DEVELOPMENT OF PHOTOSWITCHABLE RHODOPSIN MIMICS: SPECTROSCOPY, STRUCTURAL STUDIES AND WAVELENGTH REGULATION STUDIES OF THESE SYSTEMS

By

Meisam Nosrati

A DISSERTATION

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

Chemistry – Doctor of Philosophy

2015

ABSTRACT

DEVELOPMENT OF PHOTOSWITCHABLE RHODOPSIN MIMICS: SPECTROSCOPY, STRUCTURAL STUDIES AND WAVELENGTH REGULATION STUDIES OF THESE SYSTEMS

By

Meisam Nosrati

Retinal photo pigments have been an essential part of life on the planet from bacteria to mammals. Photo isomerization of the retinal is the heart of the function of all of these pigments. Vision and circadian rhythm in higher organisms and energy production and phototaxis in bacteria are all mediated by these light absorbing proteins. Retinal isomerization has never been recapitulated in an artificial retinal protein pigment before. Reengineering of human Cellular Retinoic Acid Binding ProteinII (hCRABPII) into a photo switchable retinal pigment provided a light absorbing pigment with unique properties. The reengineered protein can isomerize thermodynamically and photochemically from a 15-*cis* retinal to 15-*trans* retinal and vice versa. The isomerization is accompanied by a 3-5 unit of pKa change from above physiological pH to lower similar to other retinal pigments. Isomerization of retinal is quantitative both in solution and more importantly in the crystalline form. High-resolution X-ray crystal structures provide a crystal clear image for the isomerization of the retinal in this system. The quantitative isomerization of the retinal in the crystalline form can be a valuable tool in characterizing the intermediates of retinal isomerization. This protein in combination with

fluorescent proteins can be used as a fluorescent quencher to develop a Reversibly Switchable Fluorescent Protein (RSFP).

Wavelength regulation of rhodopsins has been a long standing question with no solid answer that has been debated for five decades. Point charges from charged residues and the dipole moment from the polar residues especially the hydroxyl containing side chains and the conformational changes in the retinal chromophore have been suggested as the main theories for wavelength regulation. However, the effect of the polar water molecules in wavelength regulation of retinal pigments has been less investigated. Using a reengineered human Cellular Retinol Binding ProteinII (hCRBPII), we were able to systematically demonstrate a coherent effect of the polar water molecules in the wavelength regulation of hCRBPII retinal protein complexes, which can be attributed to other known retinal light absorbing pigments as well. Mutating the residues interacting with these water molecules, which are located close to the retinal polyene from the beginning to the end of it, regulate the absorption of the protein.

ACKNOWLEDGEMENTS

I am grateful and indebted to all the people that had a positive impact on my life and changed it for good and brought hope into it. I would not have been able to progress and move forward without their help and support and be where I am. They are:

My father for always caring and checking on my education and providing a peaceful environment to focus on my studies and being the most unmaterialistic person that I have ever met in my entire life and I am sure I will never find anyone like him being this much generous and giving, My mother for being the simplest woman in the world with least imaginable expectations from life who I always felt comfortable to talk about literally everything going on in my life, My brothers Emran and Ehsan and my sister Fatemeh, my sister-in-law Nahid, brother-in-law Mohammad Reza and my niece Bahar who has brought a lot of joy to our lives. They have been all always there for me and gave me their unconditional support at all times.

I have had good teachers and professors that made me interested in science and I enjoyed their mentorship. My English language teachers, in guidance school Mr.Moalla and in high school, Mr.Memarzadeh and Mr.Khalilnejad, my high school chemistry teacher, Mr.Taheri, and my god blessed biology teacher, Mr.Hasani, were all the people that I was blessed to meet and be their student in my hometown of Damghan Iran. In University of Tehran during my bachelors and masters Dr.Adib played the role in making me interested in Organic Chemistry which still is the most appealing part of the chemistry for myself afterwards till now.

Coming to the United States and adopting a new life probably was the most challenging transition, and life far away from home and family especially in the first two years was a big challenge. I am thankful to Mehrdad Mahdavi, Faramarz Vafaee, Reza Sheikhi and Roozbeh Yousefi that I have made long lasting friendships with. I have always turned to Mehrdad and Faramarz in hard times in here, and they have been always helpful to me and put time for me. They have not been less than brothers to me, and I am lucky to have them in my life.

I should especially thank professor Borhan for an initial belief in me and choosing me over a wide number of applicants to come to Michigan State University for PhD studies and his scientific mentorship during my PhD, his meticulous point of views in science and perfectionism. I have enjoyed being close to the same type of scientists in his lab like Dr.Chrysoula Vasileiou, Dr.Tetyana Berbasova, Dr.Wenjing Wang and Dr.Roozbeh Yousefi. Our collaborative work especially with Dr.Berbasova and then Dr.Yapici helped us to build an interactive and fruitful research environment, which I hope to find or build in a new environment. Also I feel now that it was unique to that time. I should give a special thank to Dr.Vasileiou for her extensive help to me and everyone else in the lab. She always had time for everyone and the door of her office was always open to the people of the lab.

The most influential person however, was my PhD academic advisor prof.James H. Geiger. He gives a lot of freedom in his research lab and accepts ideas generously that is essential for making an independent thinker.

Also, we started with a regular student advisor relationship, but overtime I found him the best person to get not only advise for my research, but for my

personal matters as well where I felt the lack of experience to approach them. I admire his stability and great deal of experience in life, his humble, social, simple and happy personality with an anger that subsides quickly and follows sympathy. Jim has been always the person that made me feel confident and inspired me to believe in myself.

I also enjoyed being labmates with Dr.Susan Strom, Dr.Remie Fawaz, Dr.Rafida Nossoni, Dr.Camille Watson, Lindsey, Janice, Zahra, Hadi, Hasan and Alireza in prof.Geiger lab. Also, we were a small lab all the time, but I found it easy to find peace, focus and relax.

I am thankful to chemistry department staff at Michigan State University that helped to get my research, the research labs and equipment going.

I had the privilege to be one of the very few Iranians with the permission to use the Advance Photon Source of Argonne National Lab at Chicago on-site. I made a lot of trips to there with my labmates and collected data with the help of super friendly staff of the LS-CAT section 21 at Advanced Photon Source.

I am also grateful to anyone else who has helped me during this time and my mind does not help me to remember them and put their name here as person or part of a group.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO SYMBOLS AND ABBREVIATIONS

Å	Angstrom
σ	Sigma
PDB ID	Protein Data Bank Identifier
GPCR	G-Protein coupled receptor
RET	Retinal
TM	*trans*-membrane
PDE	phosphodiesterase
GTP	Guanosine triphosphate
GDP	Guanosine diphosphate
GMP	Guanosine monophosphate
c-GMP	cyclic-Guanosine mono phosphate
BR	Bacteriorhodopsin
SRII	Sensory RhodopsinII
HR	Halorhodopsin
ChR	Channelrhodopsin
AR	Archaerhodopsin
RSFP	Reversibly Switchable Fluorescent Protein
NMR	Nuclear Magnetic Resonance
PSB	Protonated Schiff Base
iLBP	intracellular Lipid Binding Proteins
hCRBPII	human Cellular Retinol Binding ProteinII
hCRABPII	human Cellular Retinoic Acid Binding ProteinII
RAR	Retinoic Acid Receptors
SB	Schiff Base

UV	Ultra Violet
WT	Wild Type
IPTG	Isopropyl β-D-1-thiogalactopyranoside
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
Vis	Visible
ε	Extinction coefficient
QY	Quantum Yield
λ_{max}	Maximum wavelength
λ_{ex}	Excitation wavelength
λ_{em}	Emission Wavelength
PCL	Photo chromic ligand
PTL	Photo switchable tethered ligand
GABA	Gama amino butyric acid
DBD	DNA binding domain
LBD	Ligand binding domain
PYP	Photoactivable Yellow Protein
K_d	Dissociation constant
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
KDa	Kilo Dalton
E. Coli	Escherichia coli
s	second
min	minute
h	hour
mm	millimeter

cm	centimeter
M	molar
mM	Milimolar
µM	Micromolar
nM	Nanomolar
mol	mole
mmol	millimole
mg	milligram
ml	milliliter
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
rpm	rotation per minute
ppm	parts per million
°C	degrees of centigrade
K	degrees of kelvin
pH	Logarithmic scale of hydrogen ion activity
pKa	Logarithmic constant of acid dissociation
HCl	Hydrochloric acid
NaOH	Sodium hydroxide
Cod	Cyclooctadiene
NaH	Sodium hydride
DMF	Dimethylformamide
dtbpy	4,4′-Di-tert-butyl-2,2′-dipyridyl
B_2Pin_2	Bis(pinacolato)diboron
TMOF	Trimethylorthoformate

Sphos	2-Dicyclohexylphosphino-2′,6′-dimethoxybiphenyl
TLC	Thin Layer Chromatography
Equiv	equivalent
RMSD	root mean square deviation
Ala, A	Alanine
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartate
Cys, C	Cysteine
Gln, Q	Glutamine
Glu, E	Glutamic acid
His, H	Histidine
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: Photoswitchable proteins: Biological Significance and Applications

I.1 Retinal Based Photoswitchable Proteins

Vision is a multi step process, which starts with focusing of the light on a small spot on retina curtain in the back of the eye called macula. Absorption of photons by visual pigments in macula causes a change in the membrane potential of the cells to develop a neuronal signal to the brain for analyzing the data. These events translate into what we know as the ability for vision (**Figure I-1**) (1-3).

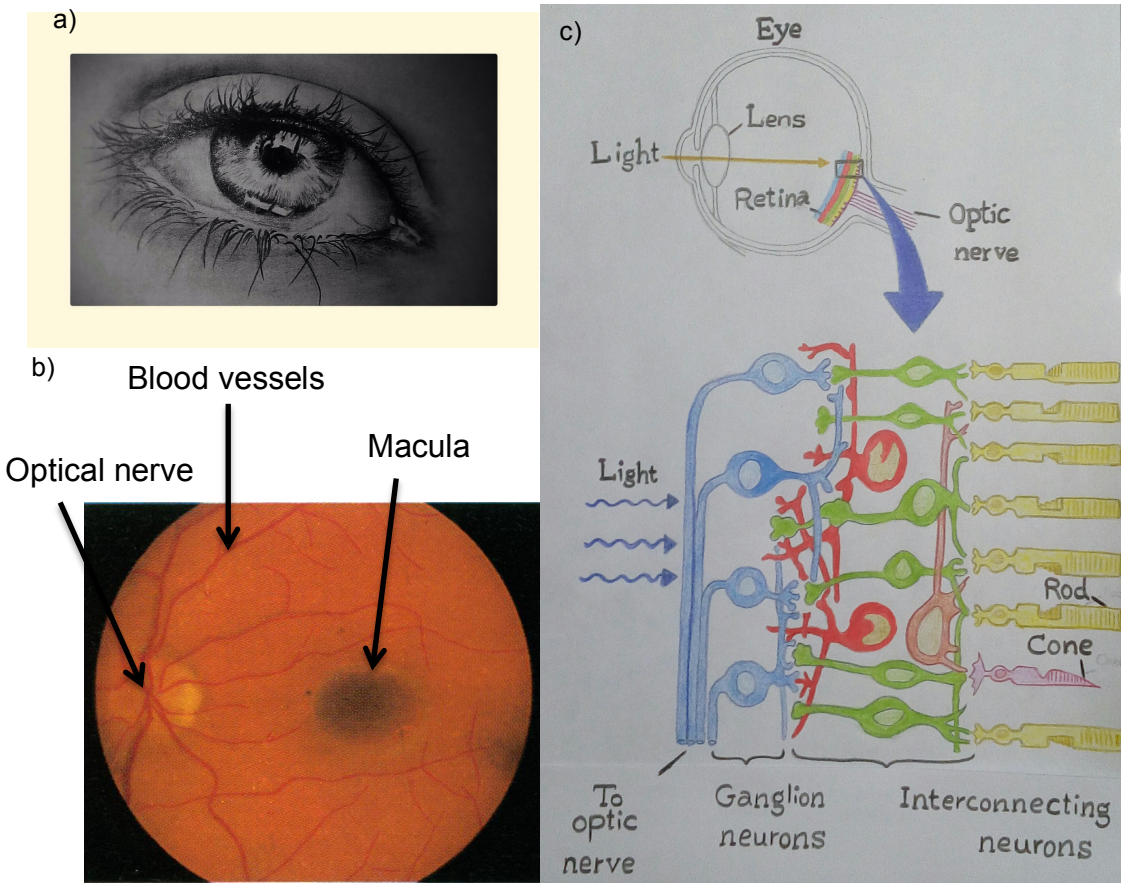


Figure I-1) a) Eye the gate way to vision b) A fundus photo of the Left eye showing the optic nerve, blood vessels an macula c) Different compartments of the eye and the cellular structure of the macula.

Macula is a small circular portion, with a 5mm diameter near the center of retina, where the first step of the visual process happens. Macula is loaded with specific type of neural visual cells called rod and cones. Rod cells are more abundant and sensitive than cone cells specialized for dark vision while cone cells are the less sensitive pigments that are specialized for color vision (**Figure I-2**) (4).

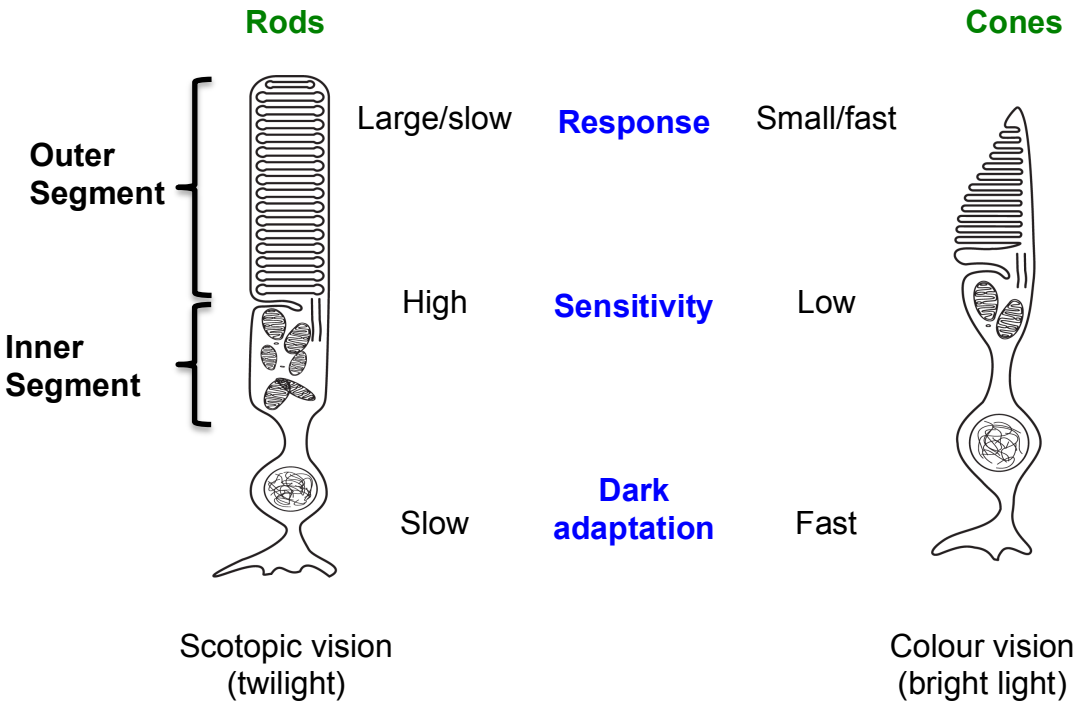


Figure I-2) a) The structure of a Rod cell. b) Comparison of the characteristics of Rod and Cone cells.

Both rod and cone cells are subdivided into two portions of outer and inner segment based on their shape (**Figure I-2**). Outer segment is mainly an extensively folded cellular membrane loaded with a trans membrane protein called Rhodopsin. There are four different types of Rhodopsins, which includes rod, blue, green and red rhodopsins. Rods are the pigment in Rod

cells. Blue, green and red rhodopsins are the pigments in the three different types of the cone cells (**Figure I-3**) (5, 6).

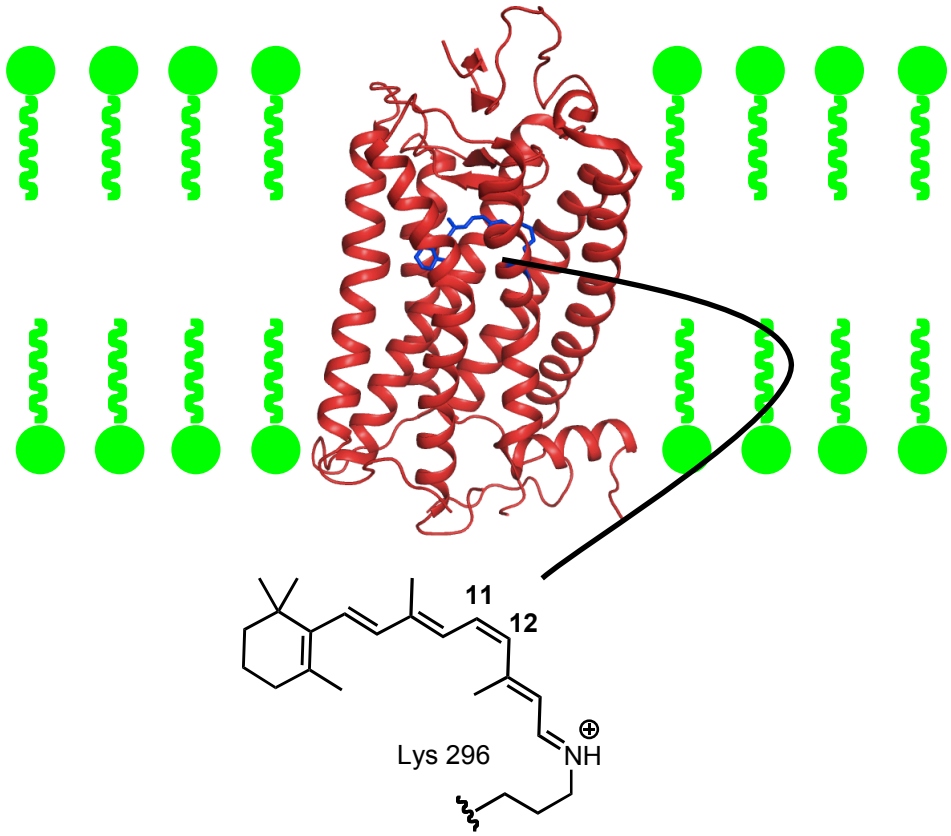


Figure I-3) The structure of a rod cell, rhodopsin and 11-*cis* retinal.

The ability of rhodopsin photo pigments to absorb light comes from the retinal chromophore inside the binding pocket of the protein. Retinal aldehyde makes a Protonated Schiff Base (PSB) with Lys 296. The delocalization of the charge along the polyene contributes to the absorption of the protein in the visible region (**Figure I-4**) (7).

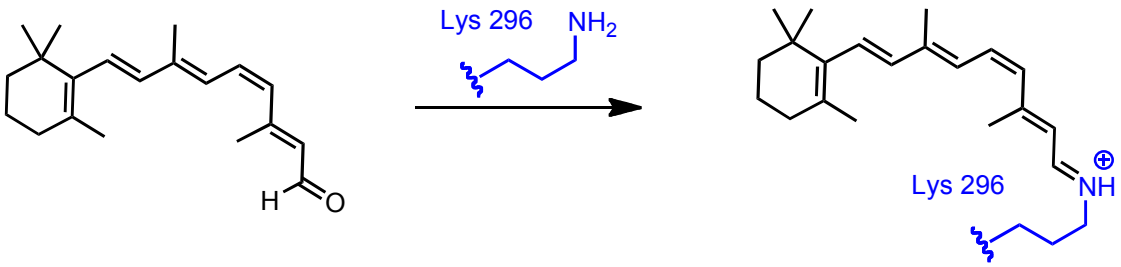


Figure I-4) Formation of a retinal iminium PSB is essential to light absorbtion of rhodopsin.

The wavelength of light absorbed by Blue, Rod, Green and Red rhodopsins are respectively 425, 500, 530 and 560nm for humans. As a photon of light hits the rhodopsin pigment, it triggers the first step in the vision process which is the photo isomerization of the 11-*cis* retinal chromophore to all-*trans* inside the protein (**Figure I-5**) (6). It should be noted that the quantum yield of the retinal isomerization in Rhodopsin is 65 percent (8).

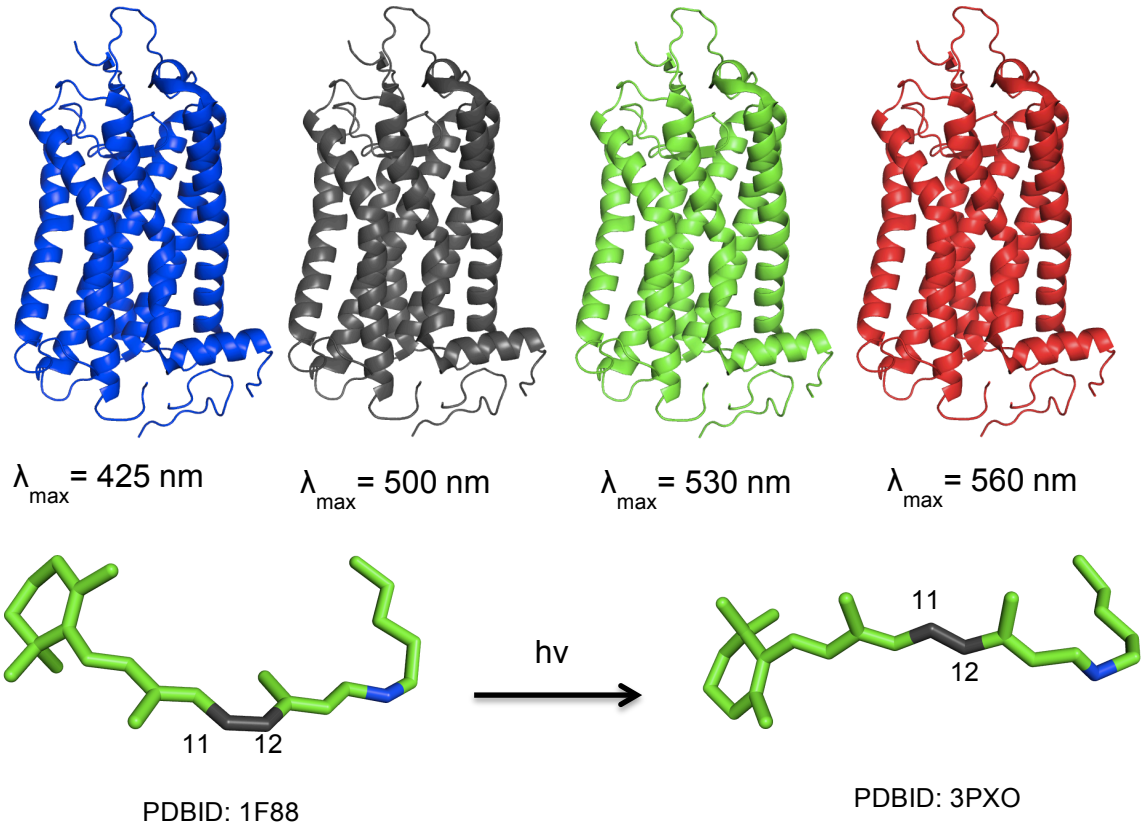


Figure I-5) Different rhodopsins absorb a photon of light with the λ_{max} highlighted to isomerize from 11-*cis* to all-*trans* retinal.

I.2 Photo cascade of Rhodopsin

Different studies have shown that after the light irradiation, Rhodopsin goes through the formation of multiple short-lived photo intermediates. All of these photo intermediates have different retinal geometries, which are accompanied

by conformational changes in the Helixes 3,5 and 6 of opsin protein to facilitate its signaling (7) (**Figure I-6**).

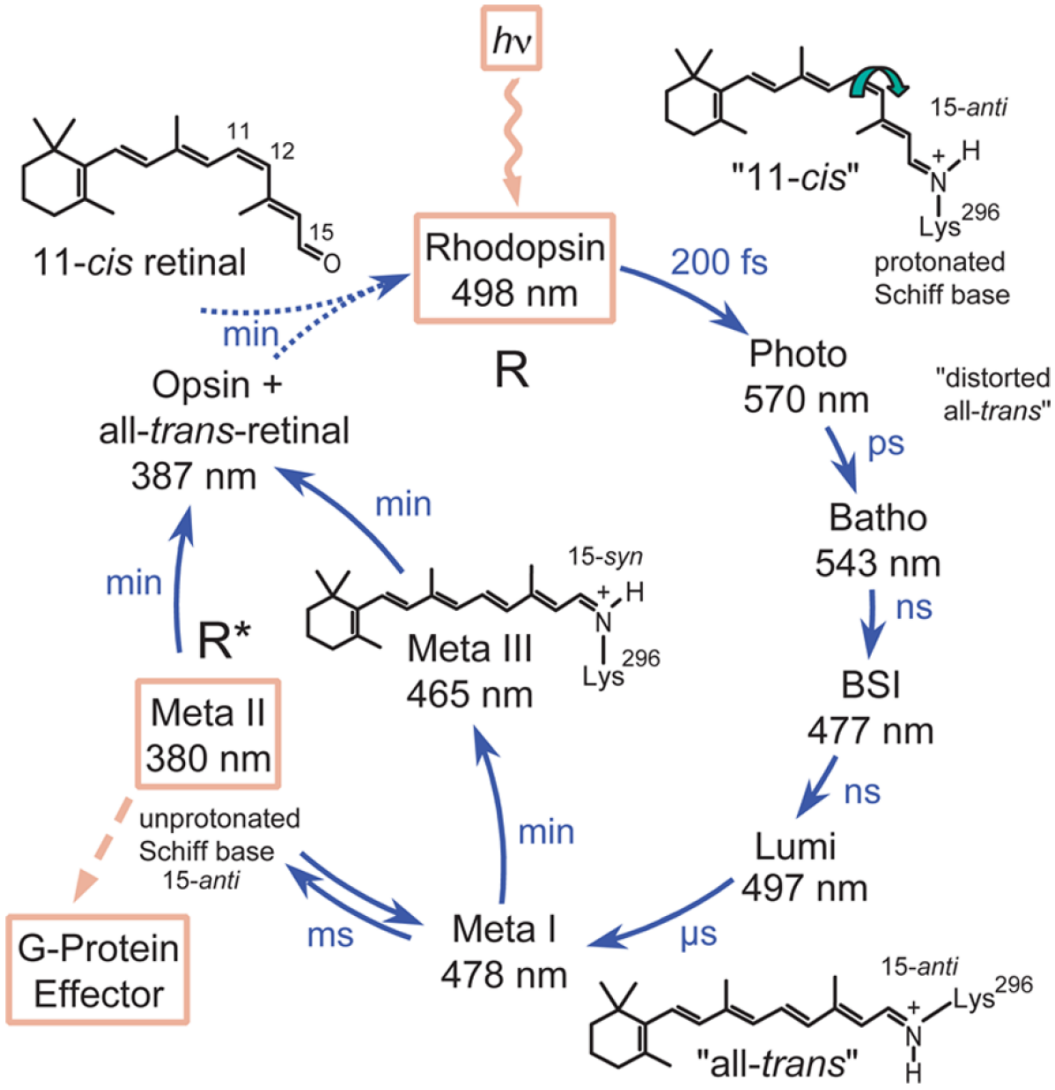


Figure I-6) Photocycle of the Rhodopsin.

The first photo intermediate is photo rhodopsin, which is red shifted with a distorted all-*trans* retinal chromophore. The decay of the Photo Rhodopsin produces Batho in nano seconds followed by the formation of Lumi. The photo cycle proceeds with the formation of MetaI and MetaII (9). MetaII is the photo intermediate in which the retinal PSB deprotonates and the Rhodopsin blue

shifts significantly to where the Retinal-SB absorbs (380 nm). MetaII is the photo intermediate that can activate G-protein and start the initial events that leads to the signaling of G-protein (10). MetaII can take two different pathways for decay. One is the generation of free retinal and opsin and the other is the isomerization to MetaIII. MetaIII has been found to have a 15-*cis* retinal conformation. MetaIII is slower than MetaII in its decay to free retinal and Opsin and it has been found that the rate of the decay can change from hours at basic pHs to minutes in acidic pHs. MetaIII is a partial agonist compared to MetaII, which is a fully active agonist. It has been suggested that the formation of MetaIII holds back the release of the free retinal and its accumulation that can be toxic for the cell (11-14).

The overlay of these two structures shows that the trans membrane α-helices can translocate significantly as the retinal expands from 11-*cis* to all-*trans* conformation (**Figure I-7a**).

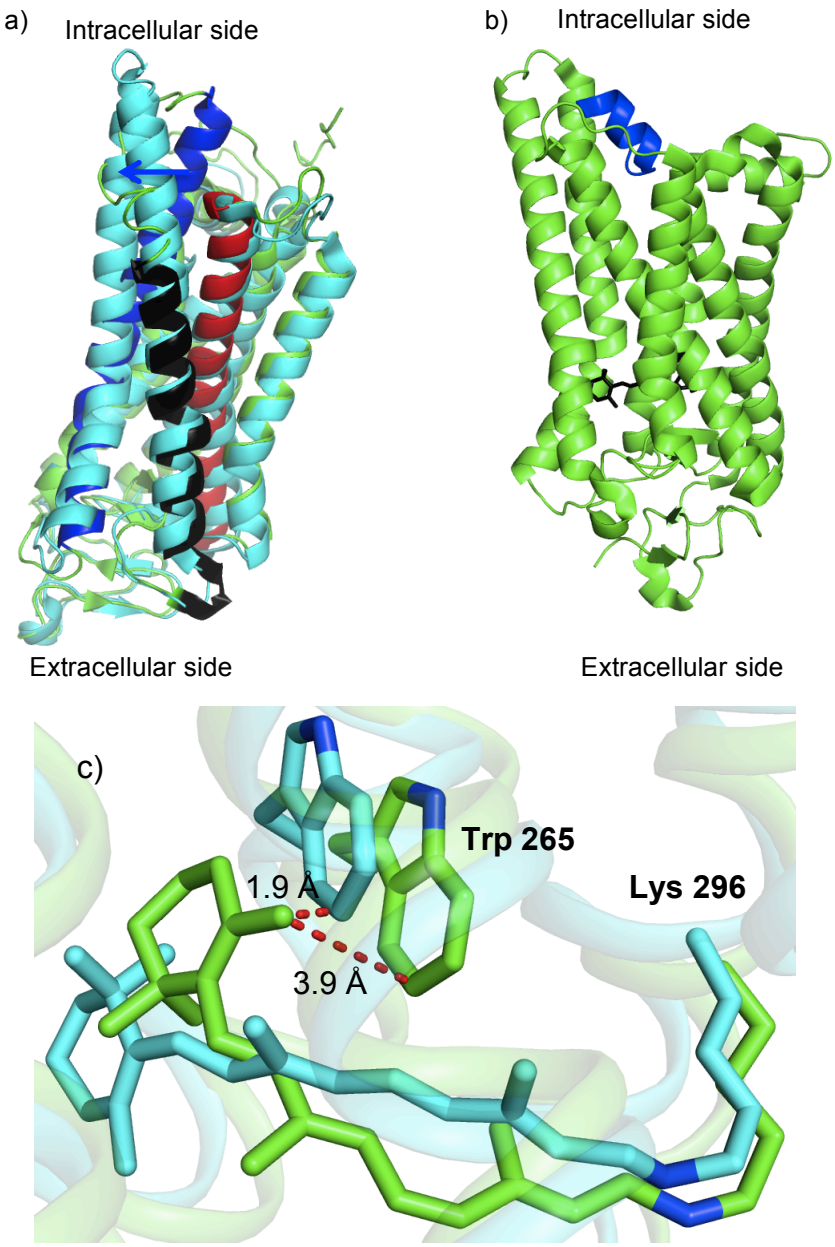


Figure I-7) a) Structural differences between the dark state (Green) and excited state (Cyan) of Rhodopsin. Red = TM3, Black = TM5 and Blue = TM6 in the ground state. b) C-terminus peptide of the Gα-subunit bound to MetaII-Rhodopsin. c) Trp265 positioning on TM6 in the ground state and excited state as retinal isomerizes.

Rhodopsins belong to a family of seven α-helical trans membrane proteins called G-Protein Coupled Receptors (GPCRs). Trans membrane proteins are hard to crystallize as their hydrophobic portions that help them to incorporate into the membrane makes it problematic for protein purification and

crystallization in aqueous hydrophilic environment. However, the improvement of the protein expression and crystallization techniques in the last decade has facilitated the crystallization of these proteins, but still they include a small number of all of the proteins crystallized so far. Rhodopsin was the first GPCR protein to be crystallized in 2000 by Paleczewski. The crystal structure shows the 11-*cis* retinal inside the binding pocket of the protein in the dark state (Ground state) (15). It took another 11 years to crystallize this protein with all-*trans* retinal (Excited state) in the MetaII state, which is one of the final photoproducts in the visual cascade of the rhodopsin (10).

These conformational changes provide a groove for the α-fragment of the G-protein to interact with the rhodopsin in order to carry on a signal, which in the next steps will translate into a membrane potential change to become a neuronal signal into the brain for analyzing the stimuli received from the outside world as colors (**Figure I-7b**). The most noticeable movement of the α-helices is related to TM6. TM6 contains a conserved residue Trp265, which changes position as the chromophore goes from 11-cis to all-trans (Figure 7c). The importance of TM6 movements in rhodopsin signaling was discovered by photo affinity labeling studies in 2000 as well (9).

The crystal structures of the Batho (16) and Lumi (17) were also solved in 2006 and the overlay structure of these proteins shows a twisted all-*trans* retinal in the Batho that is relaxed by the dislocation of the ionone ring in Lumi. Also the conformational changes of the Helixes 3,5 and 6 were reported in the Batho and Lumi structure but the large conformational changes, especially that of Helix 6 waited to be more clarified by MetaII structure. Compared to Lumi structure in MetaII both the ionone ring and the polyene tail

are translated along the retinal chromophore axis, giving it a quite different geometry compared to the ground state, batho and Lumi (**Figure I-8**).

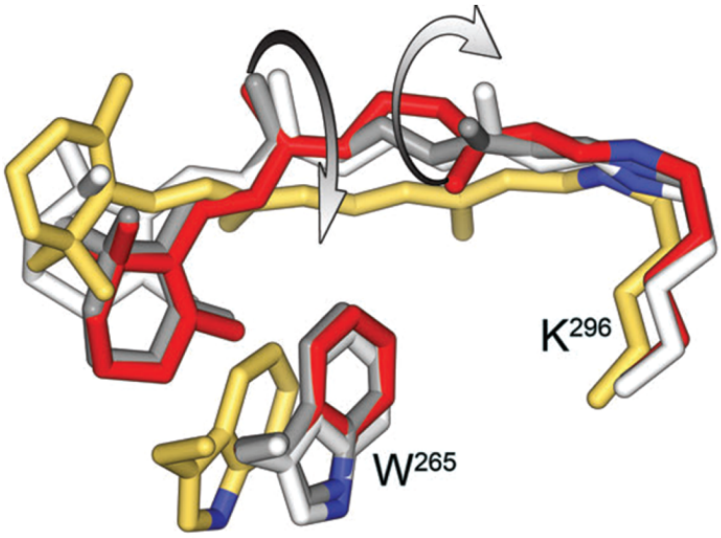


Figure I-8) The retinal conformation in Rhodopsin: Ground state (Red - PDB ID: 1U19), Batho (Gray – PDB ID: 2G87), Lumi (White – PDB ID: 2HPY) and MetaII (Yellow – PDB ID: 3PXO). The arrows show the rotation of the retinal from Lumi to MetaII.

It should be noted here that besides vision, circadian rythme, which is the body clock, is also regulated by retinal photo isomerization in ganglion cells of the retina (18-20).

I.3 Mechanism of Rhodopsin Signaling

Rhodopsin signaling is the cooperation of multiple proteins in a series of biological reactions in a matter of milliseconds. After binding of the G-protein to Rhodopsin and opsin proteins, it exchanges a GDP molecule for a GTP. Binding of GTP to G-protein dissociates the α-subunit of the protein from the β and γ subunits. The dissociated α-subunit binds to the inhibitory subunit of the cGMP phosphodiesterase (PDE) and the activated PDE hydrolyzes the cGMP to 5'-GMP. This results in the closure of the cGMP-gated calcium ion

channels and hyperpolarization of the rod and cone cells that will end up to the neuronal signal. A rhodopsin kinase protein phosphorylates the Theronine and Serine residues in the C-terminus of Rhodopsin. The phosphorylation of the Rhodopsin facilitates the binding of another protein called arrestin that prevents further interaction of the rhodopsin with G-protein.

Each Rhodopsin can activate 500 G-proteins and each PDE molecule can hydrolyze 4200 molecules of cGMP per second and the fact that multiple cGMP molecules are required to keep one cGMP gated ion channel open makes the signaling very effective (21, 22) (**Figure I-9**).

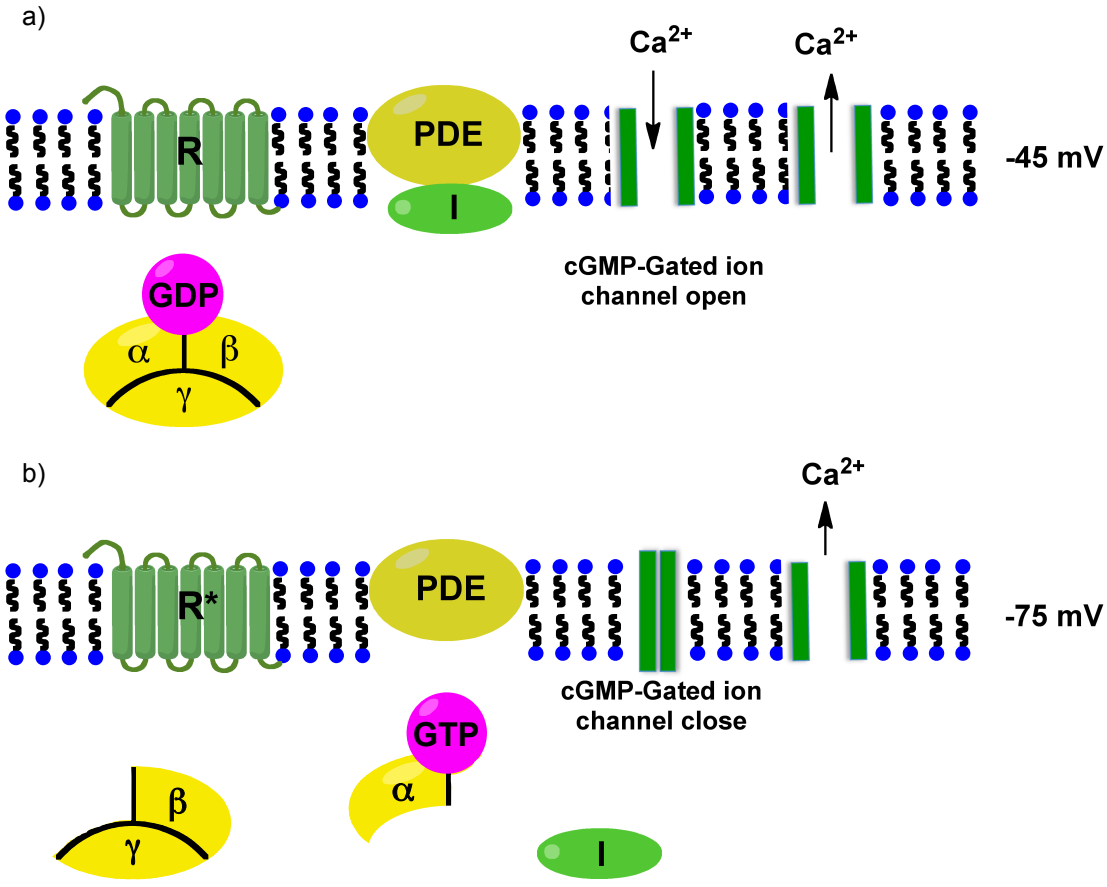


Figure I-9) Rhodopsin signaling.

I.4 Photoisomerization in Bacterial Rhodopsins

Rhodopsins in bacteria play a variety of different functions, which includes energy production by pumping H^+ outside the cells (Bacteriorhodopsin), phototaxis by transducer binding signaling (Sensory Rhodopsins), photo taxis by Na^+, K^+ and other cations pumping across the membrane (Channel Rhodopsins) and Cl^- pumping across the membrane (Halorhodopsin) (7, 23). Bacterial rhodopsins in the ground state bind to all-*trans* retinal (24) and upon absorption of light they isomerize to the 13-*cis* form (**Figure I-10**) (25). There have been a lot of studies to elucidate the photo intermediates of the isomerization process (26-30), but a lot of the structures are subject to X-ray damage during data collection on the crystals (31).

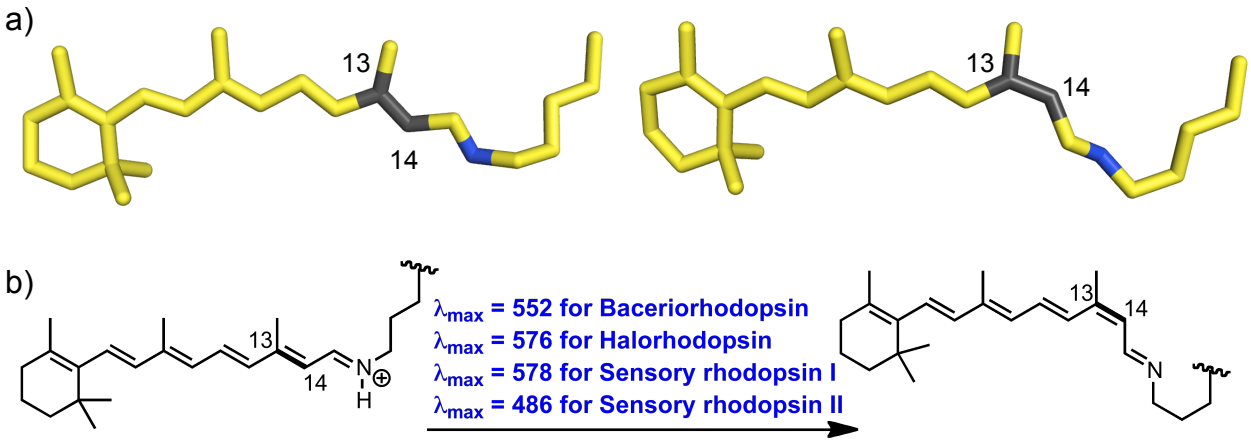


Figure I-10) a) Bacteriorhodopsin in the ground state of all-*trans* retinal (Left - PDB ID: 1C3W) and excited state of 13-*cis* retinal (Right - PDB ID: 1IXF). b) Chemical diagram of the isomerization of the bacterial rhodopsins with the λ_{max} of each indicated.

Using the photo isomerization of Channel Rhodopsin-2 (ChR-2), which is a Na^+ and K^+ pump and halo rhodopsin which is a Cl^- pump in neurons to trigger neuronal signaling in the brain is called Optogenetics. Optogenetics is a very promising field in neurobiology that helps to understand the particular function of each part of the brain. The advantage of using light in controlling the neural

activity is in the temporal and spatial firing of neurons. Specific signals will be produced right where the brain is irradiated with light and as the light source goes off the signal will be terminated. Optogenetics has been already used to understand fear learning, Cocaine dependence, Prefrontal cortex responses, restoration of vision and fibrillation (**Figure I-11**) (32, 33).

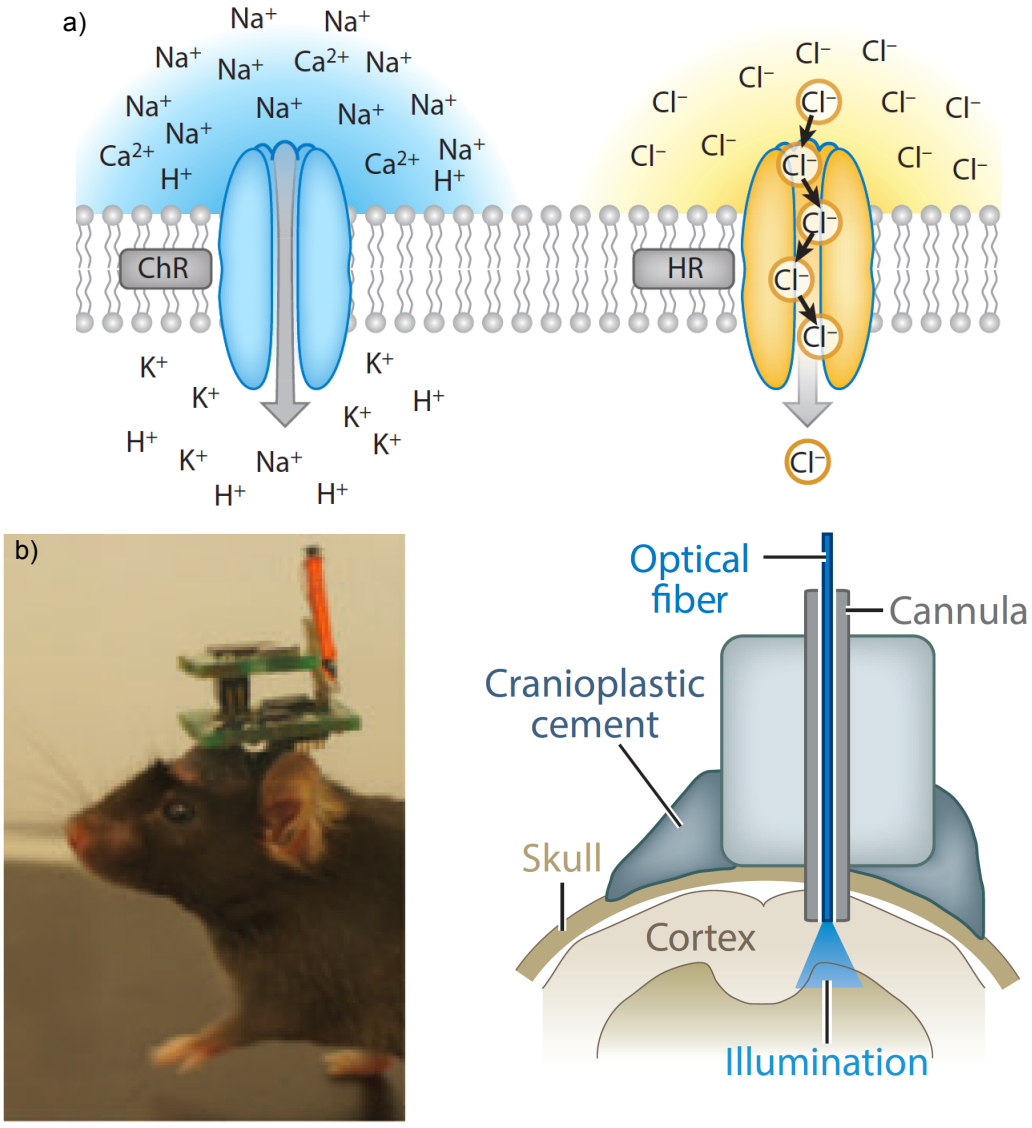


Figure I-11) a) Incorporation of Channel Rhodopsin (ChR) or Halo Rhodopsin (HR) can confer the sensitivity of light to neurons. b) The application of optogenetic tools in mice helps to understand the function of different parts of the brain by spatial and temporal activation of neurons.

I.5 Photocycle of Bacteriorhodopsin

After a photon of light hits the bacterial rhodopsins, the isomerization of all-*trans* retinal to 13-*cis* retinal goes through multiple intermediates (7). The isomerization of the bacteriorhodopsin (BR) is the one that has been studied extensively in the last two decades by different techniques including FTIR, resonance Raman, NMR, theoretical calculations and X-ray crystallography (7, 34, 35) (**Figure I-12**).

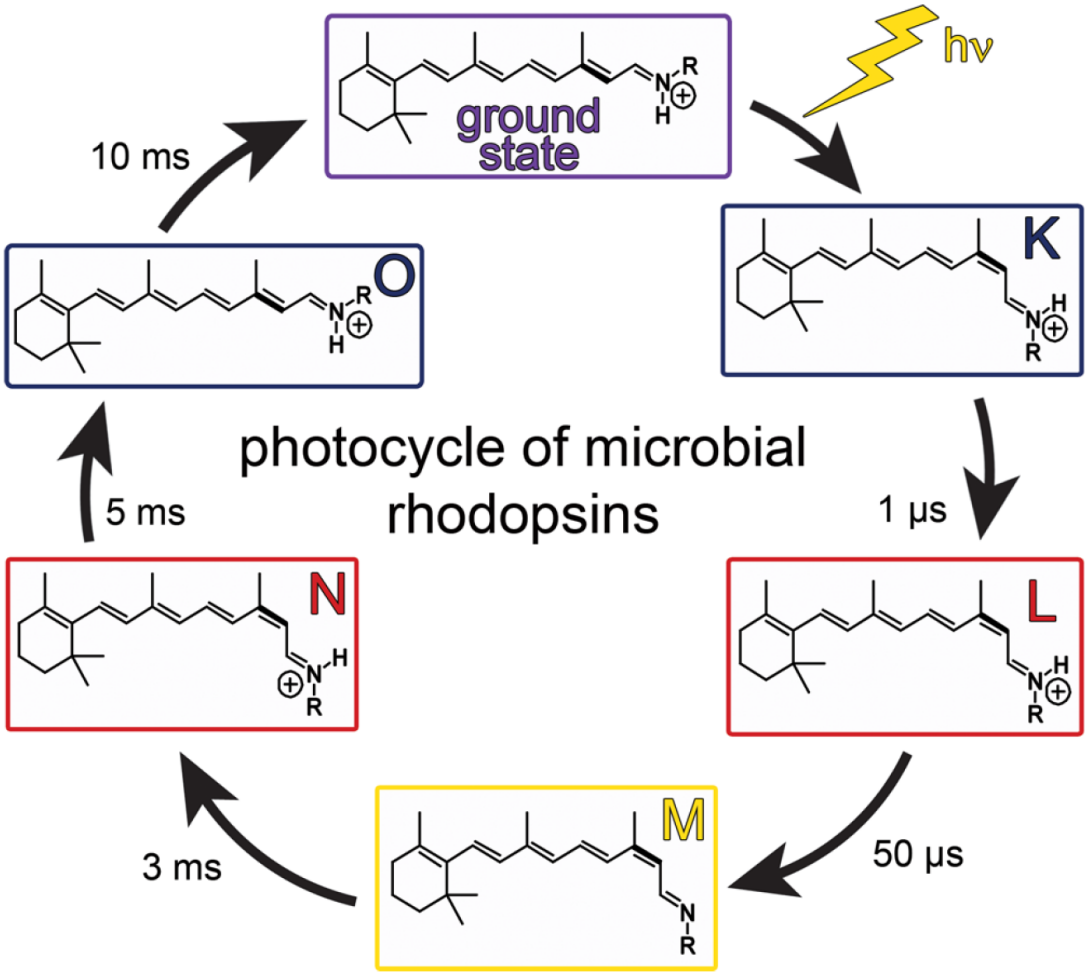


Figure I-12) Photocycle of the Microbial rhodopsins.

K and O are the red shifted intermediates and L, M and N are all blue-shifted intermediates. The deprotonation happens in the M state and the thermal

reisomerization back to all-*trans* retinal happens in the O state. It is clear from figure 9 that the isomeric state of the L and N intermediates is the same while a proton must be released from the intracellular space to the extracellular space. Multiple studies showed that the M state is subdivided into two M1 and M2 states, in which the M1 delivers the proton to the extracellular environment, and M2 faces back again toward the intracellular space. Therefore, the switch happens between M1/M2 states. The protonation and deprotonation of the Retinal PSB happens via Asp96 in the cytosolic side and Asp85 toward the extracellular side. The deprotonation of the Asp96 in the N state and protonation of the Asp85 in the M, N and O state has been confirmed in different studies (7).

Bacteriorhodopsin has seven water molecules toward the extracellular side of the protein and only two water molecules toward the cytoplasmic side. This uneven distribution of water molecules is due to the fast kinetics of proton release of bacteriorhodopsin to the extracellular environment, while the reprotonation of the protein requires conformational changes of the retinal chromophore and Helix F of the protein to completely segregate the two sides of the protein from each other for efficient proton pumping (24, 36).

Crystallographic studies of the bacteriorhodopsin photo intermediates mainly includes the irradiation of the crystals of this protein by corresponding wavelength of light that accumulates the related intermediates and point mutations that elongates their lifetime to help to trap them. Also, the crystallographic data suggests that it is the M state that the switch from cytosolic to extracellular environment happens in (28, 29), but the structural models presented for the L intermediate have been controversial (27, 30) and

X-ray radiation damages have been found as a contributing factor for these discrepancies (25, 31).

One of the interesting characteristics of the bacteriorhodopsin photo isomerization found before the crystal structures was the mixture of 13-*cis*-15-*cis* retinal along with 13-*trans*-15-*trans* retinal in its dark isomerization (37-39). In the crystal structure of the O-intermediate it was found that the crystals that are shined with green light followed by adaptation in dark the two isomers can be modeled, but when the crystals are treated with red light of 650 nm and kept irradiated during the process show a complete 13-*trans*-15-*trans* retinal model (26) (**Figure I-13**).

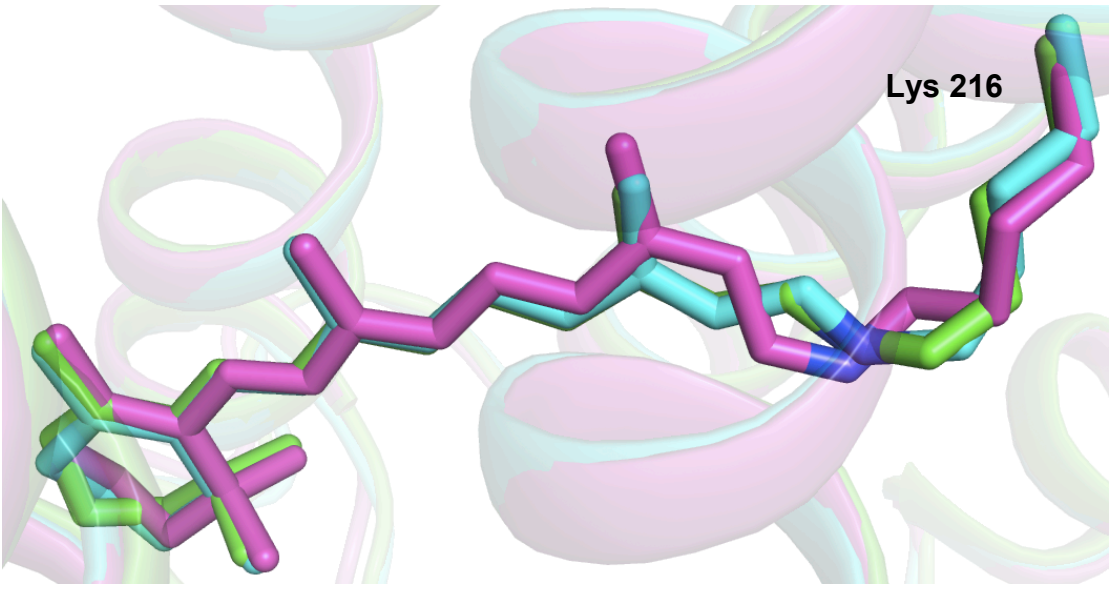


Figure I-13) The overlay structure of the bacteriorhodopsin in the: Ground state (Green – PDB ID: 1C3W), light adapted O-intermediate (Cyan – PDB ID: 3VHZ) and dark adapted O-intermediate (Magenta – PDB ID: 3VI0) shows that the light adapted form can reproduce the ground state geometry, but not the dark adapted crystal.

It should be mentioned at the end of this part that the photo isomerization of the rhodopsin and bacteriorhodosin with the highlighted conformational changes mentioned drops the pKa of the retinal PSB. In Rhodopsin the pKa of the retinal PSB drops from 16 in the 11-*cis* retinal ground state to above 8 in

the all-*trans* retinal of Metal/MetaII forms and to 5.1 in the 15-*cis* retinal MetaIII form (11, 40, 41). In bacteriorhodopsin the pKa drops from 13 in the all-*trans* retinal ground state form to below 7.5 in the 13-*cis* retinal of M form (42-44). Therefore, the pKa drop is one of the essential characteristics of the rhodopsins for their proper functioning.

The idea of combining a retinal light absorbing photo switchable pigment to a fluorescent protein has never been tried. This can make a Reversibly Switchable Fluorescent Protein (RSFP), which has been proven to be an extremely valuable biological tool.

I.6 Reversibly Switchable fluorescent Proteins (RSFPs)

Compared to fluorescent proteins that are all the time in an on-state, Reversibly Switchable Fluorescent Proteins (RSFPs) can be switched between a fluorescent and non-fluorescent state back and forth. The development of RSFPs has pushed the limits of biological microscopy from 200 nm to 30-40 nm (**Figure I-14**) (45-48).

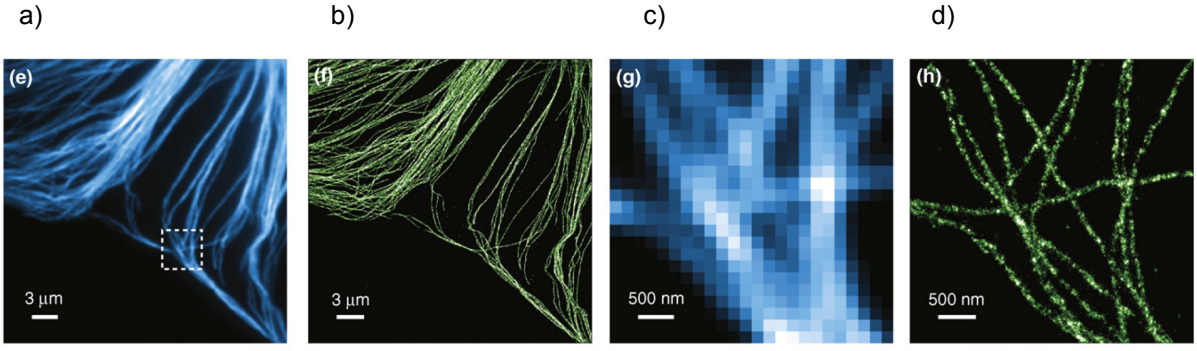


Figure I-14) Conventional immunofluorescence imaging of microtubules (a and c) and STORM images which uses RSFPs (b and d).

RSFPs are the key element to a number of growing advanced imaging techniques, which includes: STORM, PALM and FPALM (49-53),

photochromic FRET (54), optical lock-in detection (55), frequency domain imaging (56), and bio data storage (57). The 2014 chemistry Nobel prize was awarded to the development of super resolution fluorescent microscopy which requires the use of RSFPs.

The mechanism of the switch in the RSFPs has been proven crystalographically in multiple publications to a *cis*- *trans* photo isomerization of the fluorescent chromophore, which in one form the protein, is fluorescent and in the other form it is not (**Figure I-15**) (58-61).

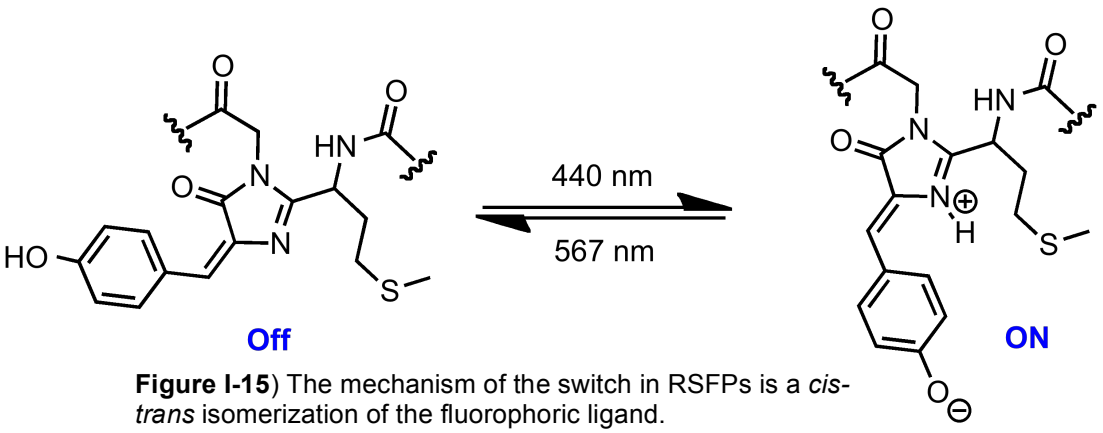


Figure I-15) The mechanism of the switch in RSFPs is a *cis*-*trans* isomerization of the fluorophoric ligand.

The loss of a water molecule that extends the conjugation of the chromophore is also another mechanism that is structurally proved for the mechanism of the switch in RSFPs. In the off state the water makes a covalent bond with the fluorophore to break its conjugation and turns the fluorescence of the protein off. In this specific switchable protein the wavelengths used for the switch are decoupled from the excitation wavelength (62) (**Figure I-16**).

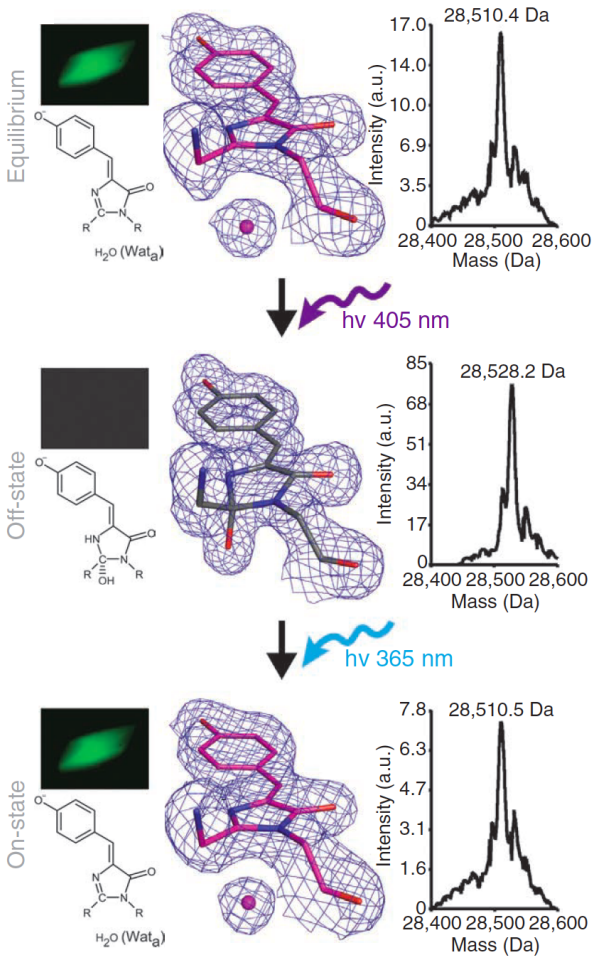


Figure I-16) The loss of a water molecule attributes to the excitation decoupled on-off switch of the fluorescent protein.

Photo bleaching is a process in which the RSFPs lose their fluorescence due to photo destruction of the chromophore. Photo bleaching limits the reversibility of the RSFPs. There are two mechanisms that are proven structurally to be the reason for this phenomenon. With regular laser powers required for the wide-field fluorescence microscopy the Met159 and Cys171 residues undergo oxidation to lock the chromophore in a protonated state. However, with stronger laser powers used in nanoscopy, the decarboxylation of the Glu212 residue, changing of the H-bonding network around the chromophore and the change of the hybridization of the methylene bridge

carbon of the fluorophore from SP^2 to SP^3 counts for the loss of the fluorescence (63).

I.7 Optochemical Genetics

Optochemical genetics uses the small synthetic photo switchable molecules in receptor proteins to give them the ability to respond to light. Like Optogenetics in Optochemical genetics the temporal and spatial control of the protein receptors become feasible by the use of these molecules. Diazo switchable

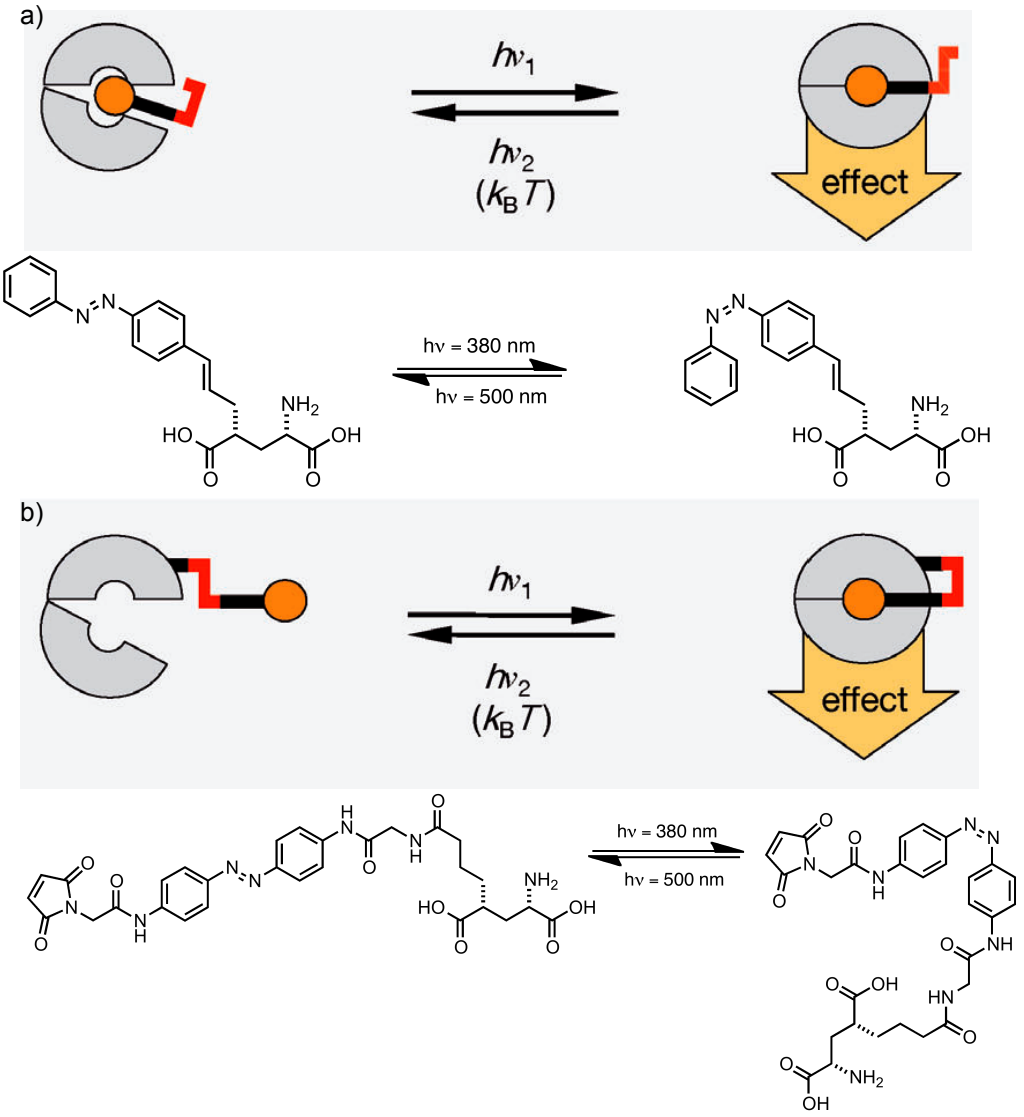


Figure I-17) a) Photochromic ligands (PTLs) in both forms can act as a ligand for protein, but just one of the two forms activates the protein. b) Photo switchable Tethered Ligand (PTL) is covalently bound to the protein, and just one form fits into the binding pocket of the protein.

molecules are the main molecules that have been used to enable trans membrane proteins and neurotransmitters to undergo photo switch by light irradiation (64, 65).

Two main categories of ligands that have been used are Photo Chromic Ligands (PCLs) and Photo switchable Tethered Ligands (PTLs) (**Figure I-17**).

In PCLs from the two *cis* and *trans* forms of the ligand, one of them is the one that is an agonist of the protein or the switch is from the antagonist in one form to agonist in the other form. Upon irradiation and isomerization the ligand goes from one form to another and activates the function of the protein. In PTLs the ligand is tethered to the protein via a covalent bond formation between the protein and the ligand. In one form the tethered ligand does not fit into the binding pocket of the protein, upon irradiation the ligand isomerizes into the appropriate length and spatial availability and locks in the binding pocket of the protein and activates the protein function.

Photo switchable azo drugs are a main category of molecules that have been used in ligand-gated neuro transmission. They have been used in GABA receptors, Neuromuscular receptors, Nicotinic acetyl choline receptors, glutamate receptors, allosteric regulation of GPCRs (66) and enzymatic enantioselectivity (67).

They have been used to control heartbeat (68) and pain sensation (69) and to restore vision (70).

I.8 Photoactive Yellow Protein

Photoactive Yellow Protein (PYP) is a small 14KDa protein, which absorbs the blue light λ_{max} = 446 nm in Halorhodospira halophila. It controls the phototaxis in bacteria to swim away from the harmful blue light and toward the more

photo synthetically suitable lights. PYP is one of the proteins, which forms crystals that respond to light in the crystalline state. PYPs photo intermediates have been extensively studied by time-resolved X-ray crystallography. It should be mentioned that just 10% of the overall population of the molecules respond to light in the crystalline form. The mechanism of the light absorption is a *cis-trans* isomerization around the chromophore of this protein (**Figure I-18**) (71).

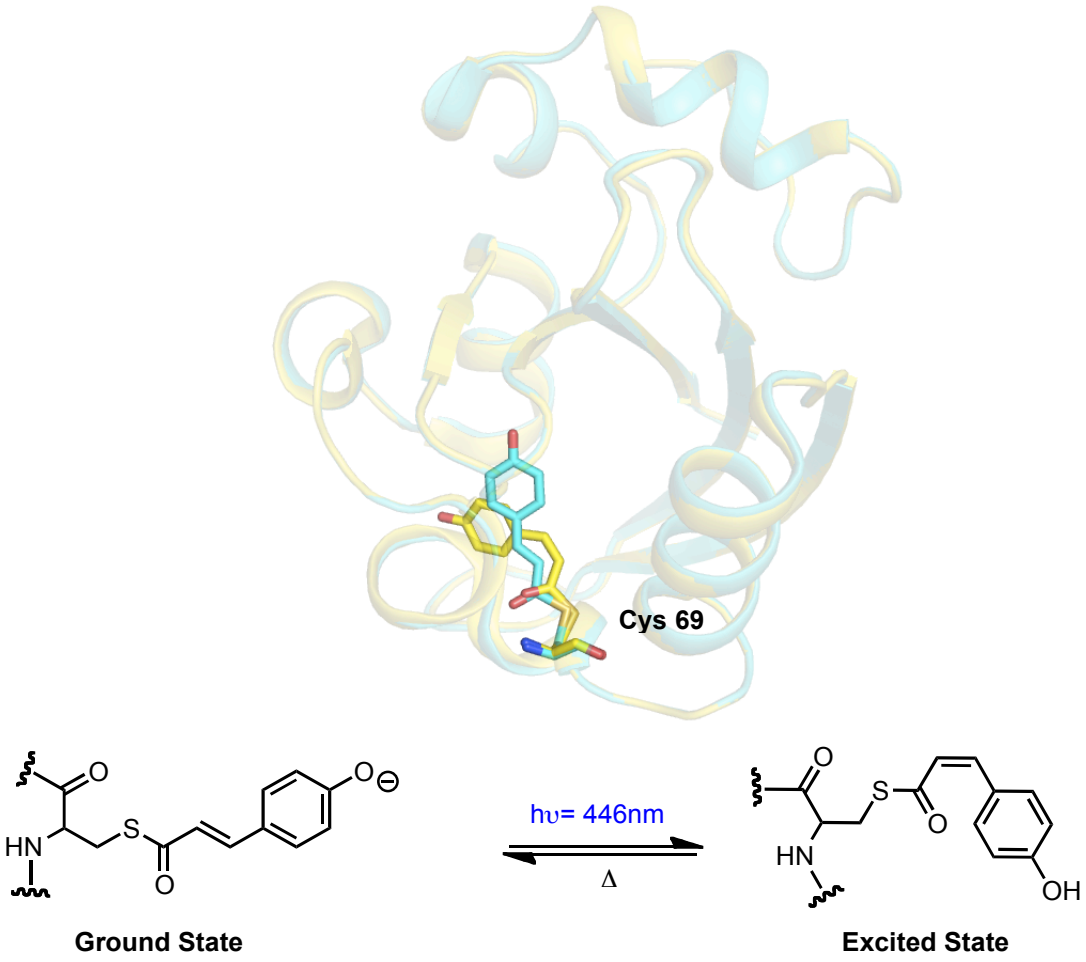


Figure I-18) Photo isomerization of PYP. Overlay structure of the Ground State (Cyan, PDB ID: 2ZOH) and Excited state (Yellow, PDB ID: 4BBV).

I.9 A photo switchable rhodopsin mimic

The result of a reaction of an aldehyde with a primary amine is a well-known reaction that makes a C=N bond known as Schiff base formation. All of the

retinal light absorbing pigments use the same strategy to make a Schiff base which upon protonation makes a Protonated Schiff Base (PSB). Delocalization of the charge along the retinal chromophore causes the absorption of light in the visible region (**Figure I-19**).

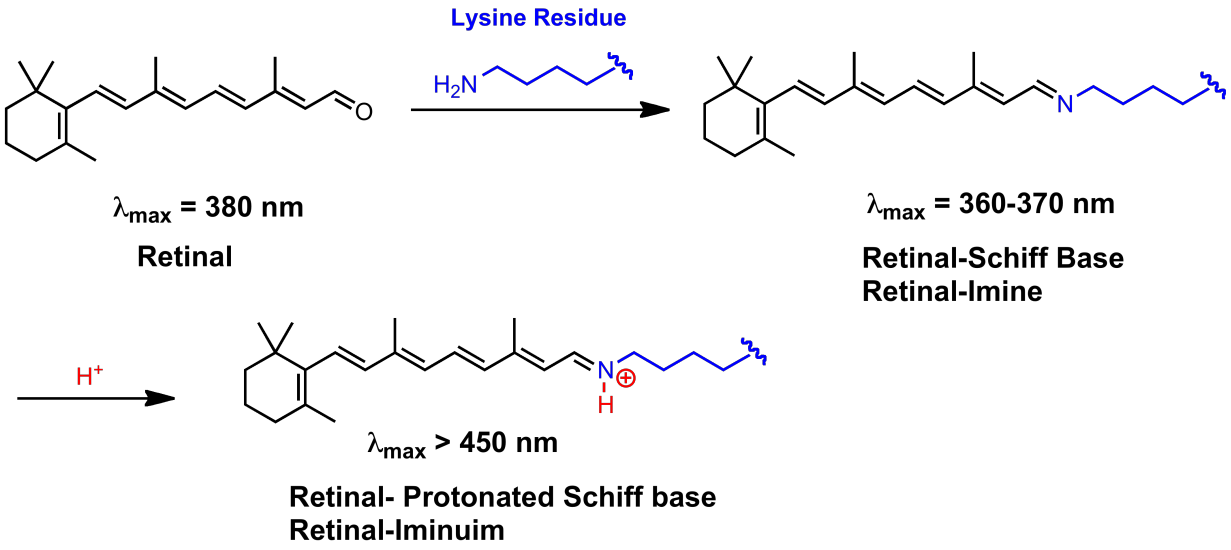


Figure I-19) Formation of a Retinal Schiff Base and Protonated Schiff Base. It is the PSB that absorbs in the visible region of light.

Working with rhodopsins is accompanied by multiple challenges that are typical of *trans*-membrane proteins. These include: expression, purification, solubility, mutagenesis and crystallization.

In the search for mimics that can be good candidates for making a retinal PSB we arrived at Cellular Retinoic Acid Binding ProteinII (CRABPII) and Cellular Retinol Binding ProteinII (CRBPII). These belong to intracellular Lipid Binding Protein Family (iLBPs). CRABPII and CRBPII are soluble small cytosolic transport proteins that can be expressed, purified, mutated and crystallized extensively in our hands (**Figure I-20**).

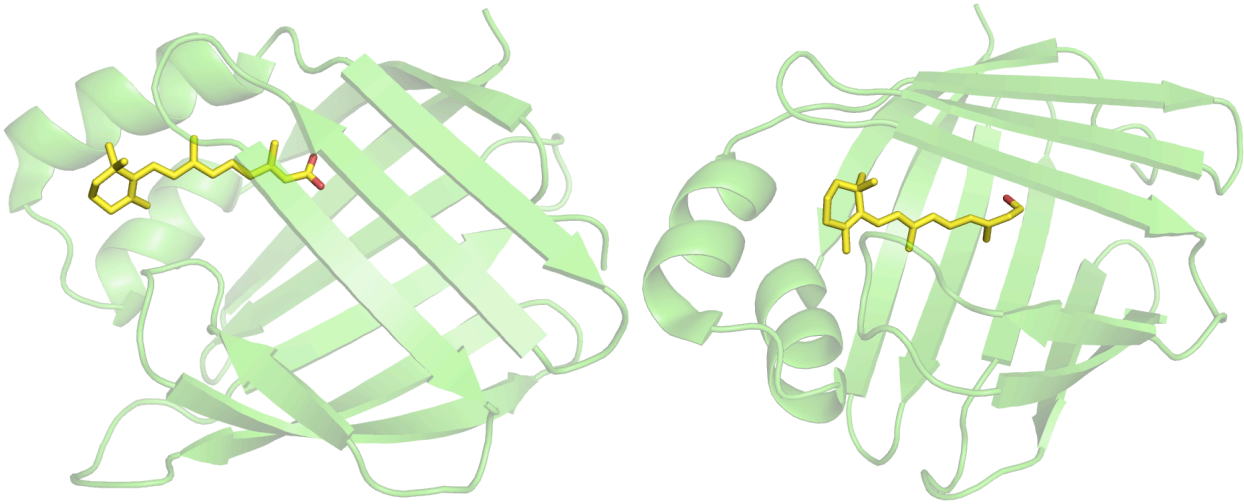


Figure I-20) human Cellular Retinoic Acid Binding ProteinII (hCRABPII) bound to Retinoic Acid (Left-PDB ID: 2FR3). Human Cellular Retinol Binding ProteinII (hCRBPII) bound to retinol (Right-PDB ID: 4QYN).

These two proteins have been reengineered to retinal light absorbing pigments successfully, spanning over 200 nm of the visible light spectrum (72, 73) (**Figure I-21**).

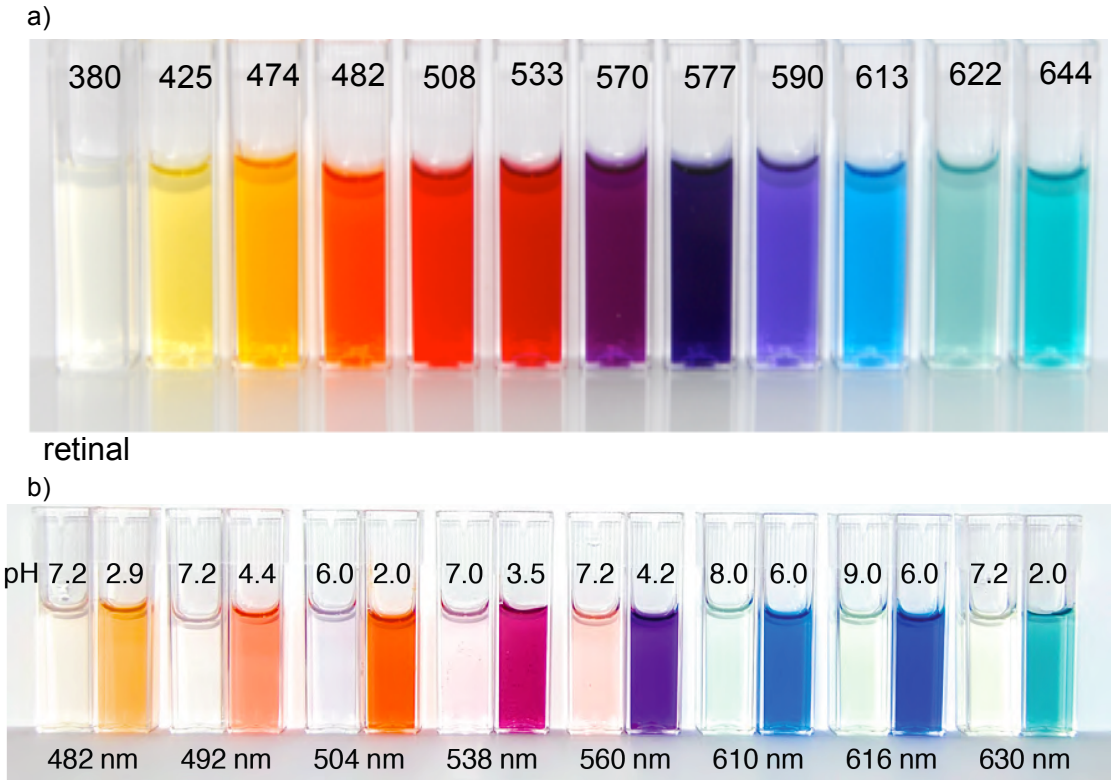


Figure I-21) a) Reengineered hCRBPII protein retinal pigments absorbing light from 425nm to 644nm. Incubated at pH = 7.3. b) Reengineered hCRABPII protein retinal pigments absorbing light from 482nm to 630nm.

We have been reengineering hCRBPII and hCRABPII for a variety of different applications besides wavelength regulation studies, which includes in vitro and in vivo Fluorescent studies and pH sensing (72-82). The common theme for all of these studies lies behind the ability of these proteins to make a Protonated Schiff Base (PSB) upon aldehyde chromophore binding.

The significance of a system that can provide fine tuned isomerization of retinal in a mimic protein intrigued us to focus our studies in this area as well.

In the next chapter the details of reengineering of hCRABPII to a protein that is able to do retinal dark isomerization and photo isomerization, the UV-Vis assays to highlight the spectroscopic characteristics of the protein and the crystallographic data with atomic resolution description of the retinal and protein conformational changes are delineated. This proteins have clean and complete dark isomerization from one isomer to another a characteristic that the rhodopsin is devoid of and in bacteriorhodopsin provides mixture of isomers. The photo isomerization also provides clean and complete photoproducts. Interestingly, these isomerizations are accompanied by 3-5 pKa unit changes, not as significant as visual and microbial rhodopsins pKa changes though.

REFERENCES

REFERENCES

1. Shichi H (1983) *Biochemistry of vision* (Academic Press, New York) pp xii, 279 p.

2. Rando RR (1996) Polyenes and vision. *Chemistry & biology* 3(4):255-262.

3. Wolin LR & Massopus.Lc (1967) Characteristics of Ocular Fundus in Primates. *J Anat* 101:693-&.

4. Shichida Y & Matsuyama T (2009) Evolution of opsins and phototransduction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364(1531):2881-2895.

5. Palczewski K (2006) G protein-coupled receptor rhodopsin. *Annual review of biochemistry* 75:743-767.

6. Bowmaker JK & Dartnall HJ (1980) Visual pigments of rods and cones in a human retina. *The Journal of physiology* 298:501-511.

7. Ernst OP, *et al.* (2014) Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. *Chemical reviews* 114(1):126-163.

8. Kim JE, Tauber MJ, & Mathies RA (2001) Wavelength dependent cis-trans isomerization in vision. *Biochemistry* 40(46):13774-13778.

9. Borhan B, Souto ML, Imai H, Shichida Y, & Nakanishi K (2000) Movement of retinal along the visual transduction path. *Science* 288(5474):2209-2212.

10. Choe HW, *et al.* (2011) Crystal structure of metarhodopsin II. *Nature* 471(7340):651-U137.

11. Mahalingam M & Vogel R (2006) The all-trans-15-syn-retinal chromophore of metarhodopsin III is a partial agonist and not an inverse agonist. *Biochemistry* 45(51):15624-15632.

12. Bartl FJ & Vogel R (2007) Structural and functional properties of metarhodopsin III: recent spectroscopic studies on deactivation pathways of rhodopsin. *Physical chemistry chemical physics : PCCP* 9(14):1648-1658.

13. Vogel R, Ludeke S, Radu I, Siebert F, & Sheves M (2004) Photoreactions of metarhodopsin III. *Biochemistry* 43(31):10255-10264.

14. Zimmermann K, Ritter E, Bartl FJ, Hofmann KP, & Heck M (2004) Interaction with transducin depletes metarhodopsin III: a regulated

retinal storage in visual signal transduction? *The Journal of biological chemistry* 279(46):48112-48119.

15. Palczewski K*, et al.* (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 289(5480):739-745.
16. Nakamichi H & Okada T (2006) Crystallographic analysis of primary visual photochemistry. *Angewandte Chemie* 45(26):4270-4273.

17. Nakamichi H & Okada T (2006) Local peptide movement in the photoreaction intermediate of rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* 103(34):12729-12734.

18. Matsuyama T, Yamashita T, Imamoto Y, & Shichida Y (2012) Photochemical Properties of Mammalian Melanopsin. *Biochemistry* 51(27):5454-5462.

19. Walker MT, Brown RL, Cronin TW, & Robinson PR (2008) Photochemistry of retinal chromophore in mouse melanopsin. *Proceedings of the National Academy of Sciences of the United States of America* 105(26):8861-8865.

20. Rinaldi S, Melaccio F, Gozem S, Fanelli F, & Olivucci M (2014) Comparison of the isomerization mechanisms of human melanopsin and invertebrate and vertebrate rhodopsins. *Proceedings of the National Academy of Sciences of the United States of America* 111(5):1714-1719.

21. Burns ME & Pugh EN, Jr. (2010) Lessons from photoreceptors: turning off g-protein signaling in living cells. *Physiology* 25(2):72-84.

22. David L. Nelson MMC (2012) *Lehninger Principles of Biochemistry* (Freeman, W. H. & Company) 6 Ed.

23. Fu HY*, et al.* (2010) A Novel Six-Rhodopsin System in a Single Archaeon. *Journal of Bacteriology* 192(22):5866-5873.

24. Luecke H, Schobert B, Richter HT, Cartailler JP, & Lanyi JK (1999) Structure of bacteriorhodopsin at 1.55 A resolution. *Journal of molecular biology* 291(4):899-911.

25. Matsui Y*, et al.* (2002) Specific damage induced by X-ray radiation and structural changes in the primary photoreaction of bacteriorhodopsin. *Journal of molecular biology* 324(3):469-481.

26. Zhang J*, et al.* (2012) Crystal structure of the O intermediate of the Leu93-->Ala mutant of bacteriorhodopsin. *Proteins* 80(10):2384-2396.

27. Lanyi JK & Schobert B (2007) Structural changes in the L photointermediate of bacteriorhodopsin. *Journal of molecular biology* 365(5):1379-1392.

28. Takeda K*, et al.* (2004) Crystal structure of the M intermediate of bacteriorhodopsin: allosteric structural changes mediated by sliding movement of a transmembrane helix. *Journal of molecular biology* 341(4):1023-1037.

29. Lanyi J & Schobert B (2002) Crystallographic structure of the retinal and the protein after deprotonation of the Schiff base: the switch in the bacteriorhodopsin photocycle. *Journal of molecular biology* 321(4):727-737.

30. Lanyi JK (2004) What is the real crystallographic structure of the L photointermediate of bacteriorhodopsin? *Biochimica et biophysica acta* 1658(1-2):14-22.

31. Borshchevskiy VI, Round ES, Popov AN, Buldt G, & Gordeliy VI (2011) X-ray-Radiation-Induced Changes in Bacteriorhodopsin Structure. *Journal of molecular biology* 409(5):813-825.

32. Warden MR, Cardin JA, & Deisseroth K (2014) Optical neural interfaces. *Annu Rev Biomed Eng* 16:103-129.

33. Yizhar O, Fenno LE, Davidson TJ, Mogri M, & Deisseroth K (2011) Optogenetics in Neural Systems. *Neuron* 71(1):9-34.

34. Lanyi JK (2004) Bacteriorhodopsin. *Annual review of physiology* 66:665-688.

35. Rothschild KJ (1992) FTIR difference spectroscopy of bacteriorhodopsin: toward a molecular model. *Journal of bioenergetics and biomembranes* 24(2):147-167.

36. Kouyama T & Murakami M (2010) Structural divergence and functional versatility of the rhodopsin superfamily. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* 9(11):1458-1465.

37. Harbison GS*, et al.* (1984) Dark-adapted bacteriorhodopsin contains 13-cis, 15-syn and all-trans, 15-anti retinal Schiff bases. *Proceedings of the National Academy of Sciences of the United States of America* 81(6):1706-1709.

38. Scherrer P, Mathew MK, Sperling W, & Stoeckenius W (1989) Retinal isomer ratio in dark-adapted purple membrane and bacteriorhodopsin monomers. *Biochemistry* 28(2):829-834.

39. Aton B, Doukas AG, Callender RH, Becher B, & Ebrey TG (1979) Resonance Raman study of the dark-adapted form of the purple membrane protein. *Biochimica et biophysica acta* 576(2):424-428.

40. Steinberg G, Ottolenghi M, & Sheves M (1993) Pk(a) of the Protonated Schiff-Base of Bovine Rhodopsin - a Study with Artificial Pigments. *Biophysical journal* 64(5):1499-1502.

41. Bartl FJ & Vogel R (2007) Structural and functional properties of metarhodopsin III: Recent spectroscopic studies on deactivation pathways of rhodopsin. *Physical Chemistry Chemical Physics* 9(14):1648-1658.

42. Sheves M, Albeck A, Friedman N, & Ottolenghi M (1986) Controlling the Pka of the Bacteriorhodopsin Schiff-Base by Use of Artificial Retinal Analogs. *Proceedings of the National Academy of Sciences of the United States of America* 83(10):3262-3266.

43. Rousso I, Friedman N, Sheves M, & Ottolenghi M (1995) Pk(a) of the Protonated Schiff-Base and Aspartic-85 in the Bacteriorhodopsin Binding-Site Is Controlled by a Specific Geometry between the 2 Residues. *Biochemistry* 34(37):12059-12065.

44. Brown LS & Lanyi JK (1996) Determination of the transiently lowered pK(a) of the retinal Schiff base during the photocycle of bacteriorhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* 93(4):1731-1734.

45. Bates M, Huang B, & Zhuang X (2008) Super-resolution microscopy by nanoscale localization of photo-switchable fluorescent probes. *Current opinion in chemical biology* 12(5):505-514.

46. Tiwari DK & Nagai T (2013) Smart fluorescent proteins: Innovation for barrier-free superresolution imaging in living cells. *Development Growth & Differentiation* 55(4):491-507.

47. Nienhaus K & Nienhaus GU (2014) Fluorescent proteins for live-cell imaging with super-resolution. *Chem Soc Rev* 43(4):1088-1106.

48. Shroff H*, et al.* (2007) Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes. *Proceedings of the National Academy of Sciences of the United States of America* 104(51):20308-20313.

49. Stender AS*, et al.* (2013) Single cell optical imaging and spectroscopy. *Chemical reviews* 113(4):2469-2527.

50. Sengupta P, van Engelenburg SB, & Lippincott-Schwartz J (2014) Superresolution imaging of biological systems using photoactivated localization microscopy. *Chemical reviews* 114(6):3189-3202.

51. Rust MJ, Bates M, & Zhuang X (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods* 3(10):793-795.

52. Betzig E, *et al.* (2006) Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313(5793):1642-1645.

53. Hess ST, Girirajan TP, & Mason MD (2006) Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal* 91(11):4258-4272.

54. Subach FV, *et al.* (2010) Red fluorescent protein with reversibly photoswitchable absorbance for photochromic FRET. *Chemistry & biology* 17(7):745-755.

55. Marriott G, *et al.* (2008) Optical lock-in detection imaging microscopy for contrast-enhanced imaging in living cells. *Proceedings of the National Academy of Sciences of the United States of America* 105(46):17789-17794.

56. Tian ZY & Li ADQ (2013) Photoswitching-Enabled Novel Optical Imaging: Innovative Solutions for Real-World Challenges in Fluorescence Detections. *Accounts Chem Res* 46(2):269-279.

57. Adam V, *et al.* (2010) Data storage based on photochromic and photoconvertible fluorescent proteins. *Journal of biotechnology* 149(4):289-298.

58. Andresen M, *et al.* (2005) Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* 102(37):13070-13074.

59. Henderson JN, Ai HW, Campbell RE, & Remington SJ (2007) Structural basis for reversible photobleaching of a green fluorescent protein homologue. *Proceedings of the National Academy of Sciences of the United States of America* 104(16):6672-6677.

60. Andresen M, *et al.* (2007) Structural basis for reversible photoswitching in Dronpa. *Proceedings of the National Academy of Sciences of the United States of America* 104(32):13005-13009.

61. Pletnev S, Subach FV, Dauter Z, Wlodawer A, & Verkhusha VV (2012) A structural basis for reversible photoswitching of absorbance spectra in red fluorescent protein rsTagRFP. *Journal of molecular biology* 417(3):144-151.

62. Brakemann T, *et al.* (2011) A reversibly photoswitchable GFP-like protein with fluorescence excitation decoupled from switching. *Nature Biotechnology* 29(10):942-U132.

63. Duan CX, *et al.* (2013) Structural Evidence for a Two-Regime Photobleaching Mechanism in a Reversibly Switchable Fluorescent Protein. *Journal of the American Chemical Society* 135(42):15841-15850.

64. Fehrentz T, Schonberger M, & Trauner D (2011) Optochemical Genetics. *Angewandte Chemie-International Edition* 50(51):12156-12182.

65. Stein M, *et al.* (2012) Azo-Propofols: Photochromic Potentiators of GABAA Receptors. *Angewandte Chemie-International Edition* 51(42):10500-10504.

66. Pittolo S, *et al.* (2014) An allosteric modulator to control endogenous G protein-coupled receptors with light. *Nat Chem Biol* 10(10):813-815.

67. Bautista-Barrufet A, *et al.* (2014) Optical Control of Enzyme Enantioselectivity in Solid Phase. *Acs Catal* 4(3):1004-1009.

68. Fortin DL, *et al.* (2008) Photochemical control of endogenous ion channels and cellular excitability. *Nature methods* 5(4):331-338.

69. Mourot A, *et al.* (2012) Rapid optical control of nociception with an ion-channel photoswitch. *Nature methods* 9(4):396-402.

70. Polosukhina A, *et al.* (2012) Photochemical restoration of visual responses in blind mice. *Neuron* 75(2):271-282.

71. Schotte F, *et al.* (2012) Watching a signaling protein function in real time via 100-ps time-resolved Laue crystallography. *Proceedings of the National Academy of Sciences of the United States of America* 109(47):19256-19261.

72. Wang WJ, *et al.* (2012) Tuning the Electronic Absorption of Protein-Embedded All-trans-Retinal. *Science* 338(6112):1340-1343.

73. Berbasova T, *et al.* (2013) Rational Design of a Colorimetric pH Sensor from a Soluble Retinoic Acid Chaperone. *Journal of the American Chemical Society* 135(43):16111-16119.

74. Crist RM, Vasileiou C, Rabago-Smith M, Geiger JH, & Borhan B (2006) Engineering a rhodopsin protein mimic. *Journal of the American Chemical Society* 128(14):4522-4523.

75. Vaezeslami S, Mathes E, Vasilelou C, Borhan B, & Geiger JH (2006) The structure of apo-wild-type cellular retinoic acid binding protein II at 1.4 A and its relationship to ligand binding and nuclear translocation. *Journal of molecular biology* 363(3):687-701.

76. Vasileiou C, *et al.* (2007) Protein design: Reengineering Cellular Retinoic Acid Binding Protein II into a rhodopsin protein mimic. *Journal of the American Chemical Society* 129(19):6140-6148.

77. Vaezeslami S, Jia XF, Vasileiou C, Borhan B, & Geiger JH (2008) 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* 64:1228-1239.

78. Vasileiou C, *et al.* (2009) Dissection of the critical binding determinants of cellular retinoic acid binding protein II by mutagenesis and fluorescence binding assay. *Proteins-Structure Function and Bioinformatics* 76(2):281-290.

79. Vasileiou C, *et al.* (2009) Elucidating the exact role of engineered CRABPII residues for the formation of a retinal protonated Schiff base. *Proteins-Structure Function and Bioinformatics* 77(4):812-822.

80. Lee KSS, *et al.* (2012) Probing Wavelength Regulation with an Engineered Rhodopsin Mimic and a C15-Retinal Analogue. *Chempluschem* 77(4):273-276.

81. Nossoni Z, *et al.* (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal. *Acta crystallographica. Section D, Biological crystallography* 70(Pt 12):3226-3232.

82. Yapici I, *et al.* (2015) "Turn-on" protein fluorescence: in situ formation of cyanine dyes. *Journal of the American Chemical Society* 137(3):1073-1080.

Chapter II: Development of a Photoswitchable engineered human Cellular Retinic Acid Binding ProteinII

II.1 Introduction

Retinoid Binding Proteins (RBPs) are a large family of binding proteins, which their function is the transfer, metabolism and regulation of the biological effects of retinoids in living organisms (**Figure II-1**) (1, 2).

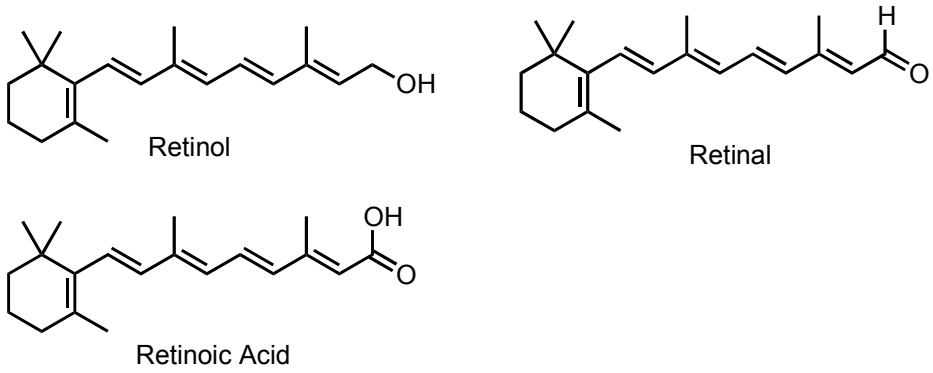


Figure II-1) Chemical structure of different retinoids.

Cellular Retinoic Acid Binding Protein II (CRABPII) is one of the proteins of this family, which binds to all-trans retinoic acid (3, 4) and solubilizes this hydrophobic molecule in the polar environment of the cells and delivers the retinoic acid through trans-localization to the nucleus to Retinoic Acid Receptors (RARs) (5, 6), which are transcription factors. Crystal structure of the wild type human CRABPII protein bound to Retinoic Acid shows the interacting residues with the ligand in the binding pocket of the protein (**Figure II-2**) (7).

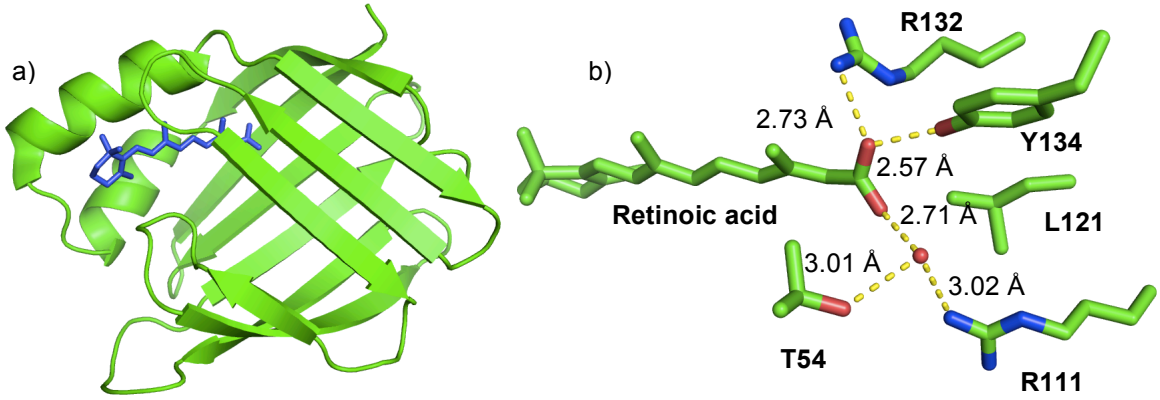


Figure II-2) a) Crystal structure of hCRABPII bound to Retinoic Acid. b) Binding pocket of hCRABPII and interacting residues with the Retinoic Acid (PDB ID: 2FR3).

In collaboration with Professor Borhan's lab we have been reengineering this protein in the last decade for a variety of different research studies, which include wavelength regulation (7, 8), pH sensing (9) and fluorescent studies (10). The first attempts of wavelength regulation studies was reengineering the binding pocket of the protein in a way that enables the protein to make a Schiff Base (SB, imine) and upon protonation a Protonated Schiff Base (PSB, iminium).This idea comes from Rhodopsins that make a PSB to absorb light (11). To this end, the Arg132 was mutated to a lysine to introduce the nucleophilic lysine residue for PSB formation, Arg111 was mutated to hydrophobic Leucine to remove the positive charge close to the retinal PSB and finally the Leu121 was mutated to a glutamic acid residue to introduce the counter anion for PSB stabilization. R132K:R111L:L121E mutant (KLE) absorbs at 450 nm with quick PSB formation completion in minutes and a pKa of 8.7. Unfortunately, the crystallization trials of different mutants of this series have not provided any retinal bound structures except the KLE mutant. The crystal structure of this mutant shows the retinal covalently bound to Lys 132 and the other interacting residues with the PSB (**Figure II-3**) (8).

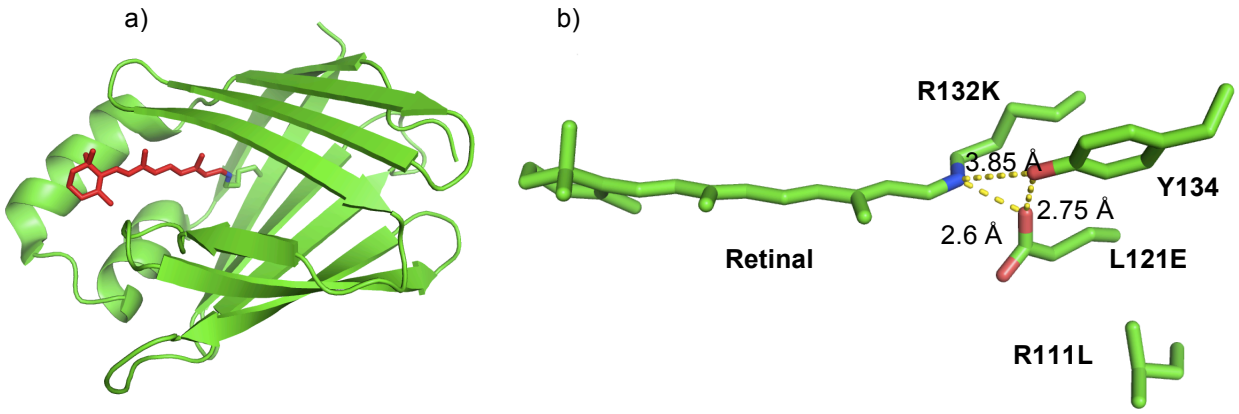


Figure II-3) a) Crystal structure of the R132K:R111L:L121E mutant of hCRABPII (PDB ID: 2G7B) shows the retinal covalently attached to the protein. b) Interacting residues with the iminium Schiff Base in this crystal structure.

Interestingly, different mutations on the entrance of the binding pocket of this protein show no effect on the wavelength of the mutated proteins. Additionally, different mutant proteins show no absorption above 450 nm (**Table II-1**).

Mutant	**λ_{max} (nm)**	**K_d**
R111K:R111L:L121E	449	1±5
R111K:R111L:L121E:R59K	444	157±24
R111K:R111L:L121E:R59H	440	18±12
R111K:R111L:L121E:R59A	442	92±8
R111K:R111L:L121E:R59L	443	10±9
R111K:R111L:L121E:R59F	440	11±9
R111K:R111L:L121E:R59Y	435	2±8
R111K:R111L:L121E:R59W	442	38±8
R111K:R111L:L121E:R59Q	444	16±8
R111K:R111L:L121E:R59D	440	73±11
R111K:R111L:L121E:R59E	450	37±6

Table II-1. Summary of the mutations on R59 position of first generation of hCRABPII mutants.

A surface representation of the crystal structure of KLE mutant protein shows that the retinal chromophore is not completely embedded in the binding pocket of the protein, leaving the chromophore exposed to the aqueous buffer, which have a much bigger dielectric constant compared to the one that

is implied to the chromophore through mutations. In other words, the aqueous media that the chromophore has remained exposed to wipes out the effect of the mutations on the wavelength of the protein (**Figure II-4**) (12). This category of proteins will be referred to as the first generation of hCRABPII mutant proteins.

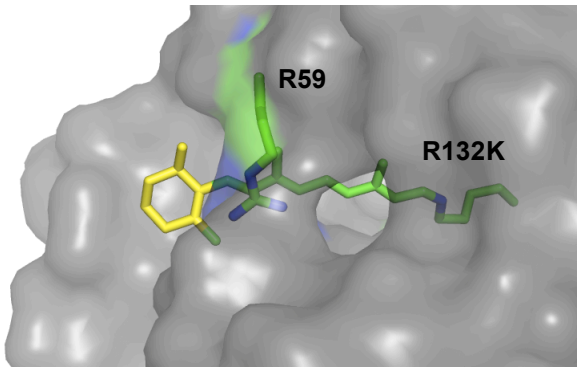


Figure II-4) Crystal structure of R132K:R111L:L121E mutant, shows the ionone ring is positioned out of the binding pocket (PDB ID: 2G7B).

To solve this problem there are three strategies that proved to be effective:

(1) Shortening the length of the chromophore by removing two double bonds of the polyene to make a C15-anolog of retinal (12).

(2) Reengineering the binding pocket of the protein in a way that it can completely encapsulate the retinal (9).

(3) Using another protein from Retinoid Binding Protein Family as surrogate (13).

With the first strategy not only the chromophore is completely embedded by the protein (**Figure II-5**), but also the mutations that originally had no effect on the wavelength of the retinal (**Table II-1**) have a clear effect on the wavelength of the C15-retinal analog (**Table II-2**).

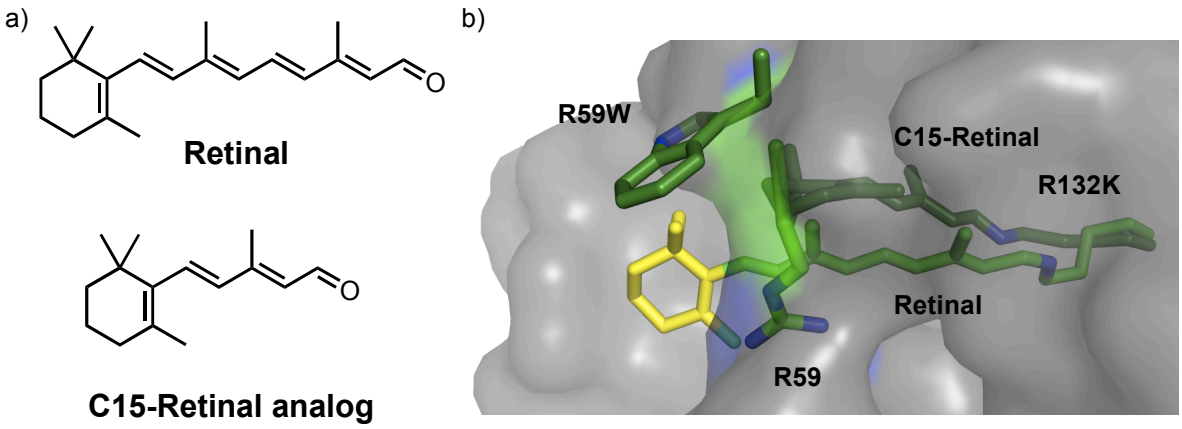


Figure II-5) a) Chemical representation of retinal C15-analog compared to retinal. b) Overlay structure of R132K:R111L:L121E-retinal complex with R132K:R111L:L121E:R59W-C15 retinal complex.

Mutant	λ_{max} (nm) – C15 analog	K_d
R111K:R111L:L121E	333	157±24
R111K:R111L:L121E:R59K	330	1390±161
R111K:R111L:L121E:R59H	398	719±96
R111K:R111L:L121E:R59A	397	236±29
R111K:R111L:L121E:R59L	391	144±19
R111K:R111L:L121E:R59F	392	279±33
R111K:R111L:L121E:R59Y	396	226±33
R111K:R111L:L121E:R59W	404	259±17
R111K:R111L:L121E:R59Q	413	259±26
R111K:R111L:L121E:R59D	409	262±35
R111K:R111L:L121E:R59E	424	1±4

Table II-2. Summary of the mutations on R59 position of first generation of hCRABPII mutants.

With the second strategy not only the retinal is completely encapsulated in the binding pocket of the protein, but with the right mutations a broad range of wavelengths has been covered, which provide proteins that are more red shifted than any retinlydene protein pigment known. This category of proteins will be referred to as second generation of hCRABPII mutant proteins (9) (**Figure II-6**).

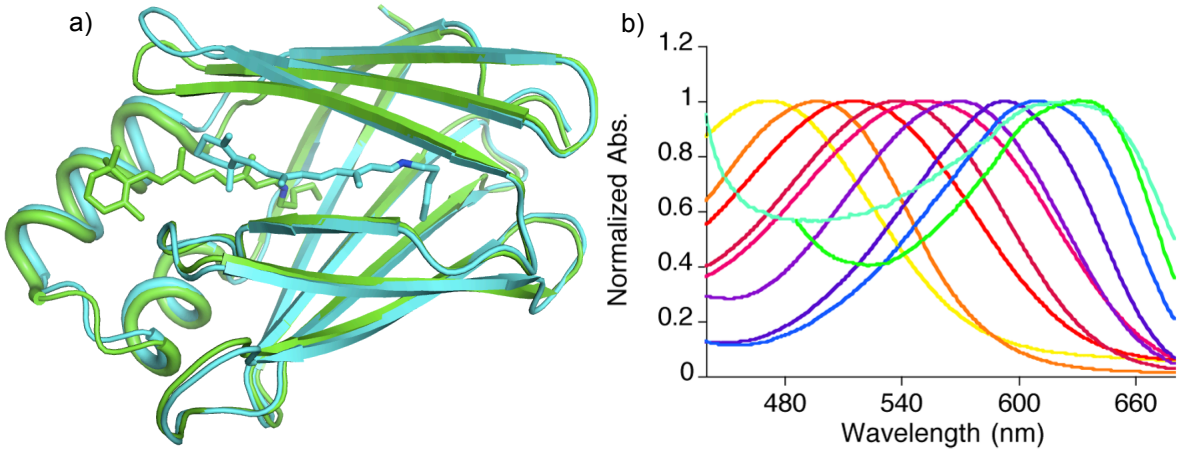


Figure II-6) a) Overlay crystal structures of R111K:R132L:Y134F:T54V:R59W and R132K:R111L:L121E shows the retinal structure deeper inside the binding pocket of hCRABPII. b) Different mutants of second generation hCRABPII absorb from 474 to 640 nm.

Two of the observed characteristics of second generation hCRABPII proteins is their low pK_as, which essentially leaves these proteins with no PSB at physiological pH 7.3 (**Table II-3**) and their quick PSB loss overtime, sometimes in minutes upon incubation with retinal (**Figure II-7**) (9).

Mutant	**λ_{max} (nm)**	**pKa**
1-R111K:R132L:Y134F:T54V:R59W:A32W	610	7.0
2-R111K:R132L:Y134F:T54V:L121Q:R59Y:A32W	538	5.9
3-R111K:R132L:Y134F:T54V:L121Q	492	5.3
4-R111K:R132L:Y134F:T54V:L121Y:R59Y:A32W	529	5.1
5-R111K:R132L:Y134F:T54V:L121Q:R59W	505	4.4
6-R111K:R132L:Y134F:T54V:L121N	495	4.2
7-R111K:R132L:Y134F:T54V:L121Q:R59W:A32W:M93L:E73A	560	4.1
8-R111K:R132L:Y134F:T54V:L121Y:R59W:A32W:M93L:E73A	530	3.0
9-R111K:R132L:Y134F:T54V:L121Y:R59W	504	2.8
10-R111K:R132Q:Y134F:T54V:R59W:A32W:M93L:E73A	630	2.6
11-R111K:R132L:Y134F:T54V:R59W:A32W:E73A	620	2.4

Table II-3. Mutants of second generation hCRABPII with pKas below physiological pH.

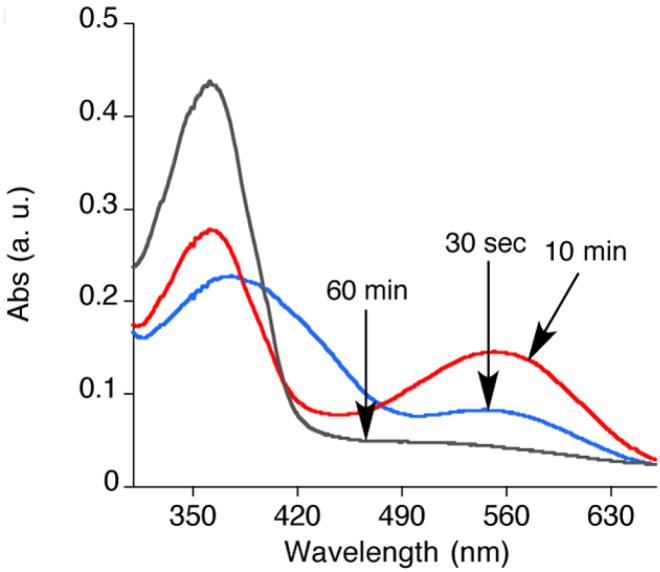


Figure II-7) PSB loss of the incubated R111K:R132L:Y134F:T54V:R59W mutant of hCRABPII with retinal at pH=5.2.

In this chapter I will discuss the underlying reasoning for these two observations and how it took us to a photo switchable retinal-based protein.

II.2 Developing of hCRABPII to a Photoswitchable Rhodopsin mimic

One of the other members of iLBPs that we have been reengineering is human Cellular Retinol Binding Protein (hCRBPII), which also shows a wide range of absorption spectrum for the mutants of this protein (**Figure II-8**).

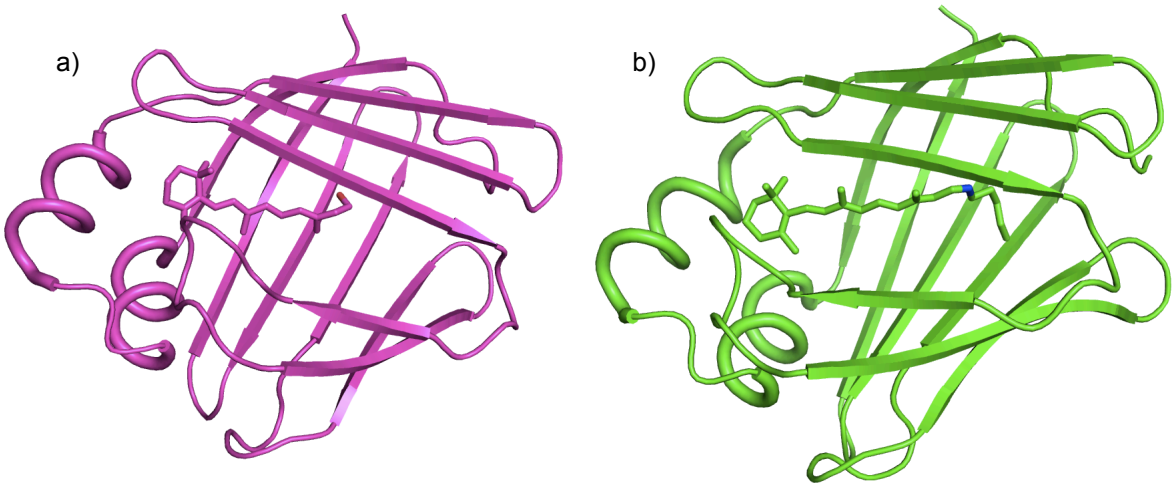


Figure II-8) a) Retinol inside the binding pocket of wt-hCRBPII (PDB ID: 4QYN). b) Retinal PSB covalently bound inside the binding pocket of Q108K:K40L mutant of hCRBPII (PDB ID: 4RUU).

Unlike the second generation of hCRABPII proteins, hCRBPII proteins have high pK_a mutants with pK_as as high as 10.2 (13) (**Table 4**).

Mutant	λ_{max} (nm)	pKa
Q108K:K40L	508	8.3
Q108K:K40L:R58Y	535	9.5
Q108K:K40L:T51V:R58Y	555	10.0
Q108K:K40L:T51V:R58Y:Y19W	561	10.2
Q108K:K40L:T51V:T53C:R58W:T29L:Y19W	591	8.2
Q108K:K40L:T51V:T53C:R58W:T51V:Y19W:Q4A	613	7.0

Table II-4. Different mutant of hCRBPII with high pK_as.

This led us to compare the binding pocket of these two proteins to figure out the differences, and to see if the problem of the low pK_a of hCRABPII second generation mutants can be solved by mimicking the binding pocket of hCRBPII into hCRABPII.

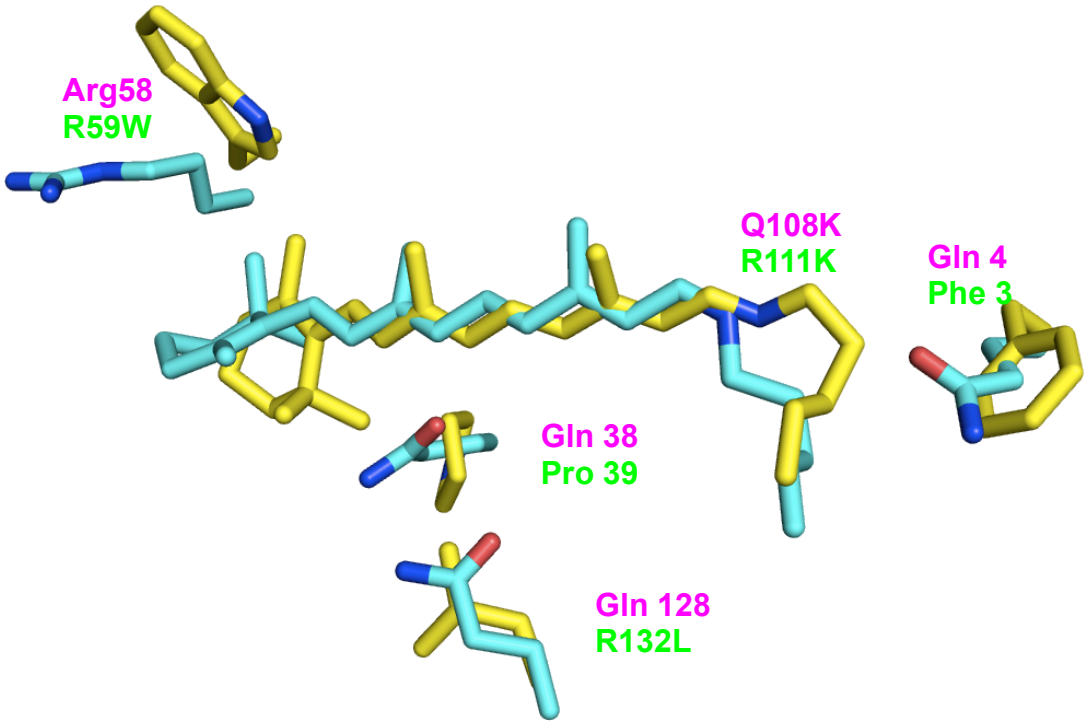


Figure II-9) The structures of the hCRABPII mutant R111K:R132L:Y134F:T54V:R59W (Yellow, PDB ID 4I9S) and the hCRBPII mutant Q108K:K40L (Cyan, PDB ID 4RUU) were overlaid and the region around the retinal binding site is shown. Note the three Glutamines in hCRBPII are substituted with hydrophobic residues in hCRABPII.

A comparison of the binding pocket of hCRBPII with hCRABPII shows three glutamine residues in hCRBPII in the vicinity of the chromophore that are not conserved in hCRABPII (**Figure II-9**). In each case mutation of these residues to Gln in hCRABPII would result in a more hydrophilic binding cavity, which we suspected might lead to stabilization of the retinal PSB. In addition, mutation of Arg58 (Arg59 in hCRABPII) to Tyr in hCRBPII also reliably resulted in PSB's with higher pKa's. (**Table II-4**). This led us to the R111K:Y134F:T54V:**R132Q:P39Q**:R59Y (**M1**) hCRABPII mutant. Incubation of M1 with retinal at pH=5.0 showed complete PSB formation in 3 hours and 30 minutes. However, the acid-base titration of this mutant revealed a broad curve, ranging from pH of 4.3 to 10.5, inconsistent with a single pKa. Other mutants containing P39Q and R132Q mutations gave similar results (**Figure II-10**).

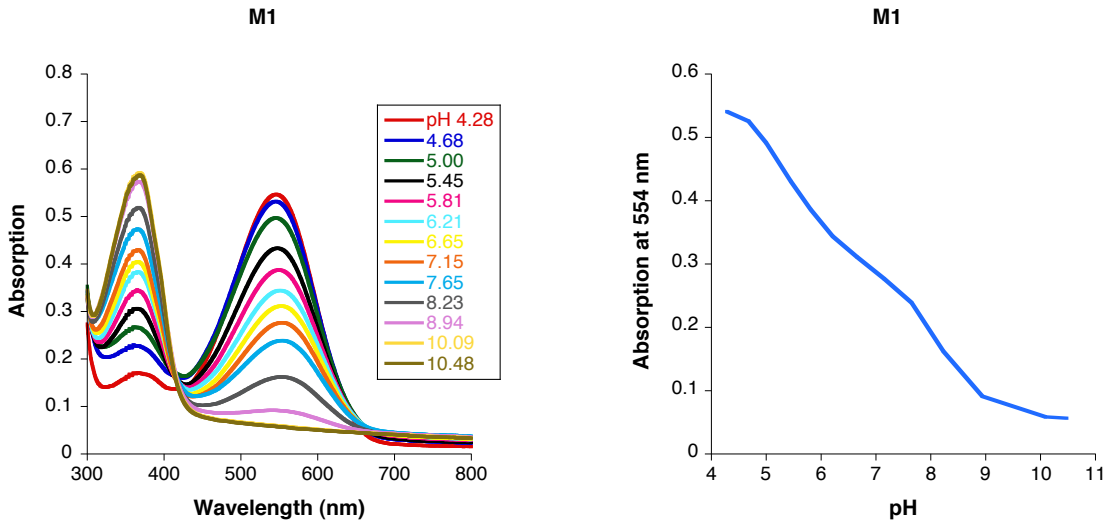


Figure II-10) pKa titration of M1 mutant in citrate buffer at pH=5. Protein concentration is 50μM incubated with 0.5 equivalent retinal.

A possible explanation for this behavior came from our previous studies of retinal-bound hCRABPII mutants, where the PSB was lost over 1-2 hours or in some cases minutes (**Figure II-7**). Subsequent acidification regenerated the

PSB and acid-base titration showed the conversion of the retinal PSB from a higher to a lower pKa form, suggesting a change of environment in the vicinity of the iminium nitrogen. Such a change in environment could be attributed to an isomerization of one of the double bonds in the chromophore polyene, analogous to that seen in the natural rhodopsin systems, where light isomerization of retinal, most commonly of the 11 or 13 double bonds, leads to the change in the position and therefore pKa of the iminium (11). If correct this would suggest a thermal analog of a functional rhodopsin system, where isomerization is occuring thermally and results in a significant change in the pKa of the iminium (**Figure II-11**). It should be mentioned that the loss of the PSB can not be attributed to retinal loss of the protein, since after the loss of the PSB it can be recovered by acidification to form a new PSB.

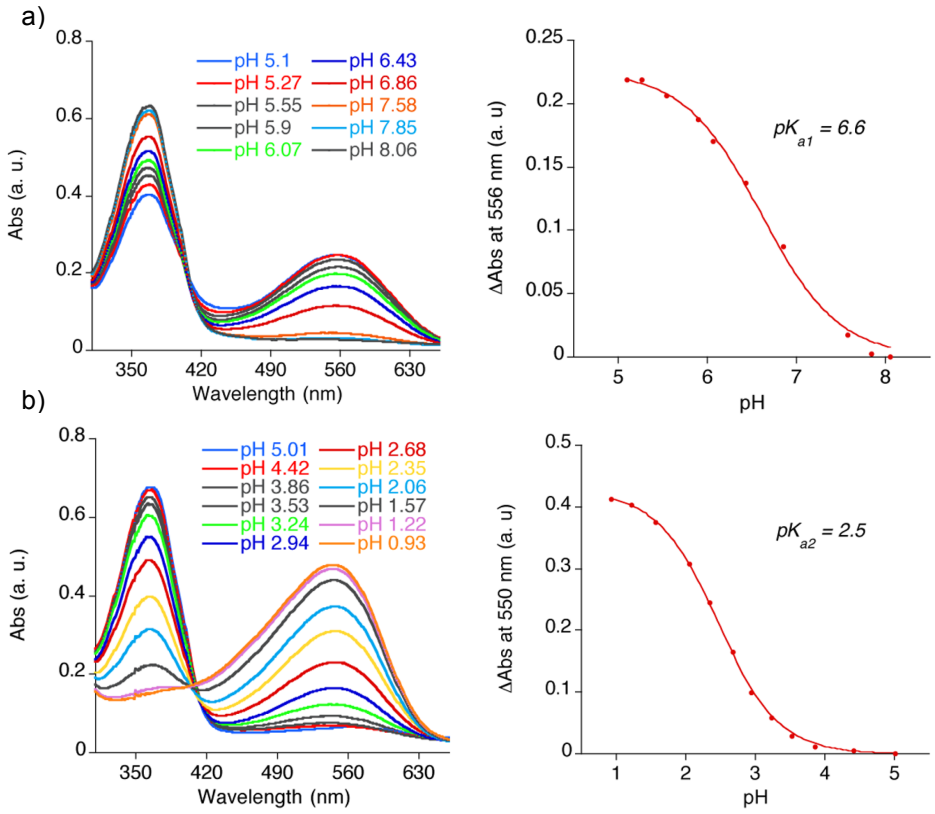


Figure II-11) a) Base titration of kinetic product b) acid titration of thermodynamic product of R111K:R132L:Y134F:T54V:R59W mutant of hCRABPII.

The most possible explanation for this event comes from our already present structural data, which was the retinal iminium double bond isomerization itself. We have shown in our hCRBPII structural studies that different mutants of this protein can adopt a *cis*-PSB retinal (15-*cis* retinal) when the Gln4 residue is preserved and a *trans*-SB retinal (15-*trans* retinal) when this residue is mutated (**Figure II-12b**, **Figure II-12c**). Also the structural data from the early mutants of the hCRABPII, which the P39Q and R132Q mutations are not implied the refinement of both isomers of 15-*cis* retinal and 15-*trans* retinal into the electron density map is feasible (**Figure II-12d**).

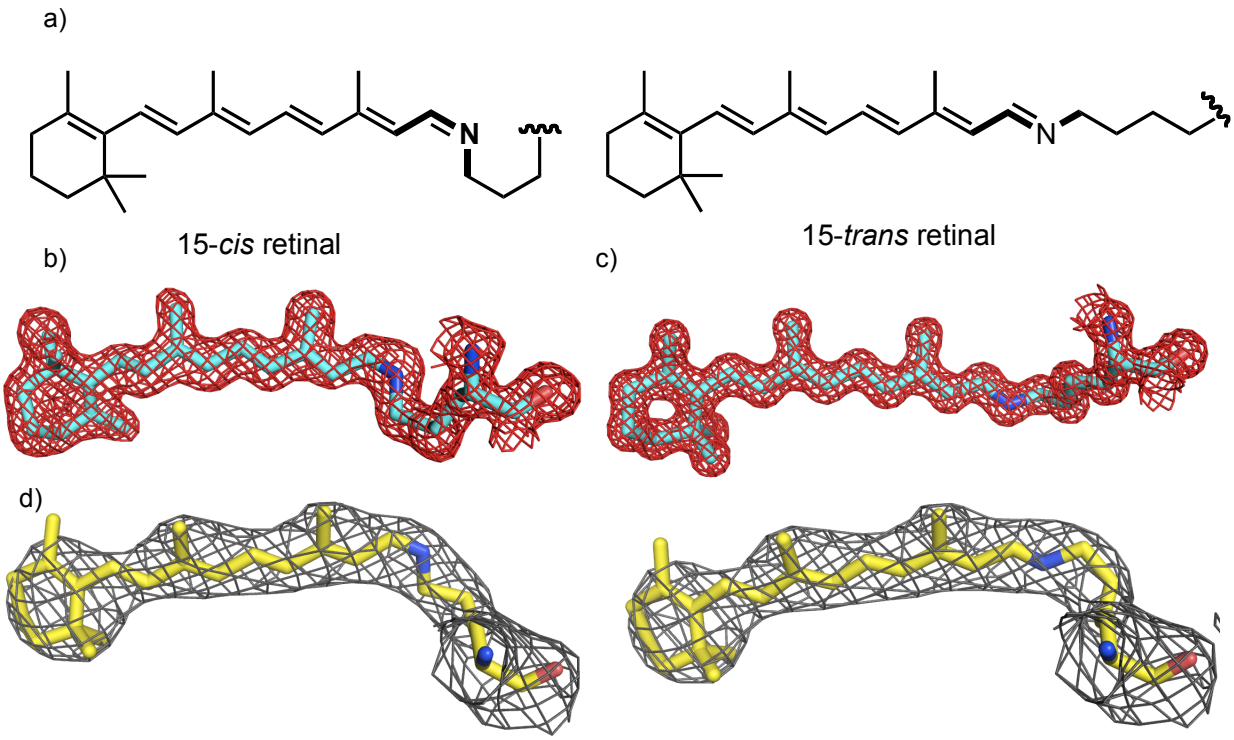


Figure II-12) a) a chemical representation of 15-*cis* retinal (left) and 15-*trans* retinal (right) b) Retinal density in crystal structure of Q108K:K40L:T51V:R58Y:Y19W mutant of hCRBPII shows a *cis*-PSB density for retinal c) Retinal density in crystal structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H mutant of hCRBPII shows a *trans*-SB density for retinal. d) Retinal density in crystal structure of R111K:R132L:Y134F:T54V:R59W:A32W mutant of hCRABPII shows both 15-*cis* retinal (left) and 15-*trans* retinal (right) which can be fit into the density of retinal (all the maps are contoured at 1σ).

Like the other second generation hCRABPII mutants, M1 also converts over time from a higher to lower iminium pKa form. However, the rate of

interconversion between high and low pKa forms is substantially slower than that seen in previous mutants (Compare **Figure II-7** and **Figure II-13a**), which may allow for the trapping and unequivocal identification of each form.

In an attempt to identify the final product, crystallization of M1 was initiated only after all of the PSB had disappeared (24 hours). This resulted in a clear electron density map consistent with an all *trans* conformation for all the chromophore's double bonds, including the iminium double bond (**Figure II-13b**). In this isomer the environment around the iminium nitrogen is very hydrophobic, consistent with its low pKa (2.8 pH units lower than a potential 15-*cis* isomer).

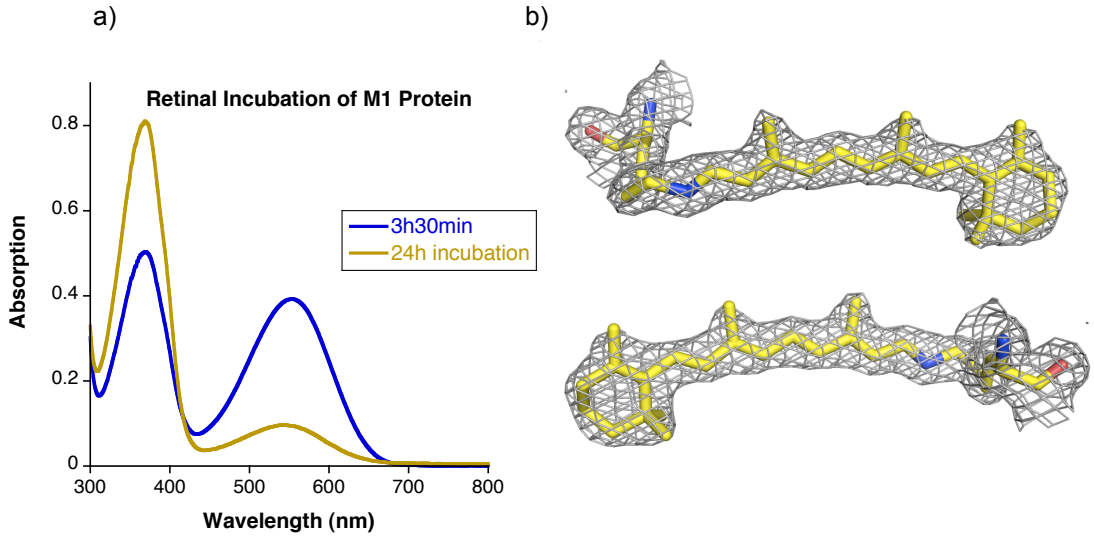


Figure II-13) Thermal interconversion in M1 a) UV/Vis spectra of the M1/retinal complex 3.5 hours and 24 hours after addition of retinal to the protein. b) Retinal density in the crystal structure of R111K:R134F:T54V:R132Q:P39Q:R59Y at 1.83 Å contoured at 1σ shows a *trans*-SB (PDB ID: 4YBP).

However, attempts at trapping the high pKa form of M1 by initiating crystallization trials immediately after maximum PSB formation resulted in crystals with ambiguous electron density for the chromophore, especially in the vicinity of the iminium region. There are two potential problems with this approach. First and most obvious is that partial conversion may occur before

crystal nucleation, resulting in an ambiguous mixture of the two isomeric forms. The other problem is the comparatively long incubation time required for the PSB to fully form in the first place (3 hours and 30 minutes). This probably results in a situation where the meta-stable form is never the exclusive form in solution. The only way to overcome this problem is to significantly accelerate the formation of the PSB, so that the high pKa form is fully formed before it has time to significantly degrade to the low pKa form. Table 5 summarizes the mutants made in pursuit of an ideal system for this study. It should be mentioned that R132Y mutation fails to accelerate the PSB formation, P39N mutation drops the pKa of the protein and in some cases is accompanied by protein expression problem and P39E mutation does not provide stable proteins. Clear from the data is that the presence of both Gln39 and Gln132 leads to both slow PSB formation and slow subsequent PSB loss.

Mutant*	**Low pK_a** (λ_{max} nm)	**High pK_a** (λ_{max} nm)	**PSB rise (hours)**	**PSB fall (hours)**
M1-KFVQ:P39Q:R59Y	5.4 (530)	8.4 (556)	3	24
M2-KFVQ:P39Y:R59Y	3.4 (564)	8.15 (573)	0.2	10
M3-KFVQ:P39Q:R59Y:A32W:F3Q	9.4 (557)	-	5	50 % after 24h
M4-KFVQ: P39Q: R59W	5.1 (537)	7.8 (555)	3.5	24
M5-KFVQ:P39Q: R59Y:A32Y	5.6 (554)	8.4 (575)	4	24
M6-KFVQ:P39Q:R59Y:A32W	6.0 (563)	8.3 (584)	2	24
M7-KFVQ:P39Q:R59W:A32W	6.2 (582)	8.6 (598)	2	24
M8-KFVQ:P39Y: R59Y:A32Y	3.2 (587)	8.4 (588)	0.1	16
M9-KFVQ:P39Y:R59Y:A32W	2.5 (600)	8.3 (596)	0.2	12
M10-KFVQ:P39Y:R59W:A32W	3.1 (601)	8.3 (595)	0.3	10
M11-KFVQ:P39Q:R59Y:A32Y:F3Q	9.3 (545)	-	3	-
M12-KFVQ:P39Q:R59W:A32W:F3Q	9.0 (571)	-	3	-

Table II.5 Screening for hCRABPII mutants with fast kinetic PSB formation and slow PSB loss.

- KFVQ is the abbreviation for R111K:Y134F:T54V:R132Q.

However, mutants with Tyr at position 39 by and large form the initial PSB much faster, with PSB formation complete in about 10 minutes. Based on

these observations mutant R111K:R134F:T54V:R132Q:P39Y:R59Y (**M2**) was made. M2 forms a PSB within 20 minutes of retinal addition, and slowly deprotonates over a course of 10 hours (**Figure II-14a**). Crystallization trials of retinal-bound (M2) was initiated both 20 minutes and 24 hours after retinal addition. Gratifyingly, electron density produced from crystals grown after short retinal incubation clearly showed the presence of a *cis*-PSB retinal, while crystals grown from protein incubated with retinal for 24 hours produced electron density that was unambiguously that of a *trans*-SB retinal, a result similar to that of the M1 mutant (**Figure II-14b** and **Figure II-14c**).

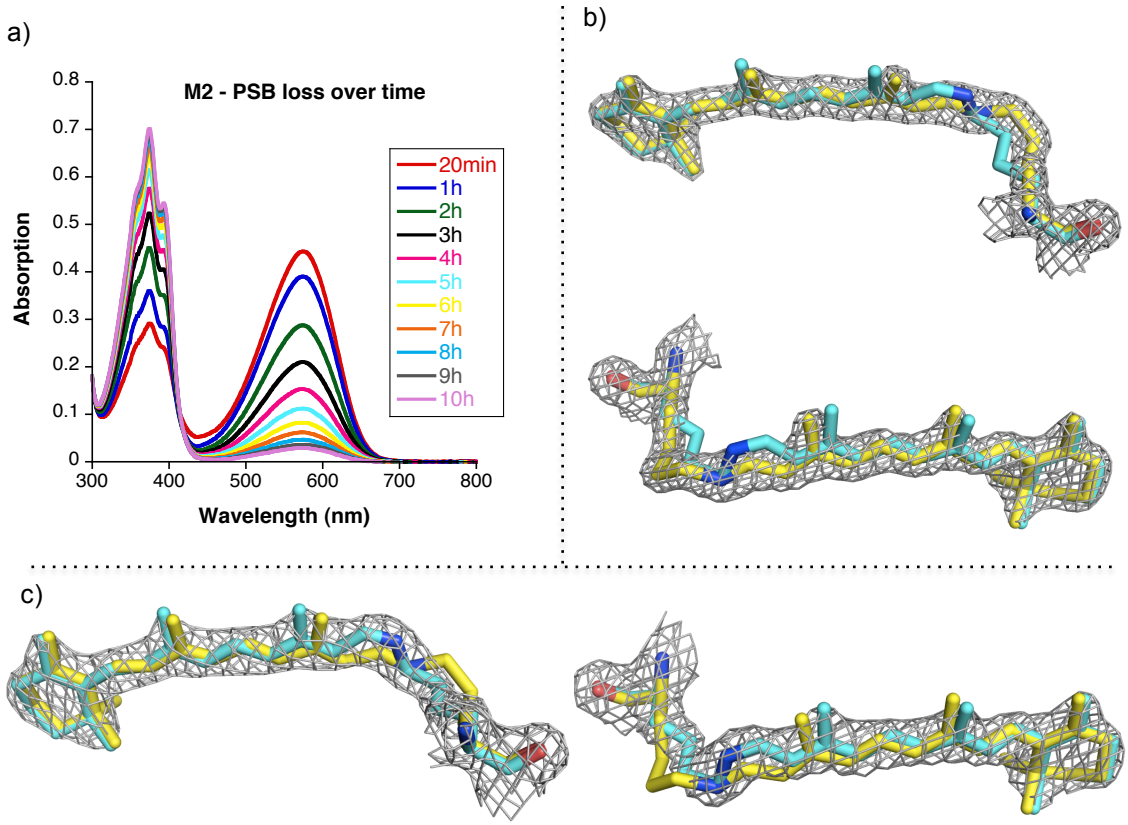


Figure II-14) Thermal isomerization of a CRABPII rhodopsin mimic. a) PSB loss over time for mutant M2. M2 was incubated with retinal, and UV/Vis spectra were taken at various time points after the addition; b) Crystal structure of M2 showing the *trans*-SB at 1.62 Å, with the electron density contoured at 2.0 σ, (Yellow, protein was incubated with 4 equiv retinal for 24 hours at room temperature before crystallization, PDB ID: 4YFQ), overlaid on the *cis*-PSB of the same mutant crystallized after a 20 minute retinal incubation (Cyan); c) Crystal structure of the M2 mutant in *cis*-PSB form at 1.97 Å at 1.5 σ, (Cyan, the protein was incubated with 4 equivalents retinal at room temperature for 20 minutes and followed by immediate crystallization, PDB ID: 4YFP) overlaid on the *trans*-SB crystal structure of the same mutant (Yellow).

These results suggest that crystal nucleation occurs rapidly, allowing the less stable *cis*-PSB retinal to nucleate before significant amounts of the *trans*-SB retinal have formed.

The M2 mutant has all the necessary characteristics of a functional rhodopsin mimic in that it can interconvert via chromophore isomerization between two distinct iminium pKa environments. However, the key characteristic of such a system is that it must be photoswitchable. A particularly attractive aspect of the M2 protein is the low pKa (pKa = 3.4) for its 15-*trans* retinal form. At a pH of 7, essentially none of this form is expected to be protonated, which means that only the high pKa form will absorb visible light. On the other hand, with a pKa of 8.2, almost all of the high pKa form should be protonated and therefore not absorb in the UV region of the spectrum. In principle, this should make it possible to quantitatively convert one form into the other depending on the wavelength of light used for irradiation. To demonstrate this light induced interconversion, we incubated M2 with retinal, and after the PSB absorbance reached its maximum, the complex was exposed to visible light. The PSB absorbance at 571 nm was immediately reduced, reaching a minimum in 1 minute of exposure. To test the system's reversibility, the sample was then irradiated with UV light, resulting in an immediate reduction in the SB peak at 370 nm with a concurrent increase in the PSB absorbance at 571 nm, indicating the system to be reversible. This can be repeated for several cycles, suggesting that the interconversion between the two forms can be repeated (**Figure II-15**).

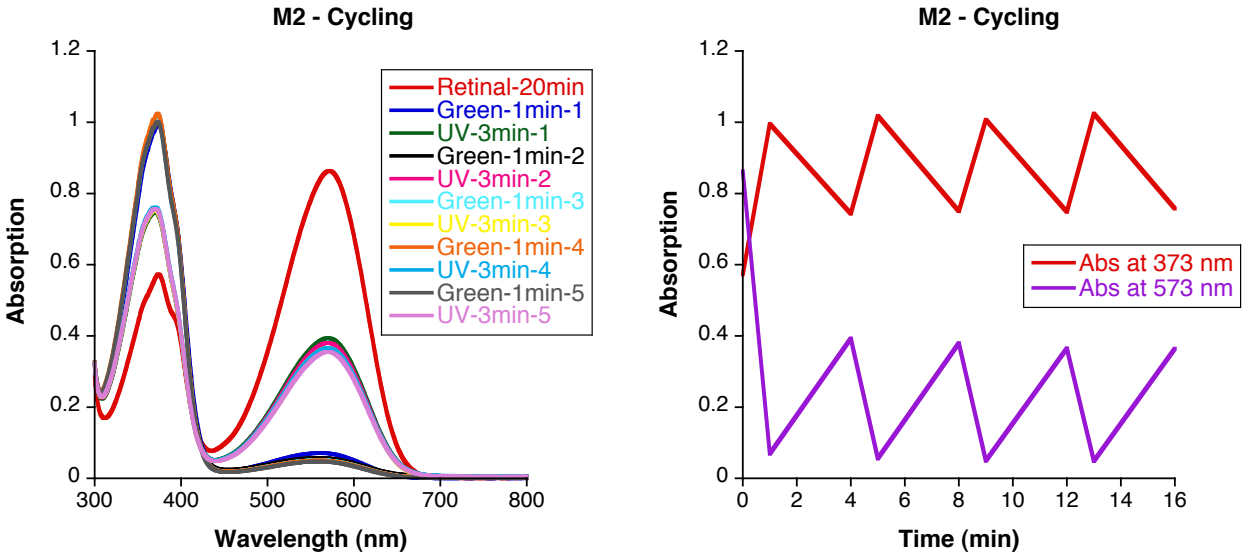


Figure II-15) Cycling of the hCRABPII proteins in solution with UV light (340-380 nm) and Green light (520-600 nm).

We next sought to identify the two interconverting species crystallographically. Different experiments including the UV irradiation of the solution before crystallization and irradiating for every 12 hours during crystallization process, UV irradiation of the crystals on the X-ray beam before data collection and UV irradiation of the crystals formed from a solution after PSB loss right before freezing them were tried to prove that the later is the most effective way of approaching this experiment. We took the crystals of retinal-bound M2 that had been grown after 24 hours of incubation, resulting in complete loss of the meta-stable high pK_a form, and soaked them in a stabilizer at pH 7.5. As shown, these crystals were completely colorless, indicating minimal absorption in the visible spectrum. As previously described, the structure obtained from such crystals shows exclusively a *trans*-SB bound chromophore (**Figure II-14b**). We then exposed these crystals to UV light, resulting in a dramatic color change of the crystals. The crystals were then

exposed to visible light resulting in the colorless crystals originally obtained. This process could be repeated several times indicating the reversibility of the photo-process in the crystal (**Figure II-16b**). The crystals exposed to UV light were immediately frozen in liquid nitrogen. The structure obtained from these crystals unambiguously shows a *cis*-PSB retinal, demonstrating that the UV light-induced isomerization involves the *trans* to *cis* isomerization of the imine double bond. For these proteins, the same process that was previously observed occurring thermally also occurs photochemically.

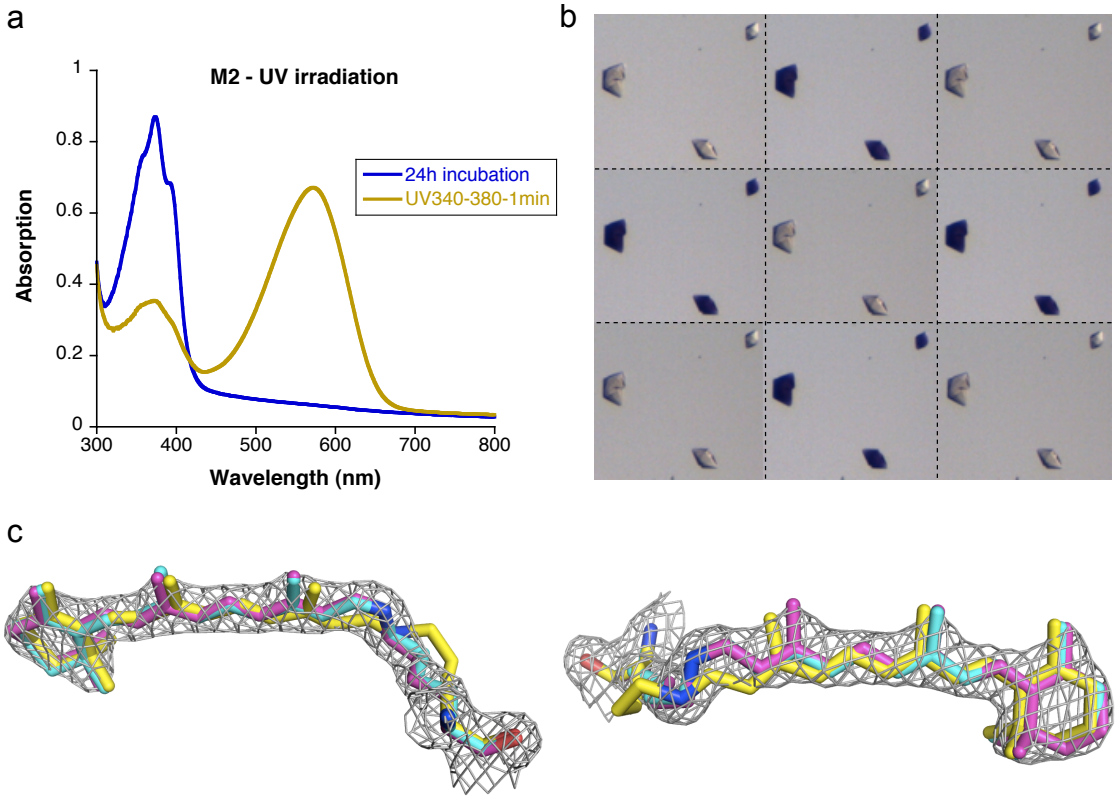


Figure II-16) Photo interconversion in both solution and crystal. a) UV irradiation of M2 after complete PSB loss shows the PSB recovery in solution; b) Cycles of UV and visible light irradiation of M2 crystals. (Each Cycle is 5 min UV light and 5 min room light at pH 7.5); c) The structure of M2 obtained from a crystal that was UV irradiated for 30 minutes at 1.95 Å. The electron density was contoured at 1.5 σ and the structure clearly shows the retinal in a *cis*-iminium conformation (Magenta, PDB ID: 4YFR) overlaid on the structure of the *trans*-SB (Yellow, PDB ID: 4YFQ) and *cis*-PSB (Cyan, PDB ID: 4YFP) crystal structures.

An overlay of the meta-stable 15-*cis*-retinal structure obtained by immediate crystallization after PSB formation and the UV irradiated crystal structure shows the two to be essentially identical, while the low pK_a 15-*trans*-retinal structure is distinctly different (**Figure II-16c**).

Crystal structures were also obtained of both UV and visible light irradiated crystals after two and three cycles of crystal irradiation for M1 and M2 proteins, to confirm structurally that the photo-isomerization reliably gives the same two products (**Figure II-17**).

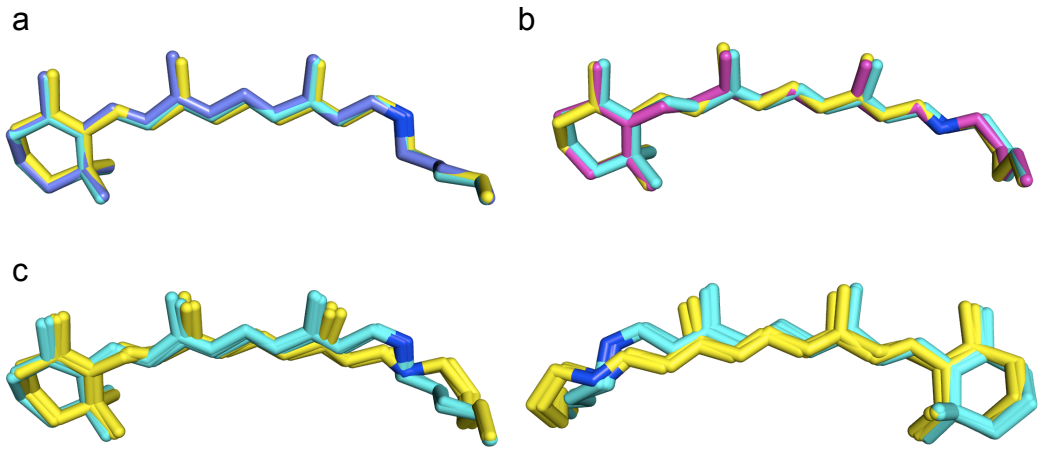


Figure II-17) Cycling of the M2 mutant crystals: a) Overlay of the retinal for the first three cycles of UV light irradiated crystals of the M2 mutant for 30 minutes. b) Overlay of the retinal for the first three cycles of visible light irradiated crystals of the M2 mutant for 30 minutes. c) Overlay of the retinal for the first three cycles of UV light irradiated crystals (cyan) and visible light irradiated crystals (yellow) of the M2 mutant each for 30 minutes.

- All of these crystals diffract between 1.6-2.2. PDB IDs in supporting information.

The overlay of the two isomeric forms of M2 indicates that the critical residues surrounding the chromophore are unchanged (**Figure II-18**). However, the most drastic difference is in the trajectory of the lysine 111, which forms the iminium. An isomerization from *trans* to *cis* is accompanied by an outward movement of the carbons of the lysine, showing that most of the motion required for the interconversion involves the change of the conformation of the

lysine side chain, not the chromophore. The last carbon of the Lys111 in 15-*cis*-retinal PSB overlays with the nitrogen of the same residue in 15-*trans*-retinal SB form, putting the protonated nitrogen of the lysine in a 3-3.5 Å distance from Trp109, close enough to make a strong π-cation interaction. This up and outward movement of the Lys 111 causes the carbons of the polyene chain and angular methyls to move up and out to accommodate to the new iminium position, being more noticeable in the first few carbons of the polyene. Towards the end of the polyene and the ionone ring, the outward movement of the chromophore becomes less significant. The only noticeable movement of the residues inside the binding pocket is related to Leu121. In the 15-*trans*-retinal form this residue has enough space to rotate for two different conformations. However, as the lysine moves out in the 15-*trans*-retinal form, one of the conformations becomes unavailable for Leu121 due to steric hindrance (**Figure II-18**).

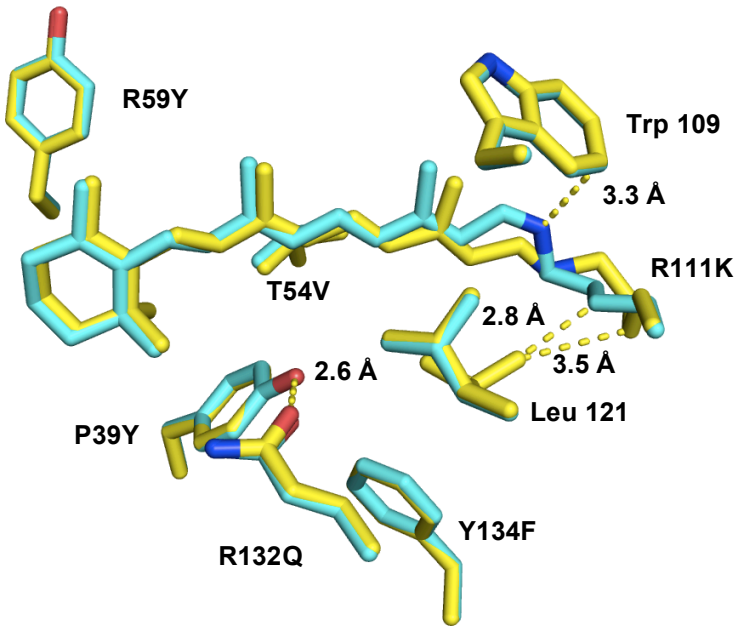


Figure II-18) The environment surrounding the iminium in M2 protein. The *trans*-SB (Yellow) and *cis*-PSB (Cyan).

Figure 19 summarizes the retinal isomerization in hCRABPII proteins. Starting with a *cis*-PSB as the initial kinetic form of the protein, a thermal isomerization to *trans*-SB form reduces the pK_a of the protein below physological pH and blue shifts the protein-retinal complex by 200 nm. This process can be switched back and forth to the *cis*-PSB and *trans*-SB form by UV light and green light irradiation, respectively.

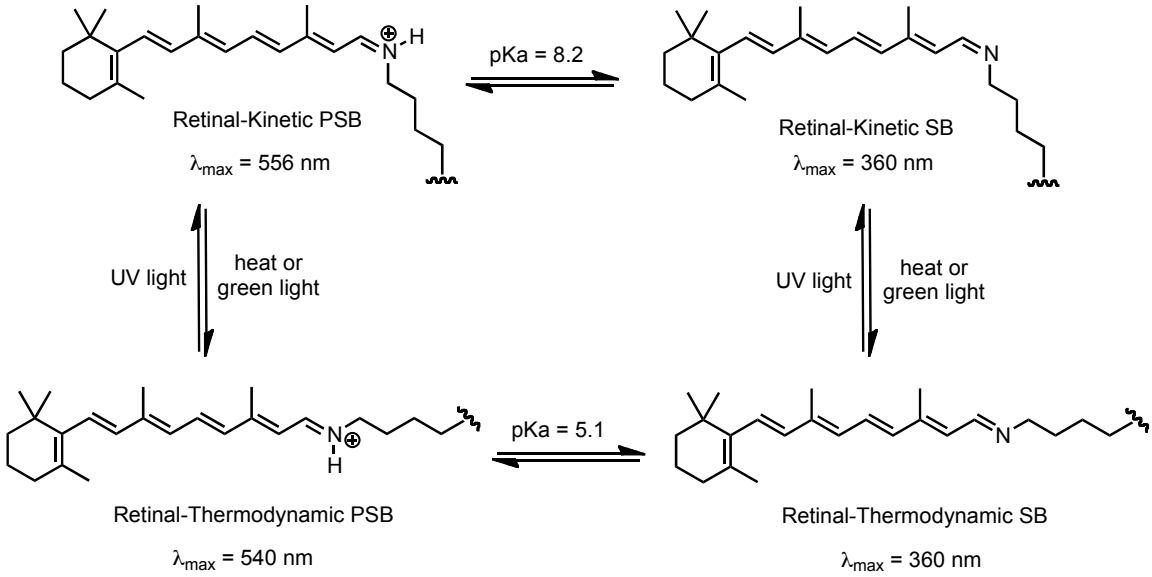


Figure II-19) Thermodynamic and light induced *cis*, *trans*-retinal iminium isomerization in hCRABPII protein.

pK_a titration of each of the *cis*-PSB, *trans*-SB and UV irradiated forms of these proteins provides a single proton pK_a titration curve. In the case of R132Q:P39Q mutants there is a 2-3 pK_a unit change between the 15-*cis*-retinal and 15-*trans* retinal forms of the protein. This difference in the pK_a increases to 5 units in the R132Q:P39Y mutants. The significant pK_a change between the 15-*cis* and 15-*trans* retinal forms indicate two different retinal iminium environments. In the 15-*cis* retinal form the Lys111 nitrogen faces towards Trp109, making a π-cation interaction (**Figure II-20a**). However, in the *trans*-SB form the Lys111 nitrogen faces hydrophobic aliphatic residues

including Val41, Ile43, Leu119 and Leu121 (**Figure II-20b**), which leads to a pK_a drop.

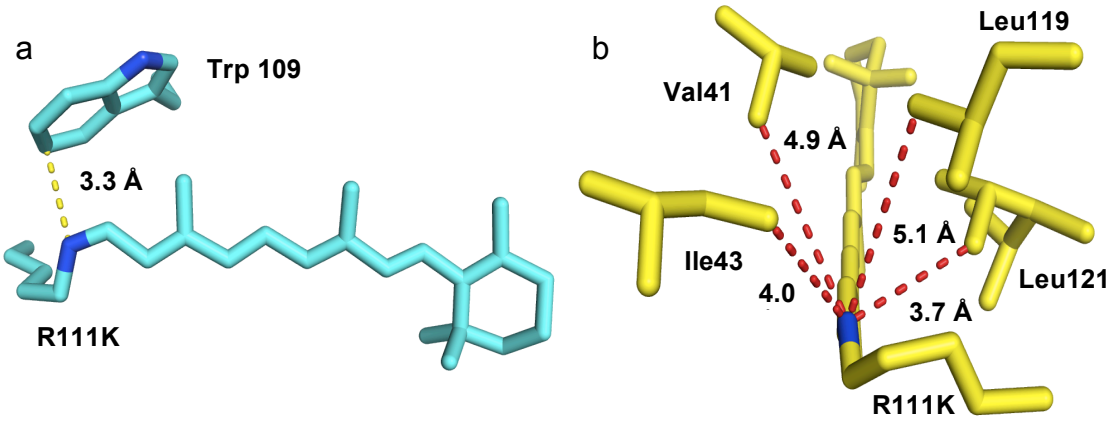


Figure II-20) M2 crystal structure indicates the difference in the two forms of the protein. a) A π-cation interaction between Trp109 and the PSB nitrogen of the *cis*-PSB contributes to the higher pK_a in this form; b) Aliphatic hydrophobic residues surrounding the *trans*-SB in M2 depress the pK_a in this form.

So far, we have successfully designed rhodopsin mimics that recapitulate the basic chemistry of a rhodopsin in that the iminium moiety can be shifted between two distinct pK_a environments via photo isomerization of a polyene double bond. However, all of these systems thermally relax to the lower pK_a all *trans*-retinylidene PSB. We next sought to create a system where the thermal product was the higher pK_a *cis*-PSB form. This is the case in many rhodopsin systems such as 11-*cis*-retinal in the visual rhodopsins, where the all *trans*-unprotonated imine is meta stable and ultimately exits the active site. To produce such a variant it is necessary to stabilize the *cis*-PSB retinal form of the protein. We again took a cue from our work on hCRBPII, where we have shown that the presence of a glutamine at position 4 (residue 3 for hCRABPII) gives the *cis*-PSB, while mutation of this residue gives the *trans*-iminium (**Figure II-12a** and **Figure II-12b**). A water mediated interaction

between Gln4 and the iminium nitrogen explains the stabilization of the *cis* conformation (**Figure II-21**).

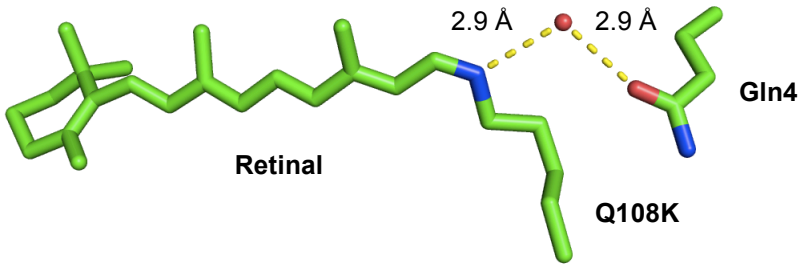


Figure II-21) crystal structure of Q108K:K40L shows that Gln4 contributes to the stabilization of a *cis*-PSB in the hCRBPII mutants (PDB ID: 4RUU).

Analogously, we added the F3Q mutation (equivalent to the position 4 in hCRBPII) to the R111K:Y134F:T54V:R132Q:P39Q:R59Y:A32W (M6) scaffold resulting in the R111K:Y134F:T54V:R132Q:P39Q:R59Y:A32W:F3Q mutant (**M3**). Gratifyingly retinal-bound M3 gave the most stable higher pK_a form so far observed, with almost 50% of the high pK_a form still present 24 hours after formation. The pK_a of this mutant was also increased by 1.1 units to 9.4 (compared to 8.3 for M6), and the absorbance blue shifted about 27 nm (compare entry M6 to entry M3 in Table 5). Exposure of M3 to visible light results in loss of the PSB absorbance with a concomitant increase in SB absorbance, but the PSB absorbance recovers in darkness, indicating thermal isomerization to the higher pK_a *cis*-iminium form (**Figure II-22**).

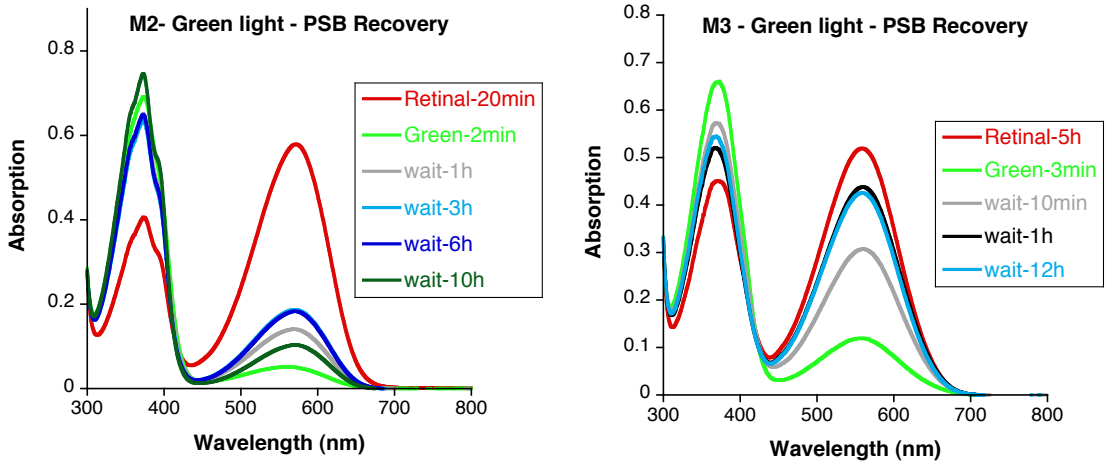


Figure II-22) Mutants M2 and M3 are shined with green light till maximum PSB loss after maximum PSB formation. PSB recovery has been tracked in dark. In 1 hour M3 mutant has significant PSB recovery, while M2 mutant spectrum shows much less PSB recovery.

Crystallization of M3 was initiated at various times after retinal addition. Incubation of the protein for over 24 hours with retinal followed by an additional 20 minutes of green light prior to the initiation of crystallization was also attempted. However, only the *cis*-PSB was seen in the structure, consistent with the extensive stabilization of the *cis* form confered by the F3Q mutation. M3 mutant loses 50% of its PSB in the course of 24 hours. We are not sure what this loss of PSB corresponds to, as the crystal structures show the same retinal geometry (**Figure II-23**).

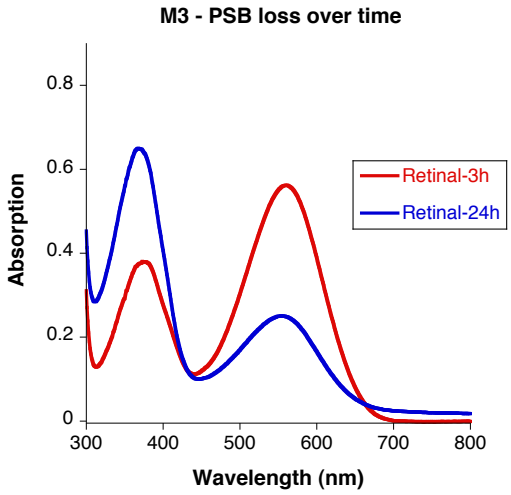


Figure II-23) PSB loss of M3 mutant over time.

In addition, the structure does not show a direct or even water mediated interaction between Gln3 and the iminium, with the distance between the Gln3 side chain and the iminium nitrogen at about 5 Å. It seems that the increased hydrophilicity in the *cis*-iminium region relative to Phe3 containing mutants is sufficient to produce the higher pKa, blue shifted spectrum, and increased stability of the *cis*-iminium (**Figure II-24**).

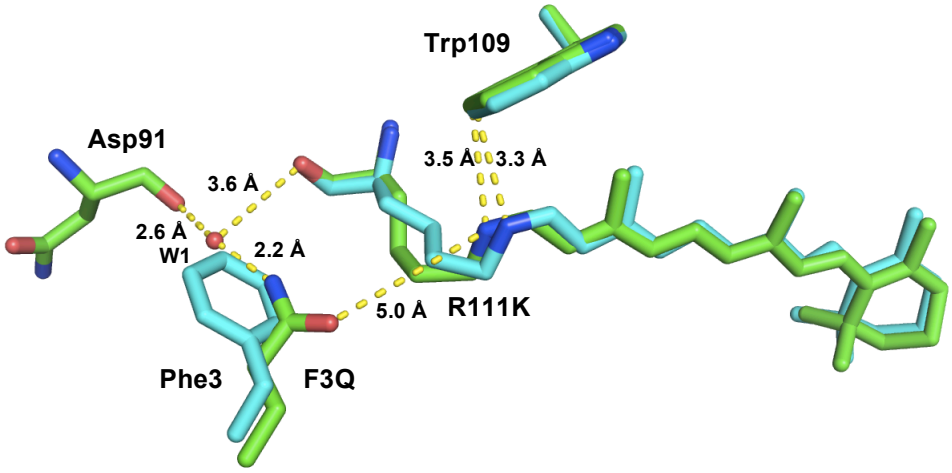


Figure II-24) Structures of retinal-bound M3 (PDB ID: 4YKM) and *cis*-PSB retinal-bound M2 are overlaid, showing the effect of Gln 3 on the environment about the *cis*-PSB, increasing the pK_a by 1 unit and altering the nature of the thermodynamic product.

Figure II-24 shows the F3Q mutant structure overlaid with one of the Phe3 containing structures to show how this position defines the hydrophobicity/hydrophobicity of the iminium environment. The introduced glutamine is fixed in its conformation via a water network to main chain carbonyls of Asp91 and Arg111, which does not exist when Phe3 is not mutated. Unfortunately, M3 crystallizes in a crystal form that cannot be photoisomerized in the crystalline state.

II.3 Isomerization of hCRABPII mutants in solution

As shown above the isomerization of the hCRABPII mutants in the solid crystalline form is very effective, and the electron density maps show the

complete conversion of the two forms to each other. Part of our efforts has been to get the same efficiency in the solution as well. The mutants that show complete conversion in the solid phase sound the reasonable ones to follow in the solution as well. For this purpose mutant M2 was targeted for following the isomerization in the solution. Protein-retinal complex was irradiated with UV light and green light for multiple cycles.

In figure 15 the cycling of the M2 mutant starting at PSB maximum formation is shown, starting by green light irradiation of the solution. It can be seen that less than 50% of the original absorption can be recovered. This indicates the low isomerization yield of the mutant M2 in solution. To solve this problem different strategies were taken. Cycling of mutant M2 was done after the PSB loss, and UV irradiating the solution, which restores a significant amount of the PSB in the first cycle, initiated isomerization. The common theme in figure 15 and (**Figure II-25a**) is that after irradiating the solution the first time with the green light, not much PSB can be recovered by UV irradiating the solution. Therefore, I thought it might be the green light source that is problematic. In the next experiment room light was used as the source of the green light. Exposing the solution after the first time UV irradiation to room light causes the solution to go colorless again in 10 minutes. However, the solution was allowed to stay in the room light for two hours to make sure that all of the protein retinal complex relaxes back to the *trans*-SB form, but it did not help to get more of the PSB back in the next cycles (**Figure II-25b**). In the next experiment the protein was allowed to relax in dark, in the absence of the green light, which proved to be effective (**Figure II-25c**).

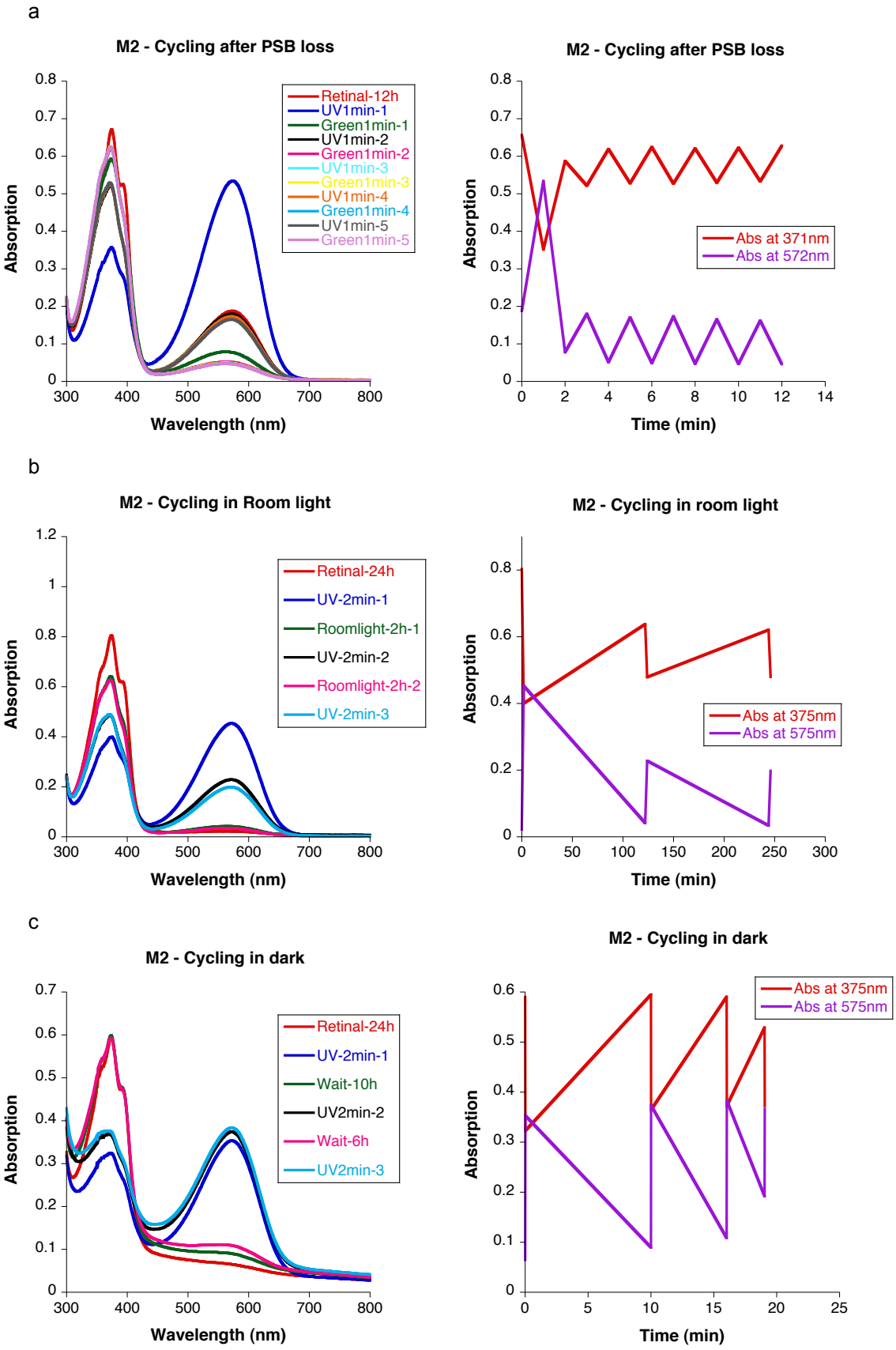


Figure II-25) Cycling of the mutant M2 after PSB loss completion. a) After the first cycle the yield of the switch drops considerably. b) Cycling with room light for visible light irradiation of the solution. c) Cycling in dark with no green light involved.

Since the *trans*-SB form of the protein has a pKa of 3.4 and the *cis*-PSB form has a pKa of 8.2, the cycling was done at pH 5.5 to check if the isomerization process is pH dependent, but the isomerization yield of the protein-retinal complex was found to be independent from the pH (**Figure II-26**).

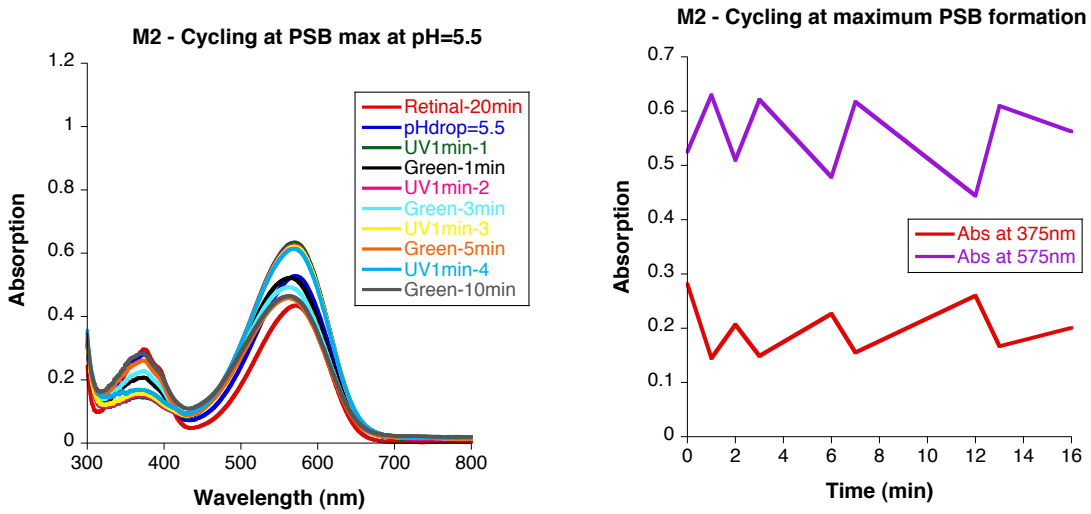


Figure II-26) Cycling of M2 mutant at pH 5.5.

The answer to these observations might be in Figure 22 that indicates mutant M2 cannot recover PSB thermally after Green light irradiation. Since the *trans*-SB is the thermodynamically stable form of the protein, after green light irradiation, part of the SB does not isomerize because it is not thermodynamically favored to go to the kinetic *cis*-PSB form.

With the same hypothesis, it is expected that mutant M3 will be able to isomerize in the solution more efficiently, because Figure 22 shows that after green light irradiation a lot of the PSB can be recovered in short time thermally proving the *cis*-PSB to be the thermodynamic form of the protein. Therefore, unlike mutant M2, UV irradiation will favor the formation of the

thermodynamic form of the protein. Incubation of the M3 with retinal till maximum PSB formation followed by green light irradiation and then UV irradiation for multiple cycles shows that more than 50 % of the absorption of the PSB can be recovered in agreement with the rationalization described (**Figure II-27**).

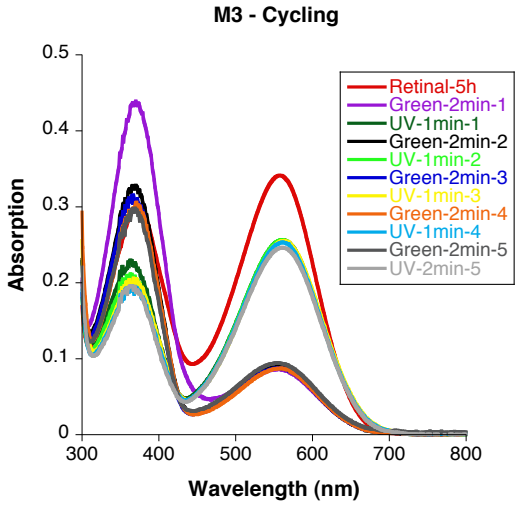


Figure II-27) Cycling of the mutant M3 in solution. More than 50% of the PSB absorption is recovered in each cycle.

II.4 The mystery of colorful crystals of M2 mutant after PSB loss

As mentioned above in order to trap the thermodynamic form of the hCRABPII second generation mutants, M1 and M2 mutants were crystallized after incubation with four equivalents of retinal and 24 hours wait period for complete PSB loss. Interestingly, the crystals that are grown from the pale yellowish solution with free retinal color, which comes from the excess retinal in crystallization solution, turned to be colorful with an intense blue color for mutant M2 at pH 5 and 6 and light blue color at pH 6.5. (**Figure II-28**).

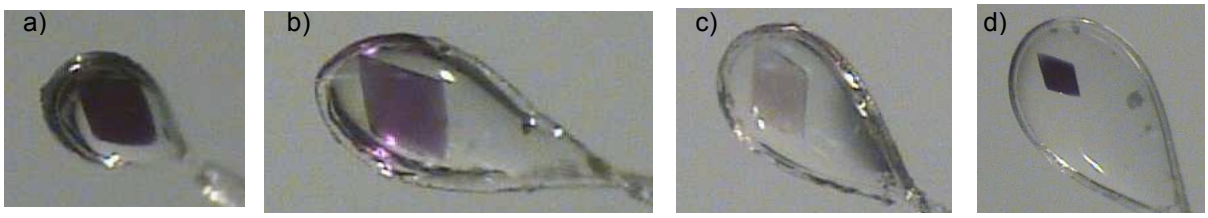


Figure II-28) a) M2 SB-crystal grown at pH 5 b) M2 SB crystal grown at pH 6.5 c) M2 SB-Soaked crystal to pH 7.5 d) M2-UV irradiated crystal after soak for 30 minutes.

For mutant M1 the color of the crystals is pale red at pH 6.0. The reason that crystals are soaked before freezing is to make sure the PSB remained in the crystals convert to SB for a fair comparison between the SB of the 15-*trans* retinal and PSB of the 15-*cis* retinal.

However, with a low pKa of 3.4 it is not expected for the crystals to be colorful at pHs of 5, 6 or 6.5. To find out the reason the pKa titration of the protein was done in 50 mM malonate, which is the buffer in crystallization condition, as a control for condition, but it did not give a different pKa than that in phosphate buffer (**Figure II-29**).

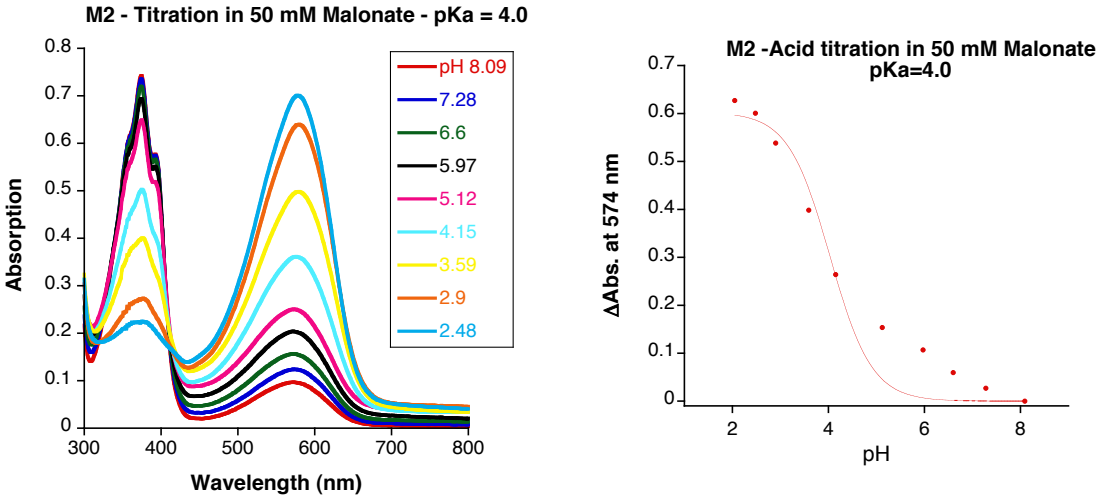


Figure II-29) pKa titration of M2 protein in 50 mM Malonate.

The next experiment was that after the protein was incubated for 24 hours, the pH of the protein solution was dropped with 50 mM malonate at pHs of 5 and 6 and the protein was allowed to sit for 24 hours (the crystallization time). The same experiment was repeated with addition of 12% PEG3350 which is the precipitant reagent in the crystallization condition (**Figure II-30**).

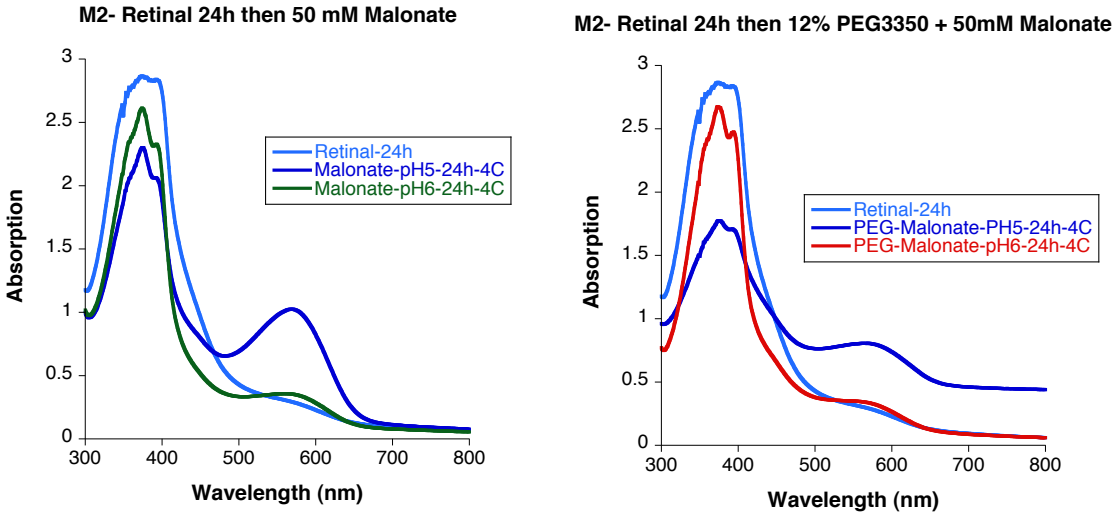


Figure II-30) M2 protein was incubated with 4 equivalents retinal and then mixed with the crystallization condition. The protein was allowed to sit in the cold room for 24h.

The result of this experiment shows that at the time that the protein forms crystals, there is considerable PSB at pH=5 and a little at pH=6, which is probably the reason why the crystals have color. At pH=5 and 12% precipitant reagent (PEG 3350) and 24 hours in the cold room a lot of protein precipitation is observed as well along with the blue color of the solution.

The UV-Vis spectrum of the crystals after the soak and after the UV irradiation was also collected. The comparison of the spectrums shows marginal amounts of PSB in the soaked crystals and considerable amounts of the PSB in the UV irradiated crystals. This provides more evidence that our

comparison of the *trans*-SB retinal in colorless crystals and *cis*-PSB in colorful crystals is valid and relevant (**Figure II-31**).

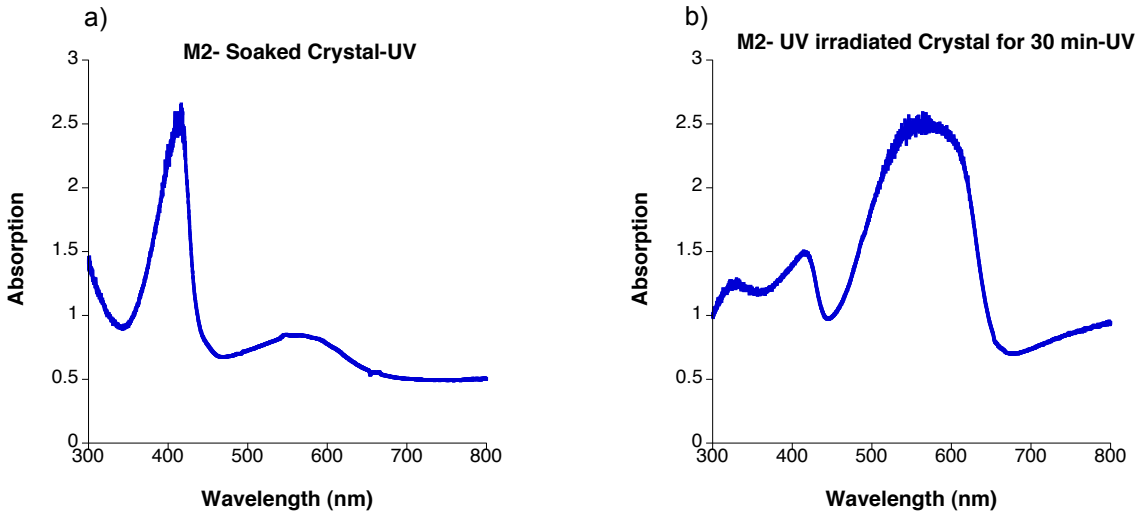


Figure II-31) a) The UV-Vis spectrum of the soaked crystal of M2 mutant b) The UV-Vis Spectrum of the UV irradiated crystal for 30 minues after soak.

The crystal UV-Vis spectrum of mutants M3 and M11 was also measured, which matches with that of solution Table 5 (**Figure II-32**).

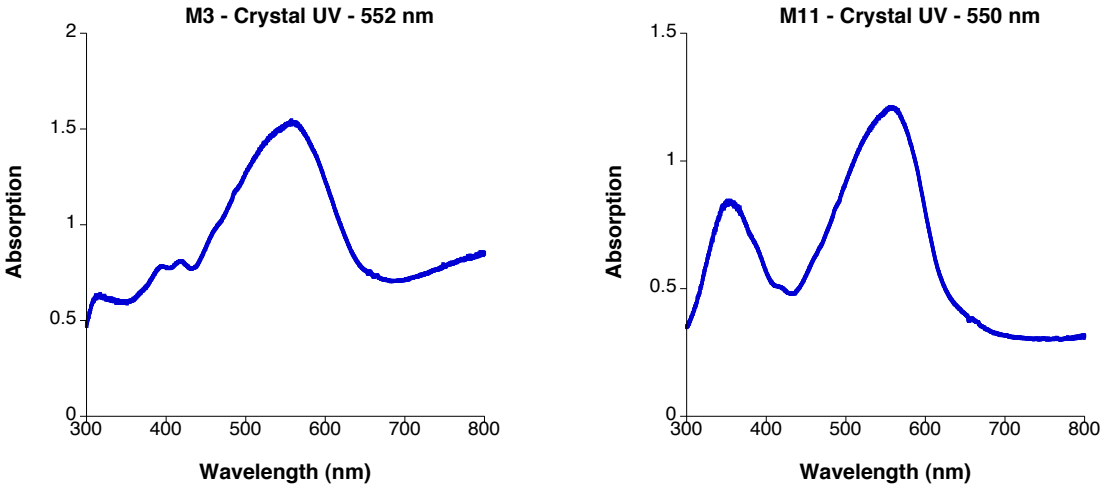


Figure II-32) Crystal UV-Vis spectrum of mutans: a) M3 b) M11.

II.5 hCRBPII isomerization solution and structural studies

Along our efforts to develop a photoisomerizing protein from hCRABPII and providing structural data for the switch, we have been aiming for the same in hCRBPII protein as well. Developing a photoswichable hCRBPII protein is important as this protein is more sutiable for in-vivo imaging than hCRABPII. Developing a photoswichable hCRBPII protein in-vitro can pave the way for the development of a photoswitchable fluorescent protein in-vivo as well. In the anaerobic condition that proteins like GFP and its derivatives does not work this can be of super importance (14-16).

hCRBPII proteins also lose their PSB over time. Like hCRABPII proteins in the presence of Gln4 this loss is not complete and up to 50% of the protein stays as PSB even after long incubations with retinal. However, when the Gln4 is mutated the PSB loss goes to completion (**Figure II-33**).

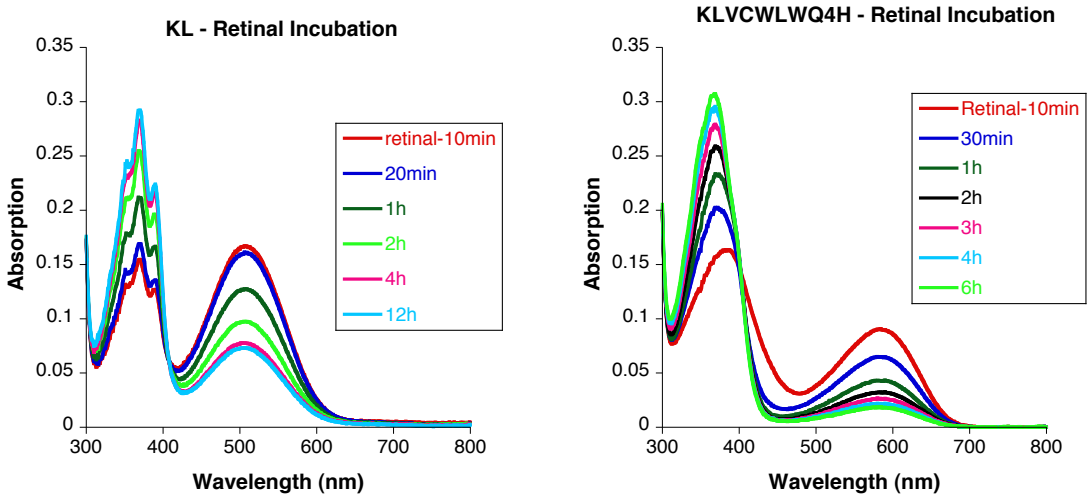


Figure II-33) Retinal incubation of the hCRBPII mutants: Left: Q108K:K40L, Right: Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H.

The mutants of hCRBPII respond to UV light and green light irradiation as well. In the presence of the Gln4 the yield of the switch is much better than when this residue is mutated (**Figure II-34** – with Dr.Tetyana Berbasova's permission). When this residue is mutated, during photo switch cycles a lot of the protein-retinal complex stays as SB and does not participate in the SB-PSB switch. This might be because of the same reasoning that discussed for hCRABPII mutants on F3Q mutation that switches the thermodynamic form of the protein to 15-*cis* retinal ratherthan 15-*trans* retinal in the absence of this mutaiton. With Gln4 the 15-*cis* retinal PSB is the thermodynamic form of the protein-retinal complex and will be recovered immediately upon UV light irradiation, but when this residue is mutated the thermodynamic form of the retinal-protein complex is the 15-*trans* retinal SB, which might not respond to the UV irradiation to go to the kinetic form which is 15-*cis* retinal PSB.

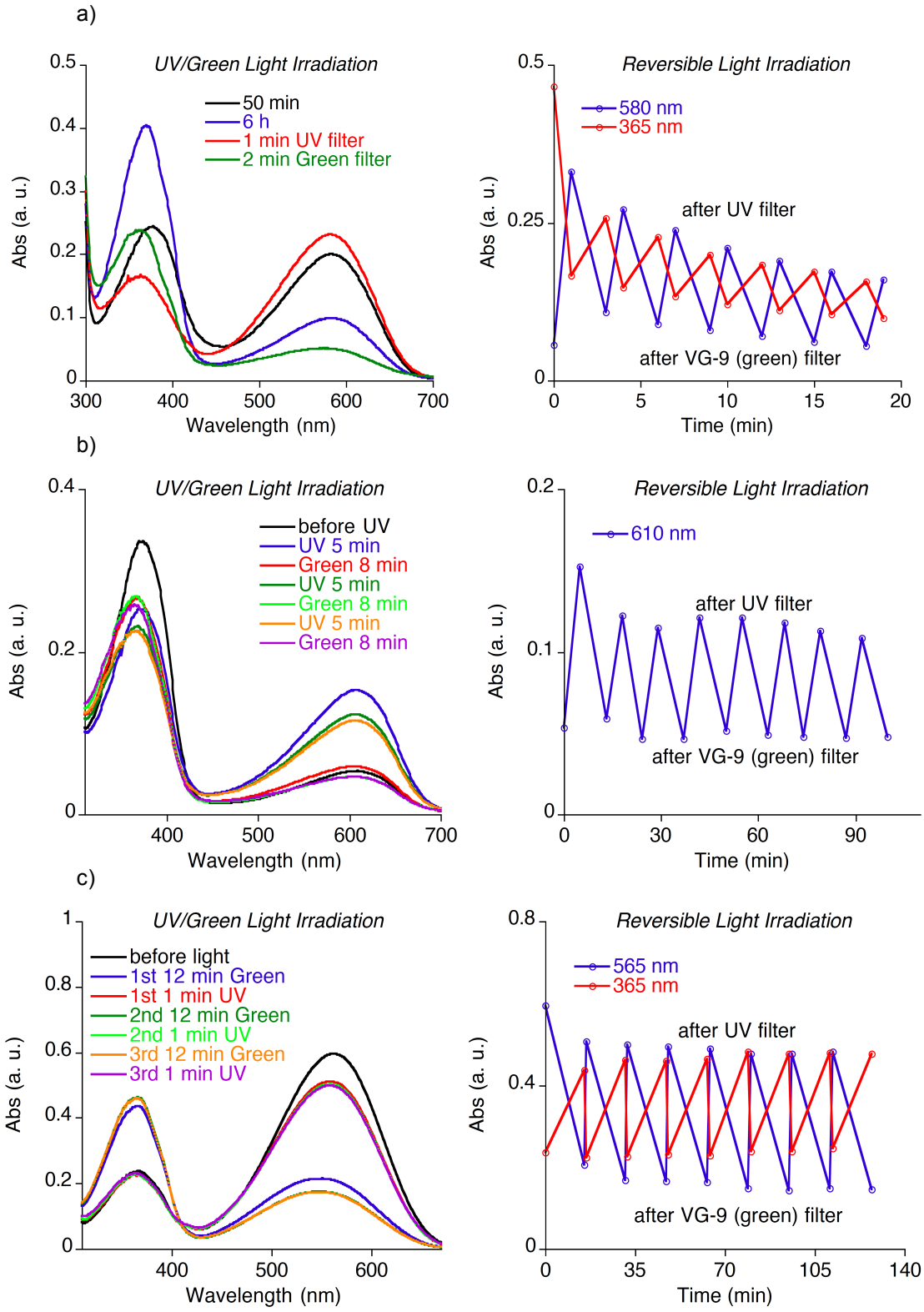


Figure II-34) UV-Green light cycling of hCRBPII mutants for: a) Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H b) Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4A c) Q108K:K40L:T51V:R58F. With Dr. Tetyana Berbosva's permission

Our extensive structural data on the hCRBPII mutant series has proven that whenever the Gln4 residue is present, the iminium geometry is *cis*-PSB retinal and whenever this residue is mutated it is *trans*-SB retinal. Gln4 stabilizes the *cis*-PSB retinal via a water molecule that stays in between the nitrogen atom of Schiff base and the oxygen of the amide group of Gln4. Therefore, the most possible structural explanation for this event is the isomerization of the iminium bond (**Figure II-12** and **Figure II-21**).

Our first attempts to structurally trap the two forms of the protein in the same mutant started with allowing the mutants with the Gln4 residue to lose their PSB to the maximum amount and then irradiating the solution with green light and then crystallizing the protein. Our hypothesis was to connect the PSB loss of these mutants to the iminium isomerization. The data collected on the mutants crystallized this way disapprovingly showed a *cis*-PSB with no isomerization along the polyene (**Figure II-35**). Also the green light irradiation indicates the PSB loss in solution, it is possible that for this mutants since the *cis*-PSB retinal is the thermodynamic form, the SB form that is made from green light irradiation recovers back to the *cis*-PSB retinal before the crystallization of the protein or during crystallization.

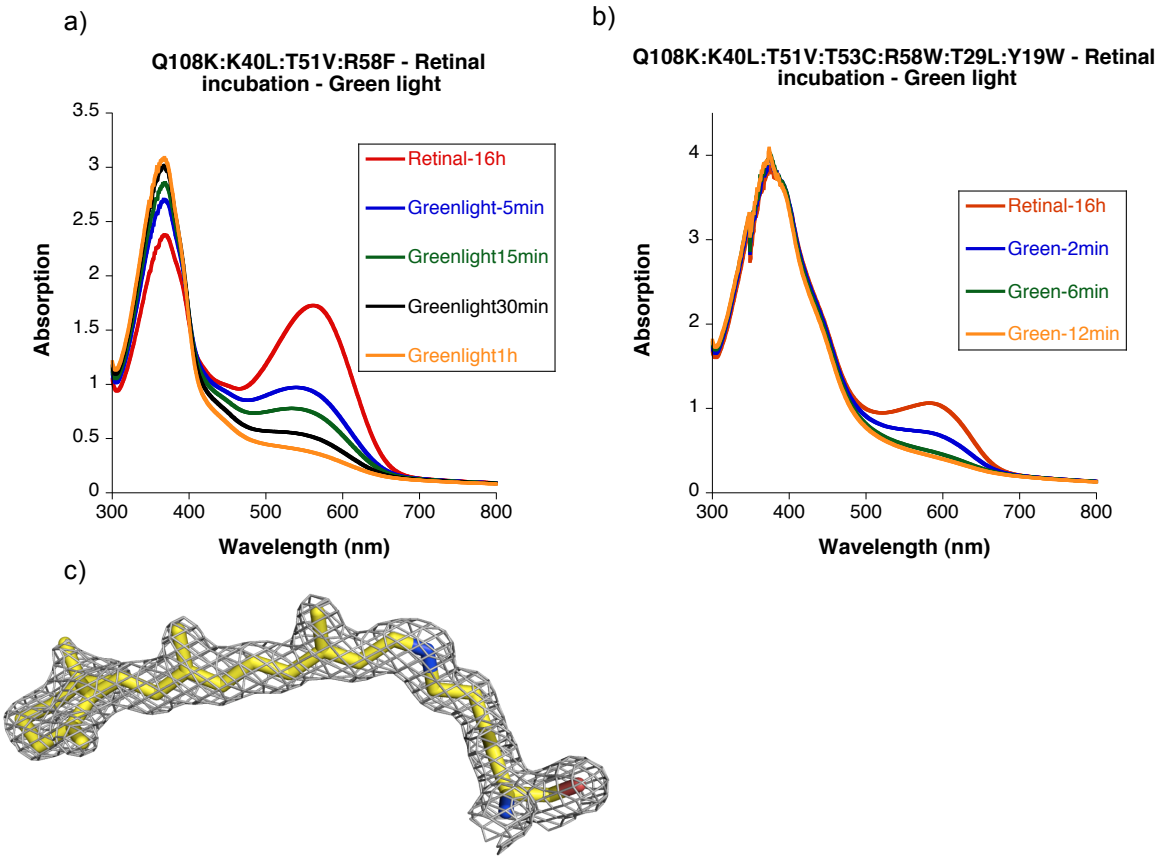


Figure II-35) Retinal incubation followed by green light irradiation for: a) Q108K:K40L:T51V:R58F mutant of hCRBPII and b) Q108K:K40L:T51V:T53C:R58W:T29L:Y19W mutant of hCRBPII. c) Electron density map of the crystal obtained from Q108K:K40L:T51V:R58F mutant contoured at 1σ shows a *cis*-iminium and not a *trans*-iminium.

- Q108K:K40L:T51V:T53C:R58W:T29L:Y19W mutant crystal structure also shows a *cis*-iminium, but it only shows the density for the first five carbons close to the iminium.

The overlay structure of the same mutants that are crystallized right after PSB maximum formation and after maximum PSB loss shows the critical residues inside the binding pocket in the same place. However, the second angular methyl of the retinal polyene tail rotates around 45 degrees. Whether this conformational change of the polyene is related to the PSB loss of the protein or not, it does not change the λ_{max} of the protein as there is no shift seen in the UV-Vis spectrums over time (**Figure II-33** and **Figure II-36**).

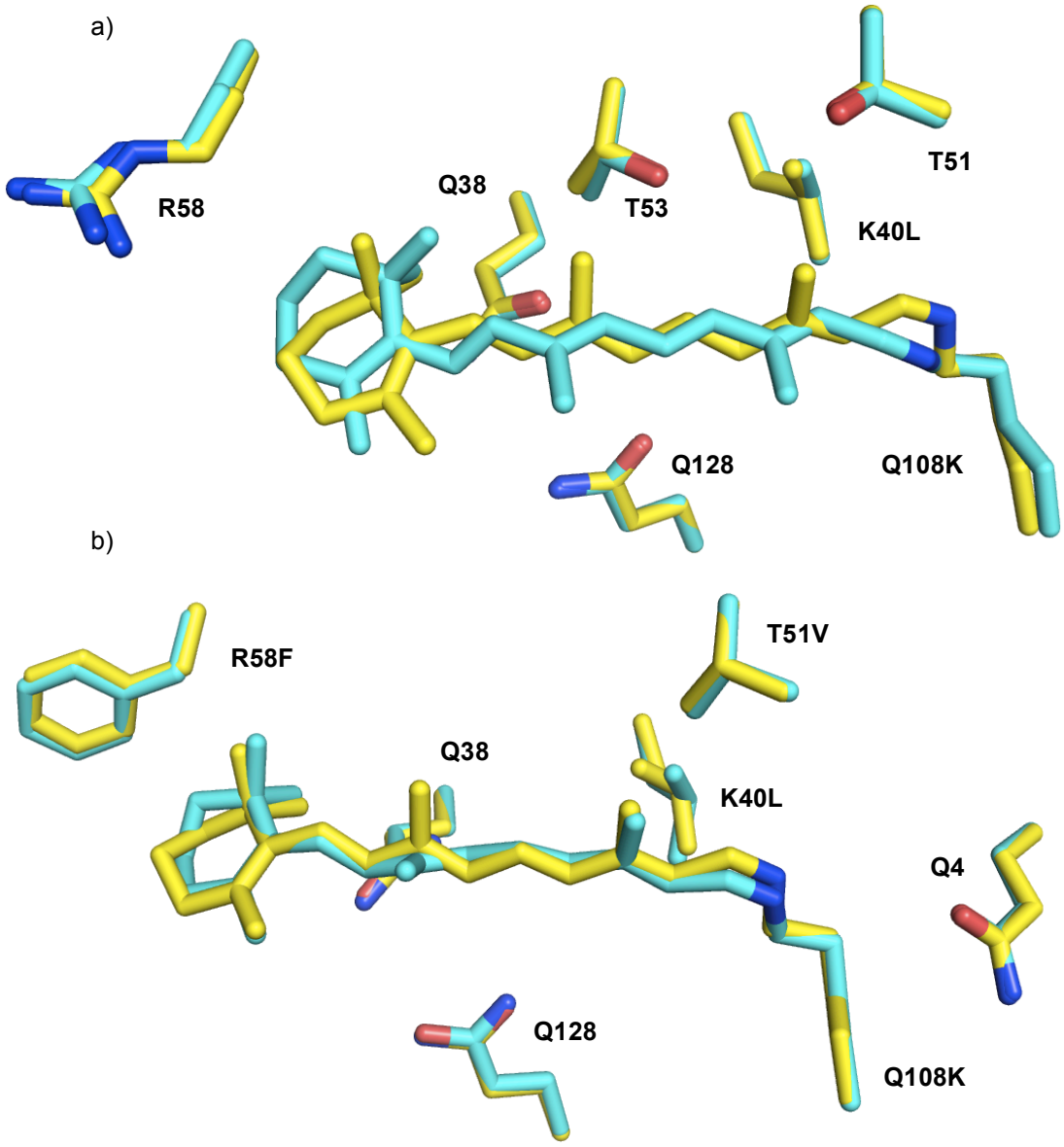


Figure II-36) a) Overlay structure of the mutant Q108K:K40L after 20 minutes incubation (Cyan, PDB ID: 4EXZ) and after 24 hours incubation (Yellow, PDB ID:4RUU). b) Overlay structure of the mutant Q108K:K40L:T51V:R58F after 20 minutes incubation (Cyan) and after 16 hours incubation and 1 hour green light (Yellow).

Being unable to correlate the PSB loss of the hCRBPII mutants with retinal isomerization, we decided to find a way to prove the conversion of the *trans*-SB to *cis*-PSB. As mentioned above mutants with Gln4 mutation show a *trans*-SB. Therefore, we decided to crystallize these mutants after complete PSB loss, soak and UV irradiate the obtained crystals and then freeze them and collect crystallography data. We were hypothesizing that the UV

irradiation will isomerize the iminium back to *cis.* This experiment was done for the Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H mutant. The crystals have a pale grey color, and after the soak they lose the color, by UV irradiation some of the color is returned back. Although the color change after the UV irradiation was recognizable, the data collection on these crystals shows the *trans*-SB retinal again.

The structural isomerization experiments done with hCRBPII prove again that the Gln4 in hCRBPII and F3Q mutation in hCRABPII thermally stabilizes the *cis*-PSB that no isomerization is seen overtime despite the partial PSB loss overtime. In hCRABPII for these mutants the structural data shows no retinal isomerization or any conformational changes. In hCRBPII although so far no isomerization is recognizable for these mutants, but the position of angular methyls change overtime (**Figure II-36**).

The PSB loss can be seen in hCRBPII mutants with Q4X mutation as well, but so far we have not been able to characterize the initial kinetic PSB form that is made.

The fact that all of the hCRBPII mutants with Gln4 conserved show a 15-*cis* retinal and all of the mutants with Q4X mutation show a 15-*trans* retinal suggests the isomerization of the imine double bond as a possible explanation for the PSB loss over time, but this isomerization has not been structurally characterized in the same mutant of hCRBPII.

The hCRBPII retinal isomerization can be tracked in solution. Different mutants of this protein with and without Gln4 mutations show the SB-PSB conversion upon irradiation with green and UV light. However, so far we have not been able to structurally elucidate the light irradiated forms of the protein.

II.6 Conclusion

Using a rational design strategy based on the Cellular Retinoic Acid Binding Protein II scaffold, we have created a series of rhodopsin family mimics that reproduce the basic chemistry in these systems, in that photoisomerization switches the retinal between two distinct pK_a regimes. We have further demonstrated the ability to independently control both pK_as, and alter the thermal product from the low to the high pK_a moiety. Both forms have been unambigously identified from high resolution crystal structures, and photo-isomerization in the crystalline state was also demonstrated, with quantitative conversion between each form shown at atomic resolution. Besides photoswiching, the clean and complete dark isomerization is unique of this protein.

This represents a retinal-based photoswitchable protein that has a variety of potential applications from a rhodopsin photoswitch model to a fluorescent quencher for further development of photoswitchable imaging techniques.

The ability to control the basic chemistry of each form will be an exceedingly valuable tool in these applications. It can be used as a valuable tool for time resolved spectroscopy and crystallography techniques to unveil the intermediates of the retinal isomerization in this system.

REFERENCES

REFERENCES

1. Bushue N & Wan YJ (2010) Retinoid pathway and cancer therapeutics. *Advanced drug delivery reviews* 62(13):1285-1298.

2. Noy N (2000) Retinoid-binding proteins: mediators of retinoid action. *The Biochemical journal* 348 Pt 3:481-495.

3. Vaezeslami S, Mathes E, Vasilelou C, Borhan B, & Geiger JH (2006) The structure of apo-wild-type cellular retinoic acid binding protein II at 1.4 A and its relationship to ligand binding and nuclear translocation. *Journal of molecular biology* 363(3):687-701.

4. Vasileiou C, *et al.* (2009) Dissection of the critical binding determinants of cellular retinoic acid binding protein II by mutagenesis and fluorescence binding assay. *Proteins-Structure Function and Bioinformatics* 76(2):281-290.

5. Dong D, Ruuska SE, Levinthal DJ, & Noy N (1999) Distinct roles for cellular retinoic acid-binding proteins I and II in regulating signaling by retinoic acid. *The Journal of biological chemistry* 274(34):23695-23698.

6. Budhu AS & Noy N (2002) Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest. *Molecular and cellular biology* 22(8):2632-2641.

7. Crist RM, Vasileiou C, Rabago-Smith M, Geiger JH, & Borhan B (2006) Engineering a rhodopsin protein mimic. *Journal of the American Chemical Society* 128(14):4522-4523.

8. Vasileiou C, *et al.* (2007) Protein design: Reengineering Cellular Retinoic Acid Binding Protein II into a rhodopsin protein mimic. *Journal of the American Chemical Society* 129(19):6140-6148.

9. Berbasova T, *et al.* (2013) Rational Design of a Colorimetric pH Sensor from a Soluble Retinoic Acid Chaperone. *Journal of the American Chemical Society* 135(43):16111-16119.

10. Yapici I, *et al.* (2015) "Turn-on" protein fluorescence: in situ formation of cyanine dyes. *Journal of the American Chemical Society* 137(3):1073-1080.

11. Ernst OP, *et al.* (2014) Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. *Chemical reviews* 114(1):126-163.

12. Lee KSS, *et al.* (2012) Probing Wavelength Regulation with an Engineered Rhodopsin Mimic and a C15-Retinal Analogue. *Chempluschem* 77(4):273-276.

13. Wang WJ*, et al.* (2012) Tuning the Electronic Absorption of Protein-Embedded All-trans-Retinal. *Science* 338(6112):1340-1343.

14. Tsien RY (1998) The green fluorescent protein. *Annual review of biochemistry* 67:509-544.

15. Chudakov DM, Matz MV, Lukyanov S, & Lukyanov KA (2010) Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiol Rev* 90(3):1103-1163.

16. Nienhaus K & Nienhaus GU (2014) Fluorescent proteins for live-cell imaging with super-resolution. *Chem Soc Rev* 43(4):1088-1106.

Chapter III: Theories of Wavelength Regulation

III.1 Introduction

In chapter 1 the mechanism of vision was explained. Four different types of Rhodopsin proteins all work with the same chromophore 11-*cis* retinal to absorb light. One of the long standing questions that in the last 50 years scientists have been trying to answer is the wavelength regulation question: how these proteins with the same chromophore absorb different wavelengths of light. As mentioned Rhodopsins work via the formation of an iminium Protonated Schiff Base (PSB). The 11-*cis* retinal chromophore by itself absorbs at 380 nm. The 11-*cis* retinal Schiff Base (SB), which is not protonated, absorbs at 365 nm and upon protonation it will absorb at 440 nm in ethanol solution. The absorption of the n-butyl retinal iminium PSB can be tuned up to 500 nm depending on the solvent (**Figure III-1**) (1, 2).

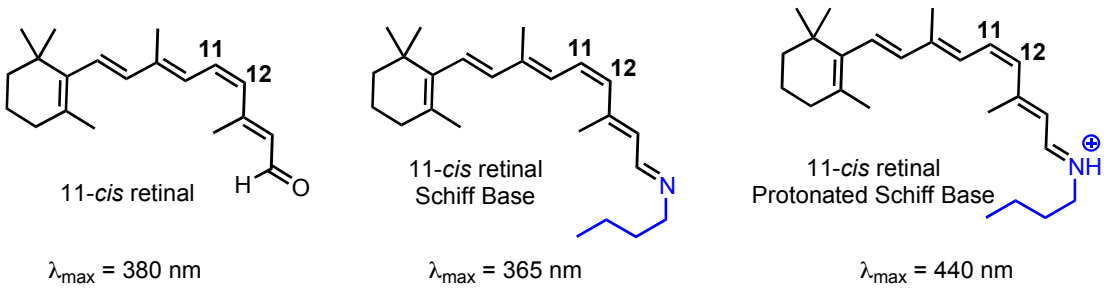


Figure III-1) structure and absorption of 11-*cis* retinal, its n-butyl amine Schiff Base and Protonated Schiff base.

This can span the absorption of the Blue and Rod Rhodopsin, but not the Green and Red Rhodopsin, which is at 530 and 560 nm respectively. The difference of the absorption of retinal PSB in protein compared to the ethanol solution (440 nm) is referred to “Opsin shift”, which is attributed to the interaction of the retinal PSB with protein environment.

The fact that none of the colored opsins is crystallized so far has hampered the development of a solid answer to the wavelength regulation question. However, extensive studies that have been done in the last few decades have elucidated a lot of the governing principles in wavelength regulation. The contributing effects have been attributed to:

(1) Point charges or dipole moments from the charged or polar residues.

(2) Conformational geometry of the retinal

(3) Polar water networks

(4) Ground state (S0)-Excited state (S1) dynamics of the retinal

(5) Excitonic coupling to Aromatic Residues

III.2 The Point Charge Theory

In brief, the point charge theory emphasizes the stabilization of the positive charge of the iminium via negative charge or negative dipole moment directly interacting with the iminium nitrogen or along the retinal polyene. If the negative charge or negative dipole moment is close to the iminium or toward the beginning of the polyene the protein is blue shifted. On the other hand, when it is toward the end of the polyene the positive charge is more delocalized and the protein is red shifted (**Figure III-2**).

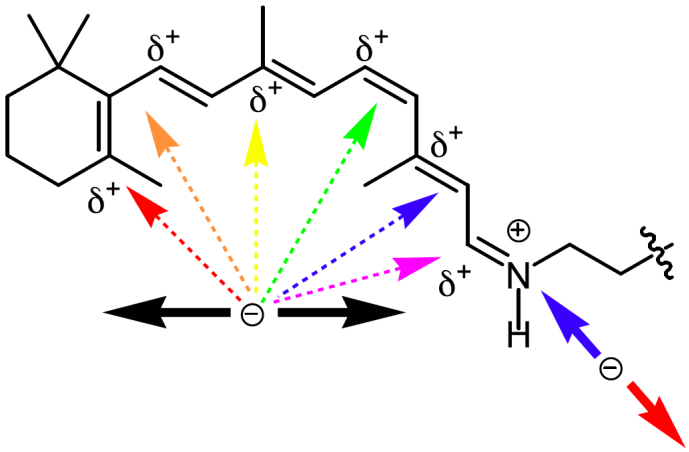


Figure III-2) The position of the putative negative charge or negative dipole moment determines the absorption of the protein.

The formation of a covalent bond between a lysine residue of rhodopsin and retinal was confirmed by different studies via reductive amination and proteolytic fragmentation decades before the crystal structure of the bovine rhodopsin (3-5). The protonation of the Schiff Base was confirmed by resonance raman spectroscopy (6), and NMR studies by Nakanishi and others proved the presence of a negative charge adjacent to the PSB positive charge to be the reason for the stabilization of the positive charge of the PSB (7-9). The pKa titration studies of retinal analogs postulated the pKa of the bovine Rhodopsin to be around 16. The pKa of 16 is nine units above the pKa of a PSB in solution (10). Oprian mutated all of the Glu and Asp residues buried in Rhodopsin to figure out that Glu113 is the only important residue for the wavelength and pKa of the Rhodopsin. E113Q mutation dropped the pKa of the protein to about 6 and blue shifted the protein to 380 nm where the SB absorbs (11). Mutating this counter anion to different residues can maximally red shift the protein by 30 nm depending on the solute anion. The red shift of the protein is expected as the removal of the negative charge of the counter anion will increase the delocalization of the charge on the retinal chromophore (12). Studying the model compounds of retinal showed that both the angle and the distance of the counter anion is important in the absorption and pKa of the retinal PSB in solution (13).

After the crystal structure of the Bovine Rhodopsin, Glu113 was found 3.45 Å away from the PSB nitrogen (14). Palczewski modeled the structures of the rest of the colored pigments based on the structure of the bovine rhodopsin (15). Besides the importance of the counter anion effect in the pKa and wavelength of the rhodopsins, other studies attempting to convert the

absorption of the colored opsins to one another shed more light into the point-charge model theory.

Mutation of three residues in Green Rhodopsin including A180S, F277Y and A285T provides a pigment like Red Rhodopsin (16, 17). These residues are all toward the ionone ring of the chromophore. Introducing these polar groups will increase the negative dipole moment toward the end of the retinal polyene and causes the red shift observed. However, for converting the Red opsin to Green there are more mutations needed which includes S116Y, S180A, I230T, A233S, Y277F, T285A and Y309F, but the majority of the spectral shift is related to the mutations at position 180, 277 and 285 (18). In Rhodopsin F261Y/A269T mutant, equivalent of F277Y/A285T in Green opsin, red shifts by 20 nm to absorb at 520 nm, while the triple mutant does not provide a detectable PSB (**Figure III-3**) (19).

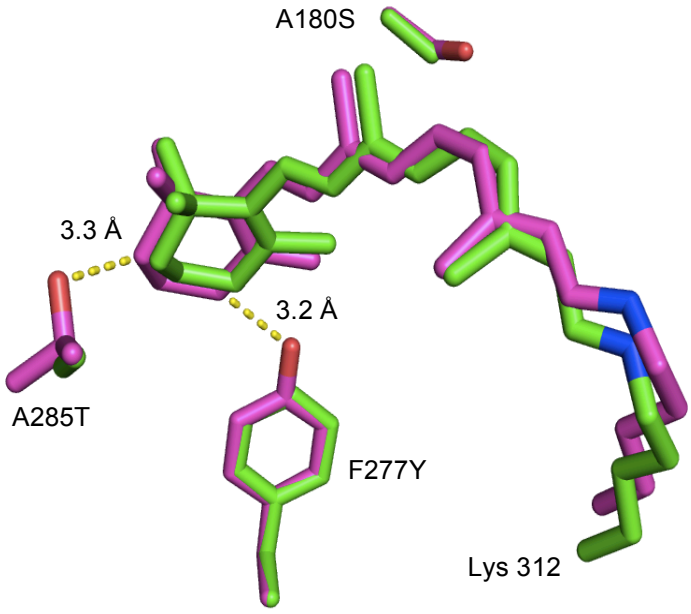


Figure III-3) Overlay structure of the models of the Green opsin (Green - PDB ID: 1KPW) and Red opsin (Magenta - PDB ID: 1KPX). Mutating the highlighted residues in Green to equivalent residues in Red will provide a Red absorbing like pigment.

In another study the important residues in blue shifting the rhodopsin from 500 nm to 438 nm were characterized. These included nine mutations of M86L/G90S on Helix 2, A117G/E122L/A124T on Helix 3, W265Y on Helix 6 and A292S/A295S/A299C on Helix 7 (20). These mutations are mainly the introduction of a polar group toward the PSB and E122L mutation removes a negative charge close to the ionone ring of the chromophore (**Figure III-4**).

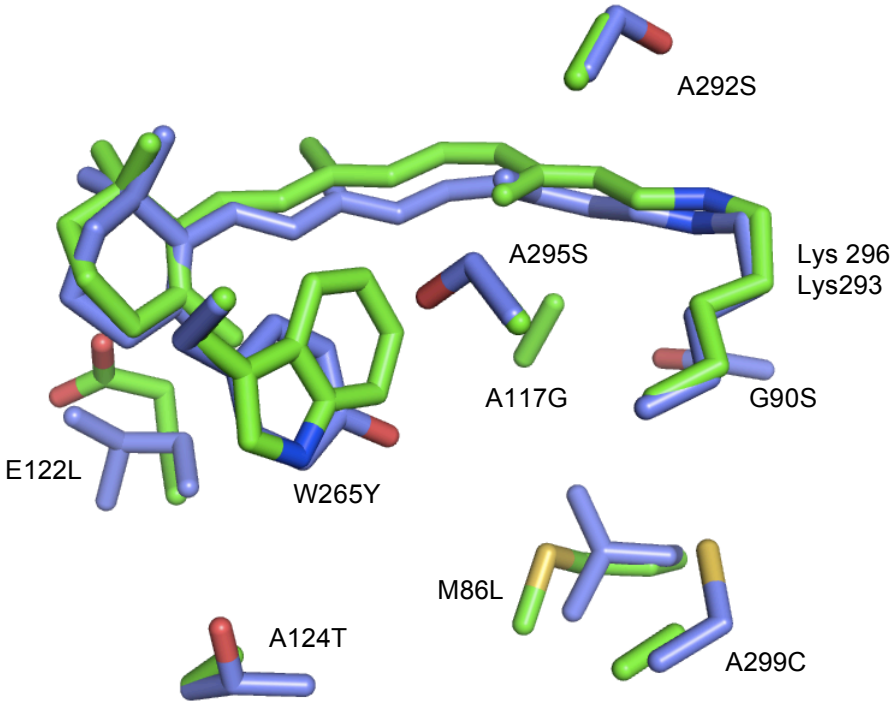


Figure III-4) Overlay structure of the Rhodopsin (Green - PDB ID: 1F88) and the model of Blue opsin (Blue - PDB ID: 1KPN). The mutation of the highlighted residues on Rhodopsin to equivalent residues of Blue provides a pigment absorbing at 438 nm. The residue numbering is that of Rhodopsin.

In another study the triple mutant of T118A/E122D/A292S was shown to be a functional mutant with 50 % rate of transducing activation compared to the wild type protein absorbing at 453 nm (21).

The difference of the absorption maxima between the Green opsin and Rhodopsin has been attributed to anion binding sites near PSB. Cl^- and Br^- has been found as the two anions that can cause these spectral shifts. The

interacting residues are identified to be His 197 and Lys 200. These residues are conserved in Green and Red opsins and absent in Rhodopsin and Blue Opsin. The increased hydrogen bonding between the PSB and waters in the absence of the chloride causes the blue shift observed in Rhodopsin and Blue opsin. The mutation of these residues in Green opsin can provide pigments that absorb at 500 nm like Rhodopsin. The mutants H197E/K200Q, H197E and K200Q absorb respectively at 500, 500 and 523 nm (**Figure III-5**) (22-24). From the studies that are done in converting the absorption maxima of the opsin retinal pigments to one another, it is clear that the polarized groups along the polyene, especially hydroxyl containing amino acids, play an important role in the wavelength regulation of these pigments. Positively charged amino acids, by providing chloride binding sites, and Glutamic acid residues, by introducing negative charge along the polyene or by attracting waters, can also regulate the wavelength absorbed by these proteins.

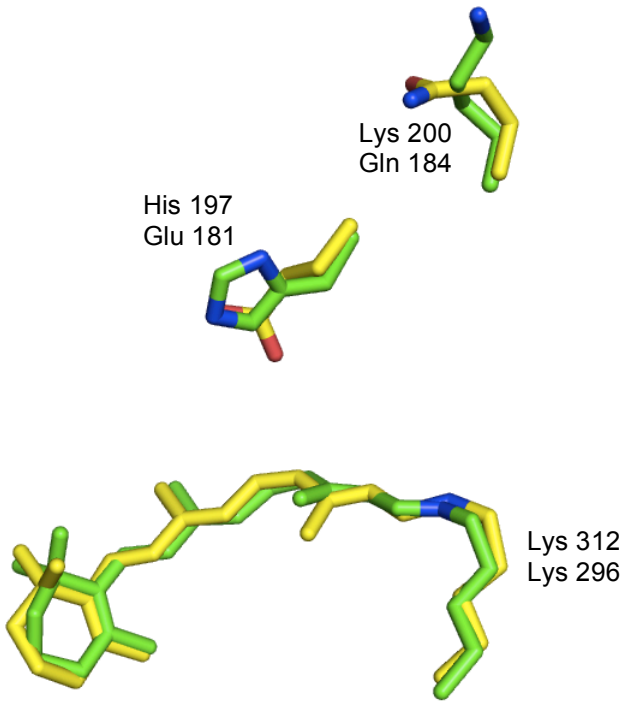


Figure III-5) The overlay structure of the model of Green opsin (Green-PDB ID: 1KPW) and Rhodopsin (Yellow-PDB ID: 1F88).

Quantum mechanical and molecular mechanical calculations on some of the mutants described above was done in 2011 to rationalize how the electronics of the ground state and excited state of the retinal PSB changes upon the mutation of residues toward the PSB or toward the ionone ring in rhodopsin (**Figure III-6**) (25). The mutations that were studied are:

(1) E113D, which is changing the counter anion to shorter Aspartic Acid residue that was shown to cause 10 nm red shift by Sakmar (12).

(2) A269T, which causes 15 nm red shift, and is toward the ionone ring (19).

(3) T118A, which causes 16 nm red shift which is toward the middle of the retinal polyene (21).

(4) A292S, which causes 9 nm blue shift and is toward the iminium nitrogen PSB (20).

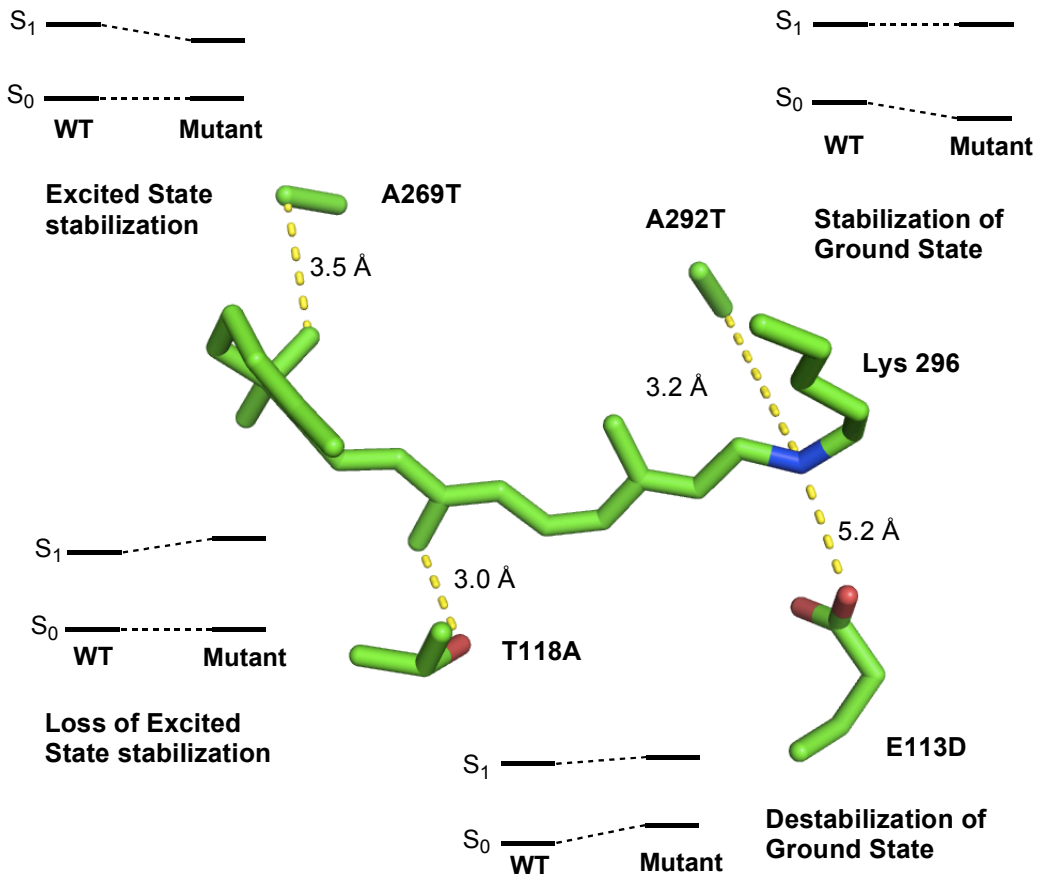


Figure III-6) The effect of mutating or introducing hydroxyl containing residues or charge residues along the retinal polyene of Rhodopsin.

In addition, Mathies and Sakmar studied resonance Raman vibrational spectra of the rhodopsin visual pigments and concluded that the main factor for the opsin shift in color vision is the interaction of the dipolar amino acids with the ground state and excited state charge distributions of the chromophore (26).

Electrostatic potential calculation for three bacterial rhodopsins including sensory rhodopsin II (SRII) 486 nm, bacteriorhodopsin (BR) 552 nm and halorhodopsin (HR) 576 nm was done for the assessment of the distribution of the charge along the retinal polyene and its correlation to the absorption of these proteins. Since the chromophore geometry is the same in these bacterial rhodopsins, the absorption difference has been attributed to changes in the distribution of the charge and dipole moments from the protein. For the sensory rhodopsin its absorption is close to 60-90 nm more blue shifted compared to the other two bacterial rhodopsins. These calculations show the most negative electrostatic potential near the PSB region of the retinal while for the other two this negative electrostatic potential is more localized toward the ionone ring. Bacteriorhodopsin and halorhodopsin differences are less compared to the sensory rhodopsin II. These calculations emphasize one more time the importance of where the negative dipole moment is set along the retinal polyene. For these theoretical calculations the electrostatic potential of the retinal has set to zero and the one projected from the protein onto the surface of the retinal is calculated (**Figure III-7**) (27).

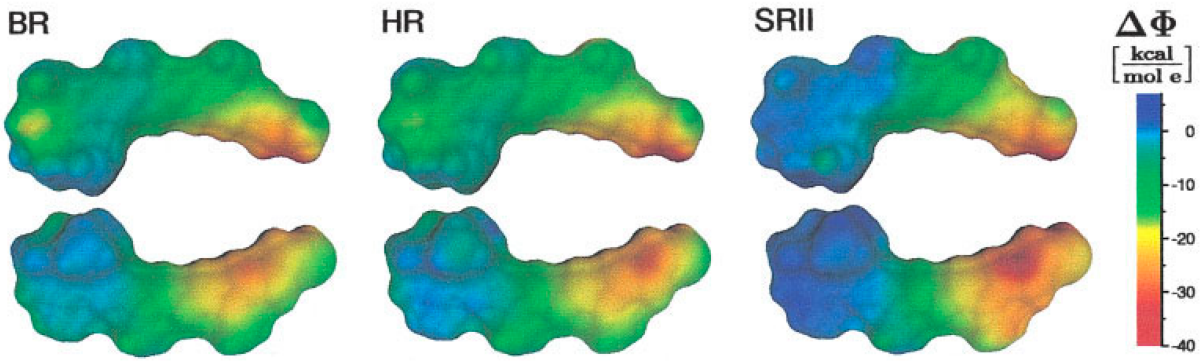


Figure III-7) Electrostatic potential of BR, HR and SRII at the van der Waals surface of the retinal. The magnitude of the potential is color coded. PDB IDs used are: BR (PDB ID: 1QHJ), HR (PDB ID: 1E12) and SRII (PDB ID: 1JGJ).

Dr.Lee in Borhan lab did the same calculation for the Rhodopsin and three other color opsins. The same ideas rationalized for the wavelength regulation concluded so far can be one more time taken from these calculations as well (**Figure III-8**) (28).

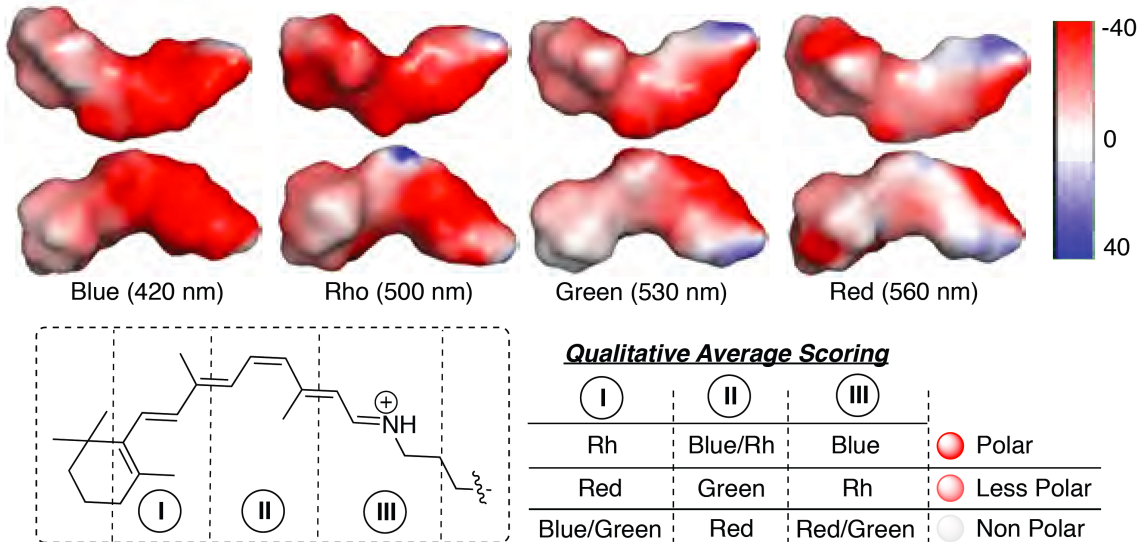


Figure III-8) Electrostatic potential calculation for Rhodopsin, blue, green and red opsin.

Among the bacterial rhodopsins sensory rhodopsinII (SRII) absorbs ~ 500 nm, while others absorb 560-590 nm. SRII is the light pigment that mediates the phototaxis in bacteria to avoid harmful blue light. In an effort to understand the blue shift of SRII, the sequence of Bacteriorhodopsin (BR) and SRII where

compared and three residues where found to be important in the wavelength shift. These three residues are conserved in the other bacterial rhodopsins, while they are not the same in SRII. They include Val108, Gly130 and Thr 204, which the equivalent residues in BR are Methionine, Serine and Ala. The single mutants of V108M, G130S and T204A were absorbing respectively at 502, 503 and 508 nm. The triple mutant V108M/G130S/T204A provided a pigment, which absorbed at 515 nm. Therefore, these residues just count for 30% of the opsin shift between the BR (560 nm) and SRII. Gly 130 and Val 108 are toward the end of the retinal chromophore, while Thr 204 is close to the PSB (29).

A closer comparison between the residues in the binding pocket of the two proteins highlighted seven residues that are different, which includes Ile 43, Ile 83, Asn 105, Val 108, Phe 127, Gly 130 and Phe 134. Mutation of these residues to equivalent positions in BR to provide I43V/I83L/N105D/V108M /F127W/G130S/F134M mutant provides a pigment that absorbs at 509 nm. In the search for more differences in the residues inside the binding pocket three more amino acids were highlighted within 5 Å distance of the chromophore. These three residues are Met 109, Ala 131 and Thr 204. The addition of these mutation to provide the mutant I43V/I83L/N105D/V108M /F127W/G130S/F134M/A131T/M109I/T204A, provided a pigment which absorbs at 524 nm. This still counts for just 44% of the Opsin shift between the SRII and BR. It was concluded that other structural factors like the angle between the β-ionone ring and polyene chain or the distance between the Schiff Base and its counter anion might be the contributing factors (**Figure III-9**) (30).

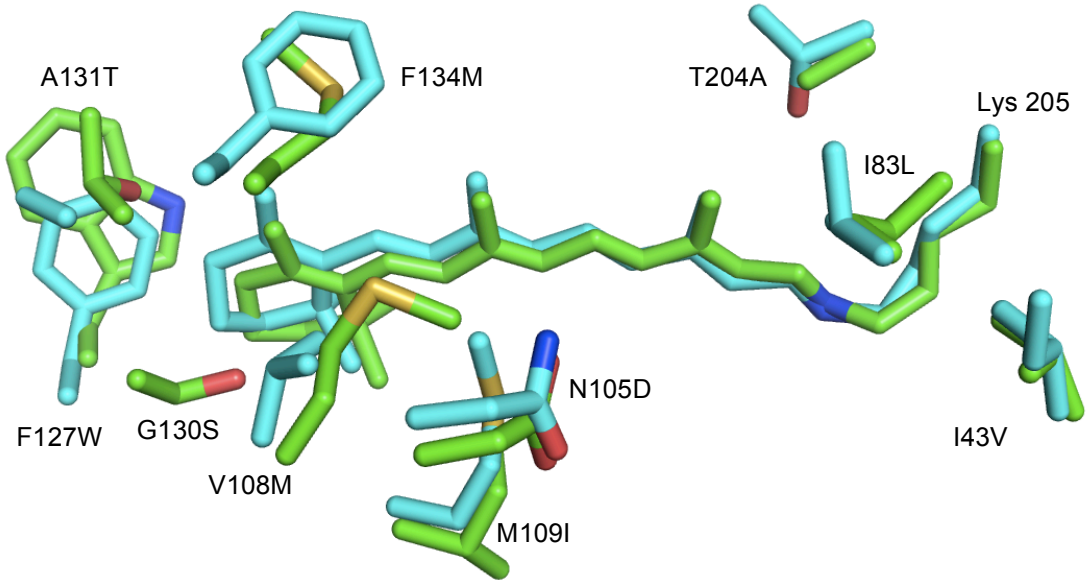


Figure III-9) Overlay structure of the BR (Green - PDB ID: 1C3W) and SRII (Cyan -PDB ID: 1JGJ). The residue numbering is that of SRII.

It was shown in other studies that L93A or M145A mutations on BR, which are the F134M and I83L mutations above, largely blue shifts the protein, changes the isomer content of the protein and affects the kinetic of the photoisomerization of the BR during its photocycle (31, 32).

Soon after these findings, the SRII crystal structure was published and the retinal polyene and its ring were found planar as BR. The two Aspartic acid residues, Asp 75 and 201 that act as counter anion for retinal PSB were also in identical positions as BR. However, three structural factors that contribute to the blue shift of the SRII are explained as:

(1) The displacement of Arg 72 by 1.1 Å and its conformational change with a rotation away from the Schiff base that strengthens the PSB and its counter anion interaction. The repositioning of the Arg 72 is attributed to the helix backbone differences between BR and SRII that changes from Phe 208 to Ile 197, Glu 194 to Pro 183 and Glu 204 to Asp 192. Also, the position of other

residues inside the binding pocket of the SRII is altered by interacting with the residues outside of it.

(2) Removal of two hydroxyl groups near the β-ionone ring, which are Ser 141 and Thr 142 in BR.

(3) Change in the tilt and position of the retinal which alters the interaction of the retinal polyene with the binding pocket, resulting in the modulation of the Ground state and Excited state energy levels.

It was rationalized that although earlier studies highlighted the differences in amino acids in the binding pockets of the two proteins and indicated the 44% of the opsin shift between the BR and SRII, but the aforementioned structural differences were not considered as part of the differences between the two proteins (**Figure III-10**) (33).

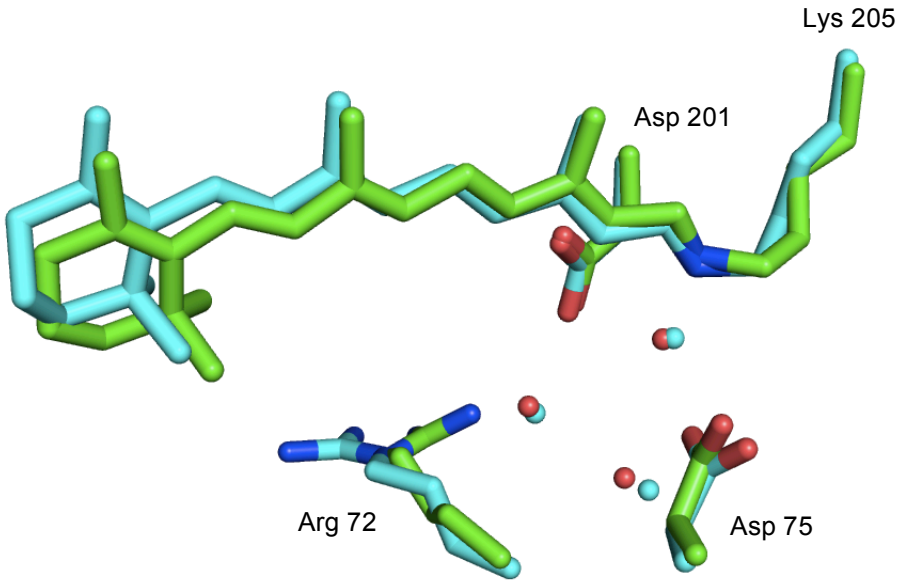


Figure III-10) Overlay structure of the BR (Green - PDB ID: 1C3W) and SRII (Cyan -PDB ID: 1JGJ). The residue numbering is that of SRII. The differences in the trajectory of the retinal and the Arg 72 residue are highlighted.

The most updated study on the point charge theory model was done in professor Borhan and Geiger lab at Michigan State University. In a 2012 science paper human Cellular Retinol Binding Protein II (hCRBPII) was

reengineered to a rhodopsin mimic. The wavelength of the protein was tuned from 425 nm to 646 nm. The introduction of counter anions close to the PSB works well for blue shifting the protein as well as the elimination of the counter anions and waters near the PSB that red shifts the protein. Counter intuitively, the introduction of negatively charged amino acids at the ionone ring of the retinal does not red shift the protein, instead the introduction of the aromatic bulky residues like Trp, Phe and Tyr not only red shifts the protein, but increases the effect of the internal mutations by a factor of 1-2. The same electrostatic calculation that was done for the microbial rhodopsins was done by the authors on the two mutant proteins absorbing at 508 nm and 622 nm. The result of these calculations indicates that the key element for red shifting a retinal pigment is not in changing the negative electrostatic potential from the PSB to the ionone ring, but in equally distributing it along the retinal chromophore (34) (**Figure III-11**).

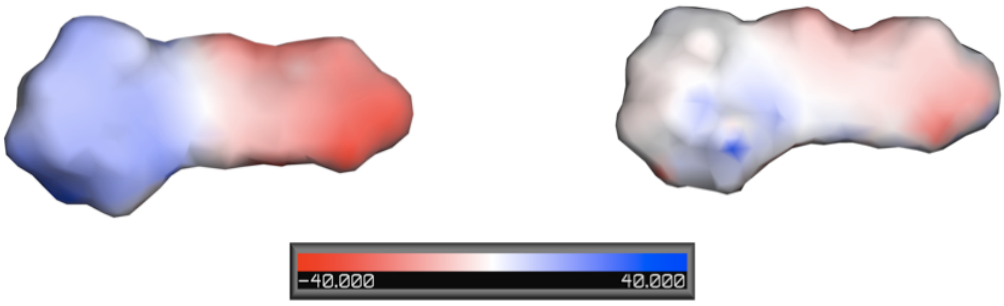


Figure III-11) Electrostatic charge distribution on the surface of the retinal projected from the protein for left: Q108K:K40L mutant of hCRBPII, λ_{max} = 508 nm. Right: Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R, λ_{max} = 622 nm.

The electronic polarization of the chromophore by aromatic residues (35-39) or the overall environment of the binding pocket (40-42) has been implicated as contributing factors in the wavelength regulation of the retinal PSB in Rhodopsins as well.

III.3 The Conformational Theory

In the conformational theory the emphasis is on the retinal trajectory and the planarity or twist of the chromophore. The binding pocket of the protein can force the retinal to keep its double bonds in plane with each other or it can force it to twist. The net result of the planarity of the chromophore is that the double bonds will stay in conjugation, and the positive charge of the PSB can travel all the way to the end of the polyene. However, when it is twisted then the double bonds will be less conjugated and the PSB positive charge will not delocalize which translates into blue shift (**Figure III-12**).

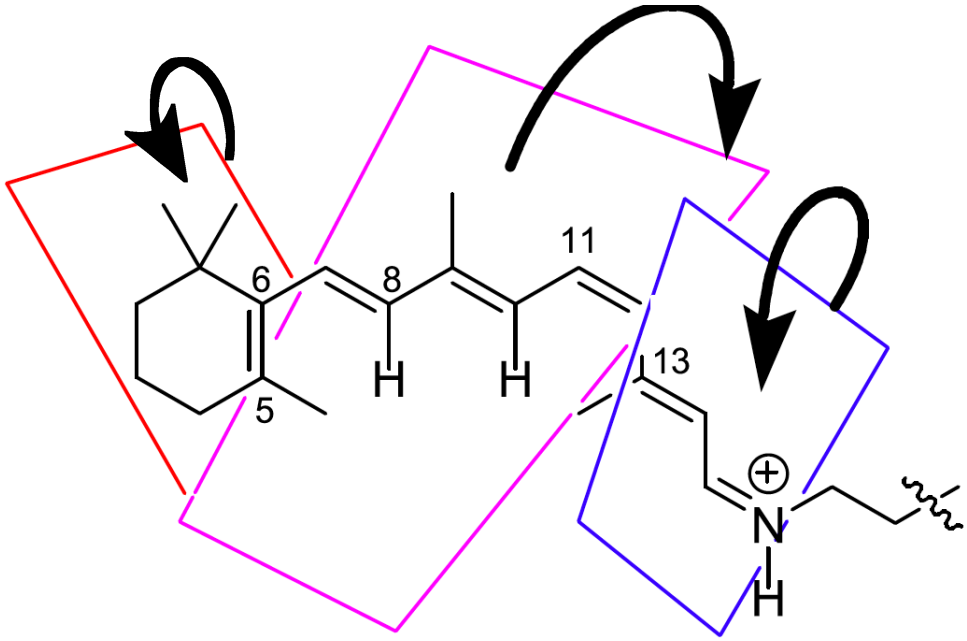


Figure III-12) The planes of rotation for 11-*cis* retinal.

The reason for the twist of the retinal chromophore is in the steric repulsion of its methyl groups with its hydrogens. As shown above for 11-*cis* retinal the C-13 and C-5 methyl groups are in steric clash with the hydrogens of C-10 and C-8, respectively. Therefore, the molecule can relieve stress by rotating around any of the planes that are highlighted. Depending on how the residues inside the binding pocket of a retinal pigment protein are arranged, the

dihedral angle of these planes can change, which will directly translate into the wavelength absorbed by the protein (43-48).

In 2010 the gas phase absorption of retinal n-butyl iminium PSB was measured by Garavelli, and showed a broad flat top absorption from 530 to 610 nm. This broad peak was attributed to conformational changes of the retinal and in particular to the β-ionone ring rotation, and the dihedral angle between the C5-C6 and C7-C8 double bonds of retinal. Retinal analogs with locked conformation of 6s-*trans* and twisted 6s-*cis* were made and interestingly, the absorption of the 6s-*trans* analog fell on the 610 nm edge and the absorption of the 6s-*cis* fell on the 530 nm edge of the spectrum. The difference in the S0 energy level between these forms of the retinal was found to be 0.6 kcal/mol and up to 4 kcal/mol for the analogs. The absorption of all of the bacterial rhodopsins that have 6s-*trans* retinal PSB is blue shifted compared to the gas phase absorption of 6s-*trans* retinal PSB. The absorption of Rhodopsin and color opsins, which have twisted 6s-*cis* retinal PSB is also blue shifted compared to the gas phase absorption of 6s-*cis* retinal PSB in the gas phase. From these it was concluded that retinal proteins mask and smoothen the blue shifting effects of the counter anion. Exceptions to these were the SRII and blue opsin, which absorb at significantly blue shifted wavelengths. Therefore, the 610 nm absorption of the all-*trans* retinal PSB in gas phase was made as the upper limit and most red shifted that a retinal PSB can absorb (45) (**Figure III-13**). Twisted 6s-*cis* retinal conformation in Rhodopsin and color opsins and 6s-*trans* retinal in bacterial rhodopsins is one of the contributing factors to the fact that the first absorb all below 560 nm while the later can absorb up to 590 nm.

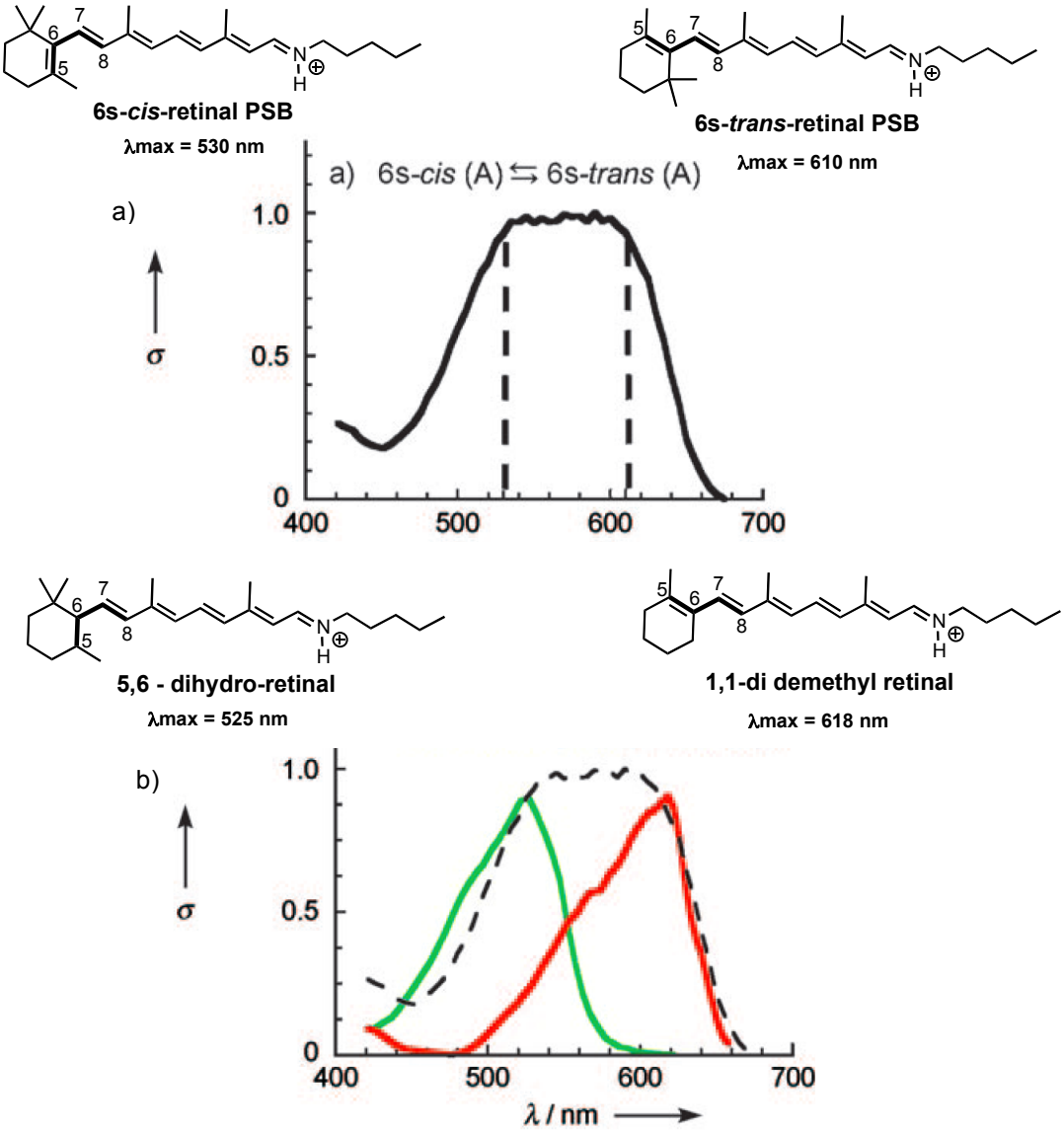


Figure III-13) The gas phase absorption of: a) retinal n-butyl iminium PSB. b) analogs of retinal n-butyl iminium PSB.

However, hCRBPII-reengineered proteins absorbing as red as 646 nm hit this barrier, and to date are the most red shifted a retinal PSB can absorb (34).

The most recent study in taking advantage of the 6s-*trans* to 6s-*cis* retinal isomerization is related to the mutagenesis studies in making blue shifted absorbing pigments of Channelrhodopsins (ChRs) and Archaerhodopsin-3 (AR-3) as optogenetic tools. It was shown that by two mutations on C1C2 ChR and four mutations on AR-3 the wavelength of the protein can shift by

100 nm. The crystal structure of the C1C2 was reported in 2012 showing a 6s-*trans* retinal inside the binding pocket of the protein like the other bacterial rhodopsins (49). However, the crystal structure of the double mutant of this protein showed a 6s-*cis* retinal structure. The double mutant absorbs at 455 nm, while the wild type protein absorbs at 476 nm. AR-3 absorbs at 550 nm and mutants of this protein by three or four mutations were blue shifted all the way to 450 nm (**Figure III-14**) (50).

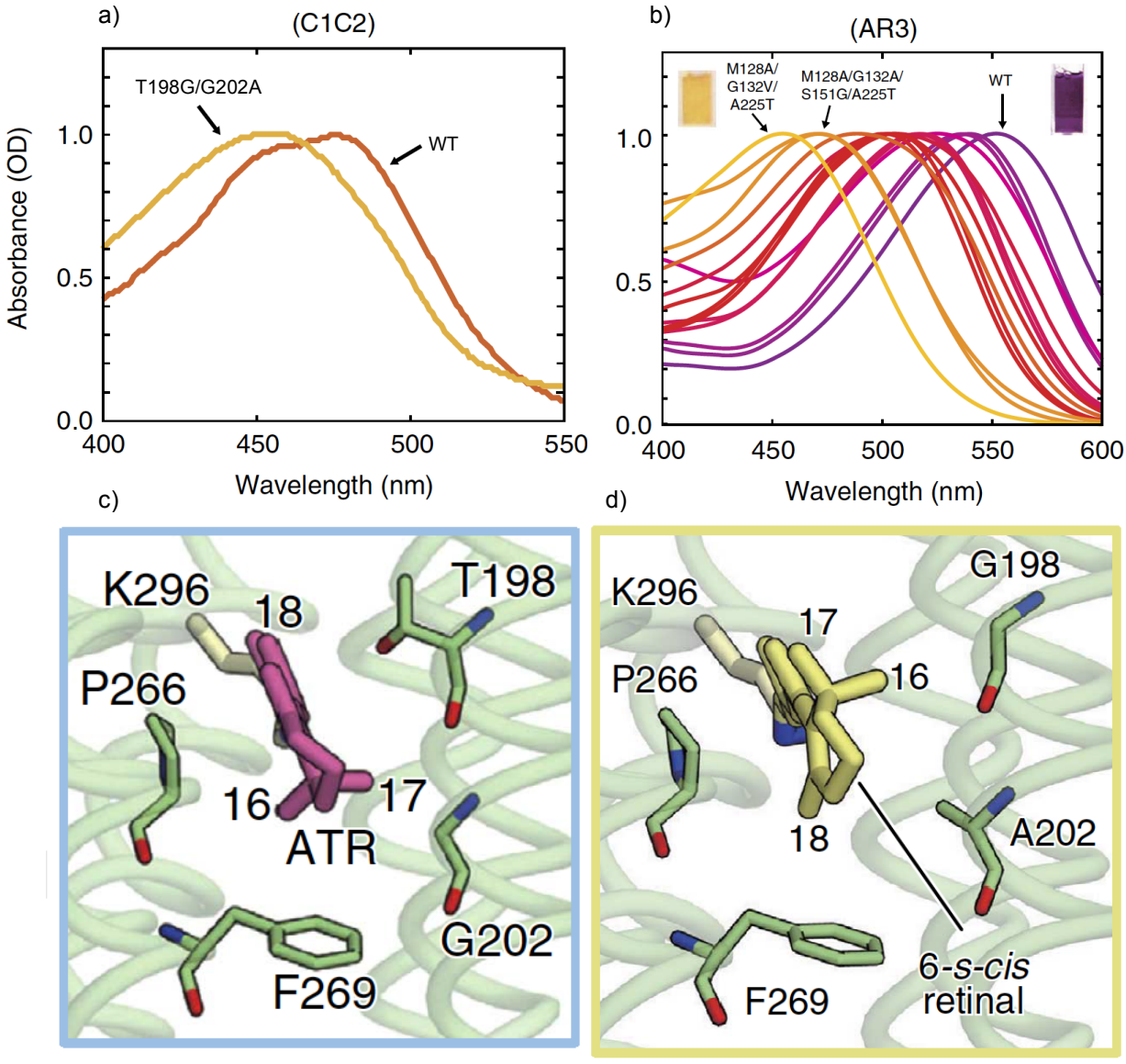


Figure III-14) The studies on developing blue shifted bacterial rhodopsin optogenetic tools: a) The UV-Vis spectrum of blue shifted C1C2 Channel Rhodopsin (ChR) proteins. b) The UV-Vis spectrum of blue shifted archaerhodopsin-3 (AR-3) proteins. c) The crystal structure of the wild type C1C2 protein shows 6s-*trans* retinal (PDB ID: 3UG9). d) The crystal structure of the T198G:G202A mutant of C1C2 shows 6s-*cis* retinal (PDB ID: 4YZI).

III.4 Water Molecules in Wavelength Regulation

Waters play an important role in protein functions, especially when interacting with conserved residues. Most of the studies on the effect of the water molecules on retinal signaling proteins have been accomplished in the last decade, and over time the effect of these waters have been more and more elucidated. It has been shown that waters play a critical role in wavelength regulation of Rhodopsin and visual pigments (51, 52) as well as their signaling (51, 53-58), wavelength regulation of bacteriorhodopsin (59) and its proton pumping activities (60, 61), bulk water in retinal-protein complex hydrolysis (62) and signaling of squid rhodopsin (63).

The first crystal structure of bovine rhodopsin at 2.8 Å resolution failed to highlight some of the water molecules important in the signaling and wavelength regulation of Rhodopsin (14). In the next few years, two more structures with 2.6 Å (51) and 2.2 Å resolution (53) showed two water molecules inside the binding pocket of the Rhodopsin near the PSB important for the wavelength regulation of this protein and theoretical calculations showed that each one can contribute up to 32 nm to wavelength shift of the protein (52). These two waters were called Wat2a and Wat2b (**Figure III-15**).

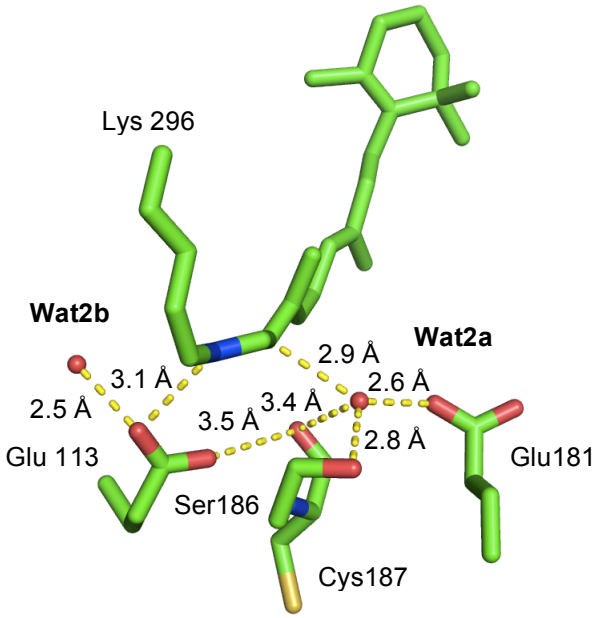


Figure III-15) The position of Wat2a and Wat2b in the binding pocket of Rhodopsin and interacting residues (PDB ID: 1L9H).

Wat2b is hydrogen bonding with Glu113, the counter anion of the Rhodopsin. It possibly contributes to further stabilization of the negative charge of the counter anion. The further stabilization of the negative charge of the counter anion might directly translate into the wavelength shift of the protein as well as the high pKa of 16.

Wat2a is hydrogen bonding with Glu181 and Ser186 and the hydrogen bonding network is extended to Glu113 indirectly via the main chain carbonyl of Cys187. The position of Wat2a is critical in terms of the color tuning of opsins, because in Green and Red opsin Glu181 is replaced by a histidine residue that facilitates the binding of Chloride anion in the binding pocket of the protein. Wat2a in rhodopsin most probably sits in this position and contributes to the blue shift of the Rhodopsin and blue opsin compared to the green and red opsin. It has been shown that an E181Q mutation can cause 10nm red shift, which might be because of the change of the polarity of the water due to the mutation (64).

It has been also suggested that Wat2a is the water molecule that facilitates the switch of the counter anion from Glu113 to Glu181 in the photocycle of Rhodoopsin. An overlay structure of the 2.6 Å resolution structure of the dark state with MetaII proves this hypothesis (53, 58) (**Figure III-16**).

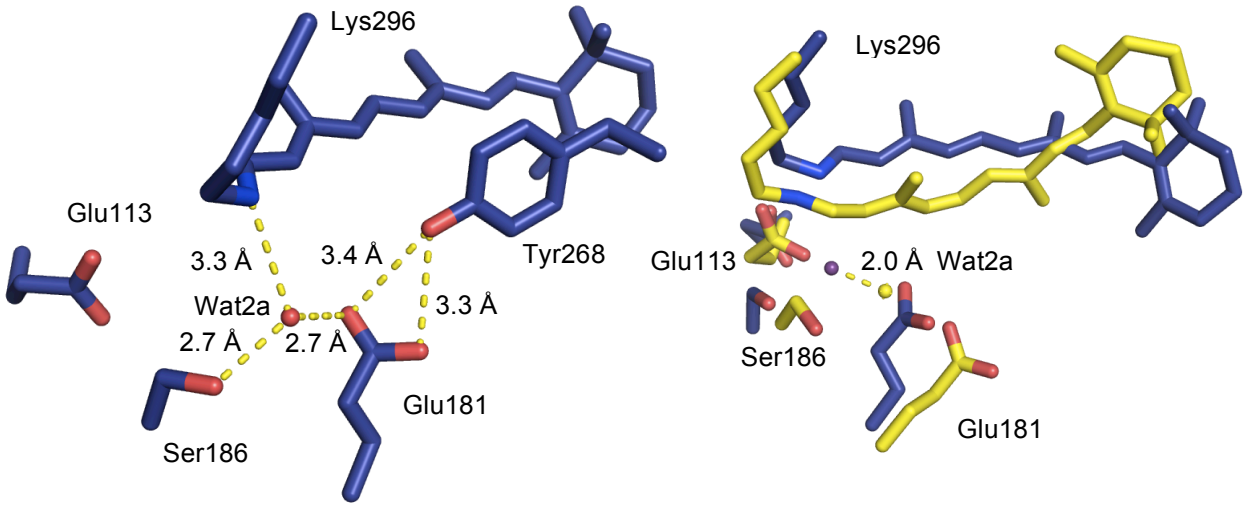


Figure III-16) The switch of the counter anion to Glu181 in the MetaII Rhodopsin-transducin complex. a) Glu181 as counter anion in MetaII Rhodopsin - transducin complex (Dark blue – PDB ID: 3PQR). b) The overlay structure of the ground state Rhodopsin (Yellow – PDB ID: 1L9H) and MetaII Rhodopsin – transducin complex (Blue - PDB ID: 3PQR).

These structures also highlighted three water molecules Wat1a, wat1b and Wat1c toward the cytoplasmic side of the rhodopsin that are important in its signaling. From these structures it was known that these waters interact with the highly conserved motif of $N_{302}P_{303}XXY_{306}X_{5,6}F$ on helix7. When the crystal structure of the Meta II rhodopsin was solved in 2012 by Ernst and co-workers, it showed a clear picture that the signaling of the Rhodopsin relies on a propagation of a water network that is made by structural rearrangement of the conserved motif on Helix7 and another conserved motif of $E_{134}(D)R_{135}Y_{136}$ on Helix3. The water network spans all the way from the binding pocket of the protein to the cytoplasmic surface of the rhodopsin where the G-protein binds (**Figure III-17**) (58).

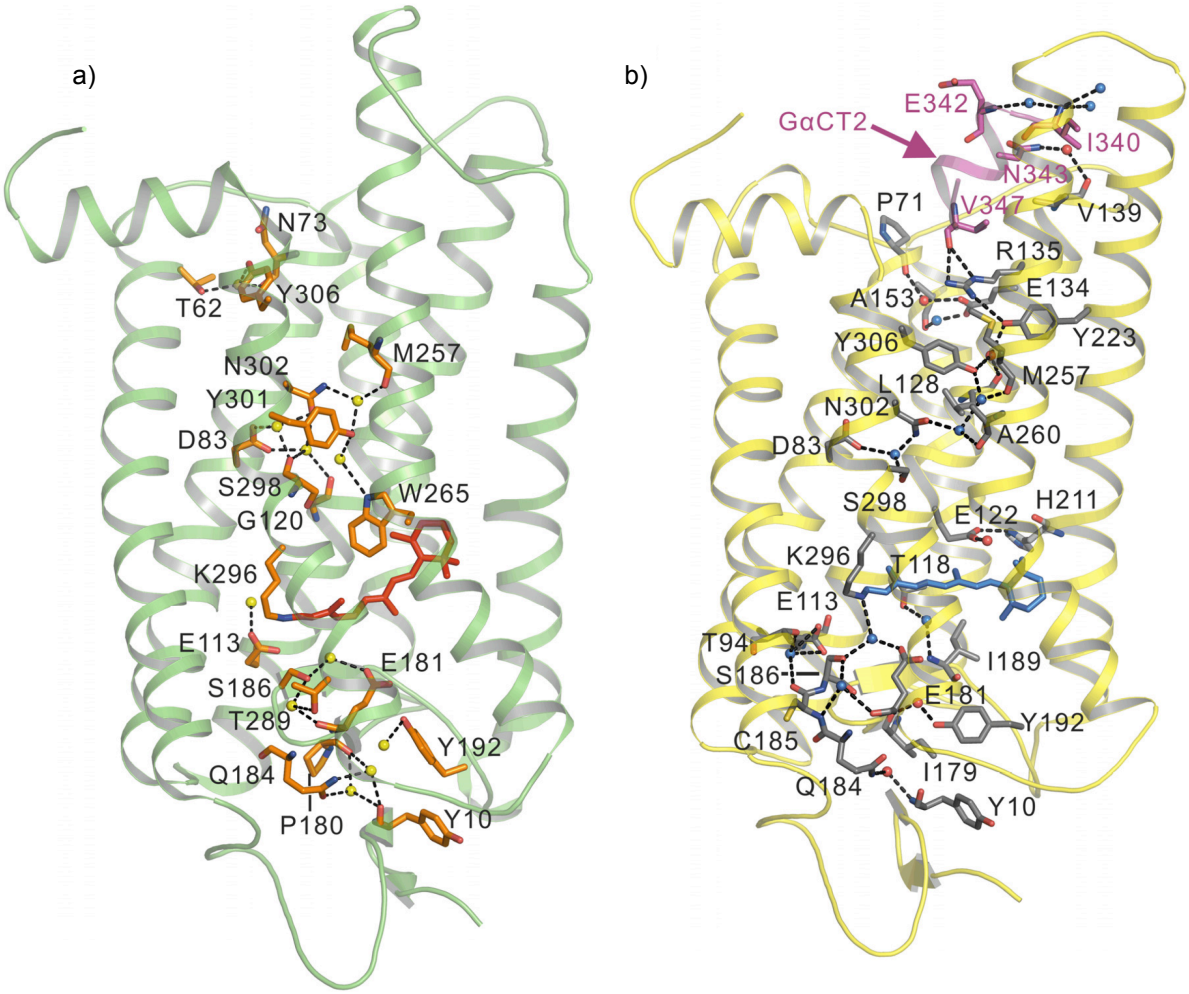


Figure III-17) a) water network in Rhodopsin structure (PDB ID: 1U19). b) Propagation of the water network all the way to the cytoplasmic surface of the Rhodopsin in the crystal structure of MetaII Rhodopsin with GaCT2 peptide fragment of G-protein (PDB ID: 3PQR).

In bacteriorhodopsin it was also shown that an L93A mutation causes 30 nm blue shift and the crystal structure of the protein showed that the cavity, which is made by this mutation, is occupied by four water molecules that contribute to the blue shift of the protein (**Figure III-18**). L93A mutation also increases the lifetime of the O intermediate of the bacteriorhodopsin photocycle by a factor of ~250 (59).

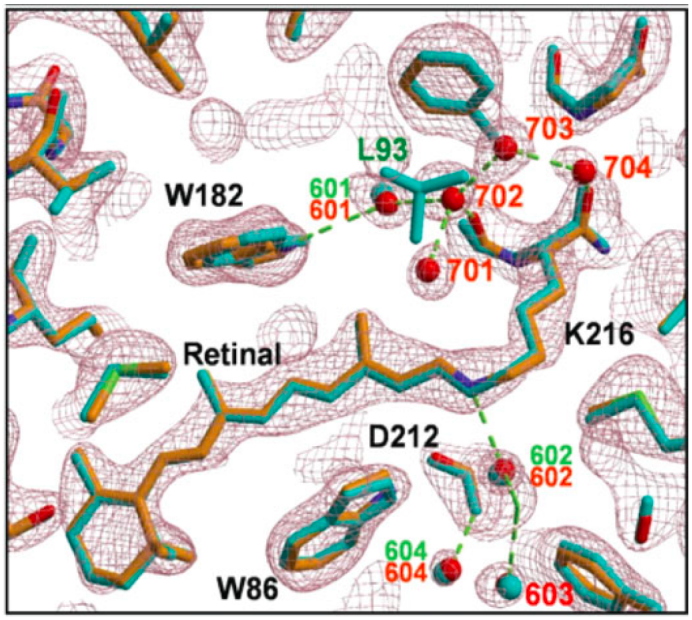


Figure III-18) Overlay structure of the L93A mutant of Bacteriorhodopsin (Gold – PDB ID: 3VHZ) with wild type Bacteriorhodopsin (Cyan – PDB ID: 1C3W). Waters 701-704 are inserted in the cavity made by L93A mutation near retinal PSB.

The importance of the water molecules in the wavelength regulation of retinal proteins was highlighted one more time in the studies of reengineered hCRBPII protein done in Professor Borhan and Geiger lab at Department of

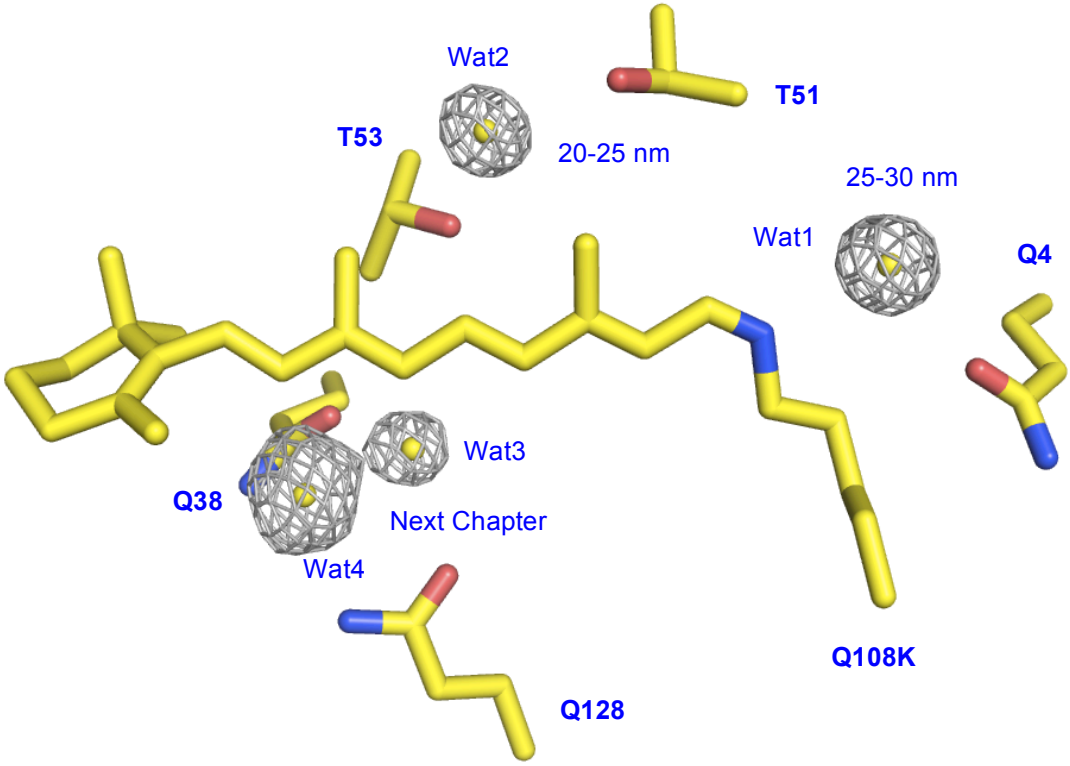


Figure III-19) The water molecules crucial in the wavelength regulation of hCRBPII (PDB ID: 4RUU).

Chemistry of Michigan State University (**Figure III-19**) (34, 65, 66).

At the end of this chapter it should be highlighted that the polar amino acids, specially the hydroxyl containing side chains, and the charged amino acids along the retinal chromophore, the twist of the retinal along its double bonds specially the ionone ring and the crucial role of water molecules are the main contributing factors to the wavelength regulation of retinal light absorbing pigments.

The next chapter will be the continuation of the reengineering of the hCRBPII protein and a more coherent story of the role of waters that puts all of the pieces of the puzzle together to prove the unanimous effect of the conserved waters in the binding pocket of hCRBPII that govern the wavelength regulation of this protein like the other retinal protein pigments.

REFERENCES

REFERENCES

1. Erickson JO & Blatz PE (1968) N-retinylidene-1-amino-2-propanol: a Schiff base analog for rhodopsin. *Vision research* 8(10):1367-1375.

2. Marti T, Rosselet SJ, Otto H, Heyn MP, & Khorana HG (1991) The retinylidene Schiff base counterion in bacteriorhodopsin. *The Journal of biological chemistry* 266(28):18674-18683.

3. Bownds D (1967) Site of attachment of retinal in rhodopsin. *Nature* 216(5121):1178-1181.

4. Akhtar M, Blosse PT, & Dewhurst PB (1968) Studies on vision. The nature of the retinal-opsin linkage. *The Biochemical journal* 110(4):693-702.

5. Wang JK, McDowell JH, & Hargrave PA (1980) Site of attachment of 11-cis-retinal in bovine rhodopsin. *Biochemistry* 19(22):5111-5117.

6. Oseroff AR & Callender RH (1974) Resonance Raman spectroscopy of rhodopsin in retinal disk membranes. *Biochemistry* 13(20):4243-4248.

7. Nakanishi K (1985) BIOORGANIC STUDIES WITH RHODOPSIN. *Pure and Applied Chemistry* 57(5):769-776.

8. Honig B*, et al.* (1979) EXTERNAL POINT-CHARGE MODEL FOR WAVELENGTH REGULATION IN VISUAL PIGMENTS. *Journal of the American Chemical Society* 101(23):7084-7086.

9. Mollevanger L*, et al.* (1987) HIGH-RESOLUTION SOLID-STATE C-13-NMR STUDY OF CARBONS C-5 AND C-12 OF THE CHROMOPHORE OF BOVINE RHODOPSIN - EVIDENCE FOR A 6-S-CIS CONFORMATION WITH NEGATIVE-CHARGE PERTURBATION NEAR C-12. *European Journal of Biochemistry* 163(1):9-14.

10. Steinberg G, Ottolenghi M, & Sheves M (1993) Pk(a) of the Protonated Schiff-Base of Bovine Rhodopsin - a Study with Artificial Pigments. *Biophysical journal* 64(5):1499-1502.

11. Zhukovsky EA & Oprian DD (1989) Effect of carboxylic acid side chains on the absorption maximum of visual pigments. *Science* 246(4932):928-930.

12. Sakmar TP, Franke RR, & Khorana HG (1991) The role of the retinylidene Schiff base counterion in rhodopsin in determining wavelength absorbance and Schiff base pKa. *Proceedings of the National Academy of Sciences of the United States of America* 88(8):3079-3083.

13. Gat Y & Sheves M (1993) A MECHANISM FOR CONTROLLING THE PK(A) OF THE RETINAL PROTONATED SCHIFF-BASE IN RETINAL PROTEINS - A

STUDY WITH MODEL COMPOUNDS. *Journal of the American Chemical Society* 115(9):3772-3773.

14. Palczewski K*, et al.* (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 289(5480):739-745.

15. Stenkamp RE, Filipek S, Driessen CAGG, Teller DC, & Palczewski K (2002) Crystal structure of rhodopsin: a template for cone visual pigments and other G protein-coupled receptors. *Bba-Biomembranes* 1565(2):168-182.

16. Neitz M, Neitz J, & Jacobs GH (1991) Spectral tuning of pigments underlying red-green color vision. *Science* 252(5008):971-974.

17. Yokoyama R & Yokoyama S (1990) Convergent evolution of the red- and green-like visual pigment genes in fish, Astyanax fasciatus, and human. *Proceedings of the National Academy of Sciences of the United States of America* 87(23):9315-9318.

18. Asenjo AB, Rim J, & Oprian DD (1994) Molecular determinants of human red/green color discrimination. *Neuron* 12(5):1131-1138.

19. Chan T, Lee M, & Sakmar TP (1992) Introduction of hydroxyl-bearing amino acids causes bathochromic spectral shifts in rhodopsin. Amino acid substitutions responsible for red-green color pigment spectral tuning. *The Journal of biological chemistry* 267(14):9478-9480.

20. Lin SW*, et al.* (1998) Mechanisms of spectral tuning in blue cone visual pigments. Visible and raman spectroscopy of blue-shifted rhodopsin mutants. *The Journal of biological chemistry* 273(38):24583-24591.

21. Janz JM & Farrens DL (2001) Engineering a functional blue-wavelength-shifted rhodopsin mutant. *Biochemistry* 40(24):7219-7227.

22. Kleinschmidt J & Harosi FI (1992) Anion sensitivity and spectral tuning of cone visual pigments in situ. *Proceedings of the National Academy of Sciences of the United States of America* 89(19):9181-9185.

23. Wang Z, Asenjo AB, & Oprian DD (1993) Identification of the Cl(-)-binding site in the human red and green color vision pigments. *Biochemistry* 32(9):2125-2130.

24. Kochendoerfer GG, Wang Z, Oprian DD, & Mathies RA (1997) Resonance Raman examination of the wavelength regulation mechanism in human visual pigments. *Biochemistry* 36(22):6577-6587.

25. Rajamani R, Lin YL, & Gao J (2011) The opsin shift and mechanism of spectral tuning in rhodopsin. *Journal of computational chemistry* 32(5):854-865.

26. Kochendoerfer GG, Lin SW, Sakmar TP, & Mathies RA (1999) How color visual pigments are tuned. *Trends in biochemical sciences* 24(8):300-305.

27. Kloppmann E, Becker T, & Ullmann GM (2005) Electrostatic potential at the retinal of three archaeal rhodopsins: implications for their different absorption spectra. *Proteins* 61(4):953-965.

28. Lee KSS (2011) 1) Probing the wavelength regulation of rhodopsin pigments via de novo protein engineering of a rhodopsin mimic and; 2) Engineering CRABPII as a retinal isomerase and a protein fusion tag. PhD (Michigan State University).

29. Shimono K, Iwamoto M, Sumi M, & Kamo N (2000) Effects of three characteristic amino acid residues of pharaonis phoborhodopsin on the absorption maximum. *Photochemistry and photobiology* 72(1):141-145.

30. Shimono K, Ikeura Y, Sudo Y, Iwamoto M, & Kamo N (2001) Environment around the chromophore in pharaonis phoborhodopsin: mutation analysis of the retinal binding site. *Biochimica et biophysica acta* 1515(2):92-100.

31. Subramaniam S, Greenhalgh DA, Rath P, Rothschild KJ, & Khorana HG (1991) Replacement of leucine-93 by alanine or threonine slows down the decay of the N and O intermediates in the photocycle of bacteriorhodopsin: implications for proton uptake and 13-cis-retinal----all-trans-retinal reisomerization. *Proceedings of the National Academy of Sciences of the United States of America* 88(15):6873-6877.

32. Greenhalgh DA, Farrens DL, Subramaniam S, & Khorana HG (1993) Hydrophobic amino acids in the retinal-binding pocket of bacteriorhodopsin. *The Journal of biological chemistry* 268(27):20305-20311.

33. Luecke H, Schobert B, Lanyi JK, Spudich EN, & Spudich JL (2001) Crystal structure of sensory rhodopsin II at 2.4 angstroms: insights into color tuning and transducer interaction. *Science* 293(5534):1499-1503.

34. Wang WJ*, et al.* (2012) Tuning the Electronic Absorption of Protein-Embedded All-trans-Retinal. *Science* 338(6112):1340-1343.

35. Irving CS, Byers GW, & Leermakers PA (1970) Spectroscopic model for the visual pigments. Influence of microenvironmental polarizability. *Biochemistry* 9(4):858-864.

36. Hays TR, Lin SH, & Eyring H (1980) Wavelength regulation in rhodopsin: effects of dipoles and amino acid side chains. *Proceedings of the National Academy of Sciences of the United States of America* 77(11):6314-6318.

37. Beppu Y & Kakitani T (1994) THEORETICAL-STUDY OF COLOR CONTROL MECHANISM IN RETINAL PROTEINS .1. ROLE OF THE TRYPTOPHAN RESIDUE, TYROSINE RESIDUE AND WATER MOLECULE. *Photochemistry and photobiology* 59(6):660-669.

38. Beppu Y (1997) Theoretical study of color control mechanism in retinal proteins .2. Orientational effects of aromatic amino acid residues upon opsin shift. *Journal of the Physical Society of Japan* 66(10):3303-3309.

39. Houjou H, Inoue Y, & Sakurai M (1998) Physical origin of the opsin shift of bacteriorhodopsin. Comprehensive analysis based on medium effect theory of absorption spectra. *Journal of the American Chemical Society* 120(18):4459-4470.

40. Wanko M, Hoffmann M, Frauenheim T, & Elstner M (2008) Effect of polarization on the opsin shift in rhodopsins. 1. A combined QM/QM/MM model for bacteriorhodopsin and pharaonis sensory rhodopsin II. *The journal of physical chemistry. B* 112(37):11462-11467.

41. Wanko M, Hoffmann M, Frahmcke J, Frauenheim T, & Elstner M (2008) Effect of polarization on the opsin shift in rhodopsins. 2. Empirical polarization models for proteins. *The journal of physical chemistry. B* 112(37):11468-11478.

42. Torii H (2002) Vibrational polarization and opsin shift of retinal schiff bases: theoretical study. *Journal of the American Chemical Society* 124(31):9272-9277.

43. Kakitani H, Kakitani T, Rodman H, & Honig B (1985) On the mechanism of wavelength regulation in visual pigments. *Photochemistry and photobiology* 41(4):471-479.

44. Cembran A*, et al.* (2005) Structure, spectroscopy, and spectral tuning of the gas-phase retinal chromophore: the beta-ionone "handle" and alkyl group effect. *The journal of physical chemistry. A* 109(29):6597-6605.

45. Rajput J*, et al.* (2010) Probing and modeling the absorption of retinal protein chromophores in vacuo. *Angewandte Chemie* 49(10):1790-1793.

46. Lugtenburg J*, et al.* (1986) MECHANISM FOR THE OPSIN SHIFT OF RETINALS ABSORPTION IN BACTERIORHODOPSIN. *Journal of the American Chemical Society* 108(11):3104-3105.

47. Van Der Steen R, Biesheuvel PL, Mathies RA, & Lugtenburg J (1986) RETINAL ANALOGUES WITH LOCKED 6-7 CONFORMATIONS SHOW THAT BACTERIORHODOPSIN REQUIRES THE 6-S-TRANS CONFORMATION OF THE CHROMOPHORE. *Journal of the American Chemical Society* 108(20):6410-6411.

48. Buss V, Kolster K, Terstegen F, & Vahrenhorst R (1998) Absolute sense of twist of the C12-C13 bond of the retinal chromophore in rhodopsin-semiempirical and nonempirical calculations of chiroptical data. *Angewandte Chemie-International Edition* 37(13-14):1893-1895.

49. Kato HE*, et al.* (2012) Crystal structure of the channelrhodopsin light-gated cation channel. *Nature* 482(7385):369-374.

50. Kato HE*, et al.* (2015) Atomistic design of microbial opsin-based blue-shifted optogenetics tools. *Nature communications* 6:7177.

51. Okada T*, et al.* (2002) Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proceedings of the National Academy of Sciences of the United States of America* 99(9):5982-5987.

52. Sekharan S (2009) Water-mediated spectral shifts in rhodopsin and bathorhodopsin. *Photochemistry and photobiology* 85(2):517-520.

53. Okada T*, et al.* (2004) The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. *Journal of molecular biology* 342(2):571-583.

54. Pardo L, Deupi X, Dolker N, Lopez-Rodriguez ML, & Campillo M (2007) The role of internal water molecules in the structure and function of the rhodopsin family of G protein-coupled receptors. *Chembiochem : a European journal of chemical biology* 8(1):19-24.

55. Angel TE, Chance MR, & Palczewski K (2009) Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proceedings of the National Academy of Sciences of the United States of America* 106(21):8555-8560.

56. Angel TE, Gupta S, Jastrzebska B, Palczewski K, & Chance MR (2009) Structural waters define a functional channel mediating activation of the GPCR, rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* 106(34):14367-14372.

57. Sun X, Agren H, & Tu Y (2014) Functional water molecules in rhodopsin activation. *The journal of physical chemistry. B* 118(37):10863-10873.

58. Choe HW*, et al.* (2011) Crystal structure of metarhodopsin II. *Nature* 471(7340):651-U137.

59. Zhang J*, et al.* (2012) Crystal structure of the O intermediate of the Leu93-->Ala mutant of bacteriorhodopsin. *Proteins* 80(10):2384-2396.

60. Gerwert K, Freier E, & Wolf S (2014) The role of protein-bound water molecules in microbial rhodopsins. *Biochimica et biophysica acta* 1837(5):606-613.

61. Furutani Y & Kandori H (2014) Hydrogen-bonding changes of internal water molecules upon the actions of microbial rhodopsins studied by FTIR spectroscopy. *Biochimica et biophysica acta* 1837(5):598-605.

62. Jastrzebska B, Palczewski K, & Golczak M (2011) Role of bulk water in hydrolysis of the rhodopsin chromophore. *The Journal of biological chemistry* 286(21):18930-18937.

63. Jardon-Valadez E, Bondar AN, & Tobias DJ (2010) Coupling of retinal, protein, and water dynamics in squid rhodopsin. *Biophysical journal* 99(7):2200-2207.

64. Terakita A, Yamashita T, & Shichida Y (2000) 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 of the United States of America* 97(26):14263-14267.

65. Sakmar TP (2012) Redder Than Red. *Science* 338(6112):1299-1300.

66. Wang W, Geiger JH, & Borhan B (2014) The photochemical determinants of color vision: revealing how opsins tune their chromophore's absorption wavelength. *BioEssays : news and reviews in molecular, cellular and developmental biology* 36(1):65-74.

Chapter IV: Water networks in the wavelength regulation of human Cellular Retinol Binding Protein II reengineered Protein

IV.1 Important waters in the binding pocket of hCRBPII protein

As our wavelength regulation studies showed little success with hCRABPII in the first generation series with the proteins absorbing maximally at 450 nm (1), one of the solutions was to use another protein from the Retinoid Binding Protein family. Since we knew that we can successfully make a Protonated Schiff Base (PSB) with hCRABPII, we decided to use a protein with the same fold and close homology to hCRABPII. human Cellular Retinol Binding Protein II (hCRBPII) was the target that closely met the requirements. hCRBPII binds, transfers and helps for the metabolism of retinol in the cellular environment (2, 3). It is a small cytosolic protein with a β-barrel fold and a smaller binding pocket than hCRABPII. The crystal structure of the wild type protein in complex with retinol shows the fold of this protein and the important residues inside the binding pocket of this protein (4) (**Figure IV-1**).

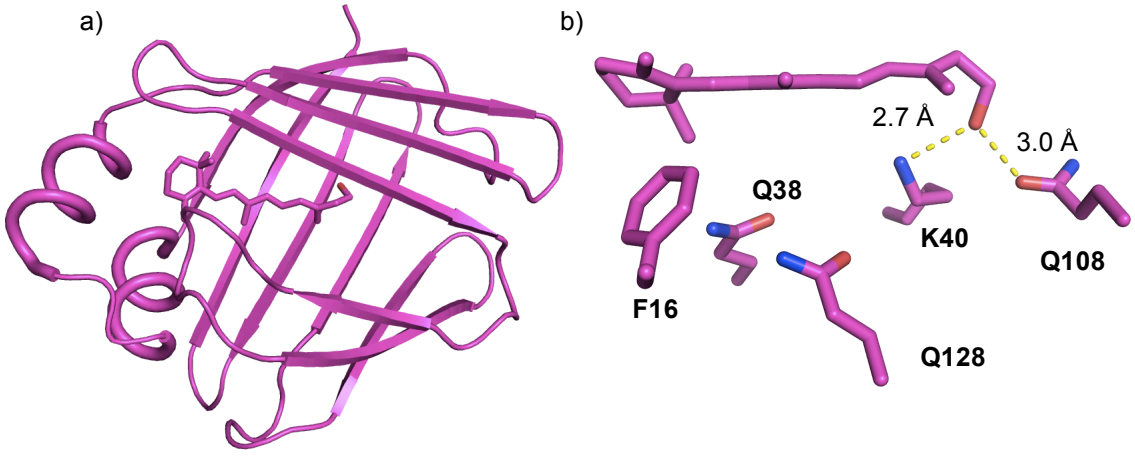


Figure IV-1) a) The crystal structure of wt hCRBPII bound to Retinol. b) Binding pocket of this protein with the important residues highlighted (PDB ID: 4QYN).

Residue Gln 108 deep inside the binding pocket of the protein was chosen for introducing the nucleophilic lysine residue. Mutant Q108K upon incubation with Retinal does not show any PSB in the physiological pH, upon acidification to pHs lower than 6 there is a slight absorption at 506 nm, but the protein can not be titrated to show a full PSB formation. Dr.Wang noticed the presence of a lysine residue Lys40 in proximity to the introduced lysine residue Q108K. In order to remove the possible positive charge of Lys40 adjacent to the formed PSB, which might be the cause of the low pKa of Q108K mutant, this residue was mutated to a hydrophobic Leucine. Q108K:K40L (**KL**) showed significant PSB formation in 10 minutes in the physiological pH with a λ_{max}= 508 nm. The pKa titration of this protein gives a pKa of 8.3 for the PSB (5). Figure 2 shows the crystal structure of this protein with retinal covalently bound to the Q108K residue (6) (**Figure IV-2**).

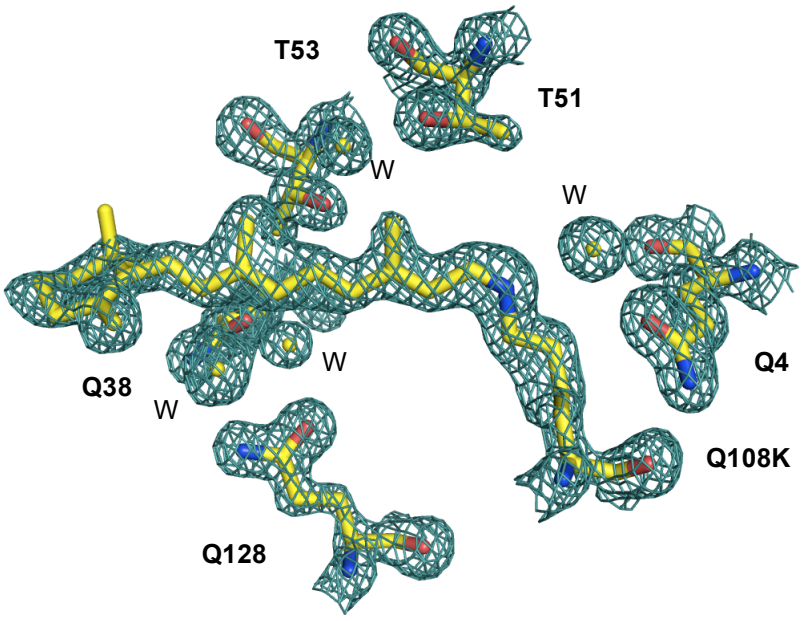


Figure IV-2) Retinal bound covalently inside the binding pocket of the hCRBPII Q108K:K40L mutant (PDB ID: 4RUU).

By extensive mutagenesis, spectroscopy and crystallization studies we were able to regulate the wavelength of this protein from 425 to 646 nm, more than

50 nm red shifted compared to the known rhodopsins, and unveil the mechanism of wavelength regulation of this protein (6-8).

Overlay of more than twenty crystal structures of hCRBPII shows conserved water molecules inside the binding pocket of this protein. These water molecules are the determinant factors that regulate the wavelength of the protein (9). These include 4 water molecules in the vicinity of the retinal chromophore that play a critical role in the wavelength regulation of this protein (**Figure IV-3**). The mutation of the interacting residues with these waters changes the wavelength of the protein at least by 20 nm.

These water molecules are:

(1) Gln4 water molecule W1

(2) Thr51 and Thr53 water W2

(3) Gln38 and Gln128 water molecules W3 and W4.

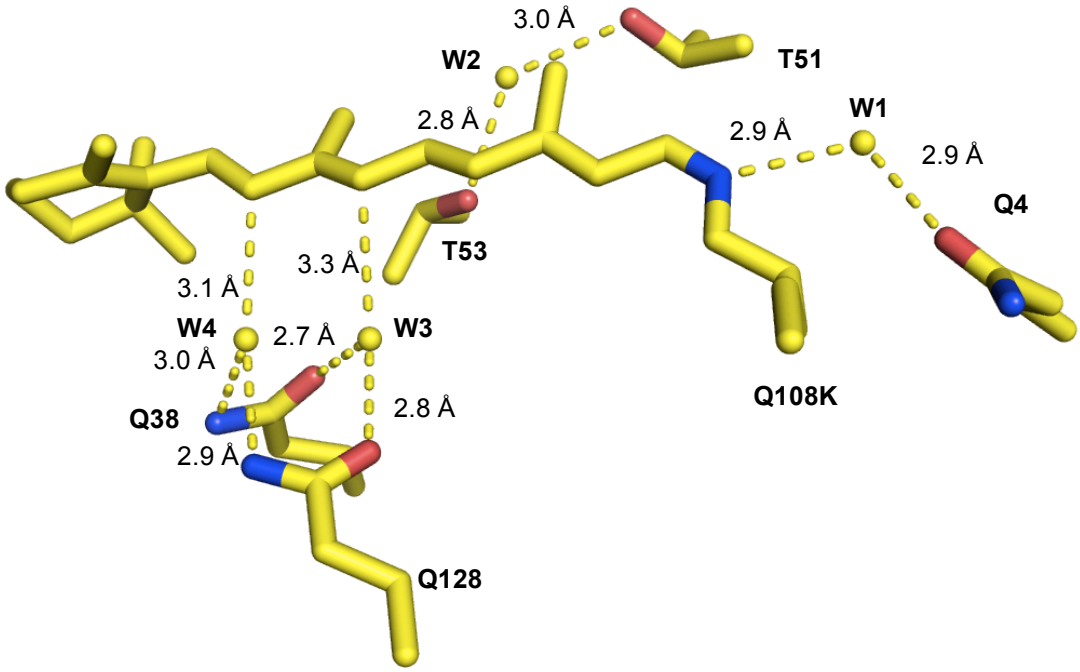


Figure IV-3) Crystal structure of the hCRBPII Q108K:K40L mutant (PDB ID: 4RUU). The important water molecules and the interacting residues with these waters are highlighted.

IV.2 Gln4 water molecule

Gln4 water molecule is in close proximity to the Schiff Base and is hydrogen bonded to the iminium. The Gln4 orientation is fixed in place by the main chain carbonyl of the neighbor residues to hold it in one fixed conformation to apply a certain dipole moment to the water molecule, which will interact with the chromophore (6) (**Figure IV-4**).

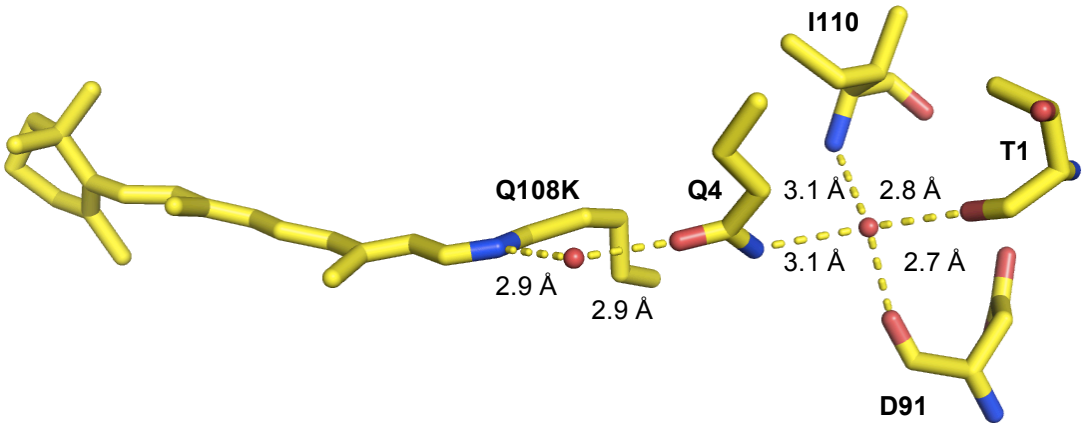


Figure IV-4) Gln4 is fixed in its conformation through a water molecule that is in interaction with the main chain carbonyl of Asp91 and Thr1. (PDB ID: 4RUU)

As expected the mutation of the Gln4 residue, especially to positively charged residues like Arginine or Lysine red shifts the protein by 25 nm. This is expected as the removal of the localized charge on the iminium and distributing it on the polyene will red shift the protein (6) (**Table IV-1**).

Protein	λ_{max} (nm)	pKa
KLVCWLW:**Q4**	591	8.2
KLVCWLW:**Q4A**	612	7.0
KLVCWLW:**Q4L**	614	7.9
KLVCWLW:**Q4F**	613	7.5
KLVCWLW:**Q4W**	613	7.7
KLVCWLW:**Q4T**	608	7.8
KLVCWLW:**Q4K**	618	7.2
KLVCWLW:**Q4R**	622	6.5
KLVCWLW:**Q4H**	585	7.9
KLVCWLW:**Q4E**	590	n.d.

Table IV-1. Gln4 mutation summary.
• KLVCWLW is the abbreviation for:
Q108K:K40L:T51V:T53C:R58W:T29L:Y19W

The crystal structure of the KLVCWLW:Q4R mutant shows that the introduced Arginine residue faces away from the positive charge of the iminium and the Gln4 water gets eliminated by this mutation (6) (**Figure IV-5**).

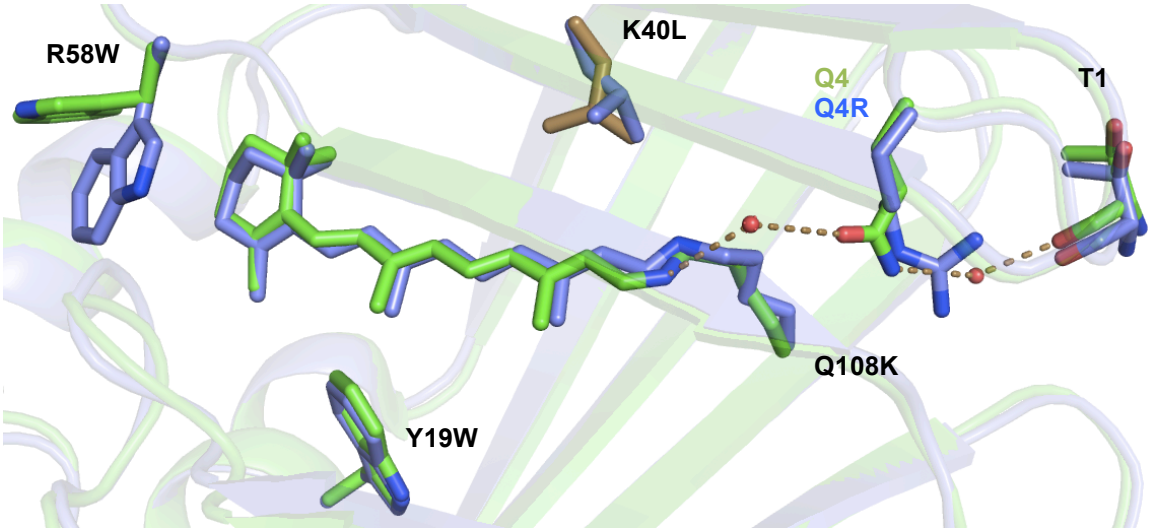


Figure IV-5) The overlaid structure of KL:T51V:T53C:R58W:Y19W:T29L (Cyan, PDB ID: 4EFG, 591 nm) and of KL:T51V:T53C:R58W:Y19W:T29L:Q4R (Green, PDB ID: 4EEJ, 622 nm).

IV.3 Thr51 and Thr53 water molecule

The water molecule in between the residues Thr51 and Thr53 is close to the beginning of the polyene tail of the retinal. One of the interesting findings during the mutagenesis of these residues is that a T51V mutation causes 25nm red shift while a T53C mutation counts for just 5nm red shift. Dr.Nossni's crystal structures of the mutants that have T51V mutation show that the water molecule is removed by introduction of this mutation, and Val51 takes a new conformation to face the hydrophobic polyene. However, the crystal structure of the mutant KL:T53C shows that the water molecule is still present. Therefore, the removal of the polarized water is important in changing the wavelength of the protein (9) (**Figure IV-6**).

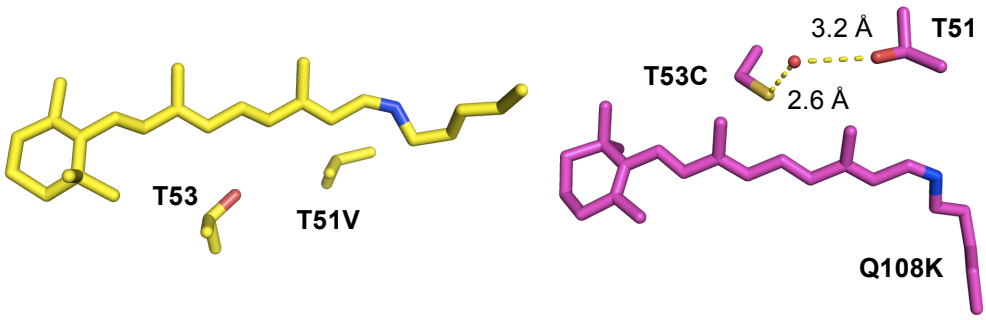


Figure IV-6) a) Crystal structure of the mutant Q108K:K40L:T51V shows the loss of W2 which counts for the 25 nm red shift of the protein. b) Crystal structure of the mutant Q108K:K40L:T53C shows the presence of the W2, which counts for the wavelength of the protein not to change.

Mutagenesis efforts proves the importance of this water molecule in red shifting the protein by 25 nm (**Table 2**) (5).

Protein	**λ_{max}(nm)**	**pKa**
Q108K:K40L	508	8.3
Q108K:K40L:**T51V**	533	8.3
Q108K:K40L:R58Y	535	9.5
Q108K:K40L:R58Y:**T51V**	563	10.0
Q108K:K40L:T53C	513	7.5
Q108K:K40L:T53C:**T51V**	539	8.4

Table IV-2. T51V mutation causes 25-30 nm red

IV.4 Gln 38 and Gln 128 water molecules

There are two water molecules in interaction with Gln 38 and Gln 128, which are located close to the retinal, toward the end of the polyene. These two glutamine residues interact with the retinal chromophore through the two water molecules (**Figure IV-2**). These two waters are highly polarized as close nearby hydrogen bonding interactions force Gln38 and Gln128 to face the waters through their amide carbonyl groups (**Figure IV-7**).

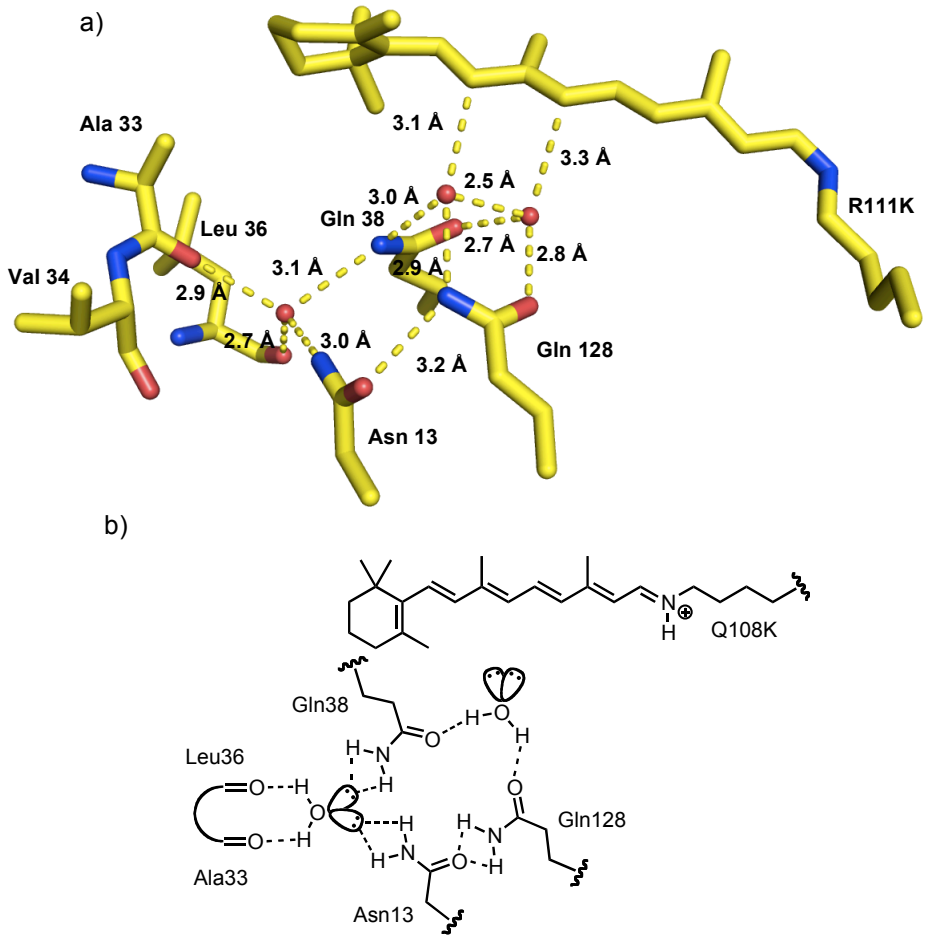


Figure IV-7) a) The Orientation of the Gln38 and Gln128 is defined by adjacent residues that fix the amide side chain of these residues in place. Crystal structure of Q108K:K40L. (PDB ID: 4RUU) b) Chemical representation of the orientation of these two rsidues.

The main chain carbonyl groups of residues Ala 33 and Leu 36 act as hydrogen bonding acceptor for a neighbor water molecule that directly interacts with Gln 38 and indirectly through Asn 13 to Gln 128. This will force the two glutamine residues to face the water molecule toward the polyene through amide carbonyl group and eventually the water (W3) will face the retinal through lone pairs to imply a negative dipole moment toward the retinal polyene. The orientation of the W4 can not be determined exactly, because the amidic $–NH_2$ group can act as a hydrogen bonding donor or acceptor.

Mutation of these residues has proved to wipe out the whole wavelength regulation of the protein by blue shifting it to the start point mutant of Q108K:K40L, which absorbs at 508 nm. A single Q38L mutation blue shifts the wavelength of the protein by up to 80 nm. The mutation of both of these residues to more hydrophobic residues produces proteins that counteracts the effects of all of the other mutations that red shifts the protein to above 600 nm back to 500 nm (Table **IV-3** and **IV-4**).

Protein	**λ_{max}(nm)**	**pKa**
KL:R58Y	533	9.0
KL:R58Y:**Q38A**	513	8.2
KL:R58Y:T51V	556	10.0
KL:R58Y:T51V:**Q38A**:**Q128A**	510	8.2
KL:R58Y:T51V: **Q38L**:**Q128L**	503	7.9
KL:R58Y:T51V:Y19W	565	10.2
KL:R58Y:T51V:Y19W: **Q38L**	522	8.7
KL:R58Y:T51V:Y19W: **Q38A**:**Q128A**	510	8.2

Table IV-3. Gln38 and Gln128 mutation on the shorter hCRBPII mutatnts.
- KL is the abbreviation for Q108K:K40L.

Gln 38 mutation is more effective than Gln128 mutation in blue shifting the mutants, but it does not drop the pKa of the protein more than 1.5 units. On the other hand, Gln128 mutation can drop the pKa of the protein by as much as 4 units, but is not as effective in blue shifting the protein as Gln 38 is (**Table IV-4**).

Protein	λ_{max}(nm)	pKa
KLW:T51V:T53C:T29L:Y19W	591	8.2
KLW:T51V:T53C:T29L:Y19W:**Q128L**	531	4.6
KLW:T51V:T53S:T29L:Y19W	600	9.0
KLW:T51V:T53S:T29L:Y19W:**Q38L**	520	8.1
KLW:T51V:T53S:T29L:Y19W:**Q38N**	571	8.3
KL:T51V:T53S:T29L:Y19W:**Q38L**:**Q128L**	502	-
KLW:T51V:T53C:T29L:Y19W:Q4R:A33W	644	7.2
KLW:T51V:T53C:T29L:Y19W:Q4R:A33W:**Q38L**	510	-
KLW:T51V:T53C:T29L:Y19W:Q4F:A33W	646	8.7
KLW:T51V:T53C:T29L:Y19W:Q4F:A33W:**Q128L**	608	
KLW:T51V:T53C:T29L:Y19W:Q4F:A33W:**Q38N**	610	4.7
KLW:T51V:T53C:T29L:Y19W:Q4F:A33W:**Q128N**	565	-
KLW:T51V:T53C:T29L:Y19W:Q4F:A33W:**Q38L**: **Q128L**	538	-

Table IV-4. The effect of the Gln38 and Gln128 mutations on hCRBPII longer series proteins.
- KLW is the abbreviation for Q108K:K40L:R58W.

Unfortunately, the drop in the pKa that accompanies the Gln38 and Gln128 mutations makes it hard to get good quality diffracting crystals from these mutants. The highest pKa mutant of hCRBPII series Q108K:K40L:T51V:R58Y:Y19W with a pKa of 10.2, after mutating both Gln38 and Gln128 produces blue shifted proteins that absorb 500-510 nm (**Table IV-3**). Although the pKa of these mutants is as high as 8.3, but still the crystallization trials provide crystals that do not diffract anywhere better than 3.5 Å. However, the crystallization of the Q108K:K40L:T51V:R58Y:Y19W:Q38L mutant provided good quality diffracting crystals in the same crystal packing form of the parent mutant to enable a comparison between the two crystal structures.

Interestingly, the comparison of the two crystal structures shows that in the crystal structure of Q108K:K40L:T51V:R58Y:Y19W in the two chains of the protein in the asymmetric unit the water molecules are present both in one and the other chain shows just one water, while in the crystal structure of

Q108K:K40L:T51V:R58Y:Y19W:Q38L mutant one of the chains shows the density for one water molecule and the other chain does not show any water molecules between the two glutamine residues proving the partial occupancy of water after the Q38L mutation. Additionally, the introduced hydrophobic leucine residue faces away from the water molecule in the chain that shows the density for it, proving the decreased dipole moment of the water compared to when that both glutamines are present. More importantly the water molecule in the Q38L structure is able to rotate in its place and is not fixed. Therefore, there is no specific dipole moment direction from the water to the retinal polyene (**Figure IV-8**).

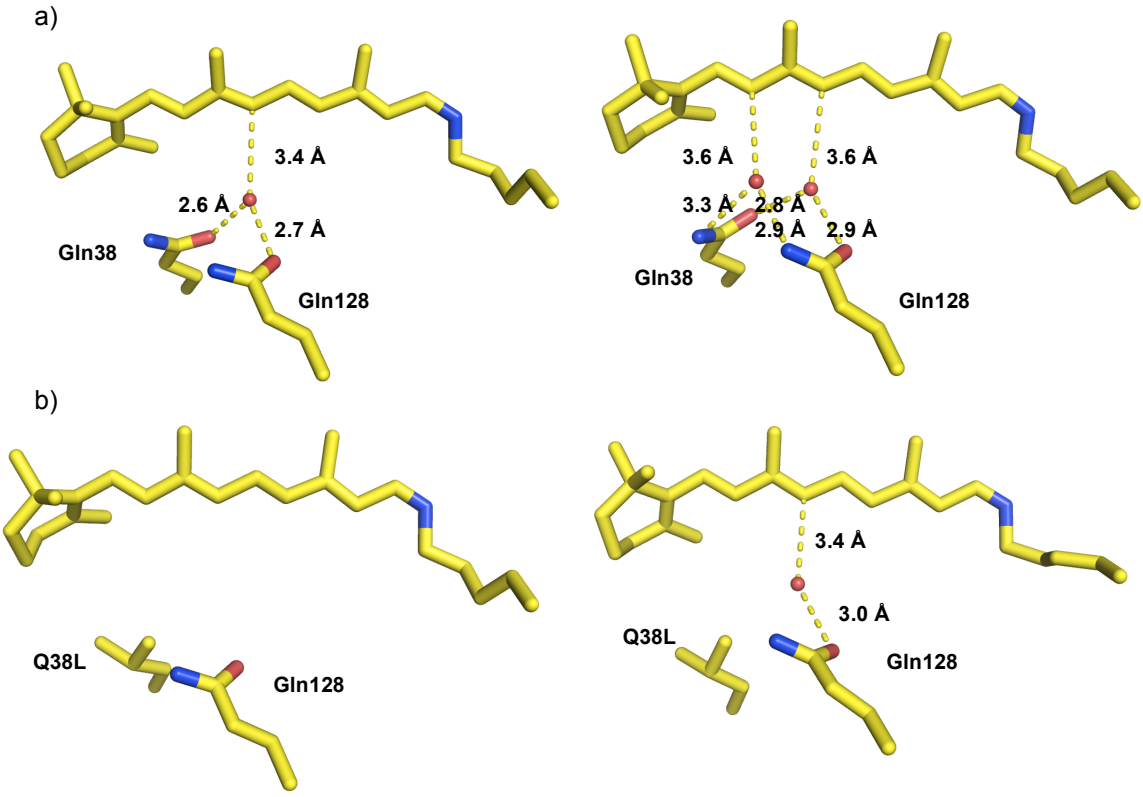


Figure IV-8) a) Gln38-Gln128 Water molecules in the two chains of crystal structure of KL:T51V:R58Y:Y19W. b) Gln38-Gln128 Water molecules in the two chains of crystal structure of KL:T51V:R58Y:Y19W:Q38L.

IV.5 Arg58 and Ala33 mutations

In hCRBPII besides the residues that directly interact with waters and their mutations affect the wavelength of the protein; two residues on the entrance of the binding pocket play an important role in the wavelength regulation of this protein as well. These include Arg58 and Ala33. Mutation of each one of these residues to aromatic hydrophobic residues can increase the absorbed wavelength of the protein by 25 nm (5, 6) (**Table IV-5**).

Protein	λ_{max}(nm)	pKa
Q108K:K40L	508	8.3
Q108K:K40L:**R58W**	519	8.7
Q108K:K40L:**R58F**	523	8.7
Q108K:K40L:**R58Y**	535	9.5
Q108K:K40L:T51V	533	8.3
Q108K:K40L:T51V:**R58W**	565	8.4
Q108K:K40L:T51V:**R58F**	561	8.7
Q108K:K40L:T51V:**R58Y**	563	10
Q108K:K40L:T51V:T53C	539	8.4
Q108K:K40L:T51V:T53C:**R58W**	585	7.4
Q108K:K40L:T51V:T53C:T29L:Y19W	533	7.9
Q108K:K40L:T51V:T53C:T29L:Y19W:**R58W**	591	8.2
Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R	622	6.7
Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R:**A33W**	644	6.5

Table IV-5. The effect of R58W and A33W on red shifting the hCRBPII

Interestingly, the addition of R58W mutation can increase the effect of the internal mutations on the wavelength of the protein by a factor of 2 or 1.5 respectively (6) (**Table IV-6**).

Entry	hCRBPII mutant	λ_{max} (nm) R58	Protein Shift (nm)	Protein Shift (CM^{-1})	λ_{max} (nm) R58W	Protein Shift (nm)	Protein Shift (CM^{-1})	Enhancement (nm)
1	KL	508	0	0	527	0	0	-
2	KL:T51V	533	25	923	570	43	1431	18 (1.7×)
3	KL:T53C	513	5	192	540	13	457	8 (2.6×)
4	KL:T51V:T53C	539	31	1132	585	58	1881	27 (1.9×)
5	KL:Y19W	513	5	192	538	11	388	6 (2.2×)
6	KL:T51V:Y19W	537	29	1063	577	50	1644	21 (1.7×)
7	KL:T51V:T53C:Y19W	538	30	1098	590	63	2026	33 (2.1×)

Table IV-6. Protein shift with reference to Q108K:K40L, R58W enhancement effect.

- wavenumbers provide a direct correlation to the change in energy. †Protein shift with reference to Q108K:K40L:R58W. Enhancement is calculated as the difference in protein shift between KL-R58W mutants and the KL mutants, and reflects the overall increased red shift in excess of that anticipated from a purely additive effect of R58W. For example, the T51V mutation leads to a 25-nm bathochromic shift (KL versus KL:T51V). A 25-nm red shift would be expected for KL:T51V:R58W versus KL:R58; however, a 43-nm shift is observed. The 18-nm difference in the level of enhancement (factor of 1.7 increase) is a result of the R58W mutation. Numbers in parentheses are relative increases of the protein shift of the KL-R58W mutant series with respect to the KL mutant series.

The mutation of R58F also has similar effects, and increases the effect of the internal residues by a factor of 1.5 (5) (**Table IV-7**).

Entry	hCRBPII mutant	λ_{max} (nm) R58	Protein Shift (nm)	Protein Shift (CM^{-1})	λ_{max} (nm) R58F	Protein Shift (nm)	Protein Shift (CM^{-1})	Enhancement (nm)
1	KL	508	0	0	523	0	0	-
2	KL:T51V	533	25	923	561	38	1296	13 (1.5×)
3	KL:T53C	513	5	192	537	14	499	9 (2.8×)
4	KL:Y19W	513	5	192	537	14	499	9 (2.8×)
5	KL:T51V:T53C	539	31	1132	571	48	1608	17 (1.5×)

Table IV-7. R58F mutation enhances the effect of the internal mutations by a factor of 1.5-2.

A surface representation of the crystal structures of the mutants which have the R58W mutation with the ones that do not, shows that there is not that much difference in how much the binding pocket is covered or in case of the Arg58 mutants the binding pocket is more covered compared to when this residue is mutated to a Tryptophan. Besides the fact that Tryptophan is more hydrophobic than an Arginine residue, Phe57 also flips out of the binding

pocket by the introduction of R58W mutation or R58F or R58Y. This causes the formation of a hydrophobic barrier that is made by R58WFY and Phe57. This hydrophobic barrier does not allow the water molecules to path easily into the binding pocket; Therefore, the effects of the high dielectric constant of the solvent, which wipes out the effect of the mutations, are furthermore shielded by this mutation (9) (**Figure IV-9**).

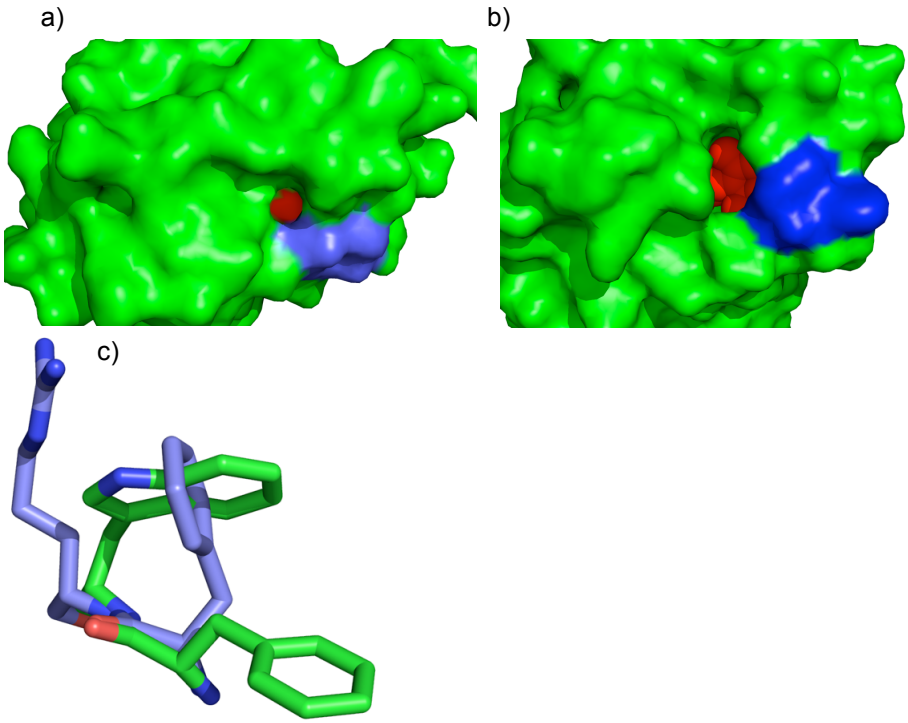


Figure IV-9) a) Surface representation of the Q108K:K40L:T51V:T53C mutant, λ_{max} = 539 nm (left) b) Surface representation of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (Right) mutant, λ_{max} = 591 nm. c) R58W mutation forces the Phe57 to flip out of the binding pocket.

The addition of A33W mutation to the mutants that have the R58W mutation also adds 15-20 nm red shift to the absorption of the protein. The overlay structure of the two proteins without and with the A33W mutation shows the effect of this residue in effectively covering the entrance of the binding cavity (6) (**Figure IV-10**).

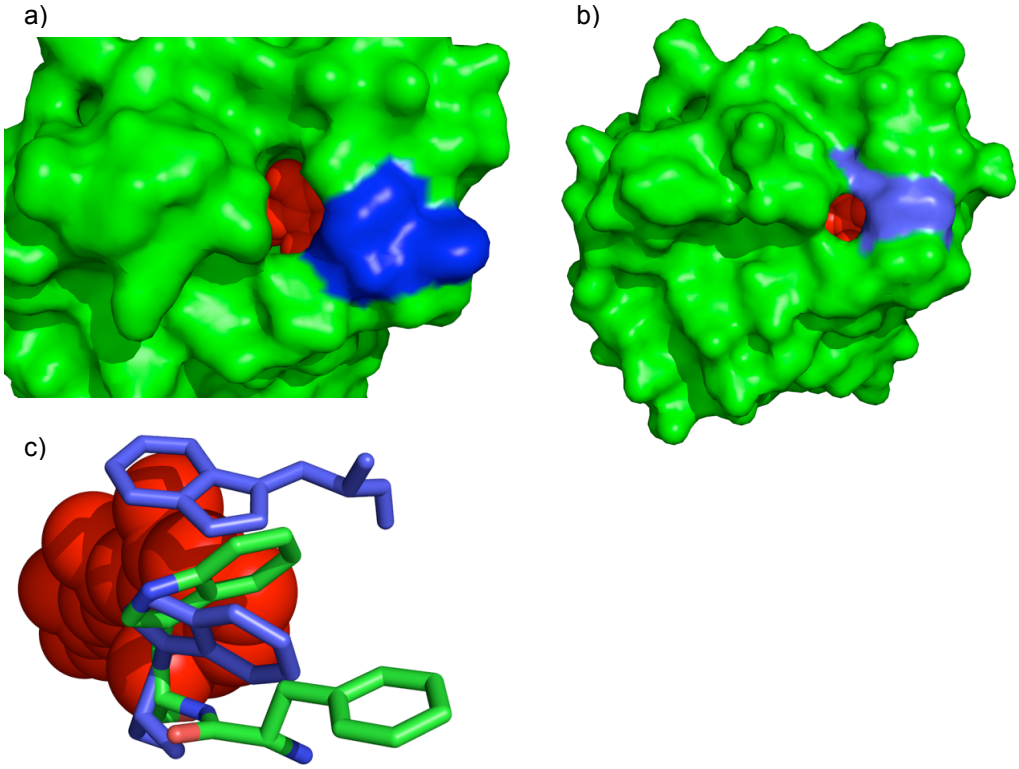


Figure IV-10) a) Surface representation of the Q108K:K40L:T51V:T53C:R58W:T29L:Y19W mutant, λ_{max} = 591 nm (PDB ID: 4EFG). b) Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W mutant, λ_{max} = 606 nm (PDB ID: 4EDE). c) A33W mutation helps to cover the entrance of the binding pocket more effectively.

IV.6 Counter anion in the hCRBPII system

One of the unique characteristics of hCRBPII retinal pigments is the pKa stabilization without a counteranion. For all of the mutants of this protein that have an absorption above 508 nm the retinal PSB has been achieved via the introduction of no counter anion. As mentioned in previous chapters in rhodopsin, residue E113 acts as a direct counter anion for the PSB rendering it with a pKa of 16 and in bacteriorhodopsin multiple Aspartic acid residues, which include the D85, 96 and 212 are involved in indirect hydrogen bonding to the PSB that increases the pKa of the PSB to 13 (10).

The water networks in hCRBPII not only govern its wavelength regulation but also the pKa of this system as well. Specially, the Q38 and Q128 residues play the most significant effect in both the pKa and wavelength absorbed by the protein. As mentioned above Q4 residue contributes to 1-1.5 pKa unit, Gln38 1-1.5 unit, Gln128 up to 4 units and R58Y mutation can increase the pKa by 1 unit. All of these mutations have a direct effect on the water binding pattern of the binding pocket (**Tables IV-1** to **IV-5**) (5, 6, 9).

In hCRBPII system pKas of up to 10.2 have been achieved by introducing no counter anion for the Schiff base, which is pretty novel (5) (Tables **IV-3** and **IV-5**).

However, achieving retinal pigments with absorption below 508 nm was not possible without the introduction of counter anion. The most blue shifted proteins in these series are the Q108K:T51D mutant absorbing at 474 nm and Q108K:T51D:L117E mutant absorbing at 425 nm. A crystal structure of the Q108K:K40D protein showed the introduced Aspartic acid residue in hydrogen bonding distance to the PSB. This residue can act as a counter anion or it probably accumulates waters near the PSB that causes the blue shift of the protein. Residue K40 stays close to the C12 of the retinal in some chains. In other chains it points right at the C14 carbon of retinal. It can be concluded that there might be some of the K40 retinal PSB in the structure as well with partial occupancy. If proven to be true with more mutagenesis and structural studies, studying the K40 PSBs can unveil novel findings on wavelength regulation and retinal binding in this system (**Figure IV-11**).

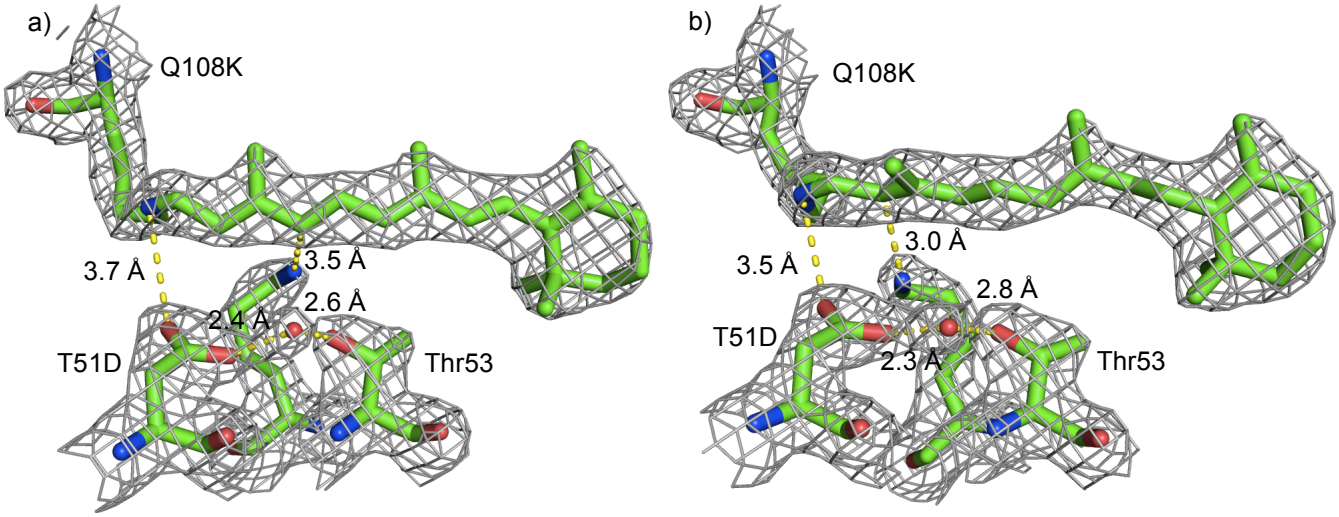


Figure IV-11) a) The retinal density in chain G of the Q108K:T51D crystal structure, Lys 40 points at C12 of retinal. b) The retinal density in chain L of the Q108K:T51D crystal structure, Lys 40 points at C14 of retinal. Map is contoured at σ = 1.0.

KD structure has a unique fold different than all of the other retinal bound structures of hCRBPII. Unlike, the other retinal bound structures that show the hCRBPII fold as a monomer, KD folds as a dimer. Two protein molecules each make an open monomer that fold together. An open monomer is the result of the severe torsion of the Phi and Psi angles of residues 50-60 which confers a different tertiary structure to the protein (9).

IV.7 Crystal growth for Neutron Diffraction

In order to understand the complete orientation of the water molecules in the binding pocket of the hCRBPII protein and their effect in the wavelength regulation of this protein, we aimed for neutron diffraction studies of this protein. Since the contribution of the hydrogens to the overall X-ray diffraction pattern is small, it is not possible to find the absolute orientation of the hydrogens from an X-ray electron diffraction pattern. However, unlike X-ray

crystallography, hydrogen's contribution to the overall neutron diffraction map is considerable and the position of hydrogens can be determined by this method. The bottleneck of the Neutron diffraction is the requirement for the growth of big crystals (11, 12).

hCRBPII grows as 0.2*0.1*0.2 mm crystals or smaller in hanging drops, which are quite small for Neutron Diffraction. For Neutron diffraction the volume of the crystals should be at least 0.15 mm^3. Besides this hCRBPII has a tendency of growing multiple lattices into each other during crystallization. Multiple lattices growing into each other become more problematic, when the formation of bigger crystals is required.

Three mutants were chosen for Neutron diffraction crystal growth studies based on their regular crystallization in H_2O hanging drops and X-ray diffraction quality. These include Q108K:K40L, Q108K:K40L:T51V:R58F and Q108K:K40L:T51V:R58Y:Y19W. Mutants with Q4X mutation were not chosen. In spite of better diffraction and better retinal density, they tend to produce less crystals that are clusters. Different crystallization techniques including hanging drop, sitting drop, crystallization under oil, micro seeding and macro seeding into hanging drop and sitting drop were tried.

IV.7-1 Q108K:K40L mutant of hCRBPII

In order to make sure that the binding of retinal to hCRBPII proteins in D_2O is not significantly different than H_2O, retinal binding assay in D_2O was done for this mutant, and as it is shown the kinetics of the binding is considerably slower, but there is no change in the protein absorption. In D_2O retinal binds slower and the PSB loss is also slower (**Figure IV-12**).

The first crystallization trials were done with 35 microliter polystyrene sitting drop micro-Bridges. 24 well VDX Hampton boxes were used and all of the solutions and buffers were prepared in D_2O. The protein solution was buffer exchanged with 10mM Tris.HCl D_2O pH=8.0 and 150 mM NaCl prior to retinal incubation and then concentrated to 6-8 mg/ml.

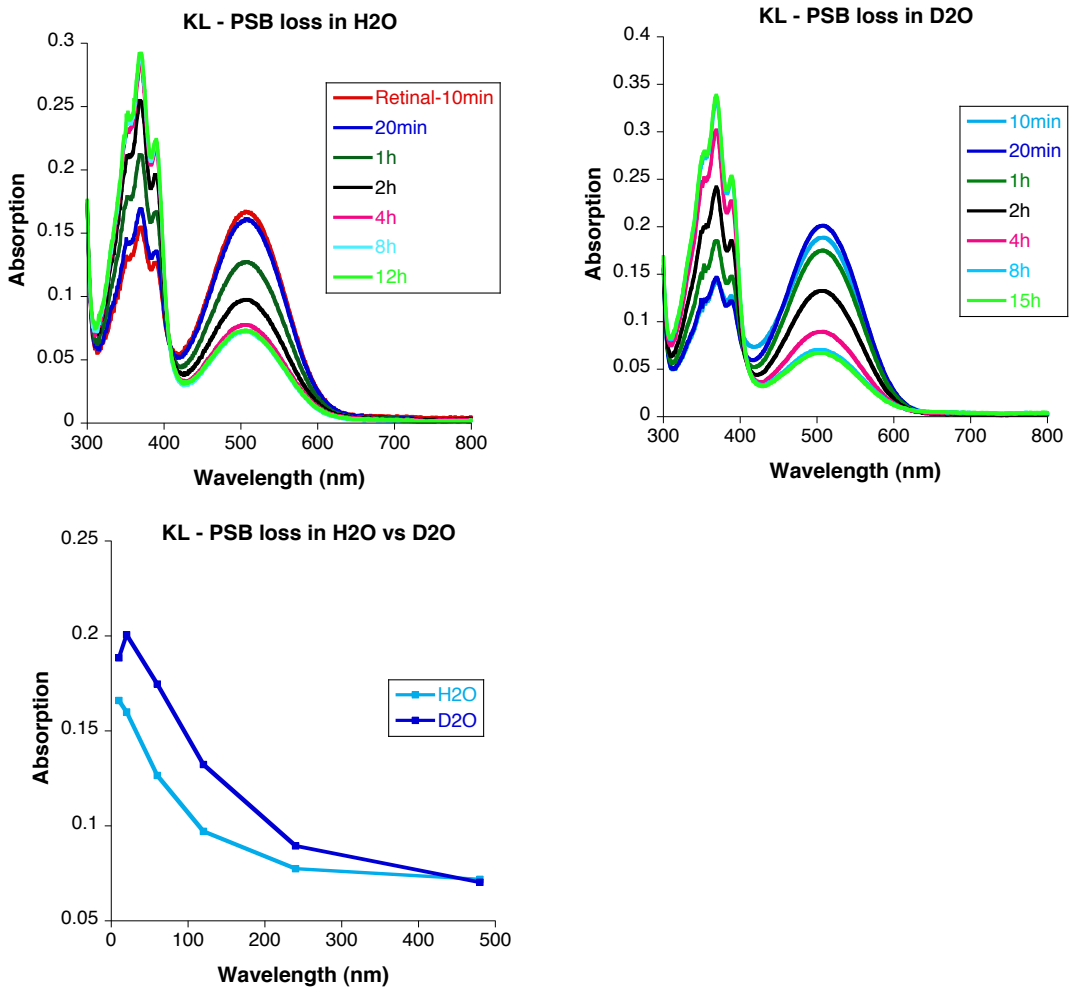


Figure IV-12) PSB loss of the mutant Q108K:K40L of hCRBPII in a) H_2O b) D_2O c) The kinetic of the PSB loss comparison in H_2O and D_2O.

The crystals were grown in 1-2 weeks. Unfortunately, the crystals grown in these drops were even smaller than the regular crystals in small hanging drops (**Figure IV-13a**). So, I decided to go back to the hanging drops and repeat the same crystallization strategies but this time with 2-3 microliter

hanging drops. The crystals grown this time were bigger than the sitting drop ones, but still not big enough for the studies (**Figure IV-13b to 13d**). The next plan was to increase the protein concentration (12 mg/ml), but screening into lower precipitation reagent (PEG 4000). This strategy proved to be effective in increasing the size of the crystals two-three times bigger than the ones with lower protein concentration and higher PEG 4000 (**Figure IV-13e**). Increasing the protein concentration to 20-25 mg/ml and screening into lower PEG with bigger drops provided a lot of crystals with the same size in 24 hours, but we were doubtful, if they are big enough for the Neutron diffraction studies (**Figure IV-13f**).

Since the original sitting drop trials were not successful, micro seeding and macro seeding experiments were done to see if we can get even bigger crystals with these methods. Unfortunately, this mutant provided just small size crystals with the micro seeding and a lot of nucleation around the original crystal in the macro seeding (**Figure IV-13g**).

Crystallization of Q108K:K40L mutant under Al's oil (Micro batch) in D_2O does not provide any better crystals compared to hanging drop or sitting drop.

At this point the only way to improve the size of the crystals seemed to be the bigger drops, which can provide more protein for a larger amount of protein deposition onto the crystals. Since the original sitting drop trials were just at 6-8 mg/ml, this time 20-25 mg/ml protein in 10,20 and 40 microliter sitting drops with polystyrene micro bridges were tried and it proved to be successful in making big size crystals. Also, it should be mentioned that a lot of these crystals still had multiple lattices grown into each other and not bigger than the hanging drop crystals. The next try was the crystallization of the protein

with larger sitting drops. For this purpose crystallization of the protein at 25-30 mg/ml with 80 microliter sitting drops on glass rods were tried. The crystals grown this way were the biggest crystals grown. They were as big as 1*1*0.5 mm in the size (**Figure IV-13h** and **IV-13i**). Some of these crystals were shot at Oak Ridge National Laboratory (ORNL). Although, the diffraction pattern of these crystals shows more than one lattice, because of the time limitations smaller crystals, which tend to be more single, were not tried (Figure **IV-13j** and **IV-13k**). Further data collection on smaller but single crystals can provide better diffraction data. Another experiment that was not tried for this mutant was the sitting drops on glass rods in H_2O rather than D_2O.

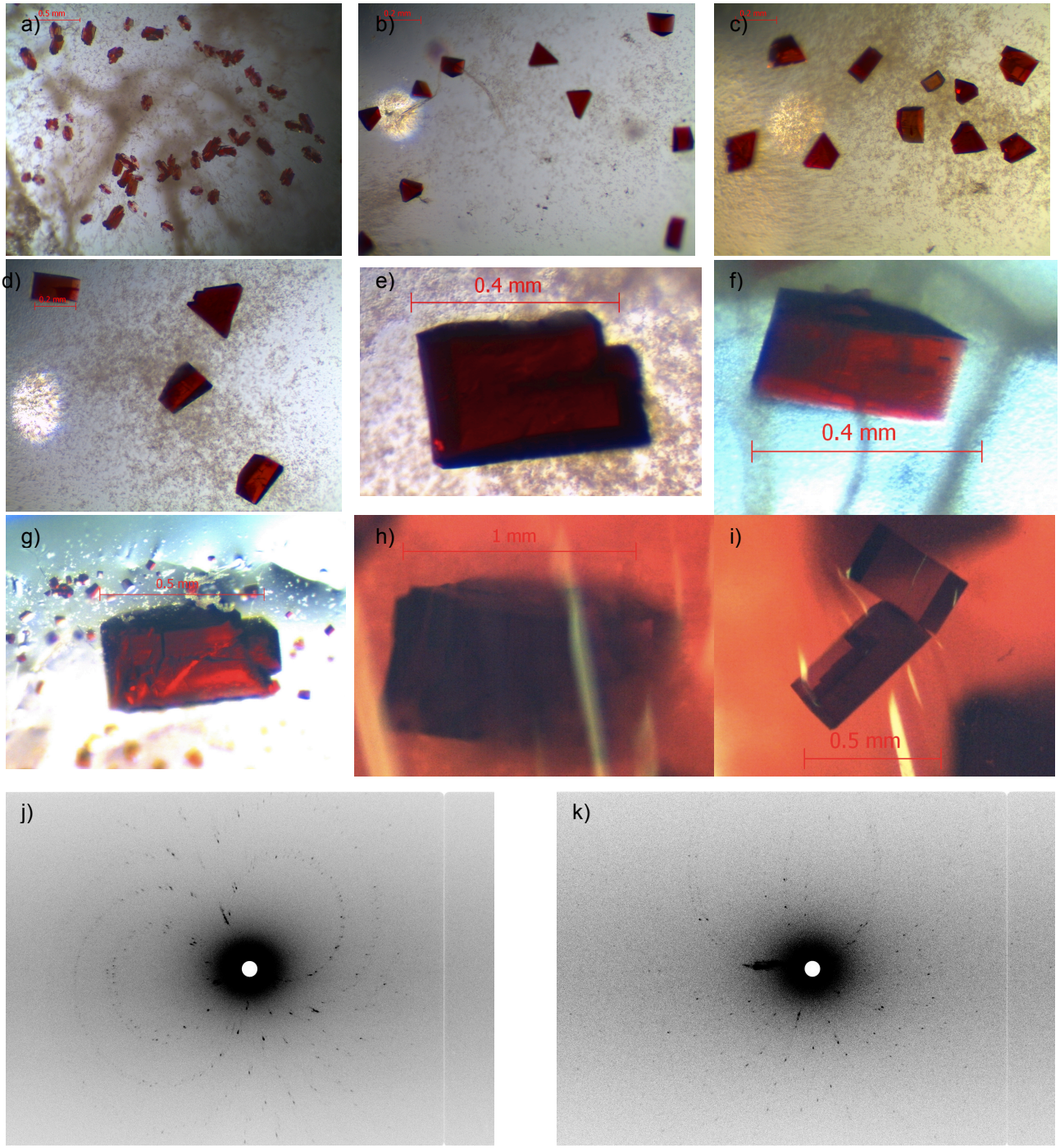


Figure IV-13) Crystallization trials of Q108K:K40L in D_2O. a) crystals from sitting drop at 6 mg/ml and 35 µl size. b,c,d) Crystals from hanging drop at 6-8 mg/ml. e) Crystals from hanging drop at 12 mg/ml by screening into lower PEG4000. f) Crystals from hanging drop at 20-25 mg/ml by screening into lower PEG4000. g) Macro seeding of the crystal in panel e. h and i) crystals from sitting drop at 25-30 mg/ml and 80 µl size. j and k) Neutron diffraction pattern of the crystal in panel h.

IV.7-2 Q108K:K40L:T51V:R58F mutant of hCRBPII

All the experimental procedures tried with Q108K:K40L mutant were tried with Q108K:K40L:T51V:R58F mutant as well. This mutant produces huge clusters with 40 and 80 microliter sitting drops in D_2O. Crystallization of this mutant under Al's oil also does not seem to produce any better crystals compared to the hanging drop crystallization.

10-15 microliter hanging drops with 20-25 mg/ml protein concentration with suitable PEG4000 concentration can provide single crystals in D_2O (**Figure IV-14a**). Micro seeding experiments in hanging drops also provided very small crystals (**Figure IV-14b**). Also, the crystallization of this mutant with sitting drops in H_2O might improve the quality of these crystals, but it was not tried.

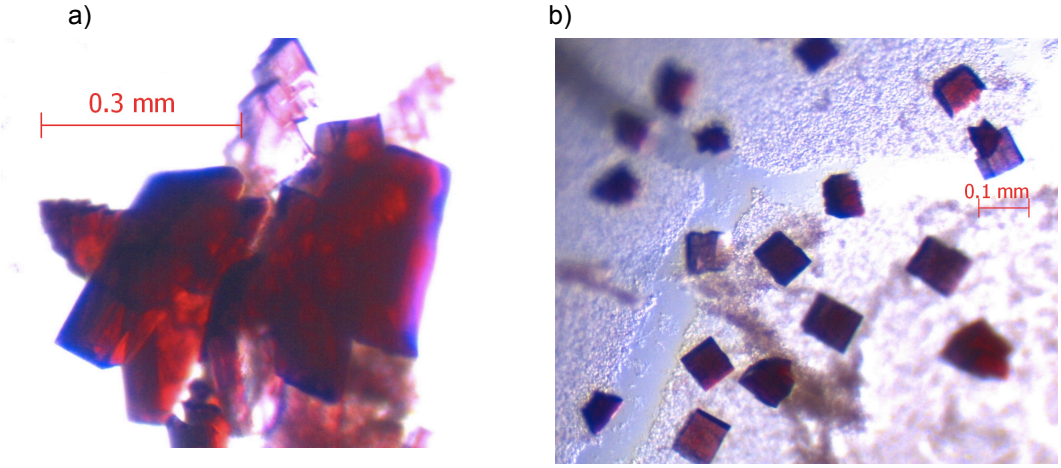


Figure IV-14) Crystallization of Q108K:K40L:T51V:R58F mutant of hCRBPII in D_2O: a) with 20-25 mg/ml protein in 10 µl hanging drops. b) microseeded drops provides smaller crystals.

IV.7-3 Q108K:K40L:T51V:R58Y:Y19W mutant of hCRBPII

Crystallization of this protein in D_2O with 40 and 80 microliter sitting drops at 20-25 mg/ml in polystyrene micro bridges and glass rods mainly is accompanied with a lot of protein precipitation. However, in H_2O with glass

rods and 80 microliter drops suitable size single crystals are formed. The Neutron diffraction of these crystals was not tried (**Figure IV-15**).

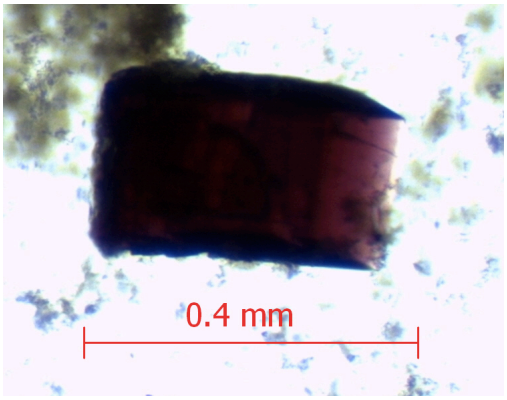
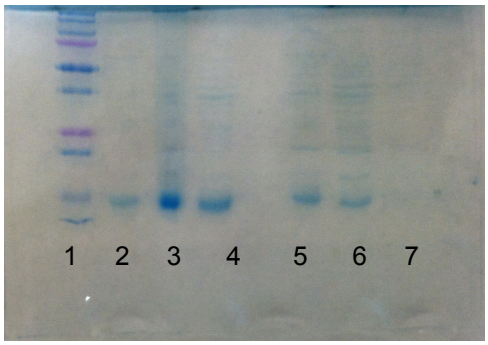


Figure IV-15) Crystal of Q108K:K40L:T51V:R58Y:Y19W mutant of hCRBPII in H_2O.

IV.7-4 hCRBPII Expression in Minimal Media

In order to deuterate the backbone of the protein for improving the signal to noise in neutron diffraction, the expression of the protein in a deuterated carbon source is required. The first step is to check the expression level in the same carbon source with hydrogen and in H_2O rather than D_2O. The expression protocol described in (11) was used. Colonies were grown in regular plates and transferred to minimal media plates for 36-48 hours for cell adaptation before transferring to liquid minimal media. Adaptation of the cells to minimal media on plates is important for protein expression. Cells were grown in 50 ml minimal media before 1 liter expression, and the OD_{600} was 2-3 before induction with 1mM IPTG for 24 hours at room temperature. The protein was then purified and analyzed by SDS-PAGE gel (**Figure IV-16**).

16a)



16b)

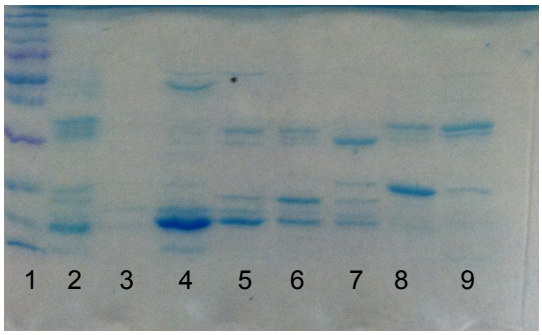


Figure IV-16) a) Expression of hCRBPII in regular and H_2O minimal media for mutant: Q108K:K40L:T51V:T53C:R58W:T29L:Y19W 1) MW marker 2,3,4) regular media, elutions from FastQ each 30ml 5,6,7) H_2O minimal media, elutions from FastQ, each 30ml 15b) Expression of Q108K:K40L in H_2O minimal media without cell adaptation 1) MW marker 2) Fast Q elution 3-9) Source Q elutions. Fraction 4 provides crystals.

IV.8 Wavelength Regulation in second generation hCRABPII mutants

IV.8-1 R59W and A32W mutations

Most of the wavelength regulation in hCRABPII comes from mutating residues toward the mouth of the binding pocket, which as described in hCRBPII makes a hydrophobic barrier for the passage of the waters into the binding pocket. Positions Arg 59 and Ala 32 are the ones that contribute significantly to red shifting the wavelength of this protein. An R59W mutation to enclose the binding cavity is accompanied by 22 nm red shift. The crystal structure of the R111K:R132L:Y134F:T54V:R59W mutant of hCRABPII showed that the entrance of the binding pocket of the protein is still wide open. To further enclose the binding cavity of the protein A32W mutation was added. Interestingly, this mutation red shifts the protein by 54 nm and the protein absorbs at 610 nm (**Table IV-8**).

Protein	λ_{max}(nm)
R111K:R132L:Y134F:T54V	534
R111K:R132L:Y134F:T54V:**R59W**	556
R111K:R132L:Y134F:T54V:**R59W**:**A32W**	610

Table IV-8) R59W and A32W mutations contribute significantly to the wavelength regulation of hCRABPII.

The crystal structure of the R111K:R132L:Y134F:T54V:R59W:A32W mutant shows that the entrance of the binding pocket is still open and not efficiently closed (Figure 17). Therefore, the importance of these mutations is not just in closing the entrance of the binding pocket. We hypothesized that the introduction of these mutations makes a hydrophobic barrier that inhibits the movement of water molecules into and out of the binding pocket of the protein, which is essential for wavelength tuning. In fact, these findings in hCRABPII has been in parallel helping us to mutate hCRBPII rationally as well as described.

These residues do not seem to be directly interacting with water molecules, but they decrease the global access of the waters from the solution into the protein (**Figure IV-17**).

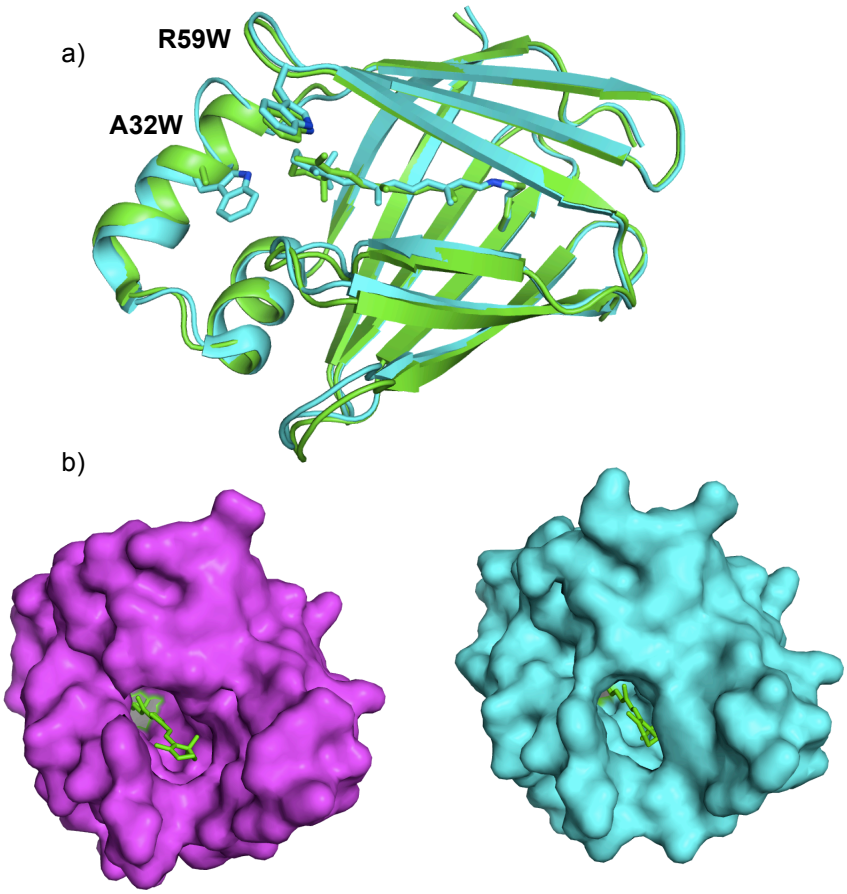


Figure IV-17) a) Crystal structure of R111K:R132L:Y134F:T54V:**R59W** at 2.6 Å (Green, PDB ID: 4I9S) λ_{max} = 556 nm and crystal structure of R111K:R132L:Y134F:T54V:**R59W**:**A32W** at 2.6 Å (Cyan, PDB ID: 4I9R) λ_{max} = 610 nm of hCRABPII overlaid. b) Surface representation of these mutants: Left: R111K:R132L:Y134F:T54V:**R59W** (magenta) and Right: R111K:R132L:Y134F:T54V:**R59W**:**A32W** (cyan).

IV.8-2 Y134F and T54V mutations

Originally the formation of PSB in second generation hCRABPII was achieved by four mutations R111K:R132L:Y134F:T54V. I tried to find the role of these mutations in lieu of P39Q and R132Q mutations described in chapter2, which improves the pKa of the protein, more stable PSBs and better diffracting crystals. The tetra mutant of R111K:R132Q:P39Q:R59Y provides complete PSB formation at pH=5.0 after 23 hours of retinal incubation with λ_{max} = 516 nm. The addition of Y134F mutation to this mutant does not change the absorption of the protein, but accelerates the PSB formation and eventually

the mutant R111K:Y134F:T54V:R132Q:P39Q:R59Y forms the PSB in 5 hours, and interestingly the addition of T54V mutation, which is the equal to Thr 51 in hCRBPII, red shifts the protein by 25 nm to absorb at 546 nm. The addition of A32W mutation also red shifts the protein by 34 nm (**Figure IV-18**).

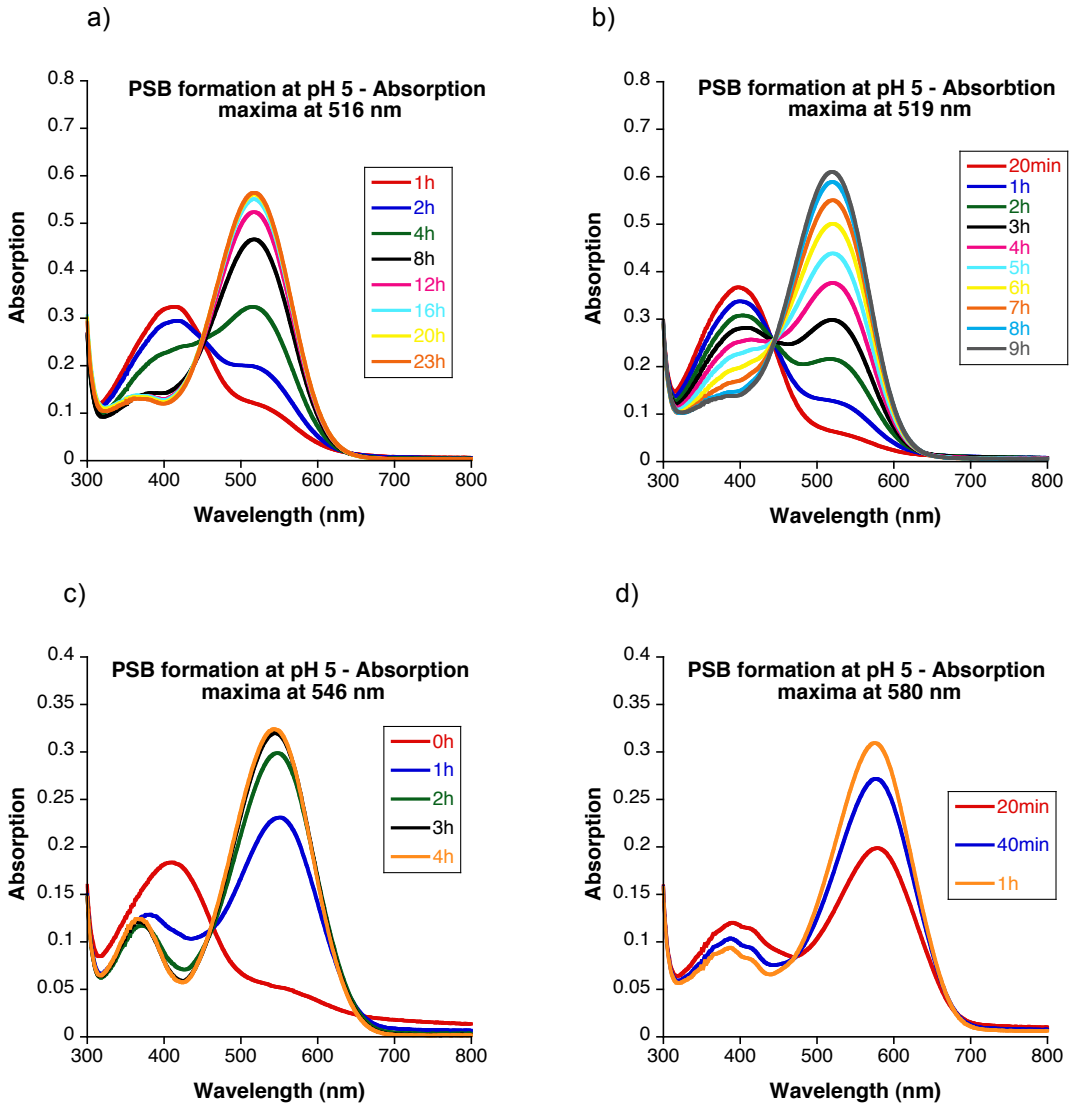


Figure IV-18) Retinal incubation of hCRABPII mutants in citrate buffer at pH5 for mutants: a) R111K:R132Q:P39Q:R59Y b) R111K:Y134F:R132Q:P39Q:R59Y c) R111K:Y134F:T54V: R132Q:P39Q:R59Y d) R111K:Y134F:T54V:R132Q:P39Q:R59Y:A32W.

It should be mentioned that the Y134F and T54V mutations favor the formation of the PSB at the physiological pH (**Figure IV-19**). This might be because of the overall decrease in the polarity of the binding pocket. Y134F mutation in the context of R132Q:P39Y mutation has an important effect in red shifting the protein by 55 nm, but no effect in the context of R132Q:P39Q (Figure **IV-18** and **IV-19**).

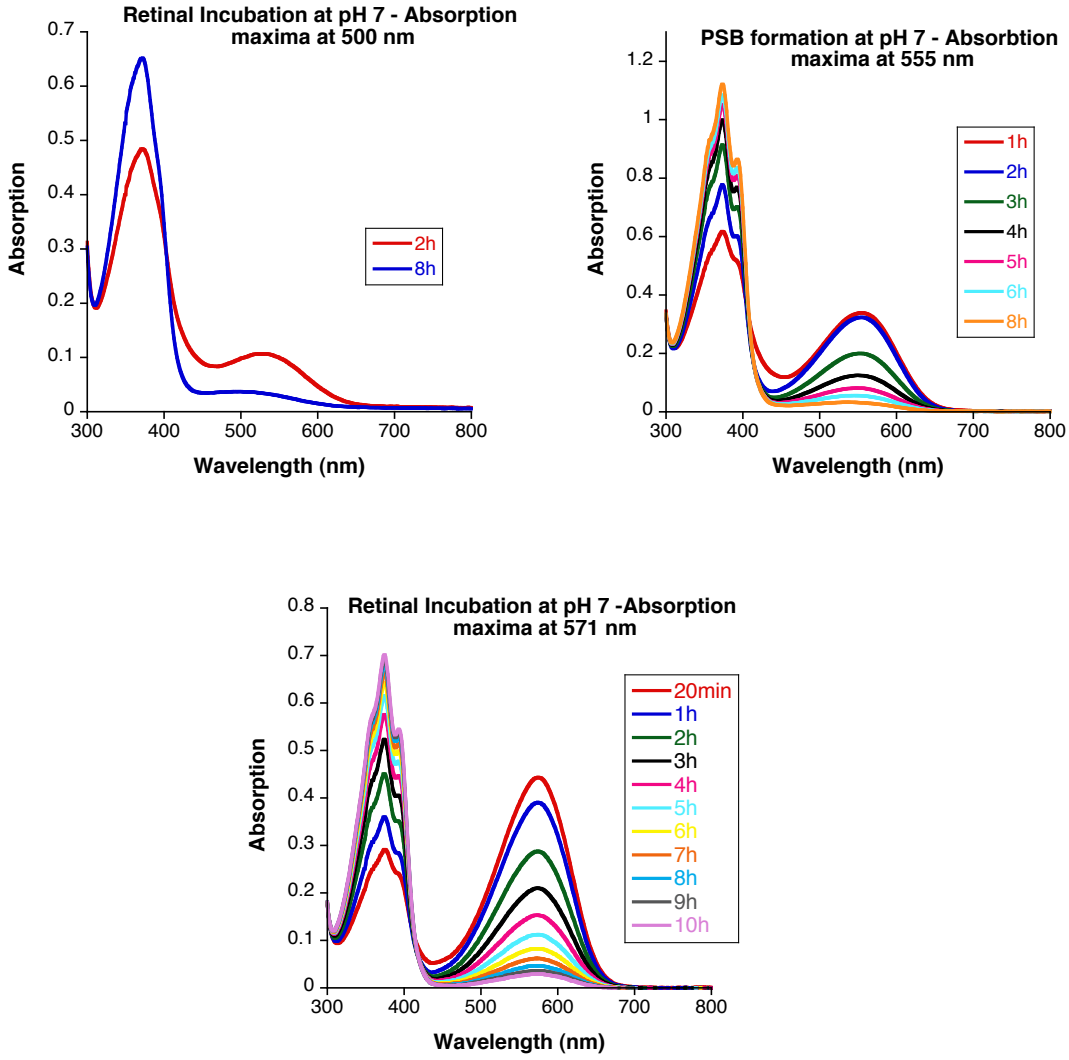


Figure IV-19) The retinal incubation of hCRABPII mutants at physiological pH: a) R111K:R132Q:P39Y:R59Y b) R111K:Y134F: R132Q:P39Y:R59Y c) R111K:Y134F: T54V:R132Q:P39Y:R59Y.

Also the acid titration of this protein shows different PSB absorbing in a range from 510 nm to 555 nm. A satisfactory answer to the effects of the Y134F cannot be achieved without further experiments. Comparing Figure 18 and 19, it is apparent that a P39Y mutation can red shift the protein by 25-30 nm compared to the mutants that have the P39Q mutation in the same position.

IV.8-3 P39Q, R132Q and F3Q mutations

The mutation of Gln 38 and Gln 128 in hCRBPII causes big blue shifts as the negative dipole moment toward the end of the polyene is removed. Since the equivalent positions in hCRABPII are Pro39 and Arg132, the mutations of these residues to Gln supposedly should increase the negative dipole moment toward the end of the polyene by the introduction of these polar residues and red shift the absorption, but counter intuitively these mutations do not cause any wavelength shift or usually blue shift the protein by 10-15 nm (**Table IV-9**).

Protein	**λ_{max}(nm)**
R111K:R132L:Y134F:T54V:R59W	556
R111K:Y134F:T54V:**R132Q:P39Q**:R59W	555
R111K:R132L:Y134F:T54V:R59Y	561
R111K:Y134F:T54V:**R132Q:P39Q**:R59Y	556
R111K:R132L:Y134F:T54V:R59Y:A32W	591
R111K:Y134F:T54V:**R132Q:P39Q**:R59Y:A32W	584
R111K:R132L:Y134F:T54V:R59W:A32W	610
R111K:R132L:Y134F:T54V:**R132Q:P39Q**:R59W:A32W	598

Table IV-9. The effect of the R132Q:P39Q mutation on the wavelength absorbed by the hCRABPII mutants.

We have learned in our hCRBPII system that equal distribution of the charge is the key to get super red shifted for the mutants of this protein, the blue shift by the introduction of the P39Q and R132Q mutations in hCRABPII might be

because of the fact that compared to the parent mutant the P39Q and R132Q mutations localize the charge and therefore get slightly blue shifted. Additionally, the crystal structures of the mutants with the P39Q and R132Q mutations shows that unlike hCRBPII, the two glutamines are separated via the water molecules in between them (**Figure IV-20d**), here the glutamine residues are close to each other for hydrogen bonding to one another. The electrostatic effects of the two glutamines on retinal will more likely cancel out as they face each other from the opposite ends of their amidic groups (**Figure IV-20**).

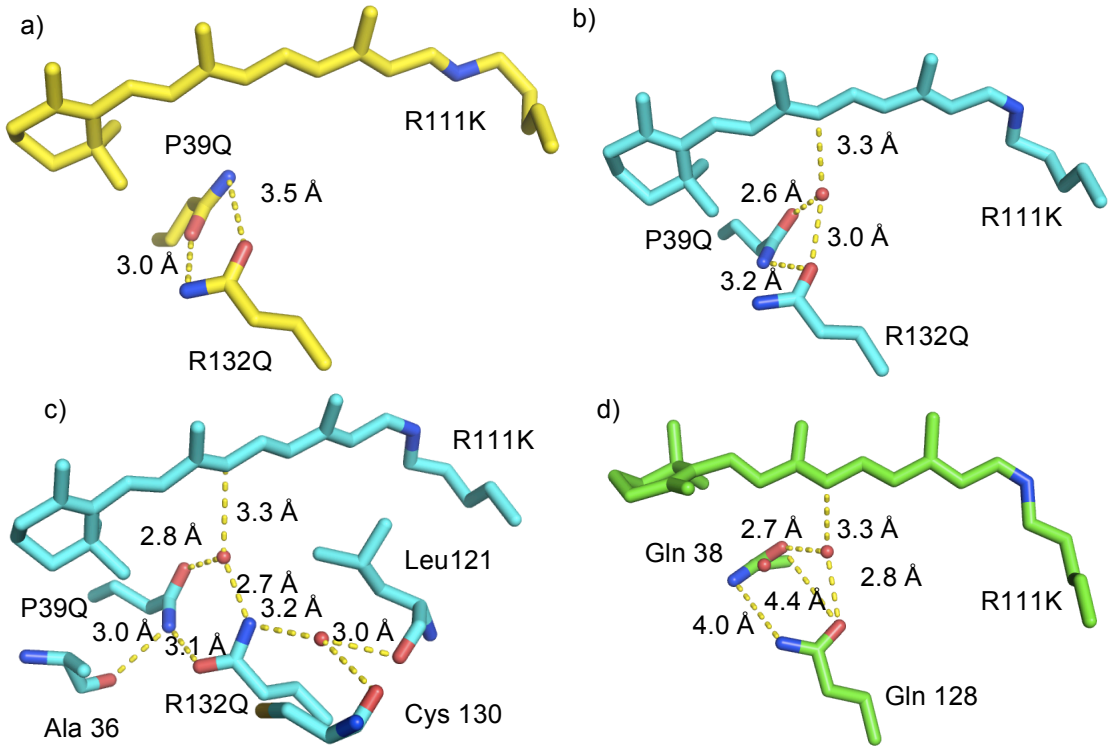


Figure IV-20) a) The orientation of the P39Q and R132Q in the R111K:Y134F:T54V:R132Q: P39Q:R59Y mutant of hCRABPII (PDB ID: 4YBP) b,c) The orientation of the P39Q and R132Q in the R111K:Y134F: T54V:R132Q:P39Q:R59Y:A32W:F3Q mutant of hCRABPII in the two chains of the protein (PDB ID:4YKM). d) The orientation of the Gln 38 and Gln 128 in the Q108K:K40L mutant of hCRBPII (PDB ID: 4RUU).

The addition of the F3Q mutation to these mutants blue shifts the protein by 25 nm, which is expected as the polar glutamine localizes the charge around the retinal iminium PSB (Table **IV-10**).

Protein	λ_{max}(nm)
R111K:Y134F:T54V:R132Q:P39Q:R59Y:A32Y	575
R111K:Y134F:T54V:R132Q:P39Q:R59Y:A32Y:**F3Q**	545
R111K:Y134F:T54V:R132Q:P39Q:R59Y:A32W	584
R111K:Y134F:T54V:R132Q:P39Q:R59Y:A32W:**F3Q**	557
R111K:R132L:Y134F:T54V:R132Q:P39Q:R59W:A32W	598
R111K:R132L:Y134F:T54V:R132Q:P39Q:R59W:A32W:**F3Q**	571

Table IV-10. The effect of the F3Q mutation on the wavelength absorbed by the hCRABPII mutants.

IV.9 Conclusion

Water molecules play a crucial role in the wavelength regulation of reengineered hCRBPII. Mutating the interacting residues with these waters around the Schiff base or at the beginning of the polyene combined with mutations at the entrance of the binding pocket to decrease the global access of the waters to the binding pocket can red shift the protein from 508 to 644 nm. Mutating the interacting residues with the water molecules toward the end of the polyene can wipe out the whole red shift of the protein from 644 to 510 nm. Neutron diffraction studies are in progress to find the absolute orientation of these waters inside the binding pocket of the protein. We have shown that the wavelength regulation in hCRABPII follows the same principles. Most of the wavelength regulation of this protein comes from the mutation of the residues at the entrance of the binding pocket of the protein. While mutating the polar residues or introducing them toward the Schiff base or at the beginning of the polyene like Y134F and T54V and F3Q changes the wavelength absorbed by the protein, introducing glutamine mutations toward

the end of the retinal polyene in hCRABPII does not change the wavelength of the protein by more than 10 nm. The crystal structures of hCRABPII shows that unlike hCRBPII, which glutamines are separated by waters and do not interact with each other, these residues directly hydrogen bond to one an other and probably cancel each other's dipole moment to have a slight effect on the wavelength of the protein.

IV.10 Other Structural Studies of hCRABPII and hCRBPII Reengineering

During my PhD, I also studied the structural studies of hCRABPII and hCRBPII in collaboration with Dr.Yapici and Dr.Berbasova. These studies included:

(1) The crystallization of hCRABPII proteins with Merocyanine in the development of an in-vivo fluorescent protein. The crystal structures provided invaluable understanding of the Quantum Yield of these pigments that correlates to the mode of the binding of the fluorophore with the protein (Figure **IV-21**) (13).

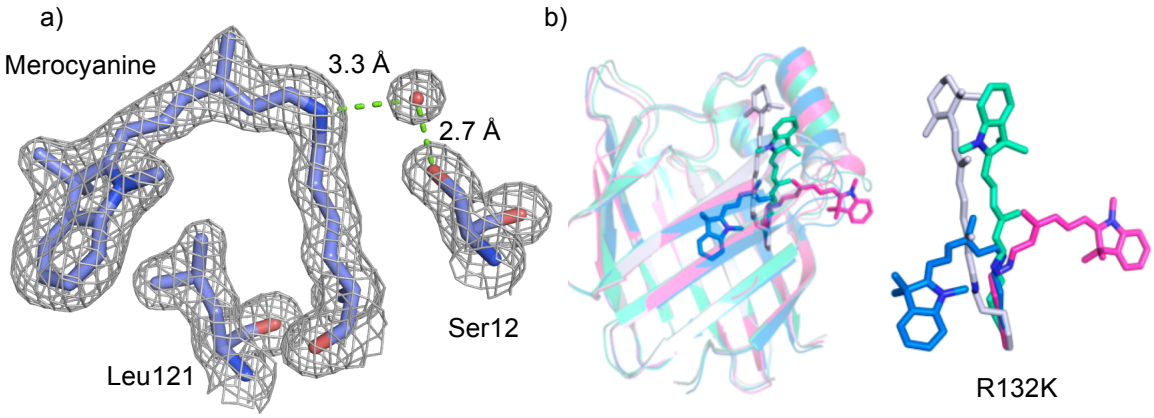


Figure IV-21) a) Crystal structure of the R132K:K111L mutant of hCRABPII with Merocyanine (PDB ID: 4QGV). b) Different modes of the binding of Merocyanine inside the binding pocket of hCRABPII (PDB IDs: Merocyanine: 4QGV, 4QGX and 3FEP, Retinal: 2G7B).

(2) The structural studies of the wavelength regulation of retinal analogs with hCRBPII. The retinal analogs were synthesized and assayed with proteins by Dr.Yapici and the structural data confirmed that these analogs bind in the predicted conformation. The focus was on the rotation around the ionone ring of the retinal and its effect on the wavelength regulation of hCRBPII protein. The 6s-*trans* retinal and 6s-*cis* retinal conformations were confirmed by high resolution crystal structures (**Figure IV-22**) (14).

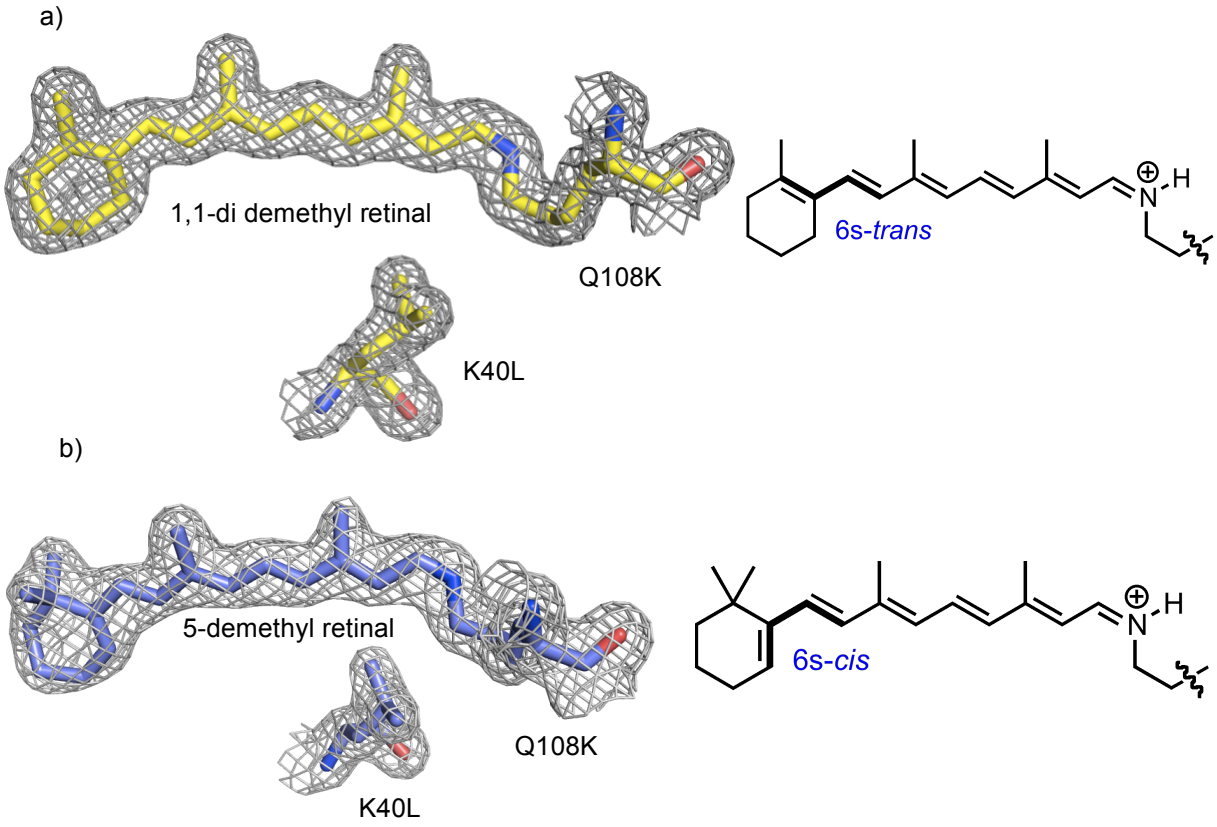


Figure IV-22) The crystal structure of the Q108K:K40L:T51V:R58Y:Y19W mutant of hCRBPII with: a) 1,1-di demethyl retinal in 6s-*trans* retinal conformation and b) 5-de methyl retinal in 6s-*cis* conformation.

(3) The structural studies of the hCRBPII protein with Fluorophore ligand Julolidine for the development of an in-vivo pH sensor. The introduction of a glutamic acid residue in the vicinity of the Julolidine PSB proved to make a

titratable residue that in the anionic and protonated form can have different absorptions. Different emission ratios based on the excitation of the protein at two wavelengths that it absorbs can be used to address pH. This can be the development of a single protein ratiometric fluorescent pH sensor (**Figure IV-23**) (15).

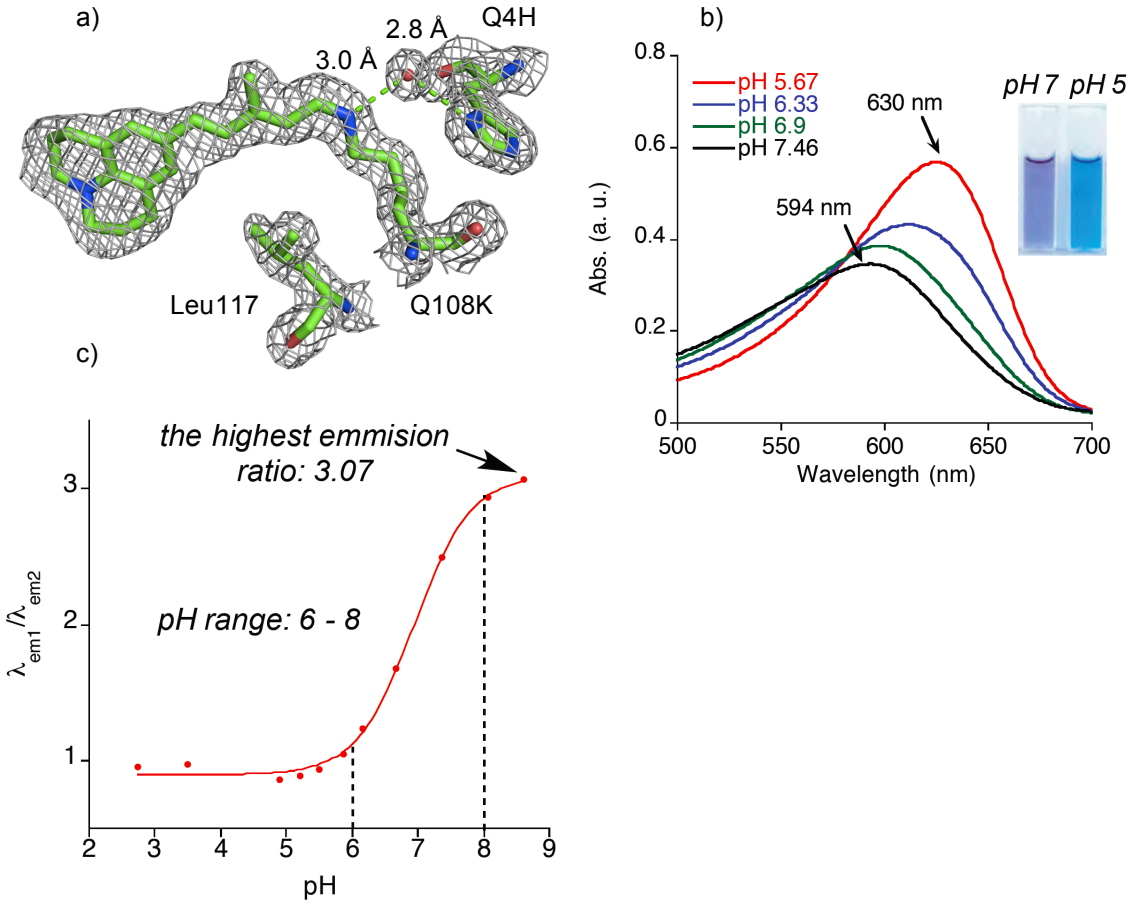


Figure IV-23) a) Crystal structure of the Q108K:K40L:T51V:T53C:R58W:T29L: Y19W:Q4H mutant of hCRBPII with Julolidene b) pH dependent change of the absorption of the protein with the introduction of L117E mutation c) pH dependent fluorescence ratio from excitation at the two absorption peaks of the protein.

REFERENCES

REFERENCES

1. Vasileiou C, *et al.* (2007) Protein design: Reengineering Cellular Retinoic Acid Binding Protein II into a rhodopsin protein mimic. *Journal of the American Chemical Society* 129(19):6140-6148.

2. Crow JA & Ong DE (1985) Cell-specific immunohistochemical localization of a cellular retinol-binding protein (type two) in the small intestine of rat. *Proceedings of the National Academy of Sciences of the United States of America* 82(14):4707-4711.

3. Vogel S, *et al.* (2001) Characterization of a new member of the fatty acid-binding protein family that binds all-trans-retinol. *The Journal of biological chemistry* 276(2):1353-1360.

4. Nossoni Z, *et al.* (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal. *Acta crystallographica. Section D, Biological crystallography* 70(Pt 12):3226-3232.

5. Wang W (2012) Protein design: Reengineering of cellular retinol binding protein II (CRBPII) into a rhodopsin mimic, functionalization of CRBPII into a fluorescent protein tag and design of a photoswitchable protein tag. PhD (Michigan State University).

6. Wang WJ, *et al.* (2012) Tuning the Electronic Absorption of Protein-Embedded All-trans-Retinal. *Science* 338(6112):1340-1343.

7. Sakmar TP (2012) Redder Than Red. *Science* 338(6112):1299-1300.

8. Wang W, Geiger JH, & Borhan B (2014) The photochemical determinants of color vision: revealing how opsins tune their chromophore's absorption wavelength. *BioEssays : news and reviews in molecular, cellular and developmental biology* 36(1):65-74.

9. Nossoni Z (2014) 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 HCRBPII. PhD (Michigan State University).

10. Ernst OP, *et al.* (2014) Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. *Chemical reviews* 114(1):126-163.

11. Meilleur F, Weiss KL, & Myles DAA (2009) Deuterium Labeling for Neutron Structure-Function-Dynamics Analysis. *Methods Mol Biol* 544:281-292.

12. Meilleur F, *et al.* (2013) The IMAGINE instrument: first neutron protein structure and new capabilities for neutron macromolecular crystallography. *Acta Crystallogr D* 69:2157-2160.

13. Yapici I, *et al.* (2015) "Turn-on" protein fluorescence: in situ formation of cyanine dyes. *Journal of the American Chemical Society* 137(3):1073-1080.

14. Yapici I (2015) Design of novel protein/chromophore complexes: Synthesis and evaluation of chromophores, fluorophores, and engineering of protein host systems. PhD (Michigan State University).

15. Berbasova T (2014) Engineering of the small cytosolic retinoid binding proteins into pH-sensing probes and novel fluorescent protein tags. Ph. D. (Michigan State University, ProQuest, UMI Dissertations Publishing).

Chapter V: Stabilizing the interaction between Cellular Retinoic Acid Binding Protein II and Retinoic Acid Receptor

V.1 Introduction

Retinoic Acid Receptors (RARs) belong to the superfamily of Nuclear Receptors. They are transcription factors that play a significant role in cell growth, differentiation, homeostasis and apoptosis (1-3). They are one of the targets for the suppression of different types of cancer (3-9). Structurally these proteins have multiple subunits in their sequence. The protein binds to the DNA through the DNA Binding Domain (DBD) of the protein and the Ligand Binding Domain (LBD) of the protein mediates the transcriptional activity (4, 10) (**Figure V-1**).

Figure V-1) Structural organization of Retinoic Acid Receptors.

The Ligand Binding domain of this protein includes eleven α-helices, which forms an antiparallel sandwich. RARs as their name indicates accept all-*trans*-Retinoic Acid as their ligand (10) (**Figure V-2**).

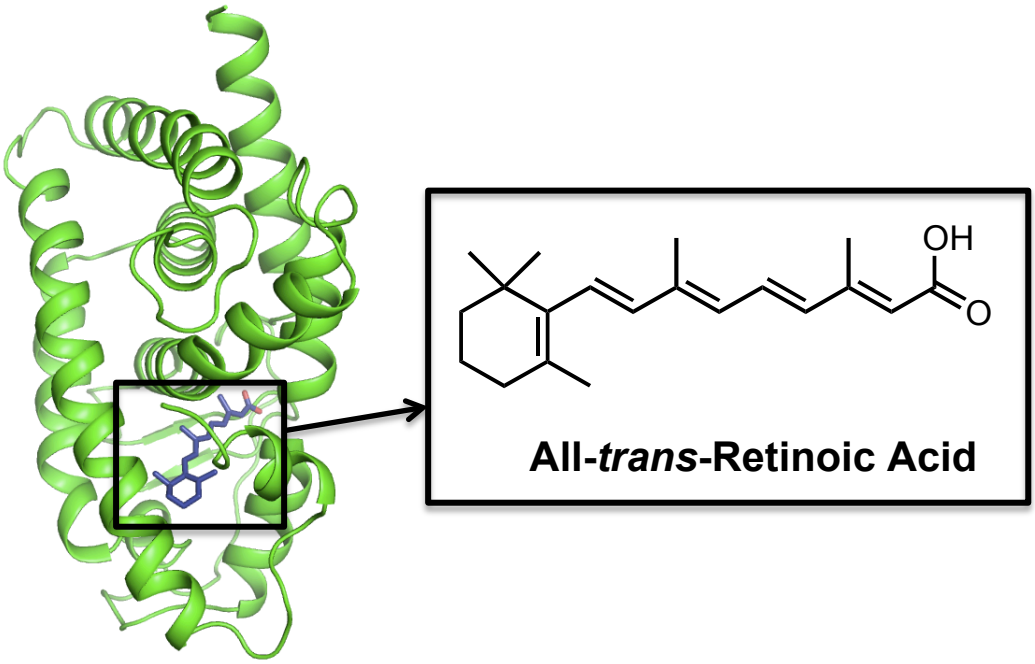


Figure V-2) a) Crystal structure of the Ligand Binding Domain of holo-human RAR-γ shows the Retinoic Acid inside the binding pocket of the protein (PDB ID: 2LBD).

RARs supply of Retinoic Acid comes from Cellular Retinoic Acid Binding Proteins (CRABP). Noa Noy studied the transfer of the Retinoic acid from CRABPs to RARs and multiple publications elucidated how this process happens. The main focus of these studies were:

(1) Elucidating the mechanism of the transfer (11-13).

(2) Identifying the important residues for protein-protein interactions (12).

(3) Stabilizing the protein-protein interaction (13).

(4) Identifying the nuclear localization signal (14) and the mechanism of the transfer from cytosol to the cell nucleolus (15).

They indicated that from the two isoforms of the CRABP, CRABPI transfers the Retinoic acid to RAR via a dissociation mechanism, independent from the RAR concentration. However, CRABPII goes through direct protein-protein interaction to deliver the ligand to RAR. Different biological assays were

performed to characterize CRABPII and RAR interaction for Retinoic acid transfer which included the chemical cross linking, electrophoresis under non-denaturing condition, fluorescence anisotropy titrations and electrophoretic mobility shift assays. The complex formation was not detected under any of these conditions (11).

Electrostatic calculations highlighted three residues that differ significantly in CRABPI and CRABPII surface potential from one CRABP isoform to the other. These include the Q75, P81 and K102 in CRABPII sequence, which the equivalent residues in CRABPI are E75, K81 and E102. Mutating these residues of CRABPI to equivalent residues in CRABPII provided a protein that was able to directly interact with RAR. On the other hand, mutating these residues of hCRABPII to equivalent residues in hCRABPI provided a protein that was not able to directly interact with RAR (12).

The translocalization of CRABPII to the cell nucleus for retinoic acid delivery to RAR was further proved by in-vivo fluorescence imaging, while the result of this experiment for CRABPI was negative. Eventually, a novel idea came in handy to be the only known experiment so far to visualize the two proteins together on an SDS-PAGE gel. The rationalization was that the CRABPII-RAR interaction is so transient that is hard to trap the complex of the two proteins, and it comes from the higher affinity of the RAR for Retinoic acid compared to CRABPII. The two proteins interact transiently that CRABPII delivers the Retinoic acid to RAR immediately and leaves. The idea was to design a ligand with high affinity for CRABPII and lower affinity for RAR. That way the two proteins will interact, but since the ligand will not partition to RAR easily, the two proteins will stay together longer. The ligand of choice was a

molecule named CD-270, which has 100 times more affinity for CRABPII compared to RAR, while Retinoic acid has nanomolar or sub nanomolar affinity for the both proteins. It should be mentioned that the in-vivo fluorescence imaging experiment of CRABPII with CD-270 shows the translocaliztion of the CRABPII to the nucleus as well.

A pull down assay in the presence of the Retinoic Acid or CD-270 was performed. The two proteins were pulled in the same fraction while CD-270 was used and different fractions when the Retinoic Acid was used (**Figure V-3**) (13).

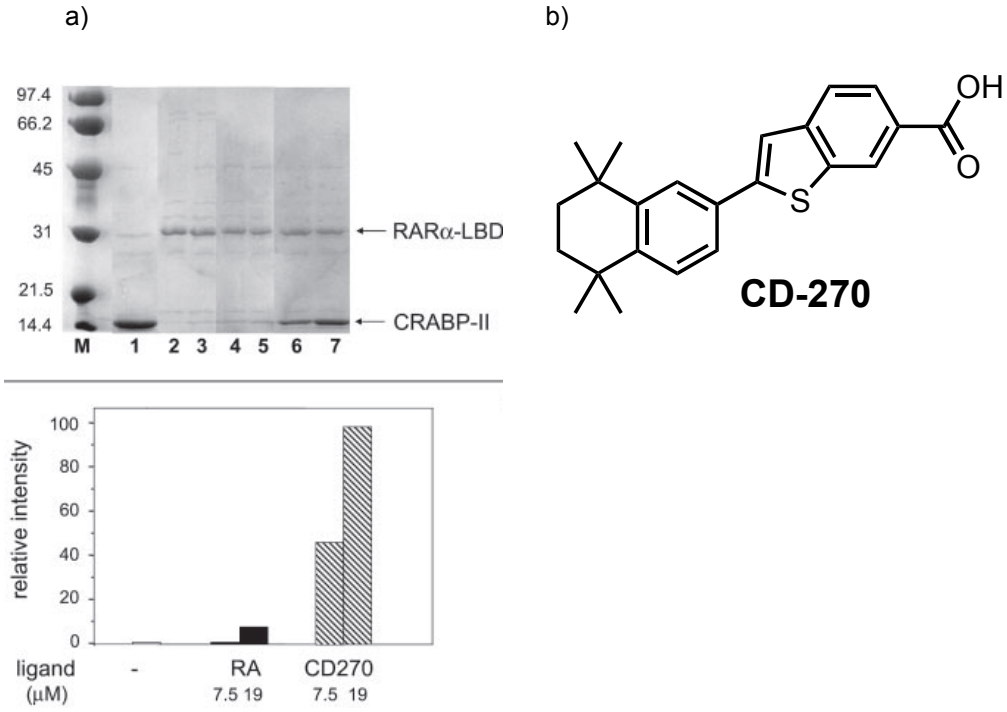


Figure V-3) a) The result of the pull down assay for the Retinoic acid and CD 270 M) MW marker 1) CRABPII at 50% percent of total input. 2) RARα-LBD on the Ni-NTA resin. 3) RARα-LBD incubated with CRABPII in the absence of ligand 4,5) RARα-LBD incubated with CRABPII in the presence of 7.5 and 18.75 μM Retinoic acid, respectively. 6,7) RARα-LBD incubated with CRABPII in the presence of 7.5 and 18.75 μM of CD-270, respectively. b) Chemical structure of CD-270.

V.2 Synthesis of CD-270

We were interested to visualize the interaction of the CRABPII and RAR with an atomic resolution description. To this end the synthesis of the CD-270 was targeted. The synthesis of the ligand from original patent pathway tried by Dr.Berbasova was not successful (**Figure V-4**) (16).

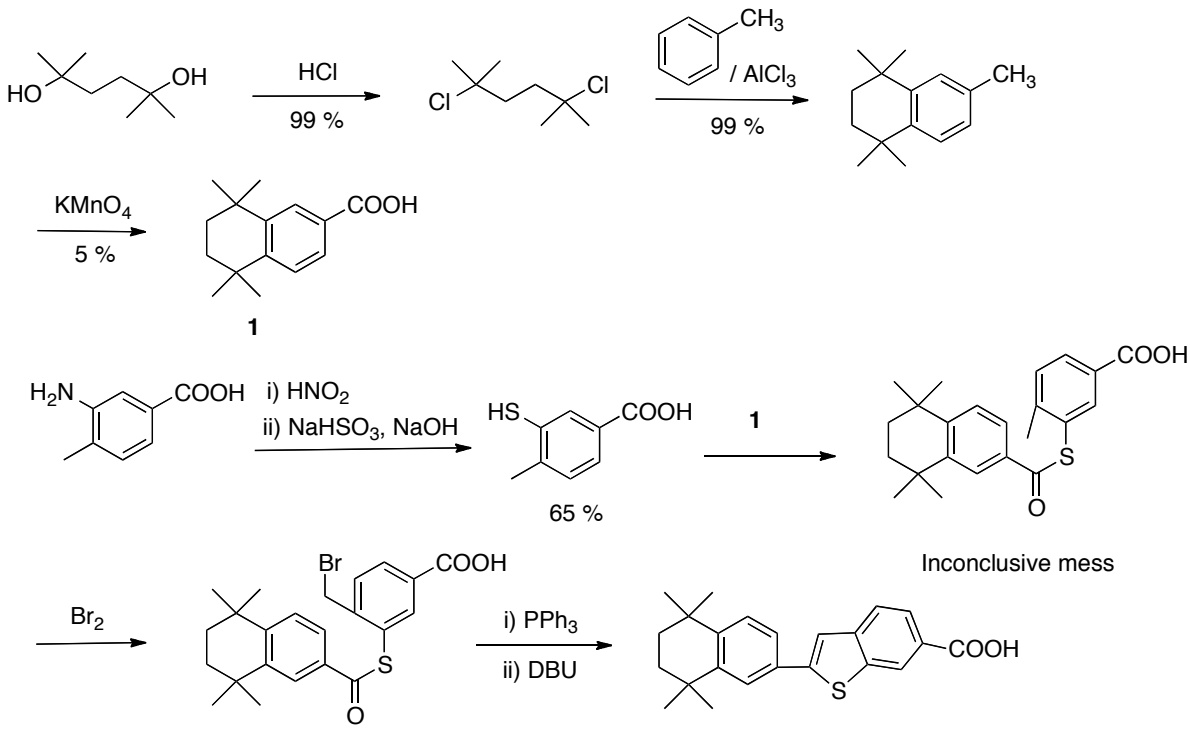


Figure V-4) CD-270 synthesis trial with the original proposed synthesis was not successful (With Dr.Tetyana Berbasova's permission).

A closer search in the literature showed the synthesis of benzothiophene molecules similar to CD-270 via the progress in the Suzuki cross-coupling reactions (17). A retro synthesis strategy was designed and the CD-270 was broken to two halves (**Figure V-5**).

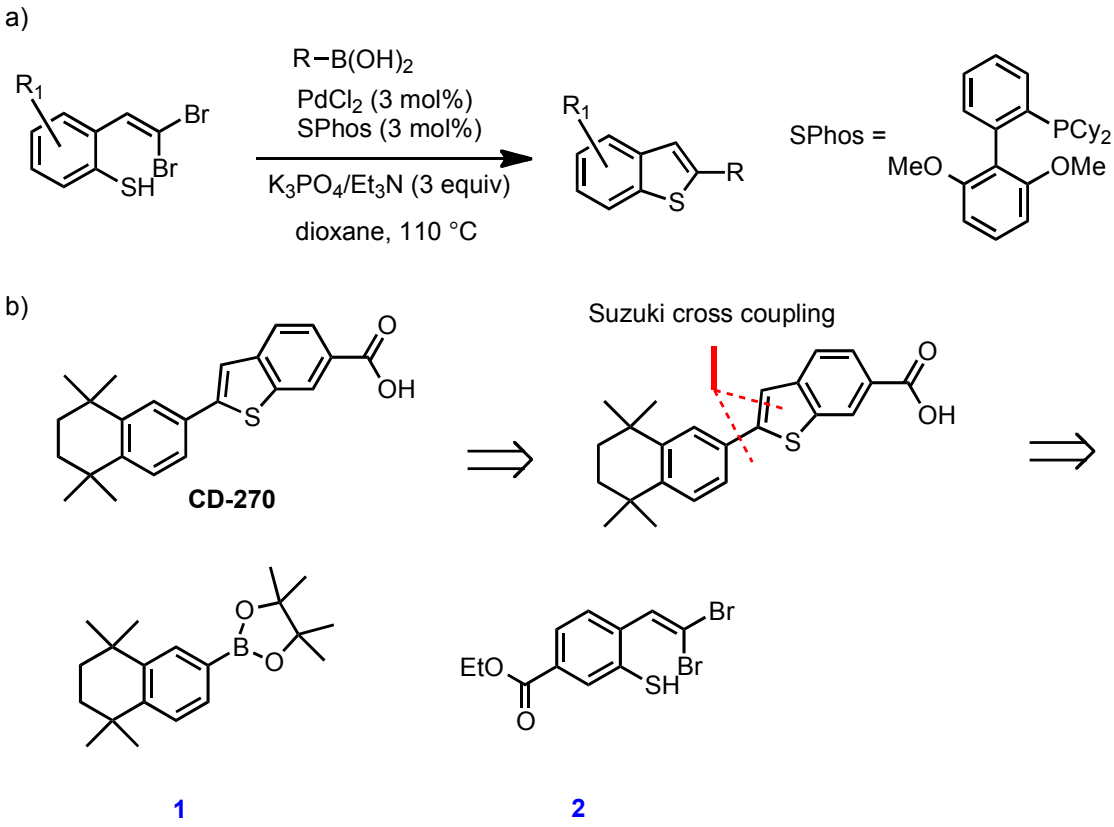


Figure V-5) a) Reported modified Suzuki cross coupling reaction b) proposed retrosynthesis for CD-270.

The synthesis of the first half was performed successfully via an iridium C-H activation reaction developed in Prof.Smith lab over the last two decades in the chemistry Department at Michigan State University (**Figure V-6**) (18-22).

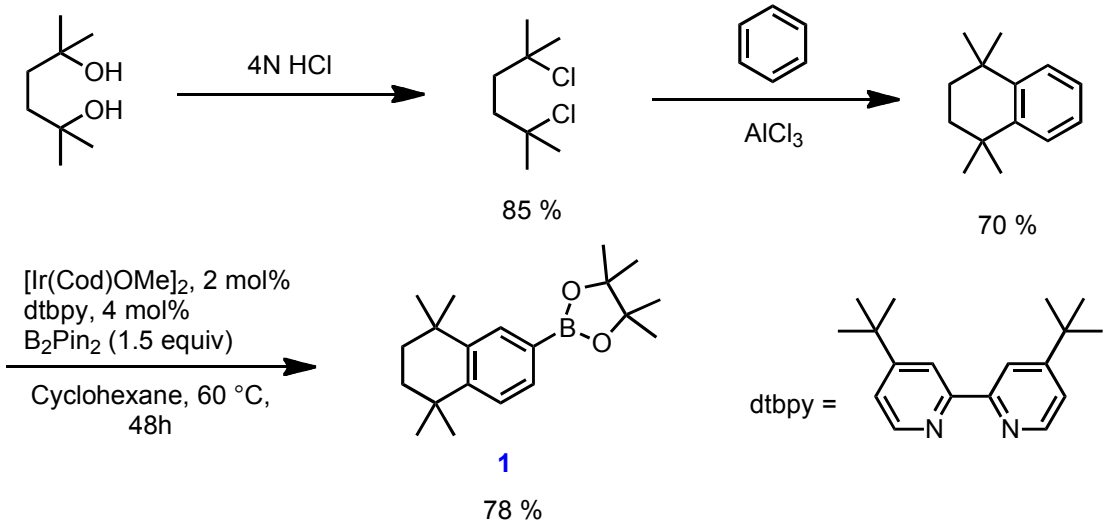


Figure V-6) The synthesis of the first fragment of the CD-270.

The synthesis of the second half proved to be more challenging. The first two steps of the synthesis of this fragment worked. Also, the protection of the aldehyde worked the best with the reagent shown (23) (**Figure V-7**).

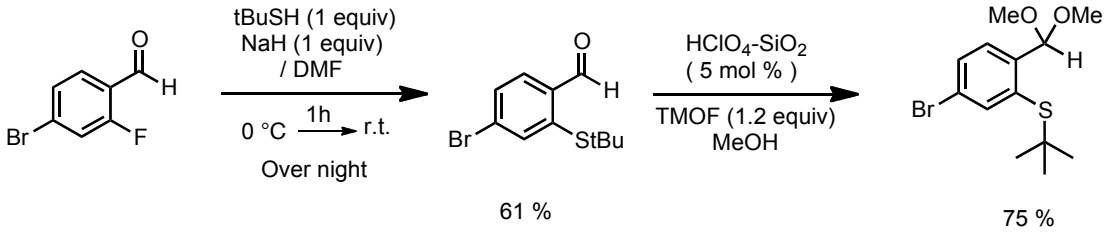


Figure V-7) The first two steps in the synthesis of the second fragment.

The rest of the synthetic plan is as shown in the following scheme, which was not tried as the first step did not work. The Grignard formation followed by carboxylation or organo lithiation followed by ethyl chloroformate addition did not work to carboxylate the aromatic ring. At this point the synthesis of the second fragment was not troubleshot extensively to find the way for the synthesis of this fragment (**Figure V-8**).

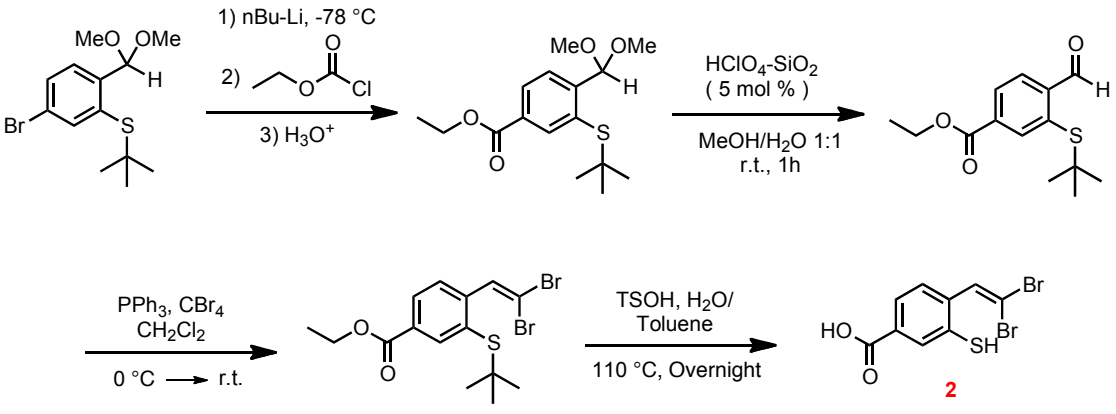


Figure V-8) The rest of the synthetic plan for the second fragment.

It should be mentioned that CRABP-RAR interaction was studied via a transporter to trap conversion as well. It was suggested that the second α-helix and βC-D sheet of CRABP proteins goes through conformational changes as the Retinoic acid binds to this protein (**Figure V-9**). Two residues, which go through the most conformational changes Ala35 and Thr57, were highlighted and mutated to Cysteine residues to convert the CRABPI protein from a transporter of Retinoic acid to a trap for it via the formation of a disulfide bond. The formation of the disulfide bond was confirmed by proteolysis fragmentation of CRABPI. Further studies showed that CRABPI penta mutant E75Q:K81P:E102K:A35C:T57C can stabilize the interaction of the CRABP-RAR complex (24, 25).

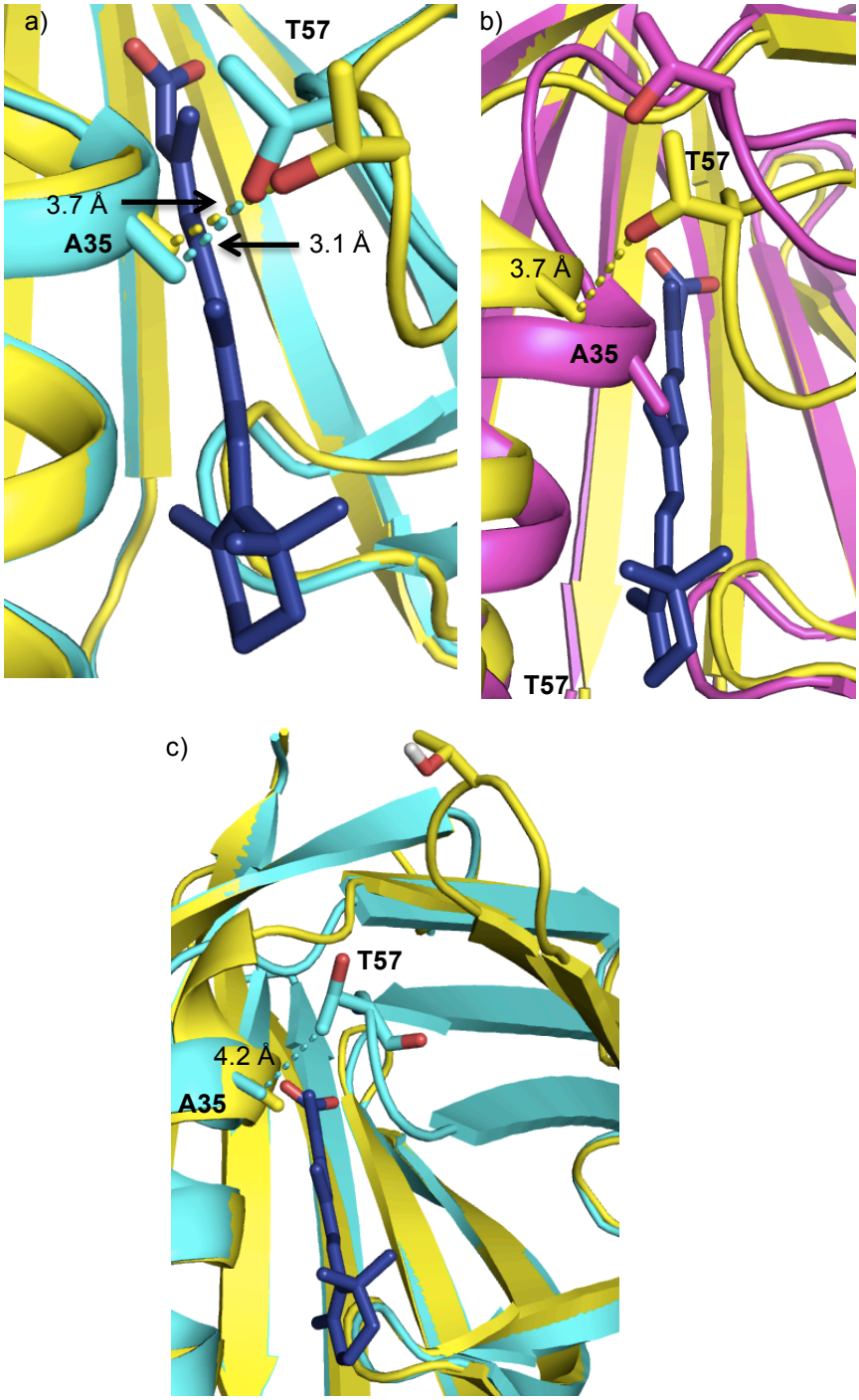


Figure V-9) a) Overlay of the structure of the Apo-CRABPII-Chain A (PDB ID: 2FS7) with the Holo structure (PDB ID: 2FR3) does not show the difference in position of Ala 35 and Thr 57. b) Overlay of the structure of the Apo-CRABPII-Chain B and Holo-CRABPII shows the two residues of Ala35 and Thr57 get close to each other upon Retinoic acid binding. c) Overlay structure of the Apo (PDB ID: 1CBI) and Holo CRABPI (PDB ID: 1CBR) shows that the A35 and Thr57 get close to each other upon Retinoic acid binding.

Dr.Xia tried a pull down assay in Prof.Geiger's lab as well. The idea was to use the R111K:R132L:L121E mutant of hCRABPII which effectively binds to retinal and take the protein-Retinal complex through a reductive amination reaction to make the retinal non-hydrolysable and non-transferable to RAR. The pull down assay was performed and the result did not show the two proteins together (26).

Also the synthesis of the CD-270 or its close analogs, and trying to crystallize the complex of CRABP and RAR is one way of approaching this question, using the library of molecules that now we have available for hCRABPII, trying them with first and second generation mutants of hCRABPII, measuring their affinity for RAR and CRABPII and trying a pull down assay with these ligands is another alternative approach.

V.3 Conclusion

The interaction between the CRABPII and RAR proteins is essential in the delivery of Retinoic acid to RARs and regulation of their transcriptional activity. We aimed at synthesizing molecule CD-270 that is the only molecule already shown to stabilize the interaction between the two proteins. We aimed for the advantage of using C-H activation and cross-coupling reactions in our synthesis. The molecule was broken into two precursors. The first precursor was synthesized via an iridium C-H activation reaction. The synthesis of the second fragment was hampered by difficulties in carboxylation of the aromatic ring.

REFERENCES

REFERENCES

1. Leid M, Kastner P, & Chambon P (1992) Multiplicity generates diversity in the retinoic acid signalling pathways. *Trends in biochemical sciences* 17(10):427-433.

2. Gronemeyer H & Laudet V (1995) Transcription factors 3: nuclear receptors. *Protein profile* 2(11):1173-1308.

3. Bushue N & Wan YJ (2010) Retinoid pathway and cancer therapeutics. *Advanced drug delivery reviews* 62(13):1285-1298.

4. Schweitzer A, Knauer SK, & Stauber RH (2010) Nuclear receptors in head and neck cancer: current knowledge and perspectives. *International journal of cancer. Journal international du cancer* 126(4):801-809.

5. Chen Y, Dokmanovic M, Stein WD, Ardecky RJ, & Roninson IB (2006) Agonist and antagonist of retinoic acid receptors cause similar changes in gene expression and induce senescence-like growth arrest in MCF-7 breast carcinoma cells. *Cancer research* 66(17):8749-8761.

6. Donato LJ, Suh JH, & Noy N (2007) Suppression of mammary carcinoma cell growth by retinoic acid: the cell cycle control gene Btg2 is a direct target for retinoic acid receptor signaling. *Cancer research* 67(2):609-615.

7. Donato LJ & Noy N (2005) Suppression of mammary carcinoma growth by retinoic acid: proapoptotic genes are targets for retinoic acid receptor and cellular retinoic acid-binding protein II signaling. *Cancer research* 65(18):8193-8199.

8. Kim H*, et al.* (2005) The retinoic acid synthesis gene ALDH1a2 is a candidate tumor suppressor in prostate cancer. *Cancer research* 65(18):8118-8124.

9. Poulain S*, et al.* (2009) Vitamin A/retinoids signalling in the human lung. *Lung cancer* 66(1):1-7.

10. Renaud JP*, et al.* (1995) Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature* 378(6558):681-689.

11. Dong D, Ruuska SE, Levinthal DJ, & Noy N (1999) Distinct roles for cellular retinoic acid-binding proteins I and II in regulating signaling by retinoic acid. *The Journal of biological chemistry* 274(34):23695-23698.

12. Budhu A, Gillilan R, & Noy N (2001) Localization of the RAR interaction domain of cellular retinoic acid binding protein-II. *Journal of molecular biology* 305(4):939-949.

13. Budhu AS & Noy N (2002) Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest. *Molecular and cellular biology* 22(8):2632-2641.

14. Sessler RJ & Noy N (2005) A ligand-activated nuclear localization signal in cellular retinoic acid binding protein-II. *Molecular cell* 18(3):343-353.

15. Majumdar A, Petrescu AD, Xiong Y, & Noy N (2011) Nuclear translocation of cellular retinoic acid-binding protein II is regulated by retinoic acid-controlled SUMOylation. *The Journal of biological chemistry* 286(49):42749-42757.

16. Darmon M*, et al.* (1988) Biological activity of retinoids correlates with affinity for nuclear receptors but not for cytosolic binding protein. *Skin pharmacology : the official journal of the Skin Pharmacology Society* 1(3):161-175 (and references there in for CD-270 synthesis).

17. Bryan CS, Braunger JA, & Lautens M (2009) Efficient synthesis of benzothiophenes by an unusual palladium-catalyzed vinylic C-S coupling. *Angewandte Chemie* 48(38):7064-7068.

18. Maleczka RE, Jr., Shi F, Holmes D, & Smith MR, 3rd (2003) C-H activation/borylation/oxidation: a one-pot unified route to meta-substituted phenols bearing ortho-/para-directing groups. *Journal of the American Chemical Society* 125(26):7792-7793.

19. Paul S*, et al.* (2006) Ir-catalyzed functionalization of 2-substituted indoles at the 7-position: nitrogen-directed aromatic borylation. *Journal of the American Chemical Society* 128(49):15552-15553.

20. Vanchura BA, 2nd*, et al.* (2010) Electronic effects in iridium C-H borylations: insights from unencumbered substrates and variation of boryl ligand substituents. *Chemical communications* 46(41):7724-7726.

21. Ghaffari B*, et al.* (2014) Silyl phosphorus and nitrogen donor chelates for homogeneous ortho borylation catalysis. *Journal of the American Chemical Society* 136(41):14345-14348.

22. Jayasundara CR, Unold JM, Oppenheimer J, Smith MR, 3rd, & Maleczka RE, Jr. (2014) A catalytic borylation/dehalogenation route to o-fluoro arylboronates. *Organic letters* 16(23):6072-6075.

23. Kumar R, Kumar D, & Chakraborti AK (2007) Perchloric acid adsorbed on silica gel (HClO4-SiO2) as an inexpensive, extremely efficient, and reusable dual catalyst system for acetal/ketal formation and their deprotection to aldehydes/ketones. *Synthesis-Stuttgart* (2):299-303.

24. Sjoelund V & Kaltashov IA (2007) Transporter-to-trap conversion: a disulfide bond formation in cellular retinoic acid binding protein I mutant triggered by retinoic acid binding irreversibly locks the ligand inside the protein. *Biochemistry* 46(46):13382-13390.

25. Sjoelund V (2009) Retinoic acid transfer and protein-protein interaction between the cellular retinoic acid binding protein and the retinoic acid receptor probed by hydrogen deuterium exchange mass spectrometry. PhD (University of Massachusetts Amherst).

26. Jia X (2008) X-ray crystallographic studies of cellular retinoic acid-binding protein II mutants designed as rhodopsin mimics. PhD (Michigan State University).

Chapter VI: Materials and Methods

UV-Vis spectra were recorded using a Cary 300 Bio WinUV, Varian spectrophotometer. All-*trans*-retinal was purchased from TRC and was used as received. All the other chemicals were purchased from Sigma-Aldrich unless otherwise specified. Fast Q anion exchange resin was purchased from Agilent Technologies. BioLogic DuoFlow (BioRad) was used for protein purifications. Source 15Q resin was purchased from Amersham Bioscience. Sonication of bacterial cells was performed using Ultrasonic Homogenizer from Biologics, Inc.

VI.1 Site-directed mutagenesis

The CRABPII-pET17b plasmid described in (1) was used for mutagenesis following the QuickChange Site-directed Mutagenesis Kit protocol (Agilent Technologies). PCR conditions for amplification of mutants are specified below:

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

Table VI-1. PCR protocol for hCRABPII and hCRBPII mutagenesis.

Table VI-1 (cont'd)

PCR Program		
1x	95 °C	30 min
	95 °C	30 sec
20x	Temprature 3-5 °C lower than primer melting temprature	1 min
	72 °C	4min 30 sec
1x	72 °C	10 min
1x	25 °C	10 min

The PCR product (5 μL) was transformed into DH5α cells (50 μL) and grown on Luria-Bertani (LB)-agar plates supplemented with Ampicillin (75 μg/mL) for 15h. A single colony was inoculated in 15 mL LB medium containing 100 μg/mL ampicillin and grown at 37 ºC while shaking, for 12 hours. DNA purification was performed using a QIAGEN Miniprep DNA purification kit. The average isolated plasmid DNA concentration was 120 ng/μl in 60 μL solution. The DNA sequence was verified by the MSU gene sequencing facility using T7 primer.

Primers

F3Q

Forward: 5'- CCAAAC**CAA**TCTGGCAACTGGAAA-3'
Reverse: 5'-TTTCCAGTTGCCAGA**TTG**GTTTGG-3'

A32Y

Forward: 5'-GAGGAAGATT**TAT**GTGGCTGC-3'
Reverse: 5'-GCAGCCAC**ATA**AATCTTCCTC-3'

A32W

Forward: 5'-GTGATGCTGAGGAAGATT**TGG**GTGGCTGC-3'
Reverse: 5'-GCAGCCAC**CCA**AATCTTCCTCAGCATCAC-3'

P39Y
For: 5' GCAGCGTCCAAG**TAT**GCAGTGG 3'
Rev: 5' CCACTGC**ATA**CTTGGACGCTGC 3'

P39Q
For: 5' GCAGCGTCCAAG**CAA**GCAGTGG 3'
Rev: 5' CCACTGC**TTG**CTTGGACGCTGC 3'

T54V
Forward: 5'-CTACATCAAA**GTC**TCCACCACCGTGCG -3'
Reverse: 5'- CGCACGGTGGTGGA**GAC**TTTGATGTAG -3'

R59Y
Forward: 5'- CCTCCACCACCGTG**TAC**ACCACAGAG -3'
Reverse: 5'- CTCTGTGGT**GTA**CACGGTGGTGGAGG -3'

R59W
Forward: 5'- CCTCCACCACCGTG**TGG**ACCACAGAG -3'
Reverse: 5'- CTCTGTGGT**CCA**CACGGTGGTGGAGG -3'

R111K
Forward: 5'-CCCAAGACCTCGTGGACC**AAA**GAACTGACCAACGATGGG-3'
Reverse: 5'-CCCATCGTTGGTCAGTTC**TTT**GGTCCACGAGGTCTTGGG-3

R132Q: Y134F
Forward: 5'- GTTGTGTGCACC**CAG**GTC**TTC**GTCCG-3'
Reverse: 5'-CGGAC**GAA**GAC**CTG**GGTGCACACAAC-3'

R132Y:Y134F
Forward: 5' GTTGTGTGCACC**TAT**GTC**TTC**GTCCGAGAG 3'
Reverse: 5' CTCTCGGAC**GAA**GAC**ATA**GGTGCACACAAC 3'

VI.2 Protein Expression and Purification of hCRABPII and hCRBPII/pET-17b Mutants

The target gene was transformed into competent E. coli BL21(DE3)pLysS competent cells (100 ng of DNA, for 100 µL of cell solution) following standard protocols and the cells were grown on Luria-Bertani (LB)-agar plates supplemented with antibiotics (Chloramphenicol: 34 µg/mL; Ampicillin: 100 µg/mL) at 37 ºC for 12 hours. A single colony was used to inoculate 50 mL of

LB medium containing 34 μg/mL chloramphenicol and 100 μg/mL ampicillin and was grown at 37 ºC, while shaking overnight. The resulting culture was used to inoculated 1 L of LB containing 27 μg/mL chloramphenicol and 100 μg/mL ampicillin and was grown at 37 ºC while shaking till OD_{600} reached 0.7-0.9. The expression was induced with addition of isopropyl-β-D-thiogalactopyranoside (IPTG, Gold Biotechnology, 1 mM). The culture was shaken at 19 ºC for 36h for hCRABPII and at room temperature for overnight for hCRBPII. The cells were harvested by centrifugation (5000 rpm, 20 min, 4 ºC) and resuspended in Tris-binding buffer (10 mM Tris, pH=8.0, 50 mL). The cells were lysed by sonication (VWR Scientific , power 70%, 3 x 1 min). The solution was spun down (12000 rpm, 30 min, 4 ºC) and the supernatant was loaded on a FastQ anion exchange column preequilibrated with Tris-binding buffer. The column was washed twice with Tris-binding buffer (2 x 100 mL) and the protein was eluted with Tris-elution buffer (10 mM Tris, 150 mM sodium chloride, pH = 8.0, 70 mL). The eluent underwent buffer exchange with Tris-binding buffer using an unltrafiltration cell under nitrogen pressure (~20 psi) (Millipore ™, Regenerated Cellulose membrane YM10, NMWL: 10,000). The desired protein was further purified using a BioLogic DuoFlow system (BioRad) equipped with a 15Q anion exchange column, according to the following protocol:

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

Table VI-2. Anion Exchange purification protocol for hCRABPII and hCRBPII.

The buffer used for above protocol is 50 mM Tris, pH is adjusted automatically. The proteins elute with 4% 2M NaCl.

VI.3 Extinction Coefficient Determination

The absorption extinction coefficients (ε) for the various CRABPII mutants were determined according to the method described by Gill and von Hippel (2).

VI.4 UV-vis Measurements

The CRABPII-PSB formation (λ_{max} > 450 nm) was followed by UV-vis in phosphate buffer (4 mM NaH_2PO_4, 16 mM Na_2HPO_4, 150 mM NaCl, pH=7.3) or 50 mM citrate buffer at pH=5.0. The experiment was performed with a final protein concentration of 50 μM, and 0.5 equivalent of retinal was added (from a stock solution of 1 mM in ethanol). Peaks with λ_{max} > 450 nm are considered PSB peaks, while deprotonated imine peaks (SB) appear at ~ 360 nm. Non-covalently bound retinal absorbs at ~380 nm.

VI.5 pKa determinations:

hCRABPII mutants: The pKa values were determined based on the plot of the absorbance change vs pH by using the curve fit as described previously for bacteriorhodopsin pKa determination.

The total absorbance change of PSB during each acidification or basification in comparison to the point with minimum PSB absorption (ΔA) is plotted against the pH (3, 4).

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

hCRBPII mutants: The points at the corresponding λ_{max} versus pH were plotted in KaleidaGraph. The curve was fit to a third degree polynomial and the pKa was determined from the second derivative of the polynomial expression set to 0 (point of inflection):

$$y = ax^3 + bx^2 + cx + d$$
$$y' = 3ax^2 + 2bx + c$$
$$y'' = 6ax + 2b$$

Base titrations in citrate buffer: A 50 μM solution of hCRABPII proteins was incubated with 0.5 equivalent of retinal in 50 mM citrate buffer at pH=5.0. The UV-Visible spectrum was followed till maximum PSB formation. The leftover SB was converted to PSB by acidifying the solution with 2M HCl before titration. Then a 1M NaOH solution was used to basify the protein solution till complete PSB to SB conversion.

Acid titrations: A 50 μM solution of hCRABPII proteins was incubated with 0.5 equivalent of retinal in PBS buffer. The UV-Visible spectrum was followed till maximum PSB loss (12-24h). Then a saturated citric acid buffer solution was used to acidify the protein for the first few points by further acidification with a 6M HCl solution till complete SB to PSB conversion. For F3Q mutants of hCRABPII the acid titration was performed at PSB maximum formation.

Base titrations in PBS buffer: A 50 μM solution of hCRABPII proteins was incubated with 0.5 equivalent of retinal in PBS buffer. The UV-Visible spectrum was followed till maximum PSB formation. Then a 1M NaOH solution was used to basify the protein solution till complete PSB to SB conversion.

UV irradiation - titrations in PBS buffer: A 50 μM solution of hCRABPII proteins was incubated with 0.5 equivalent of retinal in PBS buffer. The UV-Visible spectrum was followed till maximum PSB formation. Then a UV machine with a 340-380 light filter pass was used to convert the leftover SB to PSB (1-2 minute irradiation). A 1M NaOH solution was used to basify the protein solution till complete PSB to SB conversion. In case of **M2** the UV irradiation was after complete PSB loss.

hCRBPII titrations in PBS buffer: A 25 μM solution of hCRBPII proteins was incubated with 0.5 equivalent of retinal in PBS buffer. The UV-Visible spectrum was followed to track the PSB formation and loss. For base titration a 1M NaOH solution was used to basify the protein solution till complete PSB to SB conversion. For acid titration a saturated citric acid buffer solution was used to acidify the protein for the first few points by further acidification with a 6M HCl solution till complete SB to PSB conversion.

All of the graphs were recorded with a UV-Vis spectrophotometer. Graphs were recorded as Excel and Kaleidagraph files, and absorption at λ_{max} vs different pHs were plotted. pKa values were determine as described in section VI-5.

Entry	Mutant	Low pKa (λ_{max})	High pKa (λ_{max})	ε ($cm^{-1}M^{-1}$)	PSB rise (hours)	PSB fall (hours)
1	**R111K:R132Q**	Protein purified not assayed				
2	**R111K:P39Q:R132Q:R59Y**	516[a]		21948	20[a]	n.d.
3	**R111K:P39Y:R132Q:R59Y**	Not much kinetic PSB at pH=7.0	5.5(500nm)	22395	-	-
4	**R111K:P39Q:R132Q:Y134F:R59Y**	519[a]		22529	12[a]	n.d.
5	**R111K:P39Q:R132Q:Y134F:R59W:A32W**	561[a]		31389	n.d.	n.d.
6	**R111K:P39Y:R132Q:Y134F:R59Y**	n.d.	More than one PSB formed	22620	1	8
7	**R111K:P39Q:R132L:Y134F:T54V:R59Y**	550[a]		21066	0.2[a]	4[a]
8	**R111K:R132Q:Y134F:T54V:R59W**	Not much kinetic PSB at pH=7.0, 560		25118	-	-

Table VI-3. hCRABPII mutants.
a: Measurements were done in 50 mM in 50 mM citrate buffer, pH = 5.0.

Table VI-3 (cont'd)

Entry	Mutant	Low pKa (λ_{max})	High pKa (λ_{max})	ε ($cm^{-1}M^{-1}$)	PSB rise (hours)	PSB fall (hours)
9	R111K:P39Q:R132Q:Y134F:T54V:R59Y	5.4 (530)	7.8 (554)	20749	3.5	24
10	R111K:P39Q:R132Q:Y134F:T54V:R59W	5.1 (537)	7.6 (555)	25582	3	24
11	R111K:P39Q:R132Q:Y134F:T54V:R59Q	523[a]		19965	n.d.	n.d.
12	R111K:P39Q:R132Q:Y134F:T54V:R59Y:A32W	6.1 (563)	8.9 (584)	26681	4	24
13	R111K:P39Q:R132Q:Y134F:T54V:R59Y:A32Y	5.6 (554)	8.3 (575)	22441	2	24
14	R111K:P39Q:R132Q:Y134F:T54V:R59W:A32W	6.2 (582)	8.4 (598)	29659	2	24
15	R111K:P39N:R132N:Y134F:T54V:R59Y	Protein precipitates fast		22585	-	-
16	R111K:P39Q:R132E:Y134F:T54V:R59Y	Protein Precipitates fast		-	-	-

a: Measurements were done in 50 mM in 50 mM citrate buffer, pH = 5.0.

Table VI-3 (cont'd)

Entry	Mutant	Low pKa (λ_{max})	High pKa (λ_{max})	ε ($cm^{-1}M^{-1}$)	PSB rise (hours)	PSB fall (hours)
17	R111K:P39N:R132Q:Y134F:T54V:R59Y	Broad weak PSB peak		22580	-	-
18	R111K:R132Q:Y134F:T54V:R59Y	Not much PSB in phosphate, 557		20627	-	-
19	R111K:P39Y:R132Q:Y134F:T54V:R59Y	2.8 (564)	8.2 (573)	22487	0.2	10
20	R111K:P39Y:R132Q:Y134F:T54V:R59Y:A32Y	2.5 (587)	8.5 (588)	21881	0.1	16
21	R111K:P39Y:R132Q:Y134F:T54V:R59Y:A32W	2.5 (600)	8.3 (596)	24768	0.2	12
22	R111K:P39Y:R132Q:Y134F:T54V:R59W:A32W	3.45 (601)	8.35 (595)	30878	0.3	10
23	R111K:P39Q:R132Y:Y134F:T54V:R59W:A32W	608[a]		29567	n.d.	n.d.
24	R111K:P39Y:R132Y:Y134F:T54V:R59Y:A32Y	Protein precipitates fast		26088	-	-

a: Measurements were done in 50 mM in 50 mM citrate buffer, pH = 5.0.

Table VI-3 (cont'd)

Entry	Mutant	Low pKa (λ_{max})	High pKa (λ_{max})	ε ($cm^{-1}M^{-1}$)	PSB rise (hours)	PSB fall (hours)
25	R111K:P39Q:R132N:Y134F:T54V:R59Y	Protein precipitates fast				
26	R111K:P39Q:R132Q:Y134F:T54V:R59Y:F3Q	Protein not expressed				
27	R111K:P39Q:R132Q:Y134F:T54V:R59Y:A32Y:F3Q	9.6 (545)		21169	3	-
28	R111K:P39Q:R132Q:Y134F:T54V:R59Y:A32W:F3Q	9.65 (557)		29004	5	30h 50% PSB loss
29	R111K:P39Q:R132Q:Y134F:T54V:R59W:A32W:F3Q	9.8 (571)		29697	3	-
30	R111K:P39Q:R132Q:Y134F:T54V:R59W:F3Q	Protein not expressed				
31	R111K:P39Q:R132Y:Y134F:T54V:R59Y:F3Q	Protein not expressed				
32	R111K:P39Y:R132Q:Y134F:T54V:R59Y:A32W:F3Q	9.8 (571)		29465	2h	30h 50 % PSB loss

a: Measurements were done in 50 mM in 50 mM citrate buffer, pH = 5.0.

Table VI-3 (cont'd)

Entry	Mutant	Low pKa (λ_{max})	High pKa (λ_{max})	ε ($cm^{-1}M^{-1}$)	PSB rise (hours)	PSB fall (hours)
33	R111K:P39Y:R132Q:Y134F:T54V:R59W:A32W:F3Q	Protein not expressed				
34	R111K:P39Y:R132Q:Y134F:T54V:R59Y:F3Q	Protein not expressed				

a: Measurements were done at 50 mM citrate buffer, pH = 5.

a: Measurements were done in 50 mM in 50 mM citrate buffer, pH = 5.0.

Protein	ε280nm ($M^{-1}.cm^{-1}$)
KL:R58Y:**Q38A**	31431
KL:R58Y:T51V:**Q38A**	28945
KL:R58Y:T51V:**Q38A**:**Q128A**	28680
KL:R58Y:T51V:**Q38L**:**Q128L**	28431
KL:R58Y:T51V:Y19W:**Q38A**	33765
KL:R58Y:T51V:Y19W:**Q38L**	32920
KL:R58Y:T51V:Y19W:**Q38M**	34002
KL:R58Y:T51V:Y19W:**Q38F**	33982
KL:R58Y:T51V:Y19W:**Q38A**:**Q128A**	34556
KL:R58Y:T51V:Y19W:**Q38C**:**Q128L**	33909
KL:R58Y:T51V:Y19W:**Q38F**:**Q128L**	34006
KL:R58W:T51V:T53C:T29L:**Q38N**	33718
KL:R58W:T51V:T53C:T29L:Y19W:**Q128L**	35604
KL:T51V:T53C:T29L:Y19W:**Q128L**	31262
KL:R58W:T51V:T53S:T29L:Y19W:**Q38L**	37670
KL:R58W:T51V:T53S:T29L:Y19W:**Q38N**	37326
KL:R58W:T51V:T53S:T29L:Y19W:**Q38L**:**Q128L**	38579
KL:R58W:T51V:T53C:T29L:Y19W:Q4R:**Q128L**	35936
KL:R58W:T51V:T53C:T29L:Y19W:Q4R:A33W:**Q38L**	41977
KL:R58W:T51V:T53C:T29L:Y19W:Q4R:A33W:**Q128L**	45062
KL:R58W:T51V:T53S:T29L:Y19W:Q4F:A33W:**Q128L**	42875
KL:R58W:T51V:T53S:T29L:Y19W:Q4F:A33W:**Q38N**	-
KL:R58W:T51V:T53S:T29L:Y19W:Q4F:A33W:**Q128N**	-
KL:R58W:T51V:T53S:T29L:Y19W:Q4F:A33W:**Q38L**:**Q128L**	-

Table VI-4. hCRBPII Proteins Extinction Coefficient.
KL is the abbreviation for Q108K:K40L

Mutants	λmax(nm)	pKa
Q108K:K40L:R58Y	533	9.0
Q108K:K40L:R58Y:**Q38A**	513	8.2
Q108K:K40L:R58Y:T51V	556	10.0
Q108K:K40L:R58Y:T51V:**Q38A**	554	9.0
Q108K:K40L:R58Y:T51V:**Q38A:Q128A**	510	8.2
Q108K:K40L:R58Y:T51V:**Q38L:Q128L**	503	7.9

- Q38 and Q128 mutations on the parent mutants absorbing below 550 nm.

Mutants	λmax(nm)	pKa
Q108K:K40L:R58Y:T51V:Y19W	565	10.2
Q108K:K40L:R58Y:T51V:Y19W:**Q38A**	563	9.2
Q108K:K40L:R58Y:T51V:Y19W:**Q38L**	522	8.7
Q108K:K40L:R58Y:T51V:Y19W:**Q38M**	546	8.6
Q108K:K40L:R58Y:T51V:Y19W:**Q38F**	537	8.2
Q108K:K40L:R58Y:T51V:Y19W:**Q38A:Q128A**	510	8.2
Q108K:K40L:R58Y:T51V:Y19W:**Q38C:Q128L**	548	8.5
Q108K:K40L:R58Y:T51V:Y19W:**Q38F:Q128L**	537	7.7

Table VI-5. hCRBPII mutants.

- Q38 and Q128 mutations on Q108K:K40L:R58Y:T51V:Y19W mutant (highest pKa mutant of hCRBPII).

Table VI-5 (cont'd)

Mutants	λmax (nm)	pKa
Q108K:K40L:R58W:T51V:T53C:T29L	586	7.9
Q108K:K40L:R58W:T51V:T53C:T29L:**Q38N**	544	-
Q108K:K40L:R58W:T51V:T53C:T29L:**Q38M**[a]	513	7.5
Q108K:K40L:R58W:T51V:T53C:T29L:**Q128L**[a]	532	6.0
Q108K:K40L:R58W:T51V:T53C:T29L:**Q38M:Q128L**[a]	504	7.6
Q108K:K40L:R58W:T51V:T53C:T29L:Y19W	591	8.2
Q108K:K40L:R58W:T51V:T53C:T29L:Y19W:**Q128L**	531	4.6
Q108K:K40L:T51V:T53C:T29L:Y19W	533	8.4
Q108K:K40L:T51V:T53C:T29L:Y19W: **Q128L**	510	-
Q108K:K40L:R58W:T51V:T53S:T29L:Y19W	600	9.0
Q108K:K40L:R58W:T51V:T53S:T29L:Y19W:**Q38L**	520	8.1
Q108K:K40L:R58W:T51V:T53S:T29L:Y19W:**Q38N**	571	8.3
Q108K:K40L:R58W:T51V:T53S:T29L:Y19W:**Q38L:Q128L**	502	3.8

- Q38 and Q128 mutations on the parent mutants absorbing between 550-600 nm.

Mutants	λmax(nm)	pKa
Q108K:K40L:R58W:T51V:T53C:T29L:Y19W:Q4R:A33W	644	6.9
Q108K:K40L:R58W:T51V:T53C:T29L:Y19W:Q4R:A33W:**Q38L**	510	-
Q108K:K40L:R58W:T51V:T53S:T29L:Y19W:Q4F:A33W	646	8.7
Q108K:K40L:R58W:T51V:T53S:T29L:Y19W:Q4F:A33W:**Q128L**	608	4.7
Q108K:K40L:R58W:T51V:T53S:T29L:Y19W:Q4F:A33W:**Q38N**	610	-
Q108K:K40L:R58W:T51V:T53S:T29L:Y19W:Q4F:A33W:**Q128N**	565	-
Q108K:K40L:R58W:T51V:T53S:T29L:Y19W:Q4F:A33W:**Q38L:Q128L**	538	-

- Q38 and Q128 mutations on the parent mutants absorbing above 600 nm.

VI.8 Protein Crystallization of hCRABPII

VI.8-1 R 111K:P39Y:R132Q:Y134F:T54V:R59Y

For kinetic PSB product: To the concentrated protein solution at 20 mg/ml was added four equivalents of retinal and incubated for 20 minutes. A 24 well crystallization plate was used for vapor diffusion crystallization with the reservoir solution of 1 ml, containing 12 % PEG3350 and 0.1 M malonate at pH = 6.0. 1 µl of protein solution was mixed with 1 µl reservoir solution. Crystals appeared and grew after 24 hours at 4 °C. Crystals were flash frozen in a solution containing the mother liquor and 20 % glycerol.

For Thermodynamic SB product: To the protein solution at 1.1 mg/ml was added four equivalents of retinal. The spectrum of the protein-ligand mixture was tracked until PSB loss completion. Then the mixture was concentrated to 20 mg/ml using concentrator cells. A twenty-four well crystallization plate was used for vapor diffusion crystallization with the reservoir solution containing 12 % PEG 3350 and 0.1 M malonate at pH = 6.0. 1 µl of protein solution was mixed with 1 µl reservoir solution. Crystals appeared and grew after 24 hours. These crystals had a light blue color indicating the presence of PSB in the crystals. In order to deprotonate the PSB in the crystals, they were soaked in 12% PEG 3350 and 0.1 M malonate at pH = 7.5 for 30 minutes until crystals became colorless. The soaked crystals were flash frozen in the same solution containing 20 % glycerol.

For UV irradiated PSB product: Soaked crystals described above were UV irradiated with a hand held TLC UV lamp for 30 minutes in dark. The UV irradiated crystals showed a dark blue color at the end of the irradiation

period. These crystals were transferred to a solution of 12% PEG 3350, 0.1 M malonate at pH = 7.5 and 20 % glycerol and flash frozen immediately under dim microscope light, while still being irradiated with the UV light.

VI.8-2 R 111K:P39Q:R132Q:Y134F:T54V:R59Y

For Thermodynamic SB product: To the protein solution at 1 mg/ml was added four equivalents of retinal. The spectrum of the protein-ligand mixture was tracked until PSB loss completion. Then the mixture was concentrated to 20 mg/ml using concentrator cells. A twenty-four well crystallization plate was used for vapor diffusion crystallization with the reservoir solution containing 12 % PEG 3350 and 8 % Tacsimate at pH = 6.0. 1 µl of protein solution was mixed with 1 µl reservoir solution. Crystals appeared and grew after 24 hours. These crystals had a light red color indicating the presence of PSB in the crystals. In order to deprotonate the PSB in the crystals, they were soaked in 12% PEG 3350 and 8 % Tacsimate at pH = 7.5 for 45 minutes untill crystals became colorless. The soaked crystals were flash frozen in the same solution containing 20 % glycerol

For UV irradiated PSB product: Soaked crystals described above were UV irradiated with a hand held TLC UV lamp for 1 hour in dark. The UV irradiated crystals showed a red color at the end of the irradiation period, and were flash frozen immediately under dim microscope light in 12% PEG 3350, 8 % Tacsimate at pH = 7.5 and 20 % glycerol.

VI.8-3 R111K:P39Q:R132Q:Y134F:T54V:R59Y:A32Y:F3Q

To the concentrated protein solution at 20 mg/ml was added four equivalents of retinal and incubated for three hours. A 24 well crystallization plate was

used for vapor diffusion crystallization with the reservoir solution of 1 ml, containing 20 % PEG3350 and 0.1 M malonate at pH = 6.0. 1 µl of protein solution was mixed with 1 µl reservoir solution. Crystals appeared after 24 hours and grew in 5 days at 4 °C. Crystals were flash frozen in a solution containing the mother liquor and 20 % glycerol.

The same method was used for crystallization and cryo protection of kinetic PSB of mutant M7 in the tables after 20 minutes of retinal incubation.

VI.8-4 R 111K:P39Q:R132Q:Y134F:T54V:R59Y:A32W:F3Q

To the concentrated protein solution at 20 mg/ml was added four equivalents of retinal and incubated for three hours. A 24 well crystallization plate was used for vapor diffusion crystallization with the reservoir solution of 1 ml, containing 20 % PEG3350 and 0.1 M Bis Tris Propane at pH = 6.0. 1 µl of protein solution was mixed with 1 µl reservoir solution. Crystals appeared after 24 hours and grew in 5 days at 4 °C. Crystals were flash frozen in a solution containing the mother liquor and 20 % glycerol.

The same method was used for the crystallization and cryo protection of thermodynamic-SB of mutant M5 in the tables after 24 hours of retinal incubation.

VI.9 Protein crystallization of hCRBPII

To the concentrated protein solution at 8-12 mg/ml is added four equivalents of retinal in dark. The solution is allowed to sit at room temperature until maximum PSB formation. A 24 well crystallization plate was used for vapor diffusion crystallization with the reservoir solution of 1 ml, containing 25 % PEG 4000 and 0.1 M Ammonium Acetate and 0.1 M Sodium Acetate at pH

range 4-4.8. 1 µl of protein solution was mixed with 1 µl reservoir solution. Crystals appeared after 24 hours and grew in 5 days at 4 °C. Crystals were flash frozen in a solution containing the mother liquor and 20 % glycerol.

Most of hCRBPII Proteins also show PSB loss overtime. In some cases for crystallizing more of a homogenous solution that can improve diffraction and provide better electron density maps a solution of hCRBPII protein at 1 mg/ml protein was incubated with 4 equivalent retinal and the UV spectrum was followed until maximum PSB loss. Then this solution was concentrated to 8-12 mg/ml for crystallization trials. Waiting method was tried for all of the hCRBPII proteins reported here except Q108K:K40L:T51V:R58Y:Y19W, Q108K:K40L:T51V:R58Y:Y19W:Q38L and structures crystallized with Julolidene.

VI.10 Data Collection and Refinement

Diffraction data were collected at the Advanced Photon Source (APS) (Argonne IL) at the LSCAT (sector 21) at 1.00 Å wavelength radiation at 100 K. Data reduction and scaling were performed using the HKL2000 program package (5). The structures were solved using Molecular Replacement in CCP4 program suite (6), and refined in PHENIX program packages (7). The search model was R132K:R111L:L121E mutant of Cellular Retinoic Acid Binding ProteinII (protein data bank code 2G7B). Three cycles of refinement was considered for each run and placement of ordered water molecules was done in COOT (0.6.1) (8). The chromophore was manually fitted in the electron density at the end of the refinement. For hCRBPII protein the 2RCT was used as the model for refinement.

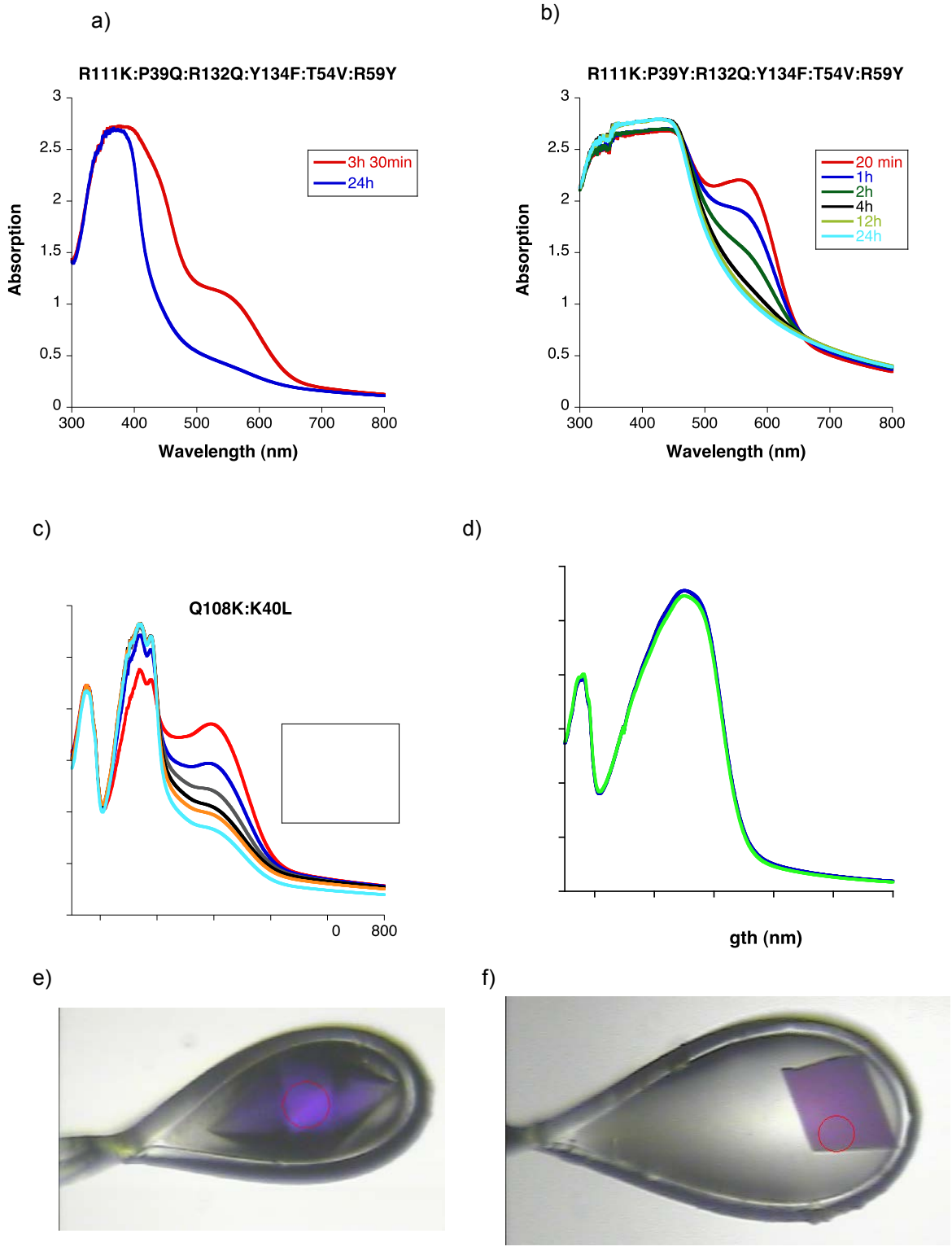


Figure VI-1) Retinal incubation for crystallization of: a) R111K:P39Q:R132Q:Y134F:T54V:R59Y mutant of hCRABPII b) R111K:P39Y:R132Q:Y134F:T54V:R59Y mutant of hCRABPII. c) Q108K:K40L mutant of hCRBPII d) Q108K:T51D mutant of hCRBPII. e) A photo of R111K:P39Y:R132Q:Y134F:T54V:R59Y crystal. f) A photo of R111K:P39Q:R132Q:Y134F:T54V:R59Y:A32W:F3Q.

Substrate	M1[b]-SB-cycle1	M1-UV-cycle1	M1-SB-cycle2	M1-UV-cycle2	M1-SB-cycle3	M1-UV-cycle3
Space group	$P3_121$	$P3_121$	$P3_121$	$P3_121$	$P3_121$	$P3_121$
a (Å)	58.552	58.326	58.599	58.488	58.509	58.655
b (Å)	58.552	58.326	58.599	58.488	58.509	58.655
c (Å)	99.392	99.936	99.854	100.402	99.688	100.352
α (°)	90	90	90	90	90	90
β (°)	90	90	90	90	90	90
γ (°)	120	120	120	120	120	120
Molecules per Asymmetric Unit	1	1	1	1	1	1
Total reflection	189007	163428	161415	161005	155486	145510
Unique Reflection	17990	15647	14985	14868	14966	13508
Completeness (%)	99.6 (99.6)[a]	99.4 (85.3)	99.5 (99.3)	100 (100)	99.4 (99)	100 (99.8)
Average I/σ	36.1 (3.78)	34.86 (3.65)	33.94 (4.69)	39.06 (4.38)	26.69 (2.00)	38.81 (4.38)
R_{merge} (%)	7.1 (52.3)	8.3 (40.5)	8.9 (44.6)	8 (54.4)	9.8 (57.3)	6.7 (55.7)
Resolution (Å) (Last Shell)	50-1.83 (1.86-1.83)	50-1.92 (1.95-1.92)	50-1.95 (1.98-1.95)	50-1.96 (1.99-1.96)	50-1.95 (1.98-1.95)	50-2.03 (2.07-2.03)
R_{cryst}/R_{free} (%)	18.73/24.20	19.74/24.45	18.99 /23.80	20.52/24.93	19.16/24.05	21.02/25.73
RMSD From Ideal Values						
Bond Length (Å)	0.007	0.008	0.007	0.008	0.008	0.008
Bond Angle (°)	1.192	1.228	1.106	1.104	1.201	1.152
Average B factor	26.90	29.8	29.76	34.71	29.99	36.24
Number of water molecules	199	141	161	123	147	106
PDB IDs	4YBP	4YBU	4YCE	4YCH	4YDA	4YDB

Table VI-6. The hCRABPII crystal table.
a Values in the parenthesis refer to the last resolution shell.
b **M1: R111K:P39Q:R132Q:Y134F:T54V:R59Y**

Table VI-6 (cont'd)

Substrate	M2[b]-Kinetic PSB	M2-SB-cycle1	M2-UV-cycle1	M2-SB-cycle2	M2-UV-cycle2	M2-SB-cycle3	M2-UV-cycle3
Space group	$P3_121$	$P3_121$	$P3_121$	$P3_121$	$P3_121$	$P3_121$	$P3_121$
a (Å)	58.935	58.869	58.865	59.199	58.779	58.975	59.031
b (Å)	58.935	58.869	58.865	59.199	58.779	58.975	59.031
c (Å)	99.954	99.71	100.441	100.038	101.228	100.443	101.123
α (°)	90	90	90	90	90	90	90
β (°)	90	90	90	90	90	90	90
γ (°)	120	120	120	120	120	120	120
Molecules per Asymmetric Unit	1	1	1	1	1	1	1
Total reflection	164293	281303	163969	179720	132920	140437	126296
Unique Reflection	15102	25940	15237	16619	12329	13048	11730
Completeness (%)	99.5 (99.4)[a]	99.4 (98.8)	99.6 (99.6)	99.8 (99.2)	99.8 (99.3)	99.9 (99.8)	100 (99.9)
Average I/σ	34.6 (3.0)	62 (4.44)	49.5 (4.41)	48.04 (4.53)	41.94 (4.62)	40.72 (5.12)	38.92 (4.1)
R_{merge} (%)	9.4 (63.9)	4.7 (47.9)	6 (52.6)	5.9 (53)	6.5 (51.6)	6.8 (53.3)	7.5 (54.2)
Resolution (Å) (Last Shell)	50-1.95 (1.98-1.95)	50-1.62 (1.65-1.62)	50-1.95 (1.98-1.95)	50-1.90 (1.93-1.90)	50-2.10 (2.14-2.10)	50-2.06 (2.10-2.06)	50-2.14 (2.18-2.14)
R_{cryst}/R_{free} (%)	19.97/24.88	19.47/22.74	20.01 /26.53	20.43/23.81	20.38/23.65	20.16/25.45	19.87/23.46
RMSD From Ideal Values							
Bond Length (Å)	0.008	0.006	0.007	0.007	0.007	0.009	0.009
Bond Angle (°)	1.301	1.158	1.144	1.123	1.097	1.228	1.167
Average B factor	35.2	29.74	35.97	32.94	39.96	36.02	40.63
Number of water molecules	108	215	116	135	90	123	92
PDB IDs	4YFP	4YFQ	4YFR	4YGG	4YGH	4YGZ	4YH0

[a] Values in the parenthesis refer to the last resolution shell.
[b] **M2: R111K:P39Y:R132Q:Y134F:T54V:R59Y**

Table VI-6 (cont'd)

Substrate	M6[b]-Thermodynamic SB	M8[c]-Kinetic PSB	M11[d]	M3[e]
Space group	P1	P21	P1	P1
a (Å)	39.716	82.839	36.737	36.574
b (Å)	65.13	45.455	39.21	39.54
c (Å)	103.743	93.963	58.402	58.183
α (°)	80.446	90	86.775	87.078
β (°)	88.391	116.004	79.881	80.016
γ (°)	71.303	90	70.831	70.511
Molecules per Asymmetric Unit	6	4	2	2
Total reflection	200218	206425	135505	169219
Unique Reflection	73056	45694	42116	44746
Completeness (%)	97.5 (97.2)[a]	95 (65.2)	93.3 (70.2)	96.0 (93.4)
Average I/σ	15.69 (4.46)	22.66 (1.92)	31 (2.55)	36.41 (2.97)
R_{merge} (%)	9.1 (25.9)	7 (45.7)	6.9 (28.8)	6.1 (47.0)
Resolution (Å) (Last Shell)	50-2.10 (2.17-2.10)	50-1.96 (1.99-1.96)	50-1.57 (1.60-1.57)	50-1.54 (1.57-1.54)
R_{cryst}/R_{free} (%)	24.19/28.66	20.77/26.23	19.63 /22.53	18.20/21.52
RMSD From Ideal Values				
Bond Length (Å)	0.008	0.009	0.006	0.006
Bond Angle (°)	1.161	1.188	1.118	1.149
Average B factor	29.83	40.51	25.56	26.85
Number of water molecules	702	355	297	340
PDB IDs	-	-	4YKO	4YKM

[a] Values in the parenthesis refer to the last resolution shell.
[b] **M6: R111K:P39Q:R132Q:Y134F:T54V:R59Y:A32W**
[c] **M8: R111K:P39Y:R132Q:Y134F:T54V:R59Y:A32Y**
[d] **M11: R111K:P39Q:R132Q:Y134F:T54V:R59Y:A32Y:F3Q**
[e] **M3: R111K:P39Q:R132Q:Y134F:T54V:R59Y:A32W:F3Q**

Substrate	Q108K:K40L-Retinal	Q108K:K40L:T51V:R58F-Retinal
Space group	P1	P1
a (Å)	29.479	30.234
b (Å)	36.57	35.704
c (Å)	64.093	64.104
α (°)	86.54	91.284
β (°)	86.601	90.912
γ (°)	64.827	114.19
Molecules per Asymmetric Unit	2	2
Total reflection	171006	136653
Unique Reflection	47535	36188
Completeness (%)	95 (92.9)[a]	96.1 (94)
Average I/σ	43.69 (3.72)	31.67 (2.17)
R_{merge} (%)	4.5 (29.2)	5 (46.8)
Resolution (Å) (Last Shell)	50-1.40 (1.42-1.40)	50-1.54 (1.57-1.54)
R_{cryst}/R_{free} (%)	20.74/23.78	19.34/22.88
RMSD from ideal values		
Bond Angle	1.096	1.081
Bond Length (Å)	0.007	0.007
Average B factor	27.09	23.76
Number of water molecules	308	237
PDB IDs	4RUU	-

Table VI-7. The hCRBPII crystal table.
[a] Values in the parenthesis refer to the last resolution shell.

Table VI-7 (cont'd)

Substrate	Q108K:K40L:T51V: R58Y:Y19W-Retinal	Q108K:K40L:T51V: R58Y:Y19W:Q38L-Retinal
Space group	P1	P1
a (Å)	29.953	29.97
b (Å)	35.721	35.722
c (Å)	64.937	64.455
α (°)	90.846	90.866
β (°)	88.451	91.328
γ (°)	65.68	114.36
Molecules per Asymmetric Unit	2	2
Total reflection	184981	160420
Unique Reflection	59171	43022
Completeness (%)	94.6 (89.2)[a]	95.5 (93)
Average I/σ	38.98 (1.9)	37.26 (3.14)
R_{merge} (%)	3.6 (41.4)	5.3 (46.5)
Resolution (Å) (Last Shell)	50-1.31 (1.33-1.31)	50-1.5 (1.53-1.50)
R_{cryst}/R_{free} (%)	20.17/23.58	20.33/25.32
RMSD from ideal values		
Bond Angle	1.138	1.158
Bond Length (Å)	0.007	0.007
Average B factor	25.30	26.22
Number of water molecules	347	219
PDB IDs	-	-

[a] Values in the parenthesis refer to the last resolution shell.

Table VI-7 (cont'd)

Substrate	Q108K:T51D	Q108K:K40L:T51V
Space group	P212121	P1
a (Å)	66.199	29.898
b (Å)	73.792	36.323
c (Å)	353.4	63.987
α (°)	90	87.17
β (°)	90	92.480
γ (°)	90	115.28
Molecules per Asymmetric Unit	12	2
Total reflection	527001	180765
Unique Reflection	108095	48161
Completeness (%)	97.2 (98.8)[a]	95.5 (92.7)
Average I/σ	22 (1.92)	37.85 (1.84)
R_{merge} (%)	7.5 (72.3)	4.5 (58.1)
Resolution (Å) (Last Shell)	50-2.06 (2.10-2.06)	50-1.40 (1.42-1.40)
R_{cryst}/R_{free} (%)	22.23/27.35	21.70/25.54
RMSD from ideal values		
Bond Angle	1.139	1.157
Bond Length (Å)	0.009	0.007
Average B factor	40.00	23.76
Number of water molecules	652	319
PDB IDs	-	-

[a] Values in the parenthesis refer to the last resolution shell.

Table VI-7 (cont'd)

Substrate	Q108K:K40L:T51V:T53C:R58W: T29L:Y19W:Q4H-Julolidene	Q108K:K40L:T51V: R58F-Julolidene
Space group	P1	P1
a (Å)	29.941	29.787
b (Å)	36.14	35.947
c (Å)	64.502	64.089
α (°)	91.78	90.506
β (°)	88.209	92.446
γ (°)	115.116	113.305
Molecules per Asymmetric Unit	2	2
Total reflection	59395	150975
Unique Reflection	21865	45939
Completeness (%)	92.2 (79)[a]	94.9 (79.4)
Average I/σ	38.65 (8.38)	42.2 (4.11)
R_{merge} (%)	4.5 (15)	4.5 (19.4)
Resolution (Å) (Last Shell)	50-1.82 (1.85-1.82)	50-1.42 (1.44-1.42)
R_{cryst}/R_{free} (%)	21.83/28.54	22.19/25.63
RMSD from ideal values		
Bond Angle	1.203	1.144
Bond Length (Å)	0.007	0.007
Average B factor	25.30	23.76
Number of water molecules	137	194
PDB IDs	-	-

[a] Values in the parenthesis refer to the last resolution shell.

VI.11 CD-270 Synthesis

VI.11-1 C-H activation reaction

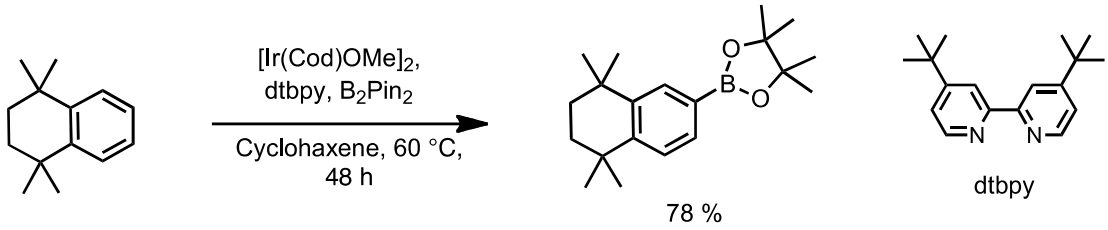


Figure VI-2) Borylation reaction.

An oven dried flask was purged with nitrogen gas. B_2Pin_2 1.9 g (7.5 mmol) and then Catalyst 66 mg (0.1 mmol, Yellow solid) were added and dissolved in 5 ml dry cyclohexane. Small amount of HBPin was added to accelerate the borylation of the catalyst. The dtbpy ligand 0.2 mmol (54 mg) was added and the change of the color of the solution to bloody red was tracked. The mixture was stirred for 5 minutes and then the solid substrate 0.941 g (5 mmol) was directly added to the flask and stirred for 48h in reflux at 60 °C. The product was purified from the starting material left over by column chromatography.

Spectral data: ^{1}HNMR (300MHz, $CDCl_3$): **δ** [PPM] = 1.26 (s, 6H), 1.29 (s, 6H), 1.31 (s, 12H), 1.66 (s, 4H), 7.30 (dd, J = 7.8 Hz), 7.56 (dd, J = 7.8 Hz, J = 1.2 Hz, 1H), 7.75 (d, J = 1.2 Hz, 1H). ^{13}CNMR (125MHz, $CDCl_3$): **δ** [PPM] = 24.86, 31.76, 31.85, 34.2, 34.5, 35.09, 35.32, 83.5, 125.92, 131.93, 133.23, 144.08.

VI.11-2 S_NAr reaction

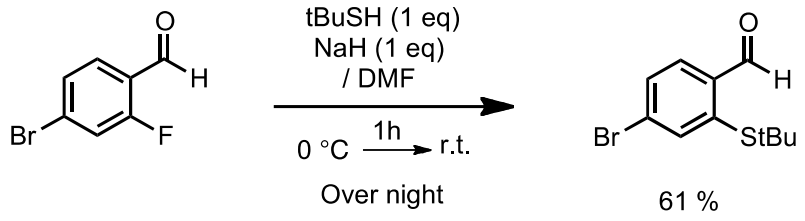
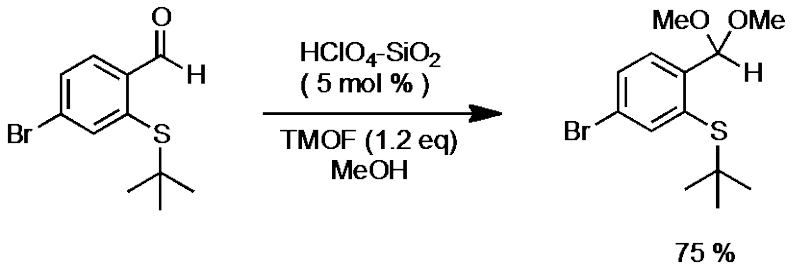


Figure VI-3) Nucleophilic aromatic substitution reaction.

To an oven dried Flask was added 1.03 g NaH (43 mmol) and 30 ml distilled DMF. The Flask was purged with nitrogen and cooled to 0 °C; 4.37 ml tBu-SH (43 mmol) was added slowly until all of the solid was dissolved (1h). In another flask the substrate was added 8.73 g in 30 ml DMF and cooled down to 0 °C. The content of the first flask was added to the second flask slowly and the reaction mixture was allowed to warm up to room temperature and stir overnight. The product was purified by extraction in 10% HCl and diethyl ether. The organic phase was dried with $MgSO_4$. (With distilled DMF more than one equivalent reagent will give the disubstituted product.)

Spectral data: ^{1}HNMR (300MHz, $CDCl_3$): δ [PPM] = 1.29 (s, 9H), 7.4 (dd, J = 5.1 Hz, J = 0.9 Hz, 1H), 7.78 (d, J = 0.9 Hz, 1H), 7.83 (d, J = 5.1 Hz, 1H). ^{13}CNMR (125MHz, $CDCl_3$): δ [PPM] = 31.21, 48.49, 128.58, 129.75, 133.15, 138.5, 18.72, 142.32, 192.74.

VI.11-3 Protection Reaction



75 %

Figure VI-4) CD-270 synthesis: protection reaction of the aldehyde.

To 0.45 g (1.65 mmol) of the aldehyde from the aromatic substitution reaction was added 1.2 equivalent trimethyl orthoformate (TMOF), one equivalent MeOH and 5 mol% dried $HClO_4$-SiO_2 reagent. The reaction was stirred for 1h and then purified.

Spectral data: ^{1}HNMR (300MHz, $CDCl_3$): **δ** [PPM] = 1.3 (s, 9H), 3.33 (s, 6H), 5.86 (s, 1H), 7.5 (d, J = 0.6 Hz, 2H), 7.69 1H (d, J = 0.6 Hz). ^{13}CNMR (125MHz, $CDCl_3$): **δ** [PPM] = 31.13, 47.68, 54.07, 101.93, 121.75, 128.71, 131.9, 133.99, 140.26, 141.79

REFERENCES

REFERENCES

1. Vasileiou C, *et al.* (2007) Protein design: Reengineering Cellular Retinoic Acid Binding Protein II into a rhodopsin protein mimic. *Journal of the American Chemical Society* 129(19):6140-6148.

2. Gill SC & Vonhippel PH (1989) Calculation of Protein Extinction Coefficients from Amino-Acid Sequence Data. *Analytical Biochemistry* 182(2):319-326.

3. Marti T, Rosselet SJ, Otto H, Heyn MP, & Khorana HG (1991) The retinylidene Schiff base counterion in bacteriorhodopsin. *The Journal of biological chemistry* 266(28):18674-18683.

4. Berbasova T, *et al.* (2013) Rational Design of a Colorimetric pH Sensor from a Soluble Retinoic Acid Chaperone. *Journal of the American Chemical Society* 135(43):16111-16119.

5. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* 276:307-326.

6. Winn MD, *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D* 67:235-242.

7. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* 66:213-221.

8. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D* 66:486-501.