as an intermediate. The monomer/dimer ratio as folding products would then depend on the relative rates of dimerization of the open monomers, versus rotation of the N and C termini together to form the “closed monomer” (Figure IV-3).¹

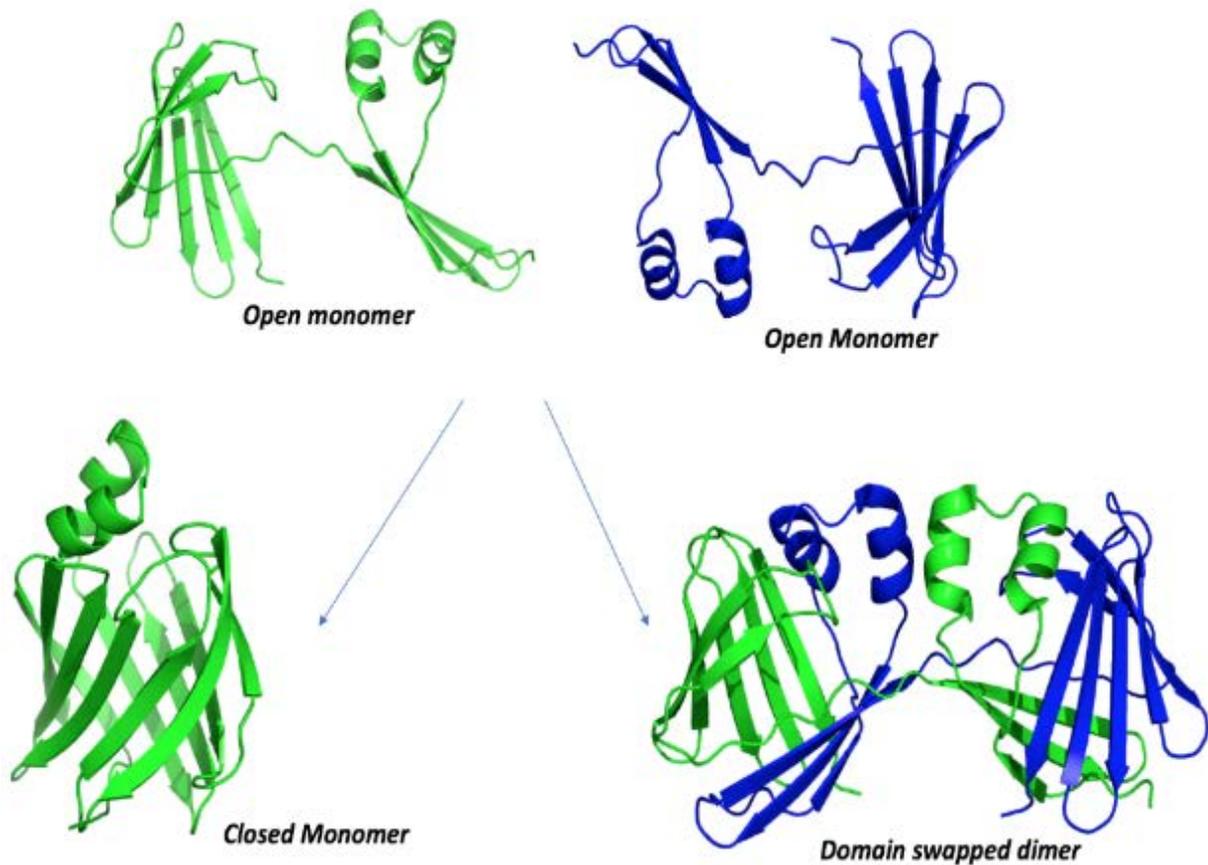


Figure IV-3: Proposed folding pathway for hCRBP II. Open monomer is as an intermediate in the folding (PDB:4ZH6). Domain swapped dimer and monomer shows as folding products.

Mutational studies on hCRBP II demonstrates the importance of some residues in formation of DSD. Studies show that with a single mutation, we can change the ratio of monomer and domain swapped dimer which are folding products in the folding pathway of this protein.^{2 1}

IV-3 INVESTIGATE THE FOLDING INTERMEDIATES OF HCRBP2 (USING CD AND FLUORESCENCE SPECTROSCOPY)

We decided to conduct the folding experiments to investigate our hypothesis about the open monomer as an intermediate. These experiments have been done by Dr. Zahra Assar and I, in collaboration with Prof. Lapidus group.⁵ As mentioned above, we suggested that the N terminus and C terminus fold independently, and they form the open monomer as an intermediate. One of the most important studies for detecting the intermediates is to detect the changes in the secondary and tertiary structure of the protein at different denaturant concentrations. In our experiments, we used different concentrations of Guanidium chloride as a denaturant to denature WT-hCRBP2. We monitored the changes in protein using Circular dichroism (CD) and fluorescence spectroscopy in each step (**Figure IV-4**). From the CD data, the ratio of α helix/beta-strand calculated in each step (220nm for β -strand and 222nm/208nm for α helices). In 1 M concentration of Gu-HCl, around 30% of β strands decreased compared to the native structure. However, α helices seemed to be very more stable compared to β strands, and only they started to decrease their intensities at 2M denaturant. At 4M denaturant, around 20% of α helices has been lost. The concentration dependence might indicate the early association of dimer during the folding. Prof. Lisa Lapidus and her group tried to find the order of the folding using the ellipticity versus denaturant (from CD experiment). As we expected, the data did not fit the two-state model ($N \rightleftharpoons U$), indicating there are one or more intermediates in the folding pathway of this protein. More calculations need to be done to fit our data into the three-state or higher-order state model.²⁰ Also, other experiments, especially H-D mass spectrometry could be used to identify the folding intermediate of the hCRBP2.

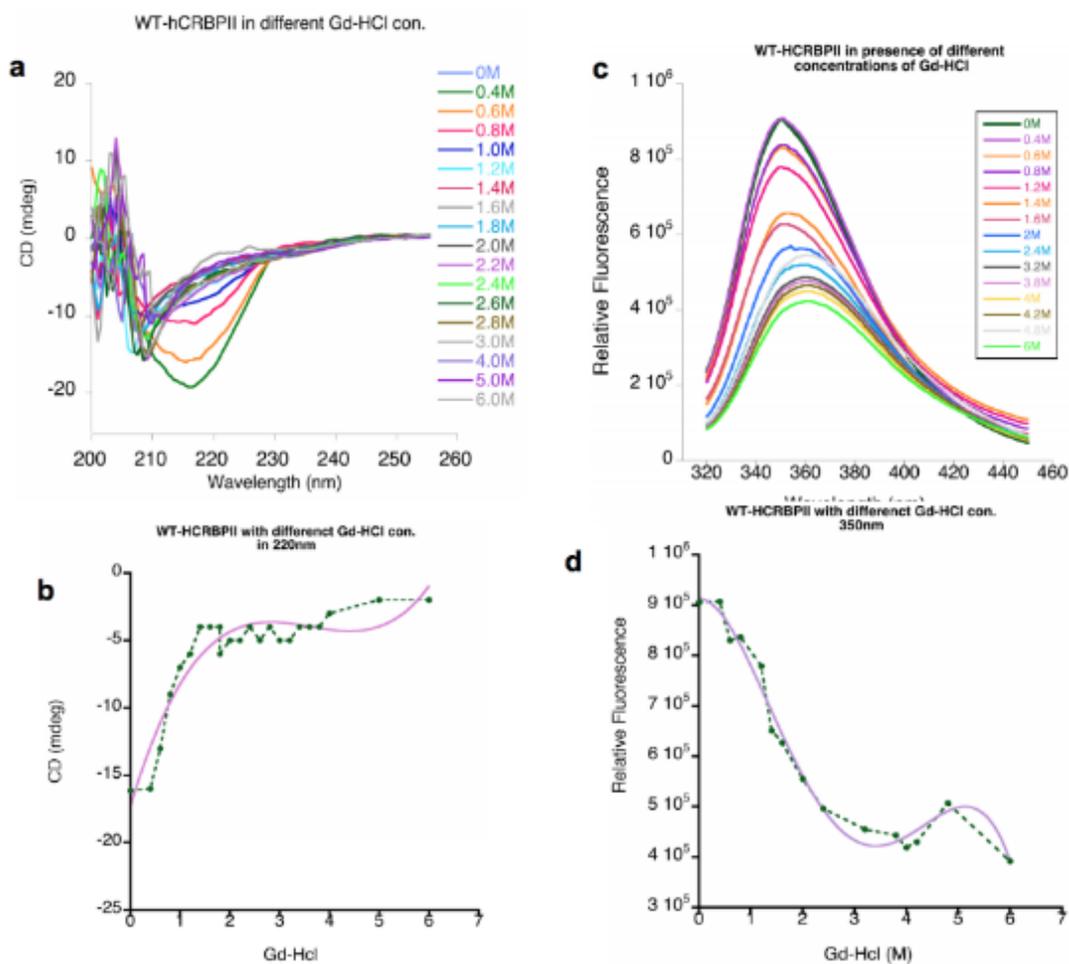


Figure IV-4: Unfolding spectra of the WT hCRBP11 in presence of different concentration of Gd-HCl. a and b) monitoring by CD measurements. c and d) Tryptophan fluorescent spectroscopy to study the intermediates in folding pathway of hCRBP11. Purple line in b and d demonstrates the third degree polynomial fit.

IV-4 IN VITRO REFOLDING EXPERIMENT TO INVESTIGATE THE EFFECT OF LIGAND BINDING ON THE FOLDING PATHWAY OF HCRBP II

Many studies have been done to find the different factors that affect the formation of domain swapped dimer in hCRBP II. In our previous studies, my lab mates elucidated the effect of protein concentration on dimerization of this protein. In vitro refolding experiments for some mutations of hCRBP II indicate that concentrations of protein, affects the dimerization. By increasing the concentration of protein during the expression, the ratio of dimer to monomer will be increased.¹ One of the interesting factors that had not been studied on dimerization before I joined the lab, was the effect of ligand binding on dimerization. I performed the in vitro refolding experiment to investigate the effect of this factor. For this experiment, I chose the Q108K:T51D construct of hCRBP II, since the structure of both the Apo and Holo dimer of this mutation have been reported by our group. The ligand which is retinal, can make a Schiff base with Q108K as discussed in the last chapters. In the trials with 6 mg/ml of protein, the chromatogram of size exclusion chromatography showed more domain swapped dimer form for Holo form of the protein. (**Figure IV-5**). These results were reproducible, but since the concentration of soluble protein decreased significantly during the refolding experiment, it needs more investigation.

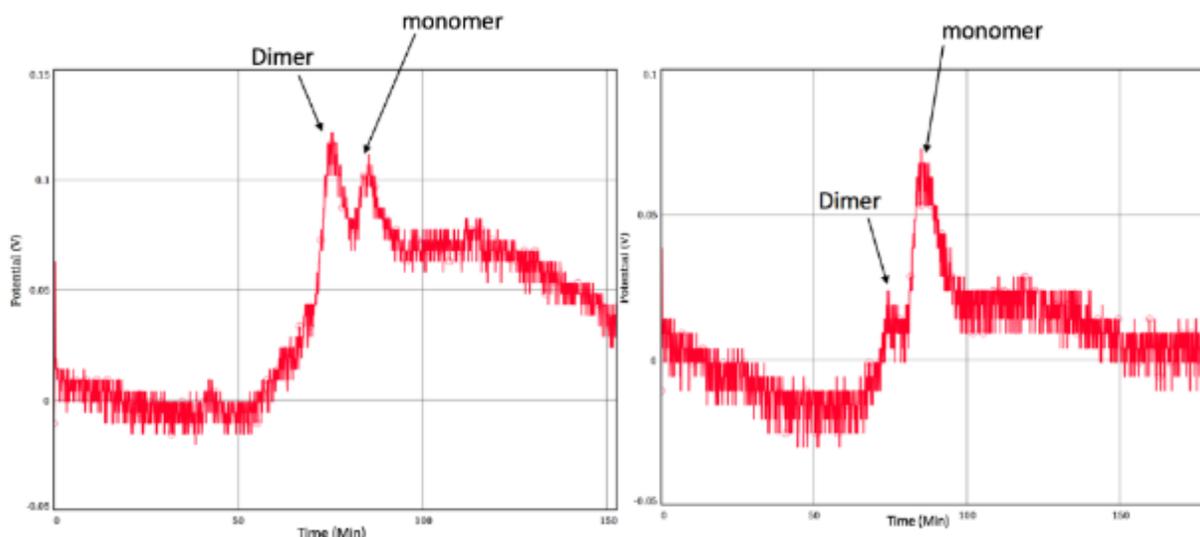


Figure IV-5: Chromatogram of size exclusion chromatography after conducting the refolding experiment in high concentration. The left figure is the chromatogram for refolding in the presence of ligand and the right one is the result of refolding in the absence of ligand.

IV-5 STUDY THE OCCURRENCE OF DOMAIN SWAPPING BY MUTATIONAL ANALYSIS OF HCRBP II (PHASE RELATIONSHIP)

Our group is successful in structure determination of many DSD hCRBP II variants. In all of these structures the loop between β strand 3 and 4 in the monomer becomes straight in the DSD, resulting in a single β strand stretching the length of the dimer. In the monomer structure, all of the odd residues in β strand 3 and even residues of β strand 4 are inside the binding pocket. However, in the dimer structure, since the β strand 3 and β strand 4 makes a single β strand, the conformation of even residues of β strand 4 should be toward the solvent this time (**figure IV-6**). This means that for the formation of an ideal single β strand in the domain swapped dimer, the residues in the C terminal strand should “re-phase” and to put the even side chains in phase with

the odd numbered side chains of strand (red label in the figure V-6). In other words, all the hydrophilic residues that have interaction with solvent should turn toward the binding pocket, and the hydrophobic residues should turn toward the solvent. Therefore, this process is not energy favorable and it may change the whole structure of the protein. Interestingly, in DSD structures, the β strand is not formed in this way and it is not ideal strand. Instead of re-phasing all the residues in C terminal, two residues have the same conformation toward the binding pocket. The resulting strand is not the ideal β strand since the side chains of two adjacent residues are on the same face of the strand, but is more energy favorable (**Figure IV-6**).⁵

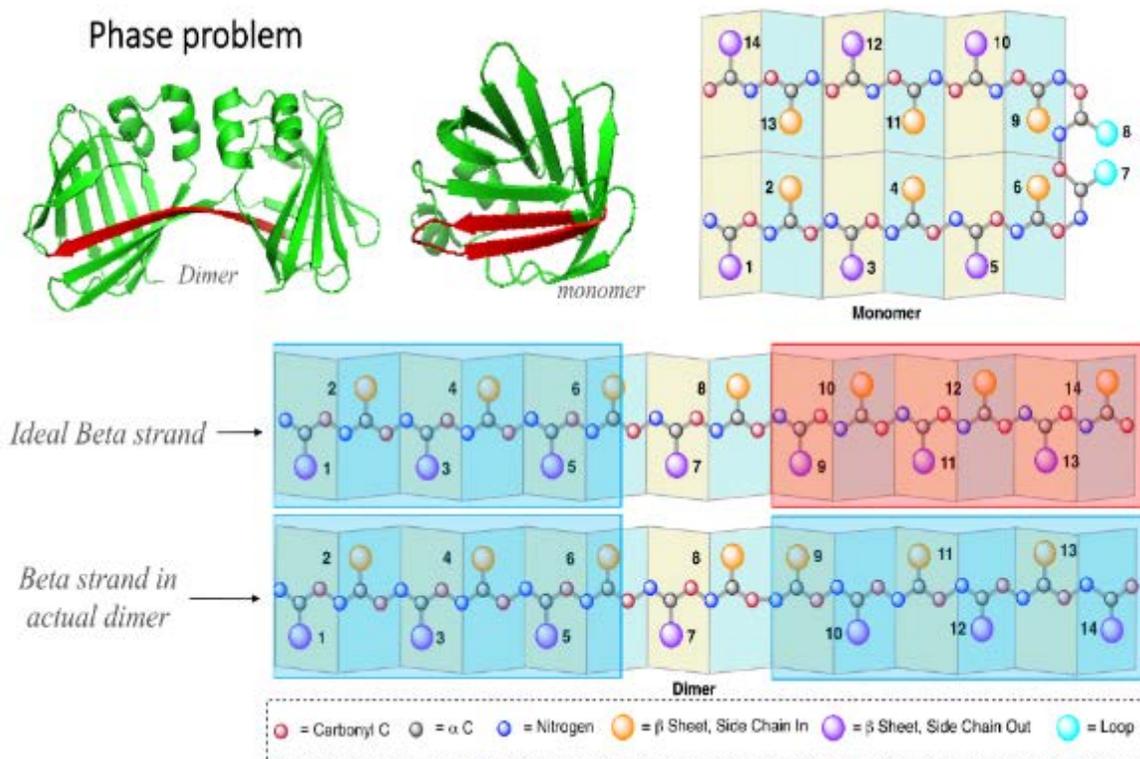


Figure IV-6: The "phase problem" in domain swap dimerization. Top figure is for monomer *hCRBPII*. In the bottom of figure, the ideal and actual β strand in DSD form elucidated.

As explained, quite a few studies have been done on domain swapping to investigate essential factors that can give rise to domain swapping. One of the most important ones was mutational

analysis. Mutational analysis on hCRBP_{II} indicates that the presence of some hydrophobic residues in the vicinity of the binding pocket can increase the domain swapping for this protein. One of these mutants is Y60L and Y60W, which increases the dimer/monomer ratio significantly. The way that these mutants, solve the phase problem might be key to the domain swapping mechanism. Another important structural detail that we thought it might be important for domain swapping is the conformation of residue Y60 and D61. In all of the domain-swapped dimers of hCRBP_{II} mutants, these two residues are pointed toward the solvent, which re-phase the strand, which is like residue 9 and 8 in schematic figure V-6. In wild type hCRBP_{II}, which express as mostly monomer, residue Y60 is inside the binding pocket. This can be another significant factor in the mechanism for dimerization in hCRBP_{II}. In the Y60W mutant of DSD hCRBP_{II}, residue 60 is perpendicular to the binding pocket, and this how it forms the β -strand in the DSD (**Figure IV-7**).

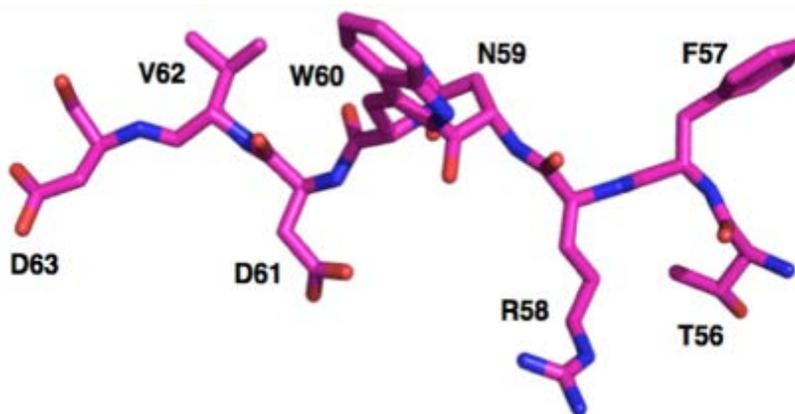


Figure IV-7: Residues 56-63 in the B subunit of Y60W-HcrbpII.

To predict the occurrence of domain swapping by amino acid sequence, Dr. Zahra Assar-Nossoni and Alireza Ghanbarpour have carried out many mutations on hCRBP_{II} (for example Y60L, Q108K:K40L: T51F, Q108K: T51D, E72A, and Q108K:K40L: T51W) and their relative orientations of amino acids have been studied by them (**Figure IV-8**).

| | | Residue Position | | | | | | | | | | | | | | | | | |
|-------------------|---------|------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| CRBP II Mutant | | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 |
| WT sequence | | N | F | K | T | K | T | T | S | Y | F | R | N | Y | D | V | D | F | T |
| Monomer | WT | O | I | O | I | O | I | O | I | L | L | I | O | I | O | I | O | I | O |
| Asymmetric Dimers | Y60WA | O | I | O | I | O | I | O | I | O | I | I | O | I | O | I | O | I | O |
| | Y60WB | O | I | O | I | O | I | O | I | O | I | O | I | S | O | I | O | I | O |
| Symmetric Dimers | WT | O | I | O | I | O | I | O | I | O | I | O | I | I | O | I | O | I | O |
| | Y60L | O | I | O | I | O | I | O | I | O | I | O | I | S | O | I | O | I | O |
| | KL:T51F | O | I | O | I | O | I | O | I | O | I | O | I | S | O | I | O | I | O |
| | K:T51D | O | I | O | I | O | I | O | I | O | I | O | I | S | O | I | O | I | O |
| | KL:T51W | O | I | O | I | O | I | O | I | O | I | O | I | S | O | I | O | I | O |
| | E72A | O | I | O | I | O | I | O | I | O | I | O | I | I | O | I | O | I | O |

Figure IV-8: Demonstrates the conformation of residues in different hCRBP II construct. *O*: out of the binding pocket, *I*: inside the binding pocket, *S*: toward the solvent.

According to these data, these data, the significant hinge motion at Thr56, subsequent proper orientation of the N and C terminus for dimer formation, and reforming the connecting strand are required for domain swapping in hCRBP II. Other critical residues for domain swapping seem to be in position 58 and 59. The relative orientation of these two residues in the domain-swapped dimer is different from the monomer. To test the effect of Asparagine 59 on domain swapping, we made a single mutant N59L: hCRBP II construct. In the domain-swapped dimer, Asparagine 59 is toward the inside of the barrel. We hypothesized that by mutating this position to a more hydrophobic residue like Leucine, we could increase the dimerization. However, the expression of this protein at room temperature was low, and it was mostly monomer. The SDS PAGE shows a small band for our protein in 15kD (**Figure IV-9**).

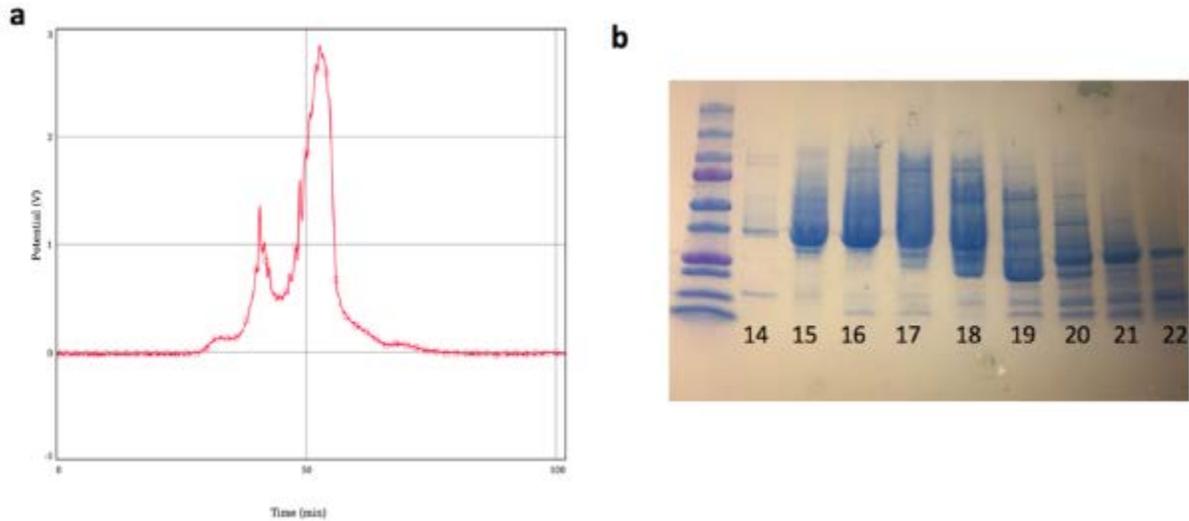


Figure IV-9: a) SEC chromatogram for hCRBP II N59L. b) SDS PAGE from fraction 14-22 of SEC.

Another critical residue was Arginine 58. We made two different mutants to both hydrophobic and hydrophilic residues, R58L and R58Q, to elucidate the effect of this residue in domain swapping. However, in both cases, only monomer hCRBP II formed (**Figure IV-10**).

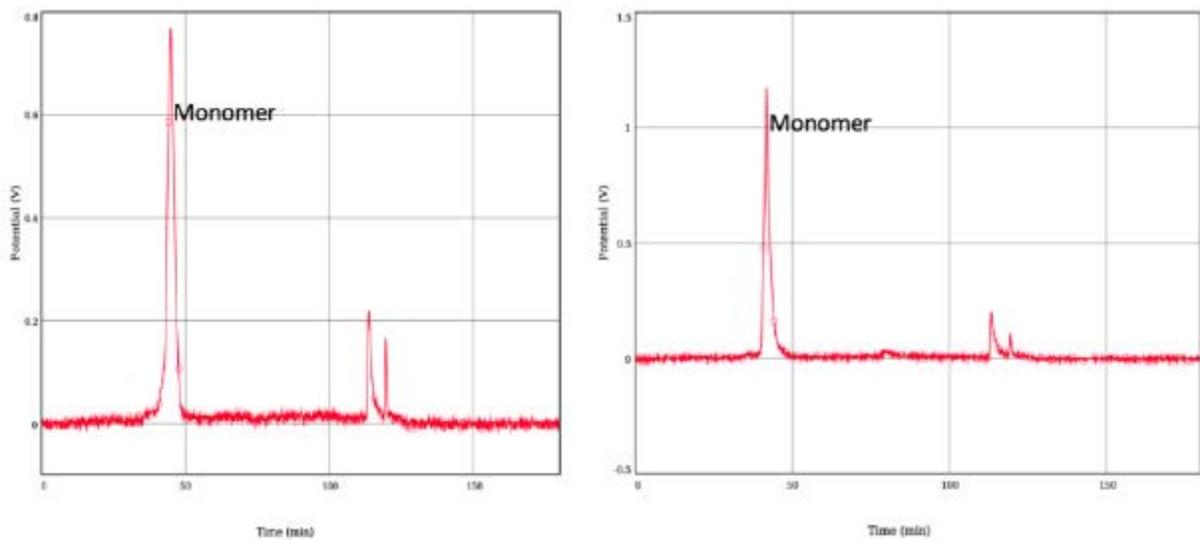


Figure IV-10: Chromatogram of Source Q Size Exclusion chromatography of a) R58L hCRBP II b) R58Q hCRBP II.

IV-6 EXPRESSION OF DSD HCRBP2 IN MAMMALIAN CELLS

IV-6-1 Expression of hCRBP2 in HeLa cells and detection using western blotting

Highly purified and low-cost expression of the majority of recombinant proteins is done by using the bacterial expression systems. However, in most of the cases, expression in mammalian cells is still needed to evaluate the function and correct post-translational modifications for the proteins.²¹ ²²Since most post translational modifications occur in the endoplasmic reticulum, mammalian expression allows for the highest level of post-translational processing and functional activity of the protein. This system is commonly used for the production of antibodies and therapeutic proteins to evaluate potential drug targets.^{23 24} Almost all of the studies on DSD hCRBP2 have been done through bacterial expression. To find out the physiological relevancy of this phenomenon, I tried to investigate the existence of the DSD form in mammalian expression. Since we had higher information on domain swapping of hCRBP2, we used this protein as our first target for this study. Different eukaryotic expression systems are available for the expression of mammalian proteins that contain unique post-translational modifications. The most common methods currently include baculovirus expression vector systems and mammalian cell systems. HeLa cell line is one of the most common mammalian cells used for mammalian culture. HeLa cells are cervical cancer cells, and the name came from a first sample taken from a woman called Henrietta Lacks. The HeLa genome was sequenced in 2013. These cells are the most commonly used human cell lines, which used for scientific research. They are one of the oldest human cell line and were used to test the first polio vaccine in the 1950s, test the infection of human cells by parvovirus in 1953, and also for testing the heptamethine dye IR-808, etc. ^{25 26}

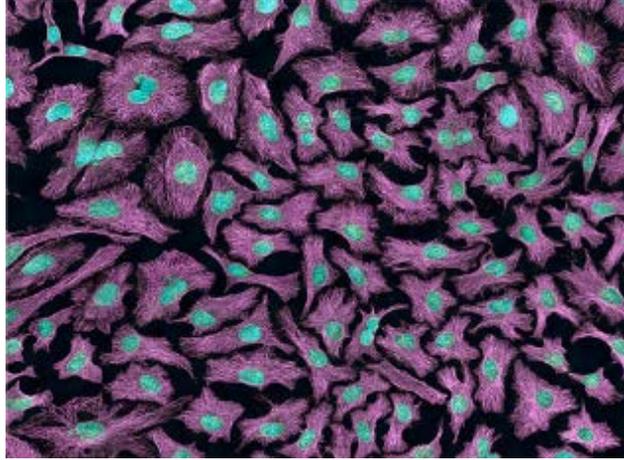


Figure IV-11: Multiphoton fluorescence image of HeLa cells with cytoskeletal microtubules (magenta) and DNA (cyan).

One of the reasons for using HeLa cells in research studies is their fast proliferation. Moreover, they have an altered version of telomerase, which can prevent aging and cell death. However, the rapid adaptation of HeLa cells to grow in tissue culture plates cause them to be challenging to control. Through improper maintenance, HeLa cells can cause contamination and interfere with other cell cultures in the same laboratory (**figure IV-11**).²⁷ After protein expression in mammalian cell culture, different analytical tools can be used to detect the target proteins. One of the analytical technique for detecting specific proteins from a complex mixture of proteins extracted from cells is western blotting.^{28 29} This method was first introduced by Towbin et al. in 1979. In western blotting, the mix of proteins from cell extract is separated by molecular weight, through SDS gel electrophoresis (SDS PAGE gel). Then, proteins are transferred from the SDS PAGE gel to a nitrocellulose or polyvinylidene difluoride (PVDF) membrane. The term "blotting" refers to this step, which means transferring protein samples from the SDS PAGE gel to a membrane. Further, the proteins can be detected by using the proper antibodies. Antibodies can be specific for each target protein, or in most cases, the antibodies can label specific protein tags (**Figure IV-12**).

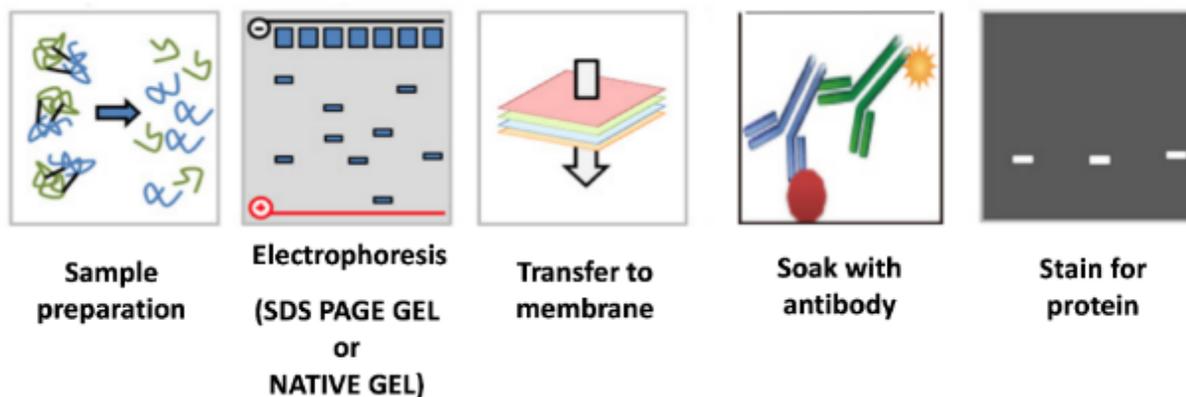


Figure IV-12: Schematic illustration of steps in western blotting.

As we discussed, the existence of domain swapping in hCRBP_{II} have huge effect on the folding pathway for iLBPs, and may lead to allosteric regulation for this protein. To check the physiological relevance of domain swapping in iLBPs, we decided to investigate the existence of domain swapping in hCRBP_{II} using the HeLa cell line for mammalian expression and Western blotting for detection. I cloned the Q108K: T51D hCRBP_{II} into Cytomegalovirus (CMV) vector with a flag tag on the N terminus of our target gene. Flag tag is a peptide tag that can help us to identify our target protein from other proteins from the extract by using a proper antibody. Q108K: T51D mutant of hCRBP_{II}, was expressed as all DSD in bacterial expression. The aforementioned plasmid was used to for transfection of 6-well culture plates containing HeLa cells. The first three wells were as controls; the first well was blank without changing the media, the second was blank with changing the media through transfection and the third one was transfected with 350 ng of empty cmv-flag tag vector. The rest of the wells were transfected with different amounts of the target construct; 350 ng, 750 ng, 1050 ng of the target plasmid, respectively (**Figure IV-13**).

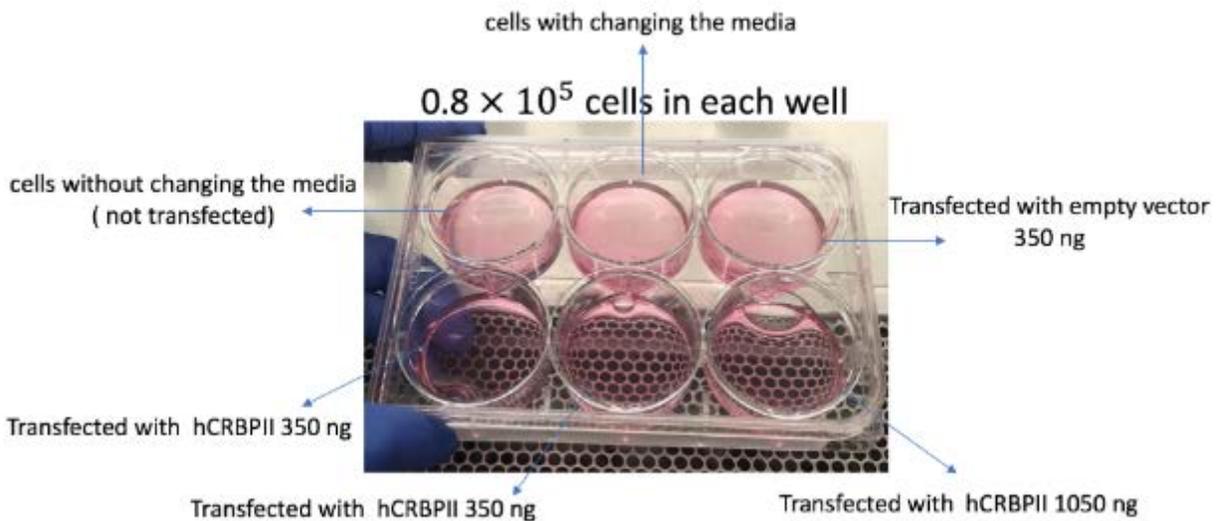


Figure IV-13: 6-well culture plate used for expression of hCRBP II. There is the same amount of *HeLa* cells in each well.

To identify the hCRBP II from a mixture of proteins in the extracts, western blotting was done. The extract of each well was collected separately and 20 ng of each extract was run through SDS PAGE and transferred to the membrane. The membrane was soaked for 24 h with anti-flag tag antibody and 1 h with goat anti mouse hrp as a secondary antibody for detection of the target protein. The membrane exposed to the stain and developed on the film. Although the antibodies should just bind to the flag tagged protein, the exposed X-ray film shows a lot of background (**Figure IV-14**).

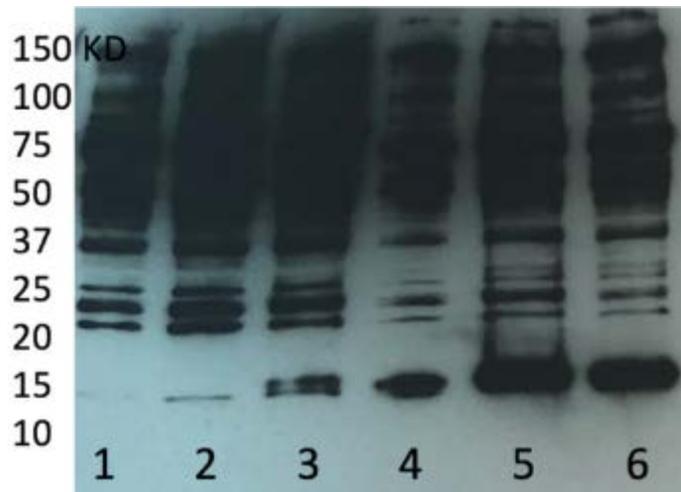


Figure IV-14: Exposed x-ray film from the western blotting. Row 1, 2 and 3 extract from controls (cells without changing media, cells with changing the media in transfection). Row 4, 5, 6 extract from transfection with hCRBP^{II} plasmid (transfection with 350 ng of extract).

Background maybe because of the high amount of extracts (20 ng), which loaded to the SDS PAGE gel. From the X-ray film from the western blotting, it can conclude that there is more of our target protein (15KD) in sample 5. 750ng of the plasmid transfected sample 5; therefore, we use the same amount of the plasmid (750ng) for our next transfection trials. For improving the background problem, increasing the specificity of the antibody to the flag tagged protein, we decided to lower the amount of extract that we inserted to SDS PAGE. In the second trial, 8 ng of the controls (one without the transfection and one transfected by empty pCMV-flag tag), 8 ng, 4 ng and 2 ng of the extract (transfected by plasmid) used in SDS PAGE. The first and secondary antibody reused from last time. Also, the membrane soaked for 1 h instead of 24 h with the first antibody. The results improved significantly, with substantially less background. These data demonstrate that by using around 4-8 ng of the extract and soaking 1h with the first antibody, we can improve the western blotting results (**Figure IV-15**).

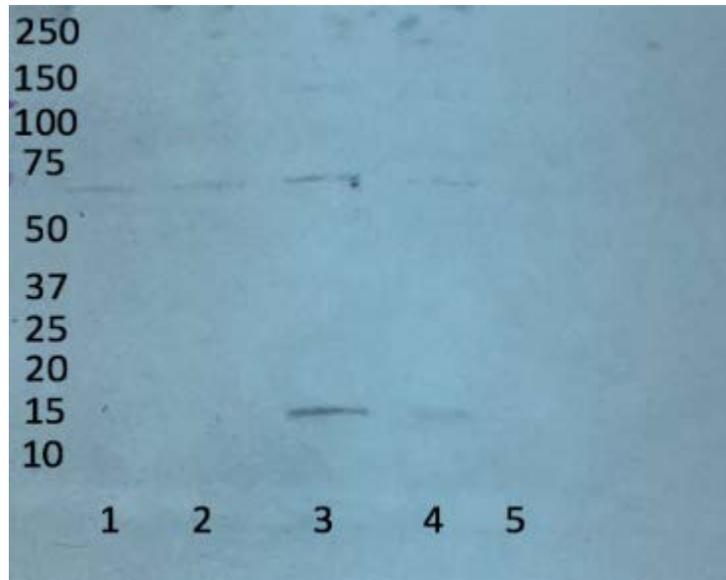


Figure IV-15: Exposed x-ray after western blotting. Sample 1 and 2 are from 8 ng extract of control and 8ng extract of empty pCMV-flag tag vector. Sample 3,4 and 5 are from 8 ng,4 ng and 2 ng extract of infection with 750ng plasmid, respectively.

IV-6-2 Investigate the size of hCRBP_{II} in mammalian expression

Our last experiments confirmed the expression of Q108K:T51D in HeLa cells. As explained before, the existence of domain swapping needs to be investigated in these expression trials. We need a way to compare the size of hCRBP_{II} from the mammalian expression with monomer and dimer form from bacterial expression. The first solution was running a native gel instead of SDS PAGE; however, since the flag tag is highly negatively charged, we could not compare it with monomer and dimer in bacterial expression without the flag tag. Another solution was running size exclusion chromatography from the extract and running the western blot from the fractions of the SEC column. We first expressed monomer and dimer hCRBP_{II} in bacteria and ran it through a superdex 75 size exclusion chromatography column, which was a different column than it used before. The monomer (Q108K: K40L: T51V hCRBP_{II} mutant) came out at fractions 24, 25, and

26. The dimer (Q108K: T51D hCRBP II mutant) came out mostly in fractions 20 and 21 (**Figure IV-16**).

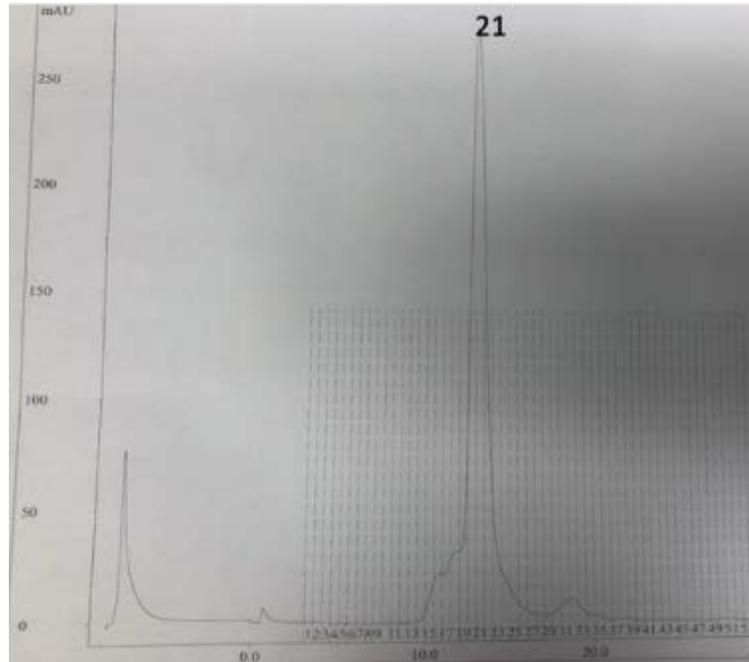


Figure IV-16: Chromatogram of size exclusion chromatography. 1.5mg of domain swapped dimer hCRBP II (Q108K:T51D mutant) ran through the superdex 75.

To run the Size exclusion chromatography from mammalian expression, more hCRBP II needs to be expressed in HeLa cells. The level of expression of hCRBP II is low in HeLa cells. Therefore, another six-well culture plate was used to transfect by 750ng plasmid. The resulting extracts from the six wells of culture plate were around 0.5mg. It is essential to clarify that the amount of hCRBP II is much lower than 0.5 mg in the extract since we did not purify the extract. Purification with flag tag may interfere with the folding of the domain-swapped dimer; therefore, we decided to only use the size exclusion chromatography as a purification system for this experiment (**Figure IV-17**).

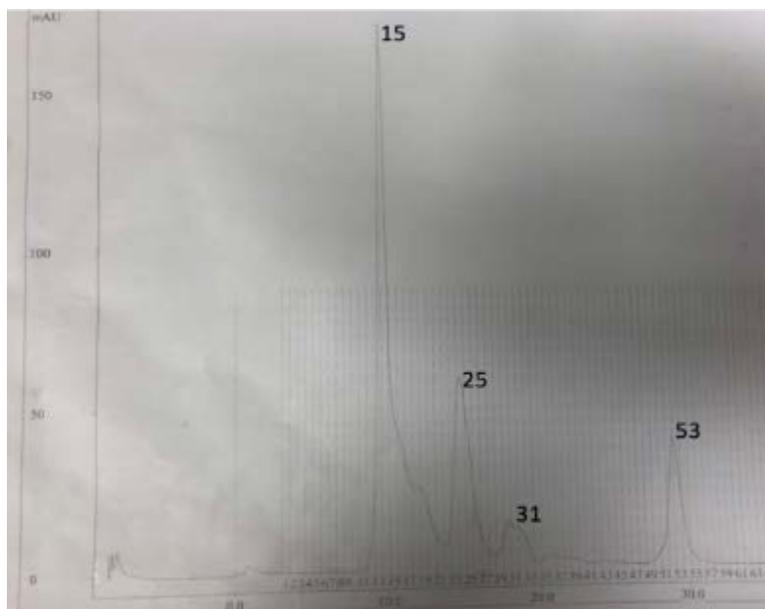


Figure IV-17: Chromatogram of size exclusion chromatography (superdex 75) for 0.5 mammalian expression extract.

We used western blotting to clarify the fractions with hCRBP_{II}. Fractions 13-28 of size exclusion chromatography were concentrated and run through SDS PAGE gel, following with the western blotting. The x-ray film surprisingly demonstrates the 15KD band only in fraction 16 (**Figure IV-18**). Fraction 16 could be a higher-order oligomer form for the hCRBP_{II}, or it can be our target protein bound to another protein. In this experiment, we did not find any clue that demonstrates the formation of monomer hCRBP_{II}. Several tests can be done to investigate the formation of DSD hCRBP_{II} in mammalian cells. Flag tag can be cloned in monomer and dimer in bacterial expression to run the native gel for comparing with the mammalian extract. To have an easier purification, His-tag can be used instead of the flag-tag for purification with Ni-column. Also, using another expression system, such as a baculovirus expression system (like SF9), may help for the level of expression.

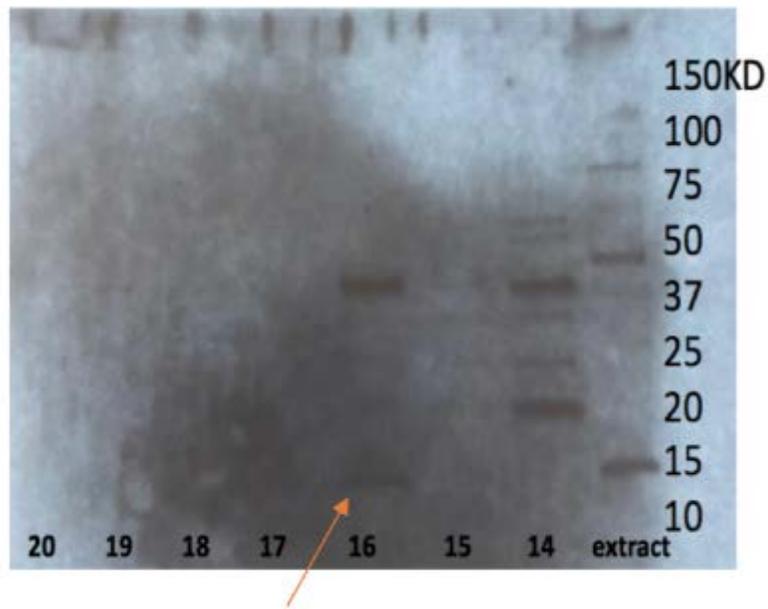


Figure IV-18: X-ray film from western blotting for extract before running the SEC column and fraction 14- 20 after SEC column.

IV-7 DOMAIN SWAPPING IN RETINOID BINDING PROTEIN 7 AND FABP2

We tried some other members of the iLBP family to investigate the domain swapping in them. Expression of Retinoid binding protein 7 does not show any formation of dimer or any oligomers.. Another member of the iLBP family which was very interesting to study domain swapping was FABP2. FABP2 has strong affinity to long chain fatty acids and may act as a lipid sensor. It also affects the cell growth and proliferation. In hCRBP2 case, mutating tyrosine 60 to the hydrophobic residue Leucine lead to almost all domain swapped dimer. Sequence comparison between WT-hCRBP2 and WT-hFABP2 shows Isoleucine instead of Tyrosine in hCRBP2. Due to the high similarity in the hinge region, we hypothesized FABP2 has a better chance to have more domain

swapped dimer in the wild type. However, expression of FABP2 shows formation of exclusively monomer for this protein (**Figure IV-19**).

```

hCRBP2  11  TWEMESNENFEGYMKALDIDFATRKIAVRLTQTKVIDQDGNFKTKTISTFRNYDVDFTV  70
          TW+++ +EN++ +M+ + ++  RK+A          I Q+G+ F  K +S FRN +V F +
FABP2   5   TWKVDRSENYDKFMEKMGVNIVKRKLAAHDNLKLTITQEGNKFTVKESSAFRNIEVVFEL  64
          +
hCRBP2  71  GVEFDEYTKSLDNRHVKALVTWEGDVLV  98
          GV F+ Y  + D  ++  + EG+ L+
FABP2   65  GVTFN-YNLA-DGTELRGTWSLEGNKLI  90

```

Figure IV-19: Amino acid sequence alignment for hCRBP2 and FABP2. The red box shows the amino acids in the hinge loop and its neighbor residues important in domain swapping.

Other members of iLBP, such as FABP4 and FABP5 have been studied. I will discuss our studies in formation of domain swapped dimerization FABP5 in the next chapter.

IV-8 CONCLUSION

We performed CD and fluorescence studies to investigate the folding intermediates in hCRBP2. According to our results, the folding of hCRBP2 does not support the two state model; therefore, we can expect at least one folding intermediate during hCRBP2 expression. More experiments, such as H-D mass spectrometry, needs to be done to investigate the folding intermediates. The phase relationship and effect of some residues for the formation of domain swapping in hCRBP2 was tested. Moreover, experiments have been done to examine the expression of DSD hCRBP2 in HeLa cells. The existence of DSD in the iLBP family during bacterial expression may have physiological relevance. Therefore, more experiments need to be done to confirm the formation of DSD in mammalian cells.

IV-9 EXPERIMENTAL

Most of the experimental part was explained in my master thesis as well.³⁰

IV-9-1 Site directed mutagenesis

For mutagenesis, we used the HCRBPII in the pET17b vector described following the Quick-change Site-directed Mutagenesis Kit protocol from Agilent Technologies (**Table IV-1**).

Table IV-1: PCR protocol for mutagenesis.

| | |
|------------------------------|----------------------------|
| Total Reaction Volume | 50μL |
| Template (DNA plasmid) | 70ng (x L) |
| Primer Forward | 20 pmol (y L) |
| Primer Reverse | 20 pmol (z L) |
| dNTP | 1 L |
| 10x pfu Buffer | 5 L |
| Pfu Turbo (DNA Polymerase) | 1 L |
| DI water | 50-x-y-z-7 L |

| PCR Program | | |
|-------------|---|-------------|
| 1x | 95 C | 30 min |
| | 95 C | 30 sec |
| 20x | Temperature 3-5 C lower than primer melting temperature | 1 min |
| | 72 C | 4min 30 sec |
| 1x | 72 C | 10 min |
| 1x | 25 C | 10 min |

For transformation, The PCR product was transformed into 50 μ L DH5alpha cells competent cells and grown on Luria-Bertani (LB)-agar plates treated with Ampicillin (75 g/mL) for 16-24h. after that, a single colony was picked from the plate and inoculated in 10 mL LB medium containing 100mg/mL ampicillin and grown at 37°C while shaking, for 12-16 hours. Then, using the QIAGEN Miniprep DNA purification kit, we did DNA purification. MSU gene sequencing facility verified the construct by using T7 primer.

IV-9-2 Primers

T59LhCRBP_{II}

Forward: 5'- CACATTCCGCTTATATGATGTGGATTTC-3'

Reverse: 5'- CTAGTGGTTTTTGTCTTG -3'

R58L hCRBP_{II}

Forward: 5'-TAGCACATTCTTAAACTATGATGTGGATTTC-3'

Reverse:5'-GTGGTTTTTGTCTTGAAG-3'

R58Q hCRBP_{II}

Forward: 5'-TAGCACATTCCAAAACTATGATGTG-3'

Reverse: 5'- GTGGTTTTTGTCTTGAAG-3'

Q108K-hCRBP_{II}

Forward:5'-CCGCGGCTGGAAGAAGTGGATTGAGGGGG-3'

Reverse: 5'-CCCCCTCAATCCACTTCTTCCAGCCGCGG-3'

K40L-hCRBP II

Forward: 5'-CTCACTCAGACGCTGGTTATTGATCAAGATGG -3'

Reverse: 5'-CCATCTTGATCAATAACCAGCGTCTGAGTGAG-3'

T51D-hCRBP II

Forward: 5'-GGTGATAACTTCAAGGATAAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTATCCTTGAAGTTATCACC-3'

T51V-hCRBP II

Forward: 5'-GGTGATAACTTCAAGGTAAAAACCACTAGCAC-3'

Reverse: 5'-GTGCTAGTGGTTTTACCTTGAAGTTATCACC-3'

IV-9-3 In vitro refolding

For initiating this experiment, the protein was expressed in *E-coli* at 16 °C for two days and purified using a fast Q column, ion-exchanged chromatography from GH Health Sciences, and size SEC column (Superdex S75 16/600 HiLoad column from GE healthcare company). Chromatogram of ion-exchanged and size exclusion chromatography shows mostly dimer formation for Q108K;T51D. After the purification step, most of the fractions that contain dimer of the protein were collected and concentrated to concentration 8 mg/ μ L. Then, four equivalents of the ligand were added to half of the protein and incubated for a couple of hours at 4°C to make

sure that the ligand was bound to the protein. After that, 1mL of the protein solution from Apo and Holo form was inserted to 6 separate dialysis cassettes (3 of them used for the Holo form of the protein and the other ones used for Apo form) (**Figure IV-20**). Also, 2mL of 8M urea was added to each cassette; therefore, the concentration of the urea was 5.4 in each cassette, which is enough to unfold all of the protein. When the cassette became ready, they were put inside the folding buffer, and every hour the buffer was replaced with the new buffer. We exchanged the folding buffer seven times to make sure that most of the urea was removed from the solution of the protein. In the last step, we loaded the refolded protein to a size exclusion chromatography column (Superdex S75 16/600 HiLoad column from GE healthcare company).



Figure IV-20: Dialysis cassettes uses for refolding experiments.

IV-9-4 Site directed mutagenesis (pCMV-Vector)

The hCRBP_{II} was cloned into pCMV-flag tag vector by using NotI and EcoRI cutting sites. The aforementioned construct was used for mutagenesis by using the Quick Change Site-directed Mutagenesis Kit protocol from Agilent Technologies company which was explained in chapter I.

IV-9-5 Amino acid sequence of hCRBP_{II} with flag tag in N terminus

MDYKDDDDKLADRMTRDQNGTWEMESNENFEGYMKALDIDFATR_KIAVRLTQTKVID
QDGDNFKDKTTSTFRNYDVDFTVGVFDEYTKSLDNRHV_KALVTWEGDVLVCVQKGE
KENRGWKKWIEGDKLYLELTCGDQVCRQVF_KKKK

IV-9-6 Mammalian expression

To perform transient transfection, all of these steps needed to be done in the laminar flow hood. 2.5 mL HeLa cells (concentration 0.8×10^5) in DMEM were added to each well in a 6-well culture plate. The DMEM contained 5% FBS and Penicillin/Streptomycin (P/S). The plate mentioned above was incubated at 37 °C for 24h (cells should be 90% confluent at this time). The next step is preparing two solutions for transfection. 14mL polypropylene tubes were labeled for A and B solutions. 250 μ L DMEM (without FBS and P/S) was added to each tube. 2.5 μ L Lipofectamine 3000 reagent (Invitrogen) company were added to solution B tubes and incubated for 15 min. Meanwhile, 1 μ L of P3000 (Invitrogen) company and appropriate amounts of DNA was added to solution A and incubated for 15min. Then, solution B was added to solution A and the mixture was incubated for another 15min. Meanwhile, the old media were removed from the wells from the culture plate, and the plates washed with 2mL PBS. We aspirated the PBS and 500 μ L DMEM

without FBS, and P/S was added to each well. Each mixture of solution A/B added dropwise to each well accordingly, and the plate was incubated in 37 °C/5% CO₂ for 5-7 hrs. The cells were harvested after 48 hrs. The media was removed from the wells. 1-1.5mL of PBS was added to each well, and the cells were scraped from the plate by scrapper. The mixture of cells with PBS was collected in Eppendorf tubes and spin for 7 min in 2000rpm. The PBS solution was removed from the cells. 50ML of the extraction buffer was added to the pellet and mixed well. Then, the mixtures were incubated in ice for 20 min and then spun for 13000rpm for 7min in the cold room. The supernatants (extract) were transferred to new tubes and were frozen in -80 °C.

REFERENCES

REFERENCES

1. Assar, Z. *et al.* Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates. *Structure* **24**, 1590–1598 (2016).
2. Ghanbarpour, A. *et al.* Engineering the hCRBP II Domain-Swapped Dimer into a New Class of Protein Switches. *J. Am. Chem. Soc.* **141**, 17125–17132 (2019).
3. Ghanbarpour, A. Functional Control of Soluble Rhodopsin Mimics Using High Resolution Structure-Based Design and Evaluation. (Michigan State University, 2019).
4. Nossoni, Z. Structural insight in to the mechanism of wavelength tuning in a rhodopsin mimic and a single mutation resulted an extensive 3D domain swapped dimerization in hCRBP II. (Michigan State University, 2014).
5. Assar, Z. Elucidation of iLBP Family Folding Pathway and Study of Reengineering Them as Fluorescent Protein Tags via Structural Analysis. (Michigan State University, 2018).
6. Bennett, M. J., Schlunegger, M. P. & Eisenberg, D. 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.* **4**, 2455–2468 (1995).
7. 3D domain swapping: As domains continue to swap - Liu - 2002 - Protein Science - Wiley Online Library. <https://onlinelibrary.wiley.com/doi/abs/10.1110/ps.0201402>.
8. Bennett, M. J., Choe, S. & Eisenberg, D. Refined structure of dimeric diphtheria toxin at 2.0 Å resolution. *Protein Science* **3**, 1444–1463.
9. Liu, Y. & Eisenberg, D. 3D domain swapping: as domains continue to swap. *Protein Sci.* **11**, 1285–1299 (2002).
10. Koharudin, L. M. I., Liu, L. & Gronenborn, A. M. Different 3D domain-swapped oligomeric cyanovirin-N structures suggest trapped folding intermediates. *Proc Natl Acad Sci U S A* **110**, 7702–7707 (2013).
11. Mascarenhas, N. M. & Gosavi, S. Understanding protein domain-swapping using structure-based models of protein folding. *Prog Biophys Mol Biol* **128**, 113–120 (2017).
12. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D. & Zhang, K. Y. Single-site mutations induce 3D domain swapping in the B1 domain of protein L from *Peptostreptococcus magnus*. *Structure* **9**, 1017–1027 (2001).
13. Fändrich, M., Fletcher, M. A. & Dobson, C. M. Amyloid fibrils from muscle myoglobin. *Nature* **410**, 165–166 (2001).

14. Cleary, J. P. *et al.* Natural oligomers of the amyloid-beta protein specifically disrupt cognitive function. *Nat. Neurosci.* **8**, 79–84 (2005).
15. Veerkamp, J. H., Kuppevelt, T. H. M. S. M. van, Maatman, R. G. H. J. & Prinsen, C. F. M. Structural and functional aspects of cytosolic fatty acid-binding proteins. *Prostaglandins, Leukotrienes and Essential Fatty Acids* **49**, 887–906 (1993).
16. Banaszak, L. *et al.* Lipid-Binding Proteins: A Family of Fatty Acid and Retinoid Transport Proteins. in *Advances in Protein Chemistry* (eds. Anfinsen, C. B., Edsall, J. T., Richards, F. M. & Eisenberg, D. S.) vol. 45 89–151 (Academic Press, 1994).
17. Lücke, C., Huang, S., Rademacher, M. & Rüterjans, H. New insights into intracellular lipid binding proteins: The role of buried water. *Protein Sci* **11**, 2382–2392 (2002).
18. Liu, Z. P., Rizo, J. & Gierasch, L. M. Equilibrium folding studies of cellular retinoic acid binding protein, a predominantly beta-sheet protein. *Biochemistry* **33**, 134–142 (1994).
19. Gunasekaran, K., Eyles, S. J., Hagler, A. T. & Gierasch, L. M. Keeping it in the family: folding studies of related proteins. *Curr. Opin. Struct. Biol.* **11**, 83–93 (2001).
20. Zwanzig, R. Two-state models of protein folding kinetics. *Proc Natl Acad Sci U S A* **94**, 148–150 (1997).
21. Khan, K. H. Gene Expression in Mammalian Cells and its Applications. *Adv Pharm Bull* **3**, 257–263 (2013).
22. Protein Folding in the Endoplasmic Reticulum.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3632058/>.
23. Nettleship, J. E., Assenberg, R., Diprose, J. M., Rahman-Huq, N. & Owens, R. J. Recent advances in the production of proteins in insect and mammalian cells for structural biology. *Journal of Structural Biology* **172**, 55–65 (2010).
24. N, P. *et al.* Recombinant Protein Expression for Structural Biology in HEK 293F Suspension Cells: A Novel and Accessible Approach. *Journal of visualized experiments : JoVE* <https://pubmed.ncbi.nlm.nih.gov/25349981/> (2014) doi:10.3791/51897.
25. Masters, J. R. HeLa cells 50 years on: the good, the bad and the ugly. *Nature Reviews Cancer* **2**, 315 (2002).
26. Rahbari, R. *et al.* A novel L1 retrotransposon marker for HeLa cell line identification. *Biotechniques* **46**, 277–284 (2009).
27. Check your cultures! A list of cross-contaminated or misidentified cell lines - Capes-Davis - 2010 - International Journal of Cancer - Wiley Online Library.
<https://onlinelibrary.wiley.com/doi/full/10.1002/ijc.25242>.

28. Towbin, H., Staehelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350–4354 (1979).
29. Mahmood, T. & Yang, P.-C. Western Blot: Technique, Theory, and Trouble Shooting. *N Am J Med Sci* **4**, 429–434 (2012).
30. Ehyaei, N. Domain Swapping in the ILBP Family. (Michigan State University, 2018).

CHAPTER V: DOMAIN SWAPPING OF FATTY ACID-BINDING PROTEIN 5 (FABP5)

The Existence of domain swapping for hCRBP2 is likely to have physiological importance.^{1 2} Therefore, Investigating the mechanism of domain swap dimerization in the iLBP family is critical. We ask if other members of the iLBP family can have undergone domain swapping.³ Among another member of the iLBP family, Fatty acid-binding protein 5 (FABP5), was reported to form a very similar domain-swapped dimer (Figure V-1) bound to AEA (endocannabinoid anandamide).⁴ Although the sequence identity between hCRBP2 and FABP5 is low, we decided to expand the research on domain swapping of FABP5 as well.

V-1 FATTY ACID-BINDING PROTEINS (FABP)

Fatty acid-binding proteins are a subfamily of iLBP that are responsible for the transport of fatty acids through the cell.⁵ FABPs are found in different parts of the body, and the nomenclature of them is based on the tissues in which they have been discovered. As explained in the first chapter, the structure of these proteins contains a β barrel as well as two α helices that play as a lead for these proteins (**Figure V-1**).^{6 7 8}

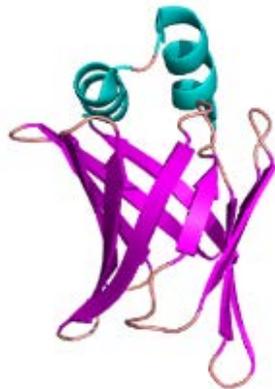


Figure V-1: Crystal structure of FABP4 (PDB;4NNS).

Fatty acid-binding protein 5 (FABP5) known as epidermal FABP5 is part of this family, and it is responsible for transferring the endocannabinoid anandamide (AEA) in the cytosol.^{9 10} This protein is found mostly in epidermal cells, brain, liver, kidney, lung, and adipose tissue. There are reported structures of Apo monomer FABP5, and monomer FABP5 bound to different ligands, including Linoleic acid and palmitic acid. **(Figure V-2) (Figure V-3).**^{11 5}

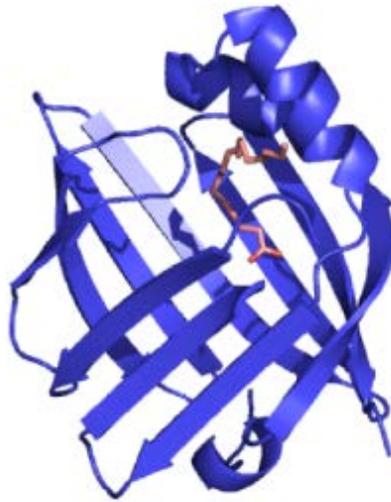


Figure V-2: Structure of human hFABP5 in complex with Linoleic acid (PDB: 4LKT).

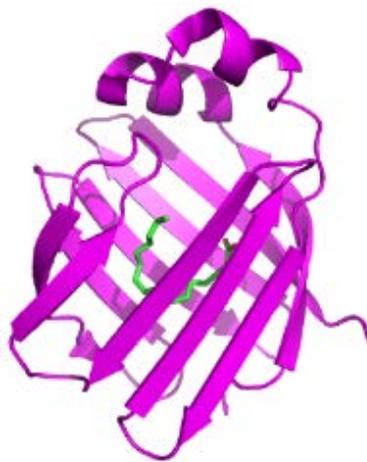


Figure V-3: Structure of human hFABP5 in complex with palmitic acid (PDB: 1b56).

Sanson et al. reported domain swapping for FABP5. Before the first domain-swapped structure for FABP5 reported by this group, all the other structures for this protein had been reported to be monomers (**Figure V-4**).⁴ The domain-swapped dimer structure of FABP5, similar to hCRBP II case, also helps us to investigate the existence of domain swapping as a natural product for the iLBP family.

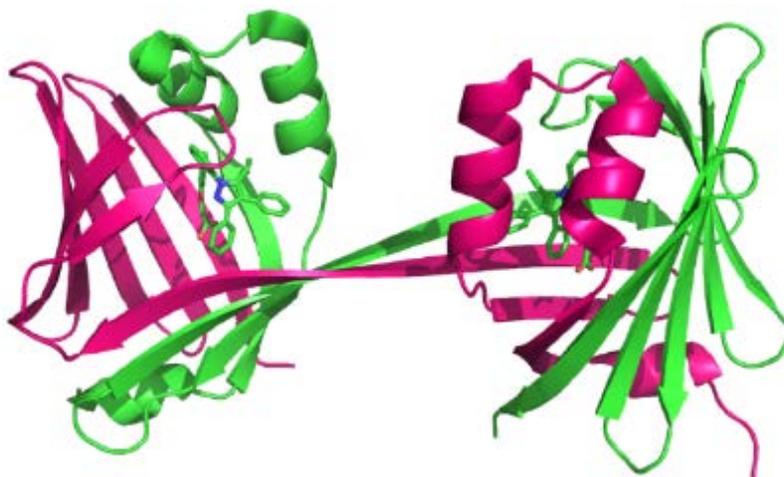


Figure V-4: Structure of domain swapped dimer for FABP5 bound to 2-AG(4AZM). Two chains are in pink and green color.

Sanson and his group demonstrate that the presence or absence of ligand does not affect the amount of dimer and monomer after expression. They suggested that conformational dynamics of FABP5 protein lead to the existence of the domain swapped dimer structure. From our hCRBP II studies, we showed that in hCRBP II the dimer and monomer do not convert to each other. In other words, domain swapped dimer appears to be a kinetically trapped product during folding and that does not convert to monomer under normal conditions.¹ Sanson and his group determined the structure FABP5 bound to the endocannabinoids anandamide (AEA) and 2-arachidonoylglycerol (2-AG) which is an inhibitor of AEA. (**Figure V-5**).⁴

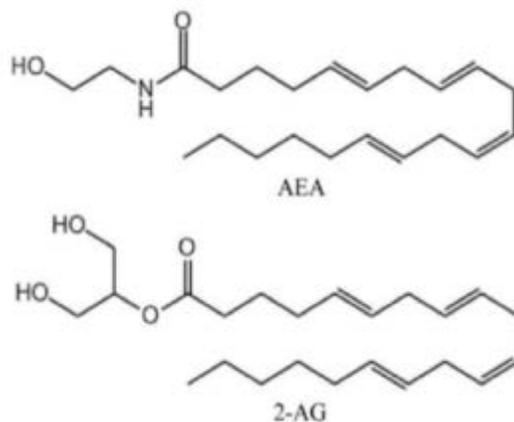


Figure V-5: Molecular structure of AEA and 2-AG.

Overlay of these two structures is illustrated in **(Figure V-6)**. The overlaid structures of FABP5 bound to AEA and 2-AG demonstrates a huge conformational change between these two structures. The conformational change between these two structures may have physiological relevant and it can lead to the allosteric regulation for this family of proteins.²

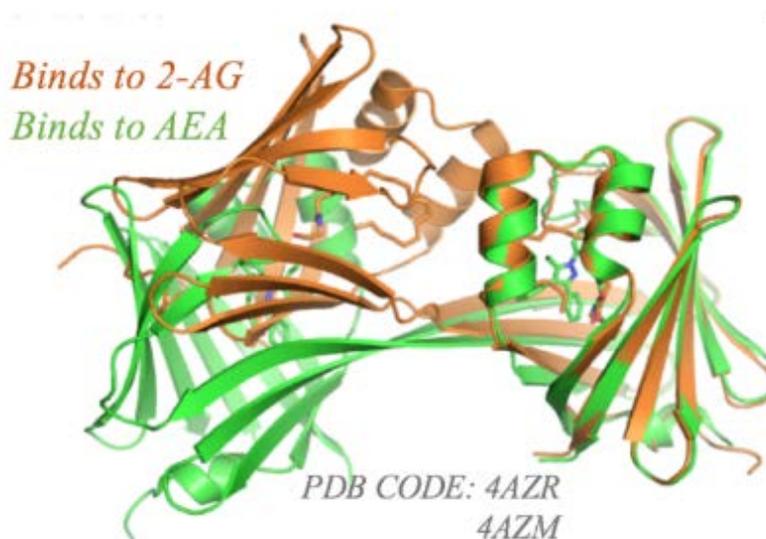


Figure V-6: Overlaid structures of FABP5 bound to 2-AG(green), and FABP5 bound to AEA(orange).

We decided to investigate the existence of domain swapping during bacterial expression without the addition of any ligands. In our studies, folding and structure of the protein is not affected by any extrinsic factors, more importantly, ligand binding.

V-2 APO MONOMER STRUCTURE OF FABP5

Dr. Zahra Assar-Nossoni and I initiated the investigation on the domain swapping in hFABP5. We used size exclusion chromatography to separate the monomer from dimer, the same method we used for hCRBP2. In most of our trials, we got a large monomer peak and a very small shoulder peak for the dimer. However, in some of the trials, size exclusion chromatography for FABP5 demonstrates different dimer to monomer ratios from each expression. The reason behind this difference in the monomer/dimer ratio is not apparent yet. In one of our first trials, we were able to produce a significant amount of dimer (dimer to monomer ratio around 30%), and we used SDS PAGE to characterize our purified protein on those monomer and dimer fractions (**Figure V-7**).

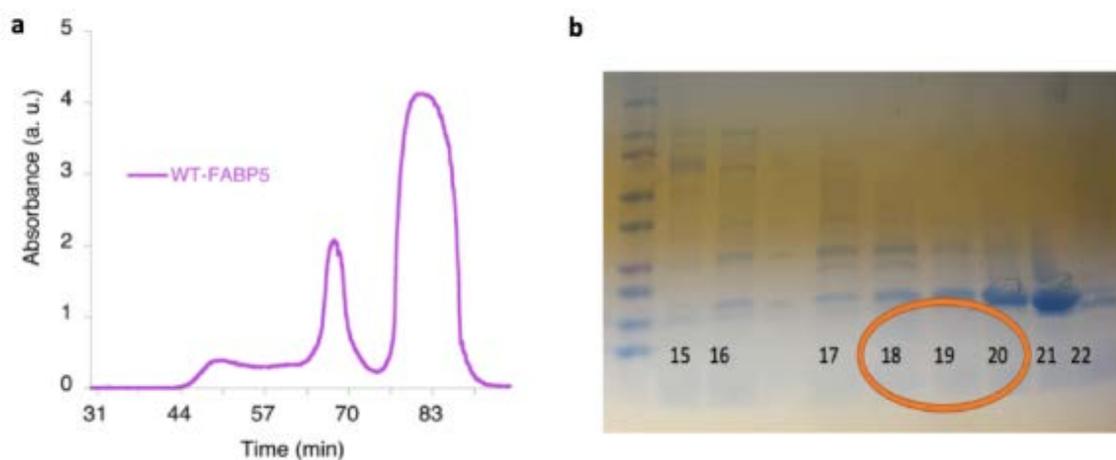


Figure V-7: a) chromatogram of size exclusion chromatography. b) SDS PAGE of fractions 15 to 22. Fractions 18, 19, 20 were mixed for putting the crystallization box.

We also tried to characterize monomer from dimer using native gel (in SEC, fraction 18, 19, 20, and 21 are expected to be our protein) and comparing it with the result with the monomer and dimer of hCRBPII. However, we were not able to say in which fractions we have dimer. We mixed fractions 18, 19, 20 that they were close to dimer size, and we set up the crystallization screen for them as explained in the experimental part. We were able to obtain crystals and determine the structure in the P212121 space group for this protein (**Table V-1**). However, unlike what we expected, the structure was apo monomer. The overlaid structure of this Apo monomer Fabp5 and past reported apo structure of FABP5 demonstrated conformational changes between residues mostly in the helices. (**Figure V-8**).

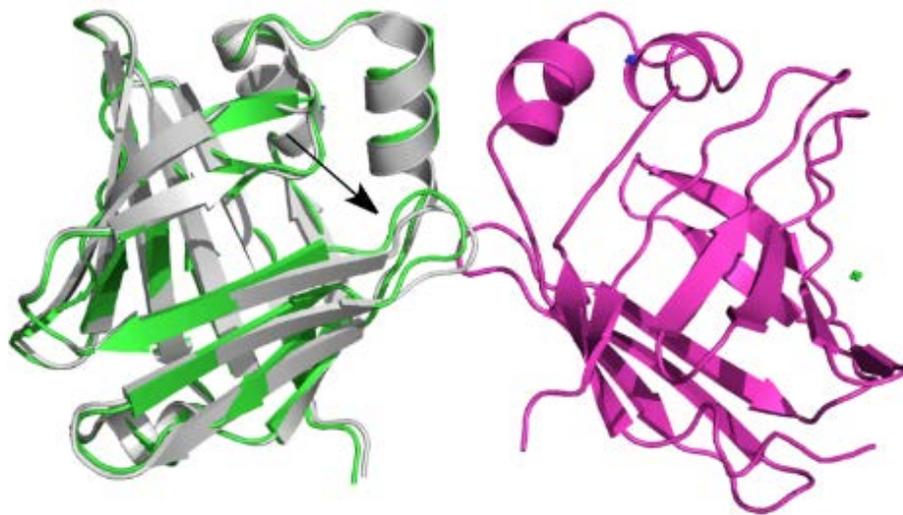


Figure V-8: *The overlaid structure of our apo monomer FABP5 in green (chain A green and chain B pink) vs reported apo monomer FABP5 in gray(PDB:4LKP).*

V-3 DOMAIN SWAPPED DIMER STRUCTURES FOR FABP5 BOUND TO PALMITIC ACID

After many expression trials, we were successful to obtain dimer peak again. From our last experience, we expect to have mostly monomer in fraction 20 and maybe 20. Therefore, we concentrated fraction 18 of SEC separately and set up crystallization plate for it. Finally, we were able to solve the domain swapped dimer of hFABP5 in P6322 space group (**Table V-2**). Although the solved structure was domain swapped dimer as we expected, there was an extra electron density in the binding pocket of the protein. The continuous electron density in the binding pocket seems to be related to a fatty acid. Since we did not add any ligand during the expression of protein, we predicted that it might be one of the common fatty acids in E-Coli, BL21 cells. Therefore, we tried to fit various common fatty acids and ligands for hFABP5. We also tried to fit endocannabinoids anandamide (AEA) to illustrate the conformation and size of this fatty acid (**Figure V-9**)

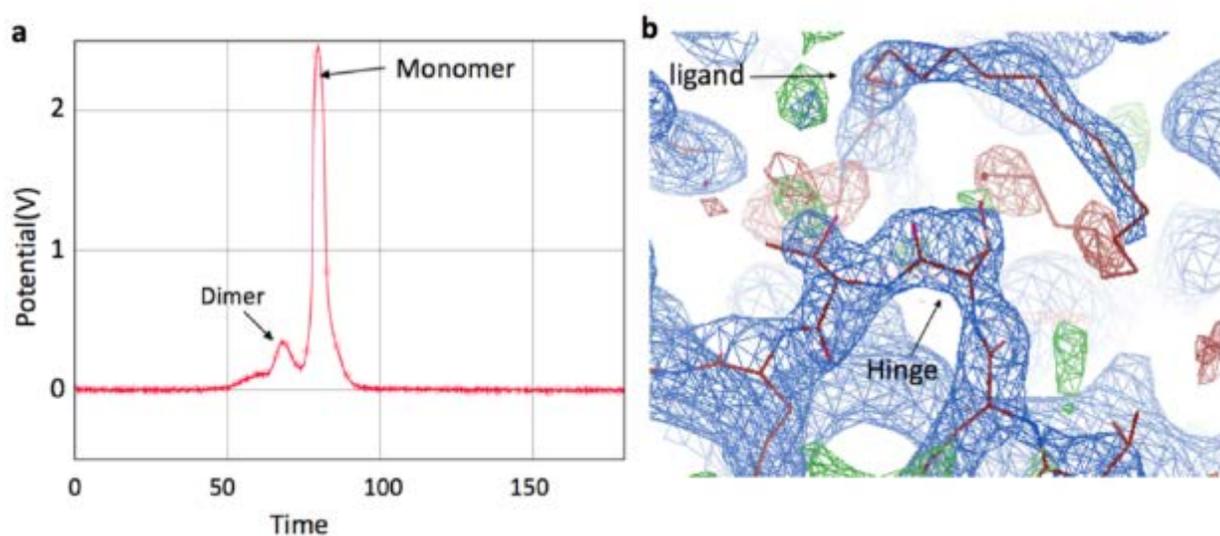


Figure V-9: a) SEC for FABP5 b) crystal structure of FABP5 domain swapped dimer with AEA as a ligand in extra electron density.

After searching through common fatty acids in bacteria and trials for fitting them in the extra electron density, palmitic acid matched perfectly to the electron density. Palmitic acid has 16 carbons and it is one of the most common fatty acids produced in bacteria. (Figure V-10).¹²

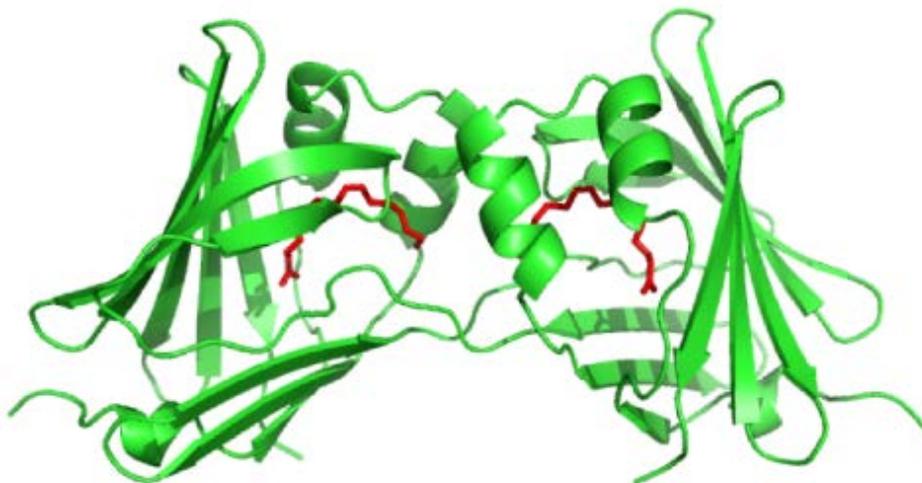


Figure V-10: Structure of domain swapped dimer hFABP5 bound to palmitic acid.

Formation of Domain swapped dimer during the expression without any other extrinsic factors might be a reason for the formation of DSD as a natural product of folding for this protein.

V-4 EFFORTS FOR GROWING APO DSD HFABP5 CRYSTALS

As mentioned, we tried to investigate the domain swapping without any external factors like ligands. At this point, we cannot be sure that Apo DSD of hFABP5 forms without the existence of the palmitic acid. Therefore, we expressed hFABP5 protein again to reproduce the dimer again and to remove the palmitic acid. This time we treated the dimer fraction, fraction 18 from the SEC column, with the lipidex column. Lipidex column is used for delipidating the non-specific lipids from the proteins. This resin contains many lipids, and based on the hydrophobic interactions, all free lipids should bind to the column. We used the lipidex column for removing the palmitic acid

from the binding pocket of hFABP5 protein. After crystallization screens, crystals were grown in a different condition in the C2221 space group (**Table V-3**). However, our new domain-swapped structure for hFABP5 showed the same extra electron density in the binding pocket as before, and it is probably palmitic acid again (**Figure V-11**).

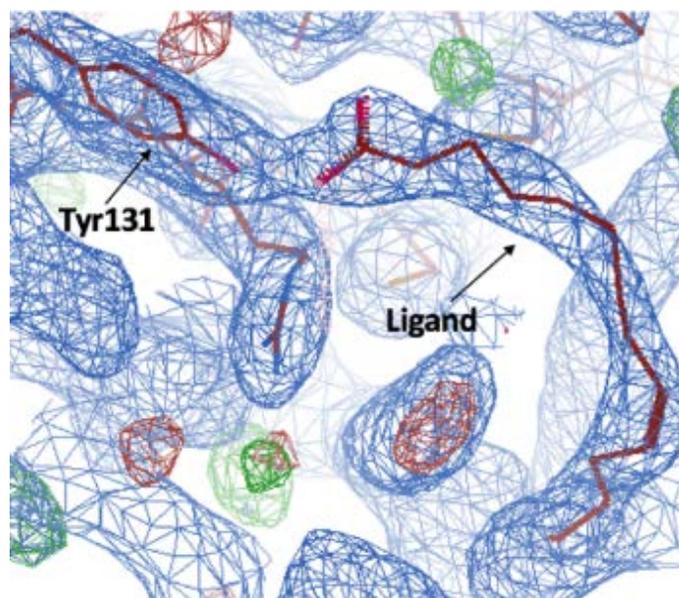
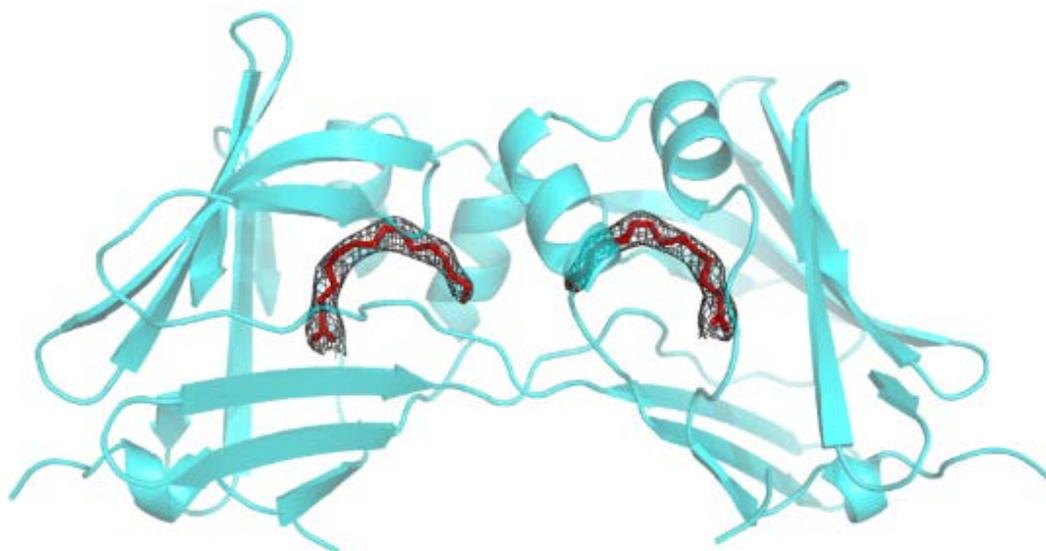


Figure V-11: Crystal structure of FABP5 with palmitic acid in the binding pocket.

Although these two DSD bound to palmitic acid grew in different conditions and one of them was treated with lipidex resin, palmitic acid has the same orientation in the binding pocket. The alignment between these two domain swapped structures shows mostly the same conformation for the residues except the residues in the hinge region (**Figure V-12**).



***Figure V-12:** Overlay of two FABP5 domain swapped structures bound to palmitic acid (first structure is in cyan and the most recent one is in orange). The arrows point to the hinge loops.*

V-5 INTERACTION OF PALMITIC ACID WITH HFABP5

Interaction between the carboxylate of fatty acid to basic residues, also hydrophobic residues near to the C chain is shown in (**Figure V-13**).

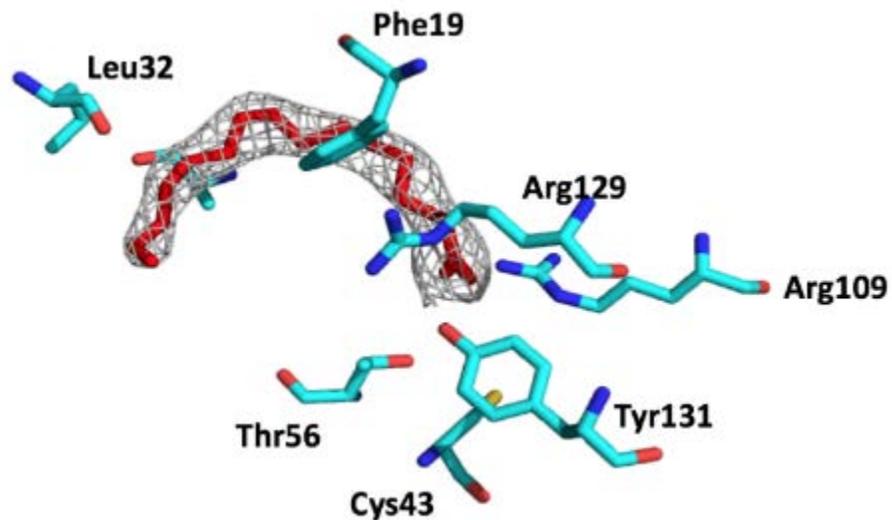


Figure V-13: Interaction of palmitic acid in the binding pocket of hFABP5.

The carboxylic group of the palmitic acid has interaction with Arg129, Tyr131, and Arg109 (Figure V-14).

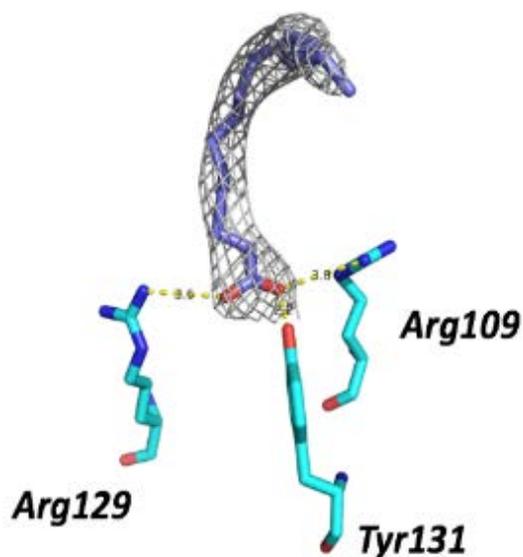


Figure V-14: Interaction of hFABP5 with the carboxylic group of palmitic acid.

Refolding has been done to make sure the Domain swapped dimer is not the effect of ligand binding. We assume after refolding and removing the hydrophobic interaction, the fatty acid is extracted from the substrate chamber of the protein and removed from the solution. SEC chromatography data and characterization through SDS PAGE gel shows the same ratio of monomer to dimer before and after the refolding experiment, which shows the formation of the apo Domain swapped dimer (**Figure V-15**). Unfortunately, crystallization trials for the dimer hFABP5 after refolding failed. Probably the DSD form of the protein maybe too flexible to form a crystal lattice without the ligand.

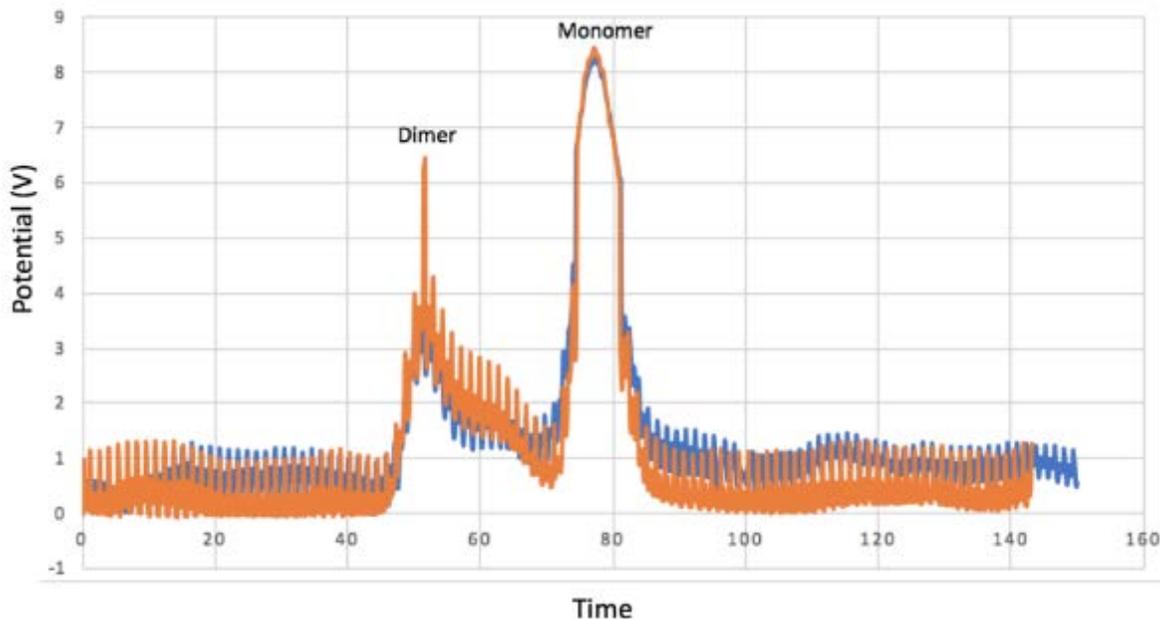


Figure V-15: SEC chromatogram for WT hFABP5 before (Orange graph) and after refolding (Blue graph). Dimer/monomer ratio remain constant.

V-6 COMPARISON BETWEEN MONOMER AND DSD HFABP5 IN COMPLEX WITH PALMITIC ACID

The monomer structure of WT-hFABP5 bound to palmitic acid has been previously reported (PDB: 1b56). From our experiments, we hypothesize that palmitic acid has a more robust interaction with DSD compare to monomer. In our trials, we could not find the structure of Apo monomer hFABP5. However, the palmitic acid in DSD seems to have strong interaction with residues of the binding pocket that even using the lipidex column could not remove fatty acid in this case. The overlaid structures of monomer hFABP5 and dimer hFABP5 bound to palmitic acid has been illustrated in **(Figure V-16)**.

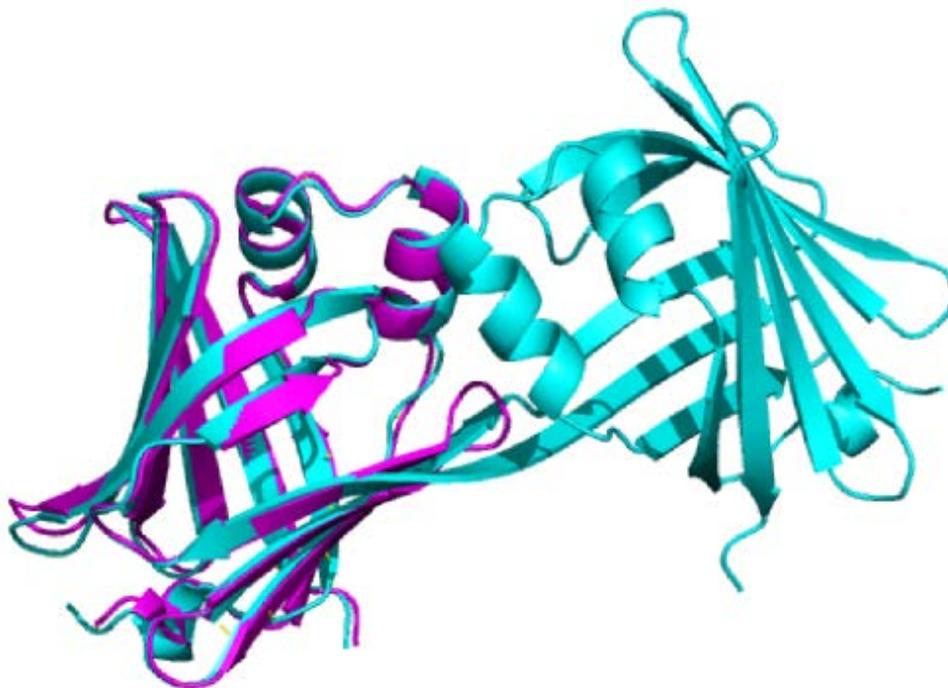


Figure V-16: overlaid structure of WT monomer hFABP5 (pink) and DSD hFABP5 (cyan) bound to palmitic acid.

Most of the conformation is the same in monomer and DSD, except the residues in the hinge loop. However, the conformation of palmitic acid seems to be more linear shape, compare to the monomer case which is U shaped (**Figure V-17**). The conformation of palmitic acid seems to be more linear shape when I compared to the monomer case which is U shaped. It was reported that in the case of FABP5 bound to Linoleic acid, LA's U-conformation correlates to its FABP5-activating form, whereas the L-conformation represents a non-activating binding mode. The formation of the U-conformation in domain swapped dimer, may be related to activation of FABP5 in DSD, similar to Linoleic acid case.

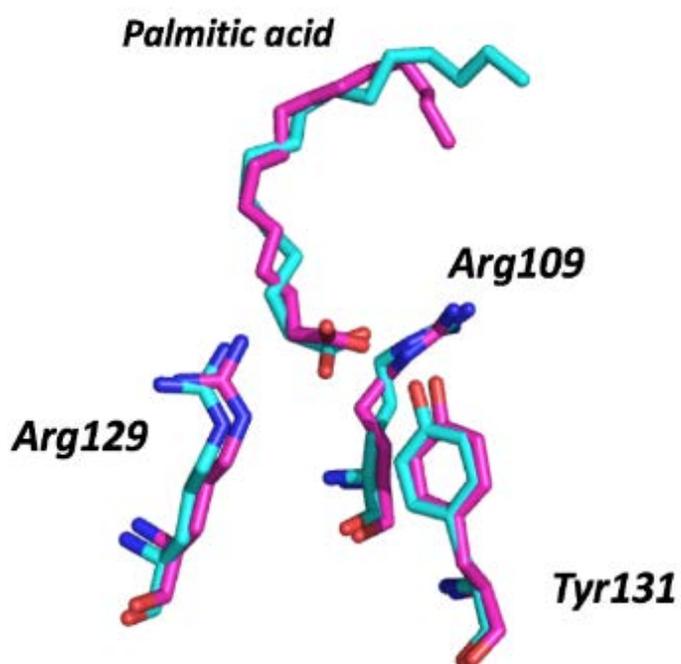


Figure V-17: Palmitic acids in overlaid structure of WT monomer hFABP5 (pink) and DSD hFABP5 (cyan).

V-7 CONFORMATIONAL CHANGE BETWEEN HFABP5 DSD BOUND TO DIFFERENT LIGANDS

As mentioned before, there is a huge conformational change between FABP5 domain-swapped structures bound with AEA and bound with 2-AG. The overlaid structures of hFABP5 domain-swapped structure with palmitic acid with these two structures is illustrated in **(Figure V-18)**.

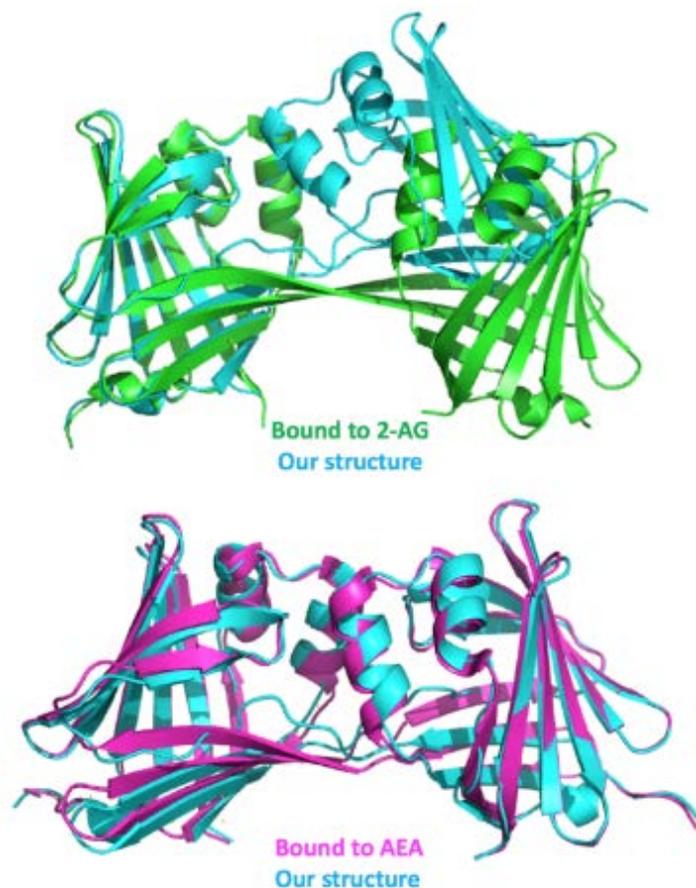


Figure V-18: Overlay between a) structures of DSD FABP5 binds to 2-AG (green) and FABP5 binds to palmitic acid (blue) b) structures of DSD FABP5 binds to AEA (pink) and binds to palmitic acid (blue).

The relative orientation of the two domains in DSD hFABP5 bound to palmitic acid is more similar to AEA, which has a closer molecular structure compared to a 2-AG molecule. These results suggest the ligand binding affect the relative orientation of domains in the DSD, undergoing conformational change. The change in the relative orientation of two domains and the hinge loop may lead to allosteric regulation for this protein. In hCRBP II case, ligand binding lead to a huge conformation change in domain swapped dimer.²

V-8 DETERMINATION OF MELTING POINT OF FABP5 BY USING THERMAL SHIFT ASSAY (TSA)

Thermal shift assays (TSA) were utilized to characterize recombinant human WT-FABP5 monomer and dimer biophysically.¹² TSA is a method for determination of melting temperature (T_m) of proteins. In this method, a fluorophore binds nonspecifically to hydrophobic surfaces, and water strongly quenches the fluorescence.¹³ By unfolding the protein, the hydrophobic surface binds to the dye, and water molecules are excluded from the fluorophore; therefore, the fluorescence will increase. In this case, we can obtain a thermal melting curve, which shows the change of fluorescence by increasing the temperature for the protein, and T_m (midpoint of the stability curve) can be obtained from this curve. In other words, this curve demonstrates the assay functions by protein denaturation over a temperature gradient, where during protein unfolding, exposed hydrophobic regions bind a dye and fluoresce due to solvent relaxation effects. In order to find the melting temperature of both FABP5 monomer and dimer, we expressed and purified the protein. After that we ran the SEC column for protein and fractions for both monomer (mostly fractions 21 22) and dimer (mostly fraction 18, 19) were collected and were sent to Dr. Zahra

Assar-Nossoni for TSA analysis in Cayman chemical company. The melting profile of WT-FABP5 Monomer (fraction 21, 22 of SEC) displays a peak at 61-63°C (**Figure V-19**).

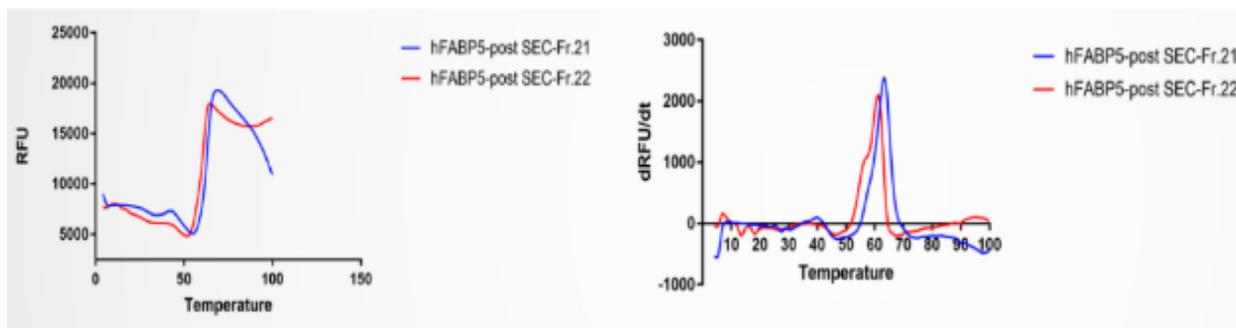


Figure V-19: Thermal stability assay for monomer FABP5 (fraction 21 in blue and 22 in red).

For fraction 18 and 19, displayed two major peaks (dRFU/dt vs. temperature), at 61°C, 42°C (**Figure V-20**).

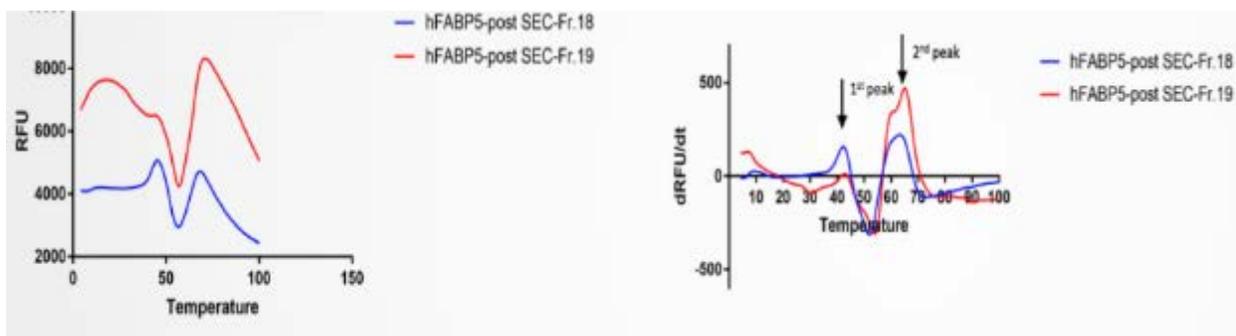


Figure V-20: Thermal stability assay for dimer FABP5 (fraction 18 in blue and 19 in red).

From SEC and our structural analysis, we know that fractions 21 and 22 are monomer fractions, and fractions 18 and 19 might be the mixture of dimer and monomer. Therefore, the melting point for monomer form estimated as 61-63, and Domain swapped dimer from 42-43°C. Also, by comparing the intensity of the peaks in fraction 18 and 19, in fraction 18 there is more dimer, and less monomer compares to fraction 19.

V-9 STUDY THE OCCURRENCE OF DOMAIN SWAPPING BY MUTATIONAL ANALYSIS OF HUMAN FABP5

Previous studies on domain swapping of hCRBP2 demonstrates some important residues that can change the ratio of monomer and domain swapped dimers. One important residue in hCRBP2 domain swapping is Tyr60, since mutation Y60L and Y60W were reported to increase the ratio of dimer to monomer significantly. Removing the hydrogen bond between Glu72 and Y60, it seems to be a reason for more DSD formation. In order to find important residues that affect the domain swapping of FABP5 project, we overlaid the structure of the monomer and dimer of hFABP5. Similar to hCRBP2 case, we try to target hydrogen bonds for finding these residues. In the monomer of FABP5, there is a hydrogen via a bond through a water network between Thr64 and Thr57; however, in the dimer there is a hydrogen bond between the water network between Q65 of one chain and T57 of the other chain (**Figure V- 21**).^{11 14}

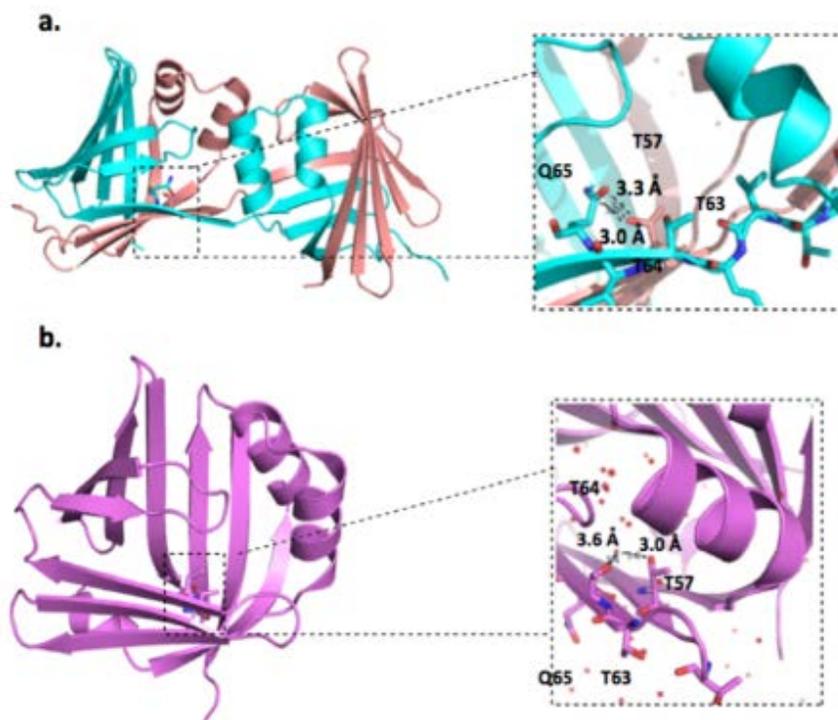


Figure V-21: a) Domain swapped dimer of FABP5 (PDB 4AZR), there is a hydrogen bond between Q65 and T57. b) monomer structure (purple, PDB ID 4LKP) Gln65 is pointing out of toward the solvent and T64 makes a hydrogen bond to Thr57 via water network.

We hypothesized that by removing the hydrogen bond between Q65 with T57 in DSD, the formation of the monomer would be more probable. Therefore, we mutated glutamine to alanine, which is a small hydrophobic residue to remove the hydrogen bond. The chromatogram of size exclusion chromatography for this mutant demonstrates the same amount of dimer compare to the wild type FABP5, which was not as we predicted (**Figure V-22**). We could obtain small crystals of this mutant, but they did not diffract. Also, we made Q65W and Q65M mutations as well, which in both mutations, only the monomer form of FABP5, was expressed. In another trial, we also applied the phase relationship for hFABP5 to increase dimerization, and we have done different

mutations like T64L and T64E. However, our hypothesis seems not to work, and we obtained a monomer form by expressing these mutations (**Figure V-23**).

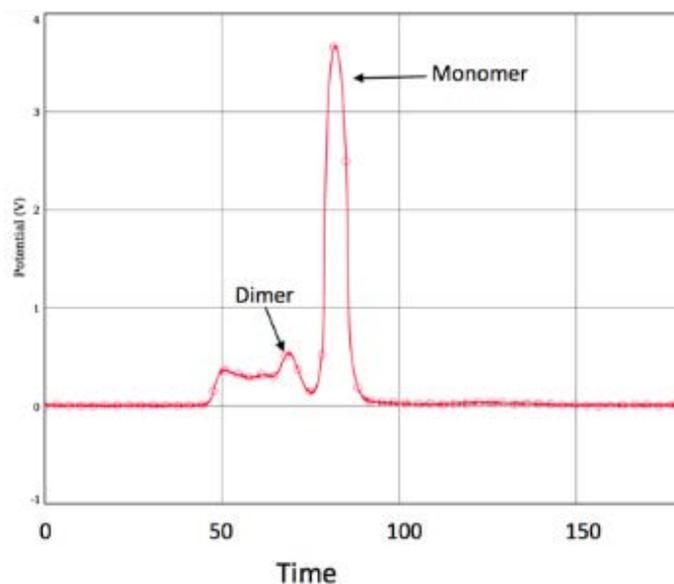


Figure V-22: Shows SEC after expression of Q65A FABP5.

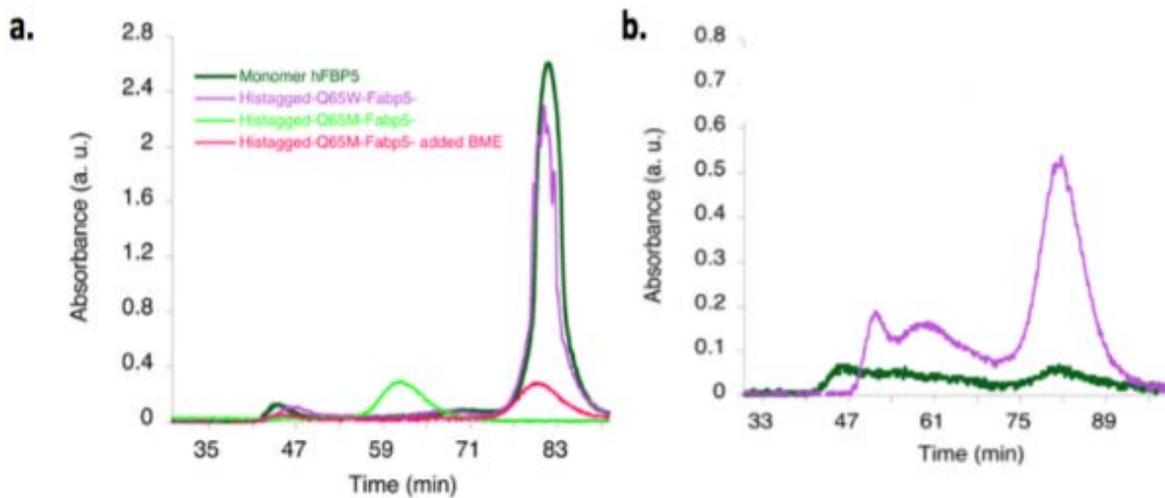


Figure V-23: a) Size exclusion chromatogram of Gln65 mutants of FABP5. b. SEC of T64 mutants of FABP5 (Thr63E pink, T63L green).

As I mentioned, each expression trial for WT hFABP5 leads to a different level of expression of DSD. Therefore, it is essential to elucidate the reasons behind the various ratios of WT hFABP5 dimer and monomer before further studies on mutational analysis.

V-10 INVESTIGATE OF THE DETERMINANTS OF DIMER/MONOMER RATIO DURING THE EXPRESSION OF WT HFABP5

Different LBs: One hypothesis was that using different source of LB for expression. was premade LB broth containing casein digest peptone, sodium chloride, yeast and tris HCl. Also, we made LB broth by mixing the tryptone, sodium chloride and yeast extract. However, in both cases the results were the same.

Changing IPTG and cell line: Changing the amount of IPTG and different cell line also did not enhance dimerization. The IPTG were added to the final concentration of 1mM, 0.5mM and 0.25mM. The difference in these concentrations seems not to have an effect on increasing the dimer FABP5. Changing the Cell line from BL21*(DE3) to BL21(DE3)pLysS did not change the expression level of dimer.

Effect of temperature: We also tried changing the temperature to increase the dimer ratio. We expressed FABP5 at 20°C and 25°C and also higher temperatures. However, the ratio of the dimer was low in both cases (**figure V-24**). Increasing the expression temperature to 37°C does not help dimerization.

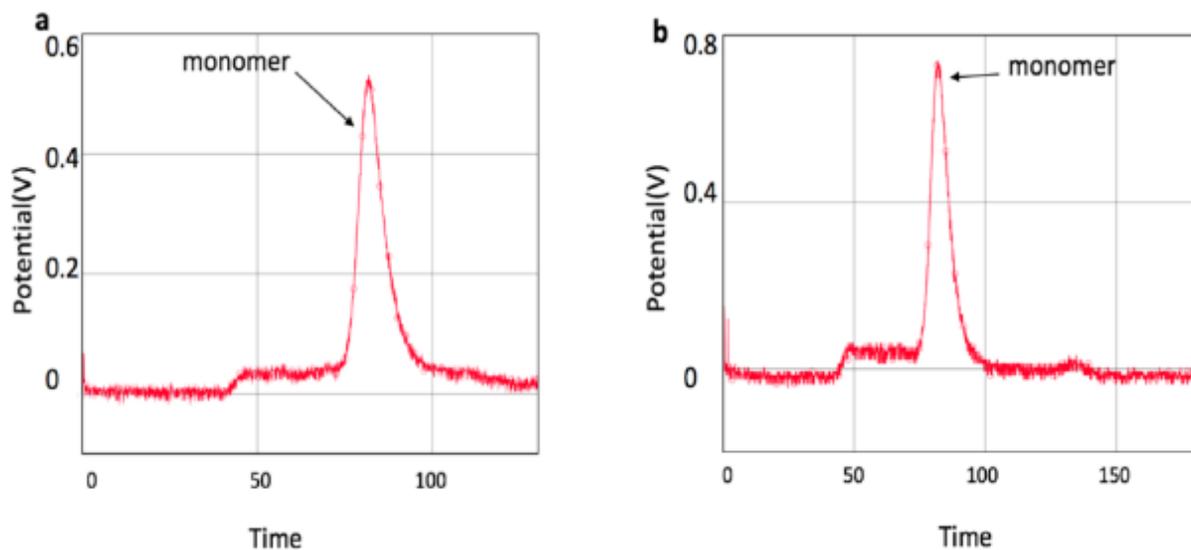


Figure V-24: Chromatogram of SEC column after a) expression of FABP5 in 25°C. b) expression in 20°C.

V-11 HFABP5 AS A NEW RHODOPSIN MIMIC SYSTEM

After elucidation of DSD of hFABP5, we decided to use this protein as a new rhodopsin mimic template. This protein has completely different binding site compare to hCRBP2 and hCRAB2 and makes it valuable for screening fluorescent proteins. As explained before, hCRBP2 crystallized in pH 4 to pH 5, which is below the physiological pH. Structural data for the hFABP5, which crystallized near the physiological pH, might be more accurate for analysis. Also, there are more reports for in vivo studies on hFABP5 compare to hCRBP2 protein, which helps us in our in vivo studies for the designed protein fluorescent tags.

Initial studies have been done to investigate the binding of WT-hFABP5 with retinal. Retinal incubated with the monomer hFABP5 for a couple of hours. Crystallization trials for the formation of crystals of hFABP5 bound to retinal failed.

We used our information on hCRBP_{II} to make hFABP5 as a mimic rhodopsin system that can make a covalent bond with a retinal, Sequence alignment of WT-hFABP5 and WT-hCRBP_{II} is shown in (Figure V-25).

| | | | |
|---------------------|-----|---|-----|
| hCRBP _{II} | 7 | GTWEMESNENFEGYMKALDIDFATR _K IAVRLTQTKVIDQDGDNFKTKTTSTFRNYDVDF | 66 |
| hFABP5 | 16 | G W + ++ F+ YMK L + A RK+ +I DG N KT ST + T | 75 |
| | 67 | VGVEFDEYTKSLDNRHV _K ALV _T WEGDVLV _C VQKGEKENRGWKQWIEGDKLYLELTCGDQV | 126 |
| | 76 | +G +F+E T D R + + + LV Q+ + + + ++ KL +E + | 133 |
| | 127 | LGEKFEETTA--DGRKTQ _T VCNFTDGALVQH _Q EWDGKESTITRKLKDGKLVVECV _M NNVT | 134 |
| | 134 | CRQVFKK 133 C +++++K | 140 |
| | 134 | CTRIYEK 140 | |

Figure V-25: Sequence alignment of WT-hFABP5 and WT-hCRBP_{II}.

Although the sequence identity is around 26%, the overall 3D structures of these two proteins, especially their binding sites, are similar. As discussed before, the Q108K mutation is essential for making the Schiff base with the retinal. By Overlaying the structure of monomer hFABP5 and monomer hCRBP_{II} bound to retinal, we hypothesized that mutating the R109 to lysine might be a good starting point to try for making the Schiff base to retinal (Figure V-26). Absorption spectra does not show any evidence of formation of SB in R:109K:hFABP5 with retinal. Lys60 might be interfering with the binding the retinal inside the binding pocket. Unfortunately, the crystals formed from the R:109K:hFABP5 with retinal did not diffract. More studies should be done to find the best construct for making the SB of retinal in this protein.

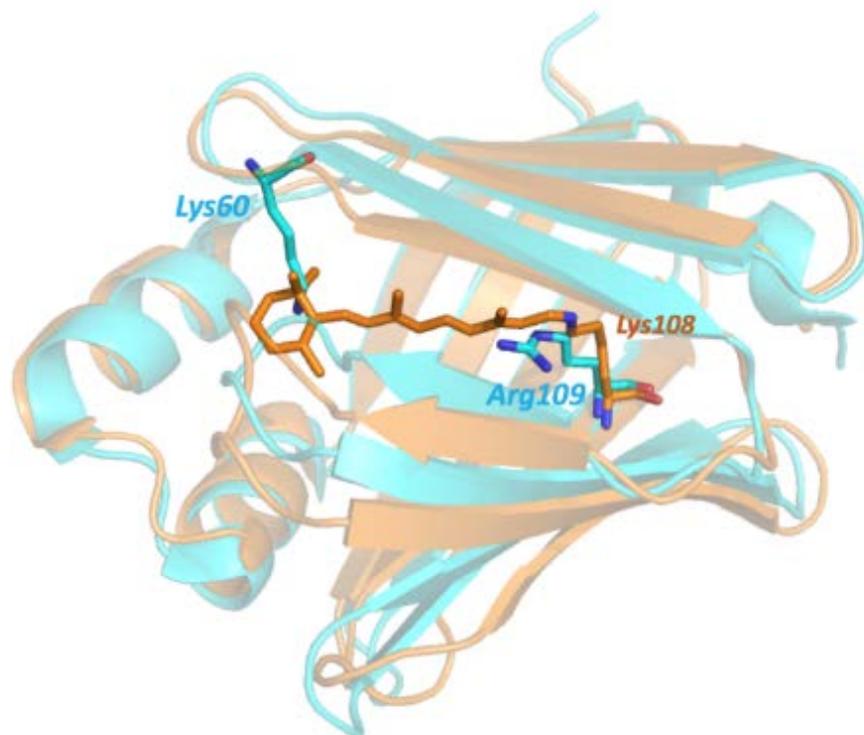


Figure V-26: Overlaid structures of monomer Q108K:K40L:hCRBP II bound to retinal (PDB: 4RUU) (Orange) with WT hFABP5 (Blue). Lys60 might interfere with binding the retinal in hFABP5.

V-12 CONCLUSION

We characterized the Domain swapped dimer for WT-hFABP5 bound to palmitic acid as a natural product during the E. coli expression. The existence of Domain swapping in FABP5 as another member of the iLBP family is another reason that indicates the formation of DSD as a natural kinetic product during the folding process, which may indicate a common folding pathway for these two proteins. Furthermore, the formation of this Domain swapped dimer lead to a conformation change in the portal region of the protein, which may lead to allosteric regulation for FABP5.

Also, since hFABP5 has a similar binding pocket as hCRBP2, and the crystallization pH is around the physiological pH, this protein can be a great target as a new rhodopsin mimic.

V-13 EXPERIMENTAL

Most of experimental part is the same as explained in my master thesis.¹⁴

V-13-1 Site-Directed Mutagenesis

The described FABP5 plasmid was ordered from IDT®. Then Dr. Zahra Assar-Nossoni cloned them into pET28a from Novagen company between BamHI and HindIII cut sites. His tag and thrombin cutting site were designed before the N terminus of the gene. The aforementioned construct was used for mutagenesis by using the Quick Change Site-directed Mutagenesis Kit protocol from Agilent Technologies company.

V-13-2 Amino acid sequence of FABP5:

HHHHHSSGLVPRGSHMGSMGSHMATVQLEGRWRLVDSKGFDEYMKELGVGIALR
KMGAMAKPDCIITCDGKNLTIKTESTVKTTQFCTLGKFEETTADGRKTQTVCNFTDG
ALVQHQEWDGKESTITRKLKDGKLVV ECVMMNNVTCTRIYEKVE

V-13-3 Primers

T64A-FABP5

Forward: 5'-CGTGAAGACGGCGCAGTTTTTCAT-3'

Reverse: 5'-GTCGATTCAGTTTTAATAGTTAAGGTTCTTA-3'

T60P-FABP5

Forward:5'-AACTGAATCGCCCGTGAAGACGA-3'

Reverse: 5'- TTAATAGTTAAGTTCTTACCATCGC-3'

T64E-FABP5

Forward:5'-CGTGAAGACGGAGCAGTTTTTCATG-3'

Reverse:5'-GTCGATTCAGTTTTAATAGTTAAG-3'

T64L-FABP5

Forward:5'-CGTGGAGACGCTGCAGTTTTTCATG- 3'

Reverse: 5'- GTCGATTCAGTTTTAATAGTTAAG- 3'

Q65W-FABP5

Forward: 5'- GAAGACGACGTGGTTTTTCATGCAC-3'

Reverse: 5'-ACGGTCGATTCAGTTTTAATAG -3'

Q65M-FABP5

Forward: 5'-GAAGACGATGTTTTTCATGCAC-3'

Reverse: 5'-ACGGTCGATTCAGTTTTAATAG-3'

V-13-4 Protein Expression and Purification of FABP5

We expressed proteins by using *Escherichia coli* BL21 cells and using the T7 expression system. After expressing the cells in LB media at OD₆₀₀ of 0.4-0.7, the IPTG were added to a final concentration of 0.4 mM. Then, the mixture was incubated at both 20°C and 25°C for 20-24 h and cells were collected by centrifugation at 5000rpm at 4°C for 20 min after that. Cells were lysed using the FABP5 lysate buffer (20mM Tris pH 8.5, 200mM NaCl), by sonication on ice. Then by 20 min centrifugation at 10,000rpm at 4°C, the supernatant was collected and loaded onto a Ni-NTA column from GE healthcare company. After mixing and incubating the protein with the Ni resin, the wash buffers containing lysate buffer and 20 mM imidazole, 30mM imidazole and 50 mM imidazole were added, respectively. The proteins were then eluted with buffer containing 100, 150 and 200 mM imidazole. For removing the lipid from the protein (delipidation), we mixed Lipidex-5000 resin with our protein and the mixture was incubated for 1 hour at 37°C. In case of removing the 6 His tagged from the N terminus, the purified proteins were incubated with thrombin from GE Healthcare Life Sciences at 10 units per milligram of protein at 4°C overnight. Then the cleaved proteins were loaded onto an Ni-NTA columns and the cleaved proteins collected. The samples were concentrated to around 6-8mg/mL and 3mL of the concentrated protein were loaded onto our SEC column which was Superdex S75 16/600 HiLoad column from GE healthcare company and equilibrated with 1CV PBS pH 8.5. The peak fractions were collected and SDS PAGEs were run to check the purity of the protein. Then, purified proteins were concentrated to for crystallization.

V-13-5 Thermal shift assay (TSA)

In order to find the melting temperature of both the hFABP5 monomer and dimer, we expressed and purified the protein. After that, we ran the SEC column for protein, and TSA analysis has done for fractions from the SEC. WT-FABP5 from each fraction at 0.5 mg/mL were mixed 1000:1 with Sypro Orange dye (Sigma Aldrich) and plated. Samples were processed using a BioRad CFX C100 Touch qPCR and run using the FRET assay settings with a heating ramp of 0.3°C/sec cycling from 4°C to 100°C. The analysis was performed using the Bio-Rad CFX manager software (Version 3.1, Bio-Rad). TSA data from the fraction 16-22 of SEC of WT hFABP5 has been shown in **(Figure V-27)**.

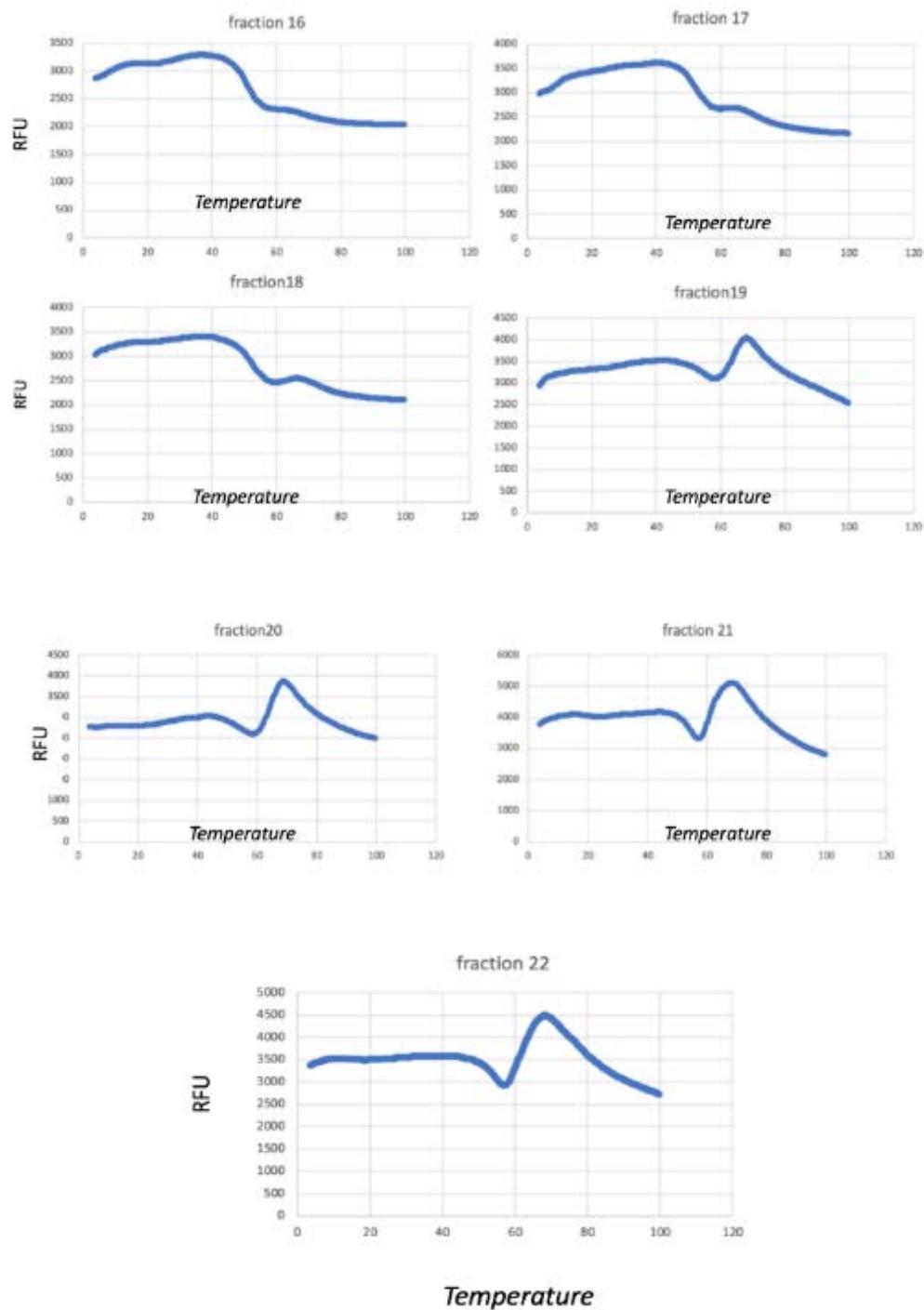


Figure V-27: TSA data for fractions 16-22 of size exclusion chromatography for WT FABP5.

Data demonstrates the gradual switch from 40 °C to 65 °C.

Calculated melting points from the TSA data:

Fraction 16: ~T_m=10C (aggregates), T_m=37C, you'll start to see Monomer peak

Fraction 17: ~T_m=30C, T_m=62C, observing Monomer peak

Fraction 18 : ~T_m=10C, T_m=35C, T_m=64.5C, observing Monomer peak

Fraction 19: ~T_m=35C, T_m=64.5C, observing Monomer peak

Fraction 20: ~ T_m=30C, T_m=62.5C, observing more Monomer peak

Fraction 21: T_m=60.5C, observing more Monomer peak

Fraction 22: T_m=60.5C, observing more Monomer peak

V-13-6 Crystallization and refinement

V-13-6-1 Crystallization of Apo FABP5

The concentrated Apo FABP5 with His tag was concentrated to 6 mg/mL in lysate buffer containing 20 mM Tris, 200 mM NaCl, pH = 8.5. We used hanging drop vapor diffusion method using 1 μ L of protein solution, 1 μ L of crystallization solution in the drop and 1 mL of crystallization solution in the reservoir for crystal grow. This protein crystallized in c. We collected our data at the Advanced Photon Source (APS) (Argonne National Laboratory IL) LS-CAT, (sector 21-ID-D, beamline F and G) using a MAR300 detector and 1.00 \AA wavelength radiation at 100K. Using the HKL2000 software, the diffraction data were indexed, processed and scaled. The structure was solved by molecular replacement using PHASER in PHENIX and FABP5 (PDB: 4AZR) as a search mode. The electron density map was produced by REFMAC5 in the CCP4 package or Phaser-MR in PHENIX. Model rebuilding and water replacement were done using COOT program. The structures were refined using PHENIX program packages.^{15 16 17 18}

V-13-6-2 Crystallization of FABP5 domain swapped dimer bound to palmitic acid (first structure)

The concentrated FABP5 with His tag was concentrated to 7 mg/mL in lysate buffer containing 20 mM Tris, 200 mM NaCl, pH = 8.5. We used hanging drop vapor diffusion method as mentioned in II-4-3-1 for crystal grow. This Holo protein binds to palmitic acid is crystallized in 25% PEG 4000, 100 mM bis-tris pH 8.0. For Data collection and refinement please refer to II- 8-3-1.

V-13-6-3 Crystallization of FABP5 domain swapped dimer bound to palmitic acid (second structure)

The concentrated FABP5 with His tag was treated with lipidex column for 1 h in 37°C. Then the protein concentrated to 6 mg/mL in lysate buffer containing 20 mM Tris, 200 mM NaCl, pH = 8.5. We used mosquito[®] Robot from TTP labtech for crystal grow. The Holo protein binds to palmitic acid is crystallized in %20 PEG 2000, 0.2M Trimethylamine N-oxide dihydrate, 0.1 tris pH= 8.5 which was different condition as before. For Data collection and refinement please refer to the first chapter (**Table V-1, V-2, V-3**).

Table V-1: Crystallographic data of FABP5 (pseudo monomer).

| | WT-hFABP5 |
|---|--------------------------------------|
| Space group | P21 21 21 |
| a (Å) | 40.89 |
| b (Å) | 55.10 |
| c (Å) | 126.36 |
| α (°) | 90 |
| β (°) | 90 |
| δ (°) | 90 |
| Molecules per asymmetric unit | 2 |
| Total reflection | 1524152 |
| Unique reflection | 4421 |
| Completeness (%) | 99.41 (92)^a |
| Average I/σ | 30 |
| | 34.329- |
| Resolution (Å) (last shell) | 1.789 (1.83-1.78)^a |
| Rwork/Rfree (%) | 25/28.2 |
| Root-mean-square deviation from ideal values | |
| Bond length (Å) | 0.009 |
| Bond angle (°) | 1.394 |
| Average B factor | 27.15 |

Table V-2: Crystallographic data of FABP5 domain swapped dimer binds to palmitic acid (first structure).

| | |
|-------------------------|--------------------|
| Space group a (Å) | P6 ₃ 22 |
| a (Å) | 97.999 |
| b (Å) | 97.999 |
| c (Å) | 65.907 |
| α (°) | 90.00 |
| β (°) | 90.00 |
| δ (°) | 120.00 |
| Number of reflections | 6870 |
| after outlier rejection | 6870 |
| Completeness | 99.88% |
| R-work | 0.2312 |
| R-free | 0.3198 |
| Rotamer outliers | 0.85 % |
| C-beta deviations | 0 |
| RMS(bonds) | 0.0081 |
| RMS(angles) | 0.99 |
| MolProbity score | 1.87 |
| Resolution | 2.50 |

Table V-3: Crystallographic data of FABP5 domain swapped dimer binds to palmitic acid.

| | |
|-------------------------|-------------------|
| Space group a (Å) | C222 ₁ |
| a (Å) | 66.265 |
| b (Å) | 114.735 |
| c (Å) | 107.729 |
| α (°) | 90.00 |
| β (°) | 90.00 |
| δ (°) | 90.00 |
| Number of reflections | 21127 |
| after outlier rejection | 21127 |
| Completeness | 99.05% |
| R-work | 0.2753 |
| R-free | 0.3107 |
| Rotamer outliers | 0.0 % |
| C-beta deviations | 0 |
| RMS(bonds) | 0.0078 |
| RMS(angles) | 0.85 |
| MolProbity score | 1.33 |
| Resolution | 2.2 |

REFERENCES

REFERENCES

1. Assar, Z. *et al.* Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates. *Structure* **24**, 1590–1598 (2016).
2. Ghanbarpour, A. *et al.* Engineering the hCRBP II Domain-Swapped Dimer into a New Class of Protein Switches. *J. Am. Chem. Soc.* **141**, 17125–17132 (2019).
3. Liu, Y. & Eisenberg, D. 3D domain swapping: as domains continue to swap. *Protein Sci.* **11**, 1285–1299 (2002).
4. Crystallographic study of FABP5 as an intracellular endocannabinoid transporter. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3940194/>.
5. Armstrong, E. H., Goswami, D., Griffin, P. R., Noy, N. & Ortlund, E. A. Structural basis for ligand regulation of the fatty acid-binding protein 5, peroxisome proliferator-activated receptor β/δ (FABP5-PPAR β/δ) signaling pathway. *J. Biol. Chem.* **289**, 14941–14954 (2014).
6. Haunerland, N. H. & Spener, F. Fatty acid-binding proteins--insights from genetic manipulations. *Prog. Lipid Res.* **43**, 328–349 (2004).
7. Bloch, G. S. K. Unsaturated Fatty Acids in Microorganisms. *Unsaturated Fatty Acids* **237**, 6 (1962).
8. Napoli, J. L. Cellular retinoid binding-proteins, CRBP, CRABP, FABP5: Effects on retinoid metabolism, function and related diseases. *Pharmacol. Ther.* **173**, 19–33 (2017).
9. Identification of intracellular carriers for the endocannabinoid anandamide | PNAS. <https://www.pnas.org/content/106/15/6375>.
10. Ogawa, E. *et al.* Epidermal FABP (FABP5) Regulates Keratinocyte Differentiation by 13(S)-HODE-Mediated Activation of the NF- κ B Signaling Pathway. *Journal of Investigative Dermatology* **131**, 604–612 (2011).
11. Hohoff, C., Borchers, T., Rüstow, B., Spener, F. & van Tilbeurgh, H. Expression, purification, and crystal structure determination of recombinant human epidermal-type fatty acid binding protein. *Biochemistry* **38**, 12229–12239 (1999).
12. Mc, L. *et al.* Evaluation of Fluorescence-Based Thermal Shift Assays for Hit Identification in Drug Discovery. *Analytical biochemistry* vol. 332 <https://pubmed.ncbi.nlm.nih.gov/15301960/> (2004).
13. Kohlstaedt, M., von der Hocht, I., Hilbers, F., Thielmann, Y. & Michel, H. Development of a Thermofluor assay for stability determination of membrane proteins using the Na(+)/H(+)

- antiporter NhaA and cytochrome c oxidase. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1112–1122 (2015).
14. Ehyaei, N. Domain Swapping in the ILBP Family. (Michigan State University, 2018).
 15. Abergel, C. Molecular replacement: tricks and treats. *Acta Crystallogr D Biol Crystallogr* **69**, 2167–2173 (2013).
 16. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25 (2010).
 17. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D* **66**, 213–221 (2010).
 18. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).