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

By

Zahra Nossoni

A DISSERTATION

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

Chemistry - Doctor of Philosophy

2014

ABSTRACT

STRUCTURAL INSIGHT OF THE MECHANISM OF WAVELENGTH TUNING A RHODOPSIN MIMIC
AND
A SINGLE MUTATION RESULTED AN EXTENSIVE
3D DOMAIN SWAPPED DIMERIZATION IN HCRBPII

By

Zahra Nossoni

In the human eye the three types of the rhodopsin (blue, green, red) are responsible for color vision. All of these color pigments bound to a single chromophore: 11-*cis*-retinal. It is not still clear that how the interaction of this chromophore with these different opsins leads to different wavelength spectrum. Since working with rhodopsins (membrane proteins) are challenging therefore in this study human Cellular retinol binding protein II (hCRBPII) small, cytosolic and soluble protein has been selected to be used as rhodopsin mimic. The first step was generating of the active lysine in the binding pocket that can bond to all-*trans*-retinal and forms the Schiff base. Systematic mutations on this protein followed by x-ray crystallography lead us conversion of hCRBPII to mimic of rhodopsin that covers the full visual spectrum. Crystal structure of several holo mutants bound to retinal at high resolution illustrate that changes in the ordered water networks (making water net work, removing eater net work or changing in dipole moment) in the binding pocket leads to tremendous changes of the absorbance of protein bound to retinal. The mutants that do not effect on the ordered water networks do not make the big differences in wavelength. The crystal structures also showed that in this system despite rhodopsin the counter

anion is not necessary and the positive charge of the protonated Schiff base can be stabilized with the other interactions.

Based on these data we proposed the mechanism of the mechanism of wavelength regulation in one rhodopsin mimic. We could also correlate the pK_a of protonated Schiff base with the water networks and wavelength. We have observed that the electrostatic interactions between the amino acids and also polar groups (water) in the binding pocket have an undeniable role on spectral tuning.

By introducing one specific mutation on hCRBPII, this protein forms a very stable dimer and domain swapped protein. This unique protein folding mechanism so far has been observed in only around 40 different proteins and is believed to be a mechanism of evolution of the dimers and oligomers. The high resolution crystal structure of the dimer and domain swapped hCRBPII shows that we could trapped or produce a very stable partially folded protein and with this data we were able to explain the mechanism of protein folding In hCRBPII.

This Dissertation is lovingly dedicated to grandfather, Agha joon

ACKNOWLEDGEMENTS

"By the power of truth, I while living have conquered the universe." Dr. John Faust

PhD time is like spring weather: sometimes stormy, sometimes rainy, sometimes sunny and beautiful. It is certainly not stable, but still beautiful. It passed much faster than I expected. I really don't know at this step if I am happy to finish it or sad to leave the lab, the office, the bench and everything I used to have for more than 5 years and move on. While my heart aches, I am glad that I was able to finish one of the toughest stages of my life. I was not alone on this journey and honestly would not have finished this without the help of those close to me. I am glad and thankful to have great supporters that were with me shoulder to shoulder. I would like to express my deepest appreciation to my advisor Prof. James Geiger. Jim taught me how to think as a scientist, how to think independently, how to face and solve the problems scientifically. Not only was Jim my academic advisor, he was also my role model. I learned from Jim how to be a brilliant "Thinker" yet still be down to earth. The transitions I've made from the person I was to who I am now I owe in great part to him. I am so glad and honored to have him as my advisor as he will be part of my life forever. Thank you Jim, for everything.

I would like to thank my second reader and committee member Prof. Babak Borhan. He was the first person who believed in me. Our weekly bio meetings with Babak were not always smooth and easy, instead often challenging. As such I learned

more in each and every meeting. Babak really changed me from a very tender and fragile person to a scientific fighter. Thank you Babak, I really appreciate it. I really do.

I would like to appreciate Dr. Kevin Walker for being on my committee and reading my dissertation. Dr Walker always gave expressed positive energy when I was walking in the hallway in the second floor.

In addition, thank you to Professor Bill Henry for also reading my dissertation. I received some great advice from him when I struggled with problems in cloning.

I want to thank Dr. Wenjing Wang for being the best lab partner that anyone would wish to work with. Wenjing is not only was a great co-worker, but also has a great personality making her a joy to work with.

I also would like to thank Dr. Chrysoula Vasileiou for her help. I knew Chrysoula before I joined graduate school in the chemistry department. My brother was working with her and always talked highly about her. Honestly, I thought he must be exaggerating about how smart and helpful she was until I witnessed it personally. Any description I might provide isn't likely to do her justice. She has such presence in the organic group and personally I cannot imagine 5th floor without her. Thank you, Chrysoula.

I consider myself to be very fortunate to be in an awesome group with great group members. I would like to thank all current and former members in Geiger's group for being friendly and helpful to me. Specifically, thank you Remie Fawaz, Dr. Camille Watson, Meisam Nosrati, Zahra Assar, Dr. Xiafei Jia, as well as Rebecca Ober, Lindsey Gilbert and Janice Chiou. I also would like to thank the bio girls in Pro. Borhan's group specifically Tetyana Berbasova and Ipek Yapici.

I would like to thank LS-CAT staff at the Argonne National Laboratory for all of their help during data collection.

I would also like to express thanks to my dear friends who supported me on this path, particularly my very best Friend Tate Schaar. When I was frustrated and unhappy during the hard days of graduate school, he was my greatest support. Thank you Tate.

Finally, I would like to thank my dear family. My mom, dad, sister (Dr. Goli Nossoni), brother (Dr. Farid Nossoni) and baby brother Hessam. Without them, their love and support, it would not be possible for me to finish this stage.

I owe my success to all of the wonderful people who support, love, and help me.

I would also like to express my thanks to Spongebob, Patrick and Piglet. During the darkest of times you kept my mind positive and were a constant reminder that some things in this universe don't have to always follow obvious paths of logic or sense.

Thank you all.

"All of the pages of this book are full but the story is still unfinished…" Saadi Shirazi, Iranian Poet of the 1200

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO SYMBOLS AND ABBREVIATIONS

Ala, A	Alanine
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartate
Cys, C	Cysteine
Gln, Q	Glutamine
Glu, E	Glutamic acid
Gly, G	Glycine
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
Å	Angstrom
MW	molecular weight

μM	micromolar
μL	microliter
Amp	ampicillin
bR	bacteriorhodopsin
cGMP	cyclic guanosine monophosphate
Clm	Chloramphenicol
CRABPII	cellular retinoic acid binding protein II
CRBPII	cellular retinol binding protein II
Da	Dalton
DMSO	dimehtylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
dsDNA	double-stranded DNA
dTTP	deoxythymidine triphosphate
Equiv	equivalent
EDTA	ethylenediaminetetraacetic acid
E. *Coli*	Escherichia coli
ε	extinction coefficient
FABP	fatty acid binding protein
FPLC	Fast protein liquid chromatography
FTIR	Fourier transform infrared spectroscopy
GDP	guanosine diphosphate

GPCR	G-protein couple receptor
h	hour
iLBP	intracellular lipid-binding protein
IPTG	isopropylthio-β-galactoside
K_d	dissociation constant
L	Liter
LB	Luria Broth
λ_{max}	maxima wavelength
Ni-NTA	Nickel-nitrilotriacetic acid
M	molar
PCR	polymerase chain reaction
PEG	polyethylene glycol
PSB	protonated Schiff base
R	rhodopsin
R-factor	reliability factor
RT	Room temperature
RMSD	root-mean-square deviation
SB	Schiff base
SDS	sodium dodecyl sulfate
SDS-PAGE	sodiumdodecyl sulfate polyacrylamide gel electrophoresis
TM	Transmembrane

7TM	Seven transmembrane
UV	Ultraviolet light
Vis	Visible light
WT	wild type

CHAPTER I

Introduction of Color Vision

I-1:Introduction

"Somewhere, something incredible is waiting to be known." Carl Sagan

Colors are part of our life, we start our life with them and sometimes we forget to ask this question of how we can see them. If we just think about the human color vision, we know that we can distinguish different absorbances from 400-700 nm. But how can we see the different colors?

But before we can answer this question we need to know about the properties and character of light. Light is a range of electromagnetic radiation that to detectable by human eye. The different color has different wavelength and as the result have different energies. The energy of each wavelength can be defined by the Plank-Einstein equation $E=h\nu$, where h is Plank's constant and ν is the frequency. The frequency is related to the wavelength by the following equation:

$$\text{Frequency} = \frac{\textit{Speed of light (C)}}{\textit{Wavelength}}$$

The energy of the light can be absorbed by the photoreceptor of the eyes and transformed in to an electrical signal to be processed by the brain. To better understand the mechanism of light detection, let's look at the structure of human eye.

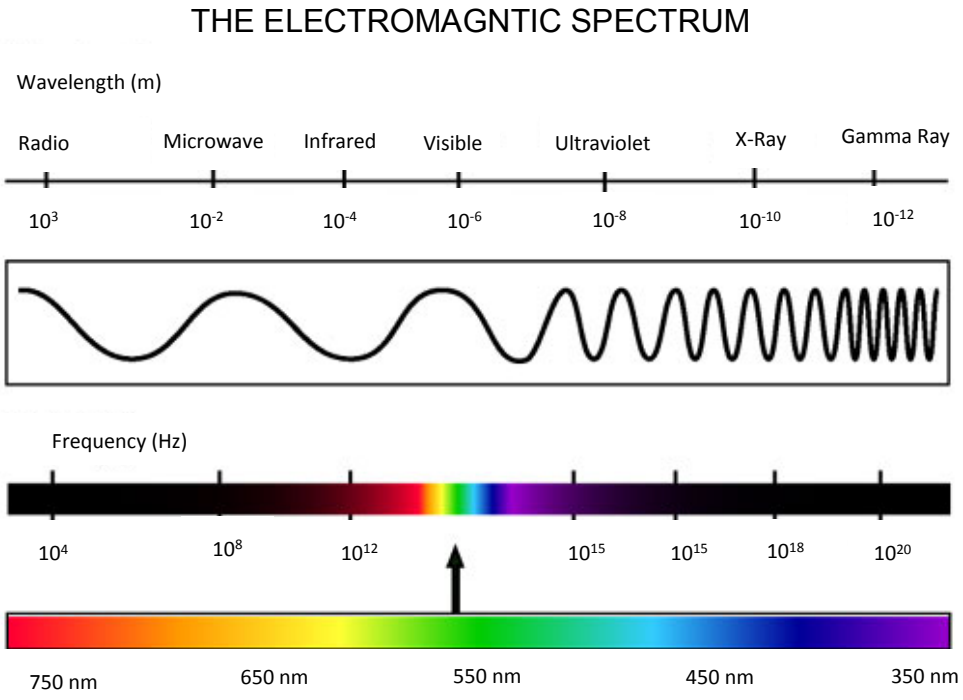


Figure I-1: The electromagnetic spectrum

Light is absorbed by photoreceptor cells in the retina, which is the most inner layer of the eye and plays the most important role in vision and color detection. There are two different types of cells are in the retina, cone cells and rod cells and as it is shown in Figure I-2 these cells are located in the outer layer of retina. The name of rod and cone cells comes from the shape of these photoreceptor cells.

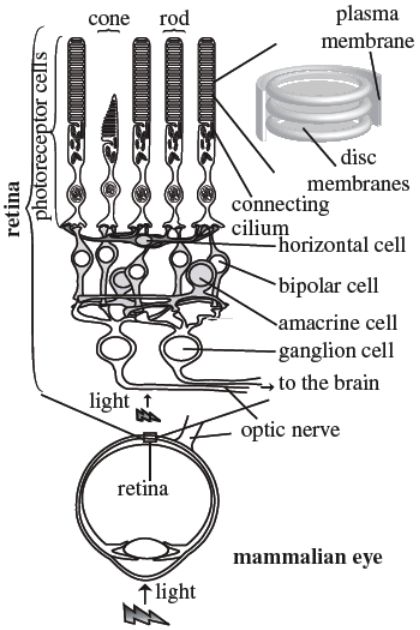


Figure I-2: Structure of mammalian eye[1]

Rod cells are responsible for the dim and black-and-white vision under dim light. These cells are responsible to see the black-and-white images in the dark. These cells are very sensitive to light but do not mediate color vision.[2,3] The other types of photoreceptors are cone cells are responsible for color vision and carry these color pigments (blue, green and red rhodopsin). The three cone cells are blue, green and red, the blue rhodopsin is sensitive to short wavelength (425 nm) while green rhodopsin is sensitive to medium wavelength (533 nm) and red rhodopsin is sensitive to long wavelength (560 nm).[2,4] in the other hand the cone cells are for the color vision and daylight and high resolution vision. These cells are mostly concentrated on the retina, also called “yellow spot”, which is located right across from the pupil of the eye and is 300-700 μm in diameter. [4] In human eye there are approximately 120 millions rod cells

while the cone cells are only 6 to 7 millions. These cells are located at the most outer layer of the retina. The cone and rode cells are neural cells that can detect the light and convert it to an electrical signal which process in the visual section of the brain. The rod cells are responsible for the dim vision and they are very sensitive the light. The optical cells are divided into two parts: the inner segment (IS) and the outer segment (OS). In both type of cells the outer segment contains the light absorption molecular structure[5]. As it is shown in Figure I-3 the outer layer contains a highly dense pack of tiny disks, these disks contains the light receptor proteins, rhodopsins[5].

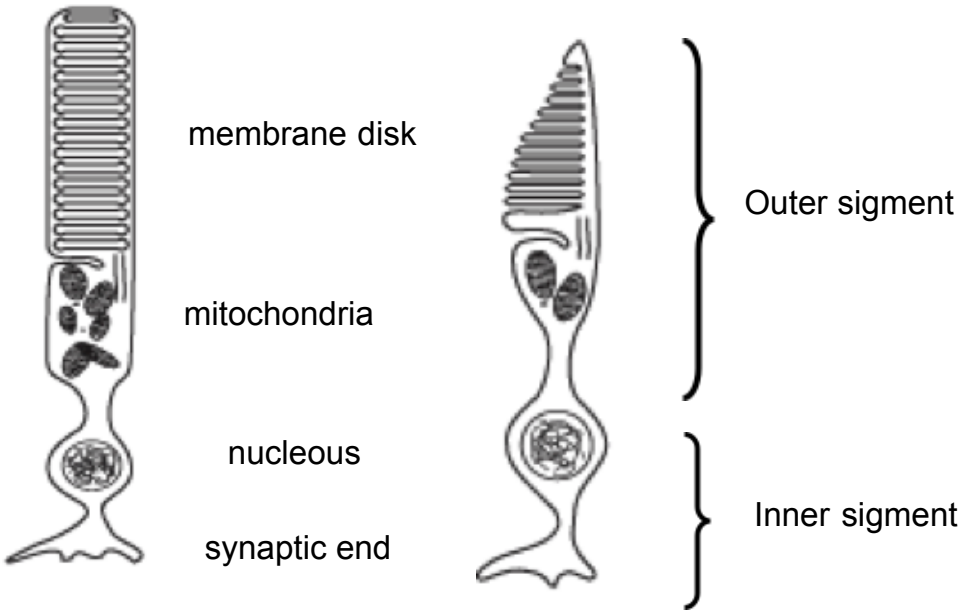


Figure I-3: The schematic structure of rod and cone photoreceptor cells

The inner segment (IS) contains the nucleus and other cell parts such as golgi body and mitochondria which are necessary for the cell functioning. Beside these parts

the inner segment is also comprised of a synapse terminus, which can send the electrical signals to the neurons in the inner layer of retina. The signals are transmitted through an elaborated array of synapses in the retina toward the visual section of the brain.[6,7,8]

As it was mentioned above both outer segment layer of cone and rode cells contains the photoactive protein, rhodopsin. In the 1930's it was discovered that rhodopsin consists of two parts: a protein part, which is called opsin and a chromophore bound to the protein, which is identified,] as 11-*cis*-retinal. [9,10]

Opsin is a seven membrane helical protein from the family of the G-protein couple receptors;[11,12] this protein can bind covalently to 11-*cis*-retinal via Schiff base to an active lysine residue. The Schiff base can be protonated and deprotonated (FigureI-4).

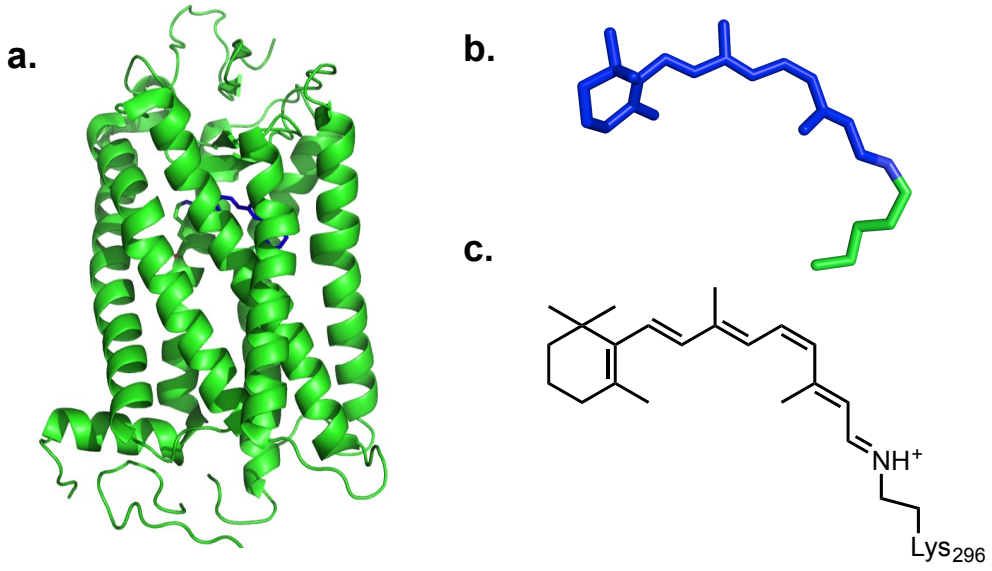


Figure I-4: Crystal structure of bovine rhodopsin (PDB entry: 1F88).

Rhodopsin belongs to the superfamily of protein, guanine nucleotide-binding protein G protein coupled receptors (GPCR).[12,13] These proteins can be activated by an external stimuli, such as light, chloride concentration, hormones, odors and etc. Activation of the G-protein initiates an enzymatic cascade reaction, which produces the neuro signals. All of the GPCRs are composed of the seven membrane helices (T7M), and they connect to each other by six loops, three of them intracellular and three extracellular. Over 90% of the protein is in the disk membrane with a very high density of 25,000 rhodopsin/μm.[14,2] The N-terminus of the protein is located at the extracellular surface and the C-terminus at the intracellular (cytoplasmic). Rhodopsin is a unique member of the GPCR family. Despite the other family members has no covalent interaction with the ligand, rhodopsin (both cone and rode) bind covalently to the ligand 11-*cis*-retinal. The only crystal structure of visual rhodopsin was determined in 2000 by Palczewski at 2.8 Å resolution (PDB ID: 1F88).[15] This structure reveals that 11-*cis*-retial has a covalent bound to Lys296 as a *cis*-Schiff base (SB). Glu113 is the counter anion and makes hydrogen bond to the SB, stabilizing the protonated state of the pigment. This feature increases the pK_a of PSB to higher than 16. The chromophore is oriented mostly parallel to the surface of the membrane. This orientation allows the chromophore to absorb the maximum amount of light and increase the light sensitivity.[18,19,20] The chromophore is deeply buried in the protein and the binding site is not accessible to the cytosolic or extracellular surface. The chromophore has interaction with the residues of the neighboring helices and therefore it is locked in the position. The crystal structure shows no gaps between the helices that would let the retinal exit

the binding pocket. However like the other ligands retinal must move out of the binding pocket for regenerating since all-*trans*-retinal must isomerize to 11-*cis*-retinal.[16]

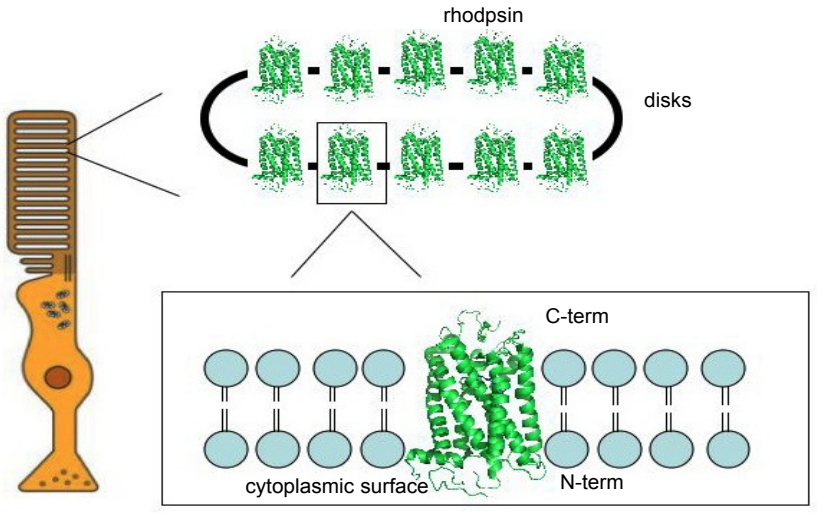


Figure I-5: The Rhodopsin spans the membrane disk of rod and cone cells.

I-1-1: Conformational Changes of 11-*cis*-retinal Upon Light Absorption

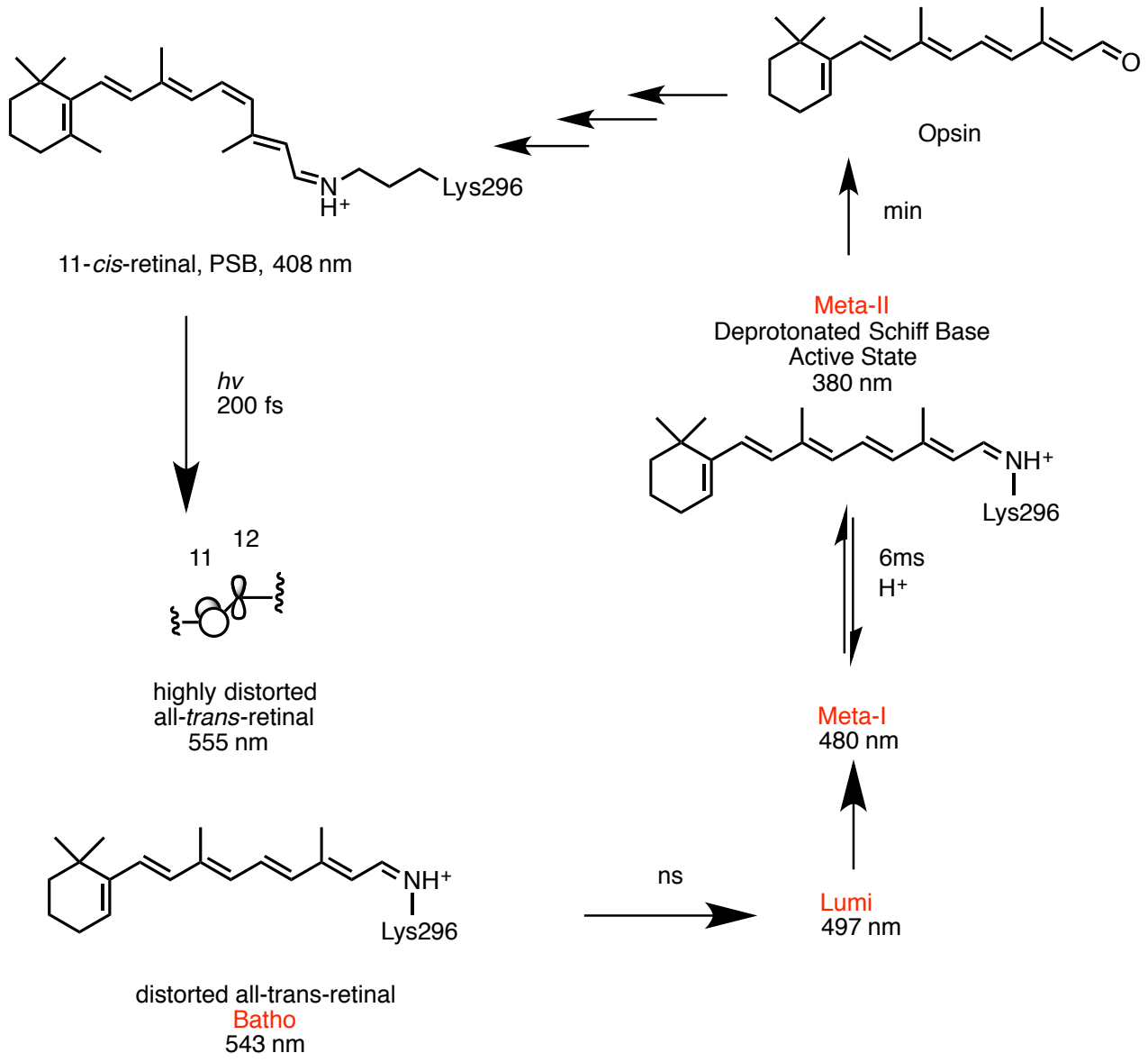


Figure I-6: The cycle of retinal conformational change after light absorption of bovine rhodopsin

Within a few picosecond of light absorption the chromophore goes through a series of conformational changes, which are described in Figure I-7 and the chromophore's structure changes from 11-*cis*-retinal to all-*trans*-retinal.[17] This conformational change of the chromophore triggers conformational changes on the protein. This retinal *cis-trans* isomerization is usually monitored at low temperature using various spectroscopic techniques such as UV-Vis.

When rhodopsin (R) absorbs the light, it converts to meta II, the active state (R*). Formation of meta II causes conformational changes in the T7M helical structure of the opsin protein and initial the biochemical cascade.[21, 22] This biochemical cascade is illustrated in Figure I-6. Upon light activation, 11-*cis*-retinyledene isomerizes to the all-*trans*-isomer.

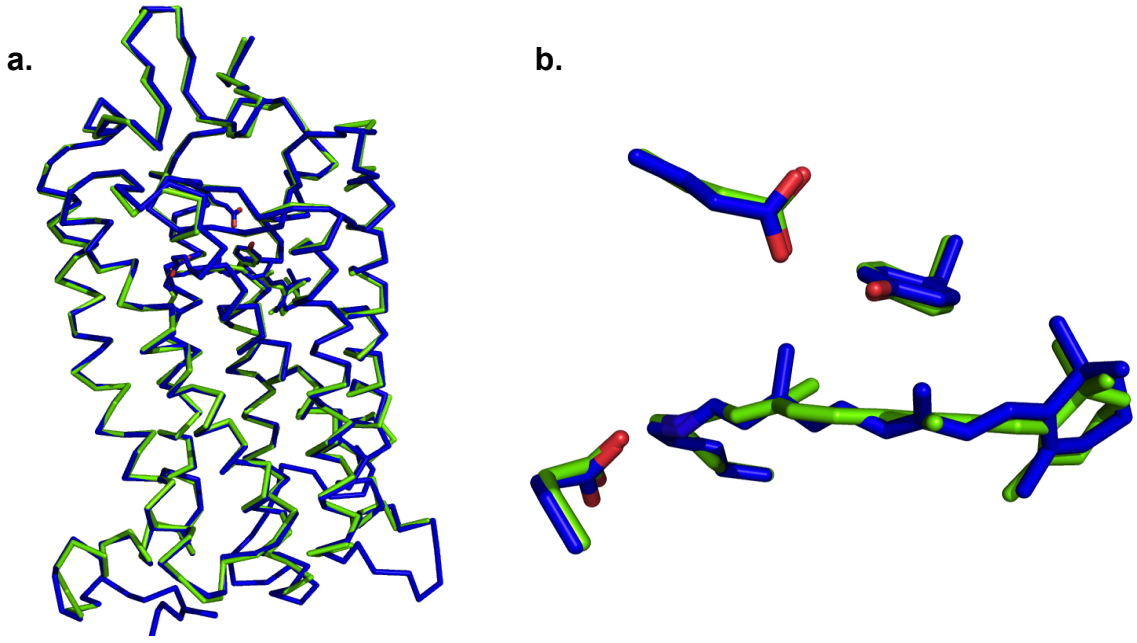


Figure I-7: The crystal structure overlay of bathorhodopsin (blue) (PDB entry 2G87) and ground state rhodopsin (green) (PDB entry 1U19)

This isomerization from *cis* to *trans* happens in picoseconds. The first stable intermediate form, which is called bathorhodopsin, has λ_{max} at 543 nm and the crystal structure of bathorhodopsin was determined by the exposure of bovine rhodopsin crystal to 488 nm light at low temperature of 98K (PDB entry 2G87). [23]

The overlay structure of the bathorhodopsin and ground state rhodopsin indicates that the overall structure of the backbone of the protein does not change at this step and only the chromophore isomerizes from 11-*cis* to distorted all-*trans*. Since in the structure of bathorhodopsin the ligand isomerizes to a highly twisted form, therefore the energy of the ground state increase and the energy gap between the ground and excited states decreases. As the result a dramatic bathochromic shift is observed from 498 nm (ground state rhodopsin) to 555 nm (bathorhodopsin). After this step rhodopsin forms two unstable intermediate Lumi and Meta I. Meta I further convert to meta II (the active state). In this step the proton transfers from the PSB to the counter anion. The crystal structure of this photo intermediate was determined in 2011 by two different groups (PDB entry 3PQR).[21,22] The crystal structure of this intermediate is available with the short Gα-helices from transducine. The crystal structure of this intermediate illustrates significant conformational changes in compared to ground state rhodopsin.

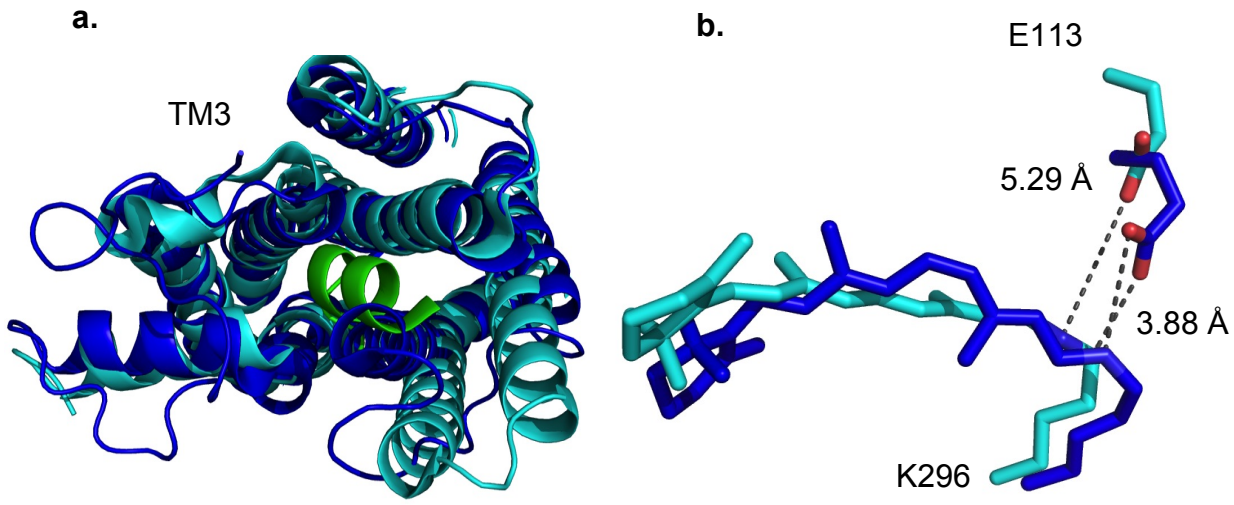


Figure I-8: The overlaid crystal structure of ground state bovine rhodopsin (dark blue, PDB entry: 1F88) with Meta II (cyan, PDB entry: 3PQR).

In Meta II, the transmembrane helix 3 moves from their position to open up the binding site for Gα peptide of transducine. As it is illustrated in Figure I-8.a due this significant structural changes in Meta II, the chromophore (all-*trans*-retinal) is bound in a very different way and this conformational changes can explain the drop of the PSB-pK_a of this intermediate. In Meta II the PSB deprotonated and the proto transfer to the counter anion Gli113. As it is shown in Figure I-8.b the distance of the counter anion, Glu113, from PSB of Meta II rhodopsin increases to 5.3 Å. Therefore the hydrogen bond between the PSB and the counter anion disturb and as the result the pK_a decreases. Among all of rhodopsin intermediates, Meta II has the longest lifetime as compared to the other rhodopsin intermediates. The conformational changing of retinal causes changes of the transmembrane helices and leads to activation of the phototransduction

cascade.[25] At the end of the visual cycle the iminum bond of Meta II hydrolyzes and Meta II breaks down to all-*trans*-retinal and opsin and the visual cycle repeats.[24,25]

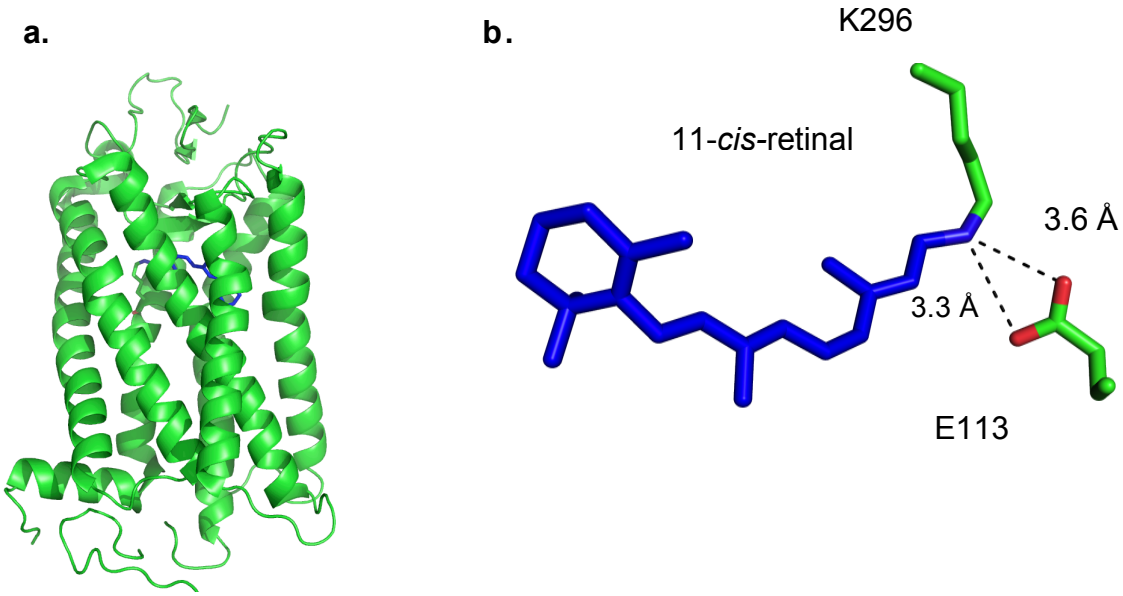


Figure 1-9: Crystal structure of bovine rhodopsin (PDB entry: 1F88)

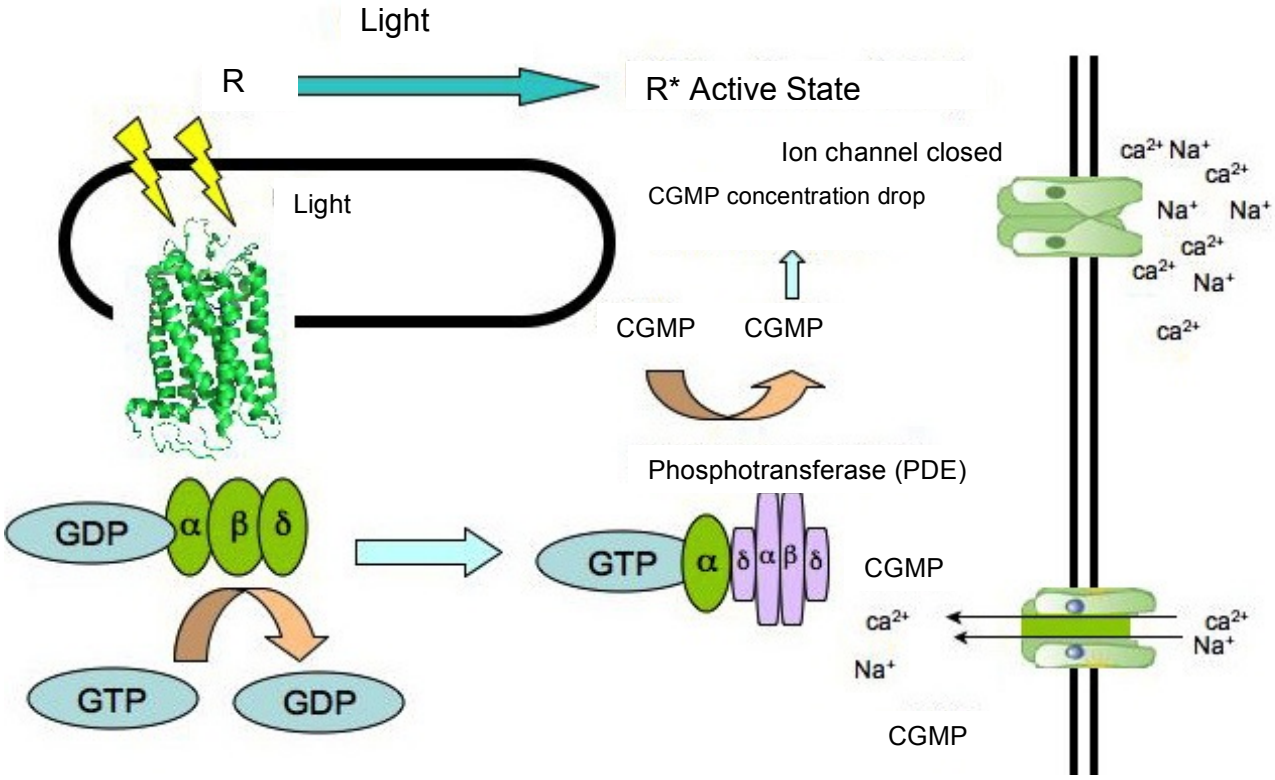


Figure I-10: The schematic diagram of Phototransduction cascade

I-1-2: The Phototransduction Cascade

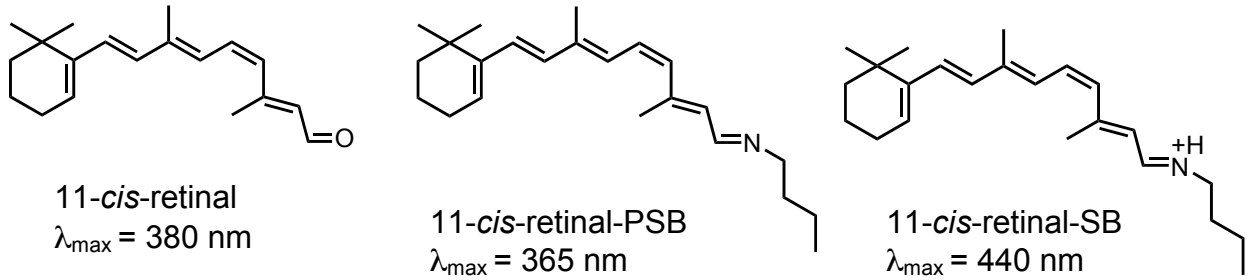
The biological cascade of phototransduction and activation of G-proteins is illustrated in Figure I-10. The rhodopsin active state (R*) binds to GDP-bound transducin (Gt), resulting of this part is changing GDP changing to GTP.[26,27] Gt has three subunits: α, β, γ. The GTP and subunit α (Gtα-GTP) dissociate from the other two subunits ($\gamma\beta$). Gtα-GTP binds cGMP PDE (phosphodiesterase) and remove the inhibitory subunit (γ). The activated PDE converts cGMP (3', 5' cyclic monophosphate)

to GMP.[24] The reduction of the level of cGMP in the cell makes cGMP gate-ion channel close and as the result is hyperpolarization of the membrane and generation the electric signal from the basis of vision.[25,28,29] The active form of rhodopsin (R*) separates from the inactive part of Gt and recycles to active more transducins. Since hydrolysis of R* is a slow process therefore R* has enough time to activate 10^2 PDEs. Each PDE is able to convert 10^3 cGMPs to GMPs, therefore each activated rhodopsin is capable to convert 10^5 cGMPs to GMPs.[72]. It starts with activation of rhodopsin by light. The α subunit of Gt separates and activates the phosphotransferase which coverts CGMP to GMP and the ion channel closes. This conversion polarized membrane leads in to neural signal. Since the hydrolysis of the imine bond between all-*trans*-retinal and the active rhodopsin (R*) is very slow R* has enough time to activate more than 10^2 of PDEs, which each of them can covert ~ 10^3 cGMPs to GMPs. Therefore one R* has the ability to covert 10^5 cGMPs to GMPs. At the end of the cycle all-*trans*-retinal dissociates from the opsin for regeneration. The first step involves a dehydrogenase that reduces the aldehyde to all-*trans*-retinol and moves it back to the retinal pigment epithelium (EP) for regeneration to the original11-*cis*-retinal. The opsin can bind to a new regenerated 11-*cis*-retinal and initiate the new cascade.

I-1-3: Color Vision with a Single Chromophore

All the different color pigments (cone rhodopsins) bind with a single chromophore, 11-*cis*-retinal. However they all have different absorbance. Free all-trans-retinal and 11-*cis*-retinal solution on the organic solvent like ethanol absorbs at 380 nm.

After formation of Schiff base (SB) with n-buthylamine (to model rhodopsin) the pigment absorbs at 365 nm. However protonated SB (PSB) absorbs at 440 nm (Figure I-11). [30] By changing the concentration and also the solvent this absorbance can shift to 500 nm.[34]



11-*cis*-retinal
λ_{max} = 380 nm

11-*cis*-retinal-PSB
λ_{max} = 365 nm

11-*cis*-retinal-SB
λ_{max} = 440 nm

Figure I-11: The different absorption of retinal as free aldehyde in methanol, bound as SB to n-butylamine and protonated form of Schiff base.

In 2005 the absorbance of all-*trans*-retinal with protonated n-buthylamine Schiff base as the model of rhodopsin complex were was measured in the vacuum and it was found the λ_{max} at 610 nm.[31,32] Since in this system there is no counter anion present that can stabilize the positive charge of the iminium, therefore the positive charge can be delocalized to the largest extent and this can lead to the moret bathochromic shift. Although formation of the PSB is enough to explain the absorption of the blue and rode pigments but it is not enough to explain the red shifted absorption for the green and red rhodopsin. [35]

The protein-chromophore interactions must play a critical role for the spectral tuning of the chromophore. [36,37,38] Therefore the single chromophore in the different rhodopsins can absorb from 400 to 600 nm. The difference between the λ_{max} of the

pigment from the PSB-butylamine model is called " Opsin Shift"[33]. This term also can define the effect of the opsin binding site on the chromophore. Although the Opsin shifts have been known for the long time,[39] still the mechanism of opsin shift is not clear. Different theories were suggested to explain the opsin shifts. One of the theories suggests that interaction of the residues inside the binding pocket of the opsin with the chromophore causes red shift as compare to the retinal in solution.[40] Unfortunately there is not any crystal structure of any of the color pigments to study on the interaction of the residues inside the binding pocket with the chromophore. As it was mentioned above the first and only available structure of the rhodopsins is bovine rhodopsin. This structure was determined at 2.8Å by Palczewski, however the higher resolution structures of this protein were also determined later. The higher resolution crystal structure at 2.2Å of rhodopsin illustrates that the β-ionone ring of the chromophore adopts the 6-s-*cis* conformation.[15] This structural data also indicates the isomerization around the C11-C12 double bond upon absorption of photons.[41] As of today there is no crystal structure of any cone rhodopsin available therefore the only information about the structure of these color pigment can be obtained through indirect methods such as NMR labeling, site mutation and advanced theoretical calculations.[42,43,44]

I-1-4: Modeling of the Cone Rhodopsin Based on the Crystal Structure of Bovine Rhodopsin

As was mentioned above so far the only available crystal structure of a visual

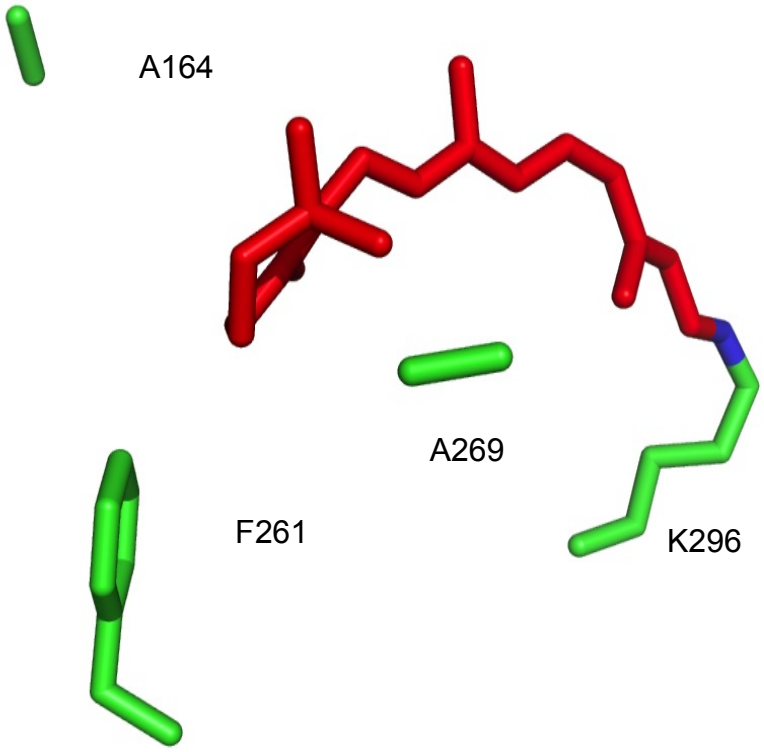


Figure I-12: The crystal structure of bovine rhodopsin with highlighted positions, which were mutated by Sakmar's group.

pigment is the structure bovine rhodopsin.

This crystal structure was used as the platform for rational mutation and modeling of the three cone rhodopsins. The crystal structure of bovine rhodopsin illustrates that the

binding pocket of this protein is made of the hydrophobic residues. Even though no crystal structure of color pigment is available, the theoretical models of the three cone rhodopsins illustrate the same fact. Nathan and his co-workers sequenced the gene of red, green and blue rhodopsin in 1986.[45] Red and green pigments are 96% identical. They are only different in 15 amino acid residues in the whole protein and they have almost identical in the binding pockets, while blue rhodopsin is very different from the other two. Both green and red rhodopsin bind to the chromophore via Lys312 (which is identical to Lys296 in bovine rhodopsin) and counter anion is Glu129 (analogous to Glu113 in bovine rhodopsin). The second glutamic acid (Glu102) is available in the green and red rhodopsin; this amino acid is absent in blue rhodopsin.

Table I-1: Sequence homology between human rhodopsins

	Percentage			
	Rod	Blue	Red	Green
Rod	100	75	73	73
Blue	41	100	79	79
Red	37	43	100	99
Green	38	44	96	100

In 1991 Neitz and his coworkers show that three amino acids that are different in the binding pocket of green and red rhodopsin are probably responsible for the 30 nm absorbance differences between green (562 nm) and red rhodopsin (530 nm) [46]. The

crystal structure of bovine rhodopsin illustrates that these three amino acid residues are located at the ionone ring region and mutation of these hydrophobic amino acid residues to polar amino acids leads to more red shifted absorption (Figure I-12). In the blue pigment Tyr262 is approximately located in the ionone region while at the same position there is a Tryptophan in both green and red rhodopsin. Mutation of Tyrosine to Tryptophan (Y262W) leads to 10 nm red shift in the absorption spectra. [47] This study illustrates the important role of Trp262 in red shifting of the wavelength. However without the crystal structure it is not clear the mechanism and the effect of each amino acid. On modulating the spectra in the protein. Oprian and his group mutated all of the 15 amino acid residues, which are different in green and red rhodopsin.[48] They observed that seven of the mutations are responsible for the wavelength for spectral tuning, whiles Neitz proposed three of these positions previously. Sakmar's group made the mutation on the three amino acids residues, which were proposed by Neitz in bovine rhodopsin. The three positions are Ala269, Phe261 and Ala164. Sakmar's group did mutations on nine different positions in the bovin rhodopsin and they could modulate the wavelength from 500 nm to 438 nm (62 nm) more blue shift with overall nine mutations.[49] These mutations are A299C, A292S, G90S, A117G, M86L, A299C, A124T, W265Y, and E122L.[51] The general conclusion from these studies was that introduction the negative charge at the PSB region and removal of a negative charge from the ionone region can lead to more blue shifted pigment. However, since the crystal structures of these mutants are not available therefore this conclusion could not be verified. In 1991 Jacobs and his coworkers demonstrated that mutation of three non-polar amino acids in green rhodopsin with polar amino acid (hydroxyl bearing) could

cause the majority of the absorption difference between blue and green rhodopsins.[46] These mutations are A108S (~4nm), F277Y (~10 nm) and A285T (~ 16 nm). Based on the theoretical model of green rhodopsin these residues are located around the ionone of the chromophore. In 1992 Sakmar's group made the same mutations at the equivalent positions of rod rhodopsin. The tetra mutant A:164S:F261Y:A269T did not bind to the chromophore. However the double mutant F261Y:A269T displays 20 nm more red shifted, while the single mutant A164S showed a slight red shift (~2nm) as compared to the wild type structure .Therefore it seems that only two amino acids, F261 and A269, are responsible for the differences between green and red pigment in human eye. These data show that spectral tuning of the chromophore by the different mutation is not exactly additive and can enhance protein-chromophore interaction.[46] Different studies have been done on the counter anion region as well. Oprian and his group mutated all of the glutamic acid and aspartic acid residues in the rhodopsin and found that Glu113 is the counter anion since the E113Q mutation leads to significantly lower pk_a of the PSB of the bovine rhodopsin.[50] After this study Sakmar's group mutated Glu113 and they showed that that by shortening the length of the counteranion by introducing an aspartic acid instead or by removing Glu113 completely and introducing the hydrophobic amino acid like alanin up 30 nm more red shift is obtained.[52]

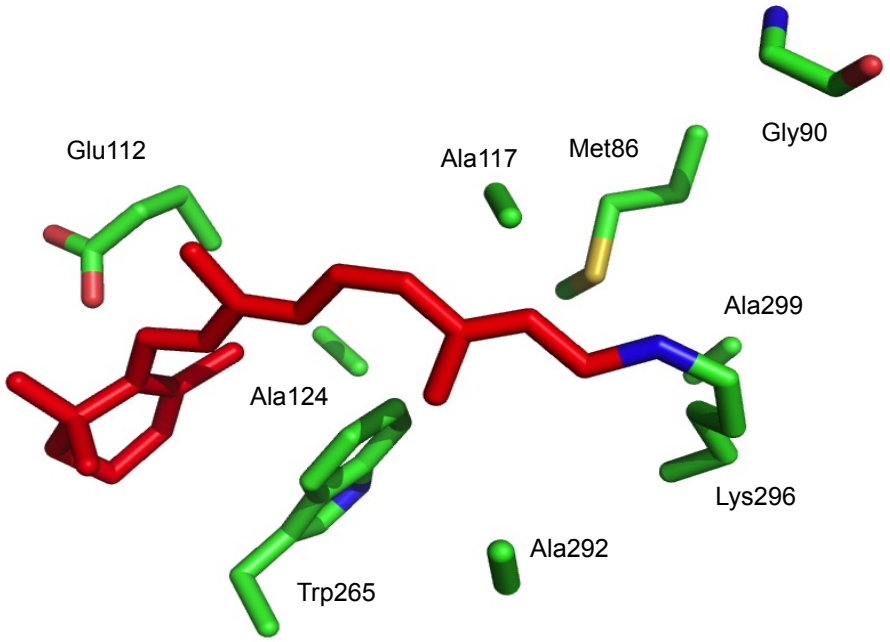


Figure I-13: The crystal structure of bovine rhodopsin PDB entry (1F88) with the highlighted amino acid residues that were mutated by Sakmar's group.

The result of this study suggest that removing the negative charge close to the PSB region leads to more delocalization of the positive charge along the chromophore and therefore red shift. The second glutamic acid in both green and red rhodopsin is believed to be responsible for the additional blue shifting in these two rhodopsins.

In both green and red rhodopsin, chloride-binding sites have been observed and it was believed that this anion is responsible for the point charge in modulation of the absorbance in the chromophore.[53] The other two positively charge amino acids (His197 and Lys200) that are conserved residues in the long-wavelength pigment are believed to forming the chloride binding site. These residues are not present in the short-wavelength pigment. Although all of these studies have resulted in significant improvement of our knowledge about the mechanism of wavelength regulation and

color vision still no crystal structure of any of these mutants rhodopsins are available. To fully understand this mechanism and study on the interactions between the amino acid residues and the bond chromophore the crystal structure of the WT opsin and their mutants are necessary.

1-2: The Proposed Theories For the Mechanism of the Wavelength Regulation

Over more than 50 years several different mechanisms have been in an attempt to explain for the mechanism of the wavelength regulation by the different color pigments. These theories are mostly focused on the delocalization of the π-electrons of the chromophore. [54] Based on the general knowledge on conjugated systems, more delocalized π electrons lead to more red shifting.

The role of the counteranion and its interaction with the PSB on the mechanism of the modulation the absorption maxima was studied. In 1993 Sheves and his coworkers used a model compound to mimic the PSB in rhodopsin and they illustrated that if the counteranion is moved away from the PSB, a significant red shift absorption maximum of the rhodopsin is obtained. They also showed that changing the angle of approach for the counter anion does not have a significant effect on the wavelength, however it could perturb the pK_a of the PSB.[55]

Another possible explanation of the opsin shift is based on the point charge model. Based on this model placement of another negative charge beside the counteranion along the polyene chain of the PSB-chromophore can change the delocalization (or conjugation) of the positive charge of the PSB along the polyene.[56,57]

It is obvious that the presence of the counteranion is necessary for stabilization of the retinal-PSB in order to achieve a high pK_a for the rhodopsin. By removing the counteranion in rhodopsin the PSB becomes very unstable. Therefore removing of the counteranion cannot explain the bathochromic shift from green rhodopsin (530 nm) to red rhodopsin (560 nm).

Nakanishi and Honig proposed the point charge is located in a different place in bacteriorhodopsin and bovine rhodopsin.[56,57] Based on this suggestion in the bacteriorhodopsin the point charge is located close to the ionone ring, while in bovine rhodopsin is closer to in the middle of the polyene chain of the chromophore. Therefore the bacteriorhodopsin (560 mm) is more red shifted than bovine rhodopsin (500 nm).

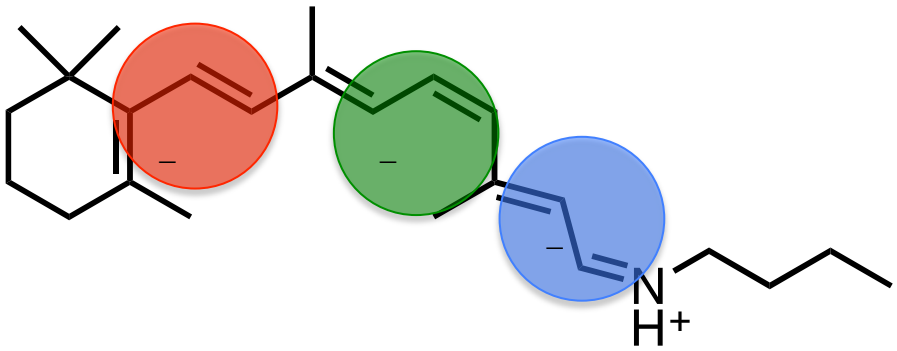


Figure 1-14: Positioning of point charge or dipoles along the polyene chain of the chromophore may cause spectral tuning of the chromophore.

This theory has lost popularity since mutagenesis studies did not have a large effect on the wavelength expect for the counteranion.[58] Besides that, the crystal

structures of bovine and bacteriorhodopsin also illustrate that the only negative charged amino acid in the binding pocket is the counter anion.[15] Since no permanent negatively charged residue was found in the binding pocket of rhodopsin, therefore the polarizability of the residues could play an important role in the opsin shift. In the other words, positioning of the polarizable groups in the binding pocket and along the chromophore can favor or disfavor the delocalization of the charge depending on the orientation and therefore cause the shift in the wavelength.[39] In 1976 Mathies and Stryer showed that at the excited state of the protonated Schiff base the negative charge transfer from the ionone ring to the PSB upon electronic excitation. In the other word at excited state the positive charge of the PSB is transferred and delocalized to the ionone ring.[54] Therefore the permanent dipoles of the amino acid residues, which have interaction with retinal could generate inducible dipoles upon electronic excitation and the maximum and optimal interaction with the ground states. This theory is supported by experiments with polarizable solvents. The data shows that the polarizable solvents like benzene or dichloromethane yield more red shifted species are compared to methanol and ethyl ether. Other studies also suggest that the presence of well-ordered water molecules in the binding pocket can increase the pK_a of the PSB.

I-2-1: Wavelength Regulation Due to the Conformation of the Chromophore

The conformation of the chromophore has been suggested to have an important role in the wavelength regulation. Based on this theory the steric factor in the chromophore may lead to twisting around the single chromophore's and therefore

change the level of the delocalization of the π-electrons or p-orbital overlap. Changing the delocalization of the π-electrons of the polyene backbone leads to changes in the degree of cationic conjugation along the polyene chain. In other word changing the level of conjugation may change the absorption of the chromophore. Although retinal would rather adopt a more planar conformation to maximize the overlap of the p-orbital, 11-*cis*-retinal is not a planer molecule. The ionone ring has a 6-s-*cis* or 6-s-*trans* conformation since steric interactions between C5-Me and C8-H as well as the C10-H and C13-Me result in twisting of the C12-C13 single bond.[60] Since there is no crystal structure of any of the color pigments available so far, there is no solid data to prove this theory.[59] One of the major conformational changes is believed to happen is twisting around the single bond C6-C7. The crystal structure of the bovine rhodopsin illustrates that the conformation of the chromophore is 6-s-*cis.* This twisting may happen because the C5-methyll and C8-H have steric interaction and therefore the ionone ring rotates around the C6-C7 single bond. The result of this twisting is less conjugation than 6-s-*trans*. As mentioned above the conformation of the chromophore in bovine rhodopsin is 6-s-*cis* while in microbial rhodopsin is 6-s-*trans*. This theory was proposed to explain why the microbial rhodopsin is more red shifted than bovine rhodopsin.[32]

Other than that the planarity of the chromophore is believed to be important for the wavelength tuning. When the ionone ring of the chromophore locked in the planar position the most red shifted pigment is to be obtained.

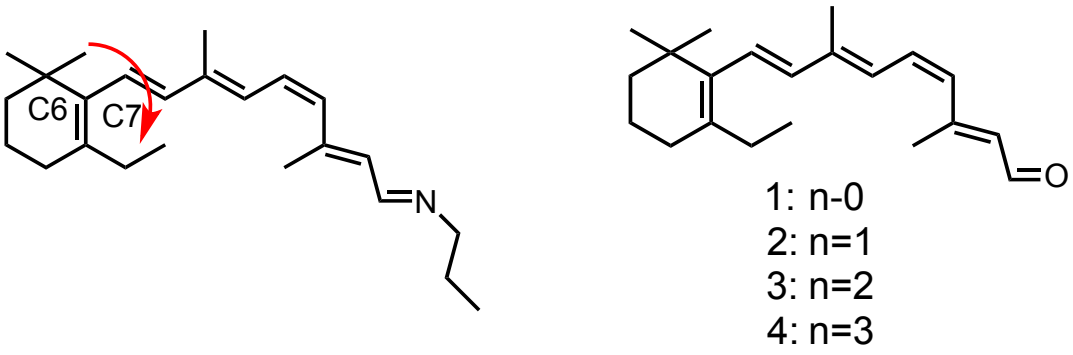


Figure I-15: The torsion angle around C6-C7 single bond of retinol-PSB

Table I-2: Characterization of different ring-locked retinal analogues.

Chromophore 11-*cis*-retinal	Absorbance (nm)		
	PSB-aldehyde	bound to rhodopsin	free
	500	444	380
1	506	539	422
2	496	546	416
3	457	503	386
4	440	483	374

As it is shown in table I-2 retinal analogs 1 and 2 are the most planar ones as compared to the 3 and 4 and therefore they are the most red shifted aldehyde in the solution. Garavelli and his group showed that 6-s-*trans*-conformer of all-trans-retinal-PSB in vacuum absorbs at 610 nm while 6-s-*cis* conformer absorbs at 530 nm. This

dramatic different in the absorbance is due to the highly twisting along C6-C7 single bond.[32] Olivucci and Garavelli computed that in gas phase 11-*cis*-retinal-PSB absorbs at 545 nm and the ionone ring of the chromophore is 68° twisted from the flat plane.[61] In their studies they also showed that in 6-s-*cis* conformer if the rotation angle (θ) between the ionone ring and the plane of the polyene chains is 90° and the conjugation of π-electrons of the polyene completely breaks from π-electron of the ionone ring, the chromophore absorbs at 535 nm.

I-2-2: Computational Studies on the Wavelength Regulation

By the rapid development in computational studies the computational simulation methods used widely in the modern studies. One of the calculations have been done by Olivucci and Garavelli on the vertical excitation energy of the S0→S1 vertical excitation of 11-cis-retinal in gas phase and they predicated the absorption with different torsion angles between the ionone ring and the polyene chain of chromophore. [62] They calculated that when the torsion angle between the β-ionone ring and the polyene chain is zero and the double bond of the ring is on the same plane as the backbone double bonds the maximum conjugation is obtained and the most red shifted absorbance is resulted, while when this angle is changed to 90° and the double bond is completely out of the plane the most blue shifted pigment is obtained (Figure I-14).

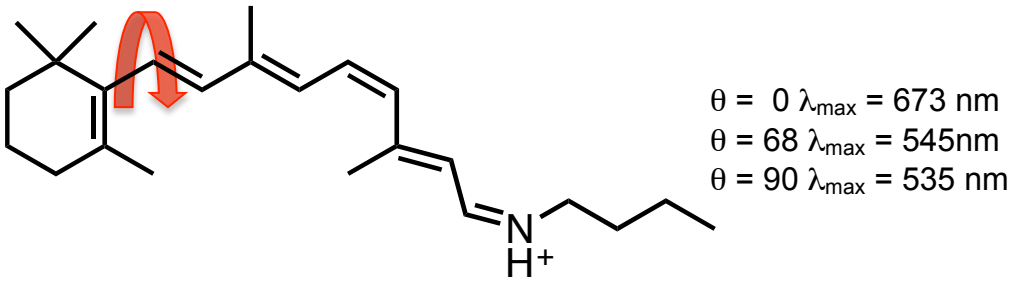


Figure I-16: The computational analysis of the gas phase visual chromophore.

Based on this data they proposed that in red rhodopsin the chromophore is more planar and has the most conjugated π-system, while blue rhodopsin is expected to have the most twisting around the C6-C7 single bond and therefore ionone ring double bond is completely out of conjugation.

Another calculation that has been done was studies on the presence of the counter anion in the structure of the rhodopsin. As it was mentioned above, the crystal structure of rhodopsin illustrate that the positive charge of the PSB is stabilized by the presence of a negatively charged residue (Glu113). Buss and his coworkers showed that by introduction of the counter anion the absorption of the retinal-PSB in vacuum from 610 nm decreases to 486 nm.[63] The different amino acids around the chromophore lead to slightly bathochromic shifts due to their charge and dipolar, while the conformational changes in the structure of the chromophore do not show any significant role in tuning the wavelength. They suggest that tuning of wavelength is dictated by the counter anion and the rest of the amino acids inside the binding pocket do not play any significant role.

However mutagenesis of the counter anion leads to just 30 nm more red shift and does not cover the whole spectrum while the pK_a drops significantly by this mutation.

I-3: Using Model Studies to Understand the Mechanism of Wavelength Regulation and Rhodopsin System

As it was explained above different theories were suggested to explain the mechanism of the wavelength regulation. However still no single conclusion has been obtained from these theories because unfortunately so far no crystal structure of any of the color pigments are available. Lack of direct information from the crystal structures made the scientists to study on the wavelength regulation with the indirect methods. Not only site mutagenesis but also biophysical methods were used to study on the rhodopsin system. Raman spectroscopy was used to study the conformational changes of retinal and the interaction of the chromophore with the residues in the binding pocket.

I-3-1: Modeling the Color Pigments Based on the Structure of Bovine Rhodopsin

The X-ray crystal structure of bovine rhodopsin was a revolution in the study on the wavelength regulation. This crystal structure provides the structural information that later on was used in the other theories which tries to explain the tuning the spectrum in human eye. Since rhodopsin is a membrane protein, therefore expression, purification and crystallization of this protein are very challenging. Therefore more flexible strategies such as computational studies and also using a small-reengineered protein as the

mimic of rhodopsin were used. Using a small protein, which is reengineered as the binding pocket of the rhodopsin, can provide us with a better platform to find out the mechanism of wavelength regulation.

Due to difficulties and challenges of working with rhodopsin, over the past few decades many different computational and modeling have been done to understand the mechanism of spectrum tuning which takes place in each of the color pigments. Although the crystal structures of the cone rhodopsins are the ultimate answer regarding the mechanism of wavelength regulation in human eye. However, due to lack of the real crystal structure we believe that a protein mimic is would be the most realistic system to study the color pigments. Professor Babak Borhan in Department of chemistry at Michigan State University and his group initiated the idea to use a soluble and relatively small protein and reengineer the binding pocket to make it a mimic of the binding pocket of the visual rhodopsin. Such a realistic model might provide more information and better platform to study the spectral tuning.

The electronic profile of retinal-PSB shows that at the ground state the positive charge of the iminium is distributed on the different carbons of the chromophore. Generally the carbon, which is directly attached to the nitrogen of the iminium, carries the highest concentration of the positive charge and by increasing the distance from the iminium the amount of the positive charge decreases. However the profile of the charge distribution in the excited state of all-trans-retinylidene is very different form the ground state. In the excited state the charged is transferred to the end of the chromophore and carbon 5 to 9 carried the most amount of the positive charge.[65]

Another electric parameter in retinal-PSB is the alternating single/double bond character. Buss shows in his calculation on the bond distances of bond form of all-trans-retinal in PSB and SB form in solution and in rhodopsin. The data clearly shows that in the most blue shifted form has the most obvious distances in the single and double bonds. By protonation of the Schiff base and red shifting from 360 nm to 440 nm these alternations of the single/double bond distances become less significant. The result of this effect is charge delocalization on the chromophore through resonance structure (Figure I-15). The data indicate that the more stable and smaller alternation differences between the bond length of single and double bonds lead to more delocalization of the positive charge of the PSB and more red shifting will obtain. The cyanine dyes can be used as an example. This fact can explain why a cyanine dye is more red shifted than retinal, although they both have the same number of double bonds. Since in the cyanine-PSB the positive charge is fully delocalized along the whole conjugated system. Based on these calculations in order to achieve the red shift, the positive charge needs to be delocalized on the polyene and the resonance structure needs to be stable through favorable electrostatic interactions with the residues in the binding pocket. The interactions of chromophore-protein in rhodopsin were studied by different groups. Matthias, Ullmann and coworkers calculated the electrostatic potential on the surface of the chromophore and the protein environment in sensory rhodopsin II (500 nm), bacteriorhodopsin (560 nm) and halorhodopsin n (570 nm). For this calculation the high resolution crystal structure is necessary, since the conformation of each residue specially the ones that have interaction with the chromophore is important.[65]

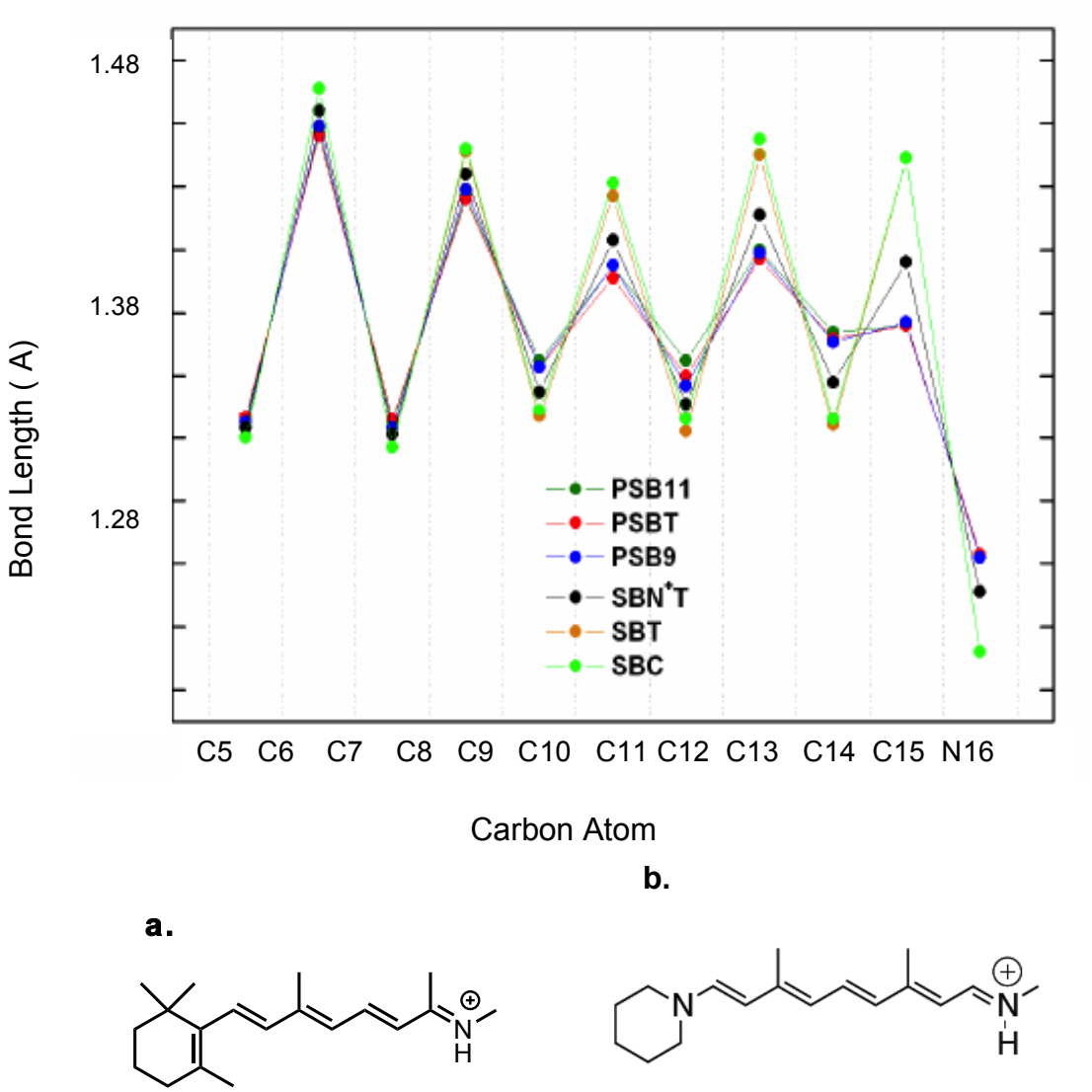


Figure I-17: The bond length of the polyene chain of chromophore.[64]
a. Retinal. **b**. Cyanine dye

Dr. Lee in Professor Borhan's lab performed the same calculation on the theoretical models of the three different color pigments.[66] The results of this calculation are summarized in Figure 1-17. The chromophore is divided into three segments; the

qualitative average score for each segment represents the overall electrostatic potentials that lead to the wavelength regulation of each opsin.[66] (calculations of blue, green, and red are based on available homology models) projected on the Van der Waals surface of retinal (the electrostatic potential of the retinylidene chromophore was set to zero as to only illustrate the electrostatic contribution of the protein).

As it is illustrated in this figure in blue rhodopsin the negative electrostatic potential is mostly concentrated on the SB region while in the red rhodopsin the negative charge is more transferred on the ionone ring area. Based on this calculation, since blue rhodopsin has more negative electrostatic potential around the iminium part therefore the positive charge of the PSB is more stabilized on that region. On the other hand in red rhodopsin the positive charge is more delocalized. This electrostatic calculation is used as the platform of the series of mutants that Dr. Wenjing Wang in Prof. Borhan's group generated on the protein mimic. In this study human Cellular Retinol Binding Protein II (hCRBPII) has been used as the platform of the rhodopsin mimic.

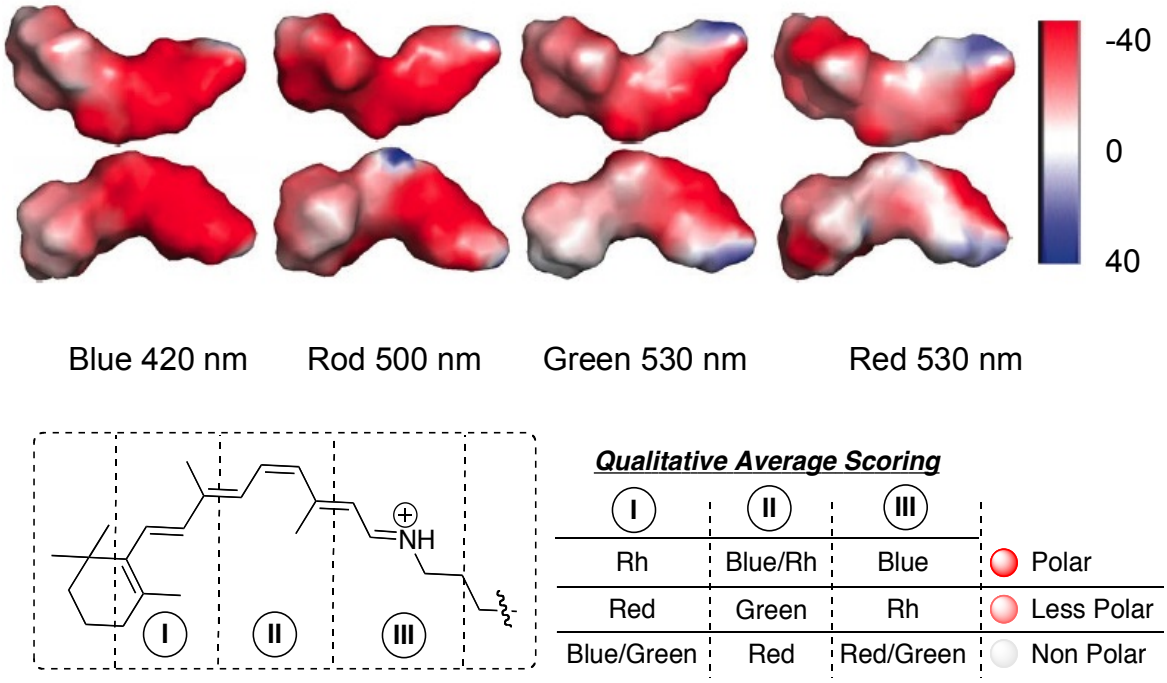


Figure I-18: Electrostatic potential calculation. Electrostatic potential calculation of blue, rod, green, and red opsin.

I-3-2: Characteristic of a Potential Protein To Be Used As a Rhodopsin Mimic

The protein that is used for this purpose must have several characteristics. (1) It must be very soluble and should not aggregate during expression and purification and therefore easy to handle and study. Additionally it should be expressed in an E.*coli* expression system. (2) Since during this study several site-mutations need to be done on the protein, therefore this protein should have a robust structure and tolerance to the mutations. (3) Also since during this study the crystal structures of different mutants are necessary, the mimic protein must be relatively easy to crystallize and easy to determine its structure. (4) The candidate protein as rhodopsin mimic also must have an embedded binding pocket. This feature of the protein makes the changes in the spectral tuning just by applying the different mutation not the effect of the solvents. The reengineering of the rhodopsin mimic is based on the electrostatic potential calculation, which was performed in Professor Borhan's group.

I-3-3: The First Generation of Rhodopsin Mimic

Human Cellular Retinoic Acid Binding Protein (hCRABPII) from the family of lipid binding proteins (iLBP) with 137 amino acid residue was initially used as the template for the rhodopsin mimic.[67] Since 11-*cis*-retinal is very light sensitive compound therefore in this study all-*trans*-retinal has been used, since it is easier to handle. As it was mentioned above the chromophore is covalently bond to the protein via an active lysine. For formation the first generation of rhodopsin mimics R132 position was mutated to a

lysine. In order to have a successful nucleophilic attack to the carbonyl group of the retinal, the Bürgi-Dunitz trajectory is important and the amine group of lysine must attack to the aldehyde plane, with an optimal angle of 107°.[68]

The crystal structure of this protein is shown in Figure 1-18.

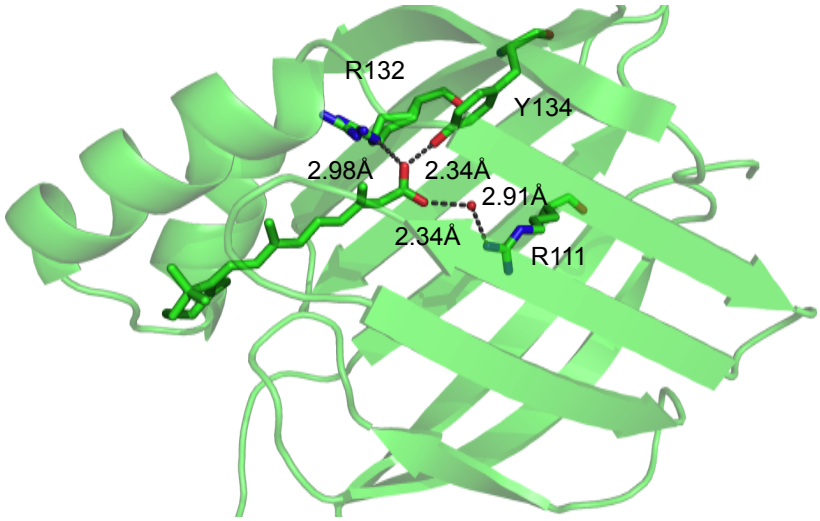


Figure I-19: The Crystal structure of wt CRABPII bound with retinoic acid with highlighted interactions of the chromophore carboxylic acid and CRABPII residues (PDB entry 2FR3)

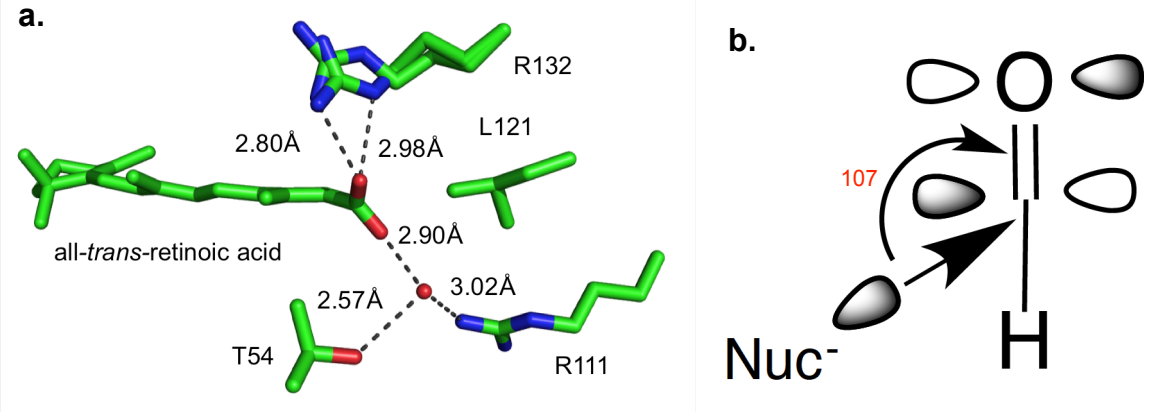


Figure I-20 a. The detail of hydrogen bonds between CRABPII and all-*trans*-retinoic acid. **b**. Illustration of the Bürgi-Dunitz trajectory.

R111L mutation is introduced to remove the observed water molecules, since this water molecule locks the aldehyde in an unfavored position. However no PSB formation have been observed with R132K:R111L mutation at the physiological pH=7.3 due to the low pK_a of the PSB of the retinal (<6.5). As it was mentioned above the pK_a of the bovine rhodopsin is higher than 16 and the crystal structure of this protein pointed to that Glu113 as a counter anion that has an important role to increasing the pK_a. Inspection of the crystal structure of CRABPII protein shows that Leu121 is a favored position for the introduction of counter anion (L121E mutation). Introducing this mutation could increase the pK_a of the PSB to higher than biological pK_a. The final pK_a of R132K:R111L:L121E was 8.7.[69,70]

However mutagenesis studies and switching the polarity of the amino acid residue (R59) which is located at the “mouth” of the protein binding pocket do not have any effect on the wavelength tuning. The crystallographic data of rhodopsin showed that the chromophore is deeply buried in the binding pocket, In CRABPII the ionone ring of the chromophore sticks out from the binding pocket and is exposed to the bulk solution and the effect of the polarity of the amino acids (specially in zone I) is buffered by the media (table I-3). Since water in the bulk solution has high polarity, it can make retinal inert to the changes of the polarity of the protein. To investigate weather exposure of the chromophore to the solution neutralized the effect of the amino acid residues, two different approaches have been used:

1) Using a shorter chromophore analogue which has the same structure but shorter, so it can be embedded in the binding pocket of the CRABPII protein.

2) Using another protein mimic, which has deeper binding pocket and reengineer it the same as CRABPII.

The first approach has been carried by Dr. Lee in Professor Borhan's Lab and the crystallographic data were collected by Dr. Jia in Professor Geiger's lab.[69,70]

Table I-3: CRABPII mutants with all-*trans*-retinal and all-*trans*-C15

CRBAP II mutants	λ_{max} with retinal	λ_{max} with C15
R132K-R111L-L121E-R59	449	n.d.[a]
R132K-R111L-L121E-R59E	450	424
R132K-R111L-L121E-R59Q	444	413
R132K-R111L-L121E-R59L	443	391
R132K-R111L-L121E-R59W	442	404

[a] due to the low pK_a of PSB and therefore no PSB peak has been observed.

The comparison of the mutagenesis studies on both all-*trans*-retinal and all-*trans*-C15 is listed in Table I-3. As it is shown in the table by changing the polarity of the R59 position no spectral tuning has been observed when the protein is bound to retinal, however with the C15 analogue mutation of the same position leads to modulation of wavelength. This data proves that for tuning the wavelength the embedding binding pocket is essential and the chromophore need to be deep in the binding pocket and not exposed to the bulk solution. Based on these data the second generation of rhodopsin mimics with deeper binding pockets were reengineered.

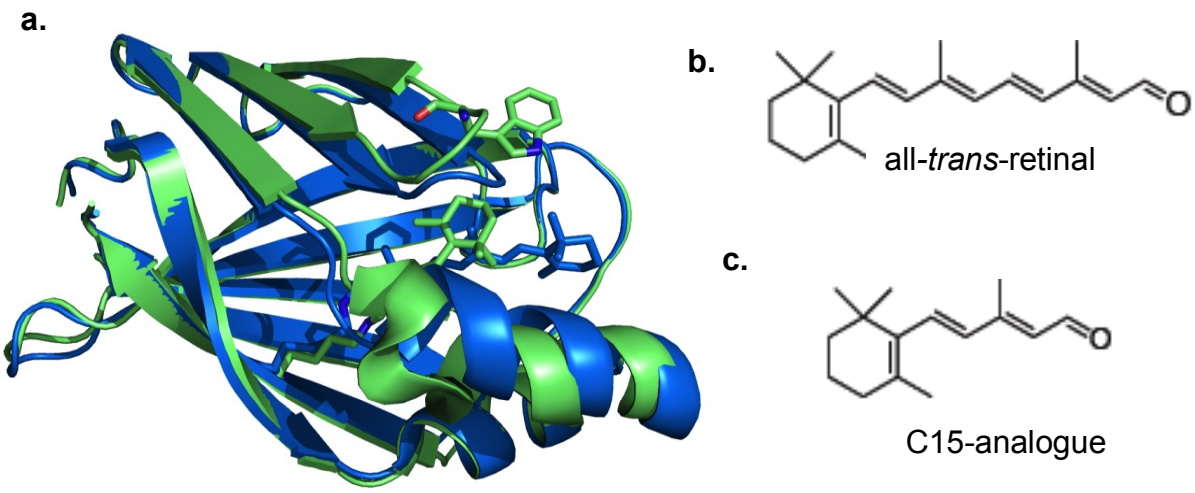


Figure I-21: a. The crystal structure overlaid of R132K:R111L:L121E (KLE, PDB entry 2G7B) bound with retinal, (blue) and structure of KLE-R59W (PDB entry 3F8A) bound to C15, with R59W position is highlighted. The overlaid structure illustrates that C15 is fully embedded within the protein binding pocket, while retinal (blue) is exposed **b**. Chemical structures of all-*trans*-retinal. **c**. all-*trans* C15 analogue.

I-3-4: The Second Generation Rhodopsin Mimic (Human Cellular Retinol Binding Protein II)

Since the data from C15 analogue illustrate the fact the embedded and covered binding pocket is an essential factor for a protein, which is used as mimic of rhodopsin, we switched to another protein from the same family of proteins, Cellular retinol binding protein (hCRBPII).[71] This is also relatively small and very soluble protein with 133 amino acid residues, which is responsible to carry all-*trans*-retinol inside the cell. The crystal structure of this protein shows that the chromophore is not exposed to the bulk solution. Since these two proteins (hCRBPII and CRABPII) have different functions in body

therefore the binding site of them are different. HCRBPII is binding to all-trans-retinal and protecting that against oxidation and other environmental damages, therefore the binding pocket is very preserved, while CRABPII is responsible for carrying the all-trans-retinoic acid and therefore that the biding pocket is more open and accessible. The sequence identity between hCRBPII and CRABPII is only 35%, however they share the same structure of the iLBPs, which is 10 β-stranded sheets with two short α-helixes that cover the binding pocket while the ligand binds deeply in the binding pocket. The crystal structure and function of CRBPII is explained in the next chapter.

The comparison of the CRABPII and hCRBPII shows that the ligand (all-*trans*-retinal) is almost 5 Å deeper in the binding pocket. Figure I-20, shows the superimposed structure of CRABPII and hCRBPII. As it is clear in this figure (b, c) the ligand is more covered in CRBPII and not accessible to the bulk solution.

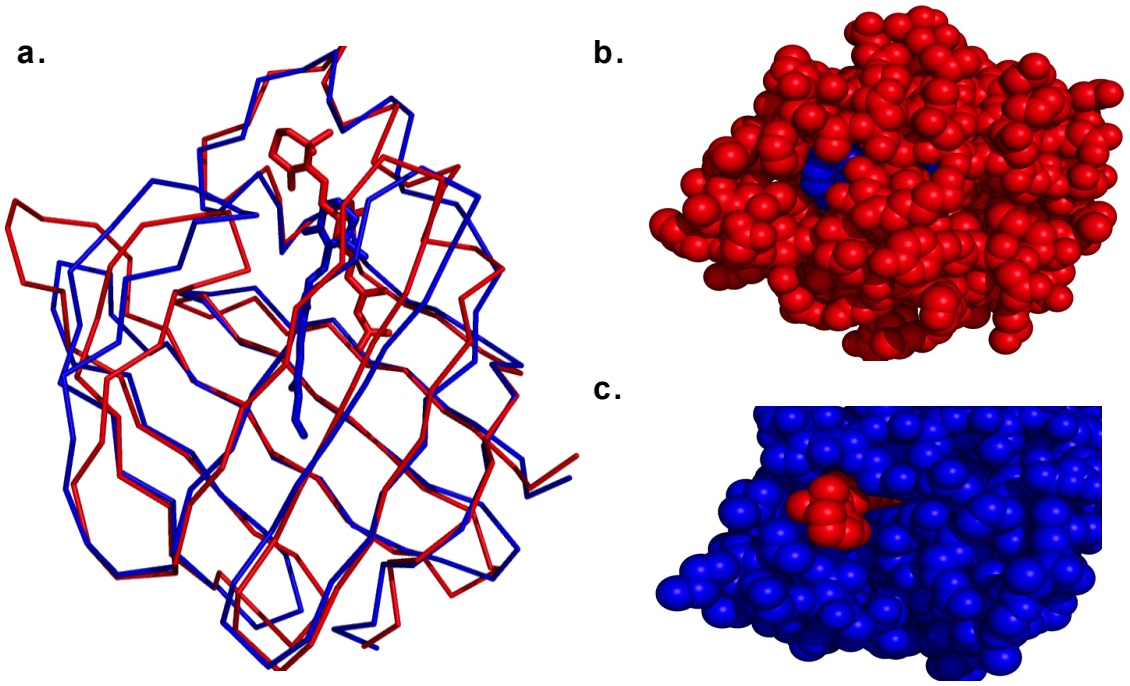


Figure I-22: a. The superimposed crystals structure of hCRBPII bound to all-*trans*-retinal (blue) and CRABPII bound to all-*trans*-retinoic acid (red). **b.** The space model of hCRBPII (the exposed part of the chromophore is blue). **c.** The space model of hCRBPII (the exposed part of the chromophore is red).

REFERENCES

REFERENCES

1. Shichida, Y.; Matsuyama, T., Evolution of opsins and phototransduction. Phil. *Trans. R. Soc. B*, **2009**, 364 (1531), 2881-2895.

2. Shichi, H*. Biochemistry of Vision, Academic Press,* Inc., London, UK, **1983**.

3. Frank, R. N., Cavanagh, H. D. & Kenyon, K. R. Light-Stimulated Photophosphorylation of Bovine Visual Pigments by Adenosine-Transphosphate. *Journal of Biochemical Chemistry*, **1973**, 248, 596-609.

4. Padgham, C. A& Saunders, J. E. *The Perception of Light and Colour*, G. Bell & Sons Ltd, London, UK, **1975**.

5. Molday, R.S. Photoreceptor membrane protein, Phototransduction and retinal degenerative disease – The Friedenwald lecture*. Investigative Ophthalmology & Visual Science*, **1998**, 39, 2493-2513.

6. Streyler, L. The Molecules of Visual Excitation. Scientific American, **1987**, 257, 42-50.

7. Ripp, H.& Weale, R. A. The Visual Photoreceptors, The Eye (Davson, H., Ed.) 2.6 vols, Academic Press, Inc., **1969**, New York, NY.

8. The Schanpf, J. Rhodopsin-Structure and Function and Genetics. Biochemistry, **1992**, 4923-4931.

9. Wald, G. Cartenoids and the visual cycle. *Journal of General physiology*, **1935**, 19, 351-371.

10. Wald, G. Pigments of the retinal I. The bull frog*. Journal of General Physiology,* **1936**, 19, 781-795.

11. Stryer, L.; Hurley, J. B.; Fung, B. K. K., Transducin - An amplifier protein in Vision. Trends Biochem. Sci. **1981**, 6 (9), 245-247

12. Gilman, A. G., G-Proteins - Transducers of receptor-generated signals. Annu. Rev. Biochem. **1987**, 56, 615-649.

13. Hofmann, K. P.; Scheerer, P.; Hildebrand, P. W.; Choe, H. W.; Park, J. H.; Heck, M.; Ernst, O. P., A G protein-coupled receptor at work:

rhodopsin model. *Trends Biochem. Sci*. **2009**, 34 (11), 540-552

14. Pugh, E. N. & Lamb, T. D. Amplification and Kinetics of the Activation Step in Phototransduction, Biochimica, Et, Biophysica Acta, **1993**, 1141, 111-149

15. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M., Crystal structure of rhodopsin: a g protein-coupled receptor. Science **2000**, 289 (5480), 739-745.

16. (A) Doukas, A. G.; Junnarkar, M. R.; Alfano, R. R.; Callender, R. H.; Kakitani, T.; Honig, B., Fluorescence quantum yield of visual pigments - evidence for subpicosecond isomerization rates. P. Natl. Acad. Sci-Biol.**1984**, 81 (15), 4790-4794; (B) Buchert, J.; Stefancic, V.; Doukas, A.G.; Alfano, R. R.; Callender, R. H.; Pande, J.; Akita, H.; Baloghnair, V.; Nakanishi, K., Picosecond absorption and fluorescence kinetic-studies of bovine rhodopsin with a fixed 11-ene. Biophys. J. **1983**, 41 (2), A338.

17. a. Winter, N. S.; Bratt, J. M.; Banaszak, L. J., Crystal-structures of holo and apo-cellular retinol-binding protein-II. J. Mol. Biol. **1993,** 230 (4), 1247-1259; (b) Tarter, M.; Capaldi, S.; Carrizo, M. E.; Ambrosi, E.; Perduca, M.; Monaco, H. L., Crystal structure of human cellular retinol-binding protein II to 1.2 angstrom resolution. Proteins **2008**, 70 (4), 1626-1630.

18. Sakmar, T, P. Structure of rhodopsin and seven-helical receptors: the same and not the same. *Current Opinion in Cell Biology*, **2002,** 14, 189-195.

19. Cahbre, M. Breton, J., Michelvillaz, M., Saibil, H., Linear Dichroism Studies in the Visible, UV and IR on Oriented Rod Suspension*, Method In Enzymology*,**1982**, 81, 605-616.

20. Nathane, J. Determinants of Visual Pigment Absorbance-Role of Charged Amino-Acids In the Putative Transmembrane Segment*, Biochemistry*, **1990**, 29, 937-942.

21. Choe, H.-W.; Kim, Y. J.; Park, J. H.; Morizumi, T.; Pai, E. F.; Krauss, N.; Hofmann, K. P.; Scheerer, P.; Ernst, O. P., Crystal structure of metarhodopsin II. Nature, **2011**, 471 (7340), 651-U137.

22. Standfuss, J.; Edwards, P. C.; D'antona, A.; Fransen, M.; Xie, G.; Oprian, D. D.; Schertler, G. F. X., The structural basis of agonist-induced activation in constitutively active rhodopsin. Nature **2011**, 471 (7340), 656- 660.

23. Nakamichi, H.; Okada, T., Crystallographic analysis of primary visual photochemistry. Angew. Chem. Int. Ed. Engl. **2006**, 45 (26), 4270-4273.

24. Leibovic, K. N. Vertebrate Photoreceptors. In Science of Vision, pp. 16-52. Springer-Verlag, **1990**, New York, NY.

25. Stenkamp, R. E., Teller, D. C. & Palczewski, K. Rhodopsin: A structural for primer for G-protein couple receptors, *Archive Der Pharmaze*, **2005**, 338, 209-216.

26. (A) Vuong, T. M.; Chabre, M.; Stryer, L., Millisecond activation of transducin in the cyclic-nucleotide cascade of vision. Nature **1984**, 311 (5987),659-661; (B) Guy, P. M.; Koland, J. G.; Cerione, R. A., Rhodopsin-stimulated activation deactivation cycle of transducin - kinetics of the intrinsic fluorescence response of the alpha-subunit. Biochemistry **1990**, 29 (30), 6954-6964.

27. Gilman, A. G., G-Proteins - Transducers of receptor-generated signals. Annu. Rev. Biochem. **1987**, 56, 615-649.

28. Filipek, S., Stenkamp, R.E., Teller, D. C. & Palcezewki, K. G protein - couple receptor rhodopsin: A prospectus. Annual Review of Physiology **2003**, 65, 851-879.

29. Stryer, L. Molecular-Basis of Visual Excitation, *Cold Spring Harbor, Symphosia on Quantitative Biology*, **1988**, 53, 283-294.

30. Blatz, P. E.; Baumgart.N; Balasubr.V; Balasubr.P; Stedman, E., Wavelength regulation in visual pigment chromophore - large induced bathocromic shifts in retinol and related polyenes. Photochem. Photobiol. **1971**,14 (4), 531-549.

31. Andersen, L. H.; Nielsen, I. B.; Kristensen, M. B.; El Ghazaly, M. O. A.; Haacke, S.; Nielsen, M. B.; Petersen, M. A., Absorption of schiff-base retinal chromophores in vacuo. J. Am. Chem. Soc**. 2005,** 127 (35), 12347-12350.

32. Rajput, J.; Rahbek, D. B.; Andersen, L. H.; Hirshfeld, A.; Sheves, M.; Altoe, P.; Orlandi, G.; Garavelli, M., Probing and modeling the absorption of retinal protein chromophores in vacuo. Angew. Chem. Int. Ed. Engl. **2010**, 49 (10), 1790-1793.

33. Motto, M. G.; Sheves, M.; Tsujimoto, K.; Baloghnair, V.; Nakanishi, K., Opsin shifts in bovine rhodopsin and bacteriorhodopsin - comparison of

2external point-charge models. *J. Am. Chem. Soc*. **1980,** 102 (27), 7947-7949.

34. Erickson, J. O. & Blatz, P. E. N-Retinylidene-1-Amino-2-Propanol-a Schiff Base Analog for Rhodopsin, Vision Research, **1968**, 8, 1397-1357.

35. Kochendoerfer, G. G., Lin, S. W., Sakmar, T, P. & Mathies, R. A. How color visual pigments are tuned. Trends in Biochemical Science, **1999**, 24, 300-305.

36. Wald, G., Molecular Basis of Visual Excitation, *Science*, **1968**, 162, 230-239.

37. Wald, G., Molecular Basis of Visual Excitation, *Nature*, **1968**, 219, 800-807.

38. Neitz, M.; Neitz, J.; Jacobs, G. H., Spectral tuning of pigments underlying red-green color-vision. *Science*, **1991**, 252 (5008), 971-974.

39. Irving, C. S.; Byers, G. W.; Leermake.Pa, Spectroscopic model for visual pigments - influence of microenvironmental polarizability *Biochemistry*, **1970**, 9 (4), 858-864.

40. Albeck, A., Livnah, N., Gottlieb, H., & Sheves, M. C-13 NMR-Studies of Model Compounds for Bacteriorhodopsin-Factors Affecting the Retinal Chromophore Chemical-Shifts and Absorption Maximum. *Journal of the American Chemical Society,* **1992**, 114, 2400-2411.

41. Okada, T., Sugihara, M., Bondar, A. N., Elstner, M., Entel, P., & Buss, V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 angstrom crystal structure. *Journal of Molecular Biology*, **2004**, 342, 571-583.

42. Sakmar, T. P. Rhodopsin: A prototypical G protein coupled receptor. *Progress in Nucleic Acid Research and Molecular Biology***, 1998**, Vol 59 1-34.

43. Schertler, G. F. X & Hargrave, P. A, Projection Structure of Frog Rhodopsin in 2 Crystal Forms*, Proceeding of National Academy of Sciences of the States of America*, **1995**, 92, 11578-11582.

44. Unger, V. M., Hargrave, P. A., Baldwin, J. M. & Schertler, G. F. X, Arrangement of rhodopsin transmembrane alpha-helices. *Nature*, **1997**, 389, 203-206.

45. Nathans, J.; Thomas, D.; Hogness, D. S., Molecular-genetics of

human color-vision - the genes encoding blue, green, and red pigments. *Science*, **1986**, 232 (4747), 193-202.

46. Neitz, M.; Neitz, J.; Jacobs, G. H., Spectral tuning of pigments underlying red-green color-vision. *Science*,**1991**, 252 (5008), 971-974.

47. Fasick, J. I., Lee, N. and Oprian, D. D. Spectral tuning in the human blue cone pigment. *Biochemistry*, **1999,** 38, 11593-11596.

48. Asenjo, A. B.; Rim, J.; Oprian, D. D., Molecular determinants of human red/green color discrimination. *Neuron*, **1994**, 12 (5), 1131-1138.

49. Lin, S. W.; Kochendoerfer, G. G.; Carroll, H. S.; Wang, D.; Mathies, R. A.; Sakmar, T. P., Mechanisms of spectral tuning in blue cone visual pigments- visible and raman spectroscopy of blue-shifted rhodopsin mutants. *J. Biol. Chem***. 1998**, 273 (38), 24583-24591.

50. Zhukovsky, E. A.; Oprian, D. D., Effect of carboxylic-acid sidechains on the absorption maximum of visual pigments. Science 1989, 246 (4932), 928-930.

51. Nakayama, T. A.; Zhang, W. M.; Cowan, A.; Kung, M., Mutagenesis studies of human red opsin: trp-281 is essential for proper folding and protein retinal interactions. *Biochemistry,* **1998**, 37 (50), 17487-17494.

52. Sakmar, T. P.; Franke, R. R.; Khorana, H. G., The role of the retinylidene schiff-base counterion in rhodopsin in determining wavelength absorbency and schiff-base pK_a. *P. Natl. Acad. Sci. Usa*, **1991**, 88 (8), 3079-3083.

53. Wang, Z,Y., Asenjo, A, B,& Oprian, D. D. Identification of the Cl^- binding site in the human red and green color vision pigments, *Biochemistry,* **1993**, 32, 2125-2130.

54. Mathies, R.; Stryer, L., Retinal has a highly dipolar vertically excited singlet-state - implications for vision. *P. Natl. Acad. Sci. Usa,* **1976**, 73 (7), 2169- 2173.

55. Livnah, N.; Sheves, M., Model compounds can mimic spectroscopic properties of bovine rhodopsin. *J. Am. Chem. Soc***. 1993**, 115 (1), 351-353.

56. Motto, M. G.; Sheves, M.; Tsujimoto, K.; Baloghnair, V.; Nakanishi, K., Opsin shifts in bovine rhodopsin and bacteriorhodopsin - comparison of 2 external point-charge models. *J. Am. Chem. Soc*. **1980**, 102 (27), 7947-7949.

57. Honig, B.; Dinur, U.; Nakanishi, K.; Baloghnair, V.; Gawinowicz, M. A.; Arnaboldi, M.; Motto, M. G., External point-charge model for wavelength regulation in visual pigments. *J. Am. Chem. Soc*. **1979**, 101 (23), 7084-7086.

58. Beppu, Y.; Kakitani, T., Theoretical-study of color control mechanism in retinal proteins .1. role of the tryptophan residue, tyrosine residue and water molecule. *Photochem. Photobiol*. **1994**, 59 (6), 660-669.

59. Wada, A.; Tsutsumi, M.; Inatomi, Y.; Imai, H.; Shichida, Y.; Ito, M., Retinoids and related compounds. part 26. synthesis of (11 z)-8,18- propano- and methano-retinals and conformational study of the rhodopsin chromophore. *J.Chem. Soc. Perk. T. 1* **2001**, (19), 2430-2439.

60. Tan, Q., Lou, J. H. Borhan, B., Karnaukhova, E., Berova, N. & Nakanishi, K. Absolute sense of twist of rhe C12-C13 bond of the retinal chromophore In bovine rhodopsin based on excitation CD spectra of dihydroretinal Analogues, *Angewandte Chemis-International Edition in English*, **1997**, 36, 2089-2093.

61. Cembran, A.; Gonzalez-Luque, R.; Altoe, P.; Merchan, M.; Bernardi, F.; Olivucci, M.; Garavelli, M., Structure, spectroscopy, and spectral tuning of the gas-phase retinal chromophore: The beta-ionone "handle" and alkyl group effect. *J. Phys. Chem. A* , **2005**, 109 (29), 6597-6605.

62. Sakurai, M.; Sakata, K.; Saito, S.; Nakajima, S.; Inoue, Y., Decisive role of electronic polarization of the protein environment in determining the absorption maximum of halorhodopsin. *J. Am. Chem. Soc*. **2003**, 125 (10), 3108-3112.

63. Sekharan, S.; Sugihara, M.; Buss, V., Origin of spectral tuning in rhodopsin - it is not the binding pocket. *Angew. Chem. Int. Ed. Engl***. 2007**, 46 (1-2), 269-271.

64. Sekharan, S.; Weingart, O.; Buss, V., Ground and excited states of retinal schiff base chromophores by multiconfigurational perturbation theory.*Biophys. J***. 2006**, 91 (1), L7-L9.

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

66. Lee, K. S. S. Ph.D Thesis, Michigan State University, East Lansing, MI, 2011.

67. (a) Vaezeslami, S.; Mathes, E.; Vasilelou, C.; Borhan, B.; Geiger, J. H., The structure of apo-wild-type cellular retinoic acid binding protein II at 1.4Å

and its relationship to ligand binding and nuclear translocation. J. Mol. Biol.2006, 363 (3), 687-701; (b) Kleywegt, G. J.; Bergfors, T.; Senn, H.; Lemotte, P.; Gsell, B.; Shudo, K.; Jones, T. A., Crystal-structures of cellular retinoic acidbinding protein-i and protein-ii in complex with all-trans-retinoic acid and a synthetic retinoid. Structure 1994, 2 (12), 1241-1258.

68. Burgi, H. B.; Dunitz, J. D., From crystal statics to chemicaldynamics. *Accounts Chem. Res.* **1983,** 16 (5), 153-161.

69. Vasileiou, C.; Vaezeslami, S.; Crist, R. M.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B., Protein design: Reengineering Cellular Retinoic Acid Binding Protein II into a rhodopsin protein mimic. *J. Am. Chem. Soc*. **2007**, 129 (19), 6140-6148.

70. (a) Crist, R. M.; Vasileiou, C.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B., Engineering a rhodopsin protein mimic. *J. Am. Chem. Soc*. **2006**, 128(14), 4522-4523; (b) Vasileiou, C.; Wang, W. J.; Jia, X. F.; Lee, K. S. S.; Watson, C. T.; Geiger, J. H.; Borhan, B., Elucidating the exact role of engineered CRABPII residues for the formation of a retinal protonated Schiff base. *Proteins*, **2009**, 77 (4), 812-822.

71. (a) Winter, N. S.; Bratt, J. M.; Banaszak, L. J., Crystal-structures of holo and apo-cellular retinol-binding protein-II. *J. Mol. Biol*. **1993,** 230 (4), 1247-1259; (b) Tarter, M.; Capaldi, S.; Carrizo, M. E.; Ambrosi, E.; Perduca, M.; Monaco, H. L., Crystal structure of human cellular retinol-binding protein II to 1.2 angstrom resolution. *Proteins,`* **2008**, 70 (4), 1626 1630.

72. Leibovic, K. N. Vertebrate Rhotoreceptors in *Science of Vision*, pp. 16-52, **1990**. Spring-Verlad, New York, NY.

CHAPTER II

The Crystal Structure of Holo-Wild-Type human Cellular Retinol Binding protein II (hCRBPII) bound with Retinol and Retinal

II-1: Structure Determination from X-ray Diffraction

One single crystal is made of huge number of molecules in the highly ordered and repetitive pattern. When a crystal is exposed to X-ray beam the scattered waves from the imaginary plane of the crystals will only take place when the path difference between the rays are equal to the integral number of wavelength. This phenomena is known as Bragg's law:

$$2d\,\mathrm{Sin}\theta = n\lambda \qquad (2.1)$$

When the d is the distance between imaginary reflected planes and θ is the angle between the incident and reflected beam (λ Is the wavelength of the incident beam). The process of the reflection is incident and reflection of the rays than make an angle θ with the fixed crystal plane. Since the electrons are the scatters in therefore each atom has the different effectiveness as the scatters. The X-ray diffraction is widely used for biological macromolecules such as proteins. The X-ray diffraction patterns reflect all the spatial distribution of the electrons of all the atoms in the crystal. The 3-D picture of the electron density of a macromolecule can be derived from the X-ray diffraction patterns with using the Fourier transformation. The scattering intensity of the individual atoms and the phase of the wave from each scattering source is known as the individual structure factor (fj) and the structure factor of each plane (hkl) is the summation of the

individual structure factors for individual atoms, fj, times a phase factor for each atom α(hkl). The structure factor can be represented by this equation:

$$F(hkl) = \Sigma f j e^{2_{\pi} i\,(hxj+kyj+lzj)} = F(hkl) e^{i_{\alpha}\,(hkl)} \qquad (2.2)$$

In this equation F(*hkl*) is wave function with frequency while α(*hkl*) is phase angles of the reflections. The frequency of F(*hkl*) is identical to the X-ray source and the amplitude is proportional to the square root of the intensity of the reflection, which can be measured from the X-ray diffraction data. In data collection the x-ray detector can only record the intensities and not the phase and therefore phase is lost and only the intensity of the diffracted beam is recorded. The intensity of each spot can is given by:

$$I(hkl) = [\ \mathbf{F}(hkl)]^2 \qquad (2.3)$$

The Fourier transform takes the structure factor, which is function of the electron density ρ(r) and invert the functional dependence and therefore the electron density is expressed as the function of the structure factor. The atomic coordination can be extracted from it.

$$\rho(r) = \int_{\text{diffraction space}} F(s) e^{-2_{\pi} ir.s}\, d\nu_s \qquad (2.4)$$

$d\nu_s$ is a small unit of volume in the diffraction space. The integration can be replaced by summation since F(s) is not continues and is non-zero only at the reciprocal lattice points. Therefore:

$$\rho(xyx) = 1/V \sum_{h} \sum_{k} \sum_{l} F(hkl) e^{2\pi i (hx+ky+lz)} \qquad (2.5)$$

$$= 1/V \sum_{h} \sum_{k} \sum_{l} F(hkl) e^{i_{\alpha}(hkl)} e^{2\pi i (hx+ky+lz)}$$

In this equation ρ(xyx) is electron density at any point (x,y,z) and F (hkl) is the amplitude and proportional to the square root of the measured intensity of each reflection. As it was mentioned above the problem is missing the phase. There are few methods to solve the phase problem. One of them is known, as Molecular replacement is a known structure is used to find the phase of the unknown structure.

Molecular Replacement: One method to solve the phase problem is molecular replacement. In this method one know structure is used to calculate the electron density map of a biological entity that X-ray diffraction data has been collected. The two structures need to share at least 25% sequence Identity, therefore the structure can be used as the initial search model. We could easily used molecular replacement to determine the initial phase of each of the structures of CRBP II mutant complexes with all-trans-retinol. For finding the correct orientation of the protein structure the two following steps are used by molecular replacement:

Rotation Function: In rotation function, the Patterson map of the unknown structure is compared to the Patterson map of the initial searching model and in the different orientations.

The Patterson function is calculated form the amplitude using the formula:

$$P(u,v,w) = \Sigma |Fhkl|^2 \, e^{-2\pi(hu+kv+lw)} \qquad (2.6)$$

The Patterson function is essentially is Fourier transform of the intensities. Evaluation of the rotation function is only necessary over the asymmetric unit in rotation space. The rotation function is usually calculated as the function of Eulerian angles, α, β and γ.

Translation Function: The translation function places the oriented unknown biological molecule in the defined asymmetric unit to the best fit. Crowther and Blow defined the standard Patterson map based on the translation function:

$$T(t) = \int P_{obs}(u) \, . \, P_{calc}(u,t) \, du \qquad (2.7)$$

In translation function the initial model moves and the new Patterson map is calculated. The calculated Patterson map compares to the unknown structure.

Therefore this is basically defined the similarity of observed and calculated Patterson map of the observed and calculated Pattersons.

Structure Refinement: After the phase problem has been solved and the solution if found. The atomic location of each atom was defined in the electron density, the high speed computer is need to be used for adjust some parameters to minimized and reduced some parameter such as $\Sigma (F_{obs} - F_{cal})^2$. Each amino acid residue of the protein is fitted to the density until the bias introduces by the initial model is reduces. A correct structure should have a reasonable R_{factor} with absolutely no major unexplained discrepancies. The satisfactory R_{factor} value depends on the resolution of the structure.

Usually at the higher resolution the lower R_{factor} can be gained. Before starting the refinement 10% of the data randomly is separated and the refinement is performed on the 90% of the data. Based on the separated data another R_{factor}, which is called R_{free}, is calculated. As the R_{factor} by continuing the refinement should decrease the R_{free} should concurrently decrease. Usually the R_{factor} is lower than R-free however there is a limitation between their difference and this value should not be higher than 6-7 %. The value of R_{factor} which is also called R_{work} are the agreement between the values of the observed structure factor and calculated by the following equations:

$$R_{work} = \frac{\sum || F_{obs} | - | F_{cal} ||}{\sum || F_{obs} |} \qquad (2.8)$$

For all the structures with the resolution range between 1.2-1.8 Å the R_{factor} and R_{work} bellow 20% is acceptable but bellow 18% is more desirable.

II-2: The Crystal Structure of wt hCRBPII Bound with Retinol and Retinal

II-2-1: Introduction

Vitamin A or retinol is necessary for vision, cell growth, reproduction, metabolism, function and function. Because of low solubility in aqueous solution and instability of this compound, therefore retinol needs a carrier and soluble protein for transporting. Different proteins in different tissues were identified and characterized to play important role in transport, storage and metabolism of vitamin A. They are Retinol Binding

Proteins (RBP), Cellular Retinoic Acid Binding Proteins (CRABP), Cellular Retinol Binding Protein (CRBP), Cellular Retinal Binding Protein (CRALBP), Interphotoreceptor Retinol Binding Protein (IRBP).[23] In all of these proteins the hydrophobic ligand and protein have non covalent interactions.[24]

CRBPs are relatively small cytosolic protein with 133 amino acid that play important role in the metabolism of vitamin A in the cells. Four isoforms of human Cellular Retinol Binding Protein (CRBP) have been identified: I, II, III and IV. [1,2] They belong to the superfamily of intracellular Lipid Binding Proteins (iLBP). This family also includes another two group members, which are besides Cellular Retinoic Acid Binding Proteins (CRABP), and Fatty Acids Binding Proteins (FABP).[2] However in CRBPs subfamily only CRBPI and II are in the process of intracellular retinol. Despite the fact that these proteins have low sequence identity, surprisingly the structures of them are highly homologues.

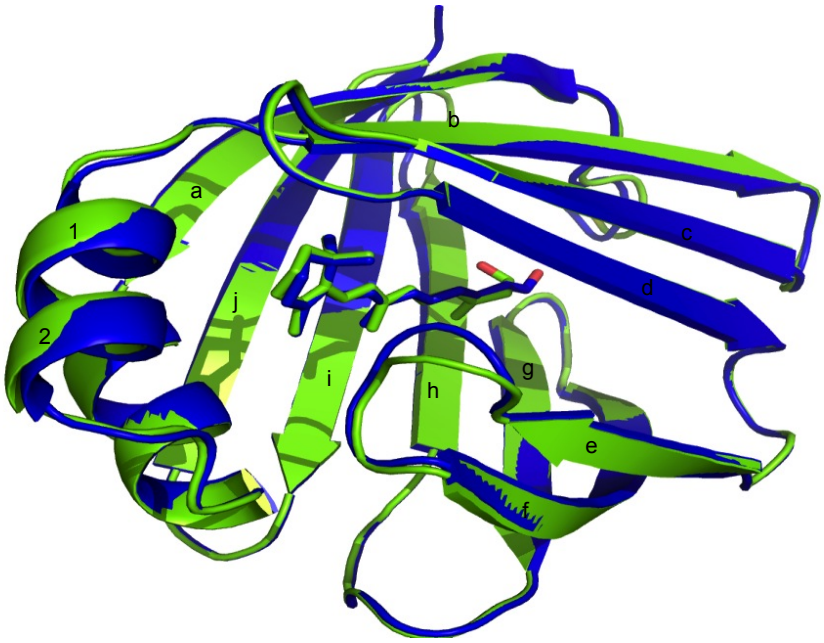


Figure II-1: The Crystal structure of overlay wt hCRBPII with all-*trans*-retinal (blue) and all-*trans*-retinol (green).

They share 10 β-sheet domains (a-j) and two short α-helices (1, 2) (Figure II-2). These two iso-forms of CRBP share 56% sequence identity[3]. CRBPI and II bind both -*trans*-retinol and 13-*cis*-retinol, while neither bind 9-*cis*-retinol[4]. These two proteins are localized in different organs. CRBPI is found predominantly in the liver and all kidney while CRBPII is more prevalent in the small intestine[5,6,7]. It was reported that CRBPII regulates retinoid metabolism in intestine and facilitation of reduction of retinal to retinol and also subsequent of esterification of retinal to retinyl ester.[25]

The different binding constant of the two CRBPs beside the different distribution of them in human body make them specific for the different propose.

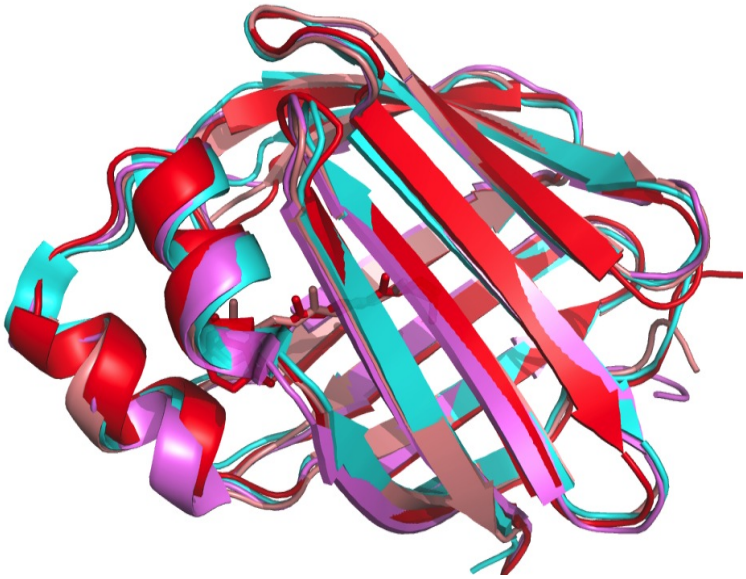


Figure II-2: The crystal structure overlay of hCRBPII (salmon red), solution structure of human CRBPI (magenta), 1KGL, crystal structure of hCRBPIII (red) 1GGL, crystal structure of hCRBPIV (cyan) 1LPJ

The first crystal structures of rat apo and all-*trans*-retinol bound CRBPII were published in 1992[8]. Comparison of the two structures showed that no significant differences were observed upon ligand binding. (The crystal structures of human CRBPII and zebra fish CRBP bound with all-*trans*-retinol at high resolution were published in 2007 and 2002 respectively.[9,10] In all three structures the main interactions with the ligand is with Gln108 and Lys40. These two residues is responsible for hydrogen bond in CRBPI and II while in CRBPIII and IV, His108 makes this interaction.

CRBPIII and CRBPIV have weaker binding constant with all-*trans*-retinol compare to the mammalian CRBPs (I and II).[2]

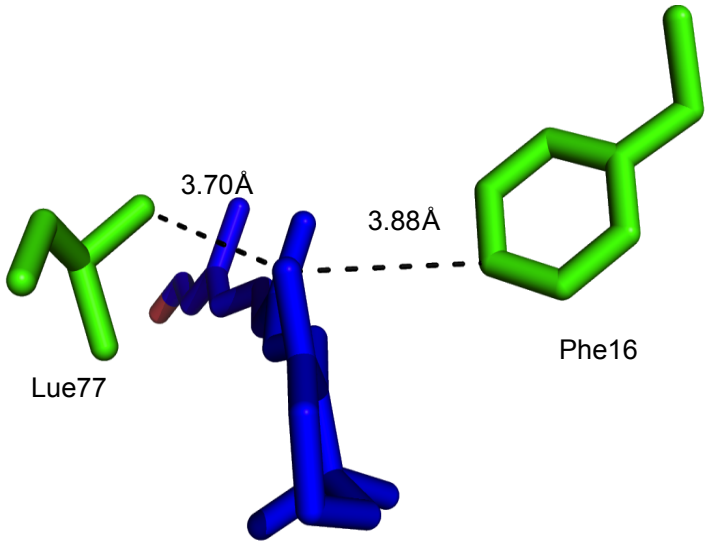


Figure II-3: The crystal structure CRBP with retinal and interaction with Phe16 and Leu77

By comparison of the binding constant between CRBPI, II with III and IV, it can be concluded since Gln108 is able to make two strong hydrogen bonds with retinol Therefore the dissociation constant of CRBPI and II for retinol is higher than CRBPIII and IV because the later proteins both have Histidine instead of Gln at the 108 position. Both of these residues (Gln108 and Lys40) make hydrogen bonds with the hydroxyl group of the ligand. All CRBP family members consist of 10 anti parallel β-strands, which together form a beta barrel. The hydrophobic ligand is deeply buried inside the binding pocket formed by this barrel. Two short α helixes cover the binding pocket like a lid, isolating the ligand from bulk solution.

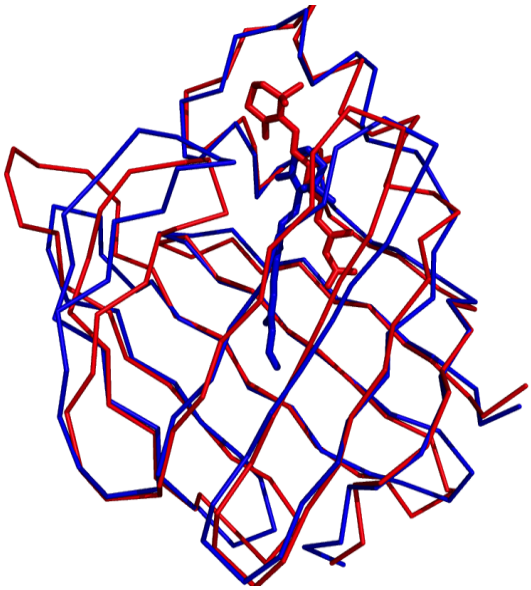


Figure II-4: The crystal structure overlaid of hCRBPII (blue) and CRABPII (red). The chromophore is highlighted in these two structures.

In addition to these two helices, the binding pocket opening is further covered by the side-chain of Arg58, which is located on the loop connecting two β-strands (C and D) and is held in position by a hydrogen bond between its guanadinium group and the carbonyl group of Ser77. The ligand is stabilized within this binding pocket, allowing it to preserve its yellow color at room temperature and in light for more than three days, while retinol solution in ethanol lost ¼ of the activity when exposed to the air.[26] CRBP II binds both all-*trans*-retinol and all-*trans*-retinal with similar affinity (dissociation constant (K_d) of CRBPII with all-*trans*-retinal is 23±10 nM, while with retinol it is 10 nM)[11,12] We have determined the first crystal structure of wt hCRBPII with retinal at 1.6Å.

The complex crystallizes in the P1 space group with four molecules per asymmetric unit. Electron density corresponding to the bound ligand was very clear in three of the four chains (A, B and C). All of the crystallographic data is listed in table II-1. The position of retinal is almost superimposable with retinol. In both ligands the β-ionone ring adopts the 6-s-*trans* conformation. Two hydrophobic residues, Phe16 and Leu77, surround the ionone ring of the ligand in hCRBPII. This position of these residues lock the chromophore in its position restricting rotation of the ionone ring about the C6-C7 single bond, locking it in the 6-s-*trans* conformation. The distances between the C5 methyl group of retinal and Leu77 and Phe16 are 3.70Å and 3.83Å respectively (Figure II-3). It is clear that rotation of the ionone ring would lead to steric clashes between the gem-dimethyl group and these two residues. Sequence alignment of hCRBP II with the other human CRBPs, rat CRBPI[13] and II, and zebrafish CRBP[10] shows that the ionone ring is surrounded by bulky hydrophobic residues in all (Figure II-5). Despite CRBPs the crystal structure of the other family member of iLBPs In contrast CRABPII has a more open binding pocket, allowing its ligand (all-*trans*-retinoic acid) to adopt a 6-s-*cis* conformation[14] which is 0.6 Kcal/mol more stable than the *trans* conformation[15].

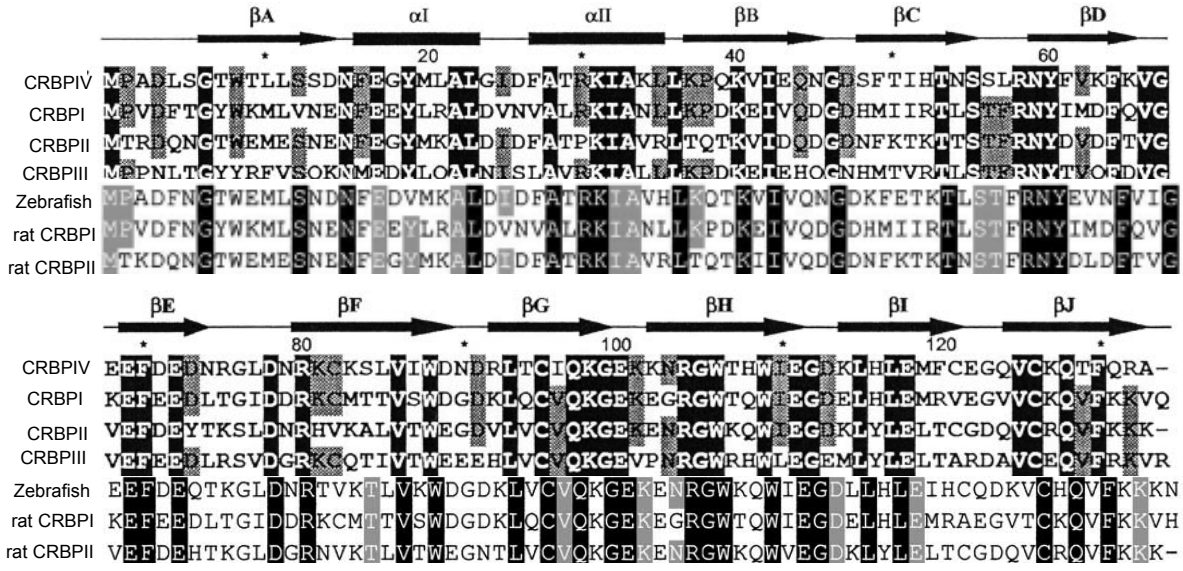


Figure II-5: The multiple sequence alignment between human CRBPs, rat CRBPI and II and zebrafish CRBP

II-2-2: Crystal Structure of Holo wt human Cellular Retinol Binding Protein II (hCRBPII) Bound with Retinol and Retinal

The overall structures of hCRBPII bound to all-*trans*-retinal and all-*trans*-retinol no significant differences (Figure II-4). However inspection of the details of binding of all-*trans*-retinol and all-*trans*-retinal show they bind differently in the binding pocket. The most important interaction in both structures is between the ligand's oxygen and the Gln108 side chain inside the binding pocket. The side chain of Gln108 is fixed in position by a small network of water molecules and residues, beginning with Thr1, continuing to Gln4, and concluding with interactions to Gln108 (Figure II-6). This water network is observed only in CRBPII, not in zebra fish CRBP, rat CRBPII and human

CRBPI, III[10] and IV[2]. This water molecule makes hydrogen bonds with the main chain carbonyls of Thr51 and Asp91. This water buttresses the orientation of Gln4 in human CRBPII (Figure II-8). Since water makes hydrogen bonds with two carbonyl groups of the main chain, only a lone pair on the oxygen of water is left to make the hydrogen bond to the amide NH_2 of Gln4, therefore defining its orientation. Therefore it is the carbonyl group of Gln4 that makes a hydrogen bond, necessarily with the amide of Gln108, defining its orientation as well. This water molecule is only seen in the all-*trans*-retinal-bound structure and is not observed in any of the CRBP members bound to retinol. As shown in Figure II-6 the carbonyl group of retinal does not make a hydrogen bond with Gln108. It is obvious that carbonyl group of retinal do not make any hydrogen bond with carbonyl of Gln108 and therefore only NH_2 group of the side chain is available to make a hydrogen bond, however the distance between NH_2 of Gln108 and the carbonyl group of retinal is about 3.6Å, which is longer than a moderate hydrogen bond. However, a water mediated interaction between Gln108 and the retinal carbonyl is formed. The hydrogen bond between this water and the retinal carbonyl is 2.49Å, which is an ideal distance for a strong low-barrier hydrogen bond. This water makes an additional hydrogen bond with the ε-amino group of Lys40, which is fixed in space by two hydrogen bonds. An overlay of the crystal structure of retinal-bound hCRBPII and the solution structure of hCRBPI (PDB entry 1KGL)[16] illustrate that in hCRBPII the ligand is more than 1Å deeper in the binding pocket. The same translation of the ligand is observed in rat CRBPI (PDB entry 1CRB)[13] and CRBPII (PDB entry 1OPB).[17] The identity of the amino acid at position 51 appears to be the deciding factor for ligand position. The sequence alignment of CRBPI and II of rat and shows that position 51 in

the binding pocket of CRBPI (rat and human) is Isoleucine while this position is Threonine in CRBPII. The larger steric bulk of Isoleucine prevents deeper penetration of the ligand in the binding pocket. CRBPI has weaker binding affinity with retinal. Inspection of the structures of human and rat CRBPI (PDB entry 1CRB)[13] shows that retinal can not bind in CRBPI in the same way as hCRBPII, possibly explaining the weaker binding affinity of CRBPI.

The side chain of Isoleucine also disrupts the water network seen in CRBPII. The CRBPII T51I mutant was produced and its binding affinity (K_d) for retinal measured. The dissociation constant of this mutant for all-trans-retinal was two fold weaker (45±10 nM) relative to wt CRBPII, but was still much smaller than all-*trans*-retinal binding to CRBPI,

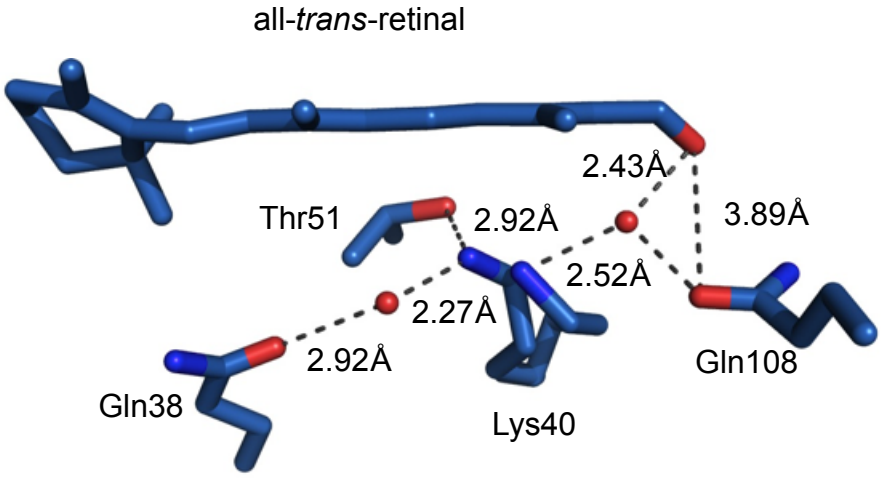


Figure II-6: The details of the hydrogen bonds in hCRBPII bound to all-*trans*-retinal.

indicating that position 51 is not solely responsible for the differences in affinity for retinal in the two proteins. Unfortunately attempts at crystallization of CRBPII T51I bound all-*trans*-retinal was not successful.

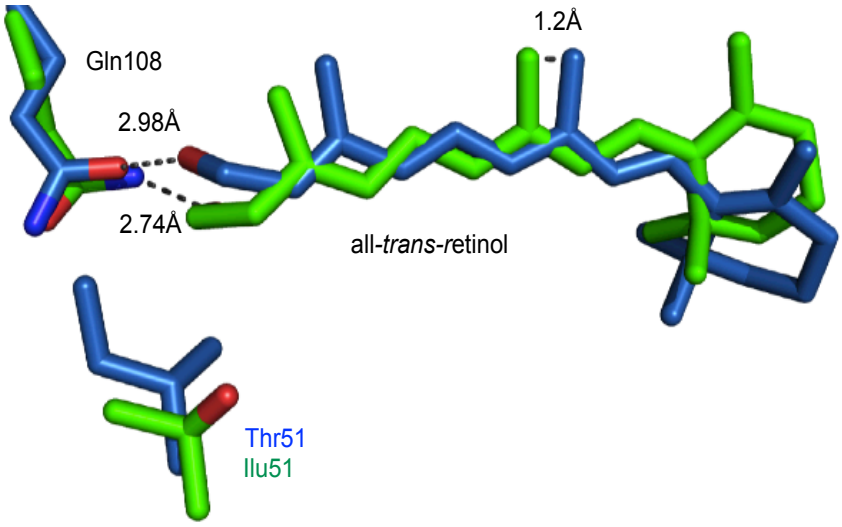


Figure II-7: The overlaid structure of CRBPII with all-*trans*-retinal (green) and solution structure of CRBPI (blue, PDB entry: 1KGL)[16]

The crystal structure of hCRBPII (PDB entry 2RCT)[9] at 1.2Å shows that the hydroxyl group of retinol adopts two different conformations (PDB ID: 2RCT), one of which having a torsion angle about the C13-C14 bond of 113 degree which is far from the 180 degree angle expected for all-*trans*-retinol. Surprisingly, and consistent with the previous observation, the C13 of retinol does not indicate the expected planar sp^2 configuration but instead appears to be tetrahedral, consistent with sp^3 hybridization. The two different conformations for the hydroxyl group were also observed in Zebrafish CRBP (PDB entry 1KQW)[10] as well. In our high-resolution crystal structure of all-*trans*-retinol-bound CRBPII, we observed a single conformation, which is equivalent to the more highly occupied position found in the previous structure of the complex (Figure II-9).

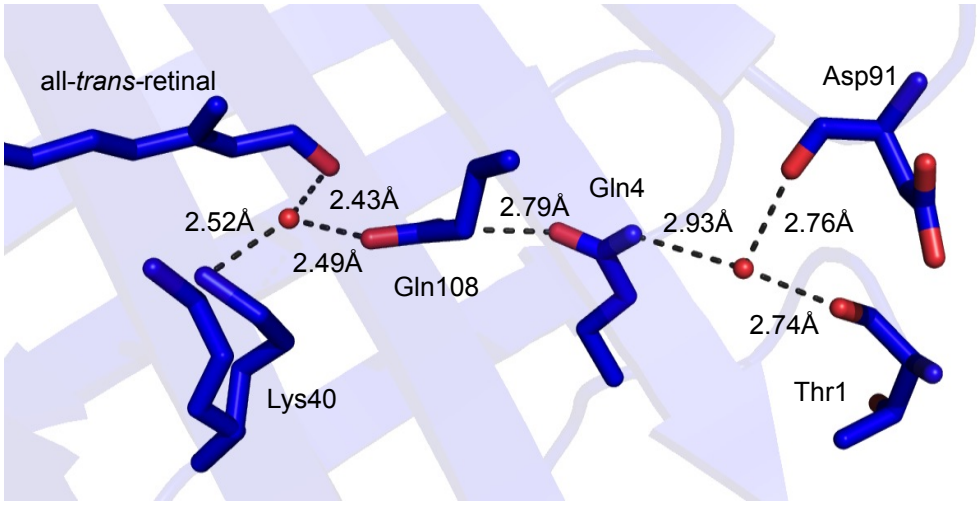


Figure II-8: The water network and the binding site in hCRBPII with all-*trans*-retinal

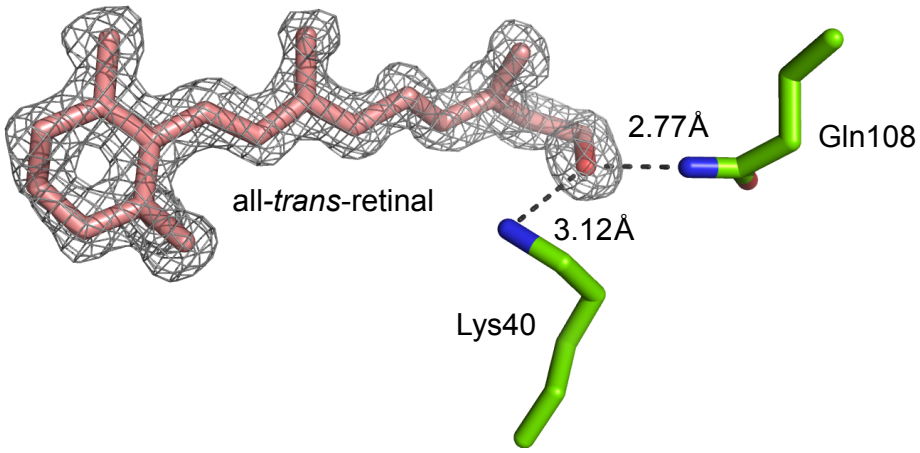


Figure II-9: The crystal structure of wt hCRBPII with all-*trans*-retinal

It is clear in our structure that the double bond between C13-C14 is has also been reduced as both centers indicate sp^3 hybridization and the position of C15 of retinol also shows that the torsion angle about the C13-C14 bond is about 50 degrees,

not the 180 degrees expected for all-*trans*-retinol. As the figure illustrates the water molecule that hydrogen bonds with the carbonyl, and bridges to Gln108 and Lys40 in our all-*trans*-retinal structure has been replaced by the hydroxyl group of retinol, which now makes direct interaction with Gln108 and Lys40. In the crystal structure wt hCRBPII with retinal this dihedral angel is only 3.95°, which is consistent with the expected sp^2 hybridization. It is obvious that the C13-C14 double bond in retinol has been reduced to a single bond. At first we suspected an impurity in the retinol as a result of retinol degradation over time. Therefore we produced new crystals with fresh retinol. The H-NMR of retinol was taken just prior to crystallization and confirms the purity of this compound. Also the absorbance of retinol solution in ethanol was measured to be 325 nm as previously reported.[22] The crystals were grown and flash frozen within three days. The structure was redetermined, this time in a P1 space group with eight-molecules per asymmetric unit at 1.7Å resolution. We observed the same retinol conformation as previously described in all eight molecules in the asymmetric unit. Wt hCRBP or not, T51V mutant was generated. The crystal structure of this mutant in P1 space group with two molecules per asymmetric unit at 1.9Å resolution was determined. Further the structure of the CRBPII T51V mutant was also determined and the retinol electron density in this structure also shows the same puzzling distortion (Figure II-10). As mentioned above this distorted structure for retinol was observed by at least two different research groups for both zebra fish and human[10,17]. Unfortunately no crystal structure of CRBPI in complex with all-*trans*-retinol is available, though the solution structure of this complex shows a single conformation for retinol consistent with all-*trans* retinol.

A possible solution to this puzzle came during data collection of these complexes at the synchrotron. It was observed that the yellow color of the crystal changed to dark orange after the first few frames of data collection and this color is maintained to the end of data collection. It is therefore possible that some rearrangement occurred during the synchrotron data collection leading to the ligand structure observed. All of the structures that display the severely distorted retinol conformation were collected at synchrotron sources.

In contrast, the first reported crystal structure of rat CRPPII bound with all-*trans*-retinol (at 1.9Å resolution) shows a single conformation for retinol consistent with all trans[13]. Rigaku RU-200. The most likely explanation for the anomalous retinol structure is that

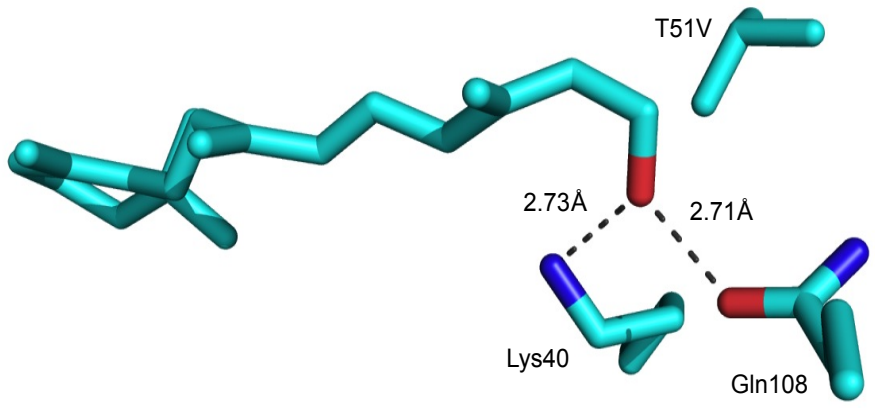


Figure II-10: The crystal structure of single mutant T51V with all-*trans*-retinol

The diffraction data for this structure was collected using CuKα X-rays at home source. This data indicates that retinol by was degraded during synchrotron data collection

Based on the rat CRBPII crystal structure we can conclude that real structure of hCRBPII is similar to the rat CRBPII and hydroxyl group makes two hydrogen bonds with both the carbonyl and amino groups of Gln108.

II-3: Experimental

II-3-1: Material and Method

The hCRBPII gene was purchased from ATCC Company and cloned in pET17b vector, with using the NdeI and XhoI restriction enzyme. 10 μl of the resulting plasmid were transformed in 100 μL of XL-1 Blue super (Novagen®) competent cells. The cells were incubated for 30 min in the ice and heat shocked at 42 °C for 45 seconds, then 1 mL of Luria-Bertani broth (LB) was added and the cells were incubated at 37°C for an hour. The result mixture were spread on the LB agar plate with (ampicillin 100 μg/mL and incubated at 37°C for 16 hour. A single colony was picked from the plate and inoculated in 5 mL of LB media contains 100 μg/mL ampicillin. The cell culture was grown over night (12-16 hours) at 37°C, then the media were centrifuged at 15000 rpm for 1 min. DNA extraction and isolation from the cell pellet was done by using the standard procedure of Promega Wizard. Plus SV Miniprep (A1330) DNA purification kit. The target gene was transformed into BL21(DE3) pLysS (Invitrogen™) E.coli competent cells for protein expression. A single colony was picked and inoculated in 5 mL of LB, containing 100 μg/mL ampicillin antibiotics at 37°C for 12-16 hours. The media transferred to 50 mL of LB with the same antibiotic at 37°C overnight. This media

transferred into 1L of LB media with ampicillin (100 mg/L) and incubated at 37°C until OD600 reached 0.5-1.0. The protein expression was induced with 1 mM isopropyl β-D-1 thiogalactopyranoside (IPTG) (IPTG, purchased from Gold Biotechnology) overnight at 22°C. The cells were harvested by centrifugation at 5000 rpm for 20 min. The harvest cells were resuspended again with Tris buffer (10 mM Tris, 10 mM NaCl pH = 8.0, 50mL). The suspended cells were lysed by sonication and the lysed cells were centrifuged at 4°C (14,000 rpm, 20 min).

The supernatant were loaded on Fast Q anion exchange resin, which was equilibrated with (10 mM Tris, 10 mM NaCl pH = 8.0). The after the binding, the resin was washed three times with 50 mL of (10 mM Tris, 10 mM NaCl pH = 8.0). The protein was eluted with 40 mL of (10 mM Tris, 100 mM NaCl pH = 8.0).

The eluted protein was desalted by using Centriprep® centrifugal filter at 2500 rpm. The protein was applied on BioLogic DuoFlow system loaded with 15Q anion exchange resin for the second step of purification. The purity of the protein was determined with SDS-PAGE.

K_d Determination via Fluorescence Titration: All-*trans*-retinol and all-*trans*-retinal were purchased from Sigma. The dissociation constant K_d was determined by fluorescence titration with previously reported method[18]

II-3-2: Crystallization and Structure Determination

The pure protein was concentrated down to 5-10 mg/mL. The complexes of protein with all-*trans*-retinal and all-*trans*-retinol were prepared by adding 3-4 equiv of retinal and retinol solution (30 mM retinal and retinol in ethanol). The final concentration of ethanol in protein solution should be lower than 10% V/V. The mixture of protein and ligands were incubated at room temperature in Black LiteSafe™ Microcentrifuge Tubes for an hour. The crystal of the complex of protein and ligands were prepared at room temperature by hanging drop vapor diffusion method in Limbro plates and plates were wrapped in aluminum foil to prevent light-initiated degradation of the retinal and retinol. The best crystals were grown by using Hampton Research Screens contains (30-35% PEG 4000 (Sigma-Aldrich), 0.1 M sodium acetate (Columbus Chemical Industry) pH 4.6-4.8 and 0.1 M ammonium acetate (J. T. Baker) with 1 µL of hCRBPII-ligand complex. The crystals appeared after 3 days and reached to the maximum size in one week. For preventing degradation of ligand at room temprature the crystals were flash frozen by liquid nitrogen and using cryoprotectant solution (30% PEG 4000, 0.1 M sodium acetate pH 4.6-4.8, 0.1 M ammonium acetate, 10-20% glycerol) after three days.

II-3-3: Data Collection, Refinement and Solution

All of the diffraction data were collected at beamline 21-ID-D, LS-CAT (Argonne National Laboratory, Advanced Photon Source, Chicago, IL) using a MAR300 detector and 1.00Å wavelength radiation at 100K. The diffraction data were indexed using HKL2000[19] software package. The initial phase were solved by MOLREP using human CRBPII Protein data bank, accession code 2RCQ as model in the CCP4 program package and REFMAC5 in the CCP4 suite produced the initial electron density maps[20,21]. The entire rebuilding, placement of ordered water molecules was done manually using COOT. The chromophore was created using jligand and manually fitted in the electron density at the end of the refinement. All refinements were carried out using REFMAC5 program.

Table II-1: X-ray crystallography data and refinement statistics of wt hCRBPII

	Holo-WT-CRBPII-retinol	Holo-WT-CRBPII-retinol	T51V-retinol	Holo-WT-CRBPII-retinal
Space group	P21	P1		P1
a(Å)	34.65	36.35	29.37	35.44
b(Å)	75.14	54.94	35.96	54.89
c (Å)	54.65	128.16	63.82	68.71
α(°)	90.00	92.19	90.73	107.75
β(°)	100.78	92.74	94.19	97.66
δ(°)	90.00	103.21	112.81	103.08
Molecule per asymmetric unit	2	8	2	4
Total reflection	306503	181745	36507	241247
Unique reflection	87302	89048	18179	76879
Completeness (%)	99.8(84.0)[a]	92.0(93.1)[a]	86.3(95.9)[a]	91.5-(83.9)[a]
Average I/σ	20.19-2.04	33.40(1,98)[a]	34.24-6.48	51.85-4.13
Resolution (Å) (last shell)	50.00-1.19(1.21-1.19)	20.00-1.83(1.83-1.80)	50.00-1.90(1.93-1.91)	50.00-1.50(1.53-1.50)
R_{merge} (%)	9.1(43.4)[a]	5.5(46.7)[a]	3.7(27.4)	3.2(32.1)[a]
Wavelength (Å)	0.97872	1.0781	0.97872	0.97856
R_{work}/R_{free} (%)	21.53-18.50	35.82/27.47	29.57/21.77	25.65/20.40
		RMSD from ideal value		
Bond length (Å)	0.0297	0.0230	0.0205	0.0240
Bond angle (°)	2.3815	2.0600	1.8208	2.200
Number of water molecule	530	790	150	320

REFERENCES

REFERENCES

1. Folli, C., Calderone, V., Ottonello, S., Bolchi, A., Zanotti, G., Stoppini, M., and Berni, R. **2001**, *Proc. Natl. Acad. Sci*. U. S. A. 98, 3710–3715

2. Folli, C., Calderone, V., Ottonello, S., Bolchi, A., Zanotti, G., Stoppini, M., and Berni, R. **2002**, *J.Biol.Chem*. 277, 41970-41977

3. Li, E., Demmer, L. A., Sweetzer, D. A., Ong, D. E., and Gordon, J. I. **1986***, Proc. Natl. Acad. Sci*. U. S. A. 83,5779-5783

4. MacDonald, P.N., Ong, D.E. **1987***, J.Biol.Chem*, 262, 10550-10556

5. Crow, J. A., and Ong, D. E. **1985**, Proc. Natl. Acad. Sci. U. S. A. 82, 4707-4711

6. Ong, D. E. **1984**, J. Biol. Chem. 259, 1476-1482

7. Chytil, F., and Ong, D. E. **1984**,The Retinoids (Sporn, M. B., Roberts, A. B., and Goodman, D. S., eds) Vol. 2, pp. 89-123, Academic Press, Orlando, FL

8. Inagami S, Ong DE.,**1992**, J Nutr 122(3):450-6

9. Tarter, M., Capaldi, S., Carrizo, M.E., Ambrosi, E., Perduca, M., Monaco, H.L. **2007,** Proteins 70: 1626-1630

10. Calderone V, Folli C, Marchesani A, Berni R, Zanotti G . **2002**, J. Mol. Biol.321, 527-535

11. Li, E., Qian, S. J., Winter, N. S., d'Avignon, A., Levin, M. S. & Gordon, J.I. **1991** J. Biol. Chem. 266, 3622–3629

12. Malpeli, G., Stoppini, M., Zapponi, M. C., Folli, C. & Berni, R. **1995**, Eur. J. Biochem. 229, 486–493.

13. Cowan, S. W., Newcomer, M. E. & Jones, T. A. **1993**. J. Mol. Biol. 230, 1225-1246

14. Kleywegt, G.J., Bergfors, T.,Senn, H., Le Motte, P., Gsell,B., Shudo, K, Jones, T.A. **1994** Structure 2: 1241-1258

15. Rajput, J.; Rahbek, D. B.; Andersen, L. H.; Hirshfeld, A.; Sheves,

M.; Altoe, P.; Orlandi, G.; Garavelli, M., Probing and modeling the absorption retinal protein chromophores in vacuo. Angew. Chem. Int. Ed. Engl. **2010**, 4910, 1790-1793

16. Franzoni, L., Lucke, C., Perez, C., Cavazzini, D., Rademacher, M., Ludwig, C., Spisni, A., Rossi, G.L., Ruterjans, H. **2002**. J.Biol.Chem. 277: 21983-21997

17. Winter, N.S., Bratt, J.M., Banaszak, L.J. **1993** J.Mol.Biol. 230, 1247-1259

18. Vasileiou, C., PhD. Dissertation, Michigan State University, Eas Lansing, 2005

19. Z. Otwinowski, W. Minor, Macromolecular Crystallography, Pt A 276, 307 1997

20. S. Bailey, Acta Crystallographica Section D-Biological Crystallography 50, 760,1994

21. A. Vagin, A. Teplyakov, J. Appl. Crystallogr. 30, 1022,**1997**

22. Dawson, R.M.C. et al., eds. Data for Biochem. Res. 3rd edition, Clarendon Press, Oxford, 136,**1987**

23. David E. Ong and David L. Page, *J. Lipid Res*. **1987**. 28: 739- 745.

24. Veerkamp, J. H., Peeters, R. A., and Maatman, R. G. H. J. ,**1991**, Biochim. Biophys. Acta 1081, 1–24

25. Vogel.s, Mendelsohn. C, Mertz.J, Piantedosi R, Waldburger**i** C, Gottesman**i**. M and Blaner.W, *Journal of Biological Chemistry*, **2001**, 276, 1353-1360

26. Sigma-Aldrich Production & Quality Control

CHAPTER III

The Structural Insight of hCRBPII as Rhodopsin Mimic

III-1: The Design of a Rhodopsin Mimic Based on a Heterogonous Protein *System*

As was mentioned in Chapter I, in this study human Cellular Retinol Binding Protein II (hCRBPII) was selected for use as a rhodopsin mimic to study wavelength regulation. Since in rhodopsin the chromophore, which is 11-*cis*-retinal, has a covalent bond to the protein via an active lysine in the binding pocket and makes a protonated Schiff base therefore for making the mimic of this protein generation the active lysine is necessary. Inspection of the crystal structure of wt hCRBPII shows that Gln108 makes a hydrogen bond with the hydroxyl group of all-*trans*-retinal. This residue seemed to be in the best position for mutation to make this active lysine so that it would have a favorable Bürgi-Dunitz trajectory to attack the aldehyde. Therefore Dr. Wenjing Wang in Professor Borhan's group generated the first mutant (Q108K). However the UV-vis spectrum of the Q108K mutant incubated with all-*trans*-retinal shows a very small peak at ~506 nm and the major peak was observed at 356 nm. This data shows that the formed Schiff base (SB) was not protonated in this mutant at a physiological pH. Inspection of the binding pocket of wt hCRBPII by Dr. Wenjing Wang illustrated that positively charged amino acid Lys40 is close to the PSB and may decrease the pK_a of the Q108K mutant, perturb the pK_a of the retinal-PSB and prevent formation of the PSB due to charge repulsion. To improve the pK_a of the PSB two different strategies were tried: 1- Introduction of a counter anion or a negatively charged residue close to the SB to stabilized the positive charge. This strategy was successfully applied on the first

generation of rhodopsin mimics (CRABPII). 2- Removing the positive charge near the PSB by mutation of Lys40 to a hydrophobic residue like leucine. The two mutants were generated by Dr. Wang. Both of these mutants have higher pK_a than the physiological pK_a (Table III-1). Q108K:K40L not only has a reasonable pK_a but also is more red shifted compared to Q108K:T51D and therefore the double mutant Q108K:K40L was chosen as the platform for all further mutations.

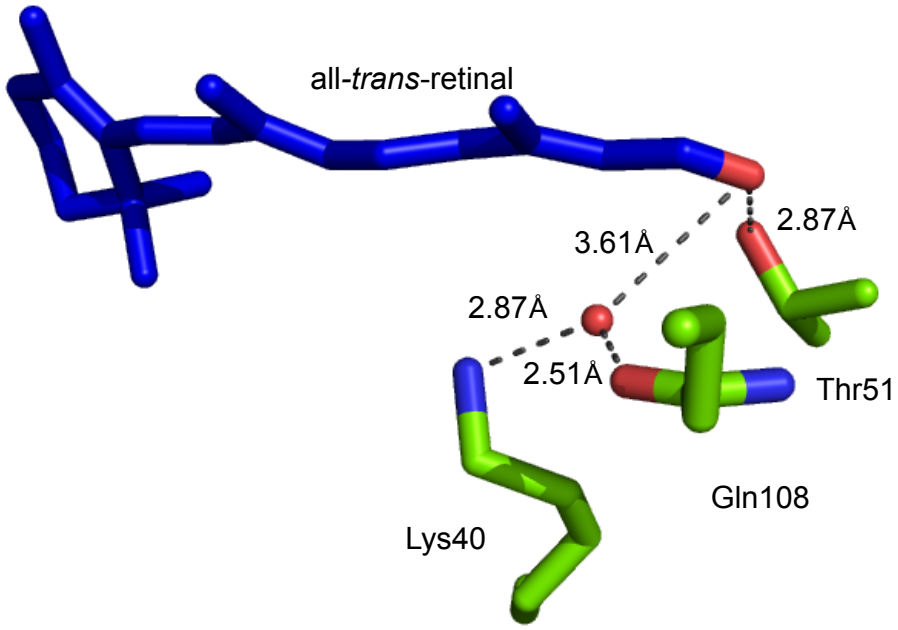


Figure III-1: The crystal structure of wt hCRBPII bound to all-*trans*-retinal with the highlighted amino acid residues that are important for generation of the platform of all further mutations.

Table III-1: The generation of high pK_a mutants

Mutants	λ_{max} (nm)	pK_a
Q108K	506	5.2
Q108K:K40L	508	7.9
Q108K:T51D	474	9.2

III-1-1: The Crystal Structure of Q108K:K40L (KL) Bound to Retinal

The crystal structure of KL was solved by using the Molecular Replacement (MR) method and the structure wild type hCRBPII was used as the initial model. The structure was refined at 1.7Å in P1 space group with two molecules per asymmetric unit and with crystallographic R_{factor} = 21.22% and R_{free} = 26.72%. The density of the ligand was clear in one chain. The overlaid structure of this mutant with wt hCRBPII shows that the plane of the polyene chain of the chromophore rotated in hCRBPII mutants relative to the wt hCRBPII retinal and retinol-bound structures. However the crystal structures show that the ionone ring of retinal is located in almost the same place, meaning that the polyene is rotated relative to the wt hCRBPII retinol structure (Figure III-2).

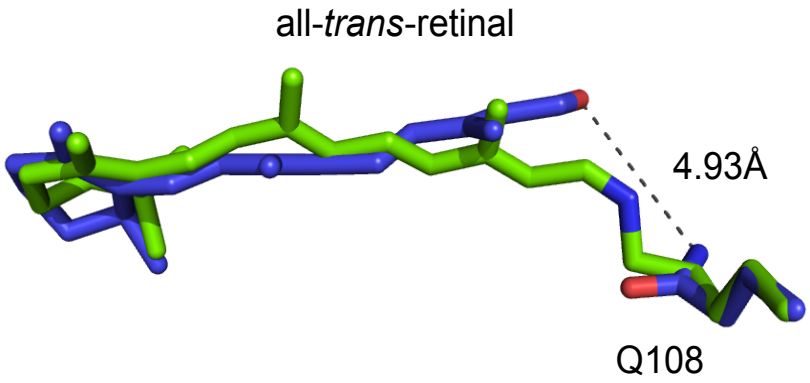


Figure III-2: The overlay of Q108K:K40L mutant with wt hCRBPII with all-*trans*-retinal.

The crystal structure of this mutant demonstrates that in this system the positive charge of the PSB appears to be stabilized in two different ways. First an ordered water molecule interacts with both Gln4 and the iminium of the PSB (Figure III-3). This water molecule was observed in all of the mutant structures that contained Gln4. Mutation of this residue disturbs this water network and the iminium adopts *trans* conformation and dropping the pK_a of retinal-PSB. The second interaction that stabilizes the pK_a is the π-cation interaction between the PSB and the side chain of Trp109, which is almost 3.8Å away from the protonated Schiff base. This interaction has been observed in all of the hCRBPII mutants. Like the structure of wt hCRBPII, only a few well-ordered water molecules were observed in the protein binding pocket.

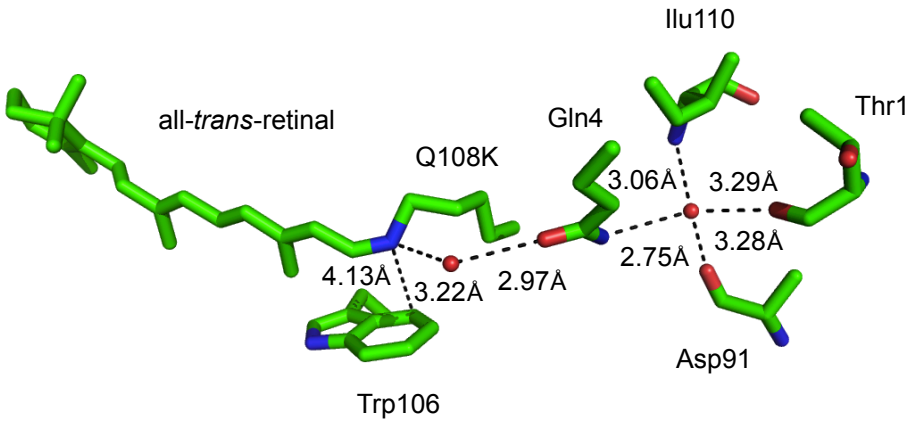


Figure III-3: The π-cation interaction and water mediate stabilize the positive charge of the iminium.

One of them is located between Gln38 and Gln128, which is 3.6Å away from the polyene of the chromophore (Figure III-4). The other one is between the hydroxyl group of Thr51 and Thr53. The binding pocket of bovine rhodopsin is made of hydrophobic residues while in hCRBPII, the binding pocket is relatively non polar.

The next step was to make systematic mutations of the polar residues inside the binding pocket to the non-polar amino acids. Inspection inside the binding pocket of the structure of both wt hCRBPII and KL showed that Thr51 is 5.4Å away from the Schiff base (Figure III-4), the closest polar residue to the PSB. One water molecule was observed between Thr51 and Thr53. This water molecule is very well ordered and observed in the structure of wt hCRBPII as well (PDB entry 2RCT). This water molecule is near the middle chromophore backbone (4.17Å), 6.1Å away from C13 of retinal and almost 7Å away from the PSB.

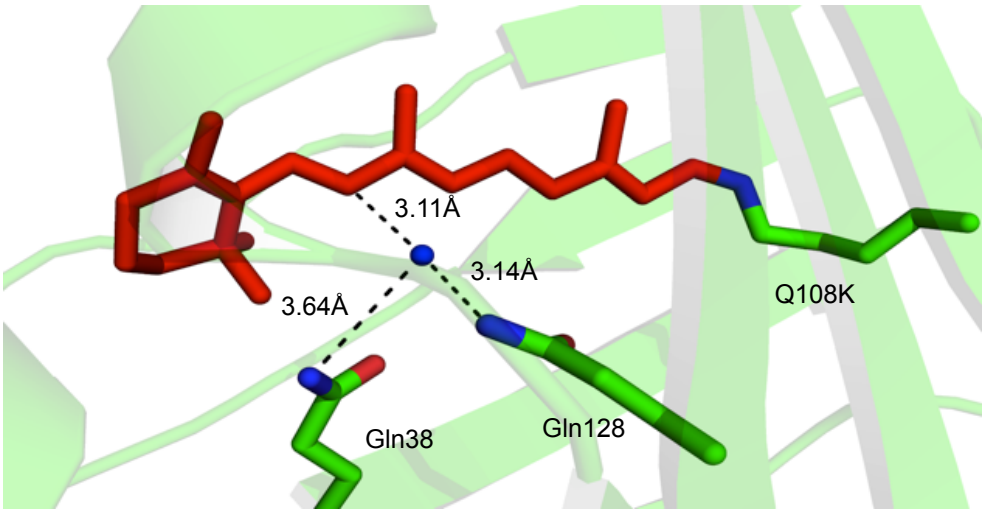


Figure III-4: The water network inside the binding pocket of hCRBPII, close to the chromophore.

.

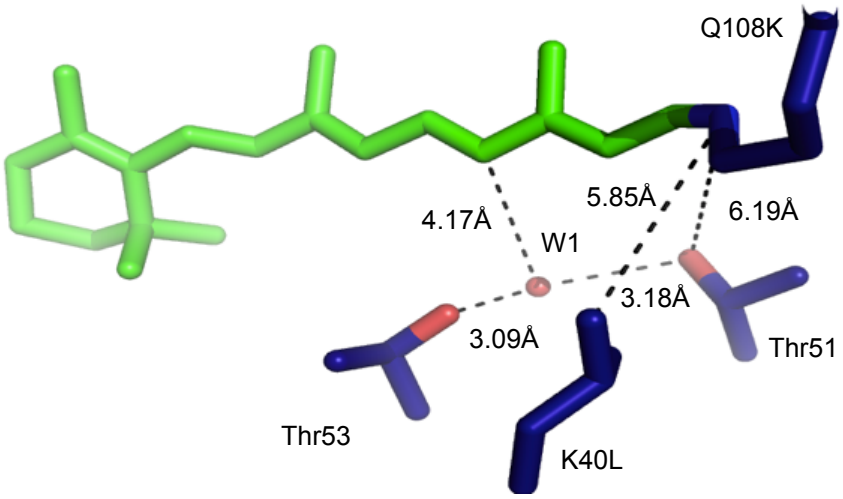


Figure III-5: The crystal structure of Q108K:K40L (KL) shows a water molecule forming a hydrogen bond with residues Thr51 and Thr53.

III-1-2: The Crystal Structure of Q108K:K40L:T51V (KLV) Bound to Retinal

The crystal structure of wt hCRBPII and Q108K:K40L shows the binding pocket of hCRBPII consists primarily of polar amino acids. Theoretically, removing the more polar groups from around the PSB-retinal can help the delocalization of the positive charge of the chromophore and lead to more red shifted absorbance. The closet polar residue to the PSB is Thr51. As shown in Figure III-4, the hydroxyl group of this residue is about 6Å away from the PSB; however the hydroxyl group does not have direct interaction with the PSB. The crystal structure shows that the hydroxyl group of Thr51 is pointing in the direction of the ionone ring of the chromophore and makes water mediated interaction with the Thr53 hydroxyl group. Mutation of Thr51 to a non polar amino acid valine (T51V mutation) was performed by Dr. Wenjing Wang. This mutation leads to a significant bathochromic shift (25 nm) moving the λ_{max} from 508 nm to 533 nm. The crystal structure of this mutant was determined in the same condition as Q108K:K40L mutant and refined at 1.7Å with crystallographic R_{fac} = 21.85% and R_{free} = 24.64%. All of the structural data is listed in table IV-1. The space group is P1 and the unit cell is virtually identical to the Q108K:K40L structure. The overlaid structures of Q108K:K40L and Q108K:K40L:T51V (Figure III-6) show that the overall structures are quite similar. However the water molecule between Thr51 and Thr53 has disappeared in both chains due to the T51V mutation.

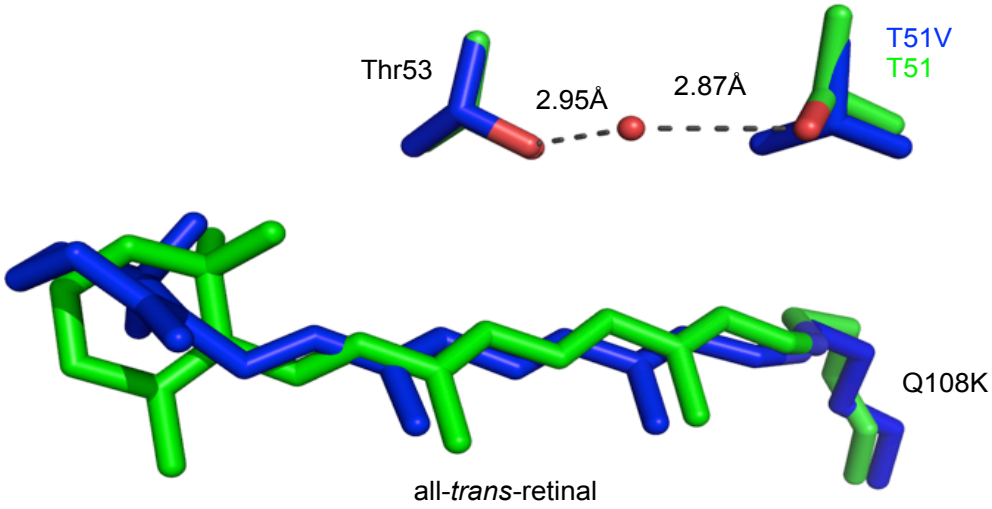


Figure III-6: The crystal structure overlaid of Q108K:K40L mutant (green) and Q108K:K40L:T51V mutant (dark blue). The water mediate between Thr51 and Thr53 belongs to Q108K:K40L mutant.

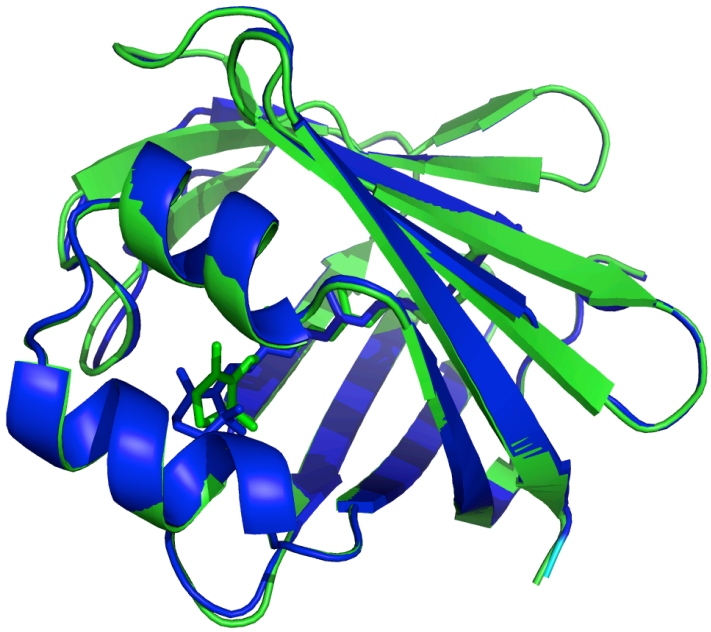


Figure III-7: The overlaid structure of Q108K:K40 (green) andQ108K:K40L:T51V (dark blue) bound to all-*trans*-retinal.

No other major differences between these two structures were observed. The pK_a of Q108K:K40L:T51V mutant was measured to be 8.3 which is 0.4 unit higher than Q108K:K40L and therefore the data shows that the PSB hydrophilicity of Thr51 did not seem to play a role in stabilization of the pK_a .

Removing this polar residue by T51V mutation has two effects: 1: Removing the polarity of the amino acid residue; 2: removing a water molecule in the binding pocket that is 7.09Å away from the PSB. The combination of these two effects leads to 25 nm red shifting compared to Q108K:K40L. The spectroscopic data and the crystal structure show that less polarity in the vicinity of the PSB of the chromophore is needed for more red shifting absorbance.

III-1-3: The Crystal Structure of Q108K:K40L:T53C (KLC) Bound to Retinal

Theoretical calculations suggested that introducing a polarizable amino acid may lead to more red shifting.[1,2,3] Since the positive charge of the protonated Schiff base is transferred from the iminium region to the ionone ring during the electronic excitation, a polarizable environment could lower the energy level of the excited state and decrease the energy gap between the ground and excited states, leading to more red shifted spectra. The crystal structure of hCRBPII shows several polar residues in the binding pocket of hCRBPII. Thr53, situated near the middle of the polyene chain, was chosen to test this hypothesis. Tryptophan, other aromatic residues and cysteine are all relatively polarizable. However since the binding pocket of hCRBP is narrow, Dr. Wenjing Wang

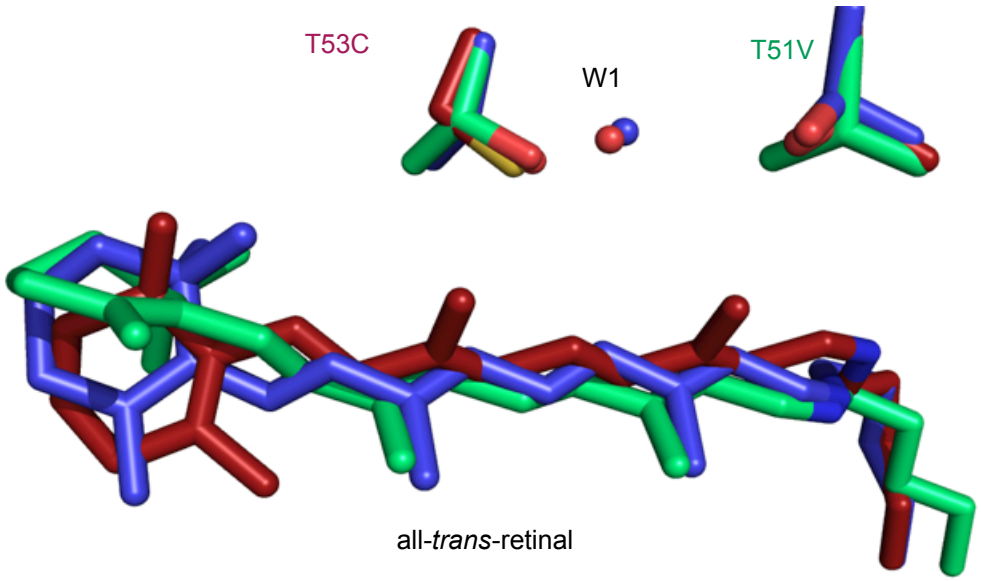
generated the T53C mutation combined with Q108K:K40L to give Q108K:K40L:T53C (KLC).

Table III-2: Comparison of mutations at the Thr53 position

Mutants	λ_{max} (nm)	pK_a
Q108K: K40L:T53	508	7.9
Q108K:K40L:T53V	503	7.3
Q108K:K40L:T53C	513	8.3

Crystals of Q108K:K40L:T53C (KLC) are quite similar to the previous two mutants (almost identical P1 space group with two molecules per asymmetric unit). The structure was refined at 1.4Å (R_{fac} = 21.20% and R_{free} = 28.20%). Again the overall structure, backbone, etc. of this mutant shows little deviation from the Q108K:K40L and Q108K:K40L:T51V structures (Figure III-7). However the conformation of the bound chromophore is considerably altered. The plane of the polyene chain is rotated almost 90° relative to KL that has 6-s-*trans* conformation, resulting in a 6-s-cis conformation. Since the position of the b-ionone ring did not change the conformational change is due to the rotation of the polyene chain. It is not clear the T53 mutation caused this conformational change. Though this is one of the mutations that proved a platform for the mutations leading to wavelength tuning, this distorted structure was not observed in any of the other T53C containing mutants. The conformation of the iminium is cis which is the same as Q108K:K40L and Q108K:K40L:T51V, and the water mediated interaction

between the PSB and Gln4 is present in both chains in this structure. In this structure the water molecule between Thr51 and Thr53 is still observed in both chains. The side chain of Cys53, which is pointing in the direction of the PSB, makes a hydrogen bond to the aforementioned water molecule. The absorbance of Q108K:K40L:T53C is 513 nm, only 5 nm more red shifted compared to Q108K:K40L (508 nm). It seems that the polarity of the amino acid in the middle of polyene does not have a significant effect on the wavelength tuning, compared to the effect of the water molecule located near the middle of the polyene chain, which seems to have a far more important role in tuning the wavelength. Since water is the most polar group is in the binding pocket, having a larger dipole moment than virtually any amino acid side chain, removing or at least changing the polarity of the water which is located almost in the middle of the polyene chain and 4.17Å away from the back bone of the chromophore can have a very significant effect on the spectral tuning. Since in Q108K:K40L:T53C the position of the water did not change, only a 5 nm red shift was observed, while in the T51V mutation, which resulted in complete removal of the water molecule, a 25 nm bathochromic shift was observed.



F**igure III-8:** Overlay of the structures of three mutants: Q108K:K40L (blue), Q108K:K40L:T51V (green) and Q108K:K40L:T53C (red). As it is shown in this figure the plane of the backbone of the chromophore rotates almost 90°. The water molecule between Thr51 and Thr53 is available.

III-1-4: The Crystal Structure of Q108K:K40L:T51V:T53C(KLVC) Bound to Retinal

Combination of the Q108K:K40L mutations with the T51V and T53C mutations resulted in tetra mutant Q108K:K40L:T51V:T53C that absorbs at 539 nm. The λ_{max} of this mutant is the combination of the 25 nm from T51V and 5 nm from T53C by adding to Q108K:K40L (508 mm). In other words the absorbance of this mutant is only controlled by these mutations. This mutant also crystallized in the P1 space group with two molecules per asymmetric unit. The data was refined at 1.6Å with crystallographic R_{fac} = 21.31% and R_{free} = 26.46 %. This structure is superimposed on the other three previous structures (RMS = 0.153Å for 133 amino acid residues) (Figure III-8). The similarity of this combination with Q108K:K40K indicates that different mutations do not

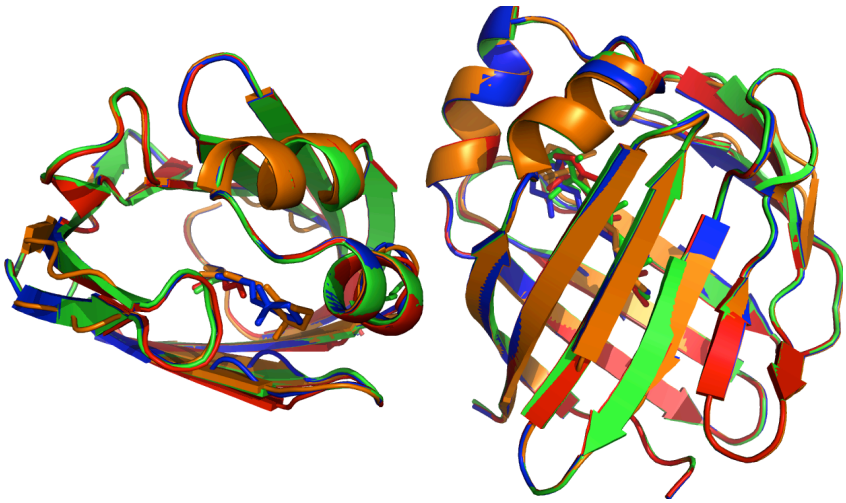


Figure III-9: The crystal structure overlay of four different mutants in P1 space group: Q108K:K40L (green), Q108K:K40L :T51V (blue), Q108K:K40L:T53C (orange) and Q108K:K40L:T51V:T53C (red).

change the integrity of the structure. The conformation of the chromophore in Q108K:K40L:T51V:T53C is located in the same place as in the other mutant structures with only minor conformational change relative to KL and KLV. The side chain of Cys53 adopts two different conformations, one of them pointed to the ionone ring and the other one pointed to the Schiff base region. The Occupancy between these two conformations is about 50%. The water molecule located between Thr51 and Thr53 is completely removed in this mutant. The different conformation can be explained by removing the water molecule, since there is no H-bond available to hold the Cys53 in a specific location.

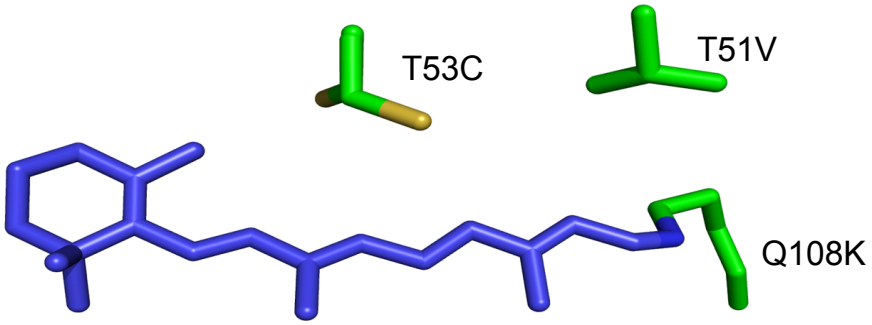


Figure III-10: The crystal structure of KLVC with the highlighted mutation positions. Cys53 adopts different conformations.

Based on the crystal structure of these four mutants and comparison of their absorbances in table III-3 we can conclude that two parameters have important roles for tuning the wavelength to more red shifted absorbance: 1- Removing the polar residues close to the PSB region. This effect can help the positive charge to be more delocalized

along the polyene chain. 2- Removing the polar groups in the binding pocket which potentially can have an interaction with the chromophore. These polar groups can be any polar or charged molecules in the binding pocket for example in hCRBPII water molecules. The effect of removing or changing the polarity (decrease or increase) of these groups has an important role in tuning the wavelength.

Table III-3: the effect of removal of the polar residues around the chromophore

Mutants	λ_{max}(nm)	pK_a
Q108K:K40L	508	7.9
Q108K:K40L:T51V	533	8.3
Q108K:K40L:T53C	513	8.4
Q108K:K40L:T51V:T53C	539	8.4

III-1-5: The Crystal Structure of Q108K:K40L:T51V:T53S (KLVS) Bound to Retinal

To find out how much the polarity of the amino acid residues in the polyene chain effects the wavelength regulation, Dr. Wenjing Wang generated other mutants at the Thr53 position. In one of these mutants serine was introduced instead of Thr53. Serine is less polarizable than cysteine but is a more polar amino acid. The lmax of this mutant is 534 nm which is 5 nm blue shifted compared to Q108K:K40L:T51V:T53C (539 nm) with a pK_a of 8.7. Unlike the other crystal structures of hCRBPII mutants the crystal structure of Q108K:K40L:T51V:T53S (KLVS) was determined in the P2 space group with two molecules per asymmetric unit. The data were refined to 1.4Å with

crystallographic R_{free} = 22.77% and R_{fac} = 18.57%. The crystallographic data of this mutant is summarized in table V-1. This space group was unique for this mutant. The structure of the bound chromophore shows a major conformational difference when compared to Q108K:K40L, Q108K:K40L:T51V and Q108K:K40L:T51V in the ionone ring region with the ionone ring almost perpendicular to the polyene chain. In this structure the polyene chain remains in essentially the same place (Figure III-10-b). In this structure similar to Q108K:K40L:T51V and Q108K:K40L:T51V:T53C, the water molecule between Thr51 and Thr53 is removed completely. However despite the Q108K:K40L:T51V:T53C structure, where Cys53 adopts two different conformations, Ser53 has just one conformation, pointing toward the ionone ring area. The difference in conformation in the chromophore is the result of the T53S mutation. Ser53 makes a strong hydrogen bond with Gln38 and the result of this interaction moves the side chain of Gln38 from its original position that has been observed in all other structures including wt, and this movement leads to conformational changes in the structure of the chromophore, with the chromophore adopting a 6-*s*-*cis* conformation. Otherwise, as shown in Figure III-10-b, steric clash happens between the side chain of Gln38 and methyl5 of retinal. The conformational difference in the structure the chromophore indicates that this 6-s-cis conformation does not lead to a major hypsochromic shift. Since in Q108K:K40L:T51V:T53C the chromophore adopts the 6-s-*trans* conformation while Q108K:K40L:T51V:T53S the structure of the ionone ring of the chromophore is *cis* and just 4 nm hypsochromic shift has been observed.

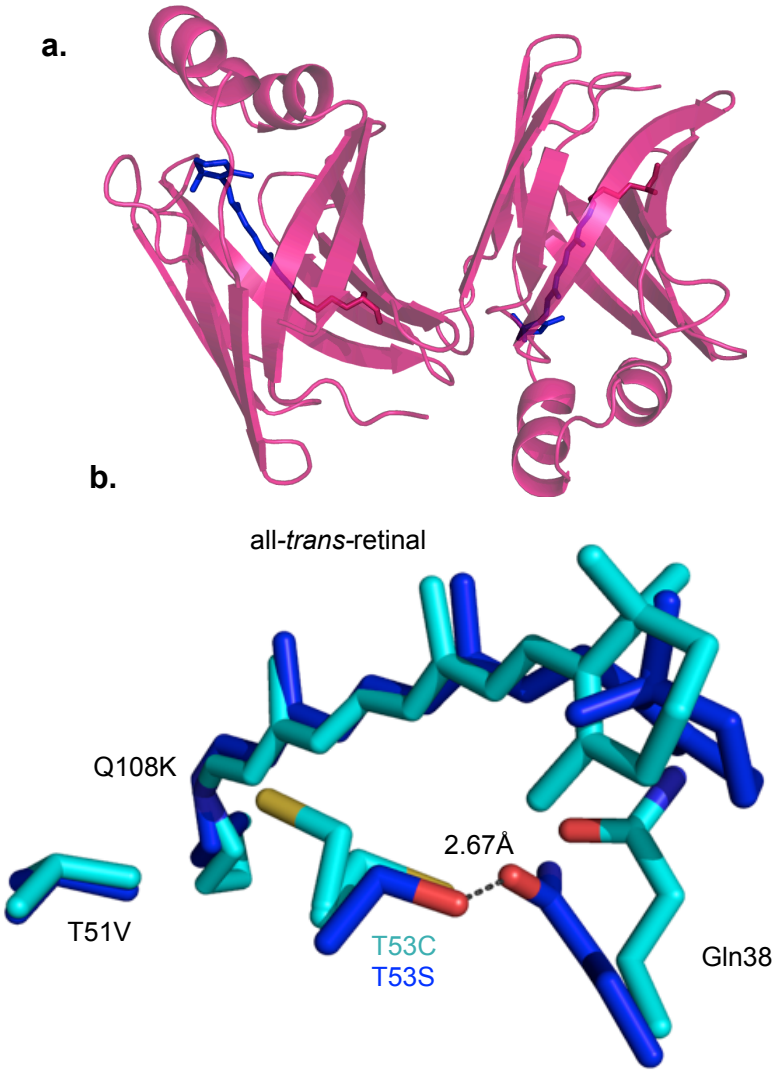


Figure III-11: a. The crystal structure of KLVS mutant at 1.4Å resolution bound to Retinal. **b.** The superimposed KLVS and KLVC mutants. As it is shown in this figure Ser53 makes a strong H-bond to Gln38, which causes changing the conformation in the chromophore.

In contrast to all other hCRBPII crystal structures, in the Q108K:K40L:T51V:T53S crystal structure the water network between Gln38 and Gln128 is not observed. The removing of this water molecule is a result of the T53S mutation. However the by the dislocation of Gln38 positions the hydroxyl group of this residue locates almost in the same position, as the water molecule was located. Removing of the water molecule is a result of the hydrogen bond between the Ser53 side chain and Gln38. This hydrogen bond is not observed. As shown in Figure III-12, the water molecule (W1) is 3.62Å away from the chromophore, while the carbonyl group of Gln38 of Q108K:K40L:T51V:T53S is 3.66Å away form all-*trans*-retinal. At the end of this chapter it will be explained that these water molecules are crucial both for wavelength regulation and also for controlling the pK_a of the PSB.

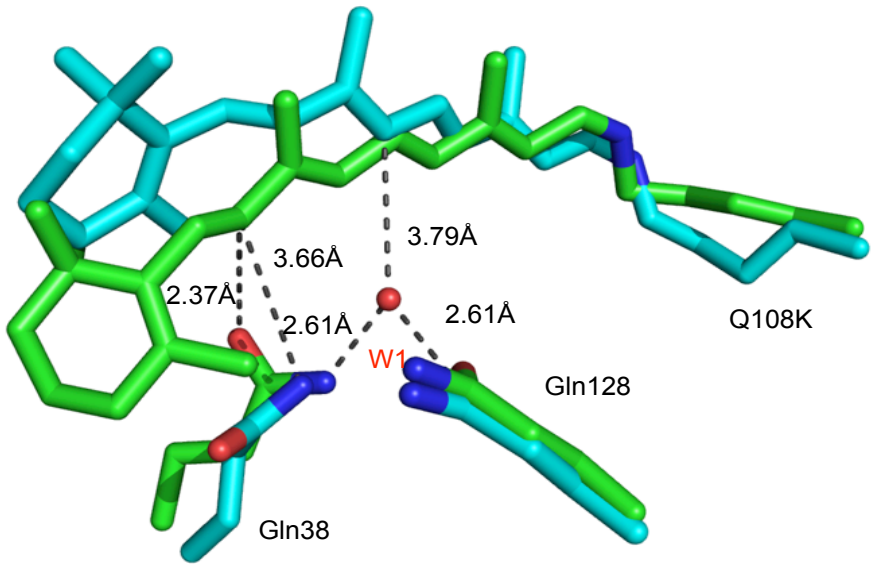


Figure III-12: The superimposed structure of KLVC (cyan) and KLVS (green). W1 (water molecule) belongs to the KLVC structure.

Removing one polar group (water) and replacing it with another polar residue (the side chain of Gln38) appears not to have a significant effect on the wavelength and pK_a. However new studies in our group (performed by Meisam Nosrati in Prof. James Geiger's group) illustrate clearly that the absence of a polar group leads to the most hypsochromic shift in the wavelength.

III-2: The Effect of Introduction of Tryptophan in the Middle of Polyene Chain and Rigidifying the Chromophore

The crystal structure of bovine rhodopsin indicates that several Tryptophan residues are present in the chromophore binding pocket.[4] Tryptophan is an electron rich and polarizable group and has the ability to encourage the delocalization of the positive charge of the PSB along the chromophore. Since introducing Cys53 leads to more red

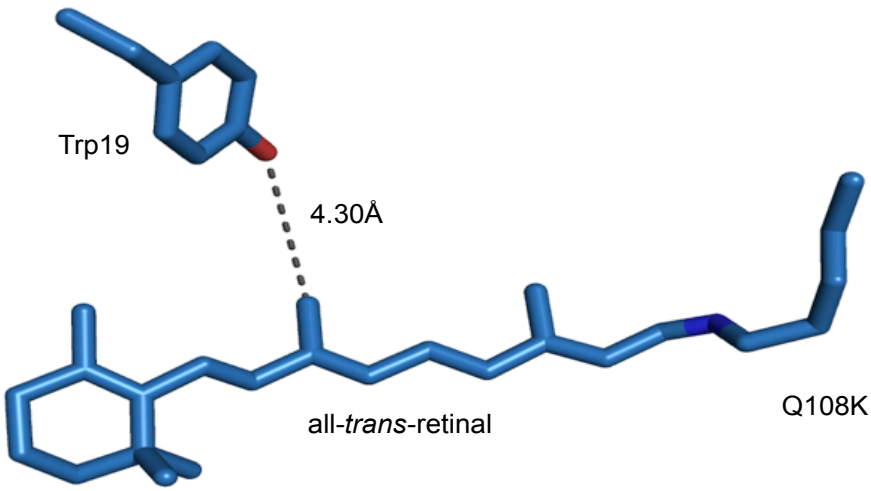


Figure III-13: The crystal structure of Q108K:K40L mutant with highlighted Y16W position

shifted absorbance, therefore Tyr19 which is near the middle of the polyene chain, about 5Å form the C9-methyl group and on the opposite side of the chromophore from Thr53 (Figure III-12), was mutated to . tryptophan and Y19W mutations were generated. The combination of this mutation with Q108K:K40L (508 nm) leads to a modest 5 nm red shift, similar to that seen for the T53C mutation. It seems that removing the more polar group (tyrosine) and replacing it with an electron rich and polarizable amino acid (tryptophan) can lead to a small red shifting.

Table III-4: Combination of Y19W mutation with the other mutations

Mutants	λ_{max}(nm)	pK_a
Q108K:K40L:Y19	508	7.9
Q108K:K40L:Y19W	513	8.9
Q108K:K40L:T51V:Y19W	537	9.3
Q108K:K40L:T51V:R58W:Y19	565	8.3
Q108K:K40L:T51V:R58W:Y19W	577	10.3

However adding this mutation to Q108K: K40L: R58W: T51V (565 nm) leads to 12 nm more red shift. The effect of the R58W mutation will be explained in the next section.

Even The effect of the Y19W mutation is not significant in short series of mutations still the red shifting can be related to two different parameters: 1.The planarity of the chromophore, which leads to more delocalization of the positive charge of the PSB. The crystal structure of all of the mutants with Y19W shows that retinal is more linear compared to the ones without this mutation like Q108K:K40L. 2. Removing the tyrosine

from this position that contains hydroxyl group and replacing of that with the softer polarizable group leads to better π-π interaction and more red shifted has been achieved.

III-2-1: The Crystal Structure of Q108K:K40L:T51V:Y19W:R58W (KLWW) Bound to Retinal

The crystal structure of Q108K:K40L:T51V:Y19W:R58W (KLVWW) was determined in P1 space group with two molecules per asymmetric units at 1.5Å resolution. The data were refined with R_{fac} = 23.16% and R_{free} = 31.04%. The λ_{max} of this mutant is 577 nm. As shown in table III-4 the effect of the Y19W was 12 nm in the presence of R58W. The overall structure of this mutant is the same as the other mutants. However the details of the binding pocket of this mutant shows a translation in the position of the chromophore. In this structure the “ionone ring part” of the chromophore is pushed more to the opposite side and leads to a more planar chromophore. Introduction of this mutation causes the chromophore to be pushed by the bulky side chain of Tryptophan therefore the free motion of the chromophore is reduced (Figure III-13). This effect rigidifies the position of the chromophore and may make the polyene chain and the double bond of the ring more conjugated. In all of the crystal structures that contain the Y19W mutation the translation in the location of the chromophore was observed. Table III-3 shows the λ_{max} absorbance in the different mutation. Another important difference between the structure with and without the Y19W mutation is the distance from Gln38. By applying the Y19W mutation the distance

between the chromophore and the side chain of Gln38 decreases and this residue possibly can help the delocalization of the positive charge and therefore lead to greater red shift (Figure III-14).

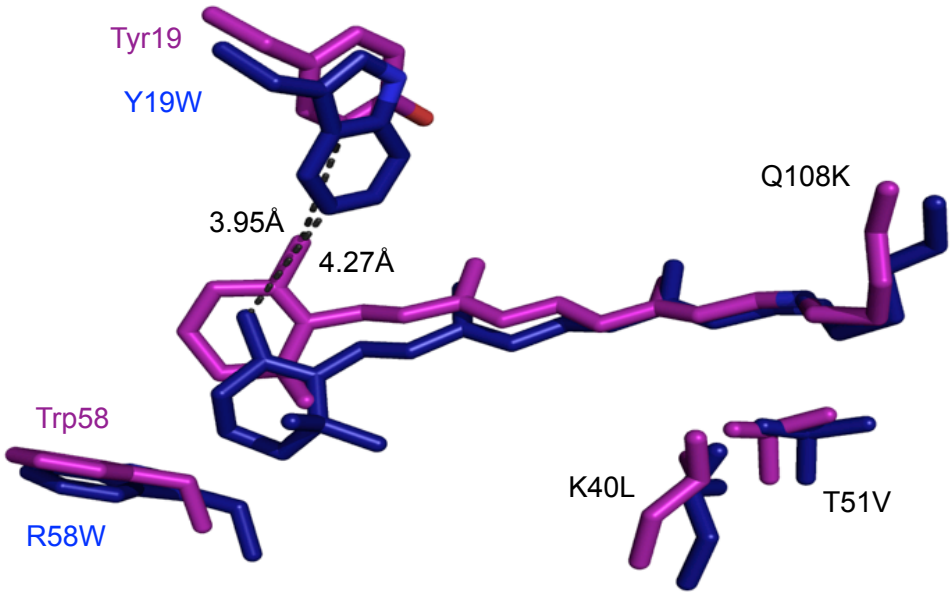


Figure III-14: The crystal structure overlay of KLVF (purple, 561 nm at 1.4Å) and KLVWW (dark blue, 577 m, at 1.5Å)

The details of the ordered water molecule in the binding pocket of this mutant illustrates that in one chain of this mutant (chain B) an electron density (2.4Å) away from the hydroxyl group of Thr53 is observed. This electron density could be a water molecule that makes a strong hydrogen bond with the hydroxyl group of Thr53.

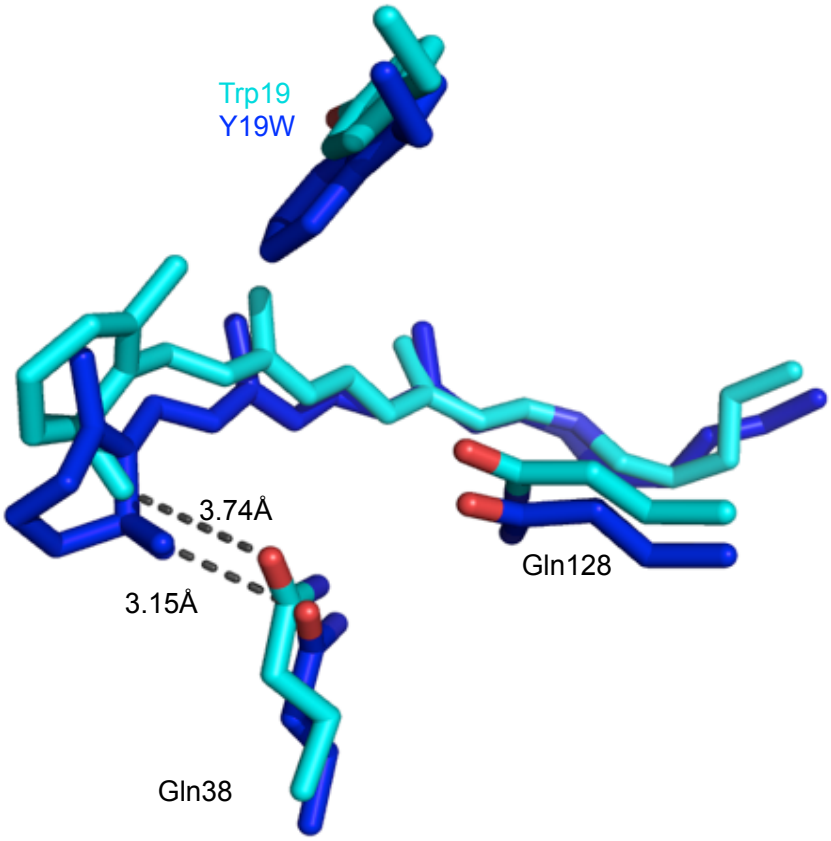


Figure III-15: The superimposed crystal structure of Q108K:K40L:T51V:R58W:Y19W (dark blue) and Q108K:K40L:T51V:R58F (cyan) with the highlighted Gln38 and 128 with the chromophore distance.

III-3: Effect of Mutation at R58 Position

The binding pocket of wt hCRBPII is covered by two short helices. In addition, because the guanidium group of Arg58 makes a hydrogen bond with the main chain of Ser55, the aliphatic side chain of Arg58 is stretched over the mouth of the binding pocket. These features make the binding pocket quite covered (Figure III-15).

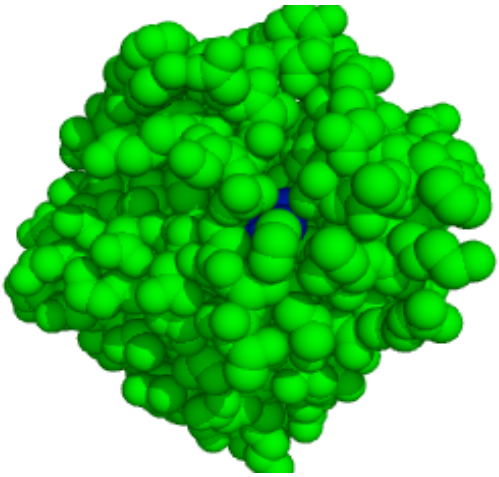


Figure III-16: The space-filling model based on the crystal structure of wt hCRBPII. The chromophore is shown in blue.

Arg58 is located at the “entrance” of the ligand to the binding pocket and mutation of this residue to aromatic amino acid residues leads to tremendous red shifting in the absorbance. As shown in table III-5 introducing the rigid and aromatic amino acids (Trp, Try and Phe) leads to red shifting absorbance, while mutation of this position to the other amino acids (polar and non polar) causes a hypsochromic shift (table III-5). In this

section the crystal structure of the different mutants that have R58 mutations and the effect of them in tuning of the wavelength will be described.

Table III-5: The effect of mutation of R58 on the wavelength

Mutants	λ_{max}(nm)	pK_a
Q108K:K40L:R58	508	8.3
Q108K:K40L:R58D	500	8.6
Q108K:K40L:R58A	499	8.4
Q108K:K40L:R58L	500	8.1
Q108K:K40L:R58Q	499	8.4
Q108K:K40L:R58E	500	8.1
Q108K:K40L:R58F	523	7.9
Q108K:K40L:R58Y	535	9.5
Q108K:K40L:R58W	527	7.9

III-3-1: The Crystal Structure of Q108K:K40L:T51V:R58F (KLVF) Bound to Retinal

The crystal structure of this mutant was determined in the P1 space group with two molecules per asymmetric unit. The data was refined to 1.4Å resolution with R_{fac} = 21.48% and R_{free} = 27.46%. Unfortunately the density of the chromophore was clear only in one of the two chains (chain B) but the density of retinal in chain A is very clear. The overall structure of this mutant is again the same as the others. Figure III-15-c

shows the superimposed structure of Q108K:K40L:T51V with Q108K:K40L:T51V:R58F. As mentioned above the side chain of Arg58 almost covers the “ mouth” of the protein in hCRBPII structures and retinal is buried inside the binding pocket (The structure of the wt hCRBPII protein has been explained in chapter II). Applying the R58F mutation makes the side chain of Phe57 flip from its original location (Figure III-15-c) and makes the binding pocket of hCRBP more exposed. The effect of R58F enhances the effect of other mutations. This effect is also observed in the R58W mutants but the enhancement factor is much more uniform (factor of 2).[5] As shown in Figure III-17, the position of Tyr58, Phe58 and Trp58 is similar and therefore the effect is similar for Tryptophan and phenyl alanine, while no enhancement has been observed for R58Y. However the mutation of Arginine, which has a positive charge, to phenylalanine or any other aromatic amino acid leads to more delocalization of the positive charge along the polyene chain and the result of this mutation is more red shifted absorbance (Table III-6). This effect has been also observed for the other mutations that have R58Y and R58W as well.

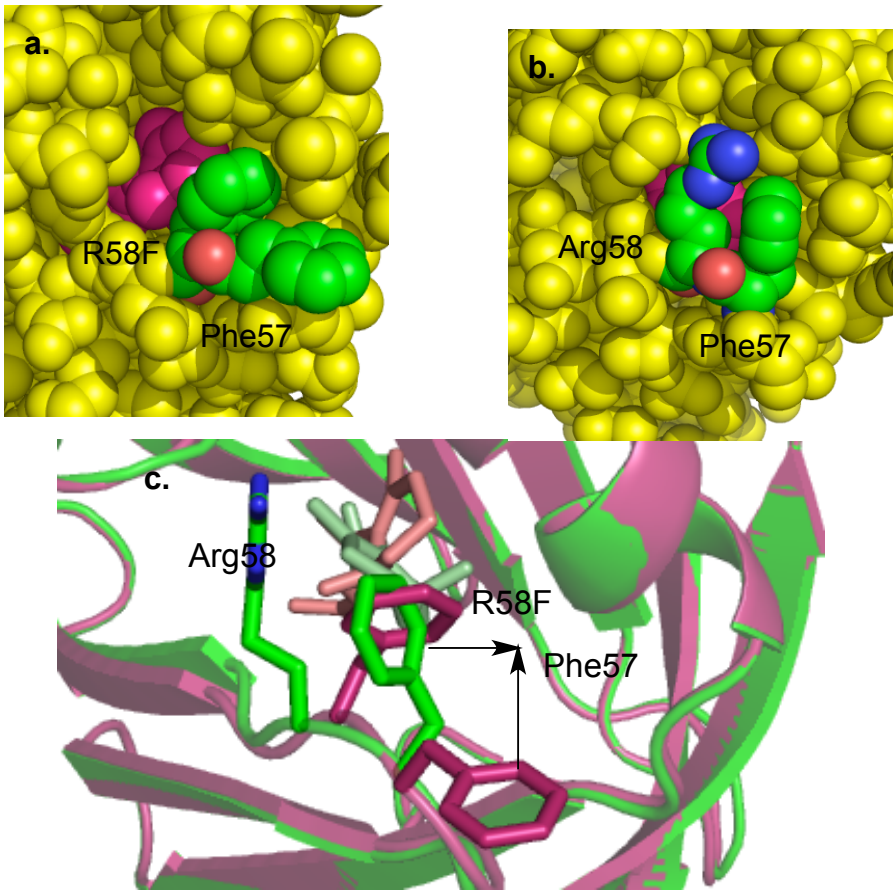


Figure III-17: a. The space filling model of Q108K:K40L:T51V. The chromophore is shown in pink. **b.** The space filling model of Q108K:K40L:T51V:R58F. **c**. The superimposed structure of Q108K:K40L:T51V (green) and Q108K:K40L:T51V:R58F (purple).

However the enhancement of the bathochromic shifts is almost two times in the longer series (combination with T51V:T53C:Y19W) with R58W mutations. This can be related to the size of the Tryptophan *vs.* phenylalanine side chain (Table III-6, III-7).

Table III-6: The effect of R58F mutation in the presence of the other mutations

Mutants	λ_{max}(nm)	pK_a
Q108K:K40L:R58F	523	8.6
Q108K:K40L:T51V	533	8.3
Q108K:K40L:T51V:R58F	561	8.5
Q108K:K40L:Y19W	513	8.9
Q108K:K40L:Y19W:R58F	537	9.3
Q108K:K40L:T53C	513	7.3
Q108K: K40L: T53C:R58F	537	8.5

Table II-7: The enhancement of red shifting based on R58F mutation

Mutant	λ_{max}(nm)	Protein Shift (nm)	R58F	Protein shift (nm)	Enhancement (nm)
Q108K:K40L	508	0	523	-----	-----
Q108K:K40L:T51V	533	25	561	38	13 (1.5x)
Q108K:K40L:T53C	513	5	537	14	9 (2.8x)
Q108K:K40L:Y19W	513	5	537	14	9 (2.8x)
Q108K:K40L:T51V:T53C	539	31	571	48	17(1.5x)

Table III-8: The enhancement of red shifting based on R58W mutation

Mutant	λ_{ax}(nm)	Protein Shift (nm)	R58W	Protein shift (nm)	Enhancement (nm)
Q108K:K40L	508	0	527	-----	-----
Q108K:K40L:T51V	533	25	570	43	18 (1.7x)
Q108K:K40L:Y19W	513	5	540	13	8 (2.6x)
Q108K:K40L:T53C	513	5	538	11	6 (2.2x)
Q108K:K40L:T51V:T53C	539	31	585	58	27 (1.9x)
Q108K:K40L:T51V:Y19W	537	29	577	11	21 (1.7x)
Q108K:K40L:T51V:T53C:Y19W	538	30	590	63	33 (2.1x)

The crystal structure overlaid of three mutants with R59F, R58W and R58Y is shown in the figure III-16. It is clear the positions of these three side chains are superimposable and therefore a similar enhancement was observed for R58Y and R58F

mutants. The electron rich and more rigid side chain of these three aromatic residues lead to more delocalization of the positive charge.

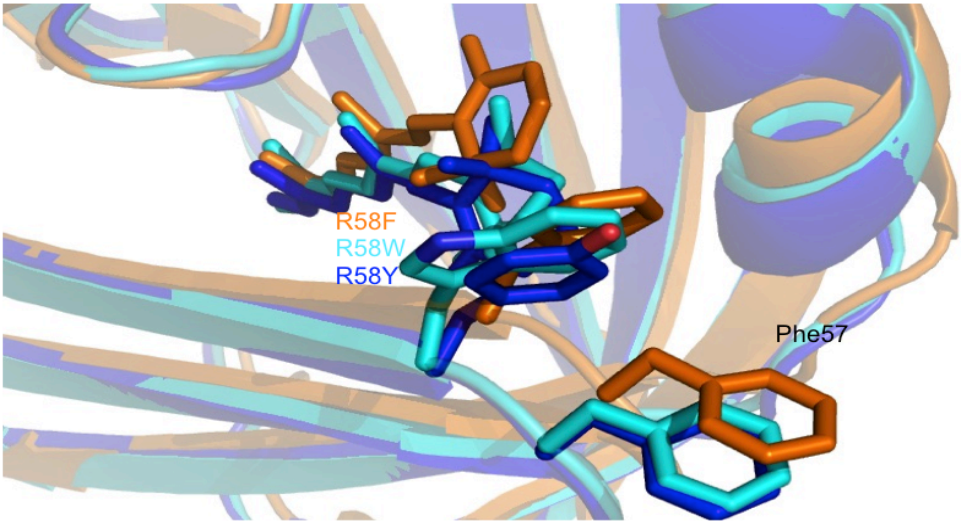


Figure III-18: The crystal structure overlay of KL:T51V:R58F (orange, 561 nm), KL:T51V:Y19W:R58Y (cyan, 577 nm and KL:T51V:Y19W:R58W (dark blue, 565 nm)

III-3-2: The Crystal Structure of Q108K:K40L:T51V:R58Y:Y19W (KLVYW) Bound to Retinal

The bound form of this mutant with all-*trans*-retinal absorbs at 565 nm while the parent mutant Q108K:K40L:T51V:Y19W, absorbs at 537 nm; thus the R58Y leads to a 28 nm bathochromic shift. This mutant also has the highest pK_a of all of the generated mutants (10.3). Though the overall conformation of the protein backbone did not change, the conformation of the chromophore changes from 6-s-*trans* to 6-s-*cis* (Figure III-17). In contrast to the previously described structures (Q108K:K40L:T53C) the ionone ring itself rotates while the polyene is similarly oriented relative to the other

structures. In the other word in Q108K:K40L:T53C mutant the plane of the polyene chain rotated about 90° while in Q108K:K40L:T51V:Y19W, ionone ring rotates about the C6-C7 single bond (Figure III-18). As mentioned above the mutants with R58Y, R58F and R58W have the higher pK_a when compared to mutants that do not have an aromatic group at position 58. Among all the R58 mutants, the ones with R58Y have the highest pK_a. It seems that the hydroxyl group of the side chain makes it even more electronegative compare to the other aromatic residues, increasing the pK_a. It is clear from these studies that stabilization of the protonation state can be achieved by manipulation of the polyene environment far from the site of protonation. Though not definitive, comparison of the Q108K:K40L:T51V:Y19W:R58Y structure to others may proved an explanation for the cis conformation of the ionone ring. The Phe16 side chain is moved in this mutant (Figure III-18.b). A explained above in section III, the position of two amino acid residues (Phe16 and Leu77) lock the ionone ring of the chromophore in the *trans* conformation. Study on another crystal structure Q108K:K40L:T51V:Y19W:R58W, shows that the chromophore of has trans conformation same as the most of hCRBPII mutants. Since these two mutants are just different in R58 mutation therefore the comparison has been done on these two structures. In Q108K:K40L:T51V:Y19W:R58Y structure the side chain of Phe16 is moved (~ 0.5 Å) therefore there is more space available for the rotation around the C6-C7 single bond. The detail is shown Figure III-19.b.

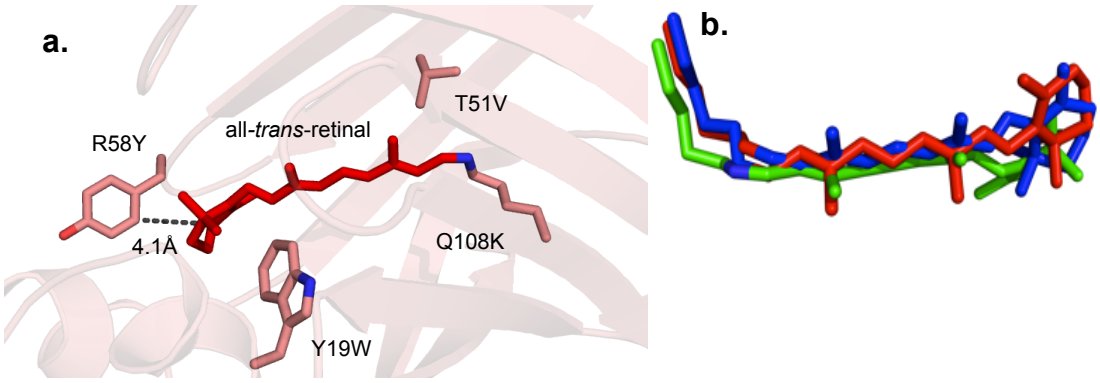


Figure III-19: a. The crystal structure of Q108K:K40L:T51V:Y19W:R58Y mutant with all-*trans*-retinal. The conformation of the ionone ring of chromophore is 6-s-*cis*. **b**. The overlaid structure of all-*trans*-retinal in Q108K:K40L (blue), Q108K:K40L:T53C (red) and Q108K:K40L:T51V:Y19W:R58Y (green).

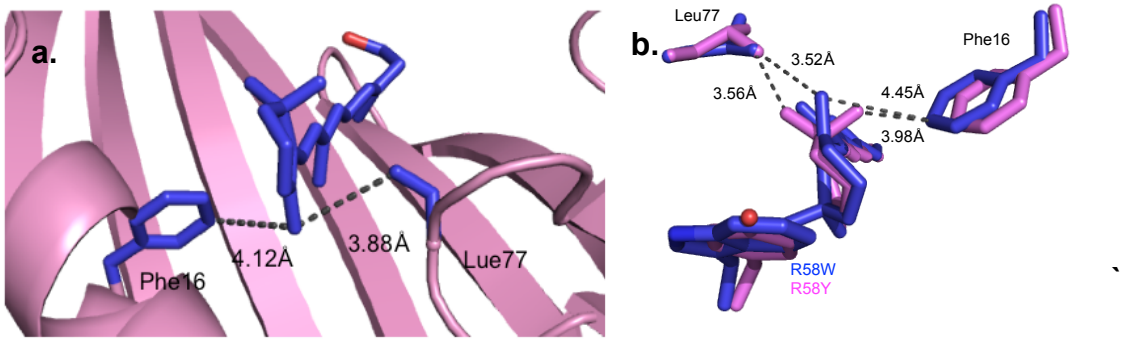


Figure III-20: a. The crystal structure of wt hCRBPII with all-*trans*-retinol. **b.** The superimposed crystal structure of KLVR58WY19W (pink) with KLVR58YY19W (blue).

However the overall structure of the protein did not shows any different at all and even Trp and Tyr side chains at 58 position are exactly in the same place. However it is not clear that why the side chain of Phe16 translated from the original position but we can conclude that in presence of more space, the C6-C7 single bond freely can rotates and therefore chromophore adopts the structure which is more stable (6-s-*cis)* (Figure III-18). Even though the chromophore in Q108K:K40L:T51V:Y19W:R58Y mutant adopts with 6-s-*cis* conformation it absorbs at 565 nm while the Q108K:K40L:T51V:Y19W:R58W mutant which the ionone ring of the chromophore adopts the trans conformation 577 nm (12 nm) more red shifted. As it was predicated applying R58W mutation leads to more bathochromic shift. It seems that changing from 6-s-*trans* to 6-s-*cis* does not lead to significant hypsochromic shift.

As mentioned above introducing of the aromatic residues (Tryptophan, Tyrosine and phenyl alanin) at R58 position leads to the red shifted absorbance and at the same time these mutations enhanced the effect of the other mutations (T51V, T53C, Y19W). This enhancement effect in R58F mutation is around 2x (Table III-7) of the original red shifting as well as R58W. On the other hand, introducing R58Y despite R58F and R58W does not enhance the effect of other mutants and the red shifting effect is additional. It seems that by mutation of Arg58 to tryptophan and phenyl alanine (both are not polar residues), the hydrophobicity at the binding pocket of protein increase, in the other word, creation of hydrophobic environment around the protein binding pocket can lead to isolating of binding cavity from the bulk solution. Tyrosine does not have the same effect since the hydroxyl group can make the protein “entrance” more polar. It seems that by increasing the size of the aromatic residue at the “ month”, of the protein, the

binding pocket is covered more efficiently and the chromophore is more isolated form the aqueous solution, decreasing the polarity of the binding pocket. Introduction of the R58 mutation to the aromatic residues in the short series of mutants (Q108L:K40L) for tryptophan and phenyl alanine shows a similar effect.

Table III-9: The enhancement of the red shifting based on R58Y mutation

Mutant	λ_{max}(nm)	Protein Shift (nm)	R58Y	Protein shift (nm)	Enhancement (nm)
Q108K:K40L	508	0	535	-----	-----
Q108K:K40L:T51V	533	25	563	28	3 (1.1x)
Q108K:K40L:T53C	513	5	540	5	0 (1.0x)
Q108K:K40L:T51V:Y19W	537	29	565	30	1 (1.03x)
Q108K:K40L:T51V:T53C	539	31	576	30	-1 (0.96x)

However the R58Y variant leads to more red shifted absorption in the short series (table III-8), while in the longer series the tryptophan variant gives the larger bathochromic shift, which might related to the size of the side chain. It seems that in the shorter series the electron rich Tyr hydroxyl group plays an important role in regulation, while in the more red shifted, longer series of mutants the better solvent sequestering resulting from the larger Trp residue dominates.

Few water molecules have been observed in the binding pocket of hCRBPII and these water networks have crucial effect on the modulation the wavelength in hCRBPII. The dielectric constant of water is 78, while by changing the environment to more hydrophobic area this number decreases dramatically and reaches to 2 to 10.

Inspecting the crystal structure of mutants with R58W mutation and comparison of them with the structures that R58 position is not mutated illustrates that the well ordered water molecules in the binding pocket did not change at all, however the dielectric constant of them can be affected by the polarity of the binding pocket. (Covering the binding pocket with bulkier hydrophobic groups like Tryptophan or/and creating the more hydrophobic environment in the entrance of protein binding site can reduce the effect of dielectric constant of water molecules and as the result enhance the effect of the more non-polar mutations).

Mutation of R58 to other amino acids such as R58E, R58A or R58L results in blue shifted spectra. It seems that mutation of Arg58 to any other shorter amino acids open the binding pocket more and therefore the blue shifting will be resulted (Table III-5).

III-3-3: The Crystal Structure of Q108K:K40L:T51V:T53C:R58L (KLVCL) Bound to Retinal

The crystal structure of Q108K: K40L:T51V:T53C:R58L was determined in the P1 space group with two molecules per asymmetric units. The λ_{max} of the bound form of this mutant with all-*trans*-retinal absorbs at 531 nm while the λ_{max} of Q108K:K40L:T51V:T53C is 8 nm more red shifted (539 nm). Inspection of the crystal structure of this mutant reveals interesting information. Compared to the crystal structure of Q108K:K40L:T51V:T53C the chromophore is twisted in this mutant.

However the crystal structure of this mutant provides structural details that support the idea that the conformational of the chromophore does not result in a blue shift in this mutant. Figure III-21 shows the structural details of the binding pocket of this

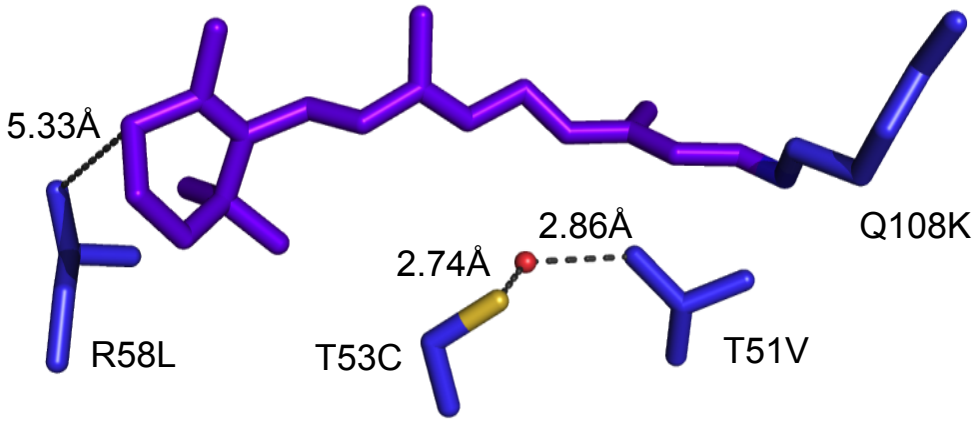


Figure III-21 The crystal structure of penta mutant Q108K: K40L: R58L: T51V: T53C with the highlighted residues.

mutant. Surprisingly and in contrast to the other structures with both T51V and T53C mutations (crystal structure of Q108K:K40L:T51V:T53C mutant), in this structure the water molecule between residues 51 and 53 is present. However in none of the other crystal structures with both T51V and T53C mutations structures) is this water molecule observed. The presence of this water molecule may explain the relatively blue shifted absorbance of this mutant compared to Q108K:K40L:T51V:T53C and it has almost the same I_{max} as Q108K:K40L:T51V (533nm versus 531 nm for KLVCR58L). It is not clear why the water molecule is still present in this mutant since neither the valine nor cysteine side chains are capable of, or in position to make a hydrogen bond with the

water and still the water molecule has the perfect hydrogen bond distance from them. It seems that because this water is able to make only one hydrogen bond, it may rotating and does not fix in the same position and therefore the dipole moment of this water is not pointing in a single direction and it's electrostatic thereby partially naturalized. This data is the evidence that the dielectric constant of water is reduced in the hydrophobic environment so this water, which is beside the valine, does not have the same dialectic constant as is close to the threonine and by decreasing the dielectric constant of water the polarity of them decrease as well.

It can be concluded that modulation of the chromophore spectrum by the protein environment is not only effected by the polarity of the side chain, but also by the ordered, and perhaps even disordered water molecules. Therefore as was shown above the only mutations that show significant red shifted spectra, are those that either directly influence ordered water molecules, or sequester the binding pocket from the external aqueous solvent.

III-3-4: The Crystal Structure of Q108K:K40L:T53V: R58W Bound to Retinal

This mutant has been generated by Dr. Wenjing Wang in Professor Borhan group to test the effect of different mutations in combination with the R58W mutation. Q108K:K40L:T53V:R58W absorbs at 541 nm which is almost the same as Q108K:K40L:T53C:R58W (540 nm). These two mutants are almost 20 nm blue shifted relative to Q108K:K40L:T51V:R58W. The crystal structure of Q108K:K40L:T53C shows that the water molecule between Thr51 and Thr53 is in exactly the same location,

between T51 and T53, the expectation was that the structure of Q108K:K40L:T53C:R58W would show the same ordered water molecule, since the R58W mutation is on the other end of the binding pocket.

Table III-10: Mutation of Thr51 and Thr53 to more hydrophobic residue

Mutant	λ_{max}(nm)	pK_a
Q108K:K40L:T51V:R58W	565	8.4
Q108K:K40L:T53V:R58W	541	9.0
Q108K:K40L:T53C:R58W	540	8.3

As previously mentioned the T51V mutation results in the absence of this ordered water molecule in all structures so far determined, and results in a large red shift in each case. The crystal structure of Q108K:K40L:T53V:R58 was determined in the P1 space group with two molecules per asymmetric unit. The crystallographic data was refined to 1.53Å resolution with crystallographic R_{fac} = 24.63% and R_{free} = 32.82 %. The electron density of the ligand was very clear in one chain (chain B), but missing in chain A. The structure of this mutant shows that the water molecule between Val53 and Thr51 is, surprisingly, not present. Further the Thr51 side chain has rotated relative to KL, and the hydroxyl group makes a water-mediated interaction with the PSB. In all the other structures with Thr51 present, the hydroxyl group is pointing toward the ionone ring and therefore cannot make a direct interaction with the PSB. This conformation would appear to be correlated with the absence of the water molecule, which allows the hydroxyl group of Thr51 to make an interaction with the PSB instead of the water

molecule lodged between residues 51 and 53 (Figure III-20). Presumably the increase in hydrophobicity of the region results in rotation of the hydroxyl group to the more hydrophilic environment in the vicinity of the iminium.

As shown in Table III-9, the pK_a of this mutant is higher than Q108K:K40L:T51V:R58W and Q108K:K40L:T53C:R58W. The increase in pK_a can be the result of the extra water mediated interaction with the PSB, which make the hydrogen bond to the PSB and therefore PSB is more stable and the pK_a increased. The presence of two water molecules interacting with the PSB is reminiscent of the bacteriorhodopsin structure, which also shows two water-mediated interactions with the PSB resulting in higher pK_a.

The flipping of the Thr51 side chain and presence of two water molecules interacting with the PSB causes almost 20 nm hypsochromic shift compared to Q108K:K40L:T51V:R58W. It is interesting to note that the KLT53V mutant is blue shifted 5 nm relative to KL to 508 nm. This hypsochromic shift may be caused by a similar water mediated interaction between T51 and the iminium in this structure, resulting in more localization of positive charge at the iminium. Surprisingly Q108K:K40L:T53V:R58W and Q108K:K40L:T53C:R58W have the same absorbance (544 nm). Comparison of the absorbances of KLT53V with KLT53VR58W shows an enhancement caused by the introduction of 58W that is much larger than that seen with the other mutants in the series (Table 9 include the KLT53V and KLT53VR58W in table).

Combination of all of the mutations so far discussed leads to the most red shifted mutant discussed thus far. These two mutants are hepta mutants

Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (absorbs at 591 nm) and Q108K:K40L:T51V:T53S:R58W:T29L:Y19W (600 nm).

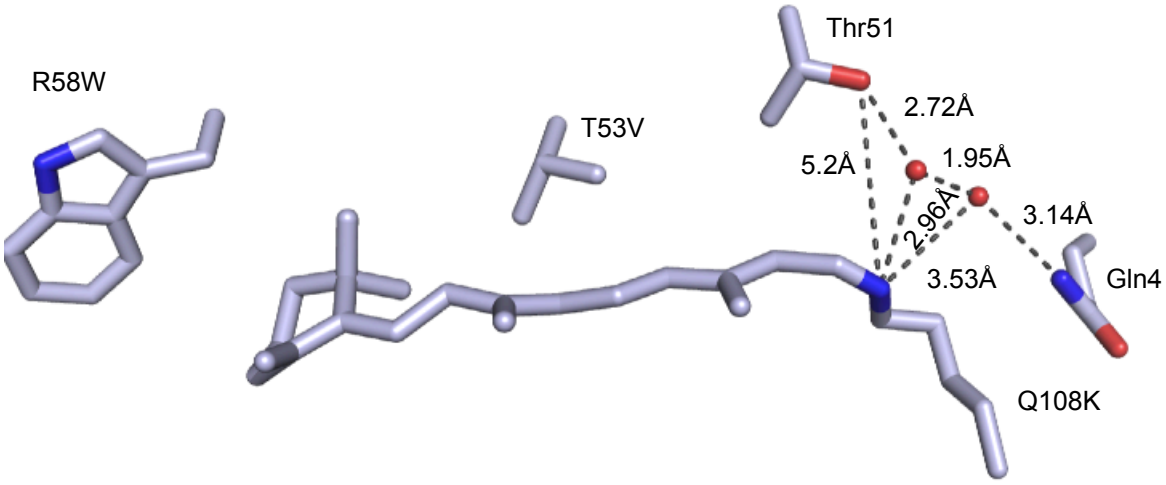
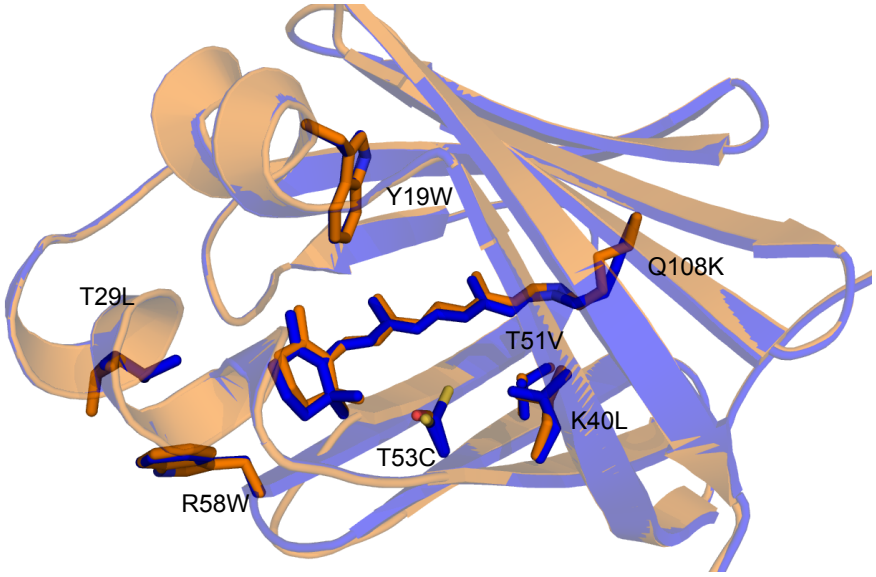


Figure III-22: The crystal structure of Q108K:K40L:T53V:R58W mutant with all-*trans*-retinal.

The crystal structures of these two mutants were determined in P1 space group with two molecules per asymmetric unit. The crystal structure of the hepta mutant Q108K:K40L:T51V:T53C:R58W:T29L:Y19W was refined at 1.5Å resolution and the electron density of the chromophore is clear on one chain. Figure III-23 shows the overlaid structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (591 nm) and Q108K:K40L:T51V:T53S:R58W:T29L:Y19W (600 nm)



FigureIII-23: the superimposed structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (blue) and Q108K:K40L:T51V:T53S:R58W:T29L:Y19W (Orange).

As shown the two proteins show little structural difference. However in the structure of Q108K:K40L:T51V:T53S:R58W:T29L:Y19W, Ser53 adopts only one conformation which is pointing toward the ionone ring, while Cys53 in Q108K:K40L:T51V:T53S:R58W:T29L:Y19W adopts two conformations. The crystal structure of these two mutants shows the chromophore to be relatively flat and locked in position, with little difference in conformation in the two structures. However the structural details of the water network inside the binding pocket of Q108K:K40L:T51V:T53S:R58W:T29L:Y19W is different from Q108K:K40L:T51V:T53C:R58W:T29L:Y19W in one chain (chain A) while in chain B is the same as the other structures (Figure III-24). In the structure of

Q108K:K40L:T51V:T53S, Ser53 makes a direct hydrogen bond to Gln38, while in Q108K:K40L:T51V:T53S:R58W:T29L:Y19W, the interaction between Gln38 and Ser53 is water mediated. As mentioned above and shown in the crystal structure of wt hCRBPII, Q108K:K40L and Q108K:K40L:T53C, a water molecule bridges Thr51 and Thr53. This water restricts the hydroxyl group of Thr53 to a single conformation. However, this water molecule is removed in Q108K:K40L:T51V:T53S and Q108K:K40L:T51V:T53S:R58W:T29L:Y19W. The hydroxyl group of Ser53 is then free to make an alternate hydrogen bond to Gln38. On the other hand, the sulfur group of Cys53 has two different conformations while the hydroxyl group of Thr53 only points toward the PSB region. Other than this conformational difference between these two hepta mutant structures, the water molecule between Gln38 and Gln128 is present in chain A as well. Another structural difference between these two chains is the conformation of Trp58. In one chain Trp58 has the same conformation as Q108K:K40L:T51V:T53C:R58W:Y19W:T29L:Y19W and the other chain this Trp flips to cover the binding pocket more. Figure III-25 shows the details of the structure inside the binding pocket of the protein. It is not very clear that what exactly causes the 9 nm more red shifted. Absorbance in Q108K:K40L:T51V:T53S:R58W:T29L:Y19W mutants, since the two chains are only slightly different. The differences in ordered water between these two structures is the most likely cause.

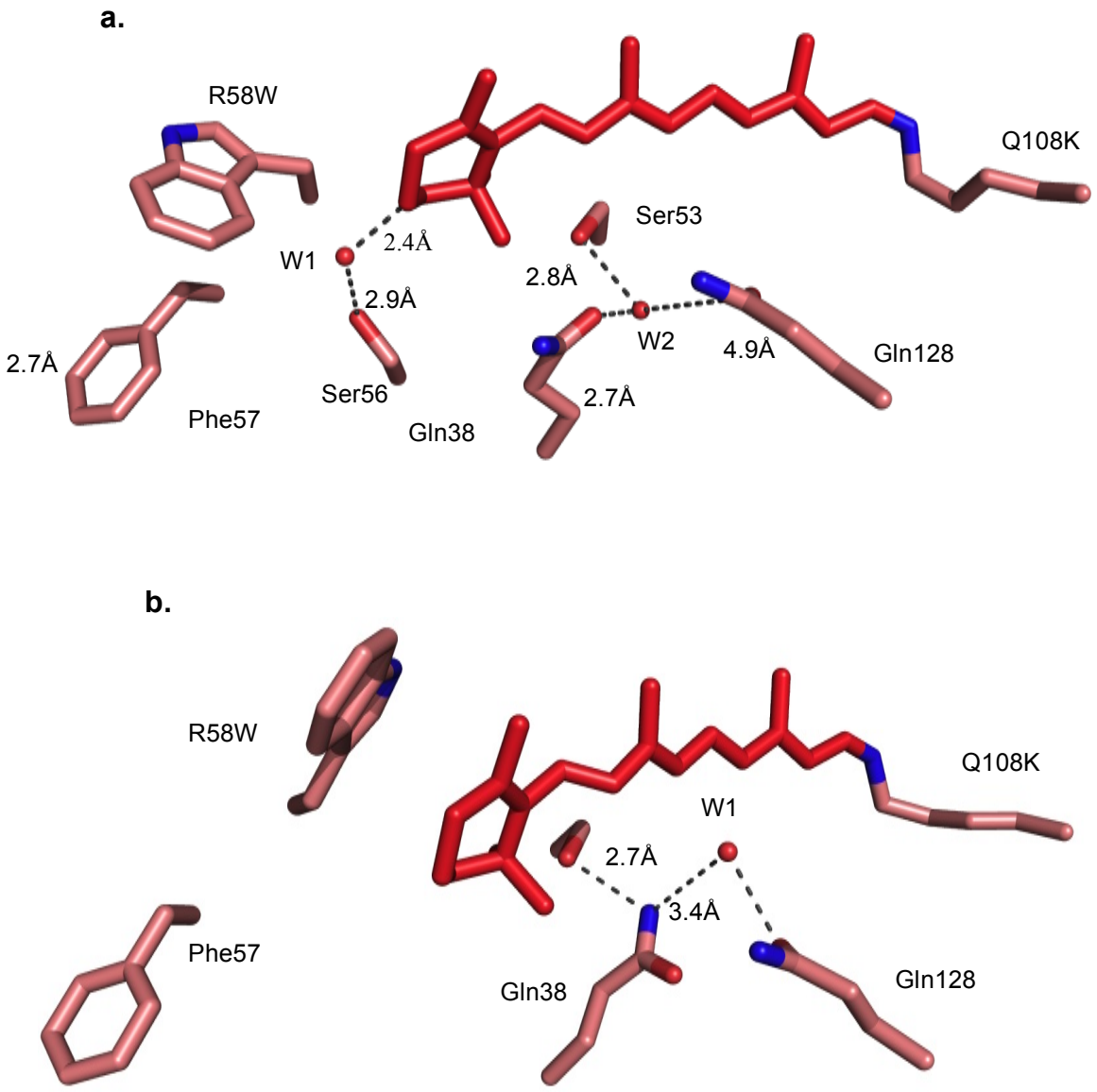


Figure III-24: a. The crystal structure of chain A of The binding pocket of Q108K:K40L:T51V:T53S:R58W:T29L:Y19W mutant. **b.** The details of the binding of chain B of the same mutant. As it is shown in this figure the position of Gln38 and Trp58 are different in two chains. Trp58 position in chain A is same as Q108K:K40L:T51V:T53C:R58W:T29L:Y19W mutant.

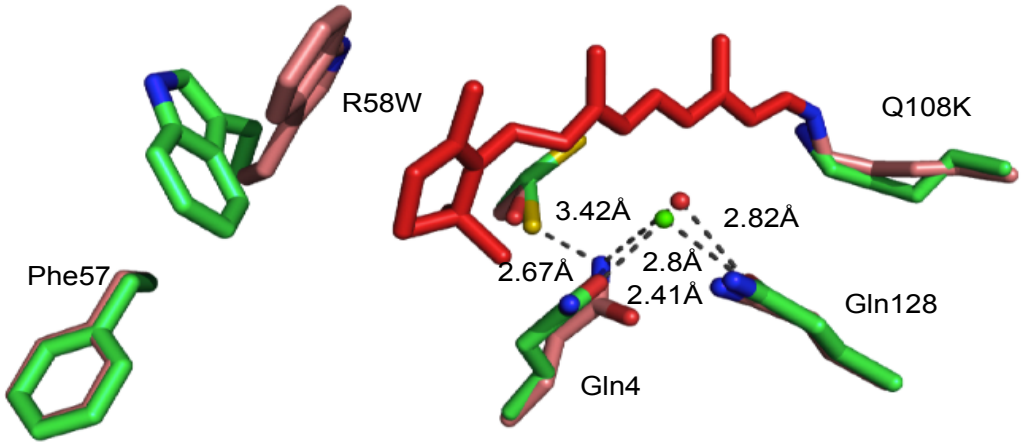


Figure III-25: The superimposed crystal structure of Q108K:K40L:T51V:T53S:R58W:T29L:Y19W (light pink) and Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (green).

III-3-5: The Crystal Structure of Apo-Q108K:K40L:T51V:T53C:R58W :T29L

The crystal structure of this hexa mutant has been determined in the $P3_12$ space group and despite most of the bound hCRBPII mutants with retinal this structure has one molecule per asymmetric unit. The unbound crystal structure is needed to compare whether conformational changing happens when the chromophore bound to the protein or not, since the goal is to study the effect of the different mutations on the wavelength regulation, therefore at least one crystal structure of one apo protein was necessary. The structural details of this protein illustrate some information, which are important for understanding the effect of each mutation. Figure III-26 shows the highlighted mutated residues of this protein.

As is shown by T51V and T53C mutations, the water molecule between them is removed and the side chain of Cys53 has only one conformation, pointing toward the "mouth" of the binding pocket and one water molecule has been observed in the binding pocket between Gln4 and Lys40. This water molecule is the one, which is responsible for stabilizing the pK_a and *cis*-conformation of the iminium. It will be explained further that removing this water molecule leads to a reduced pK_a and also a change in the conformation of the Schiff base from *cis* to *trans*. The overall structures of the apo hexa and holo (Q108K:K40L:T51V:T53C:R58W:Y19W:T29L) do not change significantly.

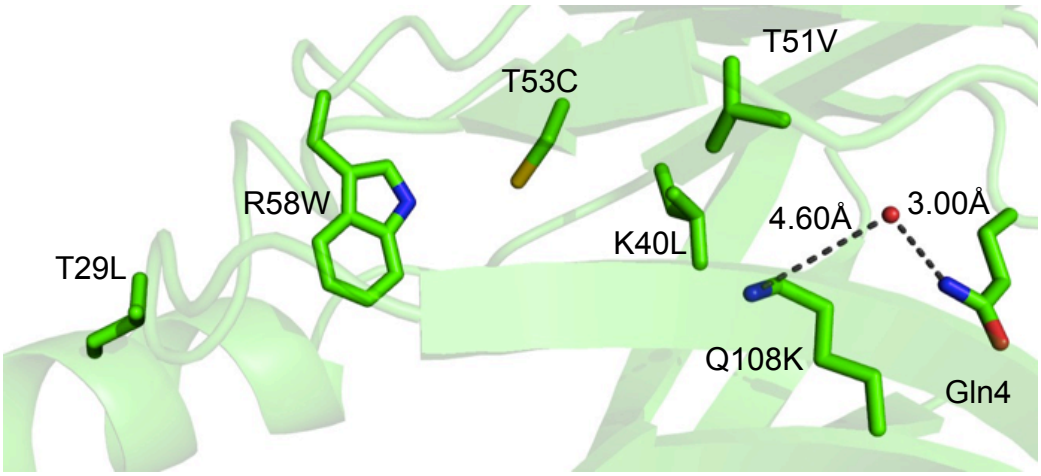


Figure III-26: The crystal structure of apo-Q108K:K40L:T51V:T53C:R58W:T29L with the highlighted mutated residues.

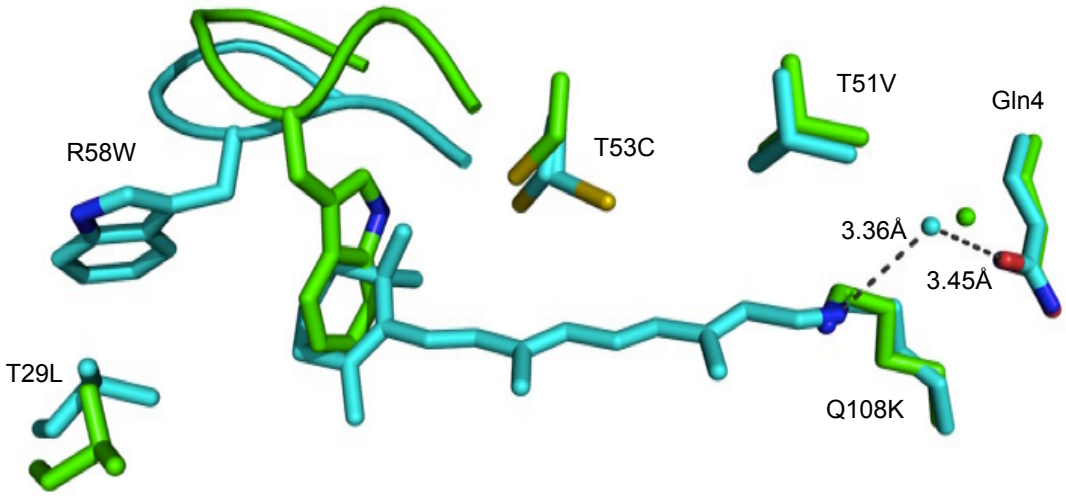


Figure III-27: The superimposed structure of apo-Q108K:K40L:T51V:T53C:R58W:T29L (green) and holo-Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (cyan)

Figure III-27 shows the superimposed structure of these two with RMS = 0.963Å for 133 amino acids residues. The main structural difference that has been observed is the conformation of the loop between two beta-strands. Trp58 is situated in this loop. When the protein is not bound to all-*trans*-retinal, the Trp58 side chain goes deeper inside the binding pocket, while in the bound protein most of the binding pocket is occupied by the ligand and therefore Trp58 does not have sufficient space inside the binding Pocket and is flipped out of the binding pocket. Otherwise there would be a clash between the chromophore and Trp58.

III-3-6: Conclusion of Mutation on Ar58

As mentioned above mutation of Arg58 to just to aromatic amino acids (Trp, Phe and Tyr) cause red shifting. However only Trp and Phe can enhance the effect of the other mutation in the binding pocket, while Tyr does not have this effect. The mutation data shows that R58E mutation despite the theories leads to more blue shift (Table III-10), while R58L does not have important effect. It seems that by mutation of Arg58 to more hydrophobic amino acid, we create a “hydrophobic tunnel” in front of the entrance of the protein and with mechanism the binding pocket isolates from the bulk solution. R58E mutation leads to more blue shift, because Glutamic acid is charged and compare to Arginine is shorter. This hydrophilic environment can increase the water exchange rate in the binding pocket and more blue shift will be resulted. Inspecting the structure of the mutation of R58F, R58Y and R58W show that Phe57 side chain flipped back to the bulk solution while in none of the other structures this effect have been observed. In the other word by flipping back the phenyl alanin in the bulk solution and creating this hydrophobic tunnel the red shifting will be resulted.

Table III-11: Mutation of Arg58 to charged and nonpolar amino acid

Mutants	λ_{max}(nm)	pK_a
Q108K:K40L:T51V:T53C:R58	539	7.4
Q108K:K40L:T51V:T53C:R58L	531	7.1
Q108K:K40L:T51V:T53C:R58E	517	8.3

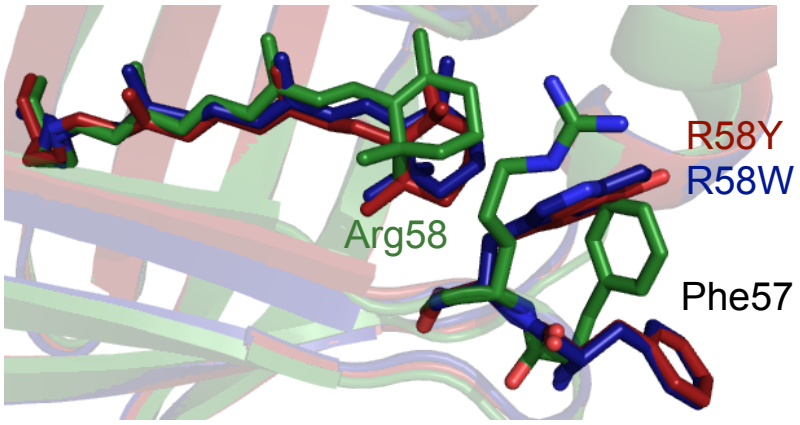


Figure III-28: The superimposed structure of KL (508 nm, green) KL:T51V:R58W:Y19W (577 nm, blue) KL:T51V:R58W:Y19W (565 nm, red).

III-4: Toward the Most Red Shifted Mutant with *Cis*-Iminium Conformation

So far, based on the crystal structures and rational mutagenesis of hCRBPII, we could show that different parameters seem to have crucial roles in the modulation of the wavelength in the hCRBPII rhodopsin mimic. We can categorized these parameters as:

1-The polarity of the binding pocket is important and removing the polar groups around the chromophore leads to more red shifting due to delocalization of the PSB charge.

2-The mutations that have the most dramatic effect on the modulation of the wavelength to more bathochromic absorbance are the ones that affect the water networks in the binding pocket.

3- Isolation of binding pocket from the bulk solution by introduction of large aromatic residue at the “entrance” of protein binding pocket.

4-the most bathochrmoic mutants are the ones where the chromophore is more linear and more locked up in the specific positions, therefore it seems that making the binding pocket more narrow and preventing the free rotation of all-*trans*-retinal is essential.

Based on this information, Dr. Wenjing Wang in Professor Borhan's group designed another mutant, which was the most red shifted one in the series of cis-iminum structures.

The crystal structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W reveals that the binding pocket of hCRBPII is not efficiently covered even by R58W mutation.

The amino acid position considered for covering the binding pocket more efficiently was Ala33. However modeling the A33W mutation in silico shows that Trp33 has two potential clashes: Trp58 or the ionone ring of the chromophore. However Tanya Berbasova in Professor Babak Borhan's group introduced the same mutation in another, homologous rhodopsin mimic, CRABPII, and observed a significant red shift. Based on these data Dr. Wang introduced the same mutation to Q108K:K40L:T51V:T53C:R58W:T29L:Y19W. The octa mutant Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W absorbs at 606 nm and leads to 15 nm more red shift, while introduction of this mutant in proteins with R58 caused a hypsochromic shift.

Table III-12: The protein shift caused by A33W

Mutant	λ_{max}(nm) with A33	λ_{max}(nm) with A33W	Protein shift of A33W (nm)
Q108K:K40L	508	498	**-10**
KL:T51V:T53C	539	533	**-6**
KL:T51V:Y19W	537	522	**-15**
KL:T51V:T53C:Y19W	538	522	**-16**
KL:T51V:T53C	539	533	**-6**
KL:R58F:Y19W	537	543	**6**
KL:T51V:T53C:R58W:T29L:Y19W	591	606	**15**

To understand the mechanism of how the A33W mutation leads to red shifting, the crystal structure of a mutant with the combination of A33W and R58W was necessary.

III-4-1: The Crystal Structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W Mutant Bound to Retinal

The crystal structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W also crystallizes in the P1 space group with two molecules per asymmetric unit. The crystallographic data was refined to 1.4Å resolution with R_{fac} = 20.15% and R_{free} = 24.15%. The electron density of the chromophore is not very well defined and is only seen in one of the two chains. The overlaid structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W and

Q108K:K40L:T51V:T53C:R58W:T29L:Y19W is show in Figure III-26.c. The superimposed structure of these two mutants shows the overall structure of protein did not change and even introduction of a large residue like Trp did not effect the position and conformation of the Trp33-containing helices. The RMSD of the two structures is 0.398 Å for 133 amino acids. However the crystal structure reveals that the chromophore is translated slightly due to the Introduction of the A33W mutation. The bulky side chain of Trp33 moved the chromophore more to the Y19W position and this residue also shifted 0.6 Å (figure III-28.c). Figure III-28.b illustrates a space-filling representation of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W with the chromophore shown in red. As predicated, A33W isolates the binding pocket more efficiently and the binding pocket of the chromophore is more embedded. The crystal structure shows the different conformation of the side chains of Trp58 and Arg58.

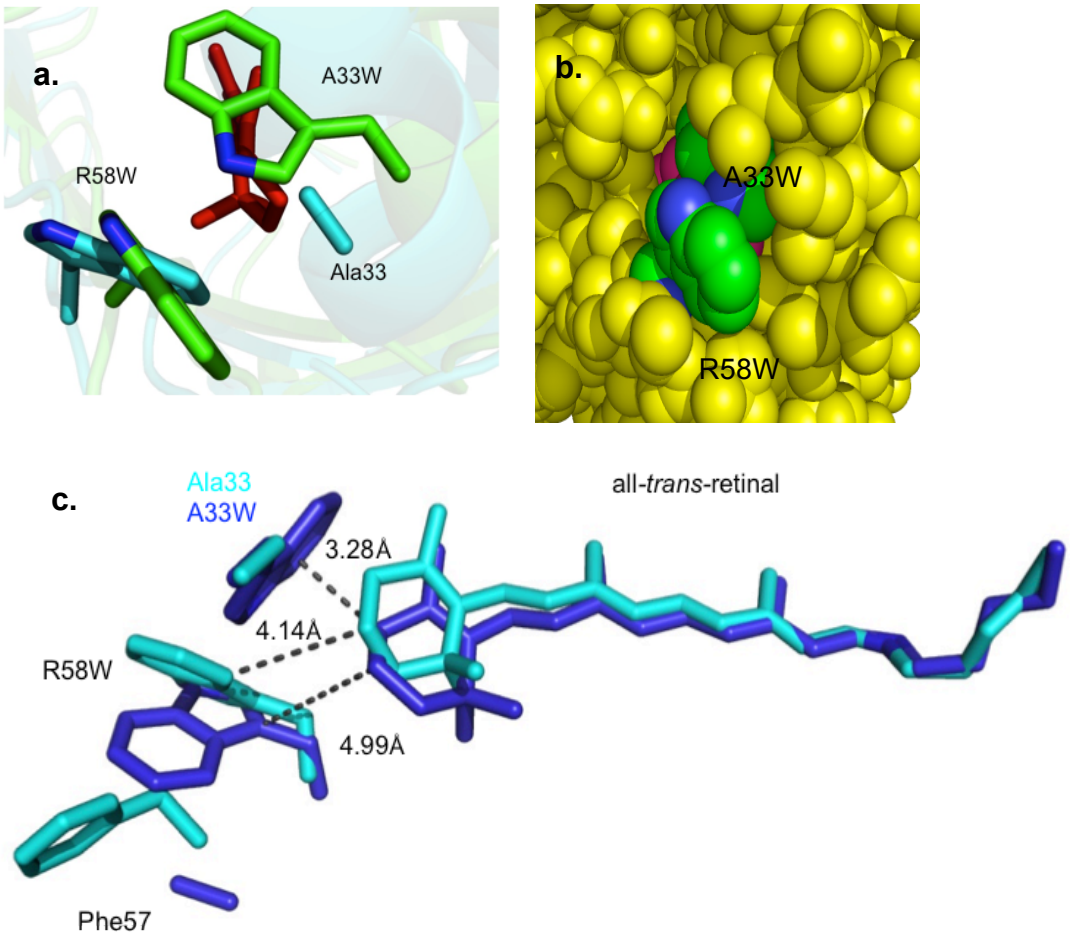


Figure III-29: a. The superimposed structure of and Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (cyan) and Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W (green). **b.** The space-filling model of the crystal structure of and Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W with the highlighted residues. **c.** The superimposed structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (blue) and Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W (cyan)

Trp58 has flipped out of the binding site to avoid a clash with Trp33 and the density of the side chain of Phe57 is weak, indicating significant disorder of this side chain. Since Trp58 is flipped back the distance of the side chain of this residue increased form the original distance (4.14Å) but the side chain of Trp33 is 3.28Å away from the ionone ring of the chromophore. This close distance of an aromatic residue can have two effects: helping the delocalization of the positive charge of the PSB and red shifting absorbance and increasing the Pka of the PSB. As shown in the Table III-11 the pK_a of the mutants with A33W is around 0.5 units higher than the ones with no mutation of A33.

Table III-13: Introduction of A33W in different hCRBP mutant

Mutant	pK_a with A33	pK_a with A33W
Q108K:K40L (KL)	7.9	8.4
KL:T51V:T53C	8.4	7.7
KL:T51V:T53C:Y19W	8.4	8.9
KL:T51V:T53C:R58W:T29L	7.9	8.1
KL:R58F:Y19W	9.3	9.7
KL:T51V:T53C:R58W:T29L:Y19W	8.2	8.7

III-5: *Trans*-iminium and Details Studies on Gln4

As described above, a water mediated interaction between Gln4 and the iminium molecule is not observed in any of the structures of wt CRBPII, nor is it seen in the apo

Q108K:K40L:T51V:T53C:R58W:Y19W:T29L mutant structure. This water-mediated interaction requires the iminium double bond to be in the cis conformation, which is the conformation seen in every structure in which Gln4 is present. In an effort to obtain even more red-shifted absorbances in hCRBPII.

Dr. Wenjing Wang in Professor. Borhan's group made various mutations at the Gln4 position. Table III-12 summarizes mutations at the Gln4 position in the context of the with hepta mutant Q108K:K40L:T51V:T53C:R58W:T29L:Y19W. Clearly mutation of Q4 to a variety of amino acids all lead to a dramatic red shift and simultaneously a significant drop in pK_a. The lone exception is the Q4H mutation (Table III-12 entry 9) Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H) which gives a small blue shift with virtually no change in pK_a. (relative to Q108K:K40L:T51V:T53C:R58W:T29L:Y19W). The maximum red shift in this series of mutants are those that introduce a positively charged amino acid at the Q4 position (Q4R (622 nm with the lowest pK_a = 6.5) and Q4K (618 nm)). To understand the effect caused by Gln4 mutation, a number of mutants were screened for crystallization. Unfortunately, most Q4 mutants failed to crystallize. This is similar to what we have seen with other mutants in that mutants with relatively low pK_a's are more difficult to crystallize, if they crystallize at all.

Table III-14: The effect of Q4 mutation

Mutants	λ_{max}(nm)	pK_a
KLVCWLW:Q4[a]	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[a]

[a]: KLVCWLW is Q108K:K40L:T51V:T53C:R58W:T29L:Y19W

Entry 10 was not stable and its pK_a could not be determined.

III-5-1: The Crystal Structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H Mutant Bound to Retinal

In contrast to the other mutants in Table III-12 Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H absorbs at 585 nm and is 6 nm blue shifted compared to Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (591 nm). The crystal structure of this protein with all-*trans*-retinal crystallized in the P1 space group with two molecules per asymmetric units and the data were refined to 1.3Å resolution with crystallographic R_{fac} = 18.18 % and R_{free} = 21.98%. The electron density of the

chromophore is very clear in both chains. The structure of the binding pocket with highlighted mutated residues is shown in Figure III-27. The structure shows that mutation of Gln4 to Histidine removed the water between Gln4 and the PSB and as a result the iminium adopts the *trans* conformation. The *trans*-imine is usually a few Kcal/mol more stable than the *cis*-imine, however the *cis*-imine in the hCRBPII mutants can be stabilized by the water-mediated interaction between the iminium nitrogen and Gln4. Further potential stabilization may come from the p-cation interaction with Trp106. Given the *trans* conformation seen in this structure, it appears that removal of the water-mediated interaction is sufficient to convert the conformation of the PSB from cis to *trans*. Indeed, in all structures that have Gln4 mutated, the PSB is found in the

trans conformation, consistent with the idea that this water-mediated interaction is critical for stabilizing the cis conformation in the binding pocket.

However, mutation of Gln4 to other amino acids abolishes the water-mediated interaction. In the structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H there are also other differences in the water network near the PSB. Figure III-31 shows the crystal structures of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H and Q108K:K40L:T51V:T53C:R58W:T29L:Y19W overlaid. The two structures did not change significantly (RMSD = 0.868Å for all 266 amino acid residues). As shown His4 makes a water mediated interaction with the main chain carbonyl of Thr1 via two water molecules while in the structure of K40L:T51V:T53C:R58W:T29L:Y19W, Gln4 makes the water mediated interaction with Thr1 using only one water molecule.

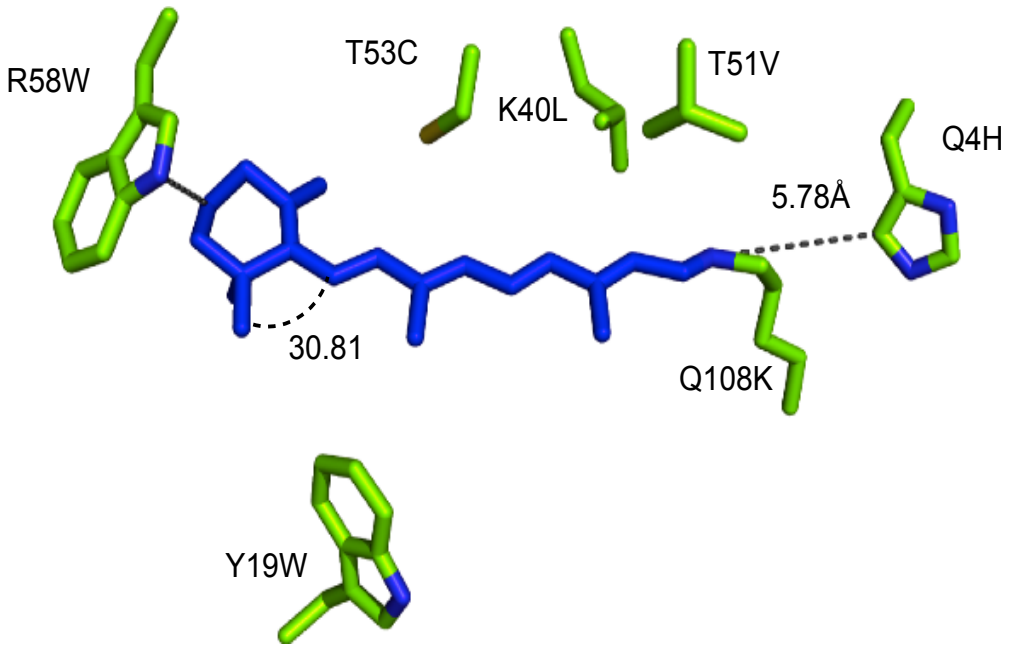


Figure III-30: The crystal structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H with the highlighted residues.

It is suspected that this extra water molecule in K40L:T51V:T53C:R58W:T29L:Y19W:Q4H may result in the small blue shift and moderate stabilization of the pK_a. This mutant has the highest pK_a of any of the Gln4 mutants (Table III-12). However this water molecule does not have any direct interaction with the PSB, and is in fact 7Å away from the iminium nitrogen.

In this structure the chromophore adopts a more planer conformation than seen in many of the other structures so far determined. For example the dihedral angle between Met5-C5-C6-C7 in this structure is 36.81°and the angle between Met9-C9-C13-Met13 is -3.68°, this is the smallest torsion angle observed in any of the structures of Retinal-bound hCRBPII mutants.

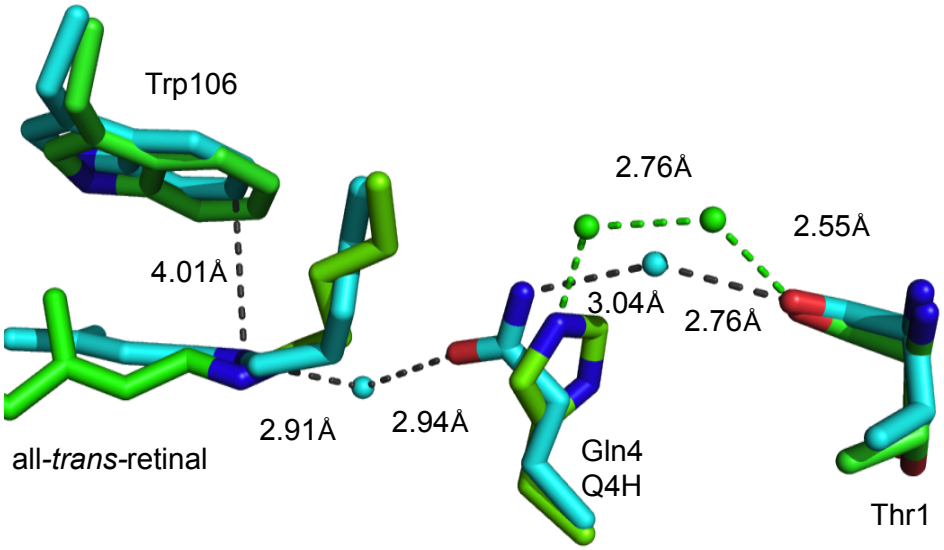


Figure III-31: The superimposed crystal structure of K40L:T51V:T53C:R58W:T29L:Y19W:Q4H (green) and K40L:T51V:T53C:R58W:T29L:Y19W (cyan).

III-5-2: The Crystal Structure of Q108K:K40L:T51V:T53C:R58F:T29L:Y19W:Q4H Mutant Bound to Retinal

As explained above mutation of Arg58 to aromatic residues leads to a large red shift, and enhances the red shift of other mutants over what would be expected form an additive effect. The different mutants with Q4H and Arg58 mutation to different aromatic residue were generated by Dr. Wenjing Wang. With these data we could see how much the size of mutant at Arg58 position has effect in the longer series of mutations and therefore Q108K:K40L:T51V:T53C:R58F:T29L:Y19W:Q4H has been generated. This

mutant absorbs at 547 nm (11 nm blue shifted relative to Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H) and has a pK_a of 7.7. The mutant was crystallized in the P1 space group with two molecules per asymmetric unit. All of the crystallographic information of all of the mutants is listed in Table I in chapter IV. The crystal structure of the two mutants with R58F and R58W do not change a major different and the RMSD is 1.654Å for 266 amino acid residues (Figure III-29). The superimposed structure of these two mutants illustrated that the only structural difference is the position of the side chain of Phe58 and Trp58. The crystal structure shows the Trp58 side chain not only cover the binding pocket more efficiently but is also closer to the chromophore. This effect seems to be result of changing the conformation of PSB from *cis* to *trans* and the chromophore is deeper in the binding pocket. This data confirms that the nature and the distance of the other polar residue even far from the PSB can stabilized the charge. The comparison of these two crystal structure shows that while Cys53 is found in a single conformation, pointing toward the ionone ring in Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H. In Q108K:K40L:T51V:T53C:R58W:T29L:Y19W, Cyc53 adopts two conformations. This is quite consistent with what has been seen in most other mutants, where the larger Trp58 leads to a more-red shifted spectrum than does Phe58.

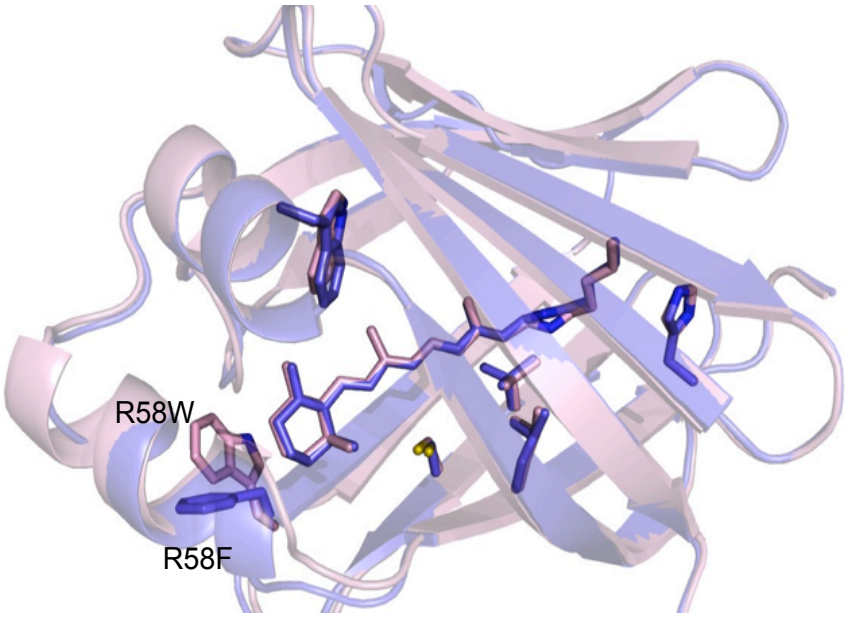


Figure III-32: The superimposed crystal structure of KL:T51V:T53C:R58W:T29L:Y19W:Q4H (purple) and KL:T51V:T53C:R58F:T29L:Y19W (pink).

III-5-3: The Crystal Structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4A and Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R Mutant Bound to Retinal

Since the Q4H mutation leads to a blue shift while virtually all other mutations at Gln4 resulted in red shifted spectra, it was necessary to confirm structurally that these mutants also give the *trans* conformation. To this end structures were sought. Since mutants with aromatic residues at position four had failed to crystallize, and the proteins were found generally to be less stable, the crystallization of a Q4A mutant was attempted, with the assumption being that less steric bulk may lead to a more stable

protein and better odds of crystallization. This was born out as well-diffracting crystals were obtained of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4A. Further, since mutation of Gln4 to a positively charged amino acid lead to even further red shifting, culminating in the most red shifted mutant obtained to date (2013), a structure of Q4R or Q4K mutant was also sought. To this end Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R was crystallized and its structure determined to high resolution. Both mutants crystallized in the P1 space group, with 2 molecules in the asymmetric unit. The crystallographic data of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R was refined at 1.5Å with crystallographic R_{free} = 20.08 % and R_{fac} = 17.07 %. The structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4A bound to retinal was determined and refined at 1.5 Å and R_{fac} = 18.70% and R_{free} = 22.49%. All four Q4 mutant structures are very similar. A *trans* iminium is observed in all four, and the conformation of the chromophore is also very similar. Taken together, it appears that Gln4 results in the iminium cis conformation while mutation of this residue to almost anything else leads to the usually more stable *trans* iminium conformation. The *trans* conformation gives a more red shifted spectrum and a lower PSB pK_a (by about 1 pH unit), due to the loss of stabilizing interactions with both Trp106 and the water mediated interaction with Gln4. Figure III-33 shows the superimposed structure of these three mutants.

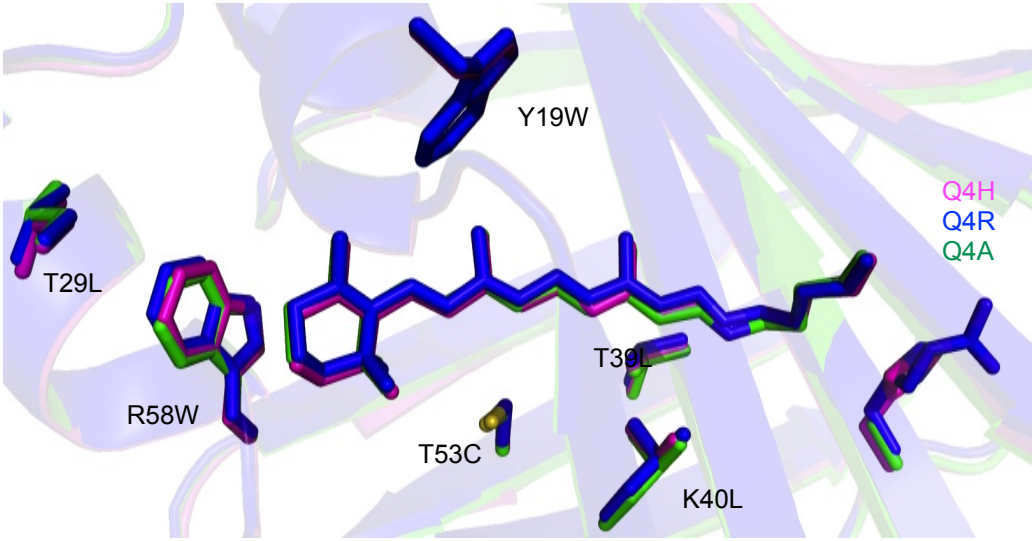


Figure III-33: The most red shifted mutants (such as Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R) also have the lowest pK_a.

Inspection inside the binding pocket of the Q4R mutant shows that Arg4 with is buried in the N-terminus of the protein and does not directly interact with the PSB, which explains why the pK_a is not severely depressed (Figure III-34). Arg58 makes hydrogen bonds with the main chain carbonyls of Thr1, Lys109 and Asp91. Several ordered water molecules, seen in virtually all other hCRBPII mutants are displaced by the position of the Arg4 guanadinium. Thus the guanadinium is buried in a relatively hydrophilic region of the binding cavity, explaining how the protein is capable of remaining relatively stable upon burial of a large positively charged side chain. Even though the guanadinium of Arg4 is relatively far from the chromophore (more than 6Å from the PSB), introduction of the positive charge leads to a significant red shift of almost 10 nm relative to the Q4W mutation (Table III-12) and a reduction in the PSB pK_a (6.5 relative to between 7.2 and 7.9 in other Q4 mutants (Table III-12)).

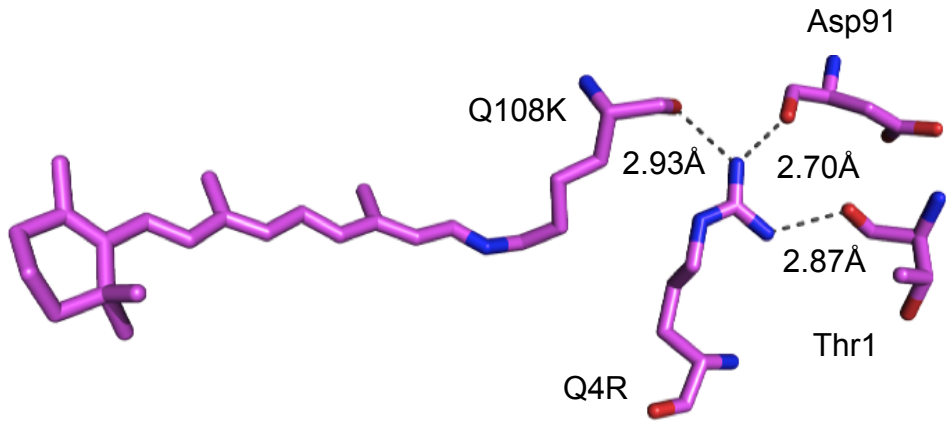


Figure III-34: The crystal structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R. The side chain of Arg4 is buried in the N-term of the protein.

Changing the conformation of the PSB from *cis* to *trans* also leads to a translation of the chromophore of about 1Å. This translation allows the side chain of Trp58 flip into the binding site, packing against the chromophore perfectly. This was also observed in all of the crystal structures with the Q4 mutations (Figure III-33). Figure III-33 illustrates that the orientation of Trp58 in Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R covers the binding pocket of hCRBP more efficiently. Further Crys53 adopts a single conformation, with its side chain pointing toward the ionone ring of the chromophore, while in all of the other mutants with both T51V and T53C mutations the side chain of Cys53 has two conformations.

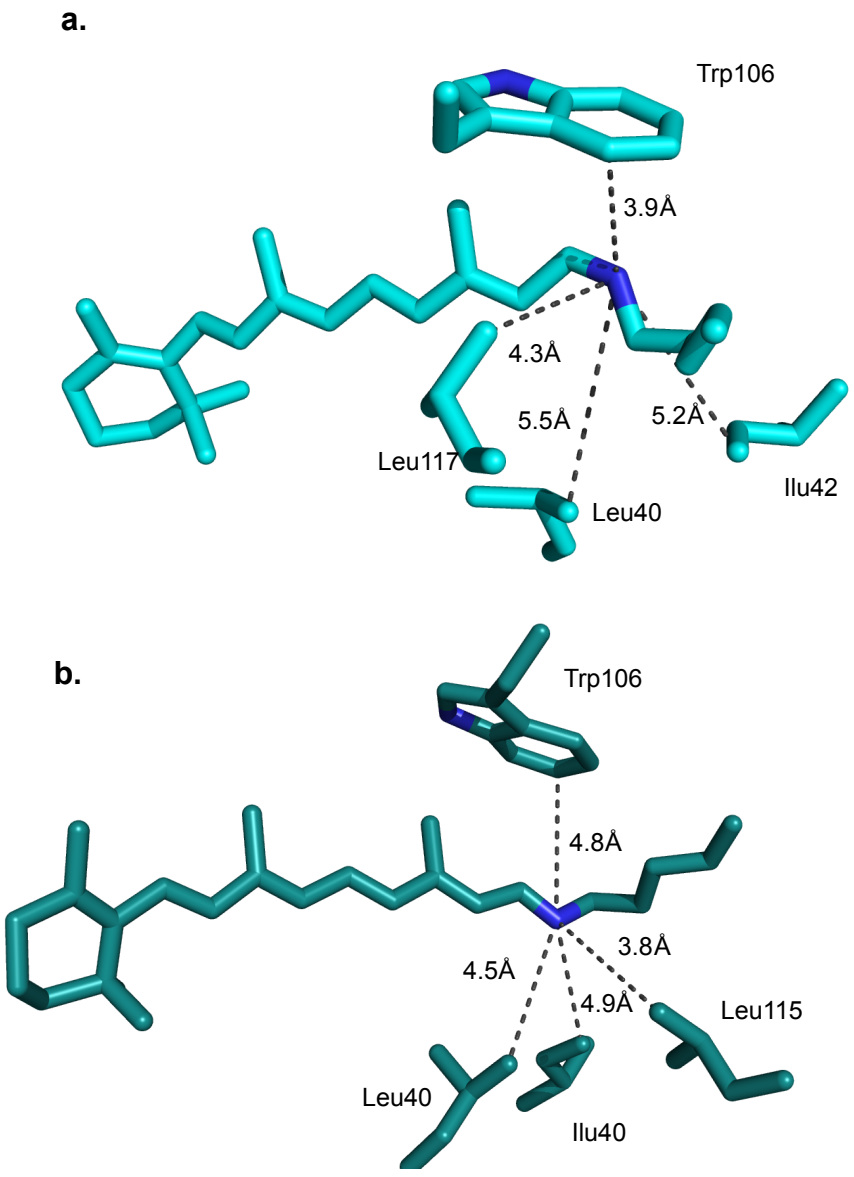


Figure III-35: a. The π-cation interaction of Trp109 with *cis*-iminium. **b.** *Trans*-iminium.

The conformation of Cys53 may be correlated to the accessibility of the binding pocket to the bulk media as the binding pocket in the *trans* mutants is even better sequestered due to the conformation of Trp58 by closing the binding pocket the side chain points to the ionone ring which probably carries more positive charge than when the binding pocket is more open to the aqueous media, then Cys53 has two different conformations. Another consistent difference seen in the Q4 mutants is in the water network around Gln38 and Gln128. In the crystal structure of Q108K:K40L and most of the other mutants with Gln4 one water molecule is found between Gln38 and Gln128.

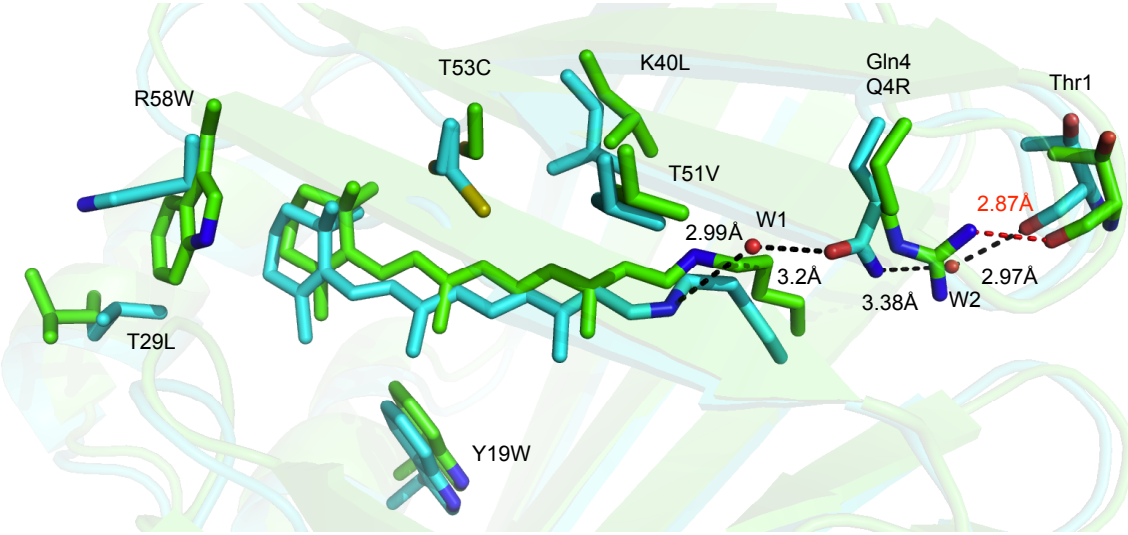


Figure III-36: The superimposed structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R (green) and Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (cyan).

However, in the structures of the Gln4 mutants, two water molecules are observed (Figure III-35). One water makes hydrogen bond with two nitrogen's of the amide while the second one makes the two carbonyl groups. This is also observed in the structure of wt hCRBPII.

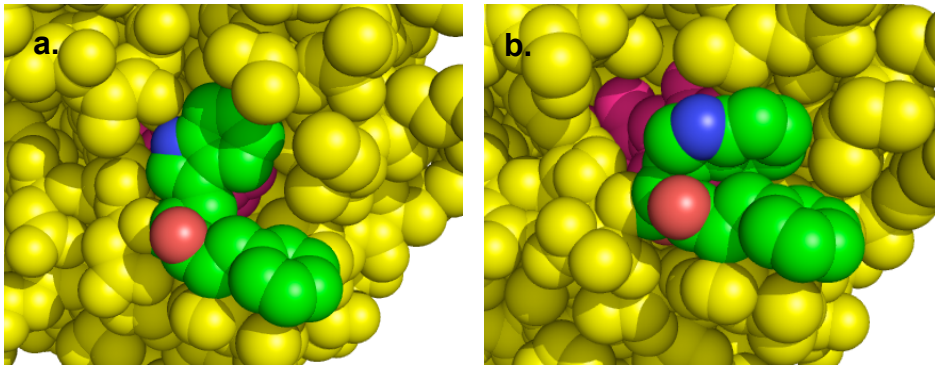


Figure III-37: a. The space filling model of the crystal structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R mutant the chromophore is red. **b.** The sphere structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W mutant.

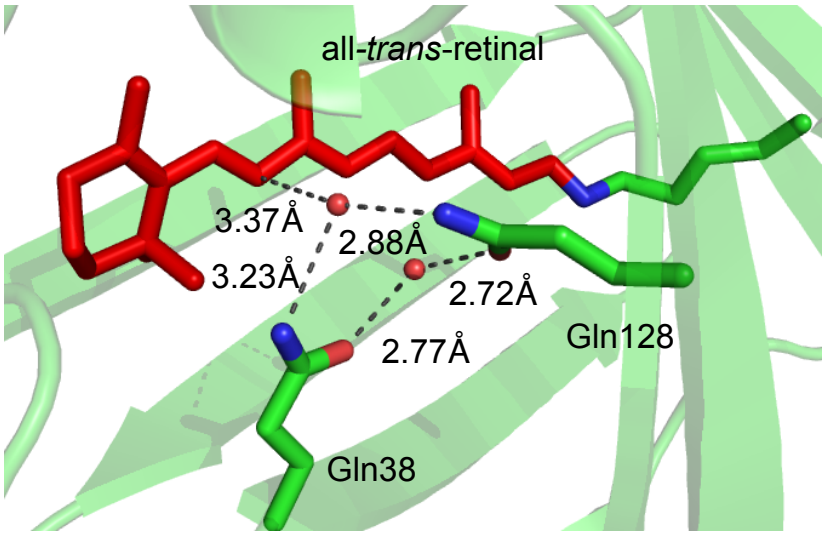


Figure III-38: The details of the water network between Gln38 and Gln128 in Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H mutant.

In the category of *trans*-iminum structures, the maximum red shifted mutant was 644 nm, while in the *cis*-iminums with Gln4; the most red shifted mutant absorbs at 606 nm. Therefore it seems that mutation of Gln4 was required to access the most red shifted spectra. With the most red shifted of these the Q4R mutations, which introduces a positive charge in the vicinity of the PSB.

Unfortunately we could not have the crystal structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4E (590 nm) which is same as hepta mutant Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (591 nm). Modeling of this mutant in Pymol suggests since the size of both glutamine and glutamic acid have the same size therefore the glutamate side chain should be in the same place of glutamine and so the water intermediate is most likely is in the same place and for this reason no red shifting has been observed (Figure III-39).

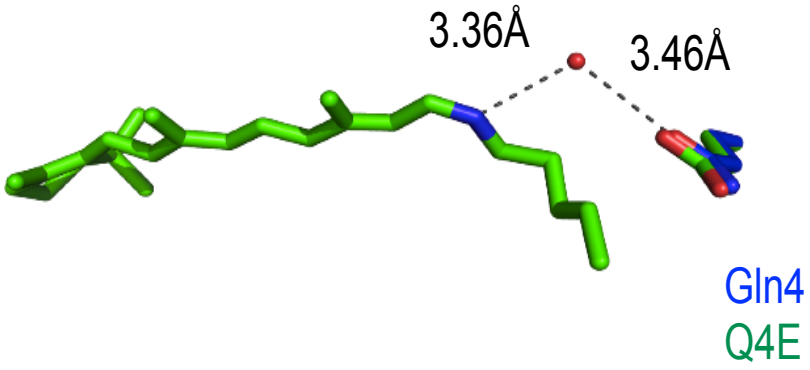


Figure III-39: The crystal structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (green) and model of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4E (blue).

It seems that removing this water molecule is necessary for the most red shifting absorbance and this strategy has been applied for the most red shifted mutant.

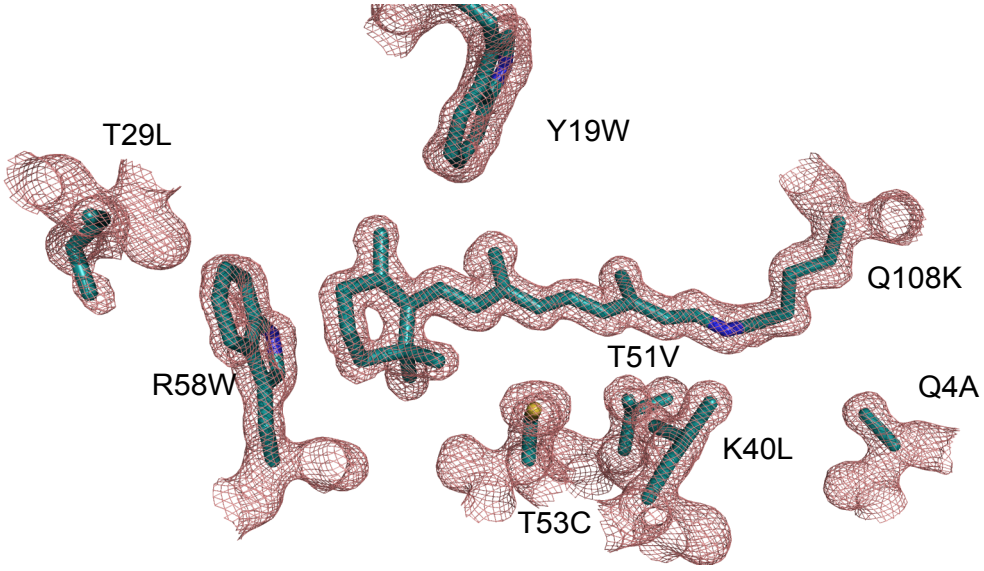


Figure III-40: The crystal structure of Q108K:K40L:T51V:T53C:Y19W:R58W:T29L:Q4A with highlighted mutated amino acids.

As previously discussed, adding the A33W mutation to the hepta mutant Q108K:K40L:T51V:T53C:R58W:T29L:Y19W (591 nm) the most red shifted mutants Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W with λ_{max} 606 nm in the category of cis-iminium was generated. The crystal structure of this mutant illustrates that A33W can cover the binding pocket of hCRBPII and better isolate the pocket from bulk solution. This isolation can prevent ions and water molecules to freely enter and exit the binding pocket and produce a more uniform environment around the chromophore has been obtained. In the category of the *trans*-iminium Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R with 622 nm was the most red shifted mutant. The combination of the most red shifted *cis* and *trans*-iminum mutant leads to the most red shifted retinylidene–bound protein complex (644 nm) and the nona mutants Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R:A33W was generated by Dr. Wenjing Wang. Unfortunately due to the difficulty of crystallization of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R. The crystal structure of the most red shifted mutant is not available. A model of the structure of this mutant (using Pymol) is shown in Figure III-41.

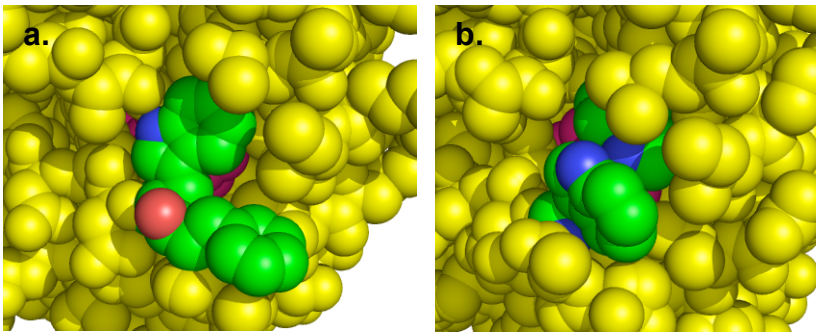


Figure III-41: a. The sphere model of the most red shifted mutant Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R. **b**. The sphere model based on the crystal structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W.

A comparison of the model of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R:A33W and the structure of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R shows the binding pocket of the most red shifted mutant to be completely covered with the chromophore deeply buried in the binding pocket and to have almost no access to the bulk solution. The pK_a of this mutant is 6.7 which is 0.2 unit higher than the pK_a of the Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R mutant (pK_a 6.5). It seems that the position of two aromatic residues (A33W and R58W) causes a small increase in the pK_a. The two important effects in the most red shifted mutants have important role:

1- Delocalization of the positive charge along the chromophore that can be achieved by removing all the polar residues around it. We also need to consider the water molecules in the binding pocket as some of the most polar groups in the binding pocket.

2- Isolating the binding pocket from the bulk solution and preventing the water or any charged ions to enter the binding pocket.

III-6: The Water Network Between Gln38 and Gln128

In the crystals structures of WT and mutant forms of hCRBPII, several ordered water molecules are found inside the binding pocket. One water network is located between the side chains ofGln38 and Gln128, which are almost 4Å away from the polyene chain of retinal. It seems that these water molecules have a very crucial role in the stabilization of pK_a and also in wavelength regulation. In the structure of WT and all of the Gln4 mutants (Q4H, Q4R and Q4R) two ordered water molecules are suited between Gln38 and Gln128, while in the structures of other mutants a single water is observed. We observed that mutation of the polar residues around the chromophore to less polar or hydrophobic groups typically lead to bathochromic shift (for example T51V, T53V, Q4W, Y19W R58W) but surprisingly, Mutations of Gln38 and Gln128 to hydrophobic residues have a dramatic effect on both pK_a and wavelength. As is listed in table III-13 mutations of one or both to hydrophobic residues lead to a decrease in the pK_a and dramatic blue shifting. In fact mutation of both Gln38 and Gln128 to non polar residues (entry 4 in table III-13), results in the largest hypsochromic shift generated to date, neutralizing simultaneously the effect of all of the mutants that give a red shift. Mutation of Gln38 to a charged or polar residue also leads to hypsochromic shifts but the effect is not as dramatic as with non polar residues. Without crystal structure it was suspected that by mutation of Gln38 or Gln128 the water

mediates as removed but fortunately the crystal structure of one of the hCRBPII mutant with 5-demethyl-retinal has been solved. This chromophore is all-*trans*-retinal analogue that has 6-s-*cis* conformation for ionone ring. This analogue was synthesized by Ipek yapici in Professor Borhan group for study on the effect of cis and *trans* conformation on the wavelength regulation. However the chromophore is not the same as retinal but as it is shown in figure III-40 they are structurally are very similar and therefore we can study on the water molecules in this structure. Figure III-39 shows the crystal structure of Q108K:K40K:T51V:R48W:Q38F (The mutant has been generated by Ipek Yapici in Professor Borhan's group) bound to 5-demethyl all-*trans*-retinal in P1 space group has been determined.

Table III-15: Mutation on Gln38 and Gn128 in hCRBPII

Mutants	λ_{max}(nm)	pK_a
Q108K:K40L:R58W:T51V:T53C:T29L	585	7.9
Q108K:K40L:R58W:T51V:T53C:T29L:Q38M	513	7.5
Q108K:K40L:R58W:T51V:T53C:T29L:Q128L	532	5.8
Q108K:K40L:R58W:T51V:T53C:T29L:Q128L:Q38M	504	7.2
Q108K:K40L:R58W:T51V:T53C:T29L:Y19W:Q4F	613	7.7
Q108K:K40L:R58W:T51V:T53C:T29L:Y19W:Q4W:Q38E	590	6.0
Q108K:K40L:R58W:T51V:T53C:T29L:Y19W:Q4F:Q38W	577	5.9

As is shown in this figure in both chains the water molecule has been observed that makes a hydrogen bond to the Gln128. This data shows at least in this mutant the

water molecules are not removed and it is possible that in the other Gln38 mutants this water is available, therefore the polarity of this water molecule seems to be important. The environment can effect on the dielectric constant of water. The dielectric constant of water drops inside the protein from 80 to 10-3.[6] The polarity of these water molecules when they make hydrogen bond to Glutamine is different when they are near hydrophobic side chain of Luecine. Therefore the polarity of these water molecules is very important, especially since it is located almost in the middle of the polyene chain, disturbing the polarity of them can leads to less delocalization of the charge on along the chromophore. More study on this water network is preformed by Meisam Nosrati in Professor Geiger's group.

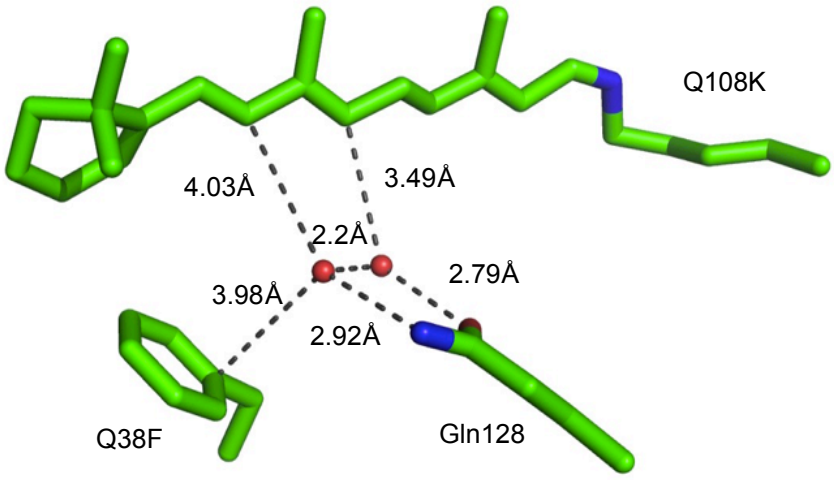


Figure III-42: The details of the water network in the binding pocket of Q108K:K40K:T51V:R48W:Q38F mutant. The water network is originally between Gln38 and Gln128.

III-7: The Effect of the Overall Electrostatic Potential and the Water Networks and Interaction with Chromophore

The different rhodopsins are able regulate the wavelength from 420 nm to 560 nm with just one single chromophore, 11-*cis*-retinal. It was the first time that by combination of mutations (maximum 9 different mutations) one protein could modulate the full visual spectrum from 425 nm to 6 nm in one chromophore (all-*trans*-retinal) Figure III-43. Combination of the rationalized mutations and X-ray crystallography provide wide range of information to explain the mechanism of the wavelength regulation in a rhodopsin mimic.

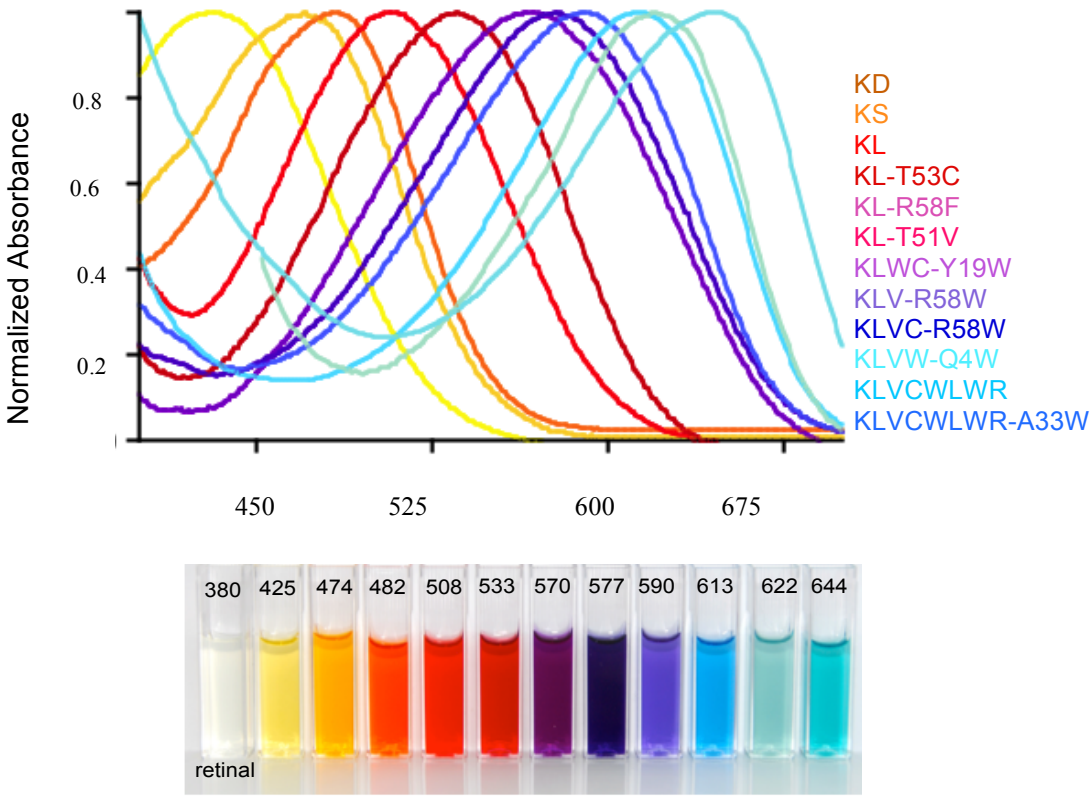


Figure III-43: Full spectrum of hCRBPII mutants. **a.** UV-vis spectra of different hCRBPII mutants bound with all-*trans*-retinal. **b.** Protein solution of different hCRBPII mutants incubated with all-*trans*-retinal.

As was mentioned before the electrostatic calculation on the different rhodopsins shows, that overall electrostatic potential of the protein around the chromophore is important to modulate the wavelength. The different crystal structures of mutants helped us to calculate the overall electrostatic potential of the protein around the chromophore. For tuning the wavelength, we need to have a closed and isolated binding pocket, which is exactly like rhodopsin structure. In hCRBPII as mimic of rhodopsin R58W mutation and additionally A33W mutations cover the binding pocket and separate the chromophore from the bulk solutions.

Therefore in these electrostatic calculations the dielectric constant of water was assigned for the mutations with and without these two mutations.

The electrostatic calculation illustrates that for at the most blue shifted structure the chromophore region has more positive charge, while the PSB has more negative electrostatic potential. However this pattern is not reversed for the “super red shifted” mutant. As shown in the Figure III-43-b for super red shifted spectrum we need to have a uniform electrostatic potential around the chromophore. The calculation the overall electrostatic potential of each mutants illustrate the different mutants with different electrostatic potential have different absorbance. However looking closely to the details of the each crystal structures shows that the mutants that have effects on the water networks in the binding pocket have a dramatic influence on the wavelength while the ones which do not have any interactions with the water networks are not able to have a large effect on the wavelength. For example removing the water network near the PSB leads to extra red shifting. Unfortunately the crystal structures do not show the direction of the hydrogen atoms in water molecules and also the polarity of the water molecules,

because as it was shown above changing the polarity of the water molecules in the binding pocket (Gln38 and Gln128 mutations) can lead to dramatic blue shifting.

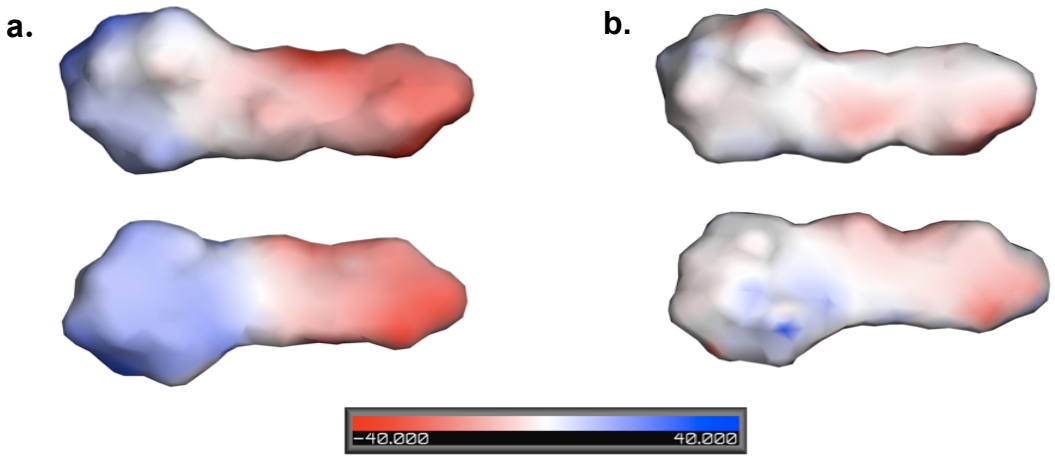


Figure III-44: Electrostatic projections ($k_bTe_c^{-1}$) using the crystal structure of **a.** KL (508 nm). **b.** KL:T51V:T53C:Y19W:R58W:T29L:Q4R (622 nm) hCBRP II mutants (for each calculation both sides of the chromophore, all-*trans*-retinal, are shown).

III-8: Effect of Conformation of the Ionone Ring on the Wavelength Regulation

As it was explained the effect of conformation of the chromophore was suggested as an important role in the different absorption maxima in PSB-bound to protein. The more planer chromophore leads to more conjugation of the p-orbitals and as the result, more red shifting is observed. One of the most common conformational differentiations is believed to happen along the C6-C7 single bond. By twisting this

single bond the double bond of the ionone ring cannot have overlap with the rest of the double bond of the backbone. Therefore hypsochromic shift will be observed.

Based on these theories Ipek Yapici in Prof. Babak Borhan group designed the series of mutations. The idea was introducing the series of mutations in hCRBPII bound to all-*trans*-retinal that can change the conformation on the ionone ring from 6-s-*trans* to 6-s-cis and sturdy in the wavelength of these mutants. Since cis conformation is slightly (0.6 Kcal/mol) is more stable than the *trans* one, therefore theoretically by providing enough space around the ionone ring, the C6-C7 single bound can rotates and 6-s-*cis* conformation will be formed. The different mutants have been generated and I solved the structure of two different mutants. Dr. Camille Watson from Professor Geiger's lab also worked crystal structures of some of those mutants.

III-8-1: The crystal Structure of Q108K:K40L:T51V:R58Y:Y19W:T53L Bound to Retinal

The crystals structure of Q108K:K40L:T51V:R58Y:Y19W mutant illustrated that the chromophore adopts the 6-s-*cis* conformation. It is not clear that why introduction of R58Y mutation leads to cis conformation, however as it was explained in the structure of this mutant the position of side chain of Phe16 is moved and enough space was provided for free rotation along C6-C7 bond. Based on these data Ipek Yapici designed Q108K:K40L:T51V:R58Y:Y19W:T53L mutant. The crystals structure of this mutant was determined in P1 space group with two molecules per asymmetric units. The data were refined at 1.2Å resolution with crystallographic R_{free} = 20.83 % and R_{fac} = 27.93%. Since

the ionone ring of Q108K:K40L:T51V:R58Y:Y19W has 6-s-*cis* conformation The idea of this mutant was introduction a bulky group (T53L) that is able to makes crowded around the gem dimethyl group, force it to rotate and lock the chromophore in the cis conformation, however the crystal structure of this mutants clearly illustrated that the ionone ring has *trans* conformation and the gem dimethyl group flipped Leu53 back. It was not very surprising to find out that the protein could not able to dictate the conformation to the chromophore and from the first time, it was not clear why Q108K:K40L:T51V:R58Y:Y19W adopts the *cis* conformation. The superimposed structure of Q108K:K40L:T51V:R58Y:Y19W and Q108K:K40L:T51V:R58Y:Y19W:T53L mutants shows that the chromophore is exactly is the same position expect the ionone ring (RMS is 0.394Å for 133 amino acid residue). The only difference in the structure of these two mutants is the position of Phe16. The side chain of this residue is 0.4Å more far from the chromophore. This distance might able to makes enough space for the chromophore to rotate.

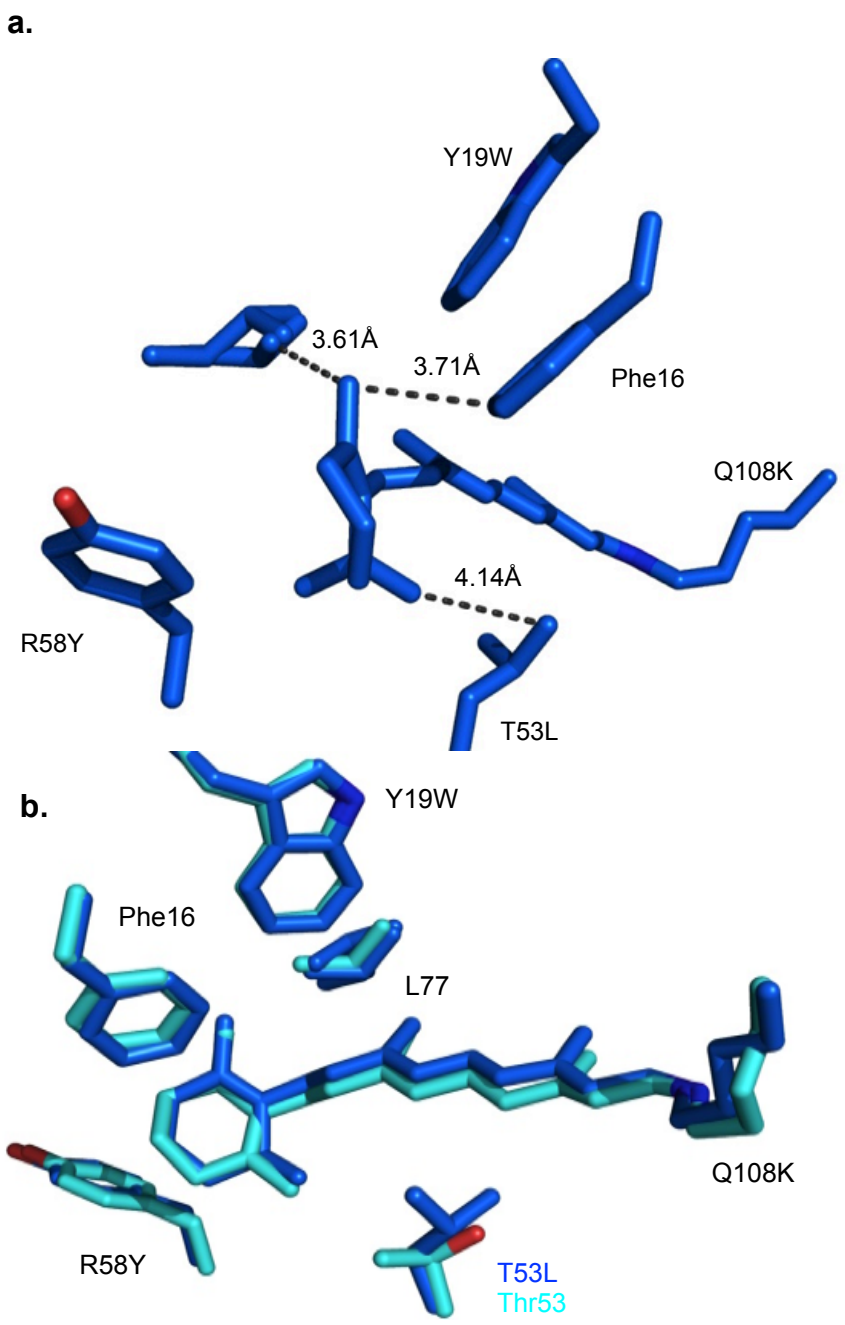


Figure III-45: a. The crystal structure of Q108K:K40L:T51V:R58Y:Y19W. **b.** The superimpose structure of Q108K:K40L:T51V:R58Y:Y19W (cyan) and Q108K:K40L:T51V:R58Y:Y19W:T53L (blue).

III-8-2: The Crystal Structure of Q108K:K40L:T51V:R58W:Y19W:T53L:L77T Bound to Retinal

The crystal structure of Q108K:K40L:T51V:R58W:Y19W:T53L:L77T mutant with all-*trans*-retinal was determined at 1.17Å in P1 space group with four molecules per asymmetric unit. This mutant was generated by Ipek Yapici in Professor Borhan‘s group. The data were refined with crystallographic data R_{free} = 20.32% and R_{fac} = 24.15%. This mutant has been generated based on the idea of forcing chromophore to adopt the 6-s-*cis* conformation. From one side mutation of Thr53 to Leucine which is more bulky group and makes steric interaction with the gem-dimethyl group and the other side mutation of Leu77 to threonine, which is less bulkier and makes enough space for the chromophore to rotate. The crystal structure of this mutant shows the chromophore adopts the same 6-s-*trans* conformation. Interestingly it was observed Leu53 side chain flipped back to the beta-sheet to avoid the clash with the gem-dimethyl of chromophore. Figure III-43 shows the crystal structure of this mutant and it is clear the chromophore adopts the perfect *trans* conformation. So far no mutant has been generated to force the chromophore to adopt another conformation. Dr. Camille Watson in Professor Geiger’s group refined other structure for this project.

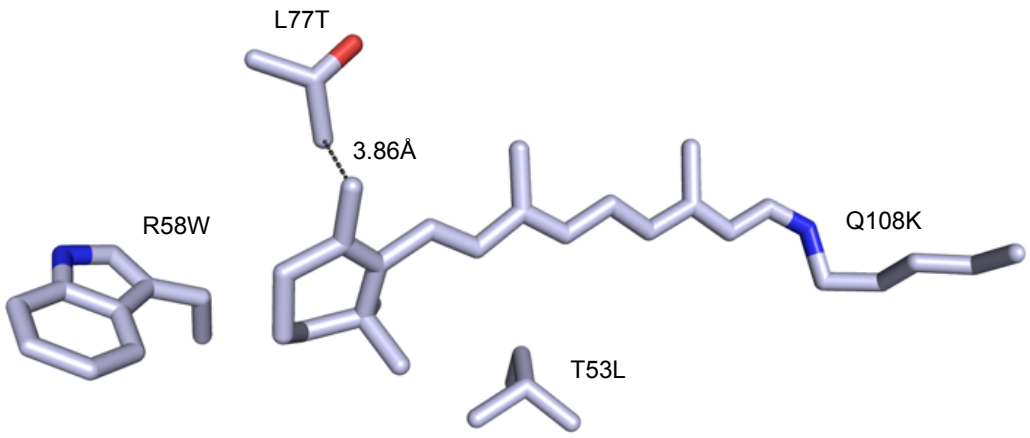


Figure III-46: The crystal structure of Q108K:K40L:T51V:R58W:Y19W:T53L:L77T with the highlighted residues.

III-9: Conclusion

In this study on the wavelength regulation in a cytosolic protein as a rhodopsin mimic and based on more than two hundred mutations and more than twenty crystal structure, we can conclude that, usually the mutations that have effect on the ordered water molecules have usually the most significant effect on the wavelength spectrum in both red and blue shifted. We could not find the relation between the conformation of the chromophore and wavelength regulation. However in this study, the most red shifted mutants also have the most planer chromophore, but there were not any correlation between the torsion angles and wavelength. Other than that based on the crystal structures of the mutants, it is concluded that the direct counter anion is not necessary for this system and the PSB can be stabilized in different mechanism such as π-cation interaction or hydrogen bond with a polar group. We were also able to show creating a

hydrophobic area around the entrance of binding pocket can leads to isolation of the binding pocket from the bulk solution. In this case study water mediate was responsible for stabilization of the PSB. The electrostatic calculation on based on the crystal structure of hCRBPII mutants clearly shows that for super red shifted spectrum, evenly distribution of charge is necessary.

REFERENCES

REFERENCES

1. Baasov, T., Sheves, M., Alteration of pka of the bacteriorhodopsin protonated schiff-base - a study with model compounds. *Biochemistry* **1986**, 25 (18), 5249-5258.

2. Rosenber.B; Krigas, T. M., Spectral shifts in retinal schiff base complexes. *Photochem. Photobiol.* **1967**, 6 (10), 769-&.

3. Baasov, T.; Sheves, M., Model compounds for the study of spectroscopic properties of visual pigments and bacteriorhodopsin. *J. Am. Chem. Soc*. **1985**, 107 (25), 7524-7533.

4. Mogi, T.; Marti, T.; Khorana, H. G., Structure-function studies on bacteriorhodopsin .9. substitutions of tryptophan residues affect protein-retinal interactions in bacteriorhodopsin. *J. Biol. Chem*. **1989**, 264 (24), 14197-14201.

5. Wang W, Wang, Wenjing, Nossoni Z, Berbasova T, Watson C, Yapici I, Lee, Kin S, Vasileiou C, Geiger J, Borhan B (**2012**) Tuning the Electronic Absorption of Protein-Embedded All-*trans*-Retinal. *Science* 338 (6112):1340-1343.

6. J. J. Dwyer et al. *, Biophys*. J. 79 , 1610, **2000**.

CHAPTER IV

A Single Mutation Resulted an Extensive 3D Domain Swapped Dimerization in HCRBPII

IV-1: Introduction

Human Cellular Retinol Binding Protein II (hCRBPII) is a stable, cytosolic, monomeric protein responsible for the intracellular transport of retinol and retinal. It is a member of the intracellular lipid binding protein (iLBP) family.[1] CRBPII and other members of this family have been shown to be structurally remarkably stable to mutation. Herein we report that a single tyrosine to tryptophan mutation causes a unique and extensive domain swapped dimer. The mutation is not in the hinge loop region but resides four amino acids away. The domain swapped region encompasses almost half the protein, representing one of the largest domain swapped regions so far discovered. The domain swapped dimer represents a novel structure with the mouths of the two binding sites facing each other to produce a new cavity that stretches across the dimer. Since the monomeric form of the mutant represents the thermodynamically lower energy species, the domain swapped dimer represents a kinetically trapped species that does not interconvert to monomer at neutral pH and room temperature. Structural studies of two forms of the domain swapped dimer suggest a possible mechanism for its formation. Three-dimensional (3D) domain swapping is a process in which two or more monomer protein molecules exchange identical structural elements to form dimers or higher-order oligomers. It was first observed in ribonuclease by the

Eisenberg's group in 1995.[2] To date over 40 different proteins are reported to show 3D domain swapping.[3] This mechanism, combined with gene duplication, may be responsible for the evolution of a number of larger protein domains from smaller ones. In a recent study 80% of protein domains were found to have potentially evolved in this way. Domain swapping can also cause aggregation, amyloid and fibril formation, which lead to protein misfunction.[4,5] It is believed that protein misfolding and aggregation is the cause of diseases such as Parkinson's disease,[6,7] Alzheimer disease,[8,9] diabetes,[11] Huntington's disease,[11] Gaucher's disease and many others. However dimerization and oligomerization of can also lead to some advantages such as increasing enzyme activity due to the larger active and binding sites, larger surface area and therefore more binding affinity, etc.[12] The mechanism of domain swapping is not completely understood, but it usually occurs during protein expression at high concentrations and when the energy barrier between an "open" and "closed" monomer for any reason decreases. This effect can be caused by pH, concentration, temperature, mutation and ligand binding. Previously it was reported that domain swapping usually occurs at high protein concentration.[14,15] The first and most studied domain swapped protein was Bovine pancreatic ribonuclease (Rnase A).[1] The first two crystal structures of this enzyme showed two distinct dimers depending on the concentration of mild acid. The minor dimer forms by swapping the N-terminal a-helixes while the major one forms by swapping the C-Terminal b-strands. So far more than 290 different structures of domain swapped proteins are available in the protein data bank (PDB) and the number of these proteins is still increasing.[13] In most cases the exchanged region is located at the C- or N-terminus of the protein. However in a few cases the exchanged part is in the middle

of the sequence. Domains swapped structures have been observed previously by introducing point mutations in some proteins. For example a single site mutation G55A in protein L from Peptostreptococcus Magnus leads to 3D domain swapping.[16]

Human cellular retinol binding protein II (hCRBPII) is a relatively small (133 residue), soluble, cytosolic, monomeric protein that plays an important role in metabolism of vitamin A and is responsible for intracellular retinol transport. A member of the intracellular lipid binding protein (iLBP) family, the structure consists of a ten-stranded b-barrel, and two short a-helices that cover the binding pocket (Figure IV-1-a). The ligand is buried deeply within the binding site.[17] The structures of both hCRBPII and other members of this family have been shown to be remarkably resilient to mutation. For example more than 150 structurally stable mutants of hCRBPII have been characterized in our lab in the course of our studies of protein-chromophore interactions.[18]

Herein we report the first domain swapped structure of a mutant of human cellular retinol binding protein II (hCRBPII). The extent of domain swapping is one of the largest reported, involving almost 50% of the protein structure. This is, to our knowledge, the first example of domain swapping caused by mutation outside the hinge loop region. Further, we show that domain swapped dimerization is favored by high expression levels *in vivo* and high concentrations *in vitro.* Further, melting studies show that the domain swapped dimer is the less stable, kinetic product of refolding.

IV-2: Results and Discussion

HCRBPII was redesigned to be used as a rhodopsin mimic to study the mechanism of wavelength tuning[18]. The first step in this process was to introduce a lysine inside the binding pocket that would react with all-*trans*-retinal to produce an Schiff base. Inspection of the crystal structure of the hCRBPII retinol complex shows Gln108 to be hydrogen bonded to the hydroxyl group of the ligand and properly positioned for nucleophilic attack when mutated to lysine, leading to the Q108K mutant. Lys40, proximal to the Schiff base, was then mutated to leucine to remove a positive charge proximal to the Schiff base and thus promote protonated Schiff base (PSB) formation. The Q108K:K40L mutant resulted in stable PSB formation with retinal (ref). The crystal structure of this mutant bound to all-*trans*-retinal at 1.7Å (PDB entry 4EXZ) shows the PSB chromophore following a trajectory through the binding pocket similar to that of Retinol in the wt hCRBPII retinol complex. The residues situated around the chromophore were then selected for systematic mutation in an effort to understand the mechanism of wavelength tuning. Tyr60 is located in beta-strand C near the loop that connects beta-strands C and D (Figure IV-2). This residue is approximately 4Å from the center of the polyene chain of all-*trans*-retinal (Figure VI-2-b).

It is a conserved residue in the CRBP family (found in hCRBPI, hCRBPII, hCRBPIII, CRBPIV, zebrafish CRBP, rat CRBPI and II. *In silico* modeling suggested that mutation of tyrosine to tryptophan, an electron rich and

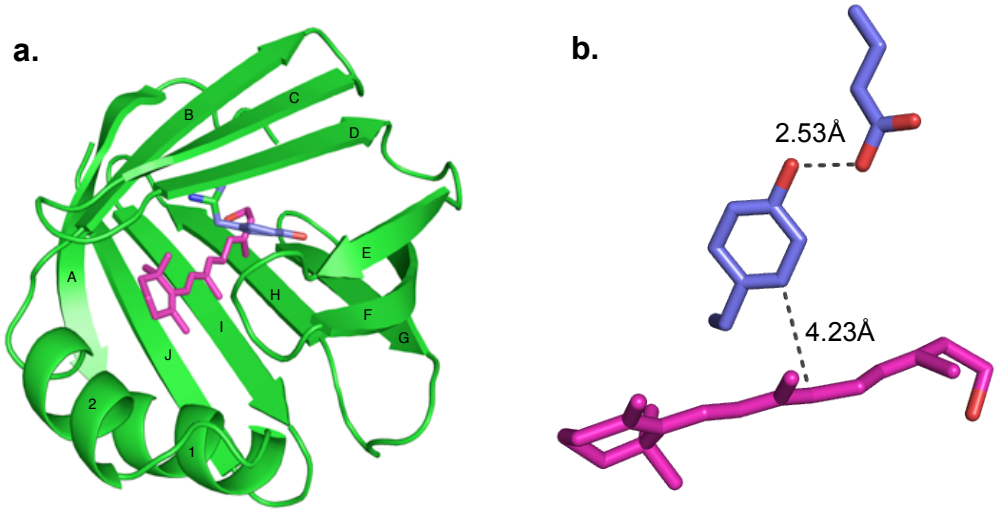


Figure IV-1: a. The crystal structure of wt hCRBPII bound with all-*trans*-retinal with the highlighted Y60 position in purple. **b.** The relative position of Tyr60 to all-*trans*-retinal.

polarizable residue, could result in a π-π interaction with the chromophore that could potentially delocalize the positive charge of the PSB along the chromophore. This modeling also suggested that the side chain of tryptophan would adopt a conformation that would be parallel to the polyene of the chromophore. Therefore the mutant Q108K:K40L:Y60W (KLY60W) was generated, expressed and purified as previously described in chapter II.

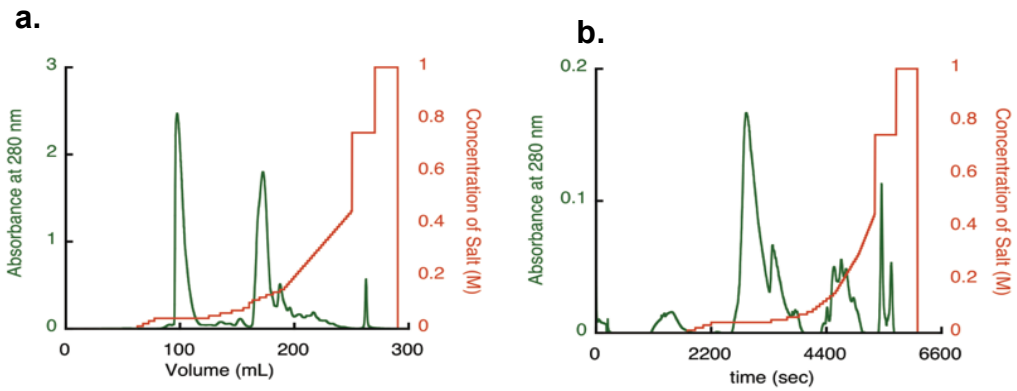


Figure IV-2: a. Chromatogram of KLY60W mutant at 280 nm, attributed to absorptions of tryptophan residues from the protein. **b.** Chromatogram of KLY60W mutant at 280 nm after refolding with urea.

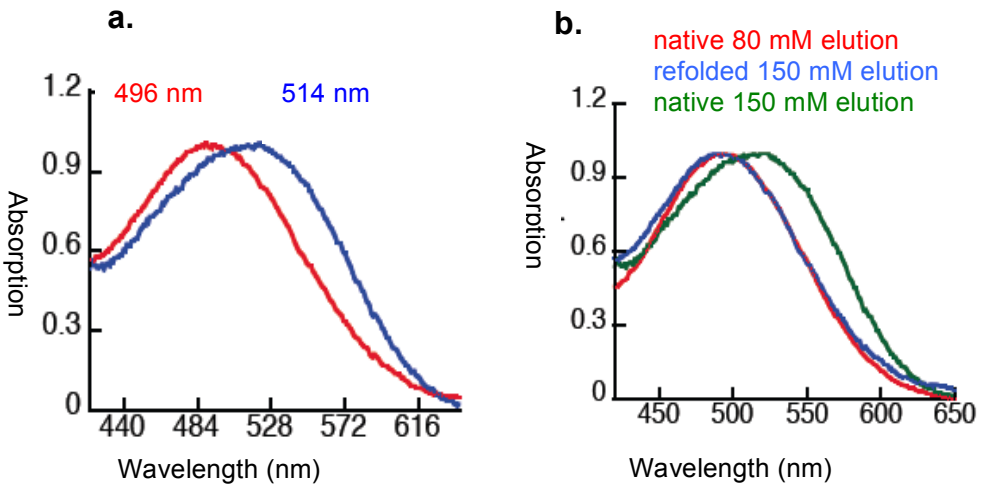


Figure IV-3: a. UV-vis spectral overlay of 40 mM salt elution (red) and 150 mM salt elution (blue) of Q108K:K40L:Y60W incubated with all-*trans*-retinal. **b.** UV-vis comparison of native 40 mM elution and 40 mM elution of refolded 150 mM elution.

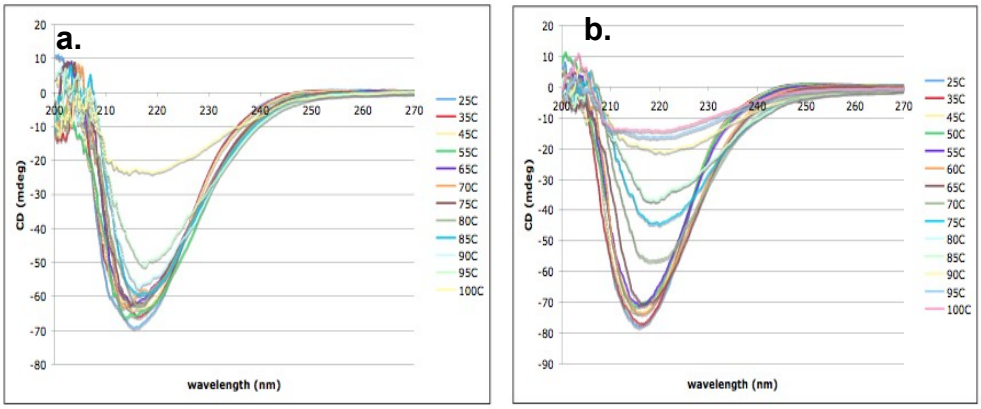


Figure IV-4: The melting point experiment in monomer and dimer of KLY60W: **a.** monomer KLY60W **b.** KLY60W dimer.

The second purification step, Source-Q ion exchange chromatography, surprisingly gave two protein fractions, one eluting at 80 mM NaCl, where hCRBPII mutants typically elute, and a second eluting at 150 mM NaCl; both fractions contained pure mutant CRBP II. The UV-Vis spectra of the all-*trans*-retinylidene PSB complex with protein from each of the fractions gave distinct spectra. The 80mM and 150 mM fractions showed a λ_{max} of 496nm and 514nm, respectively. (Figure IV-4A) Increasing the expression temperature from 16°C to 30°C resulted in higher amounts of the 150 mM fraction protein while the amount of the 80 mM fraction was unchanged.

At low expression temperature the ratio of high to low salt fractions was 80 to 20 while by increasing the temperature to 32°C this ratio reached 50% of each species.

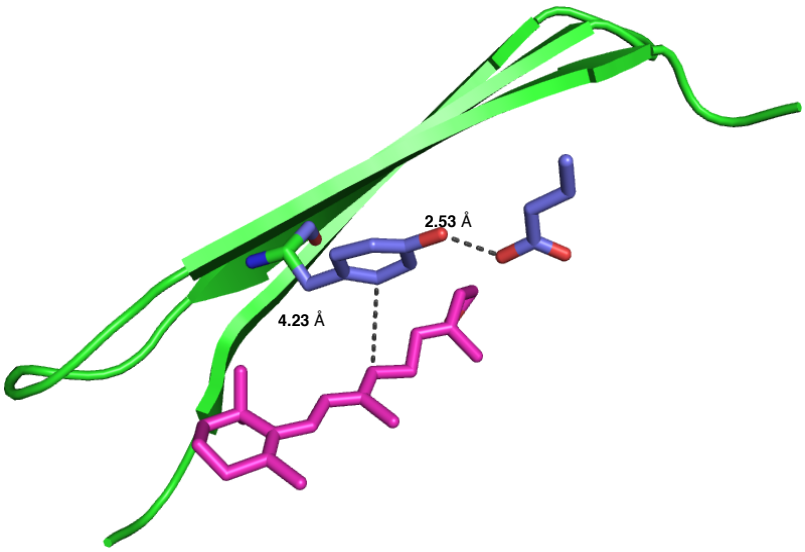


Figure IV-5: Y60 position is located on starting of the β strand D.

Size-exclusion chromatography showed that while the 80 mM protein eluted as a 15 kD monomer, the 150 mM fraction eluted as a 30 kD dimer. No equilibrium between the monomer and dimer forms was observed, even after several days at room temperature. Melting curves of each fraction showed the 150 mM fraction to be less stable than the 80 mM fraction, with the 150 mM protein denaturing at 65 °C, and the 80 mM protein stable to 95°C (Figure IV-5).

The fact that the two species displayed no evidence of interconversion at room temperature indicates the barrier to interconversion to be relatively high. Therefore, to show that the two protein species can indeed interconvert denaturation/renaturation experiments were carried out. First the dimer species was denatured in 8 M urea at a

relatively low protein concentrations (0.03 mM) and rapidly renatured by dilution and rapidly renatured by dilution. This results in renaturation occurring over a short period of time and at a relatively low concentration of protein. As shown by size exclusion chromatography, this resulted in exclusive production of the monomeric species, showing conclusively that the dimeric species can indeed be converted to the monomer. On the other hand, renaturation by dialysis at 4 °C at and high protein concentration (0.6 mM) gave a mixture of monomer and dimer, with the monomer amount about four times that of the dimer. Together, this indicates that the monomeric species is still thermodynamically favored over dimer, with the dimer species only forming at higher concentrations, because oligomerization must be reasonably fast relative to monomer folding. The UV-vis spectrum of the retinal PSB-bound refolded monomer was identical to the originally isolated protein, strongly indicating the two proteins to be structurally very similar if not identical (Figure IV-4-b).

IV-2-1:Dimerization by Applying a Single Site MutationY60W

Since both size exclusion chromatography and the crystal structure of the Q108K:K0L mutant clearly show this species to be exclusively monomeric, with dimerization only observed in the Q108K:K40L:Y60W (KLY60W) mutant, the natural question was whether the single Y60W mutation was sufficient to produce a dimer. To test this hypothesis Tyr60 was mutated to other hydrophobic and aromatic residues. To this end Dr. Wenjing Wang and Zahra Assar generated Q108K:K40L:Y60F (KLY60F) and single mutant Y60L. These two mutants were purified as previously described. In

both in both cases both proteins showed the dimer fraction and the ratio of dimer to monomer in Y60L is about 80% of the protein concentration. However, the single Y60W mutant exhibits behavior similar to that of KLY60W, eluting from the Source Q ion exchange column in two distinct salt fractions (80 and 150 mM), with the monomeric protein eluting at lower salt and the dimeric species eluting in higher salt, based on size exclusion chromatography. Together these results strongly indicate that Y60W forms a dimeric species similar to that of KLY60W.

IV-2-2: The Crystal Structure of Monomer Q108K:K40L:Y60W (KLY60W)

The crystal structure of the 80 mM species bound to all-*trans*-retinal as a PSB confirmed that it was indeed monomeric hCRBPII, with a structure very similar to that of wt hCRBPII and virtually all other mutant hCRBPII structures (Figure. 6A). The crystals have the same space group and similar unit cell dimensions to those of the other CRBP II mutants, with two molecules per asymmetric unit. Comparison of the hCRBPII KLY60W-retinal complex to the hCRBPII Q108K:K40L-retinal complex shows that though the overall fold of the two proteins is similar (with an RMSD 1.317Å) the introduction of the Y60W mutation caused significant local structural changes. Trp60 adopts a conformation similar to that seen for Tyr60 in the Q108K:K40L mutant, but the significantly larger size of the side chain has caused a significant displacement of the C-strand (by 1.35Å) compared to either the Q108K:K40L mutant, or wt hCRBPII, serving to slightly enlarge the binding cavity. This results in a rotation of the polyene portion of the chromophore by almost 90° relative to that of the Q108K:K40L mutant, while the

ionone ring remains stationary. This results in a "6-s-*cis*' conformation about the ionone ring-polyene torsion (Figure IV-6-b). Since the 6-s-*cis* conformation is thought to be more stable than the 6-s-*trans* conformation seen in Q108K:K40L hCRBPII. A possible interpretation of our results is that the extra space provided by the displacement upward of Trp60 relative to Tyr60 allows the chromophore to adopt this lower energy conformation. [24]

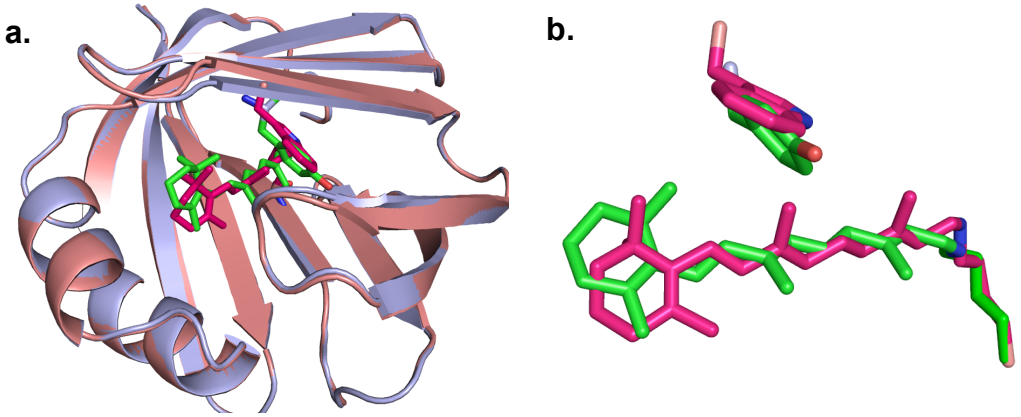


Figure IV-6. **a.** The overlaid crystal structure of Q108K:K40L mutant (light purple), the chromophore is highlighted in hot pink with Q108K:K40L:Y60W monomer (salmon) with highlighted chromophore in green. **b.** The highlighted chromophore in both structures.

Mutation of Tyr60 to Trp removes the hydrogen bonds between Trp60 and Glu72, which also has been observed in wt hCRBPII bound with retinol. It seems that the loss of this interaction and insertion of the larger amino acid may lead to displacement of the beta-C strand.

IV-2-3: Structure of KLY60W Domain Swapped Dimer

Crystals of the KLY60W dimer proved much harder to grow, appearing only after several months. Crystals that produced reasonable diffraction were rare, and only one incomplete data set was collected of this form (only 70% completion) due to decay in

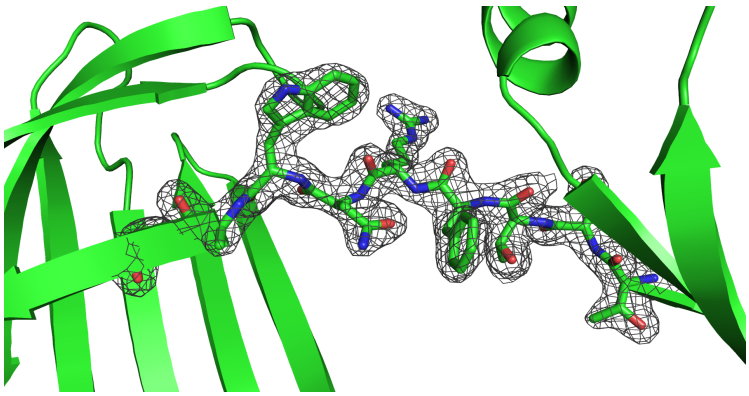


Figure IV-7: The crystal structure of Q108K:K40L:Y60W.

the X-ray beam. The structure was nevertheless determined and refined to 2.5Å resolution with two molecules per asymmetric unit. The electron density clearly shows that a domain-swapped dimer has formed, with a crystallographic two-fold axis of symmetry relating the two monomers in the dimer. Figure IV-7 shows electron density of 55-60 showing continuous density to the next molecule). Thus, the two monomers in the asymmetric unit each give rise, through a crystallographic two-fold axis of symmetry, to two independent dimers. An overlay of these two dimers (Figure IV-8-a) shows that they are very similar to one another (RMS= 0.491Å).

The resulting domain swapped dimer represents a novel structural architecture. The opening to the binding cavities are now face to face, creating a continuous internal cavity that stretches from one side of the dimer to the other (Figure IV- 8-b). The bottom of the structure forms a continuous beta sheet that also spans the length of the

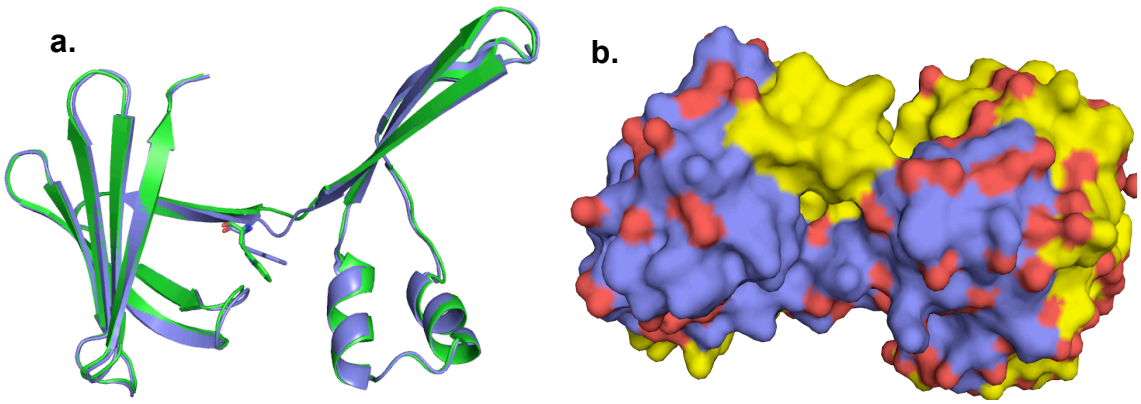


Figure IV-8: a. The two chains of the Q108K:K40L:Y60W dimer overlaid (blue, chain B, green chain A showing a minor structural difference. **b.** The surface structure of the dimer, showing the extensive binding cavity.

dimeric structure, reminiscent of the continuous beta sheet that spans the two direct repeats of the TATA-binding protein. A search of the various structural fold databases (SCOP 1.75) reveals no other protein of similar architecture.

The major conformational changes that lead to domain swapping in this structure are not confined to a single residue in this case, but involve large differences in main-chain torsion angles of several of the residues between 55-61. Interestingly, Trp60 has flipped from the interior of the protein to the exterior in both dimers, exposing the hydrophobic residue to solvent. Together the conformational change in these residues give rise to the domain swapped dimer as the remainder of the structure is virtually identical to the monomeric KLY60W structure (Figure IV-10). The domain swapping is extensive in this

case, involving almost half of the molecule, from residue 1-56 which includes 3 beta strands and two helices. This corresponds to an 83% extent of swapping (as defined in the domain swapping 3D knowledge database (http://caps.ncbs.res.in/3dswap/help.html). The structure of the KLY60W domain swapped dimer provides a potential explanation for the bathochromic shift of the absorption spectrum of the bound retinylidene PSB in the dimer versus monomer. It was previously shown that sequestration of the retinal PSB chromophore from bulk solvent invariably leads to red-shifted spectra.[18] The face to face juxtaposition of the mouths of the binding cavities in the domain swapped dimer result in a binding cavity significantly isolated from solvent relative to the monomer, likely explaining the significantly red-shifted spectrum of the dimer relative to monomer.

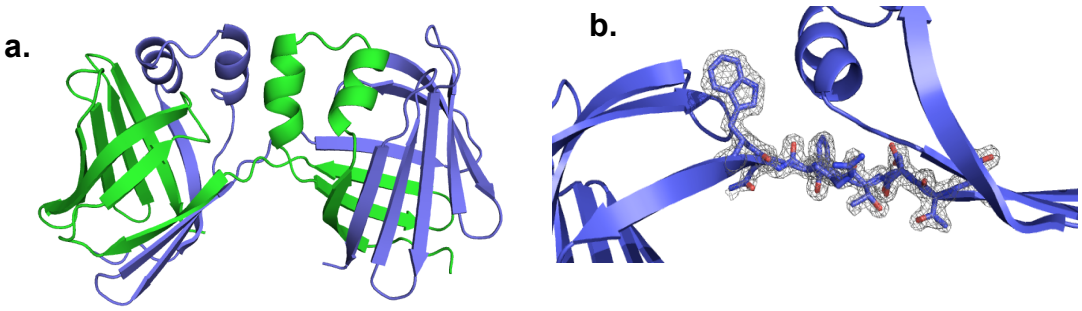


Figure IV-9: **a.** The crystals structure of the domain swapped Y60W dimer (chain A). **b.** The 2Fo-Fc electron density map around the residues between 50-60 (contoured at 1σ) illustrates the continuous electron density that crosses the dimer interface.

IV-2-4: Structure of the Y60W Dimer

To ascertain whether a single mutation could produce the dimer, the hCRBPII Y60W mutant was generated, expressed and purified.

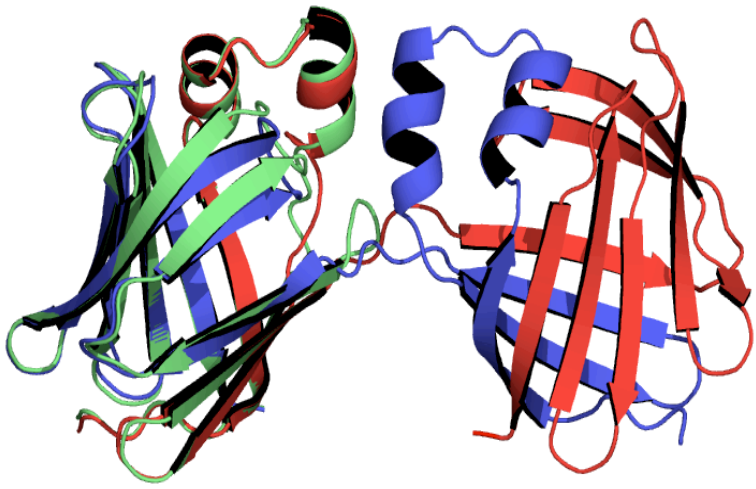


Figure IV-10: The overlaid structure of KLY60W monomer (green) and Y60W dimer, chain A in red and chain B in blue.

This mutant behaved very similarly to KLY60W, eluting in two fractions from the SOURCE Q ionic exchange column, one at 80 mM salt and a second at 150 mM salt. As previously observed, the 80 mM fraction eluted consistent with a monomeric protein by size exclusion chromatography, while the 150 mM fraction eluted exclusively as a dimer. The dimeric protein crystallized relatively easily, this time in the $P2_12_12$ space group, with one dimer in the asymmetric unit (Figure IV-9-a). The electron density map clearly showed continuous electron density across the two domains, confirming the

presence of a domain swapped dimer (Figure IV-9-b). The high quality of the crystallographic data (to 1.7Å) allowed a far more precise picture of the nature of the domain swapping. The structure of the two domains of the dimer are each very similar to monomeric KLY60W hCRBPII with RMS= 0.651Å (Figure IV-10).

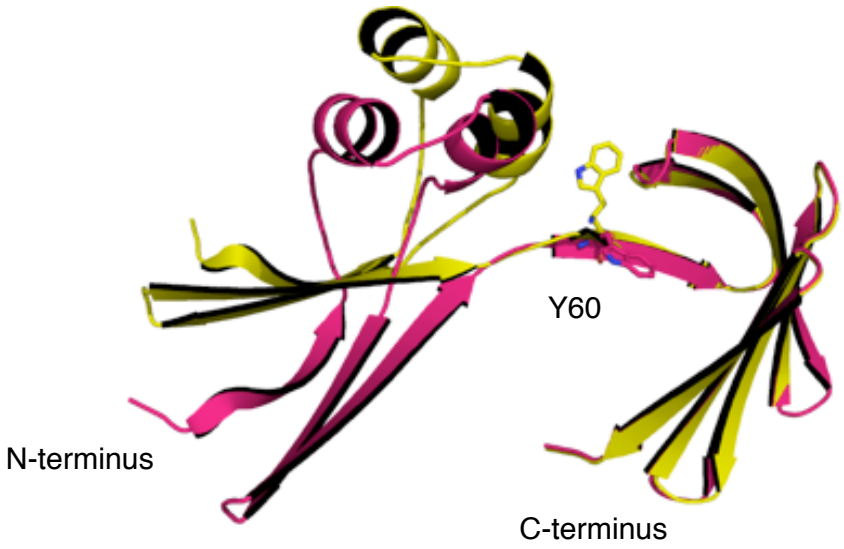


Figure IV.11: The two chains of the Y60W dimer overlaid (red, chain A, yellow chain B showing a major structural difference in these two chains.

Significant differences in the two structures are localized to the region spanning Ser55 to Trp60. The most significant change occurs in the phi/psi angle of Thr56. Surprisingly, comparison of the two dimer structures (KLY60W and Y60W) shows that the relative orientation of the two domains is significantly different in the two structures.

As shown, when the C-terminus (from 56-133) of the two domains are overlaid, the N-terminus is displaced by as much as 10Å in the two structures. Moreover, overlay of the four chains (1 from each of the two distinct dimers in the KLY60W structure and the two from the Y60W dimer) 3 of the four appear to be quite similar while the fourth (Y60W molecule A) displays a large displacement relative to the other three. Table 1 shows the backbone torsion angles in and around the hinge region for the monomer structures. As the data indicates the torsion range in the dimer structures (wt CRBP and KLY60W monomer) is not a wide range, while It is immediately evident from Table IV-2 that there are significant differences in the torsion angles that give rise to domain swapping in the four chains. Most notable is the fact that only one angle, the psi angle of Thr56, deviates significantly from its monomer value in Y60W chain A. It is therefore possible to produce the domain swapped conformation by rotation about a single torsion angle in the hinge region. It would appear that the additional large torsion angle deviations in Y60W chain B are required for mating of the two chains to form the dimer. On the other hand, several torsion angles deviate significantly from their monomeric values in the two KLY60W chains. Seven deviate substantially from the monomer values. It is important to remember that a crystallographic two-fold axis relates two KLY60W chain A's to make a single dimer, while a second crystallographic two fold axis relates two KLY60W chain B's to make a second dimer. The two additional mutations Q108K and K40L, are far from the hinge region and do not appear to contribute to the conformational differences in the domain swapped dimers (Crystal structure of Q108K:K40L, PDB entry 4EXZ).[18]

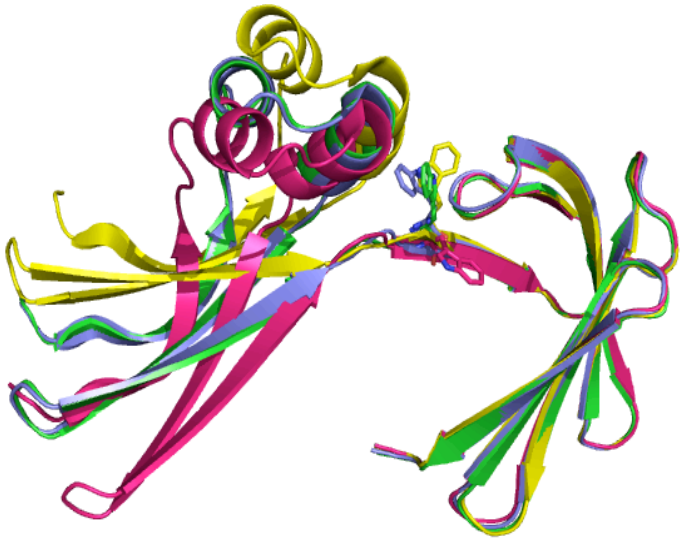


Figure IV-12: The overlaid of the four chains of the Q108K:K40L:Y60W dimer (green, chain A and blue, chain B) with Y60W dimer (yellow, chain B and red chain A).

Therefore, it seems the additional torsional rotations are the result of the crystallographically enforced two-fold axis. This leads to the conclusion that the four distinct molecules seen in the two mutant crystal structures are the result of crystal packing, allowed by the significant flexibility in the hinge region, with only the large rotation in the psi angle of Thr56 a requirement of domain swapping. Besides the difference in backbone torsion angles, there is also a significant difference in the dispensation of the mutated Trp60 residue. While Trp60 is flipped out of the binding pocket and exposed to solvent in both molecules of the KLY60W mutant structure, and in chain B of the Y60W structure, it is rotated inward in chain A of Y60W. As previously described, Tyr60 is completely buried inside the binding pocket in all known structures of hCRBPII. It is important to remember here that while there is a two-fold axis of

symmetry that relates the two chains that make up a dimer in the KLY60W structure, this is not the case in Y60W, because the two molecules that make up a dimer are quite different in this structure.

Table IV-1: The monomer psi and phi angles for 10 amino acids

Residue	Monomer range Phi		Monomer range Psi	
THR/51	-118.75	-128.77	144.78	150.60
LYS/52	-131.96	-142.20	119.02	127.14
THR/53	-106.75	-114.46	118.75	132.93
THR/54	-124.59	-130.11	146.34	152.92
SER/55	-141.03	-171.97	164.46	177.47
THR/56	-66.02	-78.04	-7.84	-28.65
PHE/57	-76.60	-96.89	-41.71	-100.34
ARG/58	-93.84	-165.78	116.95	159.20
ASN/59	-90.83	-120.79	145.26	161.72
TRP/60	-141.30	-151.03	113.78	153.91

Table IV-2: The phi and Psi angles in KLY60W and Y60W dimer structures

Residue	KLY0W-dimer chain A		KLY0W-dimer chain B		Y60W-dimer chain A		Y60W-dimer chain B	
	phi	psi	phi	psi	phi	psi	phi	psi
THR/51	-127.211	117.631	-115.367	147.721	-126.267	159.521	-121.246	143.204
LYS/52	-96.111	148.782	-124.911	131.791	-151.915	145.905	-130.250	121.593
THR/53	-138.837	115.942	-106.746	156.263	-123.728	139.693	-116.726	129.982
THR/54	-96.897	152.451	-149.731	115.200	-131.953	159.823	-121.777	135.471
SER/55	-124.717	130.491	-125.271	120.419	-145.887	136.661	-96.772	145.545
THR/56	-130.999	131.099	-150.457	-166.945	-65.386	146.007	-138.037	138.997
PHE/57	-145.711	153.215	-128.353	-160.388	-115.020	-74.831	-91.315	131.447
ARG/58	-148.754	-68.664	-128.200	-60.396	-130.341	152.976	-122.702	125.515
ASN/59	172.364	-105.767	76.781	-118.583	-124.810	136.948	-87.716	178.362
TRP/60	60.384	75.556	50.189	100.065	-136.578	141.385	-82.181	141.376

IV-2-5: The Nature of Domain Swap Dimerization in HCRBPII

Though the detailed mechanism of domain swapping in hCRBPII is not known, the data so far obtained points to several important aspects of the process. First the fact that the KLY60W monomer is significantly more stable than its domain swapped dimer indicates that in this case domain swapping is a kinetic and not thermodynamic product in the process. The fact that no dimer to monomer interconversion is ever observed for either mutant indicates that the dimer represents a kinetic trap in the monomer refolding process. Both *in vitro* refolding experiments, where monomer formation is strongly favored at low concentrations, but less favored at high concentrations, and *in vivo* expression, where the monomer is strongly favored at lower temperatures where the expression rate is slower and the concentration of protein in the process of folding is expected to be lower, indicate that high concentrations of partially folded intermediates is required for dimer formation. Indeed folding at higher concentrations often favors the production of domain-swapped oligomers. Third is the fact is mutation of Tyr60 to more hydrophobic residues is both necessary and sufficient for domain swapped dimerization, though this residue is 4 amino acids away from the hinge loop. Fourth, it is found in the non-native flipped out conformation in both domain swapped dimers. Together this data indicates that folding of the protein happens in two steps, with the first step being the independent folding of the N- and C-terminal regions independently. The N-terminus forms 4 beta-stands (bA-bD) and two helices (a1,2) and the C-terminus forms 8 beta-strand (bE-bJ). Formation of these two subdomains is thought to be the fast step of protein folding and after this step the slow step which is

formation of the tertiary structure of the protein occurs. In this step the hinge region plays the crucial role. Formation of the monomer has been slowed by the presence of Trp60 and that the rate-limiting step in folding for these mutants may be the flipping of the two sheets onto one another, allowing the extended sheet intermediate to dimerize when the concentration is high. Static structures cannot explain the details of kinetic transition states, but the presence of the flipped out conformation of Trp60 in the dimers may provide a clue. It is possible that the transition from this flipped out conformation to the flipped in conformation seen in the monomer may play a role in slowing the rate of monomer folding. Mutation of Tyr60 to smaller aromatic residues (KLY60H and KLY60F) does not result in domain swapping, indicating that the size of the residue is important to the mechanism. Further, mutation of another residue in the same region to Trp (R58W) does not lead to dimerization, indicating that the Tyr60 position may have special significance. However, this in no way implies that the native protein undergoes a similar pathway, as the presence of Trp60 may be responsible for the presence of the folding intermediate required for domain swapping. By applying this mutation the energy barrier between the open and closed states is increased and therefore the intermediate form in the open state was trapped. Nevertheless the domain swapped dimerization of Y60W CRBP II shows that a single mutation, far from the actual hinge region, can give rise to an extensive domain swapping event, involving almost 50% of the protein.

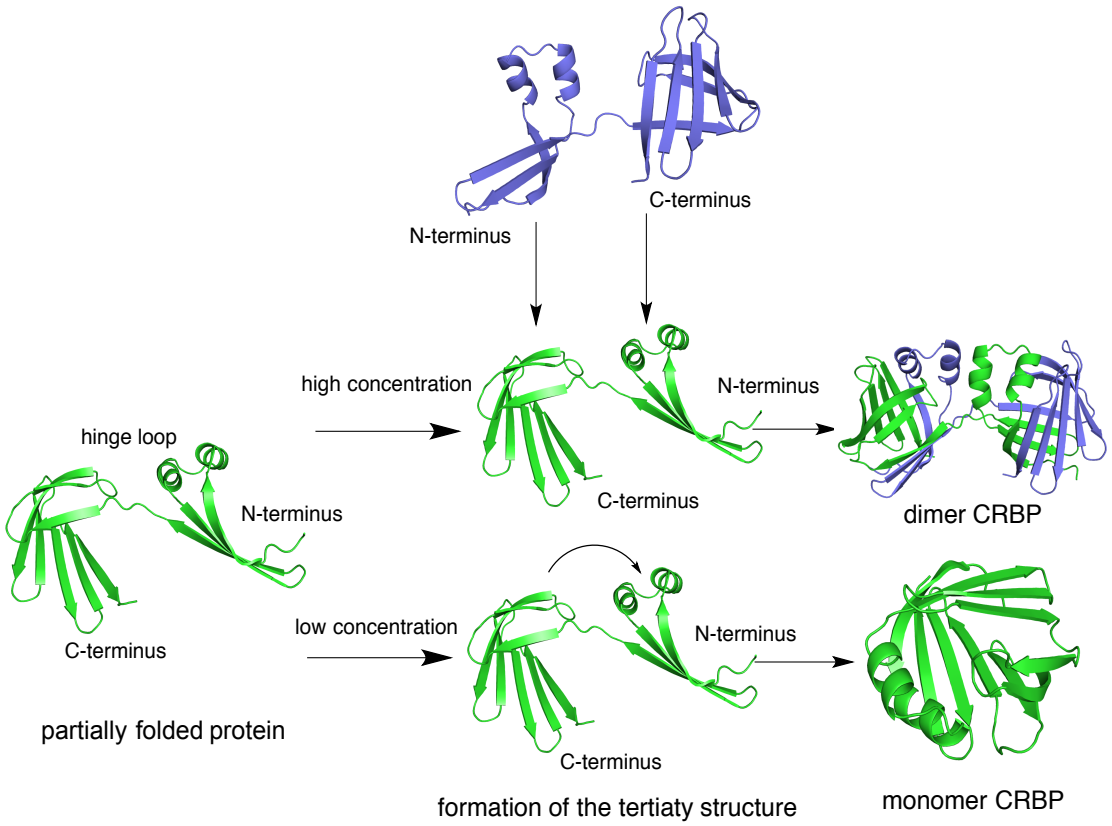


Figure IV-13: The proposed mechanism for formation of monomer and dimer in Y60W hCRBPII.

IV-3: Experimental Procedure

The cloning, expression and purification of hCRBPII mutants was performed as previously described. To increase the ratio of the dimer protein the expression temperature was increased to 32°C overnight. Size exclusion chromatography was performed using a Superdex 200 16/75 column (GE Health Sciences), packed in-house in a buffer containing 10 mM Tris, 100 mM NaCl, pH 8.0.

PCR Primers for Y60W Mutant:

Forward: 5’- CACATTCCGCAACTGGGATGTGGATTTCAC-3’
Reverse: 5’- GTGAAATCCACATCCCAGTTGCGGAATGTG-3’

IV-3-1: Crystallization of HCRBPII Mutants

Each mutant protein was concentrated to 10 mg/mL in a buffer containing 10 mM Tris, 100 mM NaCl, pH 8.0. The complex of protein with all-trans-retinal was prepared by adding 2 equivalents of retinal dissolved (30 mM retinal in ethanol) in ethanol and incubated for one hour at room temperature. The ethanol concentration should not be more than 10%. Crystals were grown using the hanging drop vapor diffusion method using 1 mL of protein solution and 1 mL of crystallization solution in the drop and 1 mL of crystallization solution in the reservoir. All crystallization trials were wrapped in aluminum foil to limit light exposure to the chromophore. The best crystals grew in 5 days and reached maximum size in one week using a crystallization solution of 30% PEG 4000, 0.1M $CH_3COONa.3H_2O$ pH 4.6, 0.1 M CH_3COONH_4. The crystals were soaked in a cryoprotectant solution (30% PEG 4000, 0.1M sodium acetate pH 4.6-4.8, 0.1 M ammonium acetate, 15-20% glycerol), flash frozen in loops in liquid nitrogen and stored in a liquid nitrogen dewar prior to data collection at the synchrotron.

IV-3-2: Data Collection and Data Refinement

All of the diffraction data were collected at the Advanced Photon Source (APS) (Argonne National Laboratory IL) LS-CAT, (sector 21-ID-D,F,G) using a MAR300 detector and 1.00 Å wavelength radiation at 100K. The initial diffraction data were indexed, processed and scaled using the HKL2000 software package.[19] The structures were determined by molecular replacement using MOLREP in the CCP4 software package and human CRBPII (PDB entry 2RCQ) as a search model.[20,21] Though the apparent space group of the KLY60W structure was P21212, due to almost perfect merohedral twinning, the actual space group was P21. The initial electron density map was produced by using REFMAC5 in the CCP4 package. All of the, rebuilding, placement of water molecules and etc. were done using COOT.[22] The chromophore was built into the monomeric KLY60W mutant structure using the jligand program to generate restraints.[23] All of the refinements were done by REFMAC (restrained refinement) until the final R-factor was between 18-20%.

Table IV-3: X-ray crystallography data and refinement statistics

	Q108K:K40L:Y60W (monomer)	**Y60W (dimer)**	**Q108K:K40L:Y60W (dimer)**
Space group	P1	P21212	P1
a(Å)	30.08	62.62	36.62
b(Å)	36.05	109.77	61.07
c (Å)	63.92	36.32	64.30
α(°)	90.94	90.00	89.82
β(°)	91.99	90.00	90.13
δ(°)	113.93	90.00	90.18
Molecules per Asymmetric Unit	2	2	2
Total reflection	124534	345311	214807
Unique Reflection	41751	55068	19239
Completeness (%)	92.8(68.0) [a]	98.6(100.0) [a]	72.9(57.6)
Average I/σ	50.30(3.58) [a]	66.13(7.71) [a]	29.48(1.38)[a]
R_{merge}(%)	4.1(29.3)	5.2(31.7) [a]	4.3(49.8)[a]
Resolution (Å) (Last Shell)	21.30-1.47(1.52-1.50)	50.00-1.70(1.73-1.70)	40.00-2.50(2.54-2.50)[a]
R_{work}/ R_{free} (%)	19.86/25.06	18.17/23.94	21.48/34.20
RMSD from Ideal Values			
Bond Length (Å)	0.025	0.023	0.012
Bond Angle (°)	2.36	2.11	1.45
Average B factor	19.98	19.92	30.17
Number of water Molecules	228	217	33

Values in the parenthesis refer to the last resolution shell

REFERENCES

REFERENCES

1. Ong De (**1984**) a novel retinol-binding protein from rat - purification and partial Ccharactrization. *Journal of Biological Chemistry* 259(3):1476-1482.

2. Bennett MJ, Schlunegger MP, & Eisenberg D (**1995**) 3D domain swapping - a mechanism for oligomer assembly. *Protein Science* 4(12):2455-2468.

3. Liu L GA (**2012**) *Protein and Nucleic Acid Folding: Domain Swapping in Proteins* (Oxford Academic Press).

4. Dobson CM (**2003**) Protein folding and misfolding. *Nature* 426(6968):884-890.

5. Selkoe DJ (**2004**) Folding proteins in fatal ways (vol 426, pg 900, **2003**). *Nature* 428(6981):445-445.

6. Shtilerman MD, Ding TT, & Lansbury PT (**2002**) Molecular crowding accelerates fibrillization of alpha-synuclein: Could an increase in the cytoplasmic protein concentration induce Parkinson's disease? *Biochemistry* 41(12):3855-3860.

7. Fandrich M, Fletcher MA, & Dobson CM (**2001**) Amyloid fibrils from muscle myoglobin - Even an ordinary globular protein can assume a rogue guise if conditions are right. *Nature* 410(6825):165-166.

8. Sunde M & Blake C (**1997**) The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Advances in Protein Chemistry, Vol 50: Protein Misassembly,* Advances in Protein Chemistry, ed Wetzel R), Vol 50, pp 123-159.

9. Cleary JP*, et al.* (**2005**) Natural oligomers of the amyloid-protein specifically disrupt cognitive function. *Nature Neuroscience* 8(1):79-84.

10. Scherzinger E*, et al.* (**1997**) Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell* 90(3):549-558.

11. Hayden MR, Tyagi SC, Kerklo MM, & Nicolls MR (**2005**) Type 2 diabetes mellitus as a conformational disease. *JOP : Journal of the pancreas* 6(4):287-302.

12. Marianayagam NJ, Sunde M, & Matthews JM (**2004**) The power of two: protein dimerization in biology. *Trends in Biochemical Sciences* 29(11):618-625.

13. http://caps.ncbs.res.in/3dswap/help.html.

14. Minton AP (**2000**) Implications of macromolecular crowding for protein assembly. *Current Opinion in Structural Biology* 10(1):34-39.

15. Ellis RJ (**2001**) Macromolecular crowding: an important but neglected aspect of the intracellular environment (vol 11, pg 114, **2001**). *Current Opinion in Structural Biology* 11(4):500-500.

16. Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus. *Structure* 9(11):1017-1027.

17. Macdonald PN & Ong DE (**1987**) BINDING SPECIFICITIES OF CELLULAR RETINOL-BINDING PROTEIN AND CELLULAR RETINOL-BINDING PROTEIN, TYPE-II. *Journal of Biological Chemistry* 262(22):10550-10556.

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

19. Z. Otwinowski, W. Minor, Macromolecular Crystallography, Pt A 276, 307 (**1997**).

20. S. Bailey, Acta Crystallographica Section D-Biological Crystallography 50, 760 (**1997**)

21. A. Vagin, A. Teplyakov, J. Appl. Crystallogr. 30, 1022 (**1997**).

22. G. N. Murshudov, A. Vagin, E. J. Dodson, J. Appl. Crystallogr. 30, 1022 (**1997**).

23. G. N. Murshudov et al., Acta Crystallogr. D Biol. Crystallogr. 67, 355 (**2011**).

24. Rajput, J.; Rahbek, D. B.; Andersen, L. H.; Hirshfeld, A.; Sheves, M.; Altoe, P.; Orlandi, G.; Garavelli, M., Probing and modeling the absorption of retinal protein chromophores in vacuo. Angew. Chem. Int. Ed. Engl. **2010**, 49 (10), 1790-1793.

CHAPTER V

Mutation, Over-expression, Purification, Crystallization, X-ray diffraction, Structural solution and Refinements of CRBPII Mutants

V-1: Material and Methods

Expects few mutations (T51I, Y60W, Q108H, T51V) the rest of mutations were done by Dr. Wenjing Wang in Prof. Borhan group at department of chemistry in Michigan State University. The all the mutations were constructed by **Stratagene's QuickChange Site-Directed Mutagenesis Protocol**, which is a high performance protocol that takes only one day for mutation. This method is widely used for insert/delete and mutations. This method is performed by pfuTutrbo DNA polymerase and temperature circle in PCR instrument. The hCRBPII was cloned into pET17b expression vector by Dr. Wenjing Wang using NdeI and XhoI restriction enzymes.

Table V-1: The PCR protocol for wt hCRBPII mutants

DNA template	Primer forward	Primer reverse	pfuTurbo DNA polymerase	dNTP	Bufer (x10)	dd water	Total volume
50 ng	20 pmol	20 pmol	1 unit	1μL	5μL	Up to 50 μL	50 μL

PCR Protocol for hCRBPII mutagenesis:

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

Tm stands for the melting temperature of the primer for each mutation, which is calculated using the following website:

http://www.basic.northwestern.edu/biotool/OligoCal.html

Wt HCRBPII Gene Sequence

ATGACGAGGGACCAGAATGGAACCTGGGAGATGGAGAGTAATGAAAACTTT

GAGGGCTACATGAAGGCCCTGGATATTGATTTTGCCACCCGCAAGATTGCA

GTACGTCTCACTCAGACGCTGGTTATTGATCAAGATGGTGATAACTTCAAGA

CAAAAACCACTAGCACATTCCGCAACTGGGATGTGGATTTCACTGTTGGAGT

AGAGTTTGACGAGTACACAAAGAGCCTGGATAACCGGCATGTTAAGGCACT

GGTCACCTGGGAAGGTGATGTCCTTGTGTGTGTGCAAAAGGGGGAGAAGA

GAACCGCGGCTGGAAGAAGTGGATTGAGGGGGACAAGCTGTACCTGGAGC

TGACCTGTGGTGACCAGGTGTGCCGTCAAGTGTTCAAAAAGAAGTGA

Wt HCRBPII Amino Acid Sequence

T R D Q N G T W E M E S N E N F E G Y M K A L D I D F A T R K I A V R L T Q
T L V I D Q D G D N F K T K T T S T F R N Y D V D F T V G V E F D E Y T K S
L D N R H V K A L V T W E G D V L V C V Q K G E K E N R G W K K W I E G D
K L Y L E L T C G D Q V C R Q V F K K K

The PCR product is treated with DpnI endonuclease. This enzyme is specific to digest the unmutated and parental DNA and methylated DNA and the target is GATC sequence. This enzyme digests and destroys the entire parental DNA and methylated ones and leaves the mutated one. The process takes for 20 min at 37°C.

The Primes Used Were as Follows:

T51I

For: GGT GAT AAC TTC AAG ATA AAA ACC ACT AGC ACA

Q108H

For: AAC CGC GGC TGG AAG CAC TGG ATT GAG GGG GAC

Q108L

FOR: GGC TGG AAG TGG ATT GAG

F57S

For: ACA AGC ACA TCC CGC AAC TAT

Transformation of Mutated Plasmid in to Competent Cells: The modified protocol by Sambrook et al has been used for all transformation. The sterile condition should be maintained during the procedure. 100μL aliquot of competent cell (DH5α) was thawed on ice. Depends on the quantity of PCR between 1 to 10μL (10-100 ng) DNA was added to the cells and mixes gently (without pipette it up and down). The competent cells should never be vortexed. The cells incubated in ice for 30 min and they were heat

shocked in the water bath at 42°C for 20 seconds. The sample immediately needs to return back to the ice for 2 min. 0.5 ml of prewarmed (37°C) LB media was added and incubated at 37°C for one hour. After one hour the solution spread on the LB/agar plate containing appropriate antibiotic (Amp 100 μg/mL). The plates finally were incubated for 16-20 hours at 37°C.

Plasmid Purification: A single colony was grown up in 3 mL LB (containing AmP) and grows over night and then purified using the Qiagen® mini plasmid purification kit with min prep protocol. The recovered DNA was eluted with 30 mL of EB buffer and analyzed by nanodrop UV-vis for determination of the purity and concentration. The DNA is sent for sequencing to The Research Technology Support Facility (RTSF) at MSU using T7promoter.

Protein Over-Expression: Human recombinant hCRBPII was expressed and purified by Wang et al described procedure. The expression plasmid pET-17b was transformed into E.coli strain BL21(DE3)pLysS cells. The transformed cells were grown at 37°C in LB/agar plates containing antibiotics (Amp^{100}). A single colony was inoculated in 50 ml of LB containing the same amount of antibiotics and grown overnight with shaking at 37°C. This culture then transferred to 1L of LB with the same antibiotics. When OD^{600} reached between 0.6 to 0.8. The sample was induced by 0.1 nM IPTG and grows over night at room temperature (22°C). The cells were harvested next day (500 rpm, 20 min, 4°C) by centrifugation. The cells can be stored at -20°C for later purification.

Protein Purification: All the steps of purification were done I cold room or using ice bath. The frozen cells (obtained from 2L) were resuspendedn In 80 mL of (10mM Tris-HCl, 10mM NaCl pH8.0). The resuspended cells were lysed by sonication (three 1 min with 1 min relaxation between each) and centrifuge for 20 min, 13000 rpm. The supernatant was subjected to FastQ, which is a strong anion exchanger (Q-sepharose Fast Flow Resin) column chromatography. (100 mL resin with capacity ~ 20 mg protein, Amersham Biosciences). The column at first per-equilibrated with the 100 mL of 10mM TRIS-HCl pH 8.0 buffer. After the flow through was collected the column was washed using the same buffer. 30 μL of sample is collected for running polyacrylamide gel electrophoresis (SDS-PAGE). The concentration of the washed solutions was measured with biorad protein assay. Finally the protein was eluted with 40 mL of 10mM TRIS_HCl, 100 mM NaCL pH 8.0 buffer. All of the samples were analyzed on 16% SDS-PAGE to test for purity of sample. The protein samples pooled together and desalted using Centriprep® centrifugal filter (10,000 MWCO) for the next step of purification. For the last step of purification FPLC system ion exchange (BioLog Due Flow, Biorad) using Source15Q (Amersham Biosciences) has been used.

Extinction Coefficient Determination of HCRBPII: For calculation the extinction coefficient of hCRBPII Gill and Von Hippel method has been used.

Theoretical extinction coefficient for denatured protein εd =

$$5690 \times Trp + 1280 \times Try + 120 \times Cys$$

in wt hCRBPII : 4 Trp, 4 Tyr, 2 Cys

and therefore the theoretical εd for hCRBPII is :

5690x4+1280x4+120x2 = 28,120

The absorption of the native protein and denatured protein (6M guanidine hydrogen chloride) at 280 nm were measured at exact same protein concentration.

The extinction coefficient for the native protein is derived from the following equation:

$\varepsilon n = \varepsilon d \times (A280 \text{ native}) / (A280 \text{ denatured})$.

V-2: Crystallization and Data Collection

All the crystals were grown using the hanging drop (vapor diffusion method). In this method a drop (2-5μL) containing the mixture of protein and reservoir solution is equilibrated against the reservoir solution (1 mL). At first around 200 different conditions have been tried and the best crystals of the wt hCRBPII were grown in one of the Hampton research crystal screen solution containing PEG 4000, ammonium acetate and sodium acetate. The best crystals were obtained by screening around this condition (30% PEG 4000, 1 M ammonium acetate and 1M sodium acetate pH 4.5) and both apo and holo crystals were grown in the same condition. Since retinal is very sensitive to the light all of the crystal boxes were wrapped in the aluminum foil to protect them against light. Expect the crystals were grown after 3 weeks and reached to their maximum size in one week. For preventing of degradation of retinal the crystals were flash frozen after one week. The crystals were cryo-protected by soaking in the cryo-protectant solution consisting the reservoir solution with 20% glycerol. The frozen crystals were taken to the Advanced Photon Source (APS) synchrotron facility (Argonne, IL) for data collection. All of the diffraction collections were collected under steam of nitrogen gas.

The collected data were integrated and scaled using the HKL2000 package at APS. The processing and refinement were done at home.

Crystallization of all Holo-mutants HCRBPII Expect Q108K:K40L:T51V:T53C:R58W:T29L:Y19Y:Q4R and Q108K:K40L:T51V:T53C:R58W:T29L:Y19Y:Q4A:

All-*trans*-retinal was purchased from Sigma-Aldrich. The complex of protein and retinal were prepared fresh before setting the crystallization box. Since retinal is a light sensitive compound the stock solution of that was kept in black eppendorf tube in freezer and all the process of incubation of that was done in dark and under the red light. HCRBPII mutants cannot tolerate more than 10% of ethanol. Therefore the saturated retinol solution in ethanol was prepared ~ 10^{-3} M. WT CRBP and the mutants of this protein are very tolerant for concentration and can be concentrated more than 20 mg/mL but usually concentration of 15 mg/mL has been chosen. The lower concentration of protein leads to longer for crystallization and also the smaller crystals were obtained. The crystals with appear at room temperature after three days and reached to the maximum size in a week. The procedure of the crystallization does not need to be in dark and the crystals can be flash freeze at the normal light. The crystals are colorful and the colors depend on the absorbance of the protein. All of the crystals without R58 mutations are perfectly rectangular, while the one with R58F, R58Y and R58W mutations form cluster. The diffraction data were collected at sector 21-LS-CAT beam line D at Argonne National lab (IL). Since the crystals are in P1 space group for each crystal 360 frames were collected with 1° oscillation. The distance of the detector to the crystals was from 100 mm to 150 mm.

Crystallization of Holo-Q108K:K40L:T51V:T53C:R58W:T29L:Y19Y:Q4R and Q108K:K40L:T51V:T53C:R58W:T29L:Y19Y:Q4A mutants: These two mutants with low pK_a did not crystallized at the same crystal condition at room temperature and the crystals grow at cold room (4°C) after more than two weeks. Because of the low pK_a of these two mutants the color of the crystals were light orange- yellow. The crystals were flash frozen in with liquid nitrogen and cryo protectant solution. The diffraction data were collected at Argonne National Lab sector 21, beam line D. for each crystal 360 diffraction data were collected with 1° oscillation.

Structure solution and refinement: All of the structure determinations and refinement were performed using CCP4 (Collaborative Computational Project, Number 4, 1994) package. The structure of the WT CRBP was determined using molecular replacement (rigid body refinement) using crystal structure of wt hCRBPII (2RCQ) as the initial model to produce the initial electron density map. The structures of the rest of mutants were solved using the high resolution of the wt hCRBPII. All of the mutation and rebuilding the structures and adding the waters were performed in COOT. All refinements were carried out using REFMAC5 in CCP4 package. JLIGAND was used for generating the ligands. The refinements statistics are reported in Table IV-2.

Table V-2: X-ray crystallography data and refinement statistics

	Q108K:K40L:T51V	Q108K:K40L:T53C	Q108K:K40L:T51V :R58F	Q108K:K40L:T51V: T53S
Space group	P1	P1	P1	P2
a(Å)	29.739	29.947	30.057	53.917
b(Å)	35.931	35.890	35.882	36.608
c (Å)	64.187	63.866	64.163	65.651
α(°)	90.78	91.01	90.91	90.00
β(°)	92.56	92.24	91.21	90.10
δ(°)	113.22	113.59	113.76	90.00
Resolution (Å)	50-1.77(1.77-1.73)[a]	50-1.64(1.64-1.59)[a]	50-1.51(1.51-1.47)[a]	50-1.57(1.57-1.53)[a]
R_{work}/ R_{free} (%)	24.68/21.75	28.20/21.20	27.46/21.17	22.7/18.57
Total reflection	82430	132782	147350	119972
Unique Reflection	31572	59796	69283	71252
Completeness (%)	92.9(94.1)[a]	97.3(84.0)[a]	97.4 (82.5) [a]	97.3(79.3)[a]
Molecules per Asymmetric Unit	2	2	2	2
I/σ	44.23(5.94)[a]	53.68(4.17)[a]	53.18(2.9)[a]	42.25(1.81)[a]
R_{merge}(%)	4.8(29.0)a	3.3(23.2) [a]	4.1(26.9)[a]	5.1(40.8)[a]
		RMSD from Ideal Values		
Bond Length (Å)	0.0332	0.0341	0.0257	0.0274
Bond Angle (°)	2.687	2.730	2.410	2.384
Number of water molecules	174	251	287	380

Note: Values in the parenthesis refer to the last resolution shell

Table V-2 (cont'd)

	Q108K:K40L:T51V:Y19W:R58Y	Q108K:K40L:T51V:Y19W:R58W	Q108K:K40L:T51V:T53S:Y19W:R58W:T20L	Q108K:K40L:T51V:T53C:Y19W:R58W:T20L:Q4H
Space group	P1	P1	P1	P1
a(Å)	29.748	29.964	29.870	30.918
b(Å)	36.030	35.914	35.691	35.978
c (Å)	64.782	64.568	64.915	64.483
α(°)	87.67	90.63	90.816	86.01
β(°)	88.06	91.28	88.726	86.56
δ(°)	64.72	113.89	66.106	65.17
Resolution (Å)	50-1.74(1.74-1.70)[a]	50-1.68(1.68-1.64)[a]	50-1.56(1.56-1.53)[a]	50-1.29(1.29-1.26)[a]
R_{work}/ R_{free} (%)	28.87/22.36	31.04/23.16	27.00/20.85	21.98/18.18
Total reflection	91114	52210	68699	216594
Unique Reflection	41402	50012	36782	96209
Completeness (%)	98.5(80.8)[a]	98.2(59.6)[a]	91.7(79.8)[a]	96.5(61.8)[a]
Molecules per Asymmetric Unit	2	2	2	2
I/σ	51.56(3.07)[a]	35.70(2.33)[a]	35.66(1.23)[a]	48.61(3.6)[a]
R_{merge}(%)	4.1(36.1)[a]	5.9(48.6)[a]	4.3(44.6)[a]	4.2(30.6)[a]
		RMSD from Ideal Values		
Bond Length (Å)	0.0118	0.0231	0.0218	0.0325
Bond Angle (°)	1.449	2.1733	2.080	2.556
Number of water molecules	270	207	282	380

Note: Values in the parenthesis refer to the last resolution shell

Table V-2 (cont'd)

	Q108K:K40L:Y19W:R58F:T53V	Q108K:K40L:T51V:T53C:R58W-Apo	Q108K:K40L:T51V:T53C:R58L	Q108K:K40L:T51V:R58Y:Y19W:T53L
Space group	P1	P3121	P1	P1
a(Å)	29.617	76.261	29.785	30.022
b(Å)	36.339	76.261	36.038	35.844
c (Å)	88.16	66.884	63.875	64.513
α(°)	90.67	90.00	89.21	90.67
β(°)	88.37	90.0	92.01	91.42
δ(°)	65.17	120.00	66.57	113.87
Resolution (Å)	50-1.77(1.77-1.72)[a]	50-1.73(1.73-1.70)[a]	50-1.84(1.84-1.76)[a]	50-1.58(1.58-1.54)[a]
R_{work}/ R_{free} (%)	32.83/24.63	24.17/21.45	29.13/21.91	27.71/20.93
Total reflection	87733	305066	88609	117570
Unique Reflection	36653	25252	60134	48728
Completeness (%)	97.7(66.0)[a]	100(100)[a]	98.9(76.0)[a]	96.8(66.8)[a]
Molecules per Asymmetric Unit	2	1	2	2
I/σ	47.58(3.8)[a]	81.08(4.3)[a]	48.49(3.21)[a]	52.26(3.15)[a]
R_{merge}(%)	4.5(22.7)[a]	3.5(50.6)[a]	6.6(25.2)[a]	4.5(29.4)[a]
RMSD from Ideal Values				
Bond Length (Å)	0.0217	0.0324	0.0315	0.0226
Bond Angle (°)	2.136	2.978	2.683	2.118
Number of water molecules	123	190	200	240

Note: Values in the parenthesis refer to the last resolution shell

Table V-2 (cont'd)

	Q108K:K40L	Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4R	Q108K:K40L:T51V:T553C:R58W:T29L:Y19W
Space group	P1	P1	P1
a(Å)	29.81	30.92	30.11
b(Å)	35.94	35.76	35.91
c (Å)	63.98	64.376	64.726
α(°)	90.83	86.40	90.93
β(°)	92.88	86.44	91.21
δ(°)	113.19	64.94	114.13
Molecules per Asymmetric Unit	2	2	2
Total reflection	73367	145141	95388
Unique Reflection	27686	35955	30286
Completeness (%)	92.2(72.5)[a]	98.7(89.3)[a]	91.0(79.8)[a]
Average I/σ	40.38(1.27)[a]	53.92(4.62)[a]	47.27(4.71)[a]
R_{merge}(%)	0.036(0.55)[a]	0.052(0.24)[a]	0.034(0.213)[a]
Resolution (Å) (Last Shell)	33.03-1.61(1.64-1.61)	32.36-1.5 (1.53-1.50)	28.91-1.58(1.61-1.58)
R_{work}/ R_{free} (%)	21.25/26.57	17.07/20.08	19.40/24.14
	RMSD from Ideal Values		
Bond Length (Å)	0.0205	0.0275	0.023
Bond Angle (°)	1.95	2.38	2.09
Average B factor	30.86	17.17	24.53
Number of water molecules	164	303	238
PDB ID	4EXZ	4EEJ	4EFG

Values in the parenthesis refer to the last resolution shell

Table V-2 (cont'd)

	Q108K:K40L:T51V:R58W:Y19W:L77T:T53L	Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:A33W
Space group	P1	P1
a(Å)	30.338	31.35
b(Å)	35.996	36.45
c (Å)	129.166	64.26
α(°)	89.92	90.04
β(°)	90.22	94.55
δ(°)	114.92	116.36
Resolution (Å)	50-1.46(1.46-1.40)	32.64-1.40(1.42-1.40)
R_{work}/ R_{free} (%)	23.09/18.70	20.04/24.24
Total reflection	622288	142087
Unique Reflection	246158	44292
Completeness (%)	97.7(93.1)[a]	93.5(80.4)[a]
Molecules per Asymmetric Unit	4	2
I/σ	50.11(1.5)[a]	47.93(2.70)[a]
R_{merge}(%)	4.7(10.6)[a]	0.082(0.358)[a]
RMSD from Ideal Values		
Bond Length (Å)	0.0304	0.028
Bond Angle (°)	2.558	2.38
Number of water molecules	530	279

Note: Values in the parenthesis refer to the last resolution shell

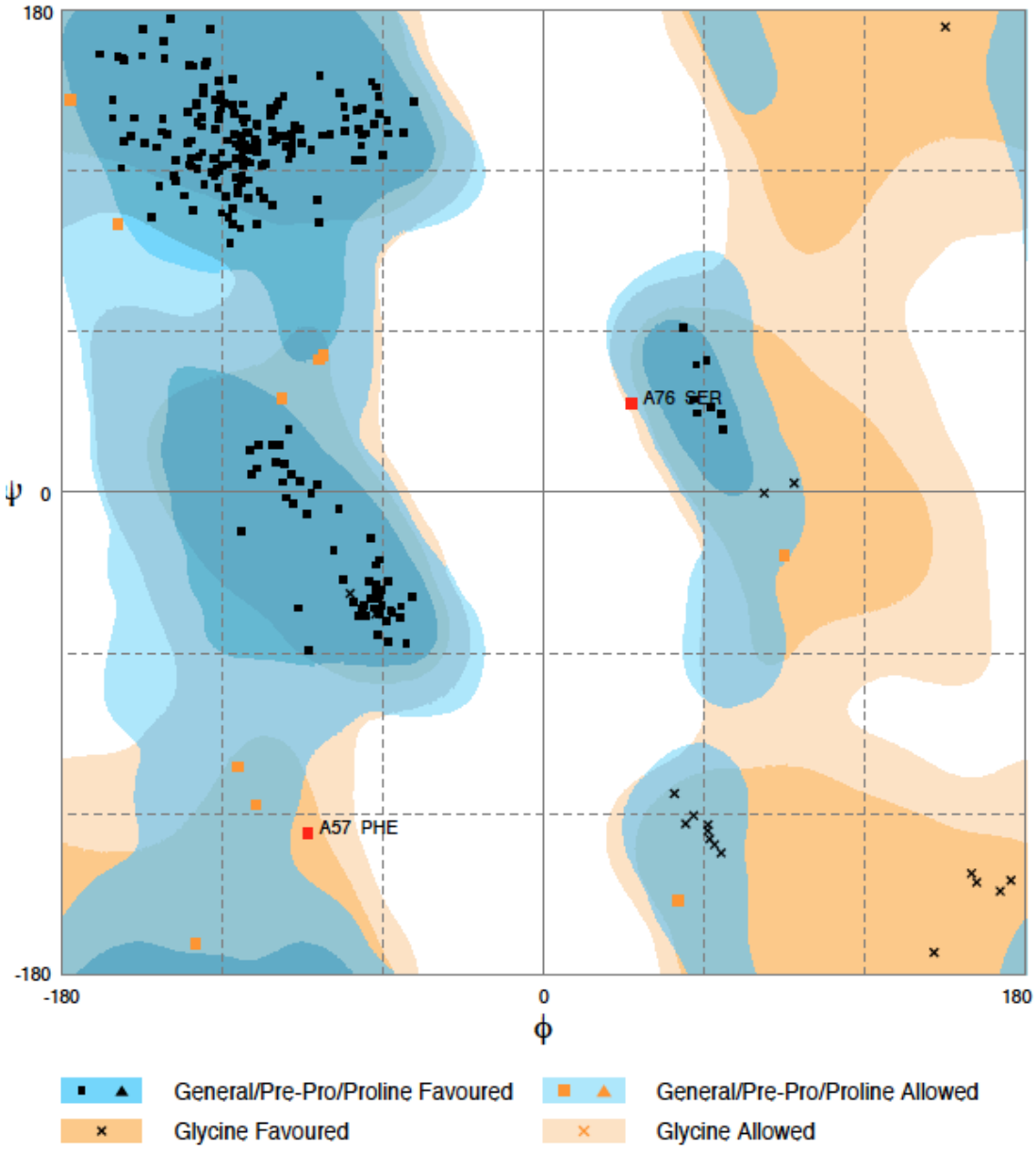


Figure V-1: The Ramachandran plot of Q108K:K40L mutant

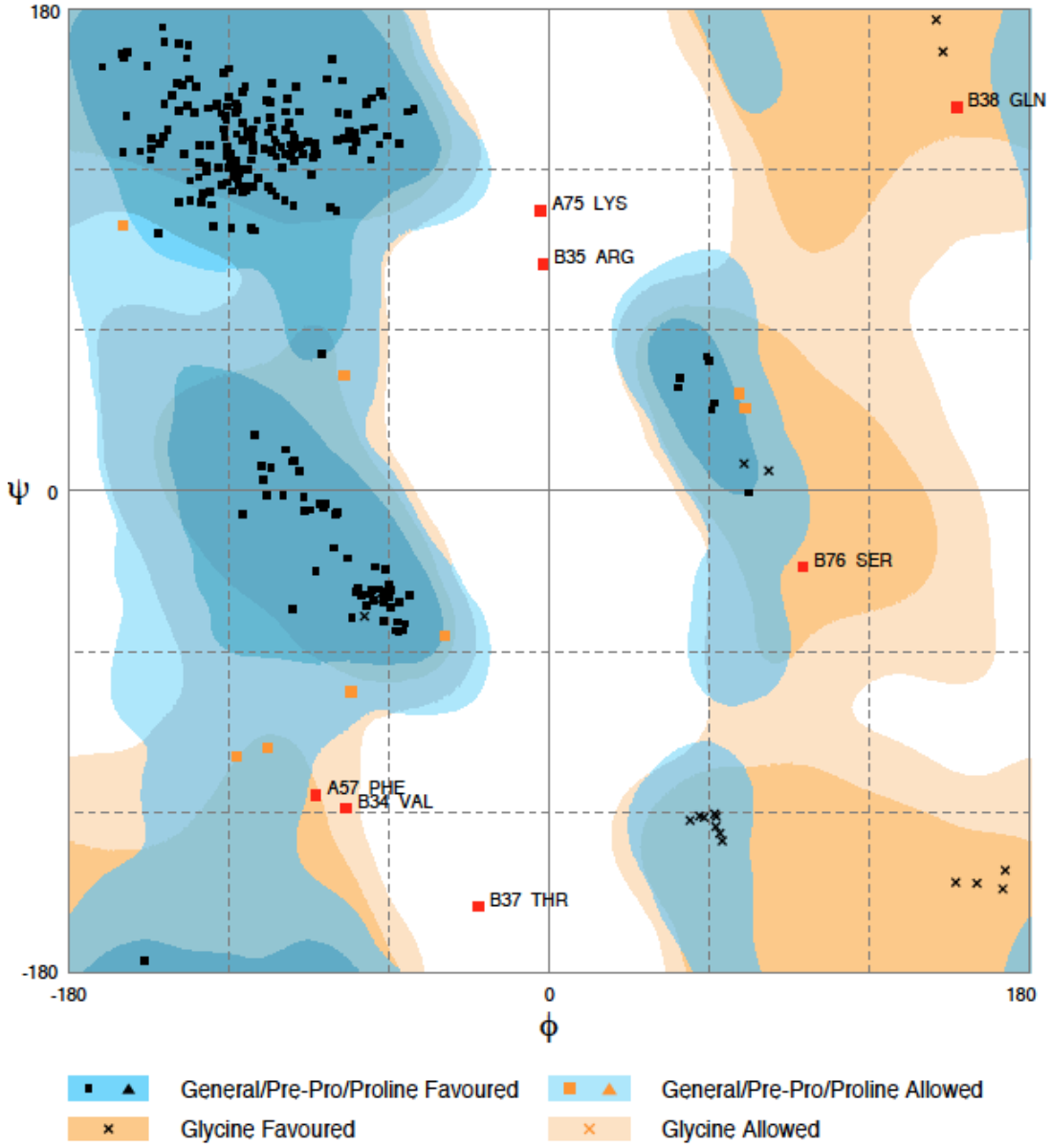


Figure V-2: The Ramachandran plot of Q108K:K40L:T53C mutant

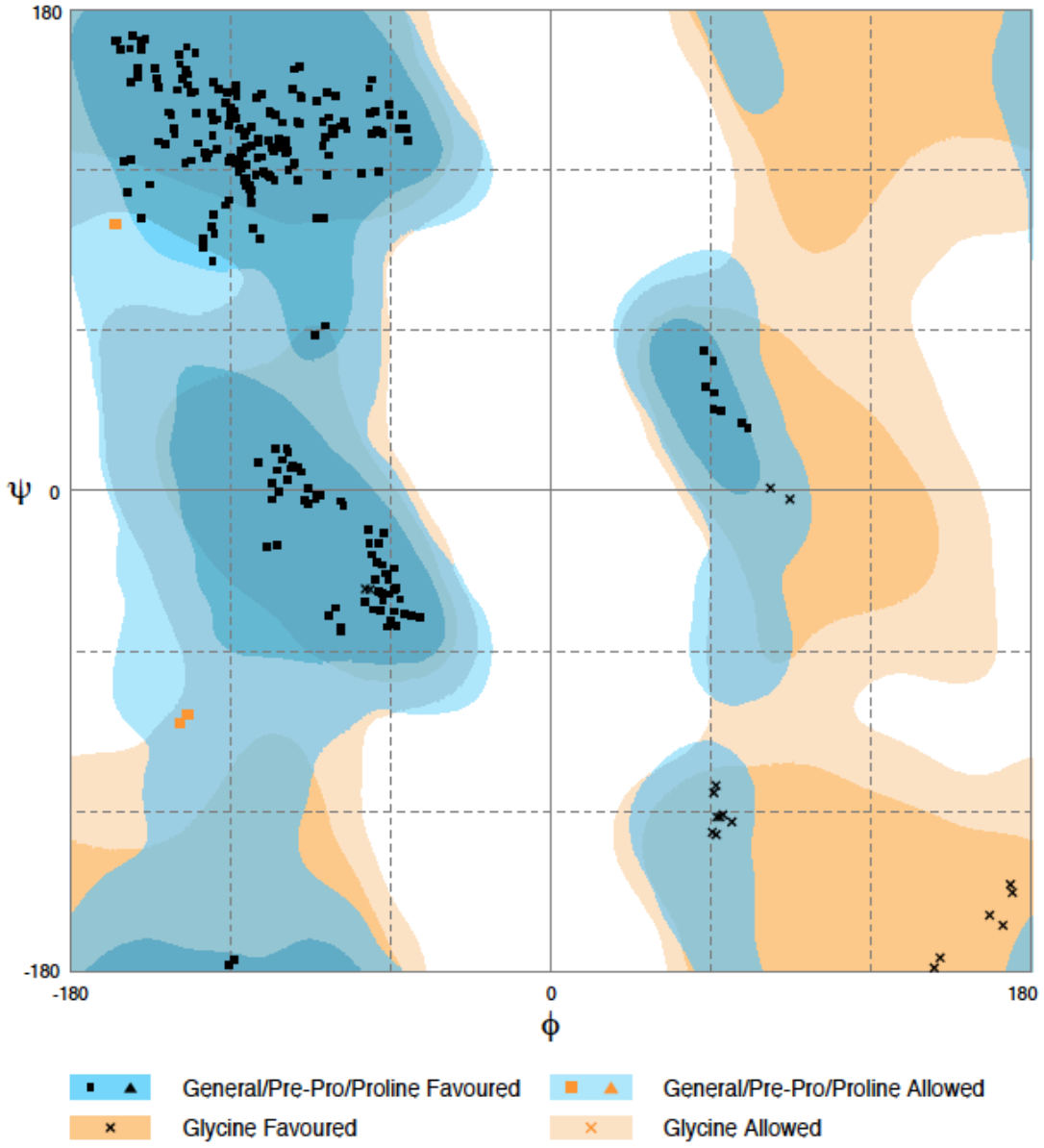


Figure V-3: The Ramachandran plot of Q108K:K40L:T51V:T53S mutant

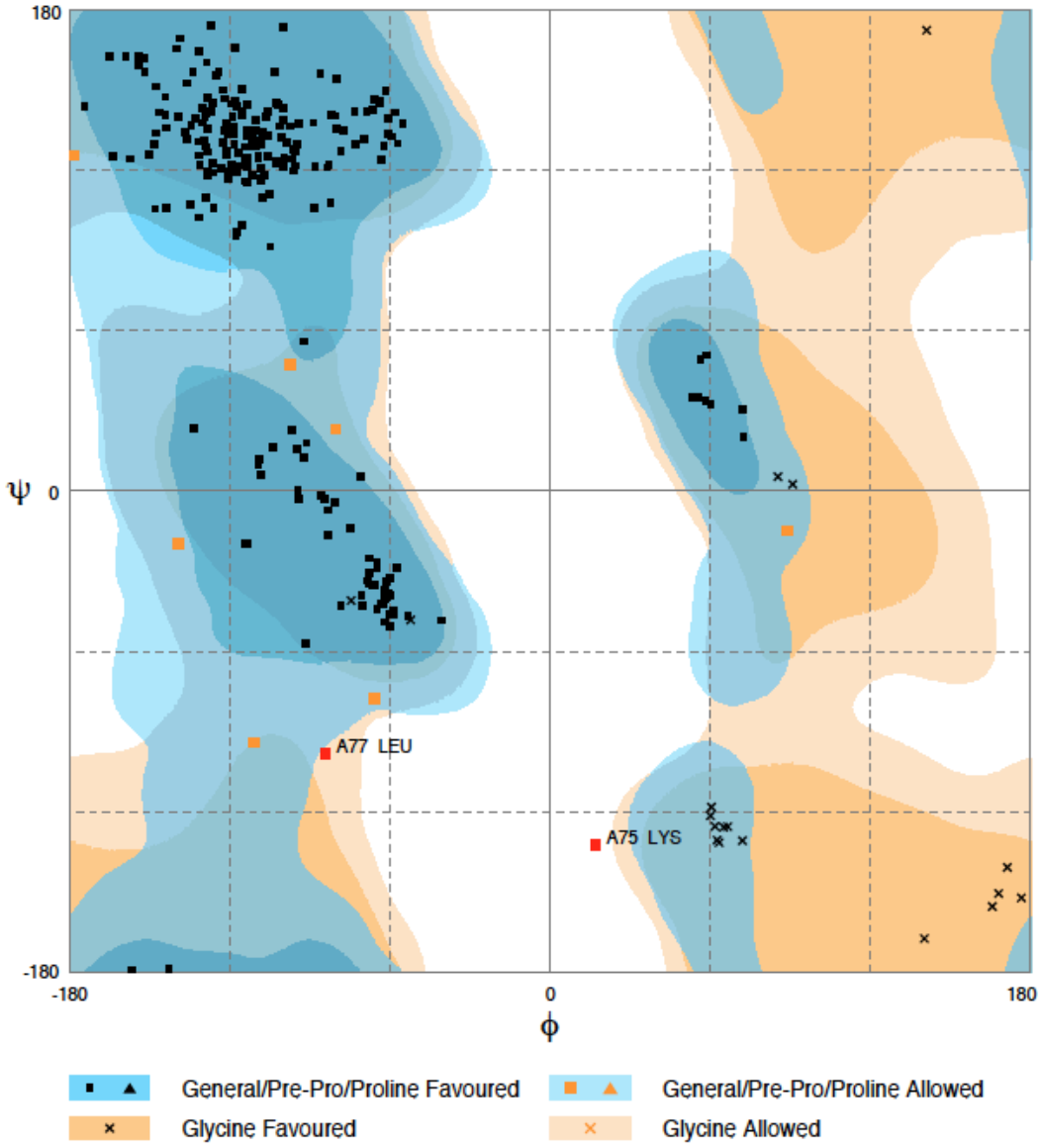


Figure V-4: The Ramachandran plot of Q108K:K40L:T51V:R58F mutant

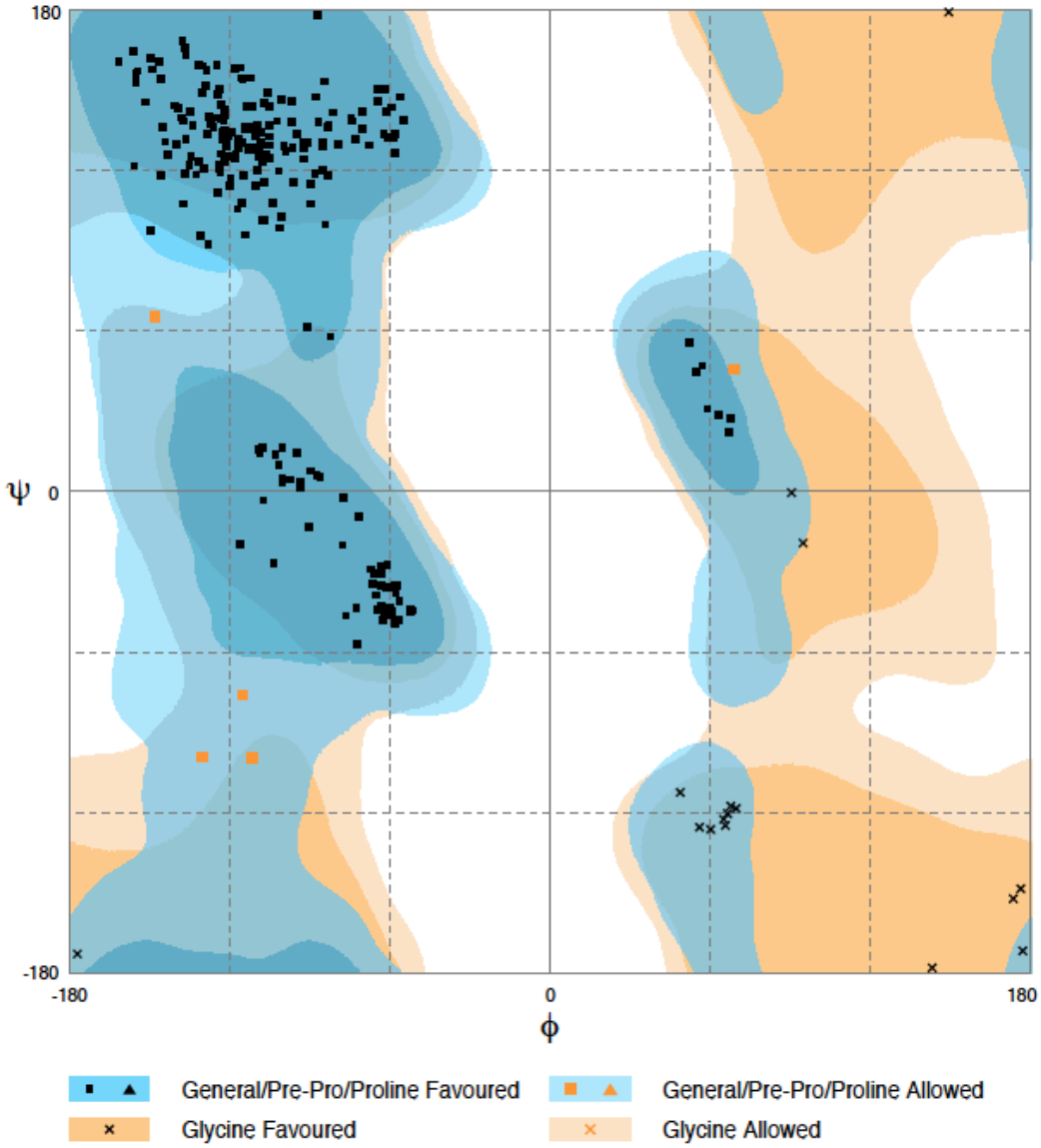


Figure V-5: The Ramachandran plot of Q108K:K40:T51V:T53S:R58W:T29L:Y19W mutant

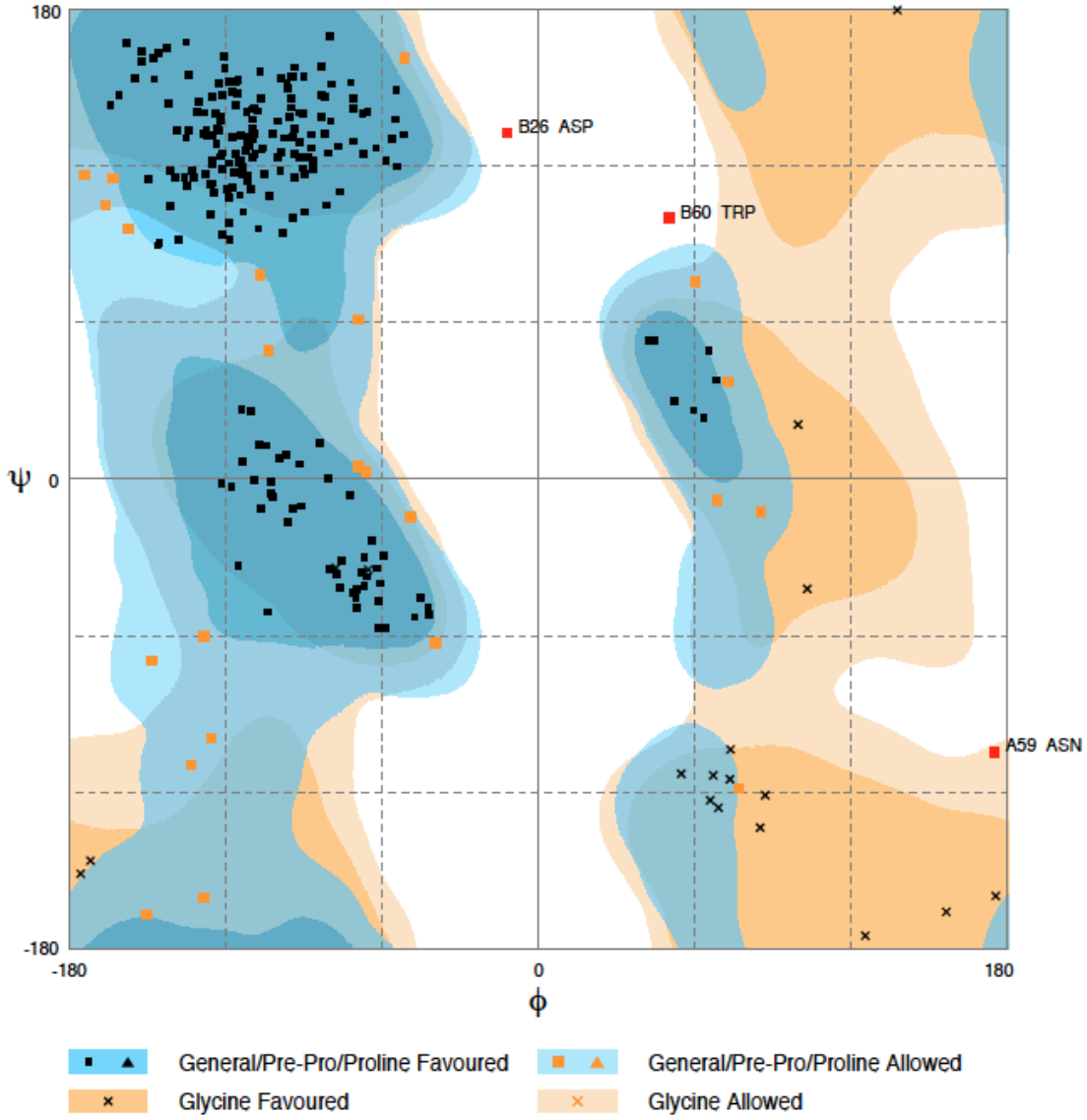


Figure V-6: The Ramachandran plot of Q108K:K40L:T51V:R58Y:Y19W mutant

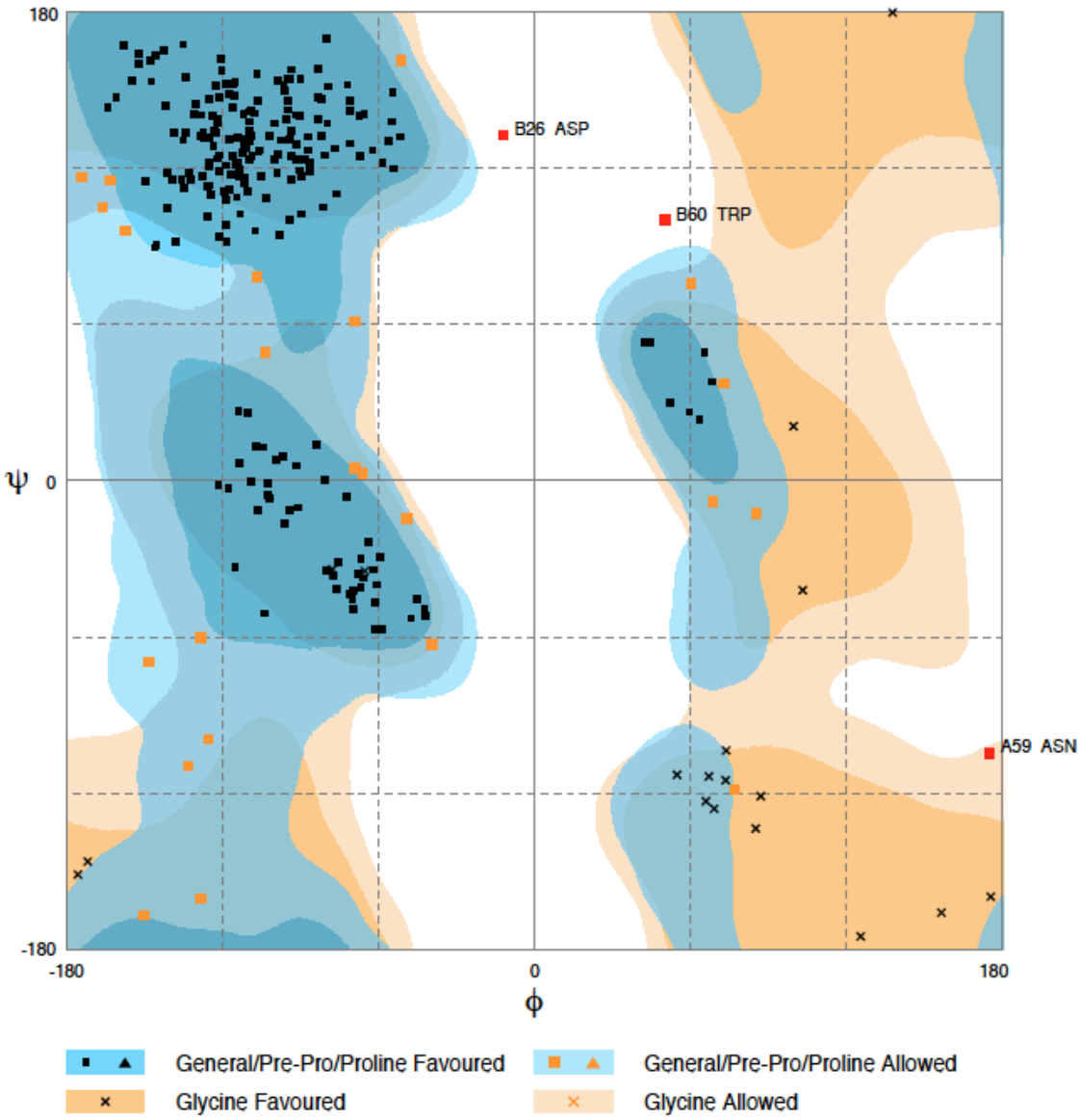


Figure V-7: The Ramachandran plot of Q108K:K40L:Y60W dimer

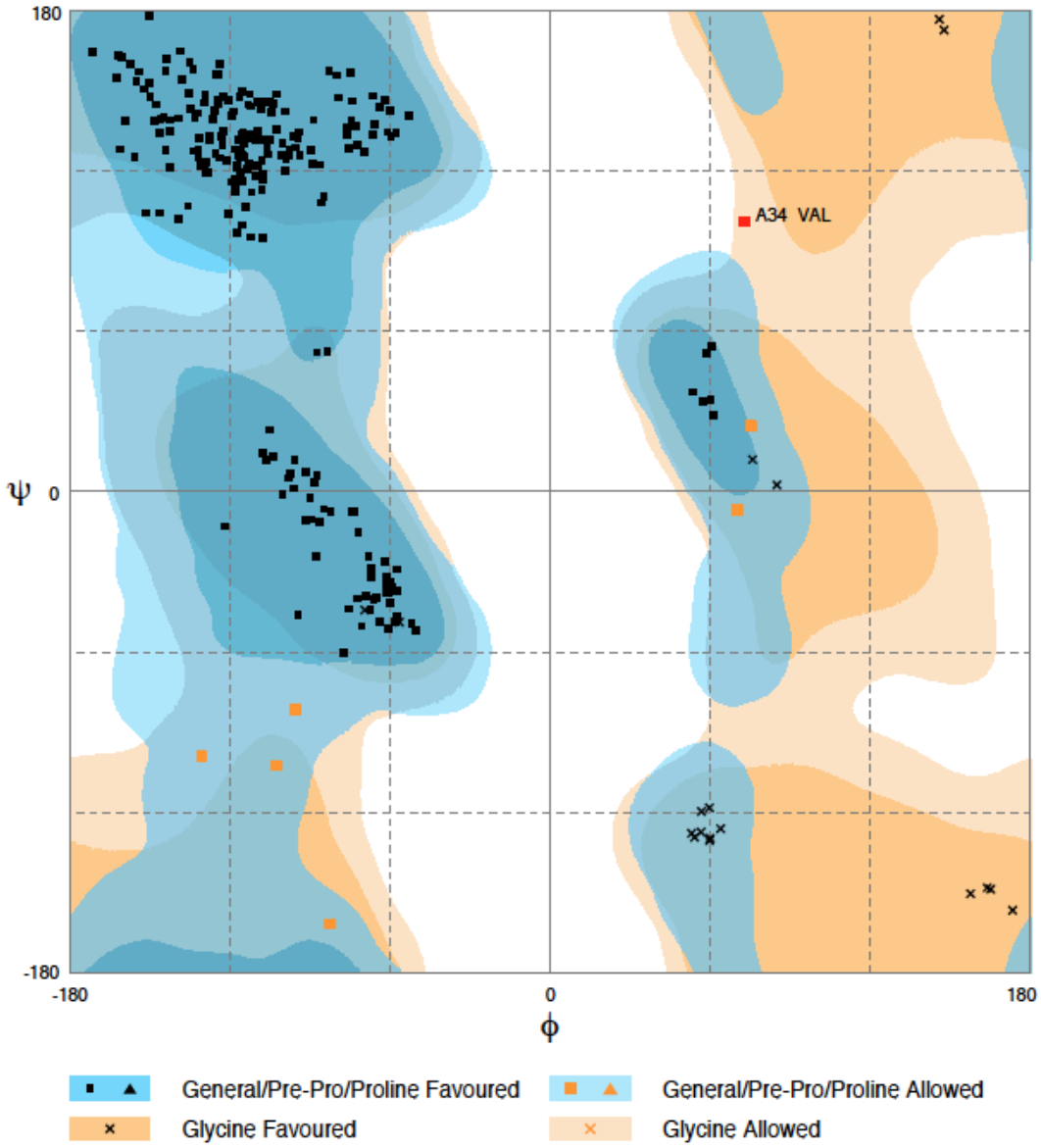
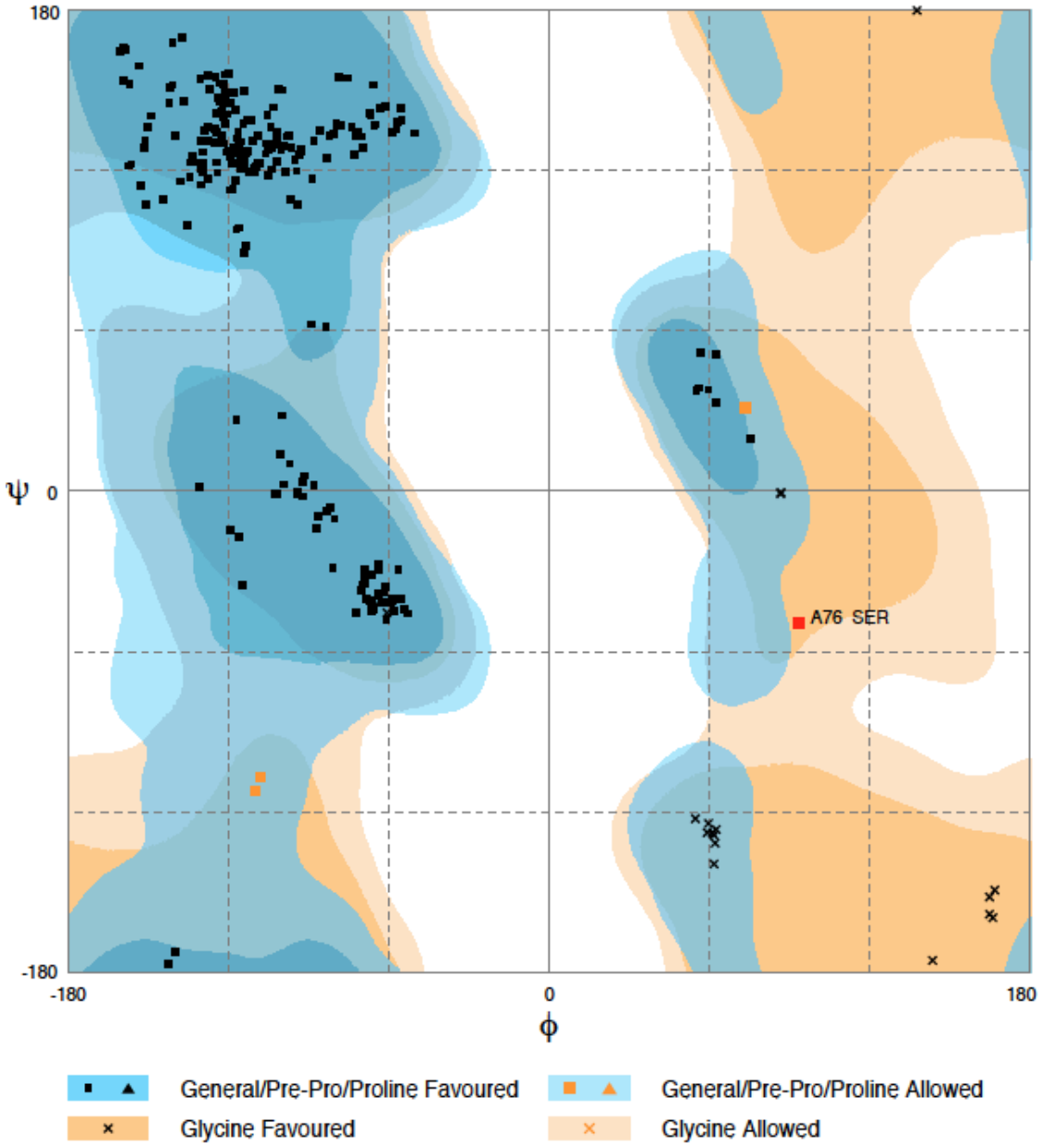


Figure V-8: The Ramachandran plot of Q108K:K40L:Y60W monomer



FigureV-9: The Ramachandran plot of Q108K:K40L:T51V:T53C:R58W:T29L:Y19W:Q4H mutant

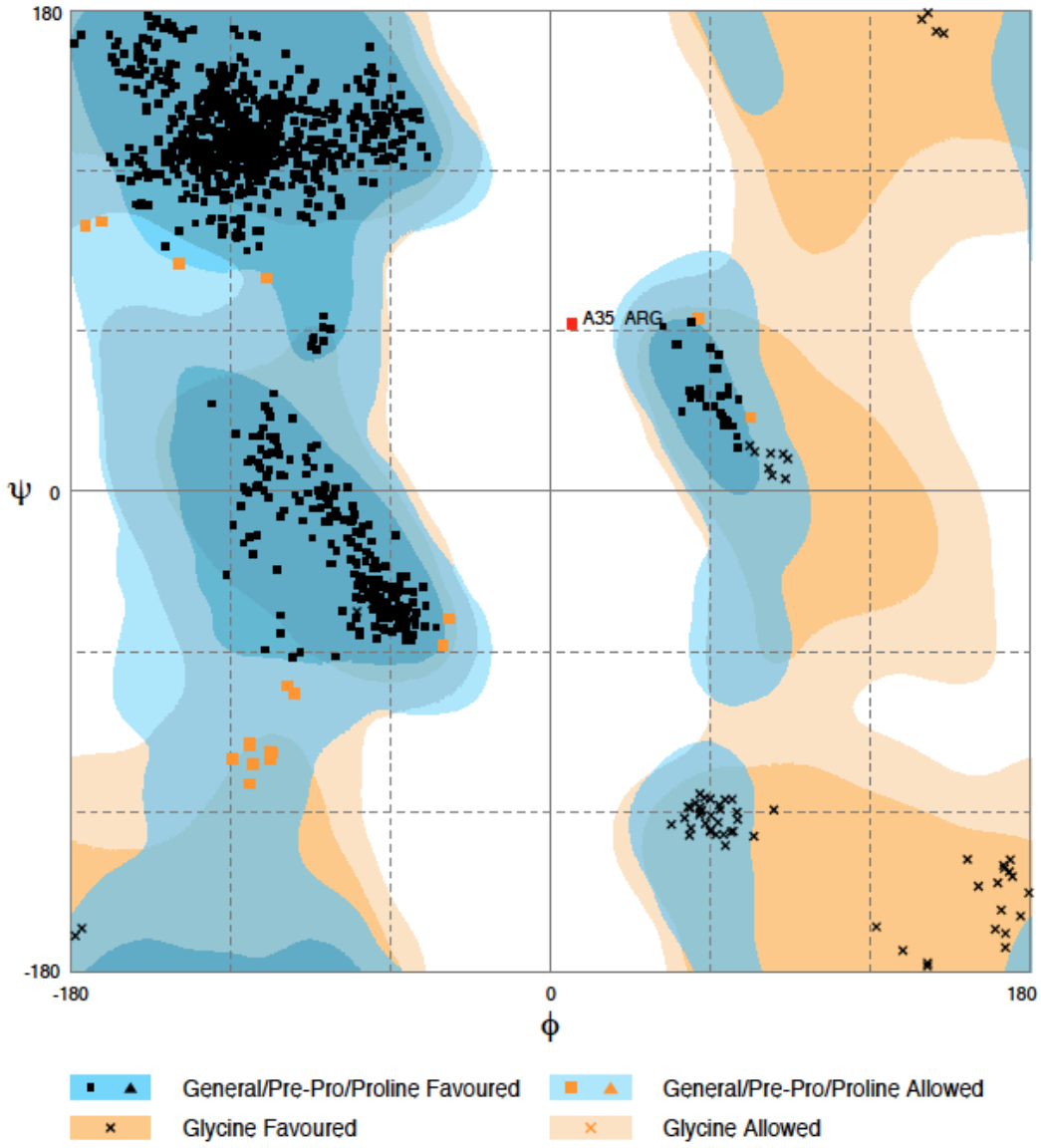


Figure V-10: The Ramachandran plot of wt hCRBPII bound to all-*trans*-retinal

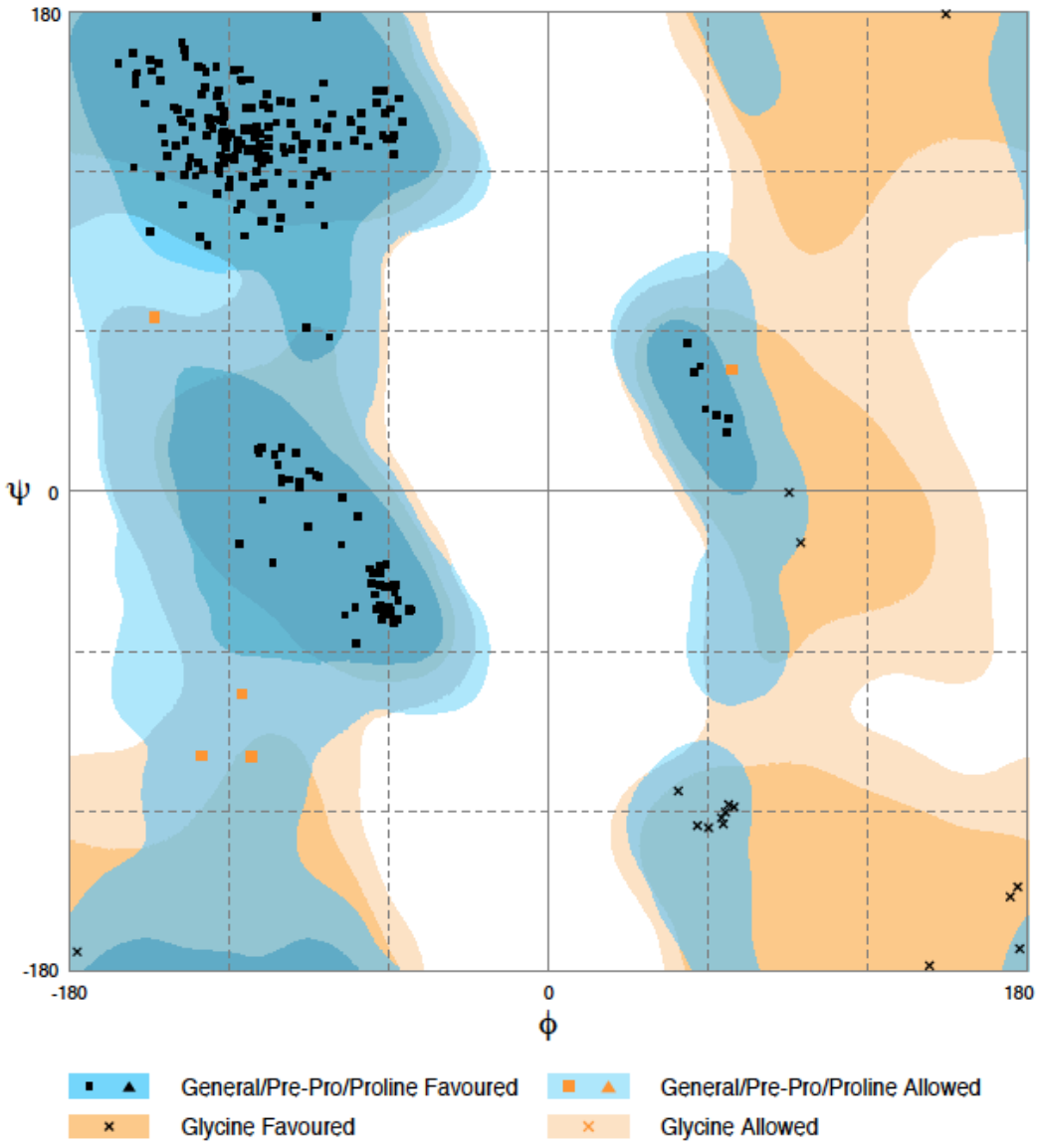


Figure V-11: The Ramachandran plot of Y60W dimer

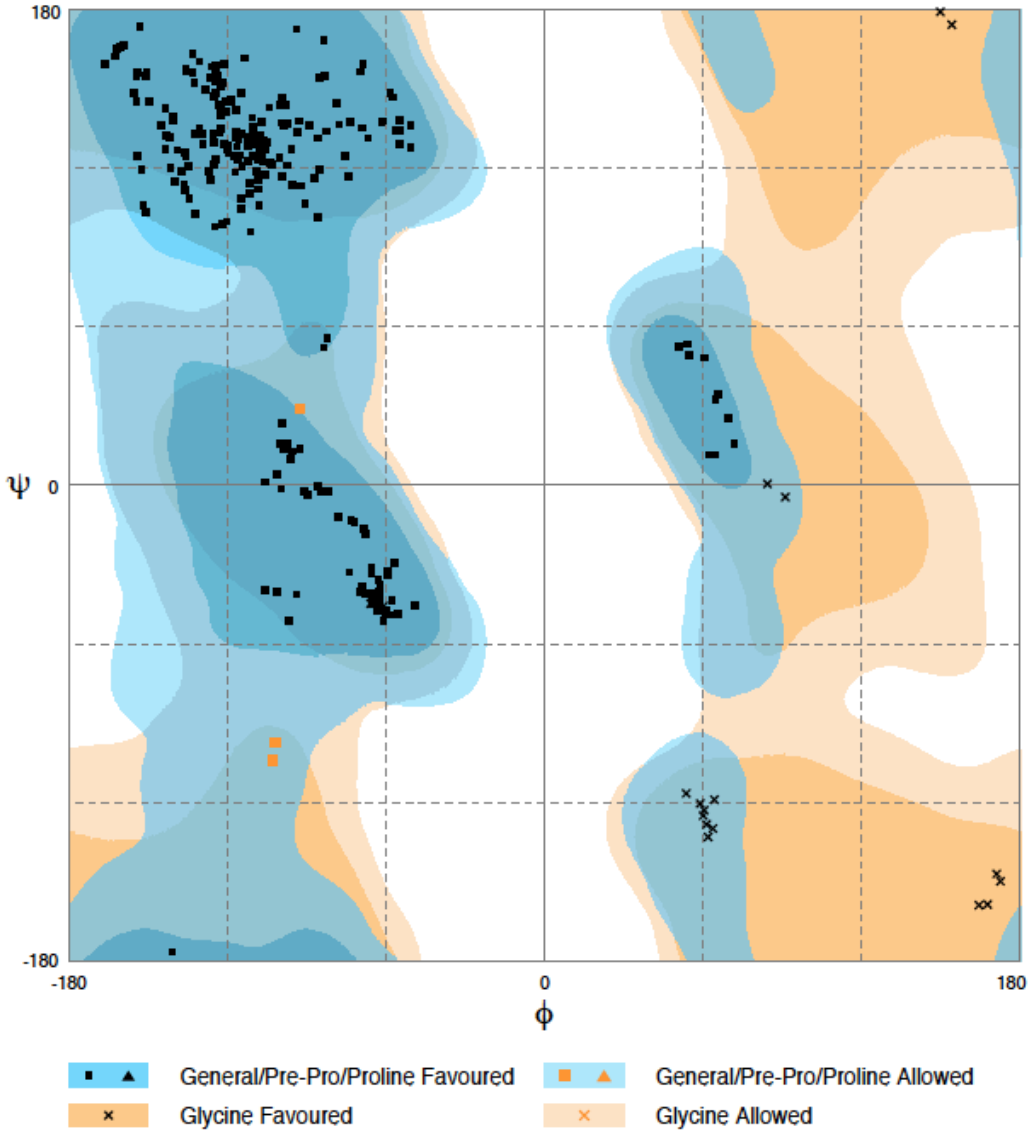


Figure V-12: The Ramachandral plot of wt hCRBPII bound to all-*trans*-retinol