DESIGN OF NOVEL PROTEIN/CHROMOPHORE COMPLEXES: SYNTHESIS AND EVALUATION OF CHROMOPHORES, FLUOROPHORES, AND ENGINEERING OF PROTEIN HOST SYSTEMS

By

Ipek Yapici

A DISSERTATION

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

Chemistry – Doctor of Philosophy

2015

ABSTRACT

DESIGN OF NOVEL PROTEIN/CHROMOPHORE COMPLEXES: SYNTHESIS AND EVALUATION OF CHROMOPHORES, FLUOROPHORES, AND ENGINEERING OF PROTEIN HOST SYSTEMS

By

Ipek Yapici

The field of protein engineering has undergone phenomenal growth from its inception approximately 20 years ago. A wide variety of topics have been addressed, including the construction of new protein folds, the introduction of metal binding sites that are both structural and catalytic, the development of novel enzymatic activity, and the creation and optimization of new ligand binding sites. However, left behind has been the issue of protein/chromophore interactions that regulate the spectroscopic properties of bound chromophores. To understand the fundamental elements that contribute to spectral tuning of a chromophore inside a protein cavity, we redesigned small cyctosolic human proteins to fully encapsulate retinylidene based fluorescent and/or nonfluorescent ligands, bound as a protonated Schiff base. Rational mutagenesis, designed with the goal to alter the electrostatic environment within the binding pocket of the host protein, enabled regulation of the absorption and emission spectra of the protein/chromophores complexes.

An initial aim of this project was to reengineer Cellular Acid Binding Protein II (CRABPII) into a fluorescent protein via coupling with a nonfluorescent cyanine dye precursor. The designed precursor, merocyanine aldehyde, is

capable of binding to variety of CRABPII mutants, resulting in stable fluorescent complexes. In the course of this study, it became evident that CRABPII/merocyanine complexes demonstrate structural variety in terms of ligand orientation and geometry that correlates well with the observed fluorescent properties.

From the point of view of structural chemical biology, the next challenge in the field would be to obtain detailed insights on wavelength changes with respect to chromophore modifications. We have focused on coupling of the reengineered hCRBPII rhodopsin mimics with different retinal analogs in order to systematically study the role of structural variations of these ligands on spectral tuning of the resultant pigments. Diverse sets of hCRBPII-retinoid complexes were evaluated with respect to their photophysical properties in light of available high resolution crystal structures. Given the results reported here, we highlight the significance of ring methyl substituents in wavelength regulation, not only as a handle for conformational issues, but also surprisingly as modulators for electrostatic interactions.

In this report, we also aimed at designing a protein system that couples with an azobenzene-based chromophore to respond to light stimulation. A synthetic azo chomophore, Azo-AA, was tested with a variety of hCRBPII mutants for affinity optimization. These protein/Azo-AA hybrids could be further engineered for studying protein-protein interactions and generating a bio-machinery for affinity-based protein separation methodology.

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my PhD advisor, Prof. Babak Borhan whose enthusiasm, encouragement and valuable ideas help me to complete this work. Babak has the attitude and the substance of a genius, knowing how to keep his students invested in the organization and give them the opportunity to learn. I am fully indebted to him for pushing me farther than I thought that I could go, and for not letting me give up under stressful conditions. During my PhD period, I did not hesitate to share my personal concerns and to ask help from him regarding to my personal issues. He was the one helping me to move out from my apartment complexes, assisting me when I needed to buy a car, providing me medicine and food while I was sick and there was no one who could take care of me, and being next to me at anytime that I needed. I can not say enough about how helpful he was, as well as personable and friendly. Besides being an outstanding advisor, he is a great friend, making me feel appreciated and special. I am also glad to know his mother, Badri Borhan, who is an incredibly wise woman with a warm heart. Her amazing cooking talents made me like Persian cuisine, in particular the dish called “tah chin”.

In addition, I would like to thank Dr. Chrysoula Vasileiou who is an extremely helpful and generous person. She welcomes everyone to her office at any time as long as we are at the department and provide us variety of snacks. She is the main character of our group who gathers people together and gives us

the spirit of being a family. I will miss the times with her while we were having conversations, eating together, and party nights at her house.

I would like to specially thank Dr. Wenjing Wang and Dr. Tanya Berbasova for being excellent lab mates over the years. It is sad that they graduated before me, but I am glad to have the opportunity to know them in person and build up long lasting friendships with them. They were always an inspiration for me to focus on my research and be motivated on my projects. I am very appreciative that they did not hesitate to share their research experience, skills and knowledge with me, and most importantly their time to assist me in my first two years. We were absolutely a solid bio-team, creating the best working conditions for ourselves and being a good example for other research groups.

I also want to thank Meisam Nosrati for all of his help in crystallizing and solving numerous protein structures, which are important for my research. His kindness, understanding and deep knowledge not only makes him an ideal collaborator, but also an outstanding researcher who will succeed in many great things in the future.

I am heartily thankful to the past and present members of the Borhan's lab, and also all of my other friends at chemistry department for their friendship and their time we spent together. I especially should thank Bardia Soltanzadeh for his great personality. He is a brother to me who exactly knew how to cheer me up and calm me down when I got upset. I will miss our tea sessions, hilarious discussions, and nonsense fights.

To my committee, Prof. Maleczka, Prof. Geiger and Prof. Arnosti, I am extremely grateful for their aspiring guidance, valuable constructive criticisms during my presentations and for proof-reading my dissertation. Their assistance and motivation conceived a direction in me and made the report a success. I would like to thank all faculty and staff members of Department of Chemistry that have been kind enough to spare their precious time for us to provide an awesome working environment.

Last but not least, I am thankful to my lovely parents, and all of my friends from back home. But most of all I want to take the opportunity to acknowledge the eight beings without whom this dissertation would have taken a different pathway: Seda, Gözde, Deniz, Şeniz, Döndü, Engin, Hakan, and Utku. I offer thanks to each of them for keeping in touch with me almost everyday and making me feel that I was never far from them. Their spiritual presence at different moments during my research was invaluable for me. Even though most of us are apart from each other in different countries and cities to continue our education and working carriers, our sincere and strong friendship has been holding us together across the miles and will continue to be there forever.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

KEY TO SYMBOLS AND ABBREVIATIONS

Symbols

Å	Angstrom
ε	Extinction Coefficient
Φ	Quantum Yield
ψ	Dihedral Angle
τ	Thermal Relaxation
δ	Chemical Shift
η	Refractive Index
$^{\circ}C$	Degree Celsius
λ^{max}	Maximal Wavelength
λ^{ex}	Excitation Wavelength
λ^{ems}	Emission Wavelength
h	Hour
min	Minute
s	second
cm	Centimeter
M	Molar
mM	Millimolar
μM	Micromolar

nM	Nanomolar
mg	Milligram
µg	Microgram
ng	Nanogram
mol	Mole
mmol	Millimole
pmol	Picomole
kDa	KiloDalton
rpm	Rotations Per Minute
ppm	Parts Per Million
Hz	Hertz
MHz	Megahertz
>	Larger than
<	Smaller than

Abbreviations of Amino Acids

Ala, A	Alanine
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartate
Cys, C	Cysteine
Gln, Q	Glutamine
Glu, E	Glutamic acid

His, H	Histidine
Ile, I	Isoleucine
Leu, L	Leucine
Lys, K	Lysine
Met, M	Methionine
Phe, F	Phenylalanine
Pro, P	Proline
Ser, S	Serine
Thr, T	Threonine
Trp, W	Tryptophan
Tyr, Y	Tyrosine
Val, V	Valine

Abbreviations of Nucleotide Bases

A	Adenine
C	Cytosine
T	Thymine
G	Guanine

All Other Abbreviations

CRABPII	Cellular Retinoic Acid Binding Protein II
hCRBPII	Human Cellular Retinol Binding Protein II
FAP	Fluorogen-Activated Peptide
FP	Fluorescent Protein

EGFP	Enhanced Green Fluorescent Protein
mRFP	Monomeric Red Fluorescent Protein
GPCR	G-Protein Coupled Receptor
7-TM	Seven-Transmembrane
BSA	Bovine Serum Albumin
iLBP	Intracellular Lipid Binding Protein
FABP	Fatty Acid Binding Protein
NIR-FPs	Near-Infrared Fluorescent Proteins
eDHFR	*Escherichia coli* Dihydrofolate Reductase
E. coli	*Escherichia coli*
PYP	Photoactive Yellow Protein
POI	Protein Of Interest
PDB ID	Protein Date Bank Identification
K_d	Dissociation Constant
K_i	Inhibition Constant
pK_a	Logarithmic Constant Of Acid Dissociation
k_2	Second-Order Rate Constant
k^{obs}	Observed Rate Constant
$t_{1/2}$	Half-life Time
n.d.	Not Determined

p.b.	Poor Binding
bp	Base Pairs
RT	Room Temperature
Rf	Retention Factor
SB	Schiff Base
PSB	Protonated Schiff Base
NMR	Nuclear Magnetic Resonance
PCR	Polymerase Chain Reaction
QY	Quantum Yield
TLC	Thin Layer Chromatography
APS	Advanced Photon Source
UV	Ultraviolet Light
UV-vis	Ultraviolet and Visible Light
WT	Wild Type
CVM	Cytomegalovirus
DNA	Deoxyribonucleic Acid
PBS	Phosphate Buffered Saline
TMP	Trimethoprim
*p*Ca	*p*-Coumaric Acid
TO	Thiazole Orange
MG	Malachite Green
RET	All-*trans*-Retinal
DEHYDRO	3,4-Dehydro-Retinal

5DEM	5-Demethyl-Retinal
GEMLESS	1,1-Didemethyl-Retinal
CYCHXE	Cyclohexene Retinal
CYCHXA	Cyclohexane Retinal
Azo	Azobenzene
Azo-AA	Azobenzene Allyl Alcohol
IPTG	Isopropyl β-D-1-thiogalactopyranoside
SDS	Sodium Dodecyl Sulfate
PEG	Poly(ethylene glycol)
BTP	Bis-Tris Propane
Tac	Tacsimate
TPA	*Tert*-butyldimethylsilyl Protected Alcohol
EDC.HCl	1-Ethyl-3-(3-Dimethylaminopropyl)Carbodiimide Hydrochloride
HCl	Hydrochloric Acid
NaOH	Sodium Hydroxide
NaH	Sodium Hydride
KH	Potassium Hydride
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DCM	Dichloromethane
EtOH	Ethanol
THF	Tetrahydrofuran

ACN	Acetonitrile
Et_3N	Triethylamine
$CDCl_3$	Deuterated Chloroform
EtOAc	Ethyl Acetate
DIBAL-H	Diisobutyl Aluminum Hydride
TMEDA	Tetramethylethylenediamine
DMAP	4-Dimethylaminopyridine
TBDMS	*Tert*-butyldimethylsilyl
PTSA	*Para*- Toluene Sulfonic Acid
$NaHCO_3$	Sodium Bicarbonate
K_2CO_3	Potassium Carbonate
KH_2PO_4	Monopotassium phosphate
Na_2SO_4	Sodium Sulfate
CCH_3I	Iodomethane
PBr_3	Phosphorus Tribromide
$(MeO)_2SO_2$	Dimethyl sulfate
HWE	Horner-Wadsworth-Emmons

CHAPTER I: ENGINEERING OF CRABPII INTO A FLUORESCENT PROTEIN VIA COUPLING WITH A PROFLUOROGENIC CYANINE ALDEHYDE

I.1 Current Approaches On Fluorescent Labeling Of Proteins

The development of fluorescent probes allows the investigation of various cellular processes in terms of visualization and quantification. The discovery of an original green fluorescent protein (GFP) back in the 1960s opened a new era in molecular biology in which there is a continuing effort to develop new fluorescent labeling tags with altered colors and improved properties.[1,2] Basically, there are three main classes of fluorescent probes that can be applied for tagging biomolecules; i) GFP and its variants are bioluminescent proteins that autocatalyze chromophore formation within their structure by using molecular oxygen, ii) near-infrared fluorescent proteins (NIR-FPs) derived from bacterial phytochromes, and UnaG, isolated from Japanese freshwater eel, are chemical-genetic fluorescent switches that can turn on upon complexation with endogenous molecules, iii) site-specific chemical labeling systems are synthetic fluorescent probes that require exogenous chromophores to couple with their acceptor peptides.

The use of GFP and its many variants has the advantage of high labeling specificity since they generate their fluorescent moiety autocatalytically with the aid of molecular oxygen. The fluorescent moiety is formed via the chemical cyclization of three amino acid residues followed by an oxidation within the β-barrel (**Figure I-1**).[3-7] Noteworthy, the tolerance on substitutions both within and outside of the tripeptide chromophore-forming moiety has led to create randomized libraries of the GFP gene

and resulted in a powerful palette of fluorescent proteins (FPs) with specialized spectral properties. Changes of the tyrosine unit in the chromophore-forming tripeptide

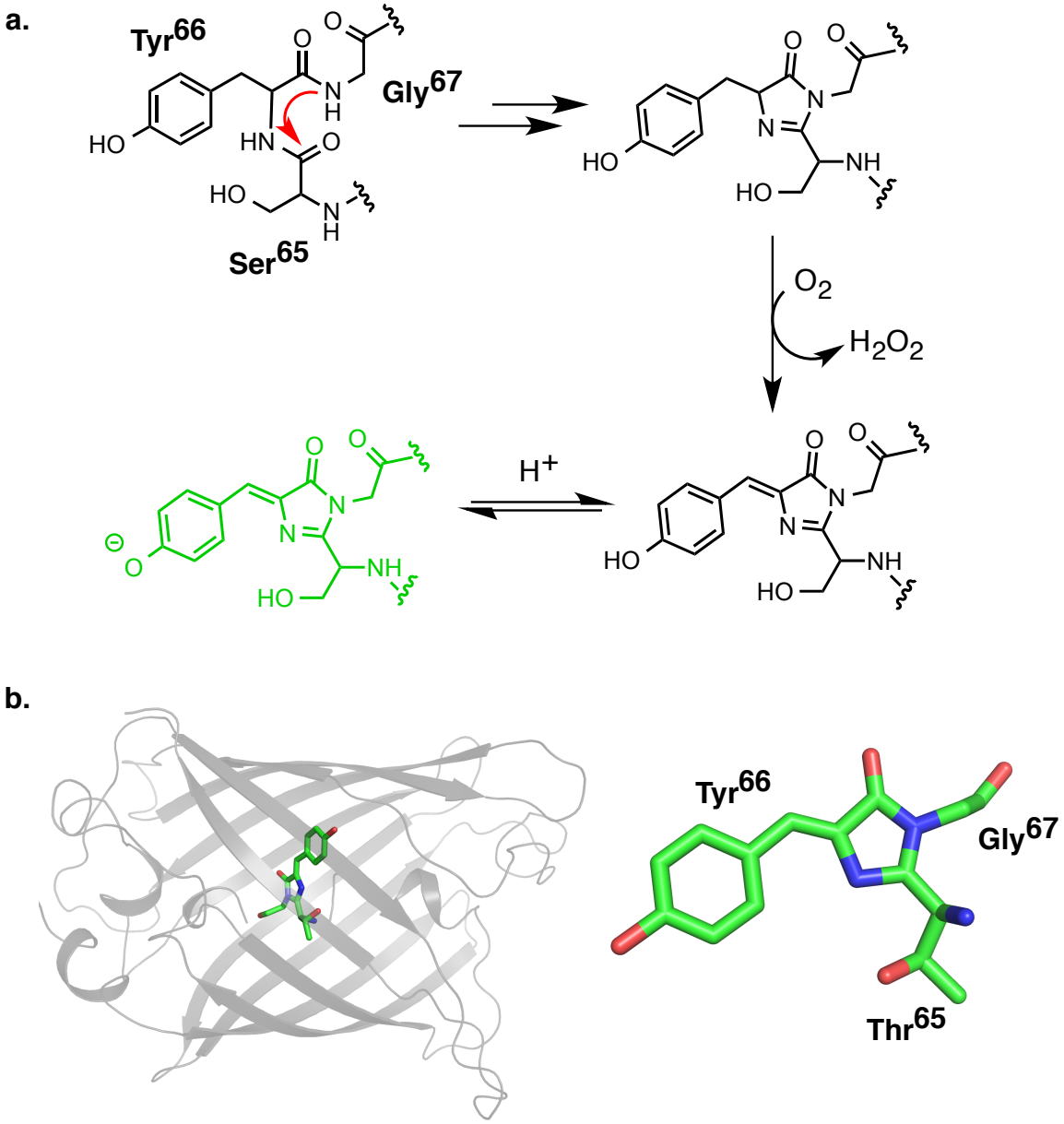


Figure I-1: Wild type GFP chromophore formation resulting from the spontaneous cyclization and oxidation of three consecutive residues (Ser^{65}-Tyr^{66}-Gly^{67}). Emissive state of the chromophore is shown in green. **b.** The crystal structure of enhanced version of GFP (EGFP) with the mutation of Ser^{65} to Thr^{65}. The expanded view of a mature chromophore is shown on right (PDB ID: 1EMA).

architecture to histidine, phenylalanine and tryptophan residues resulted in blue-shifted FPs, whereas the extension of the conjugated system through elongation of the chromophore structure provided more red-shifted pigments (**Figure I-2**).[8-15]

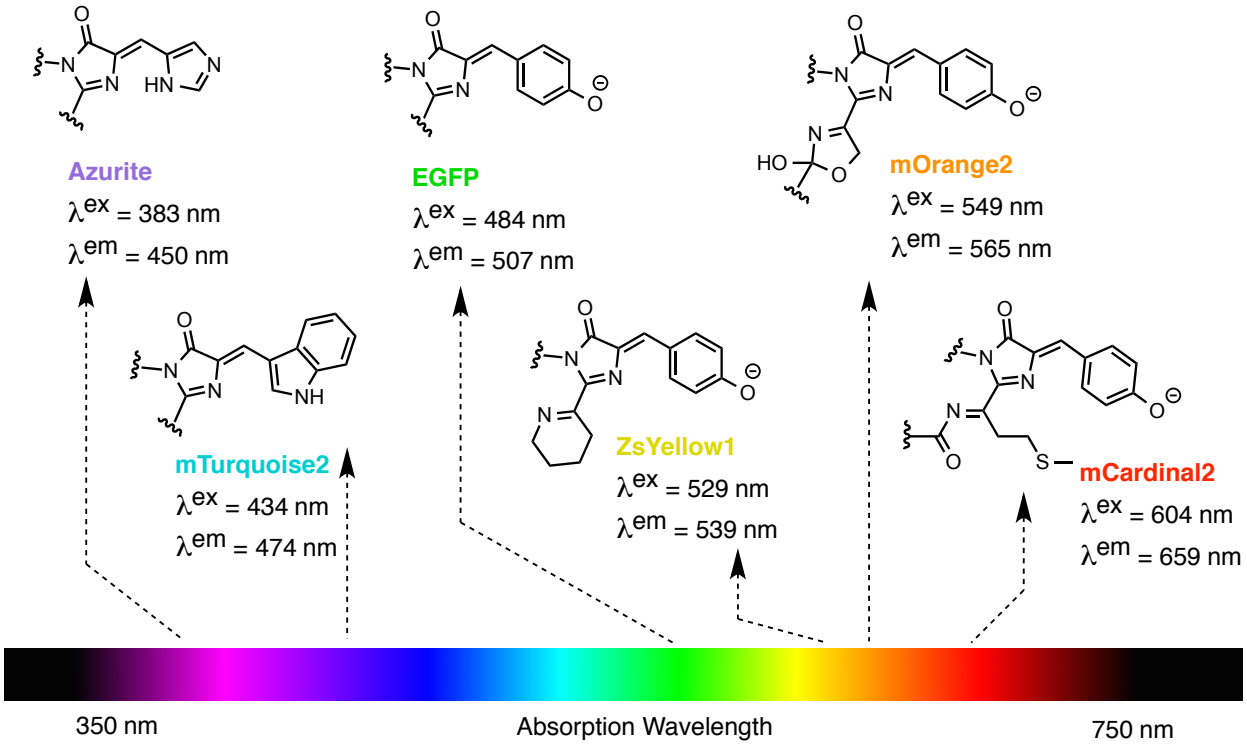


Figure I-2: Diversity of chromophore compositions found in fluorescent proteins. The spectral properties of chromophores are highlighted.

Unlike GFP-like proteins, UnaG and IRFPs are oxygen-independent fluorescent sensors. These noncanonical fluorescent proteins rely upon the complexation of endogenous cofactors that can be also added externally in order to suit the desired application. UnaG isolated from Japanese freshwater eel belongs to fatty-acid-binding protein (FABP) family which binds non-covalently and with its high specificity to bilirubin (**Figure I-3a**).[16] Although bilirubin is a non-fluorescent molecule, its sequestration within the deep UnaG binding pocket yields a bright green fluorescence at 527 nm with a

quantum efficiency of 0.51. This unique ligand-protein complex can be used as a fluorogenic marker in living cells in which bilirubin is produced endogenously. Alternatively, it can be applied as an inducible chemical-genetic fluorescent system in various cell types where bilirubin is absent by supplying exogenous bilirubin.

Similar to UnaG, NIR-FPs such as IFP2.0, iRFP and Wi-Phy are proteins developed from bacteriophytochromes, and function by incorporating biliverdin as their endogenous chromophores for yielding fluorescent complexes (**Figure I-3b**).[17-25] The covalent attachment of biliverdin via a thioether linkage to a cysteine residue in the rigid environment of the protein binding pocket yields fluorescent complexes which display excitation and emission maxima in the near-infrared region (~ 700 nm). These engineered phytochrome based proteins become a robust tool for deep in *vivo* and whole-body imaging. Nevertheless, much more work will be necessary in order to increase their quantum efficiency (~ 0.07), enhance brightness without the need for external biliverdin addition, increase photostability and expand their spectral shifts. Having such improved NIR-FPs could create new opportunities for labeling more than one tissue or organ at a time.

Beside the GFP-like fluorescent proteins and endogenously inducible noncanonical fluorescent proteins, there are many alternative approaches that allow selective labeling of proteins based on smaller chemical compounds referred to as site-specific chemical labeling systems.[26-34] One of the distinct advantages of site-specific chemical labeling is the freedom and the flexibility to use a wide repertoire of small molecules with diverse spectral properties. In a general chemical labeling method, an

acceptor peptide moiety is fused with a protein of interest and a small fluorescent or profluorogenic compound is introduced exogenously to find its target acceptor. There are some different approaches beside the general chemical labeling method for selective labeling like incorporation of unnatural amino acid acceptors in order to couple

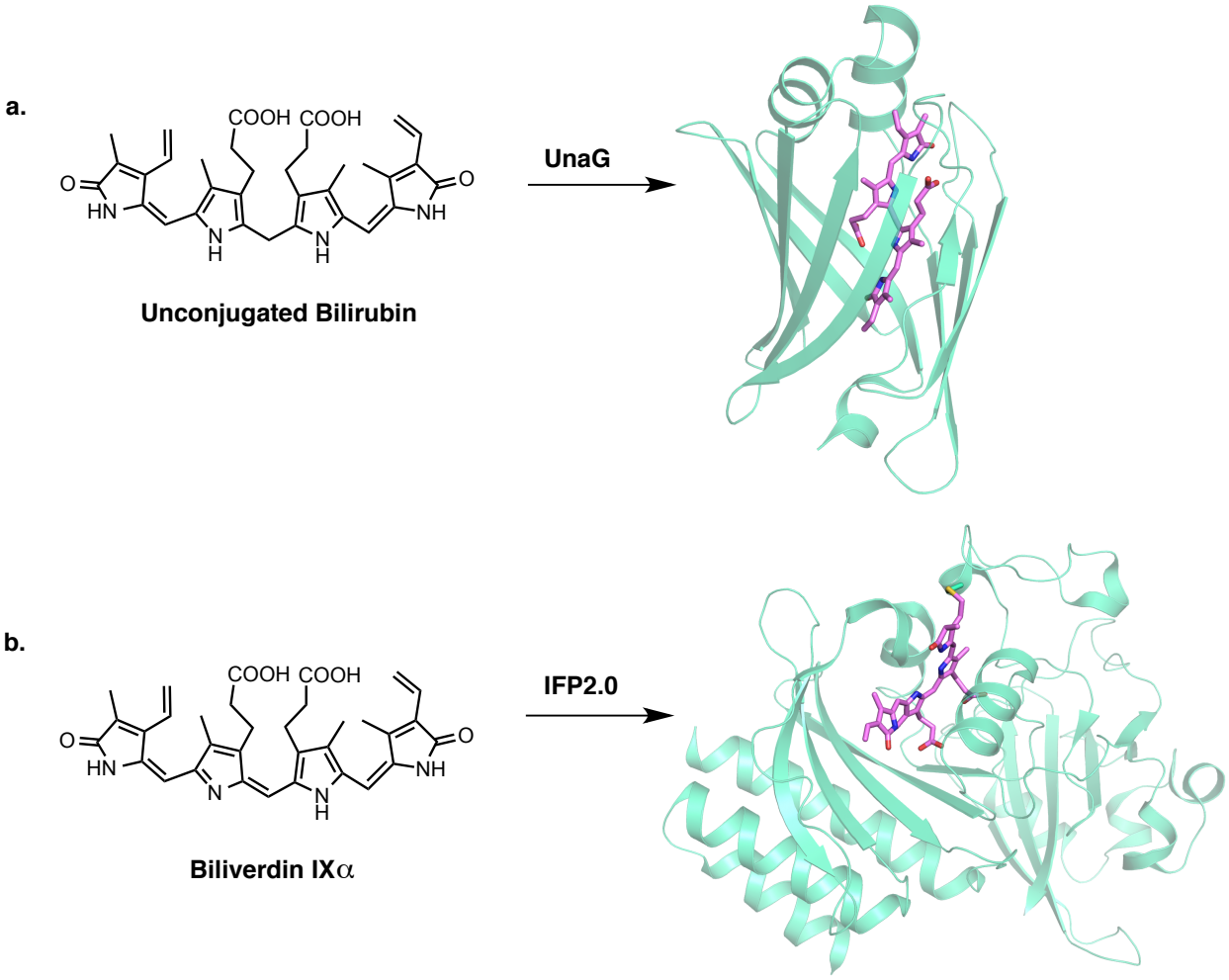


Figure I-3: Endogenous molecule-inducible fluorescence of noncanonical fluorescent proteins. **a.** The chemical structure of bilirubin chromophore and the crystal structure of wild type UnaG (PDB ID: 4I3B). The protein tertiary structure is shown in green and the bound chromophore is shown as sticks in magenta. **b.** The chemical structure of biliverdin ligand and the crystal structure of IFP2.0 (PDB ID: 4CQH). The protein tertiary structure is shown in green and the bound chromophore is shown as sticks in magenta.

with exogenous chromophores. However, in all cases, exogenous molecules are conjugated to their acceptor via one of three mainly applied modes: an enzymatic action, a direct affinity binding, or a bioorthogonal reaction. As anticipated, these small probes should be cell permeable and be nontoxic to the cells.

A variety of site-specific chemical labeling methods have been utilized with the assistance of enzymatic couplings [Halo-Tag,[35-37] SNAP/CLIP-Tags,[38-43] BL-Tag,[44-48] as examples] (**Figure I-4**). These enzymatic probes have attracted attention because they display fast labeling kinetics (rate constants, $k \sim 10^3 - 10^6$ $M^{-1}s^{-1}$) and they allow for time-controlled labeling strategies of target proteins. The exogenous ligands usually contain a fluorescent module with desired spectral features and require a washing step of the treated cells prior to microscopic analysis in order to eliminate unreacted fluorescent probe and reduce the background fluorescence from unspecific binding. However, it has been successfully shown that exogenous ligands can be designed with a quencher moiety to mask the probe fluorescence prior to coupling to their enzyme. Both SNAP-Tag and BL-Tag have utilized such fluorophore-quencher hybrid chemical probes to allow cell imaging without the need of washing steps and to provide high signal to noise ratio with a low background fluorescence (**Figure I-5**).[42-44,48-50] Nonetheless, these probes require multi-step syntheses, limiting cell permeability due to their bulky structures, and slow down the rate of labeling process.

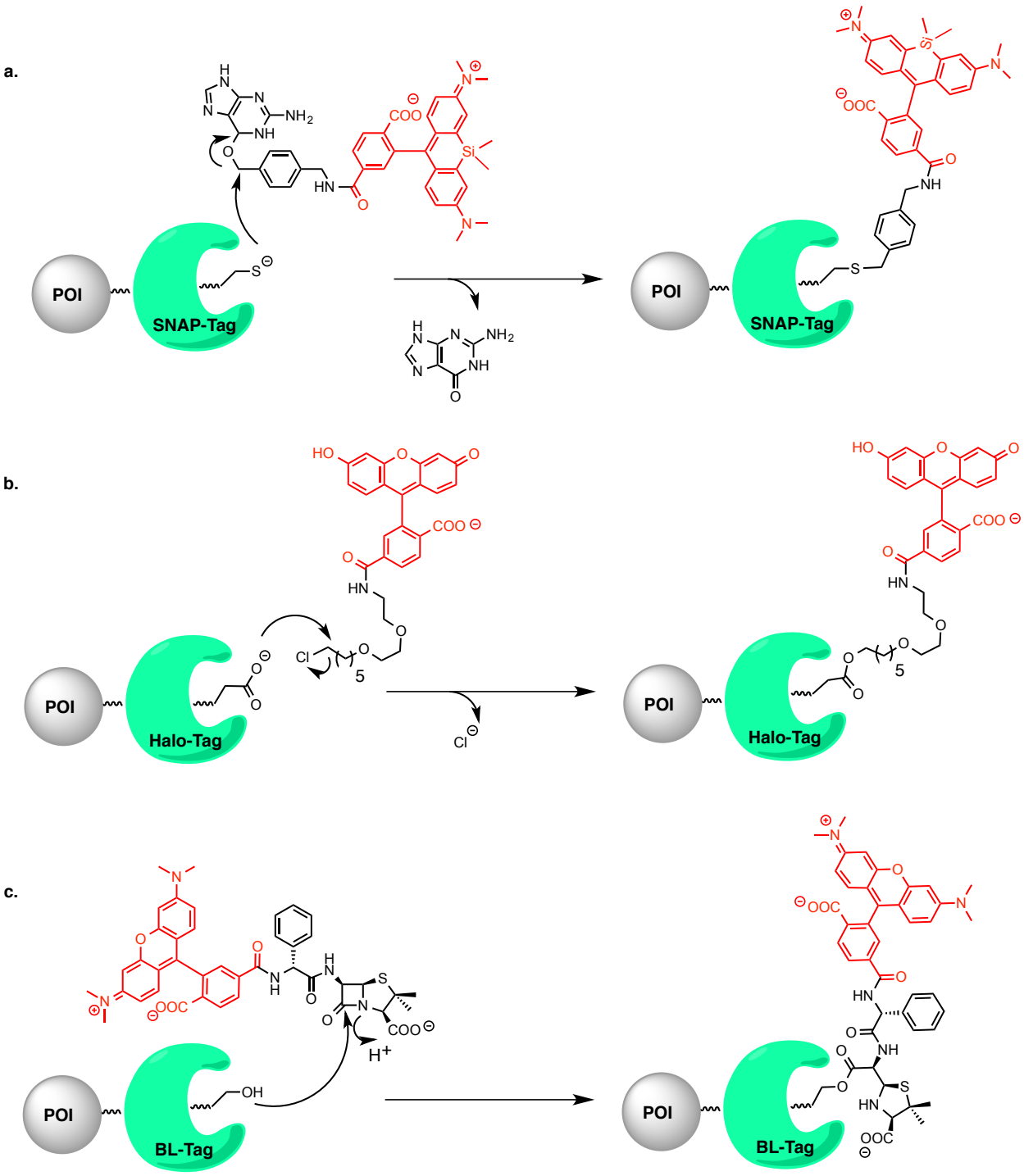


Figure I-4: General labeling mechanism of self-labeling enzymes. Enzyme tags are shown in green cartoon with a fused protein of interest (POI) in gray. The fluorescent module of the ligands is indicated in red **a.** The fluorescent O^6-benzylguanine derivative reacts with SNAP-Tag and labels the protein. **b.** Halo-Tag reacts with primary alkyl chlorides to form an irreversible ester linkage. **c.** BL-Tag opens up the *β*-lactam of ampicillin based fluorescent probes and forms a stable covalent bond.

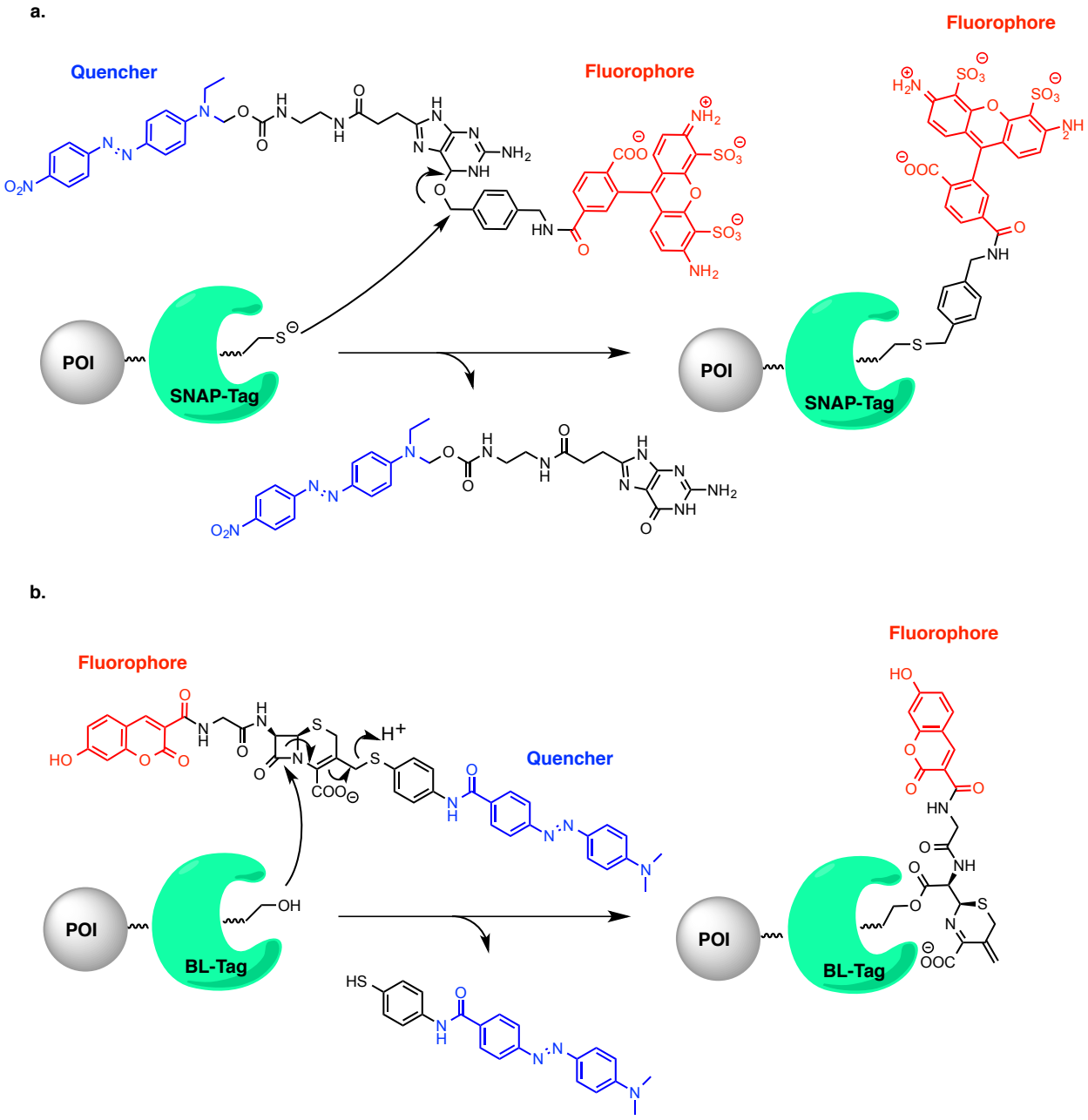


Figure I-5: Structures and labeling mechanism of the fluorophore-quencher hybrid chemical probes. Enzyme tags are shown in green cartoon with a fused protein of interest (POI) in gray. The fluorescent module of the ligands is indicated in red, while the quencher moiety is marked in blue. **a.** The installed quencher at C8 position of guanine ring is eliminated upon coupling of SNAP-Tag with its substrate. The resulting complex is fluorescent whereas the initial probe is a nonfluorescent molecule. **b.** The nucleophilic hydroxyl group of the BL-Tag attacks the amide bond of the *β*-lactam ring and discards the quencher moiety from the cephalosporin based probe to yield a fluorescent complex.

Self-labeling enzymes like SNAP-Tag (20 kDa), CLIP-Tag (20 kDa) and Halo-Tag (33 kDa) are large proteins and can cause potential problems in protein trafficking and function. To overcome such concerns, alternate enzymatic peptide labeling methods have been developed in which the enzyme (not fused with a protein of interest) plays a secondary role in labeling the target protein. In these systems, an enzyme and a protein of interest with a genetically encoded small peptide motif (13-15 amino acids) must be expressed simultaneously in the cell. Assisted by association with the enzyme, the exogenously added chromophore can be incorporated to the small peptide moiety that is expressed as a part of protein of interest. Such systems include lipoic acid ligase (LpIA) or engineered version as fluorophore ligase [51-55] and biotin ligase [56-59] (limited to *ex vivo* imaging) require active site engineering of the enzymes for enabling incorporation of different types of fluorophores and they demand multi-step labeling protocols via bio-orthogonal reactions.

The second class of conjugation between the exogenous ligand and its acceptor tag is based on a direct affinity binding in which the ligand can either form a covalent bond with its acceptor after binding or it can be tightly held on to its acceptor noncovalently with a high binding affinity. PYP-Tag (14 kDa) and TMP-Tag (18 kDa) are good examples of the fluorescent probes where the ligand covalently binds to one of the residues of its acceptor. Photoactive yellow protein (PYP) is a blue light receptor isolated from purple eubacterium, *Halorhodospira halophila* having a *p*-coumaric acid (*p*CA) as its chromophore (**Figure I-6a**).[60-63] Although the free chromophore, *p*CA, absorbs at 284 nm in solution, the formation of the thioester bond between the *p*CA and

Cys^{69} residue along with specific protein-chromophore interactions results in a strong absorption in the visible region, exhibiting a yellow PYP-*p*CA complex (λ^{max} = 446 nm).[64-66] Recently, Kikuchi *et al*. has demonstrated that PYP can be developed into a fluorescence tag with a novel design of fluorogenic ligands (**Figure I-6b**).[67,68] The authors used 7-dimethylaminocoumarin thioester derivates, which have low

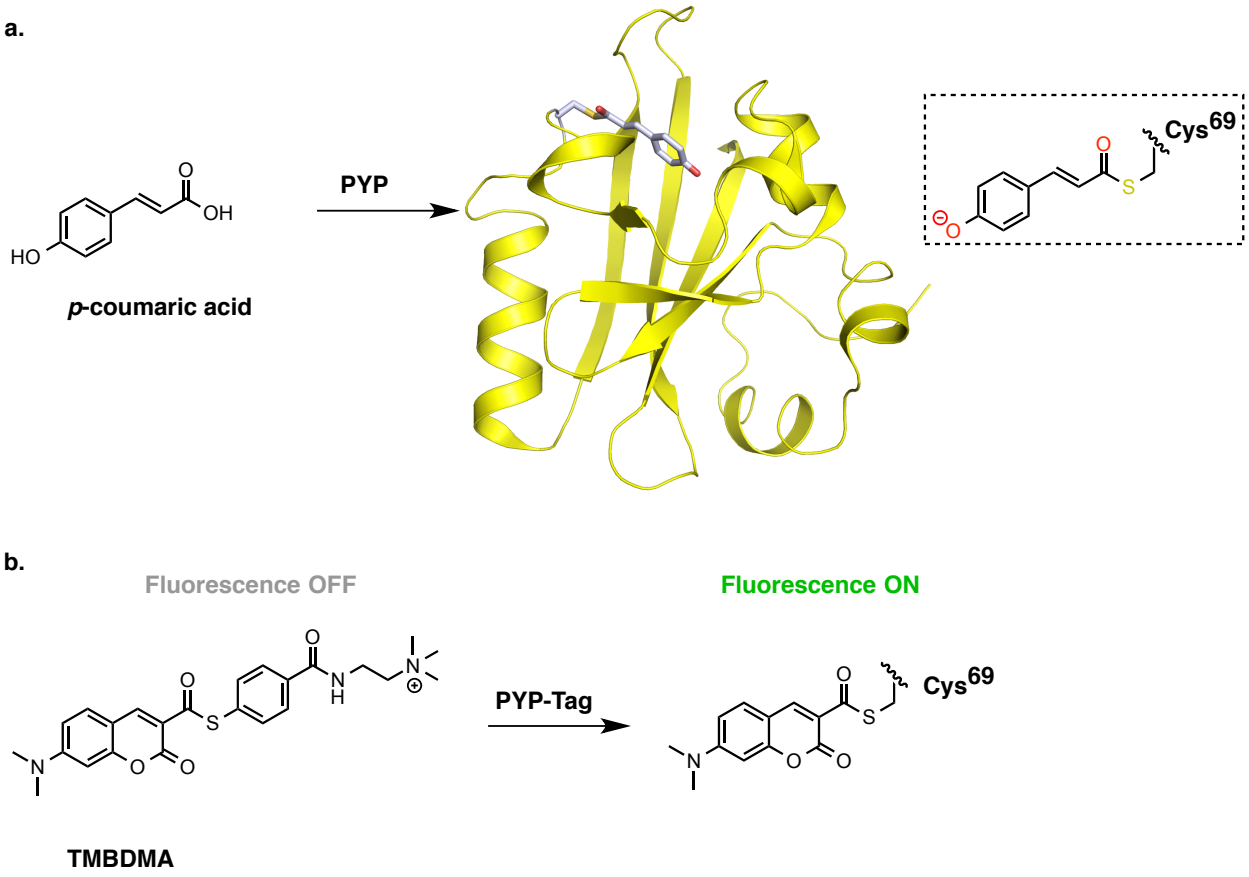


Figure I-6: Design of a PYP-Tag. **a.** The structure of the natural substrate of PYP and the crystal structure of its bound form. The protein tertiary structure is shown in yellow (PDB ID: 2PHY) with its ligand indicated in a stick model. The dashed box shows the binding mode of the ligand to PYP via a thioester linkage. **b.** The structure of a TMBDMA ligand that shows minimal fluorescence in polar environment. Upon transthioesterification, it covalently binds to PYP-Tag and yields a fluorescent complex.

fluorescence in polar solvents (quantum efficiency ~ 0.02), but become fluorescent in less polar environments like protein binding pockets. The subsequent efficient transthioesterification between the nucleophilic cysteine residue of the PYP and the designed ligand, TMBDMA, creates a low polar environment around the ligand yielding a complex (λ^{abs} = 450 nm, λ^{ems} = 487 nm) with quantum efficiency of 0.38 (**Figure I-6b**).

Similar to the PYP-Tag, TMP-Tag was used as an alternative chemical tagging system in which the ligand and the protein tag forms a stable covalent bond.[69-71] The design strategy is based on the high affinity interaction between *Escherichia coli* dihydrofolate reductase (eDHFR) and trimethoprim (TMP), which is a strong inhibitor for eDHFR (enzyme inhibition, K_i ~ 1 nM).[72] The engineered analog of eDHFR:L28C and synthetically designed acrylamide-TMP-fluorophore ligand can be covalently coupled via proximity induced Micheal reaction between the Cys^{28} residue and the acrylamide appendage. Since TMP displays high inhibition affinity only for eDHFR but not for mammalian forms of the enzyme, TMP-Tag is well suited for mammalian cell imaging with high selectivity (**Figure I-7**).

Fluorogen activating proteins (FAPs) are noncovalent peptide based fluorescent reporters. These FAPs are derived from libraries of human scFv's through a yeast display technology.[73-76] They bind profluorogenic molecules such as thiazole orange (TO) and malachite green (MG) with nanomolar affinity and generate fluorescence with high levels of brightness.[74,75,77] They have been applied in visualizing proteins on cell

surface and intracellular organelles.[74-76,78,79] FABs are a novel platform for fluorescent tagging of proteins since they couple with nonfluorescent molecules and eliminate the need of washing steps prior to imaging. However, their application and adaptability is limited as a result of requiring molecular manipulation for a particular fluorogen chosen. Additionally, they are mostly restricted to extracellular cell surface imaging or certain compartments in which a non-reducing environment is necessary for their proper folding (disulfide bond formation), stability and function. The current research with FAPs is

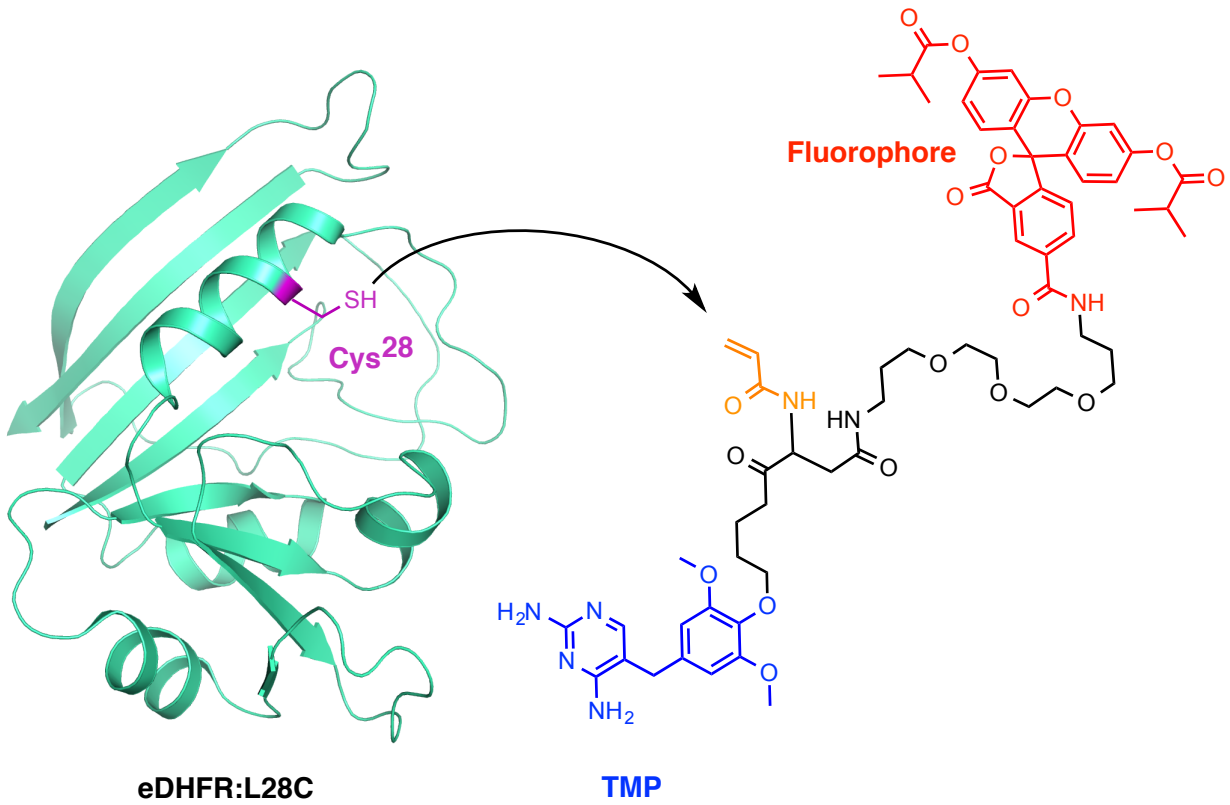


Figure I-7: Schematic illustration of the proximity-induced Michael addition between the nucleophilic thiol and the electrophilic acrylamide moiety of the acrylamide-TMP-fluorophore conjugate. The tertiary structure of the eDHFR is shown in green (PDB ID: 1RD7) and the thiol nucleophile of the cysteine at position 28 is indicated in magenta. The TMP, fluorophore and the acrylamide sections of the ligand are marked in blue, red and orange colors, respectively.

followed on developing more libraries for optimizing their utility in intracellular milieu (independence of disulfide bond formation) and for broadening the choice of fluorogenic molecules with desired spectroscopic features.[76,80,81]

The aforementioned methods for coupling a protein with its ligand require exogenous molecules either via enzyme assistance or a direct affinity binding. The third method to be discussed is a bioorthogonal method, which also requires the addition of an exogenous ligand. The ligand can be conjugated to the protein of interest in one of two ways: first, a chemical recognition peptide motif is fused to a protein of interest and the metal containing ligand is coupled with the peptide via metal-ligand chelation, and secondly, a biosynthetic machinery modifies the protein of interest within the cell environment to install a reactive bioorthogonal functional group that can react with the exogenous ligand upon addition. The first method is a relatively simpler approach and various small recognition peptide motifs can be added to the protein of interest as a genetic fusion.[82] Tetracysteine/biarsenical system,[83-87] oligohistidine/nickel-complex system,[88-92] and oligoaspartate/zinc-complex system[93-95] are examples of such recognition mediated labeling in which the selectivity is achieved by the chemical recognition between electron-rich peptide segments (6-10 amino acids) and electron-deficient metal chelates (**Figure I-8**). Various fluorescent dyes can be used to generate a wide range of fluorophore-carrying metal complexes. However, the main problem with these systems originates from the toxicity of the high dose of metal complexes and

nonspecific binding of them to similar or related sequence motifs on other proteins within the cell environment.

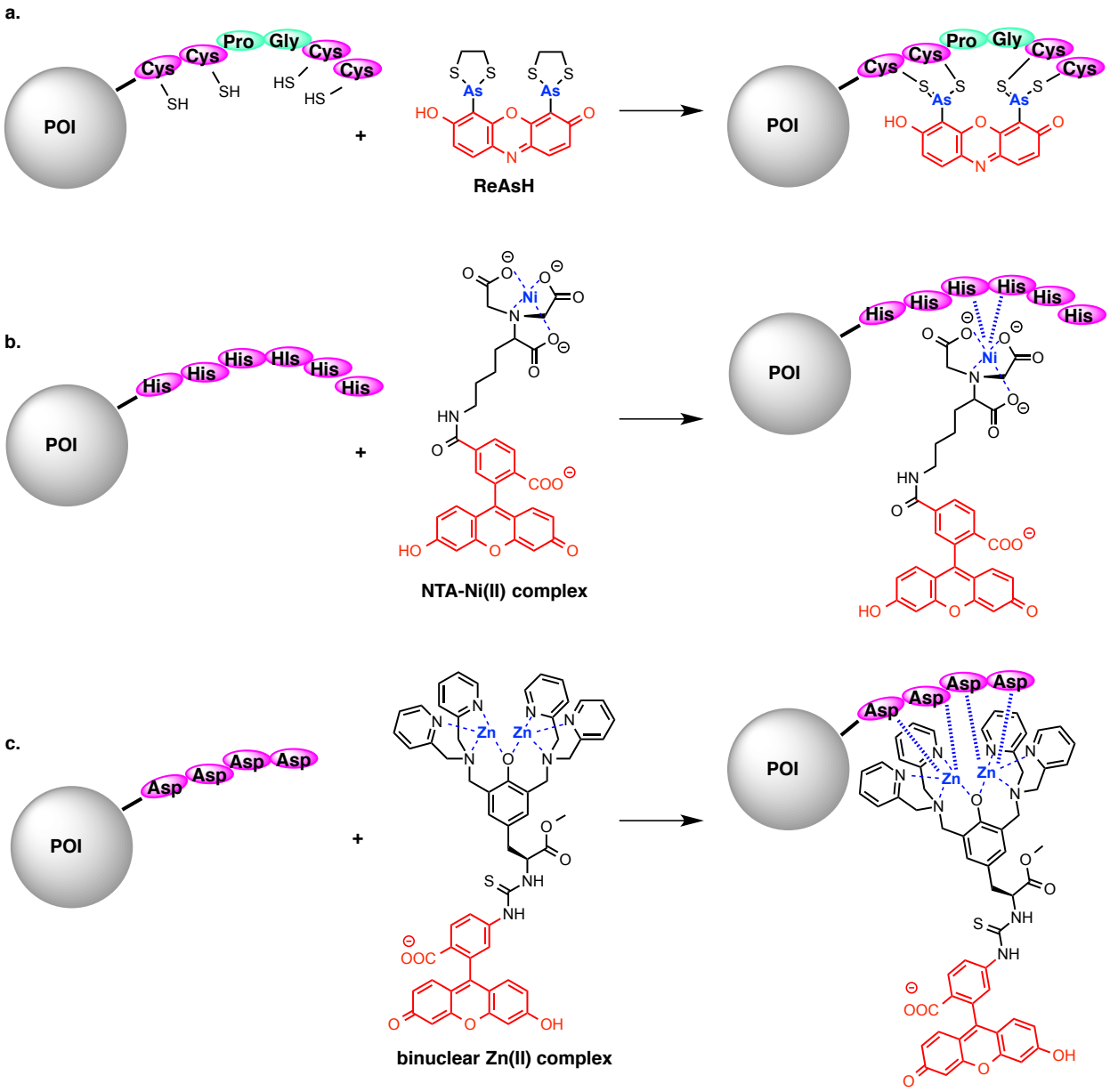


Figure I-8: Strategies for the bioorthogonal recognition mediated chemical labeling of proteins. Protein of interest (POI) is indicated in gray sphere with a sequence specific small recognition peptide motif shown in magenta ellipsoid. The structures of fluorophore-carrying metal complexes display the fluorophore moiety (in red) and the metal center (in blue). **a.** Tetracysteine/biarsenical system. **b.** Oligohistidine/nickel-complex system. **c.** Oligoaspartate/zinc-complex system.

In the second approach, the protein of interest must carry a bioorthogonal functional group within its structure to chemically couple with the exogenous ligand. The biosynthetic machinery in the cell should install this functional group either during protein translation (via unnatural amino acids) or through a post-translational peptide modification. The exogenous ligand having a complementary functional group should be added to the cell to accomplish the labeling process. A large diversity of functional groups have been used for bioorthogonal chemical labeling such as azides, alkynes, nitrones, alkenes, ketones, aldehydes and aryl-tetrazines enabling multiple methods to be used simultaneously for labeling different target proteins in the same biosystem.[32,96-99] Staudinger ligations,[100-104] alkyne-azide cycloadditions,[55,105-111] and Diels-Alder cycloadditions[112-116] are one of the few examples applied in selective labeling of proteins and in *vivo* imaging (**Figure I-9**). Advantages of these systems include the ability to design fluorogenic reporters (wash free imaging) and the ability to incorporate only a small functional group into protein of interest, which will not alter the structure and the function dramatically. Nevertheless, these systems require biological methods to install functional groups into the target proteins and slow labeling kinetics (k up to 10^4 $M^{-1}s^{-1}$).[98]

Several methods have been developed over the years for specific protein labeling based on a diverse array of chemistry on protein modification and small molecules. Yet, the design of more probes and advanced imaging technologies is an active ongoing research area, pushing both experimental and conceptual limitations in

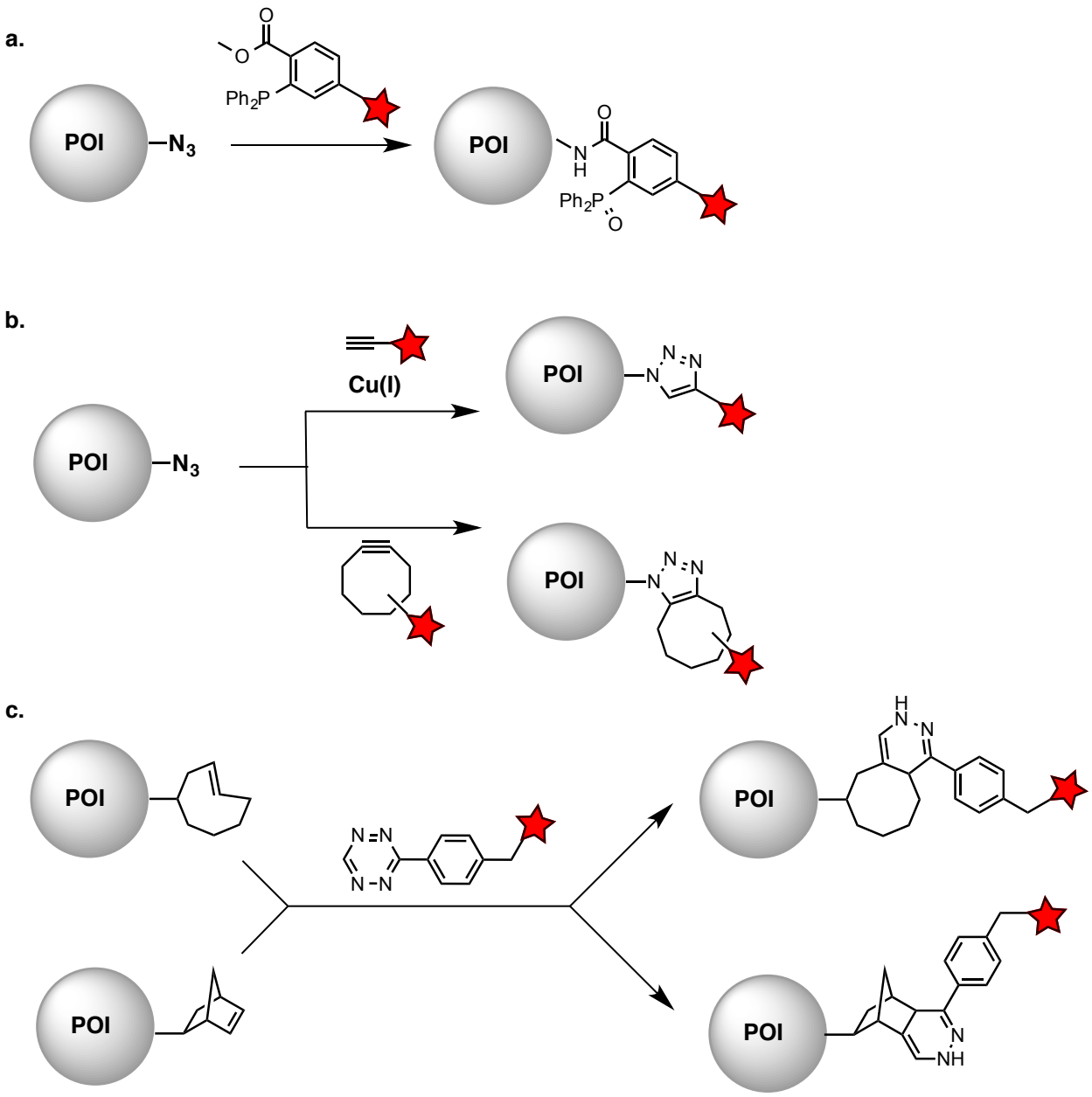


Figure I-9: Bioorthogonal reactions with different functional groups. The bioorthogonal functional group on protein of interest (POI, shown in gray sphere) reacts with the exogenous molecule carrying a complementary functional group and a fluorescent reporter represented as a red star. **a.** The Staudinger ligation of azides and triarylphosphines. **b.** [3+2] cycloadditions of azides and alkynes to form triazoles. Terminal alkynes require Cu(I) as a catalyst to achieve cycloaddition with azides (top). Cyclooctynes react with azides under physiological conditions without Cu(I) requirement via strain-promoted cycloaddition (bottom). **c.** Inverse-demand Diels-Alder reaction of *trans*-cyclooctene and monoaryl tetrazine (top). Inverse-demand Diels-Alder reaction of norbornene and monoaryl tetrazine (bottom).

order to address more complex questions in biological systems. From the compilation of recent advances in the use of fluorescence in molecular and cellular systems, one can argue that the union of a genetically addressable system (biomolecules) with the flexibility in attaining the desired spectroscopic characteristics (small molecules) would provide fundamental avenues for improvement of protein fusion tags with increasingly stringent and specialized properties. In this report, we introduced a fluorescent system that not only requires the union of a molecule and a protein for ultimate flexibility (as alluded to above), but also utilizes a nonfluorescent molecule that becomes fluorescent only upon binding and reaction with the target protein. The initial foray, disclosed herein, describes the pairing of a small dye precursor with sets of protein mutants, exhibiting red shifted emission with high quantum efficiencies and brightness.

I.2 Choice Of Protein For Engineering A Fluorescent Tag

We have recently succeeded in designing protein-embedded chromophores to control their absorptive properties by judicious manipulation of the protein architecture surrounding the chromophore.[117-123] This inspired us to consider a system that combines the strengths of protein engineering and synthetically designed fluorogenic molecules to provide a wide range of fluorescent complexes that specifically and quantitatively label the target proteins. Our previous work has demonstrated that all-*trans*-retinal, the vitamin A derivative, results in colored pigments upon complexation with small cyctosolic proteins. The iminium-based complexation is achieved with the electrophilic aldehyde moiety of the retinal and an active site nucleophilic lysine residue. The studied proteins, human Cellular Retinol Binding Protein II (hCRBPII) and human

Cellular Retinoic Acid Binding Protein II (CRABPII), are small cyctosolic proteins (~ 15 kDa) which belong to a family of intracellular lipid binding proteins (iLBPs). Similar to other members of the iLBP superfamily, hCRBPII and CRABPII have a common three-dimensional clamshell structure composed of two orthogonal β-sheets (each includes five antiparallel β-strands) and two α-helices.[124-128] Both hCRBPII and CRABPII hold

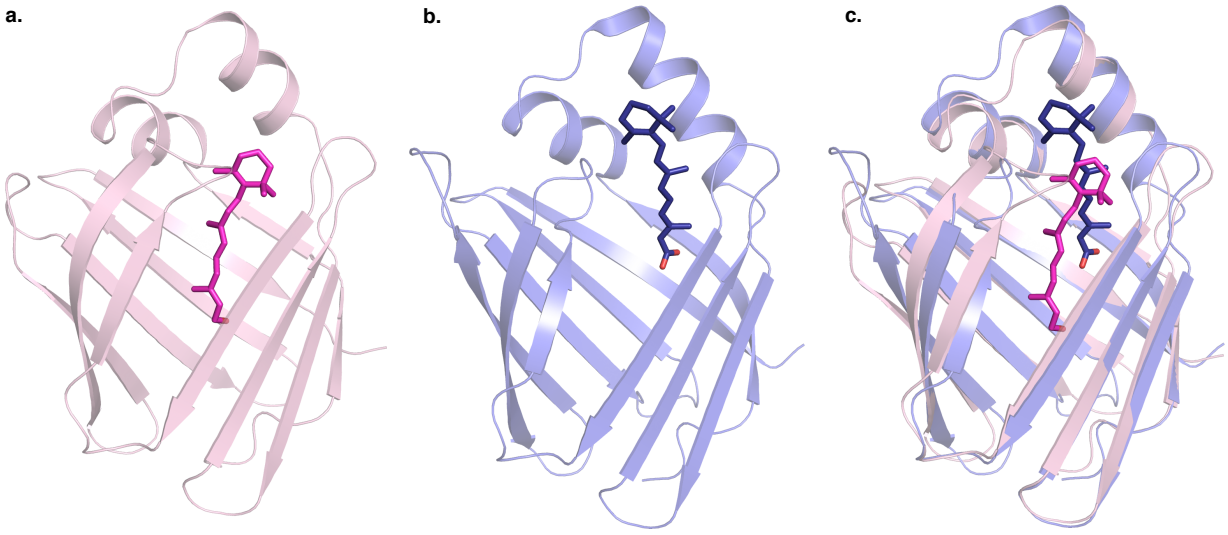


Figure I-10: Crystal structures of wild type hCRBPII and CRABPII. **a.** All-*trans*-retinol bound wild type hCRBPII (PDB ID: 4QZT). **b.** Crystal structure of wild type CRABPII bound to all-*trans*-retinoic acid (PDB ID: 2FR3). **c.** Overlaid structures of hCRBPII (pink) and CRABPII (blue) demonstrate the structural similarity between the proteins. The bound ligands, retinol (magenta) and retinoic acid (indigo) are shown as stick models.

their natural ligands, retinol and retinoic acid respectively, in their hydrophobic central cavity (**Figure I-10**).[117-121,123,129] Their remarkable tolerance of mutational studies and robust β-barrel topology with a large binding cavity (~ 600 $Å^3$) provided ideal frameworks for applications in protein re-design.

Our former lab members, Dr. Wenjing Wang and Dr. Tetyana Berbasova have engineered hCRBPII into a fluorescent tag and successfully demonstrated its utility *in*

vivo imaging studies (unpublished data). This study focuses on the reengineering of the CRABPII into a series of fluorescent proteins with expanded structural insights of chromophore-protein interactions.

I.3 Design Of A Fluorogenic Cyanine Compound

As a choice of fluorophoric precursor, we focused on cyanine dyes that have played a central role in spectroscopic applications related to molecular biology. In comparison to common fluorescent molecules, cyanine dyes are mildly fluorescent in solution, probably due to the their backbone flexibility.[130-134] However, their high extinction coefficients, large absorption cross sections and innocuous nature in biological systems make them one of the most commercialized fluorescent probes for labeling strategies.[133,135-137]

Classical cyanine dyes are cationic molecules in which two heterocyclic rings are joined together by a chain of conjugated polymethines (**Figure I-11a**). The heterocyclic rings may be identical or different providing a ring with tertiary nitrogen and another ring with quaternary nitrogen. Inspired by the cyanine dye architecture, the precursor merocyanine aldehyde **1** is synthesized (**Figure I-11b**). This profluorogenic aldehyde can form an iminium, Schiff base (SB), with primary amines such as *n*-butyl amine or an active site lysine residue within a protein binding pocket. The subsequent protonation of the merocyanine-SB yields a permanent resonating cation, resulting in a fluorophoric entity (merocyanine-PSB, **2**). This push pull system, which terminates at the nitrogen atoms at the two ends of the polyene, leads to a bathochromically shifted chromophore with an absorption profile distinct from its parent aldehyde.

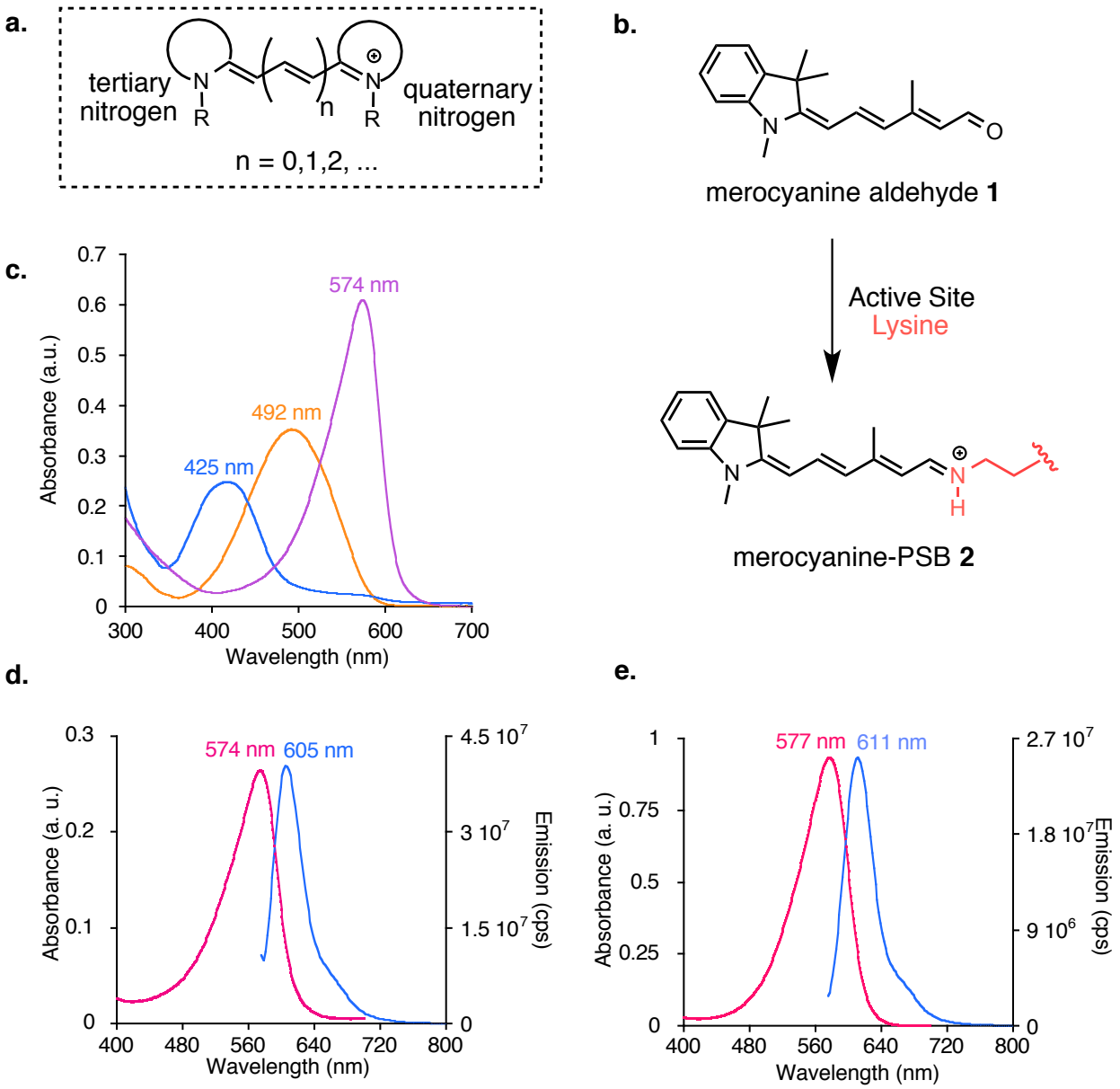


Figure I-11: Characterization of cyanine dye precursor, merocyanine aldehyde **1**. **a.** Generic structure of a classical cyanine dye, in dash box. **b**. Merocyanine aldehyde **1** binds to an active site lysine residue to generate the red-shifted fluorescent cyanine pigment, merocyanine-PSB **2**. **c.** UV-vis spectra of **1** (orange), Schiff base of **1** with *n*-butyl amine (blue), and the corresponding protonated Schiff base of **1** upon acidification (purple). All spectra are in phosphate buffer saline (PBS). **d.** Absorption and emission spectra of **2** in PBS. The emission spectrum is obtained with excitation at 565 nm and collected from 575 nm to 800 nm. **e.** Absorption and emission spectra of **2** in ethanol. The emission spectrum is obtained under identical settings in **d.**

Merocyanine aldehyde **1** absorbs maximally at 492 nm in phosphate buffer saline (PBS) while the corresponding SB (formed by reaction with *n*-butyl amine) blue shifts to

425 nm (**Figure I-11c**). Acidification results in a large red shift in the UV-vis spectrum, consistent with the formation of the iminium (λ^{abs} = 574 nm). As expected, the resulting merocyanine-PSB **2** becomes mildly fluorescent in PBS solution (quantum yield, Φ = 4 %, λ^{ex} = 565 nm, λ^{ems} = 604 nm), a result of the two nitrogen atoms being in conjugation through the polyene backbone. The change in solvent from PBS to ethanol increases the quantum efficiency of merocyanine-PSB **2** by two fold (Φ = 8 %, λ^{ex} = 565 nm, λ^{ems} = 611 nm) exhibiting environmentally sensitive fluorescence properties like other cyanine dyes (**Figure I-11d** and **Figure I-11e**).[130-132,138] Additionally, merocyanine-PSB **2** displays photophysical properties similar to sulfoindocyanines and unsymmetrical imine-based trimethine and pentamethine cyanine dye analogs which were reported previously.[136,139]

Merocyanine aldehyde **1** structurally mimics the retinylidene ligands that are already shown to bind strongly to engineered CRABPII proteins. The synthesized aldehyde has a polyene tail like retinal backbone and the indoline head group similar in size to the ionone ring. Thus, it could easily couple with CRABPII proteins and form a stable PSB. The merocyanine aldehyde **1** is additionally attractive for our studies because the subsequent PSB formation of the aldehyde within the protein cavity would result in distinct spectroscopic properties (in *situ* formation of a cyanine dye) compared to the nonfluorescent free aldehyde allowing negligible background for imaging studies. Furthermore, the sequestration of the bound ligand by the protein environment could

rigidify the molecule, leading to enhancement of fluorescence, as it was observed for many cyanine dyes in viscous solvents. This element gives an opportunity for us to design the binding pocket of the protein in order to reduce motional freedom of the bound ligand for further increase in fluorescence. Additionally, cyanine dyes display environmentally sensitive fluorescence properties, which is also taken into account as a second tunable element. In summary, we would pursue the study of merocyanine/CRABPII fluorescent complexes by rigidifying the bound chromophore and altering the electrostatic environment in protein cavity.

I.4 Complexation Of Merocyanine Aldehyde 1 With CRABPII Mutants And Structure-Guided Mutagenesis

We have successfully demonstrated that CRABPII could be reengineered into rhodopsin mimics capable of binding all-*trans*-retinal and C15 aldehyde (a shorter retinal analog) to form PSBs via a lysine residue. The nucleophilic lysine was introduced at two alternative positions within the protein cavity to generate two different generations of CRABPII mimics. In the first generation, arginine 132 was mutated to a lysine residue (as an active site lysine) and a second arginine residue at position 111 was changed into a hydrophobic side chain, leucine. This design also required the installation of a counter anion in order to obtain a stable PSB between the all-*trans*-retinal and the engineered R132K:R111L:L121E (abbreviated as KLE) CRABPII protein under physiological conditions.[118,120,121] The structural data of the first generation of CRABPII mutants showed that the aldehyde end of the retinal was deeply buried within the protein cavity, whereas the ionone ring of the chromophore was not encapsulated by the

protein, yielding a solvent exposed protein-chromophore complex (**Figure I-12a**). The resulting buffered chromophore with aqueous media was unresponsive to mutational

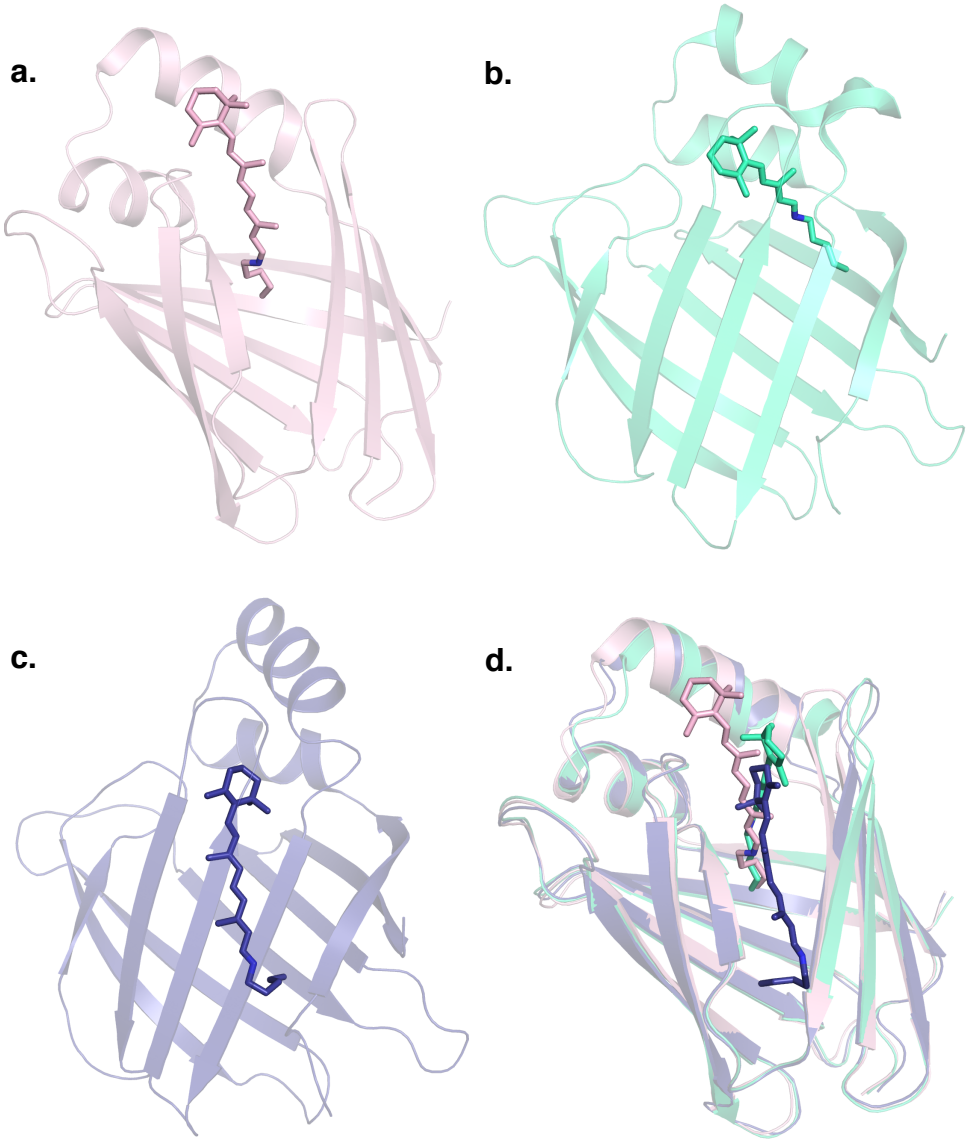


Figure I-12: Crystal structures of CRABPII rhodopsin mimics. **a.** The structure of the first generation CRABPII KLE mutant bound to all-*trans*-retinal (PDB ID: 2G7B). **b.** Crystal structure of CRABPII KLE:R59W/C15 aldehyde complex (PDB ID: 3F8A). The shorter chromophore is embedded in the protein binding pocket. **c.** Crystal structure of a second generation of CRABPII variant (CRABPII-2nd) bound to all-*trans*-retinal (PDB ID: 4I9S). The ligand is fully encapsulated by the protein. **d.** Overlay of all three crystal structures (CRABPII KLE/retinal in pink, CRABPII KLE:R59W/C15 in green, and CRABPII-2nd/retinal in dark blue) illustrating the differences in the coverage of the bound ligand by the protein scaffold.

changes of protein and not tunable in terms of absorptive characteristic. To address the problem, a shorter retinal analog, C15 aldehyde was synthesized. This short chromophore was complexed with CRABPII mutants to yield observable wavelength regulation with point mutations (**Figure I-12b**).[119] As an alternate route, the second generation of CRABPII rhodopsin mimic (CRABPII-2nd) was designed in which an active site lysine residue was placed deeper in the binding pocket of CRABPII. With this approach, the chromophore binding and the subsequent PSB formation were accomplished ~ 6 Å deeper inside the protein cavity, resulting in a substantially modulated absorption profile of the bound retinal (**Figure I-12c** and **Figure I-12d**).[117] Having in hand various CRABPII mutants, we have pursued the pairing of merocyanine aldehyde **1** with the first generation of CRABPII rhodopsin mimics to study a system that not only requires the union of a molecule and a protein for ultimate flexibility, but also utilizes a nonfluorescent merocyanine precursor becoming fluorescent only upon binding with the target protein.

From the simplest mutant to the more complex mutants, merocyanine aldehyde **1** was coupled with CRABPII variants to trigger the *in situ* formation of a cyanine dye. Incubation of the ligand with CRABPII mutants generates pigments that exhibit remarkable bathochromicity and narrow absorption bands (**Table I-1**). Time dependent binding of the R132K:R111L (KL) mutant with merocyanine aldehyde **1** shows a clear isosbestic point indicating homogenous formation of the PSB (**Figure I-13a**). Subsequent addition of retinal does not displace the merocyanine from the binding site, demonstrating that the covalent linkage between the merocyanine ligand and the protein

Table I-1. Spectroscopic data of CRABPII mutants complexed with merocyanine **1**.

Entry	CRABPII Variants	λ^{abs} / λ^{ems} (nm)	Φ(%)	ε (M.cm)$^{-1}$
1	R132K:R111L (KL)	600 / 617	18	111,700
2	R132K:R111L:L121E (KLE)	591 / 612	33	93,200
3	R132K:R111L:L121D (KLD)	591 / 612	30	97,500
4	KL:R59W	602 / 619	23	140,100
5	KL:R59Y	602 / 617	14	n.d.
6	KLD:R59W	595 / 615	38	n.d.
7	KLE:R59W	595 / 616	39	169,800
8	KLE:R59H	589 / 615	30	124,600
9	KLE:R59Y	594 / 615	32	n.d.
10	KLE:R59F	593 / 613	37	160,800
11	KLE:R59L	594 / 613	33	142,000
12	KLE:V76W	569 / 605	33	77,600
13	KLE:S37W	579 / 609	30	96,500
14	KLE:I31W	582 / 611	28	n.d.
15	KLE:L28W	581 / 611	31	58,600
16	KLE:P39W	582 / 619	17	n.d.
17	KLE:M123W	592 / 609	32	117,000
18	KLE:A32W	590 / 614	31	104,300
19	KLE:V76W:S37W	569 / 603	27	n.d.
20	KLE:V76W:L28W	571 / 609	32	67,000
21	KLE:R59W:I31W	595 / 615	39	n.d.
22	KLE:R59W:L28W	595 / 619	38	158,400
23	KLE:R59W:M123W	595 / 611	29	167,800
24	KLE:R59W:A32W	595 / 617	39	142,900
25	KLE:R59W:L28W:I31W:A32W	597 / 623	32	n.d.

Φ = quantum yield, ε = extinction coefficient of the protein/**1** complex, n.d. = not determined

is stable. Excitation of the KL/**1** complex at 565 nm yields a narrow emission spectrum that peaks at 617 nm (**Figure I-13b**). Base titration of the iminium within the protein cavity reveals a pK_a of 9.6, well above the physiological range (**Figure I-13c**). The high pK_a of the iminium helps the protein-chromophore complex to preserve its protonated stage with a long wavelength absorption band. However, at elevated pH regimes, the

iminium is deprotonated, giving rise to a short wavelength absorption at around 425 nm, which is a typical consequence of deprotonation of cyanine dyes (**Figure I-13e**).[140-142]

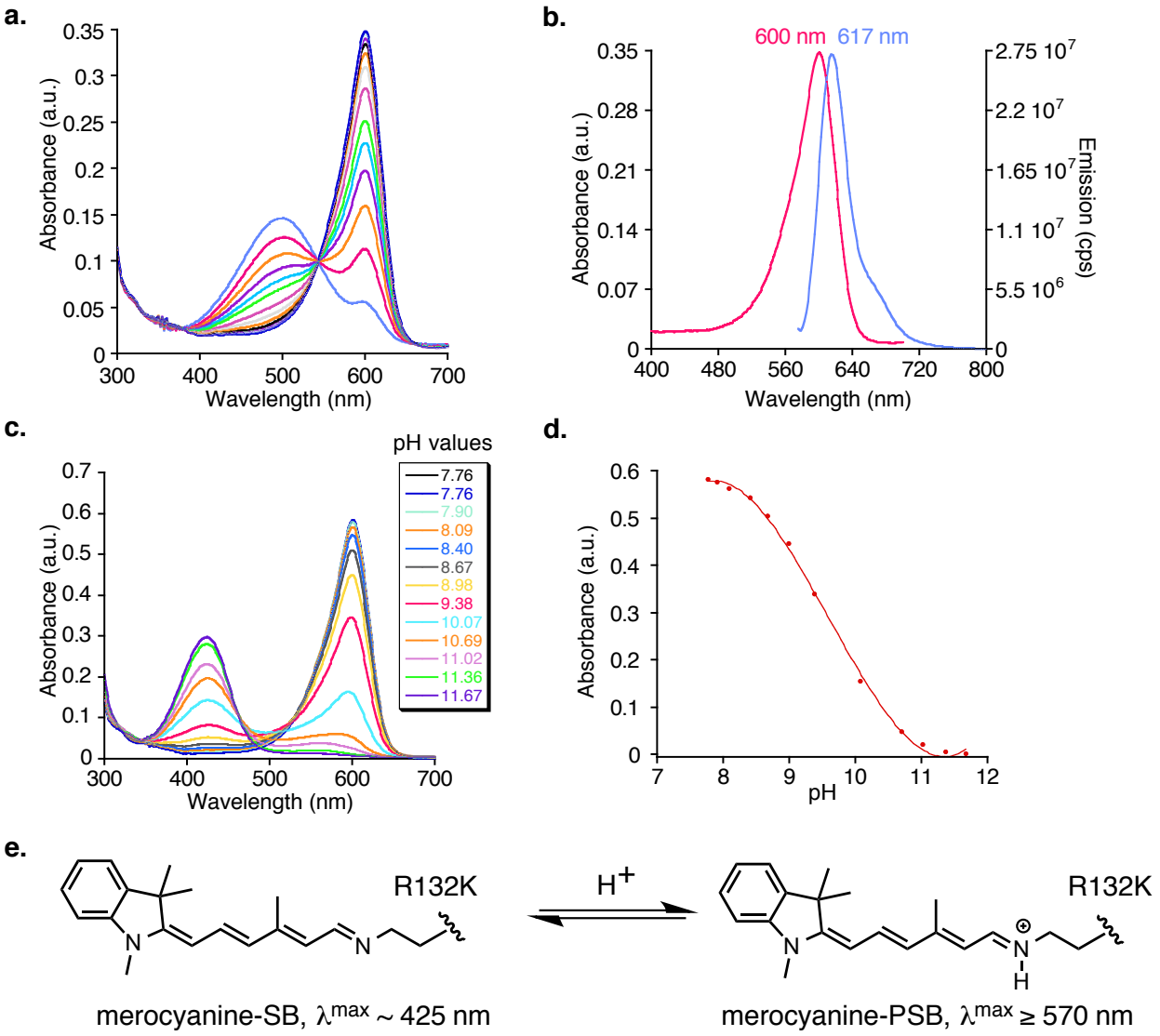


Figure I-13: Spectroscopic data of KL/**1** complex. **a.** Time course binding scan of merocyanine aldehyde **1** (0.3 equivalents) with KL mutant (20 μM). Absorption spectra were taken at 2 minutes intervals. **b.** Absorption and emission spectra of KL/**1** complex (excitation at 565 nm). **c.** Base titration of KL/**1** complex. Upon addition of NaOH, the peak at 600 nm blue shifts to 425 nm. The pH values are indicated in the box on right. **d.** Absorbance of KL/**1** complex at 600 nm as a function of pH with an apparent pK_a of 9.6. **e.** Protonation of imine (merocyanine-SB) results in merocyanine-PSB in CRABPII mutants yielding red-shifted pigments.

With a stable complex of merocyanine aldehyde **1** with the KL mutant in hand, we explored the possibility to regulate the absorption and emission spectra of the complex by using perviously established strategies in our lab. These include changing the electrostatic environment along the chromophore, altering the interactions of the counter anion for the PSB, and creating a more enclosed binding pocket for the bound ligand. To begin with, the electrostatic environment in the vicinity of the PSB was altered by introducing a negatively charged residue (glutamine) at position 121. The resulting KL:L121E (referred as KLE) showed a slight hypsochromic shift in the absorption of the protein-ligand complex (**Table I-1**, entry 2). This is due to stabilizing the cationic charge on the iminum nitrogen atom and decreasing the charge distribution along the chromophore (less delocalization). Similarly, installation of an aspartate residue as the counter anion displayed similar photophysical characteristics as the KLE mutant (KLD, entry 3, **Table I-1**). Although, we observed only small changes in absorption and emission profile of the resulting mutants, both KLE and KLD complexes showed improved fluorescence quantum yield in comparison with KL/**1** complex (see further discussion in Section **I-5.2**).

We had showed that the effect of closing the binding cavity and creating a less solvent accessible binding pocket is crucial in terms of enhancing protein-ligand interactions. The embedded and completely encapsulated bound retinoids became more sensitive towards small changes in protein environment and displayed significantly tunable absorptive properties. This phenomenon could be specifically important for the bound merocyanine dye because cyanine dyes exhibit environmentally sensitive

fluorescence properties under different conditions such as different solvent polarity, dielectric constant, ionic strength and viscosity. Previous reports have documented decreased levels of fluorescence as a function of increasing solvent polarity and vice versa.[131,132,143,144] In a similar fashion, isolation of the binding pocket from the aqueous media could have a significant impact on merocyanine-CRABPII complexes, achieving enhanced fluorescence properties. Therefore, we focus on introducing hydrophobic residues at the entrance of the protein cavity. Arginine 59 (Arg59) located at the portal of the binding site was changed to a number of different amino acids by using the KLE mutant as a template. Spectroscopic analyses of these mutants showed little change in the absorption and emission wavelength as well as in quantum yields. However, the replacement with tryptophan and phenylalanine exhibited improved quantum efficiencies (**Table I-1**, entries 6 to 11). It is possible that placement of a large hydrophobic aromatic residue provides favorable interactions with the bound ligand and better encapsulates the binding cavity, whereas more polar residues like histidine and tyrosine might retain similar solvation around the chromophore neighborhood as the original arginine, yielding comparable fluorescence properties. Furthermore, we observed the same improvement in quantum yield with tryptophan installation in the KL double mutant while mutation to tryrosine resulted in a decrease in quantum efficiency (KL:R59W vs KL:R59Y, see **Table I-1** entries 3 and 4). Interestingly, the mutants of R59W displayed significantly increased extinction coefficients of the protein-ligand complexes, leading to high fluorescence brightness (for examples see **Table I-1** entry 1 vs 4, or 2 vs 7). Together

these results indicate that the replacement of Arg59 with Trp is unique in improving fluorophoric properties of the protein-chromophore complex.

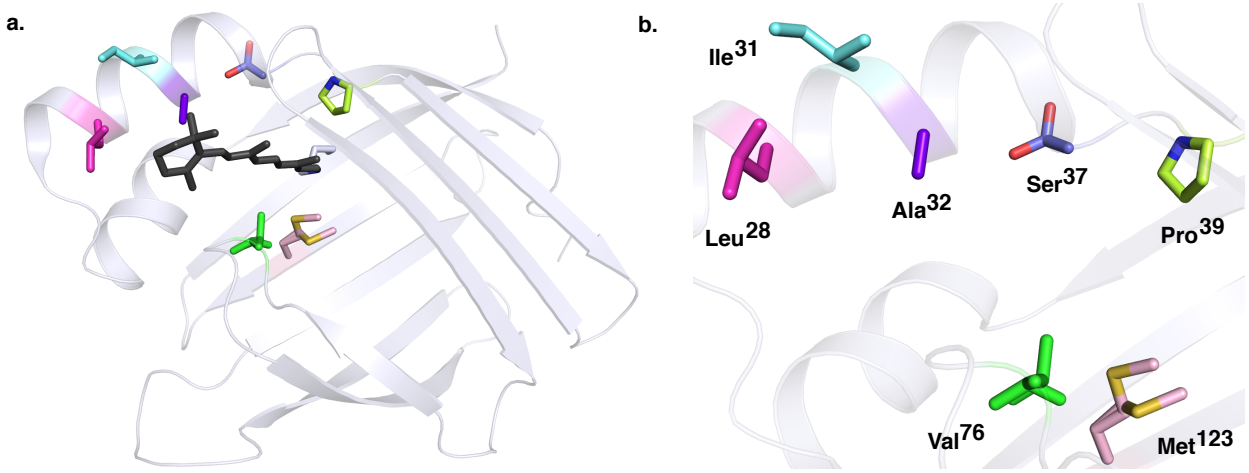


Figure I-14: Crystal strcture of KLE CRABPII bound to all-*trans*-retinal (PDB ID: 2G7B). **a.** The residues that are in close contact with the bound chromophore (in dark gray) are shown in sticks. Each residue is highlighted with a different color. **b.** Exploded view of the protein binding pocket showing the residues discussed in the text.

Further improvements in the fluorescence properties of the merocyanine bound CRABPII complexes were focused on rigidifiying the ligand through packing interactions with large amino acid side chains. It has been shown that better chromophore packing can lead to increased quantum efficiency of protein-bound fluorophores. Denaturation of GFP resulted in a destruction of its fluorescence.[145,146] Additionally, synthetic analogs of the GFP fluorophore are devoid of fluorescence, demonstrating that the rigidity of the protein structure and tight packing of residues inside the binding cavity are crucial for its high fluorescence quantum yield.[147,148] As a result, filling the space around the bound ligand is a suitable approach to constrain vibrational relaxations in order to reach enhanced fluorescence quantum efficiency. To this end, Trp with its large size (240 $Å^3$

average volume, 163 $Å^3$ van der Waals volume)[149,150] and restricted rotomeric flexibility[151-154] was installed at various positions in the binding cavity. It should be noted that the indole ring of Trp is polarizable and has a significant dipole moment, and thus it could influence the electrostatic environment of the bound ligand rather than having only steric effects. Inspection of the crystal structure of KLE CRABPII mutant suggested four positions, Leu28, Ile31, Ala32, and Ser37, located on α-helix of the protein, to investigate tighter packing via replacement by the voluminous residue Trp (**Figure I-14**). All four mutants formed a PSB upon binding to merocyanine aldehyde **1** (**Table I-1**, entries 13-15, 18). Notably, mutants of KLE:L28W, KLE:I31W, and KLE:S37W gave similar absorption/emission wavelengths with the bound chromophore, whereas KLE:A32W exhibited a more red-shifted spectroscopic profile. Unfortunately, all of these mutants gave similar quantum yields in comparison with KLE triple mutant (**Table I-1**, entry 2), showing that the substitutions of Trp residues at the selected positions did not

Table I-2. Brightness of CRABPII/**1** complexes and known red fluorescent proteins.

Fluorescent Protein	λ^{abs} / λ^{ems} (nm)	Φ(%)	ε $(M.cm)^{-1}$	Brightness (%)
mKate2	588 / 633	40	62,500	100
mCardinal2	604 / 659	19	87,000	66
KL	600 / 617	18	111,700	80
KLE	591 / 612	33	93,200	123
KLD	591 / 612	30	97,500	117
KL:R59W	602 / 619	23	140,100	129
KLE:R59W	595 / 616	39	169,800	265
KLE:R59W:L28W	595 / 619	38	158,400	241
KLE:R59W:M123W	595 / 611	29	167,800	194
KLE:R59W:A32W	595 / 617	39	142,900	223

Φ = quantum yield, ε = extinction coefficient of the protein/**1** complex, brightness is calculated as the product of ε and Φ.

result in an effective restricted environment for the bound ligand. Next, we constructed mutants in which doubly installed Trp residues around the helical region were examined. Since the R59W mutation caused an enhancement of the quantum yield, we combined R59W along with the aforementioned mutants (**Table I-1**, entries 21, 22 and 24). Sequential combinations of the R59W mutation with L28W, I31W and A32W gave quantum efficiencies identical to KLE:R59W, suggesting that the Trp residue at R59 is sufficient to provide a tighter fit for the indoline ring. Nevertheless, KLE mutants including the R59W mutation exhibit large extinction coefficients and high quantum yields with the brightness of the CRABPII complexes rivaling mKate2[13] and mCardinal2[8], two proteins with the brightest fluorescences known in this wavelength regime (**Table I-2**).

Since attempts to put steric bulk on the helical region of the protein did not result in enhanced fluorescence, we selected other positions closer to the interior of the protein and proximal around the polyene tail of the chromophore (**Figure I-14**). Pro39, Val76, and Met123 were exchanged with Trp residues to increase the steric bulk around the bound ligand. Unfortunately, KLE:P39W mutant was sluggish to yield the PSB after 24 h. We also observed this effect when Phe15 was mutated into a Trp residue (data is not shown). Presumably, placement of Trp residues at these positions blocks ligand entry, preventing the aldehyde moiety from reaching the active site Lys residue. Fortuitously however, in the case of KLE:V76W, we observed that merocyanine aldehyde reacts rapidly with the nucleophilic Lys side chain with a nearly instantaneous formation of the PSB (see further discussion in Section I.6). It was interesting that the

resulting KLE:V76W/**1** complex showed blue-shifted absorption and emission spectra as compared to other CRABPII/**1** pigments (**Table I-1**, entry 12), and also as compared to merocyanine-PSB **2** with *n*-butyl amine in PBS solution (**Figure I-11c**). A blue shift was also seen in other V76W containing mutants such as KLE:V76W:S37W and KLE:V76W:L28W in which both parent pigments KLE:S37W and KLE:L28W display a 10 nm red-shifted spectrum before the installation of V76W (**Table I-1**, entries 13 vs 19, and 15 vs 20). Another surprising observation was that these mutants have low extinction coefficients upon coupling with the merocyanine aldehyde **1**. The increased rate of PSB formation, along with an observed blue shift of absorption and low extinction coefficient for all V76W mutants may suggest that these mutants are structurally more solvent accessible (ligand is buffered in aqueous media) and flexible to accommodate the exogenously added chromophore. The substitution of Trp at Met123 (KLE:M123W) did not cause any significant change in the spectroscopic profile of the protein-ligand complex as compared to KLE mutant. Indeed, the addition of M123W mutation to KLE:R59W resulted in decreased fluorescence (Table I-1, entry 23).

I.5 Crystallographic Data Of CRABPII/1 Fluorescent Complexes

In the course of this study, it became evident that single mutations of L121E, L121D, and R59W were most effective in improving the fluorescent characteristics of KL double mutant. In an effort to understand the effect of these mutations, structures of several merocyanine bound CRABPII mutants were determined. The readily crystalizable nature of these proteins was invaluable to provide clues for the observed spectroscopic behavior that results from single mutation of key residues. As will be

evident upon review of the data below, our studies indicate that the large binding cavity and relative structural plasticity of CRABPII provides structural freedom for the bound ligand to adopt various binding conformations that correlates well with the expected fluorescent properties.

The crystals of KLE:R59W/**1** complex and the structured refinements were performed by Dr. Xiaofei Jia (Yale University). The crystals of KL/**1**, KLE/**1** and KLD/**1** complexes were determined by Meisam Nosrati and Ipek Yapici.

I.5.1 The Structure Of The KL/1 Complex

Our previous efforts in designing rhodopsin protein mimics were aided through the acquisition of numerous crystal structures of CRABPII and hCRBPII bound retinoid-protein complexes.[117-121,123,129] Although rotameric differences were observed as a result of different mutations, the bound chromophore was always found in the same location within the binding pocket. Structures of merocyanine bound CRABPII complexes defied this trend. In the KL/**1** complex, the ligand follows a trajectory that is completely distinct from previously seen ligand bound structures of CRABPII (**Figure I-15**). The overlay of KL/**1** structure with its retinal bound counterpart (KLE/retinal) shows that merocyanine ligand does not follow the same trajectory as compared to retinal chromophore. With retinal, the ionone ring is located near the mouth of the binding cavity, which is in contrast to the indoline ring of the merocyanine. The indoline ring is tucked into a hydrophobic pocket deep within, featuring significant hydrophobic interactions between Leu121 and the heterocyclic ring (**Figure I-16b**). The merocyanine aldehyde **1** forms a *cis*-imine with Lys132, similar to retinal, but the resulting PSB is stabilized by a water-

mediated hydrogen bond to Ser12 instead of a direct hydrogen bond to a counter anion (retinal iminium is stabilized by a Glu121 counter anion, see **Figure I-15**).

Close inspection of a bound ligand revealed that cyanine dye is severely twisted about the C3-C4 single bond (dihedral angle ψ = 59.1°), which adopts an s-*cis*

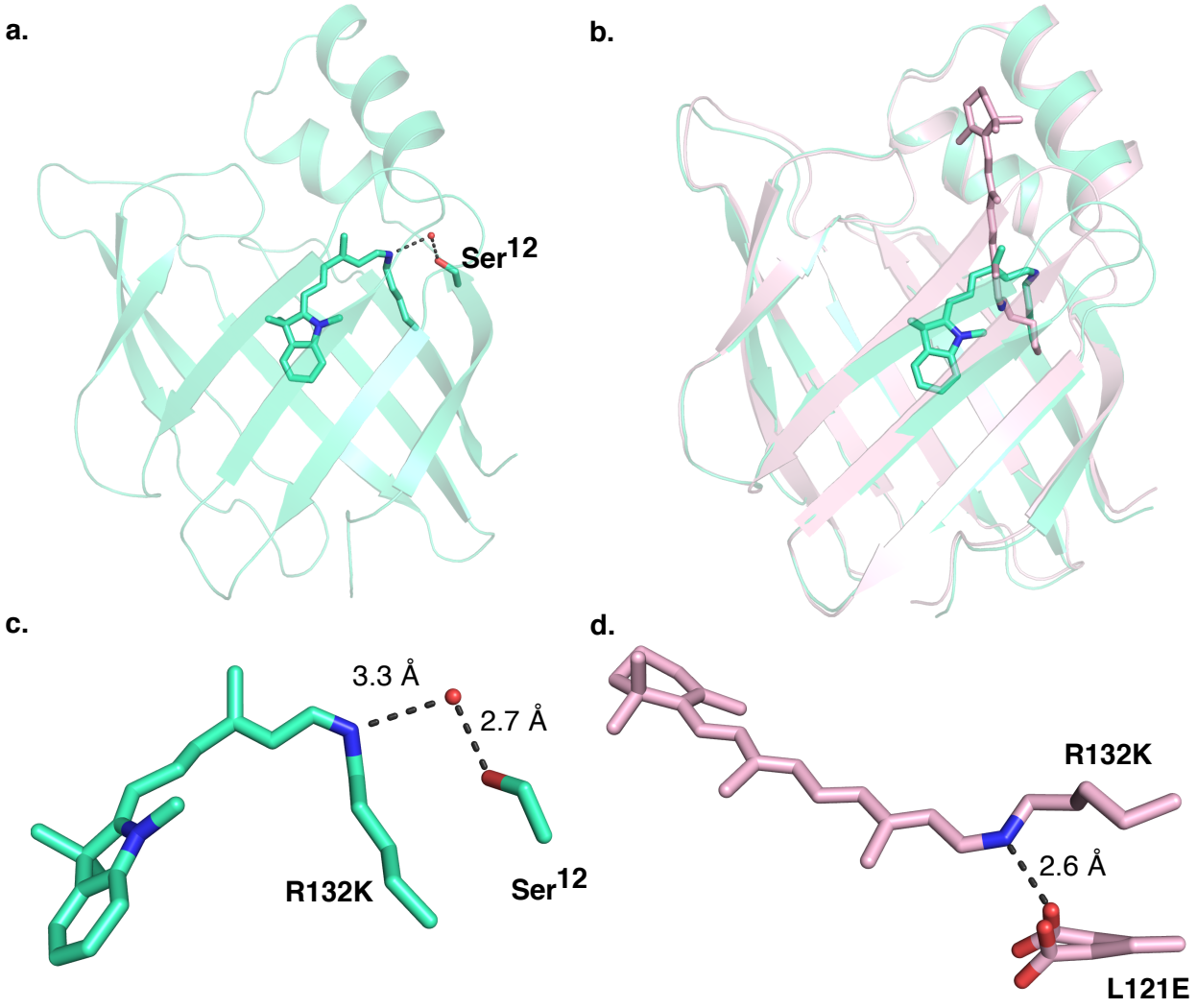


Figure I-15: Crystal structure of KL/**1** complex (1.7 Å resolution, PDB ID: 4QGV). **a.** The geometry of the bound ligand (in stick model) is shown. **b.** The overlay of crystal structures, KL/**1** complex (green) and KLE/retinal (pink, PDB ID: 2G7B) shows the distinct trajectory for the bound chromophores. **c.** The enlarged image of KL/**1** binding pocket depicts the *cis*-imine and the interaction with a water molecule (indicated by the red sphere). **d.** Retinal forms a *cis*-iminium with R132K that is stabilized by a counter anion, Glu121 residue. The crystallographic data shows that the side chain of Glu121 adopts two conformations within the binding cavity.

conformation to accommodate the binding site (**Figure I-16**). As previously reported, a twisted polyene backbone of the cyanine dyes provides an effective conduit for non-radiative decay of the excited state, yielding a loss in their fluorescence quantum efficiency.[130,155-157] This happens due to the ability of cyanine dyes to dissipate the absorbed energy by rotational freedom of their polymethine chain to adopt twisted conformation that relaxes non-radiatively to the ground state. Nevertheless, this deactivation pathway is impeded in viscous solvents or conformationally restrained environment, leading to an increase in the radiative route. The twisted cyanine dye in KL/**1** complex indicated the fact that the large binding cavity of CRABPII-KL allows the bound ligand to adopt various geometrical conformations, which can lead to the non-radiative decay of the excited state of the KL/**1** complex. It is therefore not surprising that severely twisted KL/**1** complex has a low quantum yield in comparison with other CRABPII mutants (Φ = 18 %, **Table I-1**).

I.5.2 The Structure Of The KLE/1 And KLD/1 Complexes

A radical change in binding of merocyanine is observed upon altering the electrostatic environment at position 121. Exchange of the hydrophobic Leu121 residue with either Glu or Asp yields protein complexes with merocyanine aldehyde **1** that are more fluorophoric than the KL/**1** complex (**Table I-1**, entries 2 and 3). The structures of KLE/**1** and KLD/**1** complexes show that the chromophore binds the active site Lys residue from a trajectory pointing opposite to that observed in the KL/**1** complex (**Figure I-17**). As **Figure I-16b** illustrates, Leu121 has a tightly packed interaction with the hydrophobic indoline ring of the merocyanine, lying within 3.6 Å away. The electrostatic

change, as a result of L121E and L121D mutations, plausibly disturbs the hydrophobic environment, leading to an alternate binding site, which is more energetically favorable. In the new binding orientation, the heterocyclic ring of the ligand is projected out of the β-barrel between strand 1 and helix 2, leaving it relatively exposed to the exterior of the protein. Both KLE/**1** and KLD/**1** complexes showed that the merocyanine aldehyde **1** forms a *trans*-iminium with an active site Lys residue unlike that seen in KL/**1** *cis*-

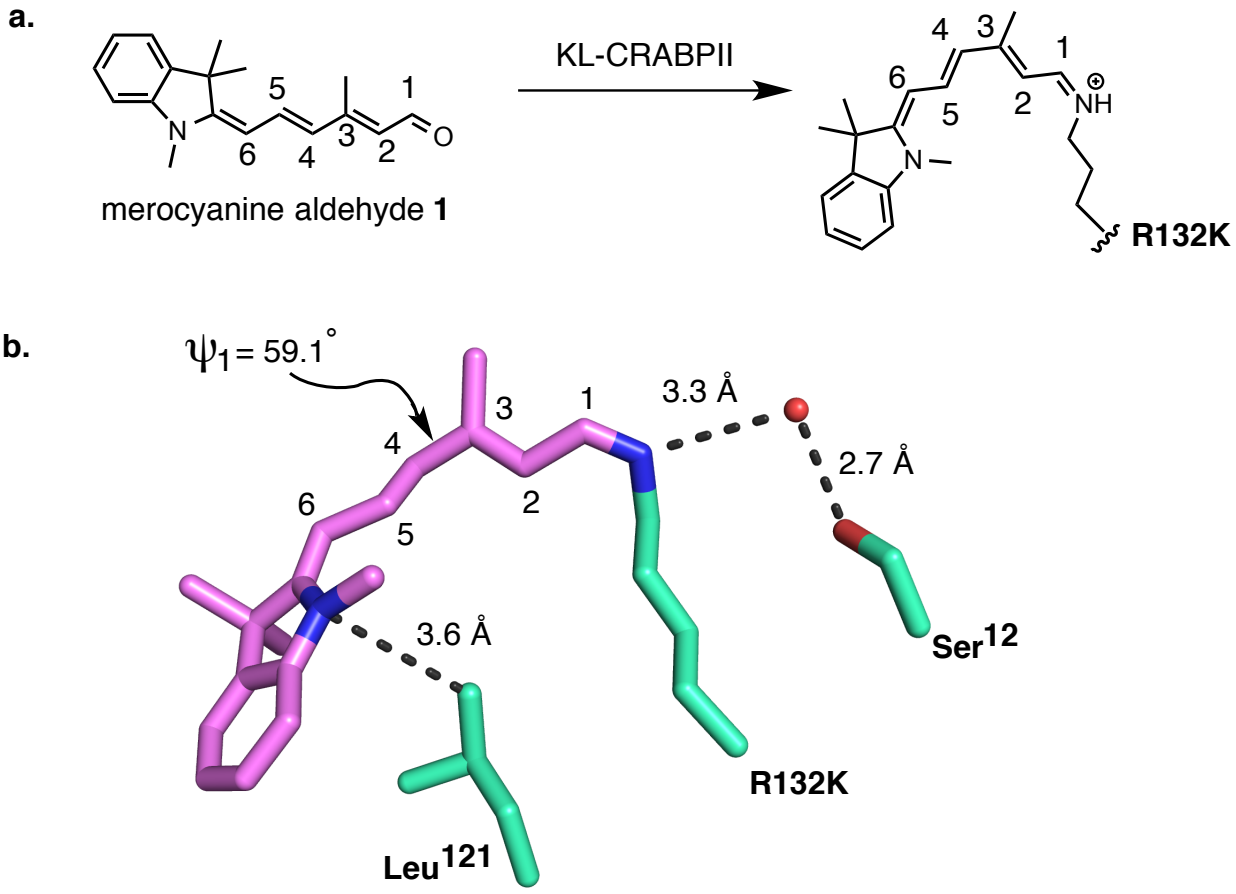


Figure I-16: Twisted conformation of merocyanine PSB in KL-CRABPII complex. **a.** The reaction of merocyanine aldehyde with an active site Lys residue and subsequent binding geometry of the bound ligand. Twist around the C3-C4 single bond results in the s-*cis* conformation. **b.** The exploded view of KL/**1** complex demonstrates the twisted conformation of the bound ligand with a dihedral angle 59.1°. The hydrophobic Leu121 resides 3.6 Å away from the indoline ring, providing hydrophobic packing interactions.

iminium complex. Also, the resulting PSBs are stabilized by direct hydrogen bonding with a side chain of Ser12 rather than a water-mediated interaction (**Figure I-17d** and **Figure I-17e**). Notably, the chromophores in KLE and KLD complexes are more planar in contrast to the KL/**1** complex. Although, they adopt the s-*cis* confomation about C3-C4, the dihedral angle is much more planar (ψ_2 = 13.9°, ψ_3 = 7.2° for KLE/**1** and KLD/**1**

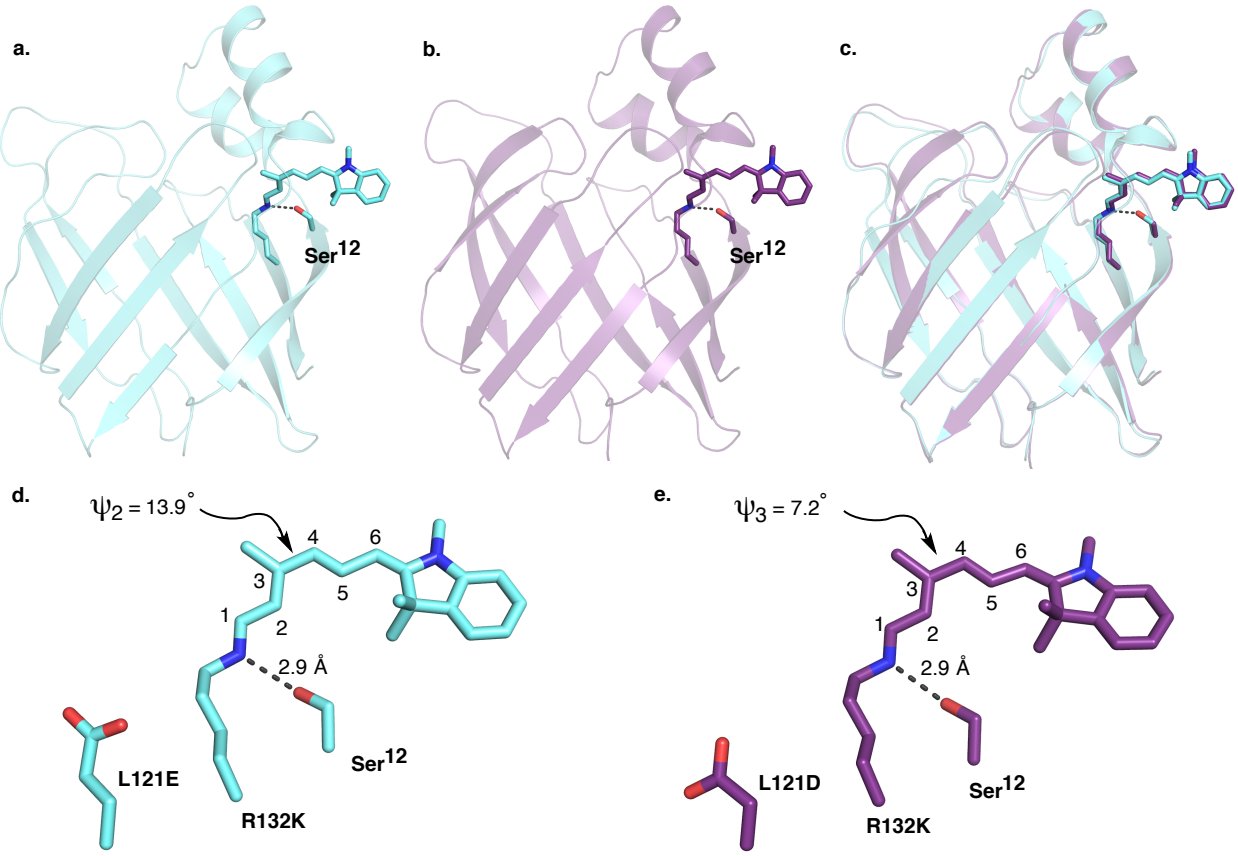


Figure I-17: Crystallographic data of KLE/**1** and KLD/**1** complexes. **a.** Crystal structure of KLE/**1** complex (1.5 Å resolution, PDB ID: 4QGX). **b.** Crystal structure of KLD/**1** complex (1.8 Å resolution, PDB ID: 4QGW). **c.** Overlaid structures of KLE/**1** (cyan) and KLD/**1** (dark purple) complexes show the identical trajectory of the bound ligand. **d.** The enlarged image of the bound chromophore in CRABPII-KLE displays the twist about the C3-C4 single bond in the s-*cis* conformation. Ser12 plays a role in stabilizing the iminium nitrogen via direct hydrogen bonding. **e.** Same illustration of **d** for the KLD/**1** complex.

complexes, respectively). As anticipated, the more planar KLE and KLD mutants complexed with merocyanine aldehyde **1** have higher quantum yields than the severely twisted KL/**1** complex (**Table I-1**, entries 2 and 3 vs 1).

I.5.3 The Structure Of The KLE:R59W/1 Complex

The structure of the KLE:R59W/**1** complex depicts the binding of the ligand in a fashion similar to those previously observed with retinal and retinoic acid bound CRABPII mutants. Mutation of both Arg59 to Trp and Leu121 to Glu results in ligand binding mode in which the indoline ring points towards the mouth of the binding cavity and the aldehyde end of the ligand is deeply buried in the interior of the protein. The Schiff base between the active site Lys residue, R132K, and the merocyanine aldehyde adopts a *trans*-imine geometry. The installed counter anion, Glu121, interacts with the iminium via a water molecule that resides 3.6 Å away from the nitrogen atom, yielding a complex that exhibits a pK_a of 10.5 for the PSB (**Figure I-18**). The conformation of the chromophore is quite planar as the entire polymethine chain stays in an s-*trans* conformation. Additionally, the chromophore is well packed within the binding cavity with restricted ability to move, in contrast to the previous structures, where it either severely deviates from planarity (in KL/**1** complex), or is less well packed and more solvent exposed (as in KLE/**1** and KLD/**1** complexes).

KLE:R59W mutant with merocyanine aldehyde **1** has superior fluorescent characteristics as compared to other mutants investigated in our study. It is noteworthy that the KLD:R59W mutant (replacement of Glu121 with Asp121) that results in similar fluorophoric properties as KLE:R59W/1 complex, thus suggesting that the shorter Asp

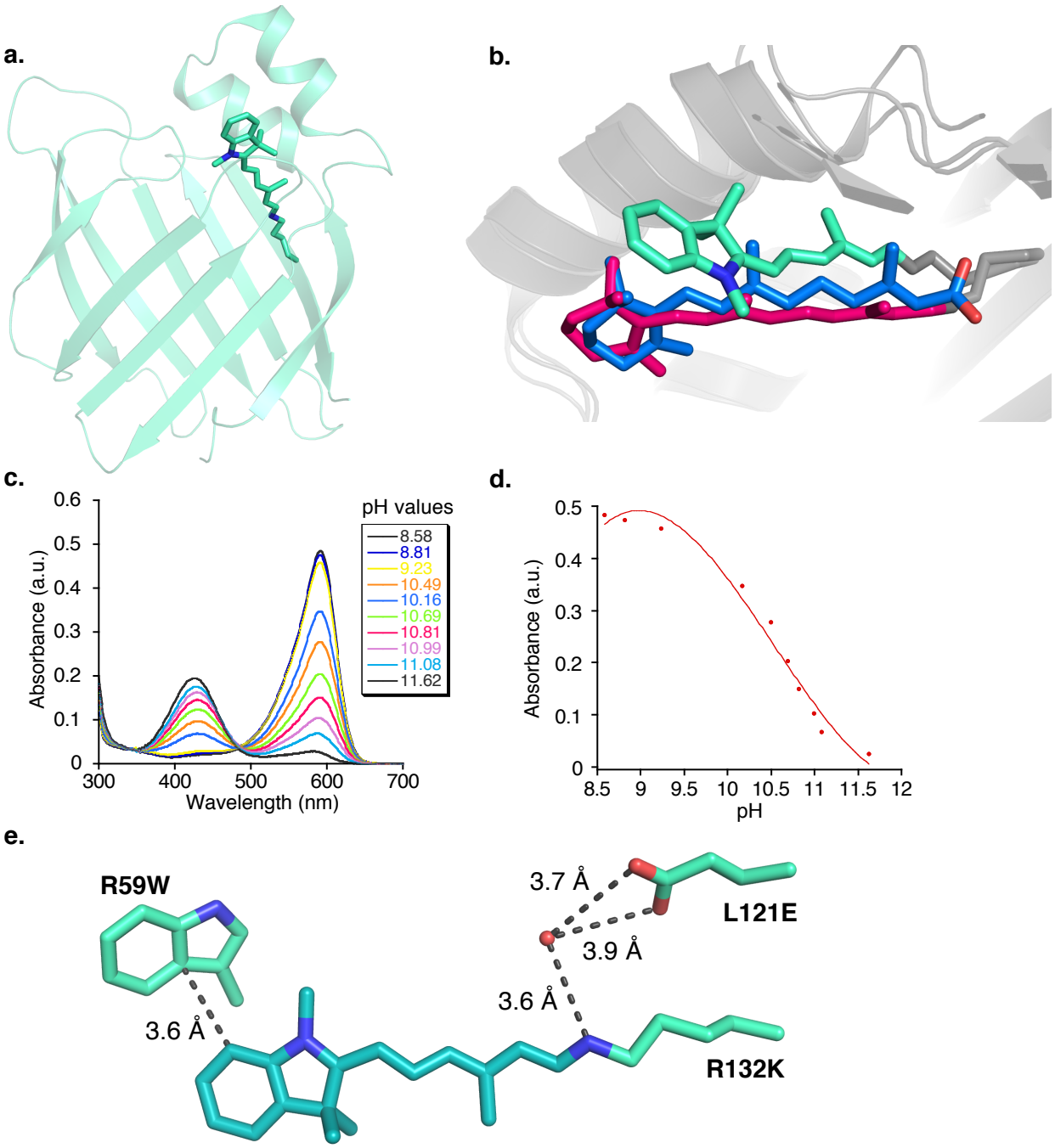


Figure I-18: KLE:R59W CRABPII mutant complexed with merocyanine aldehyde **1**. **a.** Crystal structure of KLE:R59W/**1** complex (2.7 Å resolution, PDB ID: 3FEB). The bound ligand is shown in stick. **b.** Overlay of all three crystal structures, KLE/retinal – pink (PDB ID: 2G7B), wild type CRABPII/retinoic acid – blue (PDB ID: 2FR3), KLE:R59W/**1** – green, illustrates similar binding mode for each chromophore. **c.** Base titration of KLE:R59W/**1** complex. The pH values are indicated in the box on right. **d.** Absorbance of KLE:R59W/**1** complex at 595 nm as a function of pH with an apparent pK_a of 10.5. **e.** The exploded view of the KLE:R59W/**1** complex with L121E and R59W mutations.

side chain is of sufficient length to maintain the interaction with the bound chromophore without causing a significant change in ligand geometry (**Table I-1**, entry 6). The fluorescent properties of these mutants could be the result of the following three major factors; i) the relatively flat and well-packed conformation leads to maximal overlap of the π system, ii) closure of the binding cavity with a large Trp residue further isolates the binding pocket and could also exclude water molecules that otherwise might weaken fluorescence through enhancing non-radiative decay of the excited state, iii) although the crystal structure precludes π-π stacking between Trp59 and the indoline ring of the bound ligand, its close proximity (3.6 Å) could result in a favorable electrostatic interaction that further rigidifies the cyanine chromophore.

Given all the structural information obtained, it would appear that mutation of both Leu121 to Asp or Glu, and Arg59 to Trp, is required to achieve the more orthodox ligand trajectory seen in KLE:R59W/**1**. As predicted from the crystal structures of KLE:R59W/**1**, a favorable hydrophobic packing between Trp59 and the heterocyclic ring of the ligand is gained with the R59W mutation. All in all, it seems that a combination of two conditions is necessary with merocyanine as the ligand to adopt the planar and linear geometry, “retinylidene-like binding trajectory”. First, an anionic residue is required at position 121, which destabilizes the ligand conformation seen in KL, and provides a water-mediated interaction with the iminium in KLE:R59W. Second, placement of the Trp residue at position 59, which supports hydrophobic interactions with the merocyanine ring, leads to the relatively well-packed, flat ligand conformation seen in KLE:R59W/**1** structure. As a result, these changes yielded the highest extinction

coefficient and quantum efficiency, since the chromophore is flattened and conformationally restricted. Indeed, the mutants with the best fluorescent characteristics invariably contain the mutations at both of these positions (**Table I-1**).

I.6 Kinetic Analyses Of PSB Formation Of CRABPII/1 Fluorescent Complexes

We proceeded to analyze the coupling efficiency of merocyanine aldehyde **1** with CRABPII variants. Mutations that introduced bulky residues within the binding pocket, as well as the altered binding modes of the ligand for some mutants could adversely affect the kinetics of PSB formation. We were able to monitor PSB formation by UV-vis spectroscopy since the absorption of the PSB (absorbing higher than 570 nm) is distinct from the free merocyanine aldehyde **1** (λ^{abs} = 492 nm). All CRABPII variants (20 M) were coupled with 0.3 equivalents of merocyanine aldehyde **1** at 23 $^{\circ}$C and the formation of PSB was followed at the λ^{max} of each CRABPII/**1** complex. **Table I-3** depicts the rate of PSB formation of CRABPII mutants relative to KL double mutant. The presence of Trp at position 59 decreased the rate of PSB formation for the corresponding parent KL mutant (**Table I-3**, entry 2). However, installation of L121D or L121E restores the original binding trajectory of the chromophore and accelerates formation of the PSB. Notably, the kinetics of PSB formation of merocyanine aldehyde **1** with KLE:R59W mutant is about twice faster than the KL mutant (relative rate 1.8 vs 1). The change of Glu121 into Asp121 hampers PSB formation and results in similar kinetics observed in with the KL mutant (**Table I-3**, entries 5 vs 6). The observed rate difference between Asp and Glu residues implies that the installed residues (Asp and Glu) at position 121 might participate in acid catalysis in order to activate the aldehydic

moiety for imine formation. Plausibly, the closer proximity of Glu side chain facilitates the imine formation, causing a rate enhancement more effective than the Asp residue, which perhaps stays far from the bound aldehyde.

Table I-3. Rate of complex formation for a variety of CRABPII mutants.

Entry	CRABPII Variants	Relative Rate
1	KL	1.000
2	KL:R59W	0.389
3	KLE	0.950
4	KLD	0.243
5	KLD:R59W	0.955
6	KLE:R59W	1.809
7	KLE:V76W	3.550
8	KLE:S37W	0.431
9	KLE:L28W	0.376
10	KLE:A32W	0.428
11	KLE:V76W:L28W	1.394
12	KLE:R59W:L28W	1.160
13	KLE:R59W:A32W	0.779

Most Trp mutations along the chromophore slowed down imine formation, thus extending the time for full binding and complex maturation. This is presumably due to the steric hindrance imposed by the bulky Trp residues that could interfere with the ligand entry, or reduce accessibility to the active site Lys residue. An exception was the incorporation of Trp at position 76, which led to an increase in the rate of PSB formation (**Table I-3**, entries 7 and 11). CRABPII variants in combination with V76W mutation showed remarkably faster reaction rates, yielding completely matured complexes within few minutes at room temperature. To get more insights into the counter intuitive observations with the V76W mutant (such as rate enhancement, blue shift in absorption/emission spectra, and low extinction coefficients), we pursued crystallization of the KLE:V76W mutant. Unfortunately, our attempts for crystallization of this mutant were unsuccessful so little could be said about the subtle effects of Trp76 substitution in

the absence of structural information. However, the increased rate of PSB formation, along with observed blue-shifted spectroscopic characteristics for all V76W mutants may suggest that these mutants are structurally more flexible and solvent accessible.

To better benchmark the current system with existing methods, detailed kinetic analysis was conducted to calculate the rate constant for CRABPII/**1** complex formation. Measurements of the rate constant and the order of the reaction were followed by previously established protocols.[67,158] KLE:R59W (100 nM) was reacted with an excess amount of merocyanine aldehyde **1** in PBS solution at 37 °C at different concentrations (1 to 5 μM), and the increase of fluorescence intensity was monitored for each reaction (**Figure I-19a**). The second-order rate constant for the reaction of merocyanine **1** with KLE:R59W was determined by plotting the pseudo first order rate constants against merocyanine **1** concentration, leading to a linear fit (**Figure I-19b**). The resultant slope yields the second-order rate constant (k_2 = 2356 $M^{-1}s^{-1}$) for the complex formation. This value is favorably comparable with other fast labeling methods such as PYP-Tag (k_2 = 3950 $M^{-1}s^{-1}$),[67] BGSBD/SNAP-Tag (k_2 ~ 7200 $M^{-1}s^{-1}$),[158] bioorthogonal reactions for labeling proteins (k_2 ~ 10^{-4} – 10^4 $M^{-1}s^{-1}$),[98] and enzymatic labeling techniques (k_2 ~ 10^3 – 10^6 $M^{-1}s^{-1}$),[33,98] all of which have similar or faster kinetics. Half-life binding measurements ($t_{1/2}$) were determined under stoichiometric conditions (5 μM KLE:R59W and merocyanine aldehyde **1**) in PSB solution at 37 °C. Noteworthy, KLE:R59W exhibits 1a remarkably short time of full complexation with merocyanine **1** and forms a mature

complex within 7 minutes, ($t_{1/2}$ = 39 s, **Figure I-19c**), whereas monitoring the nonfluorescent merocyanine aldehyde **1** in the absence of a protein did not result in any increase in fluorescence under the same experimental condition.

I.7 Persistence Of Fluorescence Over A Broad PH Range

It is well known that cyanine dye variants can be modified to confer pH responsive properties. When the indole nitrogen atoms (one or both) on a cyanine dye are not alkylated, the dye exhibits sensitivity to pH changes.[140-142] In this type of non-N-alkylated indolium structure, acidic environments cause protonation of the indoline nitrogen atoms and preserve the highly fluorescent state of the dye with a long-wavelength absorption band. However, these dyes become almost nonfluorescent under elevated pH conditions with a newly emerged blue shifted absorption peak. As a

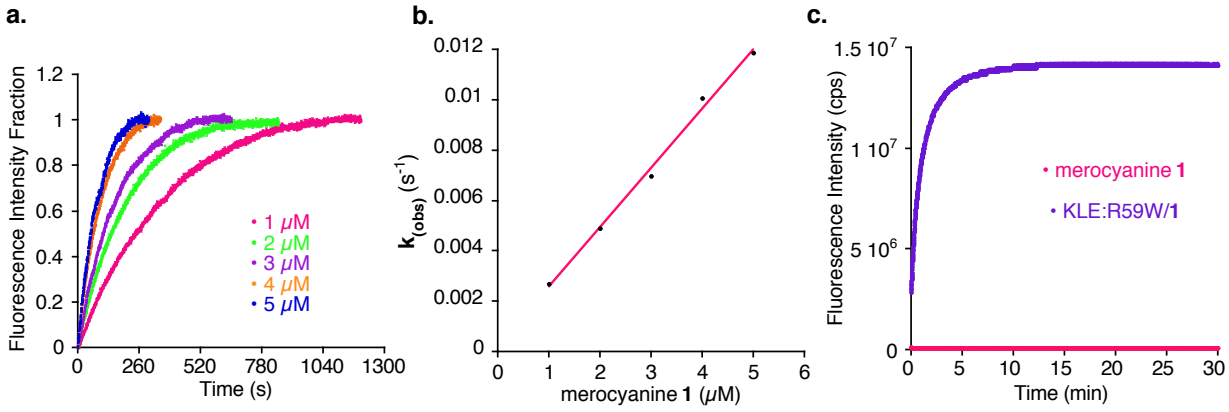


Figure I-19: Kinetic analysis of KLE:R59W complexation with merocyanine **1**. **a.** Time-course binding of KLE:R59W (100 nM) with **1** (1 – 5 μM) in PBS at 37 °C. **b.** k_{obs} values at different merocyanine **1** concentrations show a linear relationship (R^2 = 0.995), yielding a k_2 (second-order rate constant) of 2350 $M^{-1}s^{-1}$. **c.** Stoichiometric binding of KLE:R59W with **1** (5 μM each) in PBS at 37 °C was followed at 616 nm with excitation at 565 nm over 30 minutes. Merocyanine **1** does not exhibit any fluorescence, as illustrated in the plot.

result, depending on the protonation and deprotonation of the cyanine dye, fluorescence can be reversibly tuned between a nonemissive state and an emissive state.

Since the fluorescence of the protein-bound merocyanine PSB depends on the

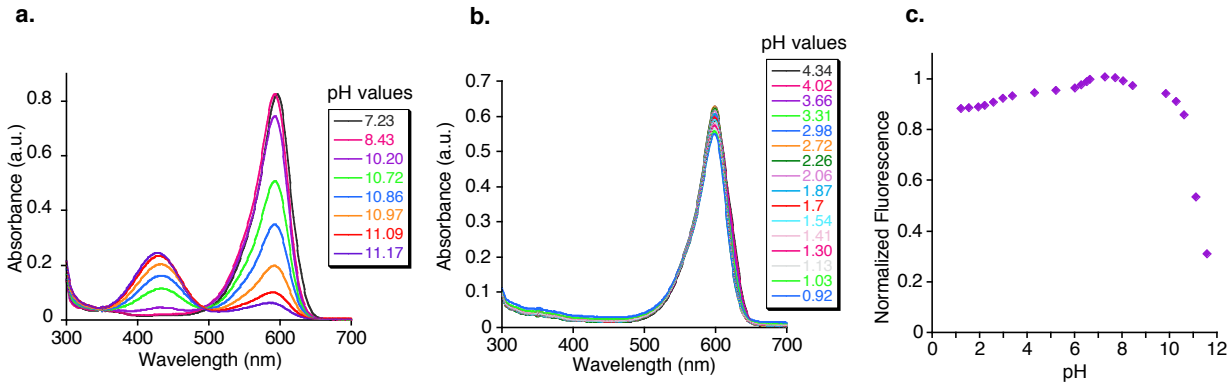


Figure I-20: Measurements of KLE:R59W:L28W/**1** complex in different pH regimes. **a.** Base titration of KLE:R59W:L28W/**1** complex with addition of NaOH in PBS, pH values are indicated in the box. **b.** UV-vis spectra of the KLE:R59W:L28W/**1** complex monitored after each addition of HCl. **c.** Total fluorescence of the complex as function of pH.

protonation state of the chromophore, we investigated the pH dependence of fluorescence for this system. **Figure I-20** depicts the UV-vis spectra of KLE:R59W:L28W/**1** complex titrated with NaOH. Basification of the solution does not deprotonate the PSB until pH ~10; further addition of base gives rise to the absorption at ~ 425 nm (Schiff base absorption). During this process negligible fluorescence is lost up to pH of 10.2. The absorption spectra reveal no change, indicating that the complex maintains its tertiary structure even under highly basic condition (pH as high as 11). As a control experiment, we conducted denaturation experiments with the addition of detergent into CRABPII/**1** complexes. We chose a blue-shifted mutant (KLE:V76W:L28W/**1**, λ^{abs} = 571 nm) and also a red-shifted mutant

(KLE:R59W:L28W/**1**, λ^{abs} = 595 nm) in order to analyzed the spectroscopic properties of the denatured analogs (**Figure I-21**). Notably, the resultant denatured complexes show an absorption band at 583 nm, resembling that of the iminium form of the merocyanine aldehyde **1** with bovine serum albumin, BSA (see further discussion in Section **I-8**). However, we did not observe this absorption band during our pH titrations, supporting that the native fold of the complexes are preserved over a wide range of pH regime.

Next, we investigated the stability of the CRABPII-merocyanine **1** complex under highly acidic conditions. Lowering the pH to 2.2 resulted in a small increase in absorption of the protein-chromophore complex with no change in λ^{max}. The slight change in absorption could be due to protonation of residues close to the chromophore, affecting protein-chromophore interactions. The increase of the solution acidity affects the emissive state of the protein-chromophore complex slightly (**Figure I-20**). Since CRABPII-merocyanine variants retain most of their fluorescence, they may find application for cellular imaging of acidic intracellular compartments without loss of fluorescence or self-quenching issues.

I.8 The Specificity Of The Merocyanine Aldehyde 1

The selectivity of the merocyanine aldehyde **1** for CRABPII is crucial in order to achieve specific labeling in a complex proteome. Since merocyanine **1** has a reactive

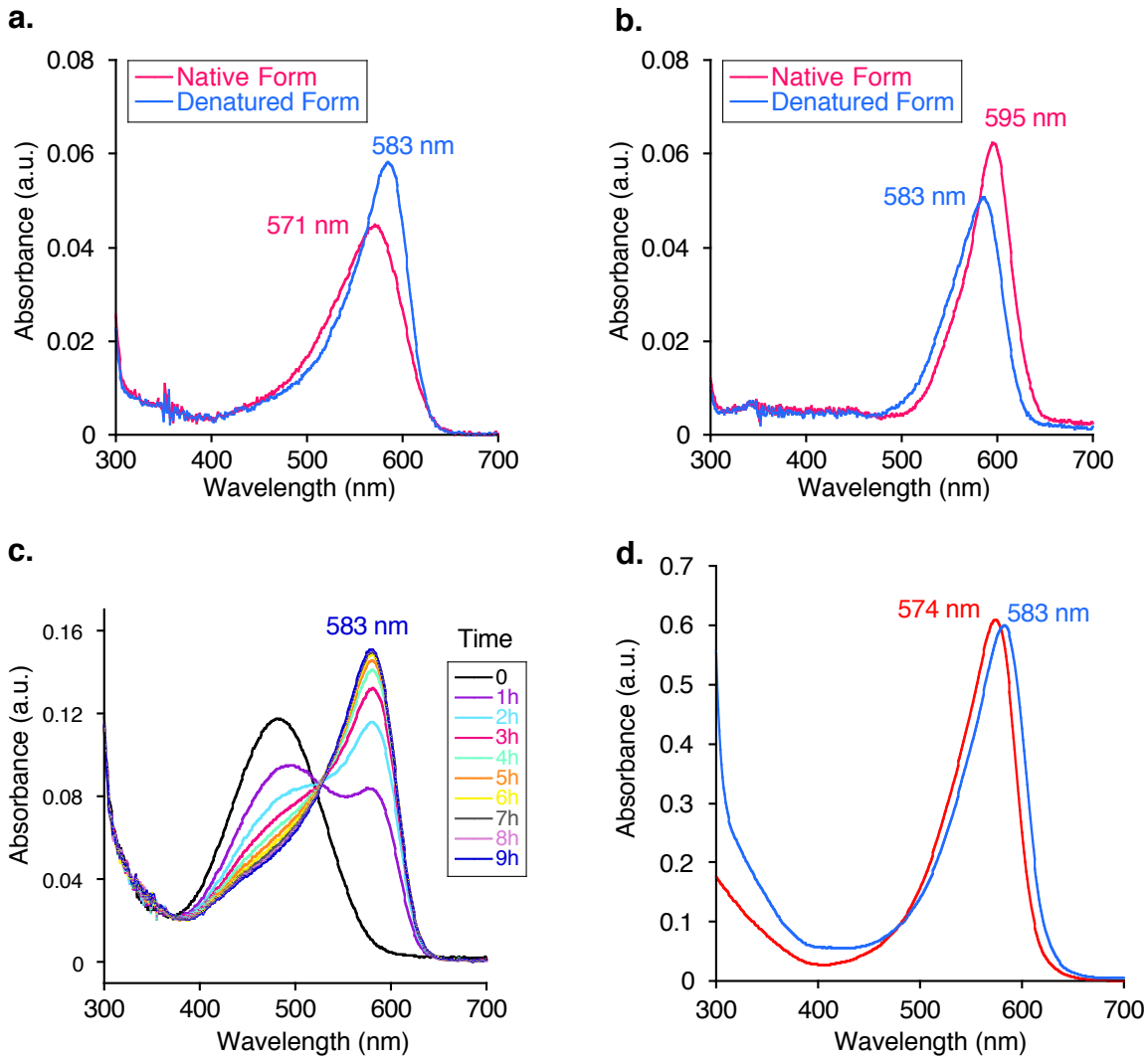


Figure I-21: UV-vis spectra of CRABPII/**1** pigments before and after denaturation with SDS, demonstrating the effect of protein environment on the absorption of the bound merocyanine ligand. **a.** KLE:V76W:L28W mutant complexed with **1**. The native form (pink) has an absorption maximum at 571 nm while the denaturation yields a peak at 583 nm (blue). **b.** KLE:R59W:L28W mutant complexed with **1**. The native form (pink) has an absorption maximum at 595 nm while the denaturation causes a shift of the peak to 583 nm (blue). **c.** Time-course binding of BSA (20 μM) with 0.3 equivalents of **1** in PBS solution at room temperature. Merocyanine aldehyde **1** reacts with Lys residues of BSA, yielding an absorption peak at 583 nm. The reaction time is indicated in the box. **d.** The absorption of merocyanine-PSB **2** at 574 nm (**Figure I-11a** and **Figure I-11b**) shifts in the presence of BSA to 583 nm. The encapsulation of merocyanine-PSB **2** in binding sites of BSA results in a slight red shift in the absorption maximum of the chromophore, yielding an identical λ^{max} that is observed in **c**.

aldehydic moiety, it could react nonspecifically with off-target lysine residues of endogenous proteins in the cell media. To test the reactivity of merocyanine **1** with other lysine residues, BSA (lysine rich protein) was incubated with the merocyanine aldehyde **1** in PBS solution and the proceeding reaction was monitored by UV-vis spectroscopy. **Figure I-21c** depicts that merocyanine ligand forms a PSB with BSA over nine hours at room temperature, yielding an absorption peak at 583 nm. However, the resultant BSA/**1** complex displays much lower extinction coefficient in comparison to KLE:R59W/**1** complex under the same experimental conditions (**Figure I-22b**).

Next, we measured the rate of PSB formation between the merocyanine aldehyde **1** and BSA. Stoichiometric amounts of merocyanine **1** and BSA (5 μM each) were mixed in PBS and formation of the PSB was monitored over time by increase of fluorescence. Also, the time required for complete PSB formation of BSA/**1** complex was about two hours whereas KLE:R59W CRABPII mutant showed remarkably fast kinetics, achieving a fully matured complex within 7 minutes (**Figure I-22c** and **Figure I-19c**). Plausibly, this rapid coupling of merocyanine probe with CRABPII would lead to minimal nonspecific binding of aldehyde **1** to off-target endogenous proteins in cells. The selectivity of merocyanine **1** was further tested by using a higher excess of BSA to better mimic competing binding of endogenous proteins in cells. First, the merocyanine ligand was reacted with 10-fold excess of BSA in PBS solution and the kinetics of PSB formation was followed by fluorescence spectroscopy. Then, the competition experiment was carried out in which excess amount of BSA was mixed with the KLE:R59W mutant (the ratio of BSA to KLE:R59W was 10:1) and PSB formation in the resulting mixture of

protein solution was monitored upon ligand addition (the ratio of ligand to CRABPII mutant was 1:1). As anticipated, the rate of the PSB formation was accelerated in the

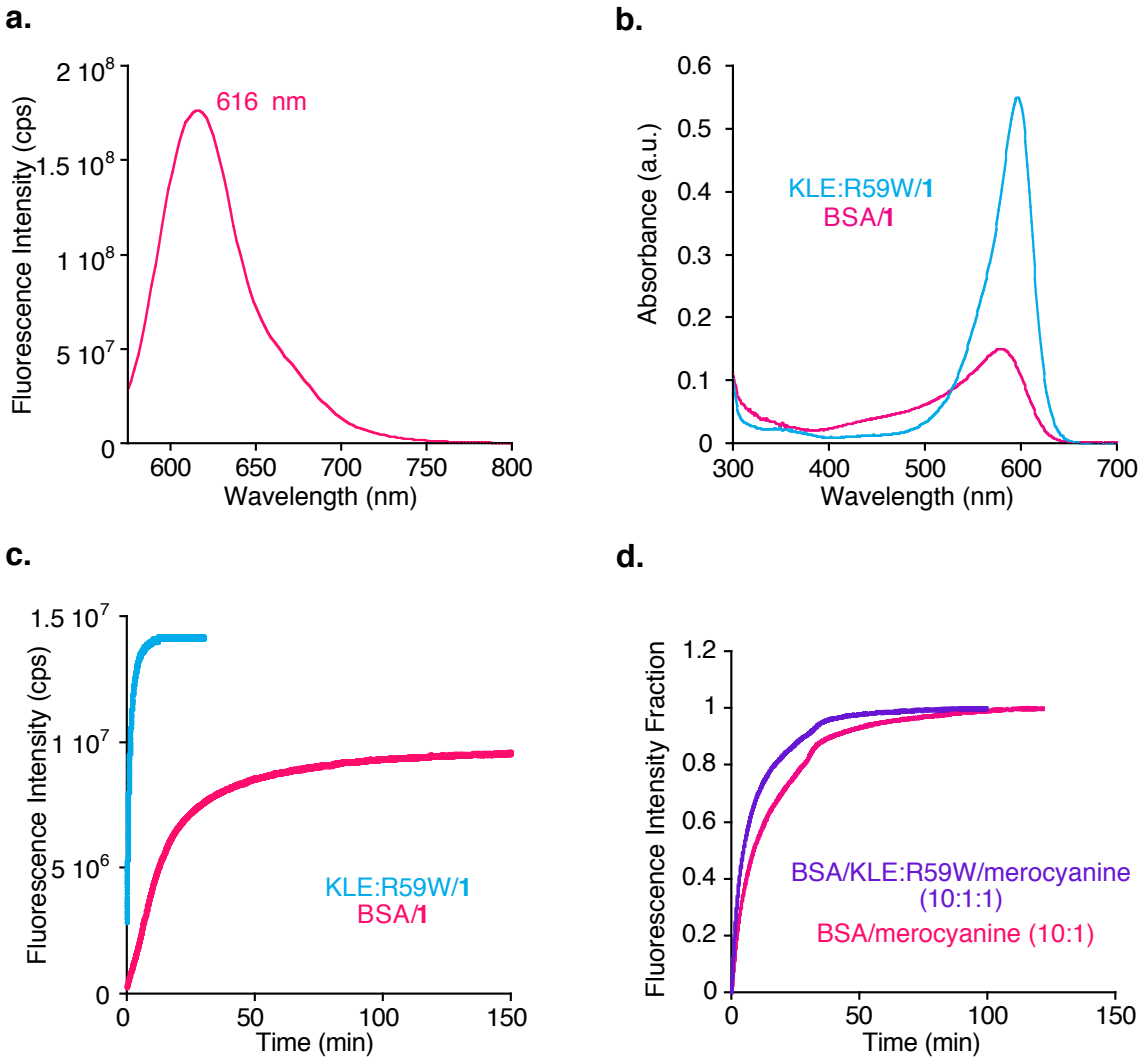


Figure I-22: Analysis of merocyanine aldehyde **1** with BSA. **a.** The emission spectrum of BSA/**1** complex with excitation at 565 nm. **b.** The overlaid UV-vis spectra of KLE:R59W/**1** (blue) complex and BSA/**1** (pink) complex. The protein concentrations are 20 μM with a 0.3 equivalents of merocyanine aldehyde **1**. **c.** The kinetics of PSB formation of KLE:R59W mutant and BSA with merocyanine. The protein and the ligand concentrations are 5 μM in each case. The increase in fluorescence at 616 nm is followed with excitation at 565 nm. KLE:R59W mutant rapidly forms PSB within few minutes (blue) whereas BSA requires over 100 minutes to form a fully complete PSB (pink). **d.** The fluorescence intensity fraction before (pink) and after (purple) addition of KLE:R59W mutant into excess amount of BSA. The ratios are indicated in parenthesis for each solution.

presence of KLE:R59W mutant (**Figure I-22d**) but it was slower in comparison with the system where KLE:R59W mutant was mixed with stoichiometric amount of ligand in the absence of the BSA.

I.9 Visualization Of The CRABPII/1 Fluorescent Complexes In Bacteria

In a proof-of-principle experiment to show feasibility for cellular work, CRABPII mutants, KL (**Table I-1**, entry 1) and KLE:R59W:L28W (**Table I-1**, entry 22) expressed in *E. coli*, were incubated with **1** and imaged. The choice of the latter two mutants was based on their differing spectral characteristics. The bacterial cells readily uptake the chromophore within a minute of addition, and instantly yield visibly colored CRABPII-merocyanine adducts (**Figure I-23**). Both CRABPII mutants exhibited intense fluorescence with excitation at 594 nm. The control cells, which were not transformed with the plasmid expressing CRABPII mutants, were devoid of non-specific fluorescence after incubation with merocyanine aldehyde, demonstrating that background fluorescence from the unbound chromophore is minimal (**Figure I-24**). It should be noted that images produced in **Figure I-24** required less than one minute of incubation with the profluorogenic aldehyde **1**.

High photostability is a desirable trait for a variety of fluorescent protein applications. It is known that cyanine dyes are sensitive towards reactive oxygen species and are not known to be particularly photostable.[159] Recently, there has been work on synthesis of new cyanine dyes with improved photophysical characteristics in order to address their weak photostability.[160-162] To determine the photostability of CRABPII/**1** complexes, photobleaching assays in *E. coli* cells were performed and

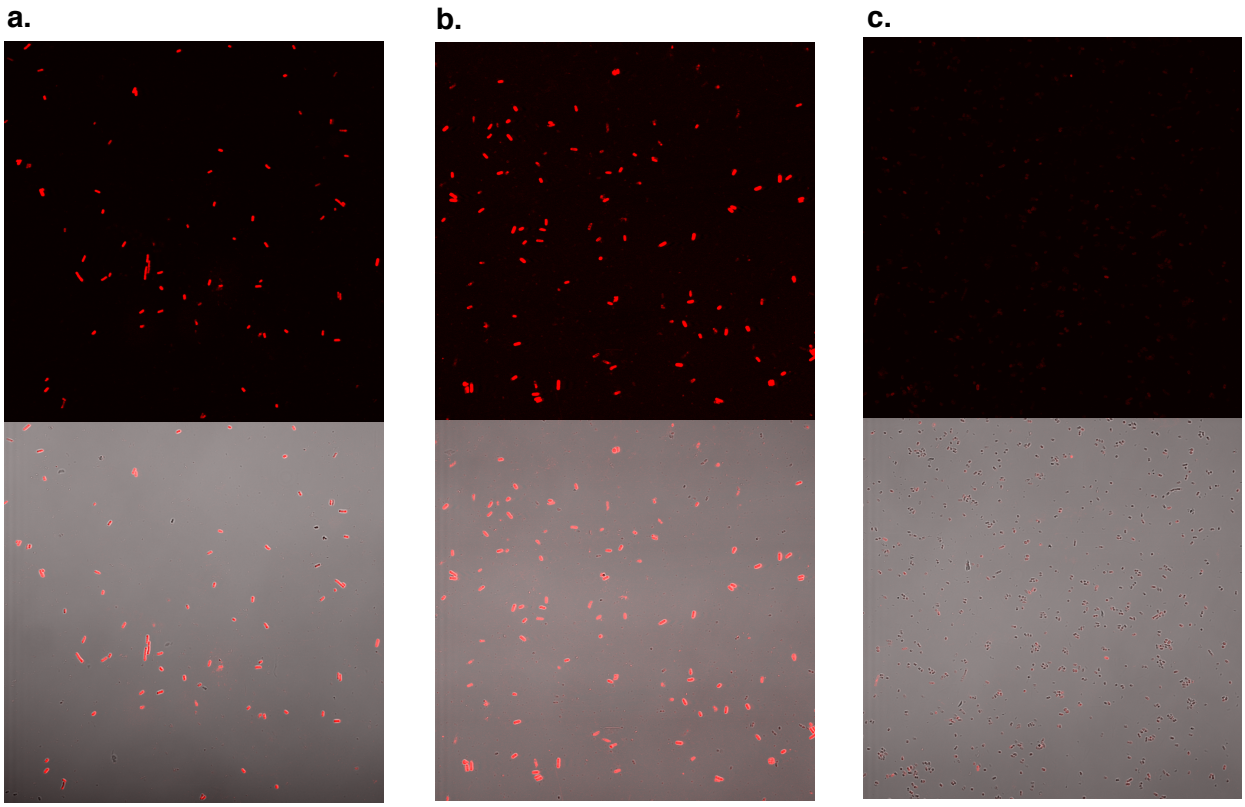


Figure I-24: Fluorescence visualization of *E.coli* cells, transformed with a vector overexpressing CRABPII mutants. **a.** Imaging of KL/**1** complex with 594 nm excitation and 615 nm long pass filter for emission. Bottom shows the brightfield image overlaid with the fluorescence of the complex. **b.** Detected fluorescence from KLE:R59W:L28W/**1** complex with the same settings mentioned in **a**. **c.** Control panel (nontransformed cells) shows negligible fluorescence after addition of merocyanine **1**.

compared to a monomeric red fluorescent protein, mRFP, under identical experimental conditions (**Figure I-25**). The mutants KLE:R59W:L28W and KLE:R59W complexed with merocyanine **1** showed faster photobleaching behavior as compared to mRFP. Both mutants showed about 40% loss of fluorescence over 250 seconds of continuous illumination while mRFP displayed about 20% loss during the same period.

All the imaging and phobleaching experiments were performed by Dr. Tetyana Berbasova. The detailed experimental protocols are provided in Chapter **IV-7.1** and **IV-8**.

a.

b.

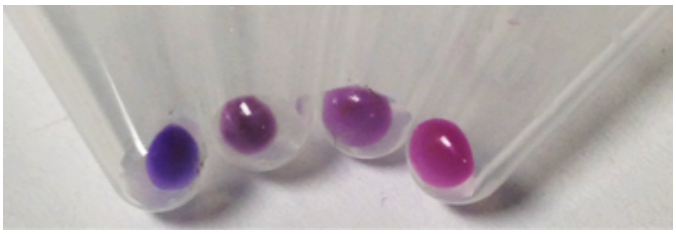


Figure I-23: Bacterial cells expressing CRABPII mutants. **a.** CRABPII mutants of KL, KLE:R59W, KLE:R59W:L28W, and mRFP from left to right, respectively. Picture taken immediately after addition of merocyanine aldehyde **1**, after 1 min centrifugation at 5,000 rpm. No chromophore added to mRFP. **b.** The enlarged image of **a**.

I.10 Visualization Of The CRABPII/1 Fluorescent Complexes In Mammalian Cell

In pursuit of application of CRABPII/**1** system as a fluorescent reporter, CRABPII gene was cloned into several mammalian cell expression vectors. To start with, a construct of EGFP-CRABPII fusion was generated in pEGFP-C2 vector system with the hope that both fluorescent reporters could be visulized simultaneously in transfected cells. As previously reported by Dr. Tetyana Berbasova, the optimized merocyanine aldehyde concentration (25 nM) was used for CRABPII labeling. **Figure II-26a** shows that the green fluorescence of EGPF was observed upon with excitation at 488 nm (505-530 band pass filter) in transfected cells. However, the CRABPII channel did not show any fluorescence at 594 nm excitation (615 long pass filter), resulting in only

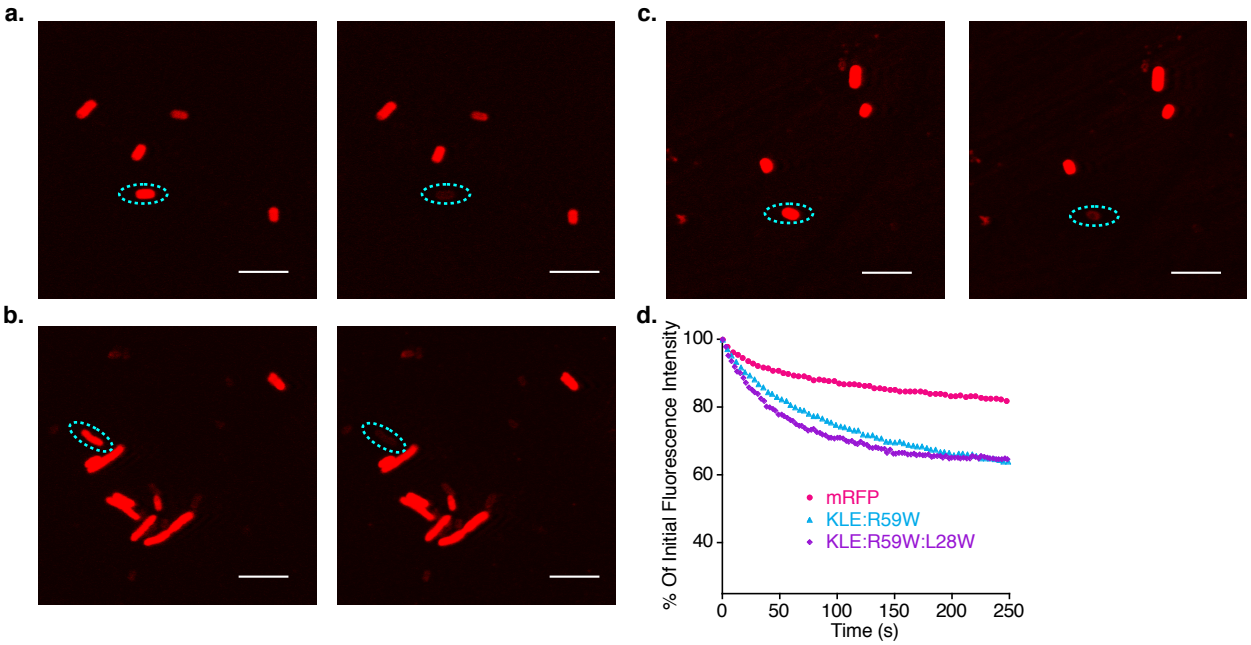


Figure I-25: Photobleaching of KLE:R59W:L28W/**1**, KLE:R59W/**1**, and mRFP with laser light at 594 nm. Images were taken before (left panel) and after (right panel) laser exposure. The dash circles highlight the selected bacterial cells under laser. The scale bar is 5 μm. **a.** Photobleaching of KLE:R59W:L28W/**1** complex formed *in situ*. **b.** Photobleaching of mRFP after its maturation in bacteria. **c.** Photobleaching of KLE:R59W/**1** complex formed *in situ*. **d.** Bleaching curves of each fluorescent protein with a normalized intensity to 100 photons/sec of initial emission.

background fluorescence in Hela cells due to non-specific binding of merocyanine aldehyde **1**. Likewise, testing of the EGFP-CRABPII construct in U2OS cells also fared poorly in specific labeling. It could be possible that CRABPII protein disfavors N-terminal tags, leading to incomplete folding upon expression. Thus, we sought to change the fusion complex by using another vector system in which ECFP is positioned at C-terminus of the CRABPII in fusion complex. Excitation at 458 nm laser produced strong fluorescence of ECFP in the transfected cells, indicating the complete expression of the fused proteins (**Figure II-26b**). Disappointingly, no specific fluorescence was detected at 594 nm excitation, where CRABPII mutant would be expressed and

complexed with merocyanine aldehyde. The possible destabilization of the CRABPII/**1** complex in the chimeric proteins led us to test the individually expressed CRABPII mutants in imaging studies. First, a stop codon was introduced at the end of the CRABPII gene in the previously used ECFP fusion template. As illustrated in **Figure II-26c**, excitation at 458 nm laser led to the absent of ECFP fluorescence, as no expression of ECFP was expected. On the other hand, intense red fluorescence was seen with laser excitation at 594 nm, associated with the off-target labeling of merocyanine aldehyde. The lack of a uniformly distributed fluorescence across the cells indicated that the complex formation of CRABPII/**1** is disfavored and is impractical for mammalian cell imagining. Similarly, trials of individually expressed CRABPII mutants failed to yield fluorescence in many other vectors and cell lines.

Since our previous lab members, Dr. Tetyana Berbasova and Dr. Wenjing Wang, have successfully shown the application of CRBPII/**1** systems in mammalian cells, we did not pursue to rectify the problems of CRABPII/**1** system in eukaryotic cell imaging. All the imaging studies were carried by Dr. Tetyana Berbasova and Elizabeth Santos.

I.11 Conclusion

In summary, we have engineered CRABPII into a fluorescent protein via coupling with a nonfluorescent cyanine dye precursor. CRABPII/**1** complexes demonstrated structural variety in terms of ligand orientation and geometry that correlates well with the observed fluorescent properties. Furthermore, it should be emphasized that engineered CRABPII variants feature a modular design that can be easily adapted to new fluorophores, spanning the entire visible–infrared spectral region. Our probe is

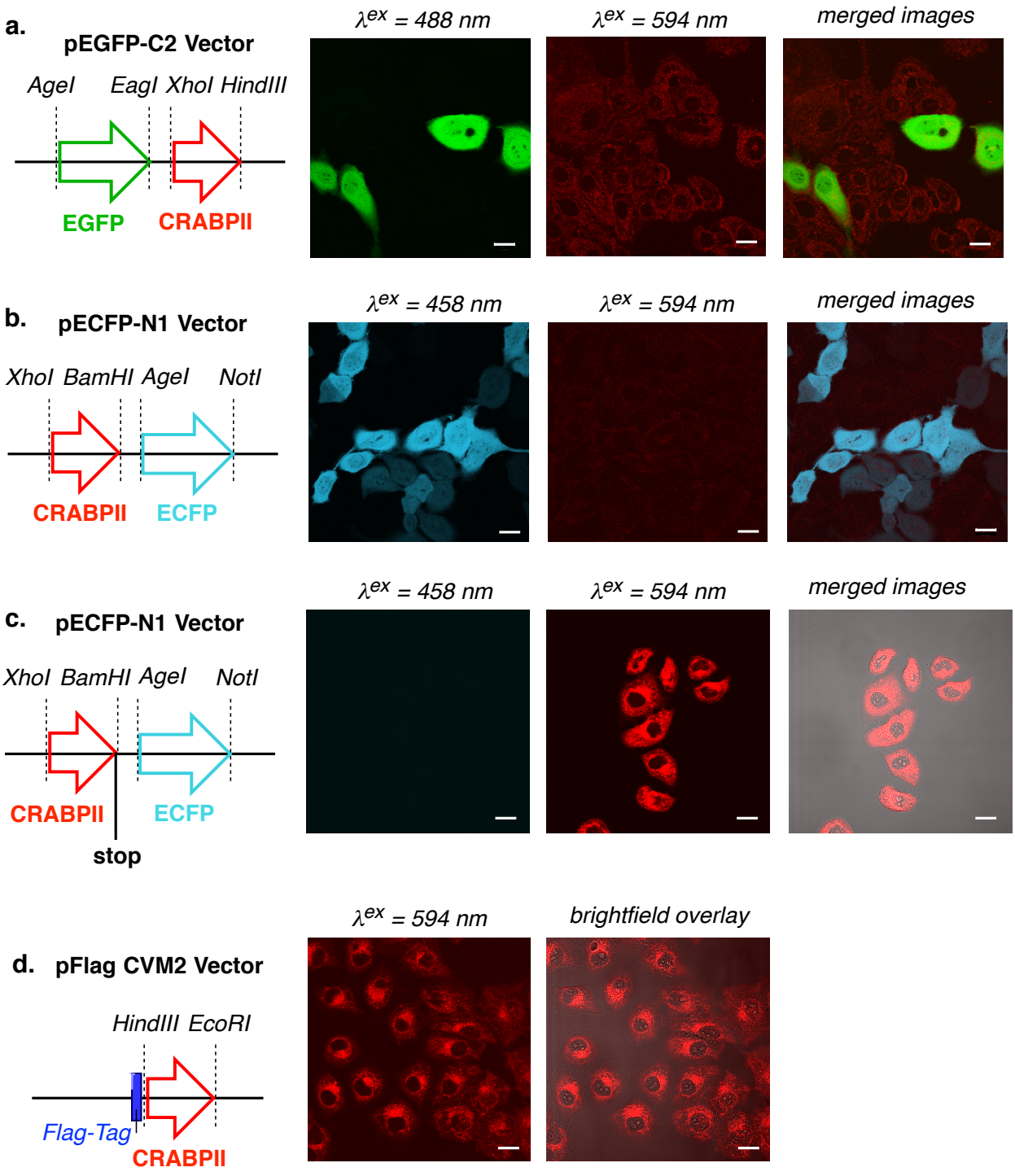


Figure I-26: Utilization of different constructs to visualize CRABPII/**1** fluorescence in Hela cells. **a.** Imaging of the EGFP-KLE:R59W:A32W fusion (PEGFP-C2 vector). **b.** Fluorescence visualization of the KLE:R59W:A32W-ECFP construct (pECFP-N1 vector). **c.** Attempts to image KLE:R59W:A32W (pECFP-N1 vector). **d.** Attempts to image pFlag tagged KLE:R59W CRABPII. All images are pseudocolored. The scale bar is 20 μm for all images.

sufficiently selective to enable live *E. coli* cell imaging. The chromophore is readily cell-permeable and well behaved in live bacteria, coupling with CRABPII variants instantaneously, and fluoresces in bacterial cells within minutes of chromophore addition. In addition, background fluorescence is minimal since the unbound aldehyde form of the chromophore has significantly different spectral properties, as compared to the covalently bound protonated Schiff base form that leads to red-shifted spectra in the protein environment. However, CRABPII does not form a functional complex with merocyanine **1** in eukaryotic cells, limiting the utility of the system in a broad application.

REFERENCES

REFERENCES

1. Morin, J. G.; Hastings, J. W., "Energy Transfer in a Bioluminescnet System" *J Cell Physiol* **1971**, *77*, 313.

2. Shimomura, O.; Johnson, F. H.; Saiga, Y., "Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from Luminous Hydromedusan, Aequorea" *J Cell Compar Physl* **1962**, *59*, 223.

3. Heim, R.; Cubitt, A. B.; Tsien, R. Y., "Improved Green Fluorescence" *Nature* **1995**, *373*, 663.

4. Ormo, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J., "Crystal structure of the Aequorea victoria green fluorescent protein" *Science* **1996**, *273*, 1392.

5. Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J., "Primary Structure of the Aequorea-Victoria Green-Fluorescent Protein" *Gene* **1992**, *111*, 229.

6. Yang, F.; Moss, L. G.; Phillips, G. N., "The molecular structure of green fluorescent protein" *Nat Biotechnol* **1996**, *14*, 1246.

7. Shimomura, O., "Discovery of Green Fluorescent Protein" *Method Biochem Anal* **2006**, *47*, 1.

8. Chu, J.; Haynes, R. D.; Corbel, S. Y.; Li, P. P.; Gonzalez-Gonzalez, E.; Burg, J. S.; Ataie, N. J.; Lam, A. J.; Cranfill, P. J.; Baird, M. A.; Davidson, M. W.; Ng, H. L.; Garcia, K. C.; Contag, C. H.; Shen, K.; Blau, H. M.; Lin, M. Z., "Non-invasive intravital imaging of cellular differentiation with a bright red-excitable fluorescent protein" *Nat Methods* **2014**, *11*, 572.

9. Kremers, G. J.; Gilbert, S. G.; Cranfill, P. J.; Davidson, M. W.; Piston, D. W., "Fluorescent proteins at a glance" *J Cell Sci* **2011**, *124*, 157.

10. Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y., "Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp red fluorescent protein" *Nat Biotechnol* **2004**, *22*, 1567.

11. Shaner, N. C.; Patterson, G. H.; Davidson, M. W., "Advances in fluorescent protein technology" *J Cell Sci* **2007**, *120*, 4247.

12. Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y., "A guide to choosing fluorescent proteins" *Nat Methods* **2005**, *2*, 905.

13. Shcherbo, D.; Murphy, C. S.; Ermakova, G. V.; Solovieva, E. A.; Chepurnykh, T. V.; Shcheglov, A. S.; Verkhusha, V. V.; Pletnev, V. Z.; Hazelwood, K. L.; Roche, P. M.; Lukyanov, S.; Zaraisky, A. G.; Davidson, M. W.; Chudakov, D. M., "Far-red fluorescent tags for protein imaging in living tissues" *Biochem J* **2009**, *418*, 567.

14. Subach, F. V.; Piatkevich, K. D.; Verkhusha, V. V., "Directed molecular evolution to design advanced red fluorescent proteins" *Nat Methods* **2011**, *8*, 1019.

15. Olenych, S. G.; Claxton, N. S.; Ottenberg, K. G.; Davidson, M. W., "The Fluorescent Protein Color Palette" *Current Protocols in Cell Biology* **2007**, *Chapter 21*, Unit 21.5.

16. Kumagai, A.; Ando, R.; Miyatake, H.; Greimel, P.; Kobayashi, T.; Hirabayashi, Y.; Shimogori, T.; Miyawaki, A., "A Bilirubin-Inducible Fluorescent Protein from Eel Muscle" *Cell* **2013**, *153*, 1602.

17. Auldridge, M. E.; Satyshur, K. A.; Anstrom, D. M.; Forest, K. T., "Structure-guided Engineering Enhances a Phytochrome-based Infrared Fluorescent Protein" *J Biol Chem* **2012**, *287*, 7000.

18. Filonov, G. S.; Piatkevich, K. D.; Ting, L. M.; Kim, K.; Verkhusha, V. V., "A bacteriophytochrome-based near infra-red fluorescent protein for in vivo imaging" *Mol Biol Cell* **2011**, *22*.

19. Filonov, G. S.; Piatkevich, K. D.; Ting, L. M.; Zhang, J. H.; Kim, K.; Verkhusha, V. V., "Bright and stable near-infrared fluorescent protein for in vivo imaging" *Nat Biotechnol* **2011**, *29*, 757.

20. Lecoq, J.; Schnitzer, M. J., "An infrared fluorescent protein for deeper imaging" *Nat Biotechnol* **2011**, *29*, 715.

21. Pastrana, E., "Near-infrared probes" *Nat Methods* **2013**, *10*, 36.

22. Shcherbakova, D. M.; Verkhusha, V. V., "Near-infrared fluorescent proteins for multicolor in vivo imaging" *Nat Methods* **2013**, *10*, 751.

23. Shu, X. K.; Royant, A.; Lin, M. Z.; Aguilera, T. A.; Lev-Ram, V.; Steinbach, P. A.; Tsien, R. Y., "Mammalian Expression of Infrared Fluorescent Proteins Engineered from a Bacterial Phytochrome" *Science* **2009**, *324*, 804.

24. Wagner, J. R.; Zhang, J. R.; von Stetten, D.; Guenther, M.; Murgida, D. H.; Mroginski, M. A.; Walker, J. M.; Forest, K. T.; Hildebrandt, P.; Vierstra, R. D., "Mutational analysis of Deinococcus radiodurans bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes" *J Biol Chem* **2008**, *283*, 12212.

25. Yu, D.; Gustafson, W. C.; Han, C.; Lafaye, C.; Noirclerc-Savoye, M.; Ge, W. P.; Thayer, D. A.; Huang, H.; Kornberg, T. B.; Royant, A.; Jan, L. Y.; Jan, Y. N.; Weiss, W. A.; Shu, X. K., "An improved monomeric infrared fluorescent protein for neuronal and tumour brain imaging" *Nat Commun* **2014**, *5*.

26. Dean, K. M.; Palmer, A. E., "Advances in fluorescence labeling strategies for dynamic cellular imaging" *Nat Chem Biol* **2014**, *10*, 512.

27. Hinner, M. J.; Johnsson, K., "How to obtain labeled proteins and what to do with them" *Curr Opin Biotech* **2010**, *21*, 766.

28. Jing, C. R.; Cornish, V. W., "Chemical Tags for Labeling Proteins Inside Living Cells" *Accounts Chem Res* **2011**, *44*, 784.

29. Jung, D.; Min, K.; Jung, J.; Jang, W.; Kwon, Y., "Chemical biology-based approaches on fluorescent labeling of proteins in live cells" *Mol Biosyst* **2013**, *9*, 862.

30. Marks, K. M.; Nolan, G. P., "Chemical labeling strategies for cell biology" *Nat Methods* **2006**, *3*, 591.

31. Rashidian, M.; Dozier, J. K.; Distefano, M. D., "Enzymatic Labeling of Proteins: Techniques and Approaches" *Bioconjugate Chem* **2013**, *24*, 1277.

32. Sletten, E. M.; Bertozzi, C. R., "Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality" *Angew Chem Int Edit* **2009**, *48*, 6974.

33. Sunbul, M.; Yin, J., "Site specific protein labeling by enzymatic posttranslational modification" *Org Biomol Chem* **2009**, *7*, 3361.

34. Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y., "Creating new fluorescent probes for cell biology" *Nat Rev Mol Cell Bio* **2002**, *3*, 906.

35. Lee, H. L. D.; Lord, S. J.; Iwanaga, S.; Zhan, K.; Xie, H. X.; Williams, J. C.; Wang, H.; Bowman, G. R.; Goley, E. D.; Shapiro, L.; Twieg, R. J.; Rao, J. H.; Moerner, W. E., "Superresolution Imaging of Targeted Proteins in Fixed and Living Cells Using Photoactivatable Organic Fluorophores" *J Am Chem Soc* **2010**, *132*, 15099.

36. Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohane, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V., "HatoTag: A novel protein labeling technology for cell imaging and protein analysis" *Acs Chem Biol* **2008**, *3*, 373.

37. Los, G. V.; Wood, K., "The HaloTag A Novel Technology for Cell Imaging and Protein Analysis" *Methods in Molecular Biology* **2007**, *356*, 195.

38. Gautier, A.; Juillerat, A.; Heinis, C.; Correa, I. R.; Kindermann, M.; Beaufils, F.; Johnsson, K., "An engineered protein tag for multiprotein labeling in living cells" *Chem Biol* **2008**, *15*, 128.

39. Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K., "A general method for the covalent labeling of fusion proteins with small molecules in vivo" *Nat Biotechnol* **2003**, *21*, 86.

40. Keppler, A.; Kindermann, M.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K., "Labeling of fusion proteins of O-6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro" *Methods* **2004**, *32*, 437.

41. Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K., "Labeling of fusion proteins with synthetic fluorophores in live cells" *P Natl Acad Sci USA* **2004**, *101*, 9955.

42. Lukinavicius, G.; Umezawa, K.; Olivier, N.; Honigmann, A.; Yang, G. Y.; Plass, T.; Mueller, V.; Reymond, L.; Correa, I. R.; Luo, Z. G.; Schultz, C.; Lemke, E. A.; Heppenstall, P.; Eggeling, C.; Manley, S.; Johnsson, K., "A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins" *Nat Chem* **2013**, *5*, 132.

43. Sun, X. L.; Zhang, A. H.; Baker, B.; Sun, L.; Howard, A.; Buswell, J.; Maurel, D.; Masharina, A.; Johnsson, K.; Noren, C. J.; Xu, M. Q.; Correa, I. R., "Development of SNAP-Tag Fluorogenic Probes for Wash-Free Fluorescence Imaging" *Chembiochem* **2011**, *12*, 2217.

44. Mizukami, S.; Watanabe, S.; Akimoto, Y.; Kikuchi, K., "No-Wash Protein Labeling with Designed Fluorogenic Probes and Application to Real-Time Pulse-Chase Analysis" *J Am Chem Soc* **2012**, *134*, 1623.

45. Mizukami, S.; Watanabe, S.; Hori, Y.; Kikuchi, K., "Covalent Protein Labeling Based on Noncatalytic beta-Lactamase and a Designed FRET Substrate" *J Am Chem Soc* **2009**, *131*, 5016.

46. Mizukami, S.; Yamamoto, T.; Yoshimura, A.; Watanabe, S.; Kikuchi, K., "Covalent Protein Labeling with a Lanthanide Complex and Its Application to Photoluminescence Lifetime-Based Multicolor Bioimaging" *Angew Chem Int Edit* **2011**, *50*, 8750.

47. Watanabe, S.; Mizukami, S.; Akimoto, Y.; Hori, Y.; Kikuchi, K., "Intracellular Protein Labeling with Prodrug-Like Probes Using a Mutant beta-Lactamase Tag" *Chem-Eur J* **2011**, *17*, 8342.

48. Watanabe, S.; Mizukami, S.; Hori, Y.; Kikuchi, K., "Multicolor Protein Labeling in Living Cells Using Mutant beta-Lactamase-Tag Technology" *Bioconjugate Chem* **2010**, *21*, 2320.

49. Komatsu, T.; Johnsson, K.; Okuno, H.; Bito, H.; Inoue, T.; Nagano, T.; Urano, Y., "Real-Time Measurements of Protein Dynamics Using Fluorescence Activation-Coupled Protein Labeling Method" *J Am Chem Soc* **2011**, *133*, 6745.

50. Zhang, C. J.; Li, L.; Chen, G. Y. J.; Xu, Q. H.; Yao, S. Q., "One- and Two-Photon Live Cell Imaging Using a Mutant SNAP-Tag Protein and Its FRET Substrate Pairs" *Org Lett* **2011**, *13*, 4160.

51. Cohen, J. D.; Thompson, S.; Ting, A. Y., "Structure-Guided Engineering of a Pacific Blue Fluorophore Ligase for Specific Protein Imaging in Living Cells" *Biochemistry-Us* **2011**, *50*, 8221.

52. Fernandez-Suarez, M.; Baruah, H.; Martinez-Hernandez, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y., "Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes" *Nat Biotechnol* **2007**, *25*, 1483.

53. Slavoff, S. A.; Liu, D. S.; Cohen, J. D.; Ting, A. Y., "Imaging Protein-Protein Interactions inside Living Cells via Interaction-Dependent Fluorophore Ligation" *J Am Chem Soc* **2011**, *133*, 19769.

54. Uttamapinant, C.; White, K. A.; Baruah, H.; Thompson, S.; Fernandez-Suarez, M.; Puthenveetil, S.; Ting, A. Y., "A fluorophore ligase for site-specific protein labeling inside living cells" *P Natl Acad Sci USA* **2010**, *107*, 10914.

55. Yao, J. Z.; Uttamapinant, C.; Poloukhtine, A.; Baskin, J. M.; Codelli, J. A.; Sletten, E. M.; Bertozzi, C. R.; Popik, V. V.; Ting, A. Y., "Fluorophore Targeting to Cellular Proteins via Enzyme-Mediated Azide Ligation and Strain-Promoted Cycloaddition" *J Am Chem Soc* **2012**, *134*, 3720.

56. Chen, I.; Howarth, M.; Lin, W. Y.; Ting, A. Y., "Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase" *Nat Methods* **2005**, *2*, 99.

57. Howarth, M.; Chinnapen, D. J. F.; Gerrow, K.; Dorrestein, P. C.; Grandy, M. R.; Kelleher, N. L.; El-Husseini, A.; Ting, A. Y., "A monovalent streptavidin with a single femtomolar biotin binding site" *Nat Methods* **2006**, *3*, 267.

58. Howarth, M.; Takao, K.; Hayashi, Y.; Ting, A. Y., "Targeting quantum dots to surface proteins in living cells with biotin ligase" *P Natl Acad Sci USA* **2005**, *102*, 7583.

59. Slavoff, S. A.; Chen, I.; Choi, Y. A.; Ting, A. A. Y., "Expanding the substrate tolerance of biotin ligase through exploration of enzymes from diverse species" *J Am Chem Soc* **2008**, *130*, 1160.

60. Borgstahl, G. E. O.; Williams, D. R.; Getzoff, E. D., "1.4 Angstrom Structure of Photoactive Yellow Protein, a Cytosolic Photoreceptor - Unusual Fold, Active-Site, and Chromophore" *Biochemistry-Us* **1995**, *34*, 6278.

61. Hellingwerf, K. J.; Hendriks, J.; Gensch, T., "Photoactive Yellow Protein, a new type of photoreceptor protein: Will this "yellow lab" bring us where we want to go?" *J Phys Chem A* **2003**, *107*, 1082.

62. Meyer, T. E., "Isolation and Characterization of Soluble Cytochromes, Ferredoxins and Other Chromophoric Proteins from the Halophilic Phototrophic Bacterium Ectothiorhodospira-Halophila" *Biochim Biophys Acta* **1985**, *806*, 175.

63. Meyer, T. E.; Yakali, E.; Cusanovich, M. A.; Tollin, G., "Properties of a Water-Soluble, Yellow Protein Isolated from a Halophilic Phototrophic Bacterium That Has Photochemical Activity Analogous to Sensory Rhodopsin" *Biochemistry-Us* **1987**, *26*, 418.

64. Devanathan, S.; Genick, U.; Canestrelli, I.; Getzoff, E.; Meyer, T.; Cusanovich, M.; Tollin, G., "Probing the active site of photoactive yellow protein (PYP) from Ectothiorhodospira halophila." *Biophys J* **1997**, *72*, Tu401.

65. Genick, U. K.; Devanathan, S.; Meyer, T. E.; Canestrelli, I. L.; Williams, E.; Cusanovich, M. A.; Tollin, G.; Getzoff, E. D., "Active site mutants implicate key residues for control of color and light cycle kinetics of photoactive yellow protein" *Biochemistry-Us* **1997**, *36*, 8.

66. Kumauchi, M.; Hara, M. T.; Stalcup, P.; Xie, A. H.; Hoff, W. D., "Identification of six new photoactive yellow proteins - Diversity and structure-function relationships in a bacterial blue light photoreceptor" *Photochem Photobiol* **2008**, *84*, 956.

67. Hori, Y.; Kikuchi, K., "Protein labeling with fluorogenic probes for no-wash live-cell imaging of proteins" *Curr Opin Chem Biol* **2013**, *17*, 644.

68. Hori, Y.; Nakaki, K.; Sato, M.; Mizukami, S.; Kikuchi, K., "Development of Protein-Labeling Probes with a Redesigned Fluorogenic Switch Based on Intramolecular Association for No-Wash Live-Cell Imaging" *Angew Chem Int Edit* **2012**, *51*, 5611.

69. Chen, Z. X.; Jing, C. R.; Gallagher, S. S.; Sheetz, M. P.; Cornish, V. W., "Second-Generation Covalent TMP-Tag for Live Cell Imaging" *J Am Chem Soc* **2012**, *134*, 13692.

70. Jing, C. R.; Cornish, V. W., "A Fluorogenic TMP-Tag for High Signal-to-Background Intracellular Live Cell Imaging" *Acs Chem Biol* **2013**, *8*, 1704.

71. Wang, T. Y.; Friedman, L. J.; Gelles, J.; Min, W.; Hoskins, A. A.; Cornish, V. W., "The Covalent Trimethoprim Chemical Tag Facilitates Single Molecule Imaging with Organic Fluorophores" *Biophys J* **2014**, *106*, 272.

72. Rosa, L. E. P.; Reddy, D. R.; Queener, S. F.; Miller, L. W., "Selective Antifolates for Chemically Labeling Proteins in Mammalian Cells" *Chembiochem* **2009**, *10*, 1462.

73. Feldhaus, M. J.; Siegel, R. W.; Opresko, L. K.; Coleman, J. R.; Feldhaus, J. M. W.; Yeung, Y. A.; Cochran, J. R.; Heinzelman, P.; Colby, D.; Swers, J.; Graff, C.; Wiley, H. S.; Wittrup, K. D., "Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library" *Nat Biotechnol* **2003**, *21*, 163.

74. Szent-Gyorgyi, C.; Schmidt, B. A.; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. A. J.; Woolford, C. A.; Yan, Q.; Vasilev, K. V.; Berget, P. B.; Bruchez, M. P.; Jarvik, J. W.; Waggoner, A., "Fluorogen-activating single-chain antibodies for imaging cell surface proteins" *Nat Biotechnol* **2008**, *26*, 235.

75. Szent-Gyorgyi, C.; Schmidt, B. F.; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. A. J.; Woolford, C. A.; Yan, Q.; Vasilev, K. V.; Berget, P. B.; Bruchez, M. P.; Jarvik, J. W.; Waggoner, A., "Fluorogen-activating single-chain antibodies for imaging cell surface proteins (vol 26, pg 235, 2008)" *Nat Biotechnol* **2008**, *26*, 470.

76. Yates, B. P.; Peck, M. A.; Berget, P. B., "Directed Evolution of a Fluorogen-Activating Single Chain Antibody for Function and Enhanced Brightness in the Cytoplasm" *Mol Biotechnol* **2013**, *54*, 829.

77. Ozhalici-Unal, H.; Pow, C. L.; Marks, S. A.; Jesper, L. D.; Silva, G. L.; Shank, N. I.; Jones, E. W.; Burnette, J. M.; Berget, P. B.; Armitage, B. A., "A rainbow of fluoromodules: A promiscuous scFv protein binds to and activates a diverse set of fluorogenic cyanine dyes" *J Am Chem Soc* **2008**, *130*, 12620.

78. Fitzpatrick, J. A. J.; Yan, Q.; Sieber, J. J.; Dyba, M.; Schwarz, U.; Szent-Gyorgyi, C.; Woolford, C. A.; Berget, P. B.; Waggoner, A. S.; Bruchez, M. P., "STED Nanoscopy in Living Cells Using Fluorogen Activating Proteins" *Bioconjugate Chem* **2009**, *20*, 1843.

79. Wu, Y.; Tapia, P. H.; Fisher, G. W.; Waggoner, A. S.; Jarvik, J.; Sklar, L. A., "High-throughput flow cytometry compatible biosensor based on fluorogen activating protein technology" *Cytom Part A* **2013**, *83A*, 220.

80. Holleran, J.; Brown, D.; Fuhrman, M. H.; Adler, S. A.; Fisher, G. W.; Jarvik, J. W., "Fluorogen-Activating Proteins as Biosensors of Cell-Surface Proteins in Living Cells" *Cytom Part A* **2010**, *77A*, 776.

81. Saunders, M. J.; Szent-Gyorgyi, C.; Fisher, G. W.; Jarvik, J. W.; Bruchez, M. P.; Waggoner, A. S., "Fluorogen activating proteins in flow cytometry for the study of surface molecules and receptors" *Methods* **2012**, *57*, 308.

82. Soh, N., "Selective chemical labeling of proteins with small fluorescent molecules based on metal-chelation methodology" *Sensors-Basel* **2008**, *8*, 1004.

83. Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y., "New biarsenical Ligands and tetracysteine motifs for protein labeling in vitro and in vivo: Synthesis and biological applications" *J Am Chem Soc* **2002**, *124*, 6063.

84. Cao, H. S.; Xiong, Y. J.; Wang, T.; Chen, B. W.; Squier, T. C.; Mayer, M. U., "A red Cy3-based biarsenical fluorescent probe targeted to a complementary binding peptide" *J Am Chem Soc* **2007**, *129*, 8672.

85. Griffin, B. A.; Adams, S. R.; Tsien, R. Y., "Specific covalent labeling of recombinant protein molecules inside live cells" *Science* **1998**, *281*, 269.

86. Nakanishi, J.; Nakajima, T.; Sato, M.; Ozawa, T.; Tohda, K.; Umezawa, Y., "Imaging of conformational changes of proteins with a new environment/sensitive fluorescent probe designed for site specific labeling of recombinant proteins in live cells" *Anal Chem* **2001**, *73*, 2920.

87. Spagnuolo, C. C.; Vermeij, R. J.; Jares-Erijman, E. A., "Improved photostable FRET-competent biarsenical-tetracysteine probes based on fluorinated fluoresceins" *J Am Chem Soc* **2006**, *128*, 12040.

88. Goldsmith, C. R.; Jaworski, J.; Sheng, M.; Lippard, S. J., "Selective labeling of extracellular proteins containing polyhistidine sequences by a fluorescein - Nitrilotriacetic acid conjugate" *J Am Chem Soc* **2006**, *128*, 418.

89. Guignet, E. G.; Hovius, R.; Vogel, H., "Reversible site-selective labeling of membrane proteins in live cells" *Nat Biotechnol* **2004**, *22*, 440.

90. Kapanidis, A. N.; Ebright, Y. W.; Ebright, R. H., "Site-specific incorporation of fluorescent probes into protein: Hexahistidine-tag-mediated fluorescent labeling with (Ni2+: Nitrilotriacetic acid)(n)-fluorochrome conjugates" *J Am Chem Soc* **2001**, *123*, 12123.

91. Soh, N.; Seto, D.; Nakano, K.; Imato, T., "Methodology of reversible protein labeling for ratiometric fluorescent measurement" *Mol Biosyst* **2006**, *2*, 128.

92. Amano, H.; Ohuchi, Y.; Katayama, Y.; Maeda, M., "A new fluorescent reagent for the detection of proteins having histidine-tag (his-tag)" *Anal. Sci* **2001**, *17*, i1469.

93. Honda, K.; Fujishima, S. H.; Ojida, A.; Hamachi, I., "Pyrene excimer-based dual-emission detection of a oligoaspartate tag-fused protein by using a Zn-11-DpaTyr probe" *Chembiochem* **2007**, *8*, 1370.

94. Nonaka, H.; Tsukiji, S.; Ojida, A.; Hamachi, I., "Non-enzymatic covalent protein labeling using a reactive tag" *J Am Chem Soc* **2007**, *129*, 15777.

95. Ojida, A.; Honda, K.; Shinmi, D.; Kiyonaka, S.; Mori, Y.; Hamachi, I., "Oligo-Asp Tag/Zn(II) complex probe as a new pair for labeling and fluorescence imaging of proteins" *J Am Chem Soc* **2006**, *128*, 10452.

96. Devaraj, N. K.; Hilderbrand, S.; Upadhyay, R.; Mazitschek, R.; Weissleder, R., "Bioorthogonal Turn-On Probes for Imaging Small Molecules inside Living Cells" *Angew Chem Int Edit* **2010**, *49*, 2869.

97. Lahann, J. *Click Chemistry for Biotechnology and Materials Science*; John Wiley & Sons Ltd: United Kingdom, 2009.

98. Lang, K.; Chin, J. W., "Bioorthogonal Reactions for Labeling Proteins" *Acs Chem Biol* **2014**, *9*, 16.

99. Milles, S.; Tyagi, S.; Banterle, N.; Koehler, C.; VanDelinder, V.; Plass, T.; Neal, A. P.; Lemke, E. A., "Click Strategies for Single-Molecule Protein Fluorescence" *J Am Chem Soc* **2012**, *134*, 5187.

100. Beatty, K. E.; Liu, J. C.; Xie, F.; Dieterich, D. C.; Schuman, E. M.; Wang, Q.; Tirrell, D. A., "Fluorescence visualization of newly synthesized proteins in mammalian cells" *Angew Chem Int Edit* **2006**, *45*, 7364.

101. Beatty, K. E.; Xie, F.; Wang, Q.; Tirrell, D. A., "Selective dye-labeling of newly synthesized proteins in bacterial cells" *J Am Chem Soc* **2005**, *127*, 14150.

102. Hangauer, M. J.; Bertozzi, C. R., "A FRET-based fluorogenic phosphine for live-cell Imaging with the Staudinger ligation" *Angew Chem Int Edit* **2008**, *47*, 2394.

103. Prescher, J. A.; Dube, D. H.; Bertozzi, C. R., "Chemical remodelling of cell surfaces in living animals" *Nature* **2004**, *430*, 873.

104. Saxon, E.; Bertozzi, C. R., "Cell surface engineering by a modified Staudinger reaction" *Science* **2000**, *287*, 2007.

105. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., "A strain-promoted [3+2] azide-alkyne cycloaddition for covalent modification of blomolecules in living systems" *J Am Chem Soc* **2004**, *126*, 15046.

106. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., "A new bioorthogonal ligation for labelling of azide-bearing glycoconjugates" *Glycobiology* **2004**, *14*, 1197.

107. Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R., "Copper-free click chemistry for dynamic in vivo imaging" *P Natl Acad Sci USA* **2007**, *104*, 16793.

108. Plass, T.; Milles, S.; Koehler, C.; Schultz, C.; Lemke, E. A., "Genetically Encoded Copper-Free Click Chemistry" *Angew Chem Int Edit* **2011**, *50*, 3878.

109. Uttamapinant, C.; Tangpeerachaikul, A.; Grecian, S.; Clarke, S.; Singh, U.; Slade, P.; Gee, K. R.; Ting, A. Y., "Fast, Cell-Compatible Click Chemistry with Copper-Chelating Azides for Biomolecular Labeling" *Angew Chem Int Edit* **2012**, *51*, 5852.

110. Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G., "Bioconjugation by copper(I)-catalyzed azide-alkyne [3+2] cycloaddition" *J Am Chem Soc* **2003**, *125*, 3192.

111. Yu, Z. P.; Pan, Y. C.; Wang, Z. Y.; Wang, J. Y.; Lin, Q., "Genetically Encoded Cyclopropene Directs Rapid, Photoclick-Chemistry-Mediated Protein Labeling in Mammalian Cells" *Angew Chem Int Edit* **2012**, *51*, 10600.

112. Blackman, M. L.; Royzen, M.; Fox, J. M., "Tetrazine ligation: Fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity" *J Am Chem Soc* **2008**, *130*, 13518.

113. Devaraj, N. K.; Weissleder, R.; Hilderbrand, S. A., "Tetrazine-Based Cycloadditions: Application to Pretargeted Live Cell Imaging" *Bioconjugate Chem* **2008**, *19*, 2297.

114. Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C. J.; Deiters, A.; Chin, J. W., "Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction" *Nat Chem* **2012**, *4*, 298.

115. Liu, D. S.; Tangpeerachaikul, A.; Selvaraj, R.; Taylor, M. T.; Fox, J. M.; Ting, A. Y., "Diels-Alder Cycloaddition for Fluorophore Targeting to Specific Proteins inside Living Cells" *J Am Chem Soc* **2012**, *134*, 792.

116. Plass, T.; Milles, S.; Koehler, C.; Szymanski, J.; Mueller, R.; Wiessler, M.; Schultz, C.; Lemke, E. A., "Amino Acids for Diels-Alder Reactions in Living Cells" *Angew Chem Int Edit* **2012**, *51*, 4166.

117. Berbasova, T.; Nosrati, M.; Vasileiou, C.; Wang, W. J.; Sing, K.; Lee, S.; Yapici, I.; Geiger, J. H.; Borhan, B., "Rational Design of a Colorimetric pH Sensor from a Soluble Retinoic Acid Chaperone" *J Am Chem Soc* **2013**, *135*, 16111.

118. Crist, R. M.; Vasileiou, C.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B., "Engineering a rhodopsin protein mimic" *J Am Chem Soc* **2006**, *128*, 4522.

119. Lee, K. S. S.; Berbasova, T.; Vasileiou, C.; Jia, X. F.; Wang, W. J.; Choi, Y.; Nossoni, F.; Geiger, J. H.; Borhan, B., "Probing Wavelength Regulation with an Engineered Rhodopsin Mimic and a C15-Retinal Analogue" *Chempluschem* **2012**, *77*, 273.

120. 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*, 6140.

121. 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*, 812.

122. Wang, W. J.; Geiger, J. H.; Borhan, B., "The photochemical determinants of color vision" *Bioessays* **2014**, *36*, 65.

123. Wang, W. J.; Nossoni, Z.; Berbasova, T.; Watson, C. T.; Yapici, I.; Lee, K. S. S.; Vasileiou, C.; Geiger, J. H.; Borhan, B., "Tuning the Electronic Absorption of Protein-Embedded All-trans-Retinal" *Science* **2012**, *338*, 1340.

124. Akerstrom, B.; Flower, D. R.; Salier, J. P., "Lipocalins: unity in diversity" *Bba-Protein Struct M* **2000**, *1482*, 1.

125. Banaszak, L.; Winter, N.; Xu, Z. H.; Bernlohr, D. A.; Cowan, S.; Jones, T. A., "Lipid-Binding Proteins - a Family of Fatty-Acid and Retinoid Transport Proteins" *Adv Protein Chem* **1994**, *45*, 89.

126. Flower, D. R.; North, A. C. T.; Sansom, C. E., "The lipocalin protein family: structural and sequence overview" *Bba-Protein Struct M* **2000**, *1482*, 9.

127. Li, E.; Norris, A. W., "Structure/function of cytoplasmic vitamin A-binding proteins" *Annu Rev Nutr* **1996**, *16*, 205.

128. Skerra, A., "Lipocalins as a scaffold" *Bba-Protein Struct M* **2000**, *1482*, 337.

129. Vaezeslami, S.; Mathes, E.; Vasilelou, C.; Borhan, B.; Geiger, J. H., "The structure of apo-wild-type cellular retinoic acid binding protein II at 1.4 A and its relationship to ligand binding and nuclear translocation" *J Mol Biol* **2006**, *363*, 687.

130. Cao, J. F.; Wu, T.; Hu, C.; Liu, T.; Sun, W.; Fan, J. L.; Peng, X. J., "The nature of the different environmental sensitivity of symmetrical and unsymmetrical cyanine dyes: an experimental and theoretical study" *Phys Chem Chem Phys* **2012**, *14*, 13702.

131. Dai, Z. F.; Lin, G. L.; Li, Z. X.; Peng, B. X., "Styryl cyanine dyes: Aggregate behaviour, electrochemical properties and photographic performance" *Imaging Sci J* **1998**, *46*, 73.

132. Dai, Z. F.; Peng, B. X., "Photophysics of indocarbocyanine dyes in organic solvents" *J Chin Chem Soc-Taip* **1998**, *45*, 237.

133. Levitus, M.; Ranjit, S., "Cyanine dyes in biophysical research: the photophysics of polymethine fluorescent dyes in biomolecular environments" *Q Rev Biophys* **2011**, *44*, 123.

134. Mishra, A.; Behera, R. K.; Behera, P. K.; Mishra, B. K.; Behera, G. B., "Cyanines during the 1990s: A review" *Chem Rev* **2000**, *100*, 1973.

135. Escobedo, J. O.; Rusin, O.; Lim, S.; Strongin, R. M., "NIR dyes for bioimaging applications" *Curr Opin Chem Biol* **2010**, *14*, 64.

136. Mujumdar, R. B.; Ernst, L. A.; Mujumdar, S. R.; Lewis, C. J.; Waggoner, A. S., "Cyanine Dye Labeling Reagents - Sulfoindocyanine Succinimidyl Esters" *Bioconjugate Chem* **1993**, *4*, 105.

137. Tatikolov, A. S., "Polymethine dyes as spectral-fluorescent probes for biomacromolecules" *J Photoch Photobio C* **2012**, *13*, 55.

138. Dai, Z. F.; Peng, B. X.; Fan, L. Z.; Li, Y. F., "Meso-chloro-substituted indodicarbocyanines: Aggregate behavior, electrochemical properties, crystal structure and photographic performance" *J Chin Chem Soc-Taip* **1998**, *45*, 643.

139. Meguellati, K.; Spichty, M.; Ladame, S., "Reversible Synthesis and Characterization of Dynamic Imino Analogues of Trimethine and Pentamethine Cyanine Dyes" *Org Lett* **2009**, *11*, 1123.

140. Hamer, F. M., "Bases of which methincyanines are the quaternary salts." *J Chem Soc* **1940**, 799.

141. Hilderbrand, S. A.; Weissleder, R., "Optimized pH-responsive cyanine fluorochromes for detection of acidic environments" *Chem Commun* **2007**, 2747.

142. Mazieres, M. R.; Duprat, C.; Wolf, J. G.; Roshal, A. D., "pH dependent spectral properties and electronic structure of benzothiazol containing cyanine dyes" *Dyes Pigments* **2009**, *80*, 355.

143. Casalboni, M.; De Matteis, F.; Prosposito, P.; Quatela, A.; Sarcinelli, F., "Fluorescence efficiency of four infrared polymethine dyes" *Chem Phys Lett* **2003**, *373*, 372.

144. Muddana, H. S.; Morgan, T. T.; Adair, J. H.; Butler, P. J., "Photophysics of Cy3-Encapsulated Calcium Phosphate Nanoparticles" *Nano Lett* **2009**, *9*, 1559.

145. Enoki, S.; Saeki, K.; Maki, K.; Kuwajima, K., "Acid denaturation and refolding of green fluorescent protein" *Biochemistry-Us* **2004**, *43*, 14238.

146. Fukuda, H.; Arai, M.; Kuwajima, K., "Folding of green fluorescent protein and the cycle3 mutant" *Biochemistry-Us* **2000**, *39*, 12025.

147. Follenius-Wund, A.; Bourotte, M.; Schmitt, M.; Iyice, F.; Lami, H.; Bourguignon, J. J.; Haiech, J.; Pigault, C., "Fluorescent derivatives of the GFP chromophore give a new insight into the GFP fluorescence process" *Biophys J* **2003**, *85*, 1839.

148. Wu, L. X.; Burgess, K., "Syntheses of highly fluorescent GFP-Chromophore analogues" *J Am Chem Soc* **2008**, *130*, 4089.

149. Baumann, G.; Frommel, C.; Sander, C., "Polarity as a Criterion in Protein Design" *Protein Eng* **1989**, *2*, 329.

150. Richards, F. M., "areas, volumes, packing and protein structure" *Annual Review of Biophysics and Bioengineering* **1977**, *6*, 151.

151. Eyal, E.; Najmanovich, R.; Edelman, M.; Sobolev, V., "Protein side-chain rearrangement in regions of point mutations" *Proteins-Structure Function and Genetics* **2003**, *50*, 272.

152. Eyal, E.; Najmanovich, R.; Sobolev, V.; Edelman, M., "MutaProt: a web interface for structural analysis of point mutations" *Bioinformatics* **2001**, *17*, 381.

153. Najmanovich, R.; Kuttner, J.; Sobolev, V.; Edelman, M., "Side-chain flexibility in proteins upon ligand binding" *Proteins-Structure Function and Genetics* **2000**, *39*, 261.

154. Zhao, S. R.; Goodsell, D. S.; Olson, A. J., "Analysis of a data set of paired uncomplexed protein structures: New metrics for side-chain flexibility and model evaluation" *Proteins-Structure Function and Genetics* **2001**, *43*, 271.

155. Ediz, V.; Lee, J. L.; Armitage, B. A.; Yaron, D., "Molecular engineering of torsional potentials in fluorogenic dyes via electronic substituent effects" *J Phys Chem A* **2008**, *112*, 9692.

156. Sanchez-Galvez, A.; Hunt, P.; Robb, M. A.; Olivucci, M.; Vreven, T.; Schlegel, H. B., "Ultrafast radiationless deactivation of organic dyes: Evidence for a two-state two-mode pathway in polymethine cyanines" *J Am Chem Soc* **2000**, *122*, 2911.

157. Silva, G. L.; Ediz, V.; Yaron, D.; Armitage, B. A., "Experimental and computational investigation of unsymmetrical cyanine dyes: Understanding torsionally responsive fluorogenic dyes" *J Am Chem Soc* **2007**, *129*, 5710.

158. Lui, T. K.; Hsieh, P. Y.; Zhuang, Y. D.; Hsia, C. Y.; Huang, C. L.; Lai, H. P.; Lin, H. S.; Chen, I. C.; Hsu, H. Y.; Tan, K. T., "A rapid SNAP-Tag fluorogenic probe based on an environment-sensitive fluorophore for no-wash live cell imaging" *Acs Chem Biol* **2014**, *In Press*.

159. Yang, S. J.; Tian, H.; Xiao, H. M.; Shang, X. H.; Gong, X. D.; Yao, S. D.; Chen, K. C., "Photodegradation of cyanine and merocyanine dyes" *Dyes Pigments* **2001**, *49*, 93.

160. Altman, R. B.; Terry, D. S.; Zhou, Z.; Zheng, Q. S.; Geggier, P.; Kolster, R. A.; Zhao, Y. F.; Javitch, J. A.; Warren, J. D.; Blanchard, S. C., "Cyanine fluorophore derivatives with enhanced photostability" *Nat Methods* **2012**, *9*, 68.

161. Renikuntla, B. R.; Rose, H. C.; Eldo, J.; Waggoner, A. S.; Armitage, B. A., "Improved photostability and fluorescence properties through polyfluorination of a cyanine dye" *Org Lett* **2004**, *6*, 909.

162. Samanta, A.; Vendrell, M.; Das, R.; Chang, Y. T., "Development of photostable near-infrared cyanine dyes" *Chem Commun* **2010**, *46*, 7406.

CHAPTER II: ELUCIDATING THE ROLE OF ALL-*TRANS*-RETINAL RING METHYL SUBSTITUENTS IN WAVELENGTH REGULATION

II.1 Overview Of Retinal Analogs Studied In Opsins

Membranes divide the biological cells into functional and structural compartments or organelles. There are many proteins associated with membranes (membrane proteins) enabling cells to communicate, receive signals, sense their environments and transport various substances across their membranes. These membrane bound proteins are part of diverse sets of protein families such as receptors, transporters, ion channels, pumps and proteins responsible for enzymatic activity or structural features.[1-4] A fascinating subclass of membrane bound proteins is the light absorbing or transducing proteins being able to function as a result of light absorption. Photosystems (light-harvesting complexes during photosynthesis), and opsins (light-transducing receptor proteins) are examples of such membrane proteins that capture and utilize the light in various cellular processes.[5-8] The latter protein family is divided into two different subfamilies: i) animal opsins or type II opsins present in higher eukaryotic organisms (vertebrates and invertebrates) and ii) microbial opsins or type I opsins found in prokaryotes, algae and fungi.[9-17] Although each opsin family serves distinct function in the host organisms, both share a common molecular protein structure (seven transmembrane helices, 7-TM helices) and require the incorporation of a light sensitive small molecule within their protein structure for light sensitivity (**Figure II-1a**). This small ligand is covalently

attached to the protein via a lysine residue as a protonated Schiff base in the binding pocket. All microbial opsins utilize all-*trans*-retinal as the light sensitive ligand whereas the chromophore in animal opsins could be either 11-*cis*-retinal, 3,4-dehydroretinal, 3-hydroxyretinal or 4-hydroxyretinal (**Figure II-1b**).[9,10,13,14,16,18] 11-*cis*-Retinal is the most commonly encountered chromophore in vertebrates and invertebrates, while other derivatives of retinal described above are present in sea animals, insects and molluscs.

The combination of different opsins with analogs of retinal isomers enables organisms to sense light at various wavelengths. Additionally, the absorbance maxima of the opsin-ligand complex could be further tuned by installing changes near the chromophore. To be able to answer the questions concerning the spectral tuning of opsins, numerous experimental[19-31] and theoretical studies[32-44] have been carried out over the years (as summarized in excellent reviews[9,12,45-48]). Broadly, two major factors are involved in spectral tuning; i) electrostatic interactions of the chromophore with the protein [partial charges or polarities along the chromophore, dispersion interactions due to nearby aromatic residues, counterion balancing effects], and ii) changes in chromophore conformation as a results of steric interactions with the protein binding pocket. These mechanisms could be combined and play a role in destabilization or stabilization of either excited electronic state or ground state of the bound ligand to result in a change of its absorption maxima. The electrostatic interaction between the protein and the chromophore is extensively investigated.

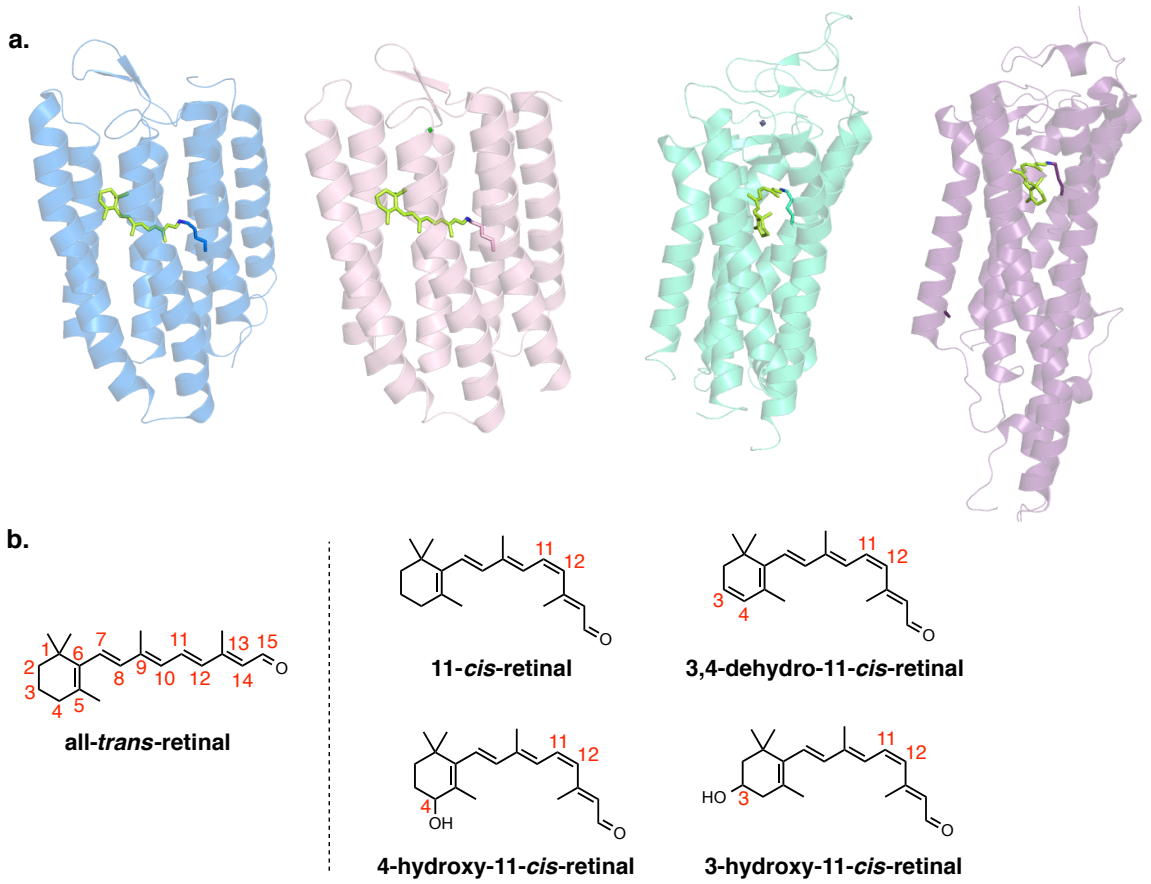


Figure II-1: Molecular structure of rhodopsins. **a.** Crystal structures of rhodopsins found in different organisms. The structures of bacteriorhodopsin (blue, PDB ID: 1FBB), sensory rhodopsin II (pink, PDB ID: 1H68), bovine rhodopsin (green, PDB ID: 1F88) and squid rhodopsin (purple, PDB ID: 2Z73) from left to right respectively. Bacteriorhodopsin and sensory rhodopsin II are microbial opsins whereas bovine rhodopsin and squid rhodopsin are isolated from the host animals. The bound chromophore is shown as a stick model (lime color). **b.** Structure of retinals used by opsins. Microbial opsins utilize all-*trans*-retinal (shown left side of the dash line) as a light sensitive chromophore. However, animal opsins preferentially bind 11-*cis* isomer of retinal or its derivatives as chromophores. The numbering of carbons is indicated in red.

Various mutagenesis studies along with the available crystal structures have led to compelling theories for the spectral tuning mechanism of opsins. Recently, we have also demonstrated that our rationally engineered rhodopsin mimics are an ideal platform to understand the effects of different protein environments on the

wavelength regulation of the bound ligand. Satisfyingly, our mimics regulated the absorption maximum of all-*trans*-retinal over 200 nm (425 nm to 644 nm) across the visible spectrum, yielding "redder than the red rhodopsin pigments".[49-51]

As a second important factor in spectral tuning of opsins, the effect of conformational changes of the bound chromophore within the protein cavity has been investigated. This phenomenon was firstly introduced by Blatz and Liebman and was named as the "twist mechanism".[52] They proposed that as the chromophore is co-planarized, the better overlapping of the π-system could be achieved resulting in a bathochromic shift in the absorption maximum. In other words, if a chromophore is twisted or distorted, then the degree of π-conjugation is decreased, inducing a blue shift in the spectrum. By systematically studying several conjugated polyenes, the enhancement in absorption of retinal vinylogs (polyenic cations) is predicted to be 62 nm per double bond addition in solution. Nevertheless, this prediction does not account for the protein environment. Indeed, the substitution of 11-*cis*-retinal with a 3,4-dehydro-11-*cis*-retinal in

Table II-1. Absorption maxima of visual opsins from different species with 11-*cis*-retinal and 3,4-dehydro-11-*cis*-retinal.

Species	λ^{max} (nm) 11-*cis*-retinal	λ^{max} (nm) 3,4-dehydro-11-*cis*-retinal	Δλ (nm)
Cattle (rod)	498	517	19
Goldfish (red cone)	579	625	46
Goldfish (green cone)	509	537	28
Goldfish (blue cone)	441	452	11
Goldfish (rod)	499	522	23
Chicken	562	620	58

Values were obtained by regeneration studies. Protein was identical for both chromophores.

various animal opsins results in different enhancement of absorption maxima (**Table II-1**).[53-57] Notably, the predicted increase in absorption as a result of each double bond addition was in close agreement with only chicken visual opsin (iodopsin), yielding a 58 nm red shift in the absorption maxima. However, binding of 3,4-dehydro-11-*cis*-retinal with other visual opsins did not correlate well with the expected shift in λ^{max} of each pigment in comparison with the values obtained with 11-*cis*-retinal. Although protein environments are identical for both ligands (presumably having same electrostatic and conformational features for both ligands), different degree of delocalization of the π-system along the bound ligand in red-absorbing pigments versus blue pigments could be considered as a possible source of absorption changes. Thus, association of 3,4-dehydro-11-*cis*-retinal with red-absorbing pigments such as goldfish red-cone or chicken iodopsin results in a large wavelength shift since the π-electrons are highly delocalized in those pigments and exhibit less bond alterations. On the other hand, 3,4-dehydro-11-*cis*-retinal bound to blue-absorbing pigments (rods, blue and green cones) shows small wavelength shifts due to the less delocalized π-system (weak dependence of λ^{max} on increase of double bonds in the polyene chain).[53,58,59]

It is well known that retinal is not a planar molecule and it is twisted about the single bonds to alleviate steric hindrance due to C5-CH_3/H8 and C13-CH_3/H10 (in case of 11-*cis* isomer) interactions. Crystal structures of both

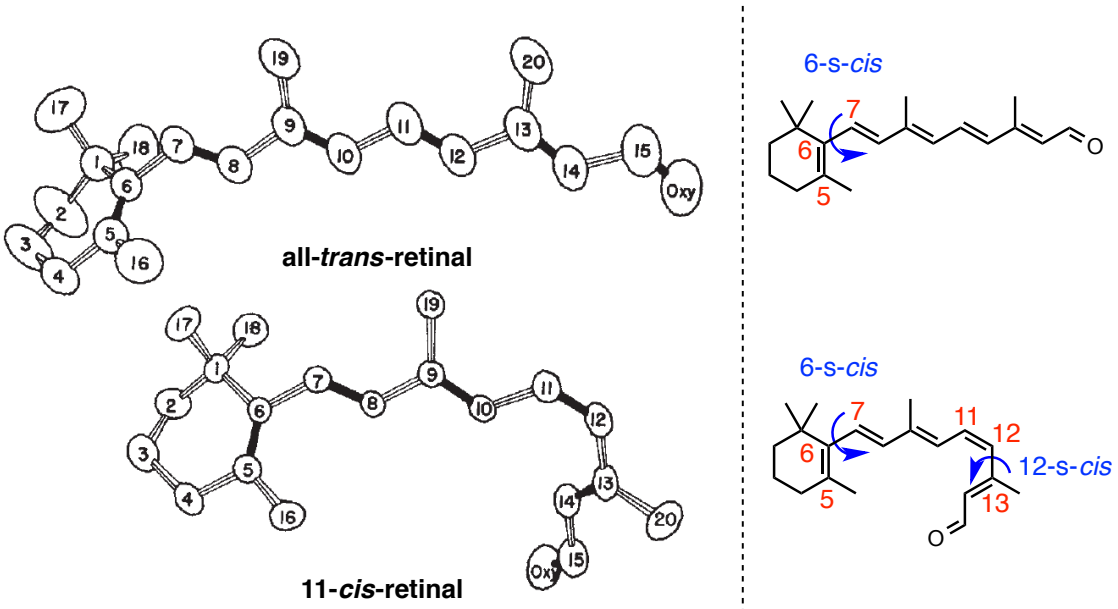


Figure II-2: Solid state crystal structures of all-*trans*-retinal (top) and the 11-*cis*-retinal (bottom).[60] The drawings on the right hand side depicts the torsion along the chromophore (shown by blue arrows).

isomers of retinal show that the ring-chain junction around the C6-C7 single bond adopts a 6-s-*cis* conformation with 59° and 40° torsion angles for all-*trans*-retinal and 11-*cis*-retinal, respectively.[60-62] Furthermore, there is a 39° torsion about the C12-C13 bond of the 11-*cis*-retinal in order to relief the steric hindrance of H10/H14 in 12-s-*cis* conformation (**Figure II-2**). It is worth mentioning that the binding of retinal aldehyde to opsin protein induces conformational changes on the retinal structure. The crystal structures of microbial opsins such as bacteriorhodpsin and sensory rhodospsin II revealed that the conformation around the C6-C7 single bond is 6-s-*trans* for the covalently attached all-*trans*-retinylidene chromophore, which is in a completely planarized geometry unlike that of seen in the solid state crystal structure of all-*trans*-retinal (**Figure II-2 top**

versus **Figure II-3a**).[9,63,64] Inspection of bovine rhodopsin and squid rhodopsin crystal structures also exposes the structural changes on the bound ligand in

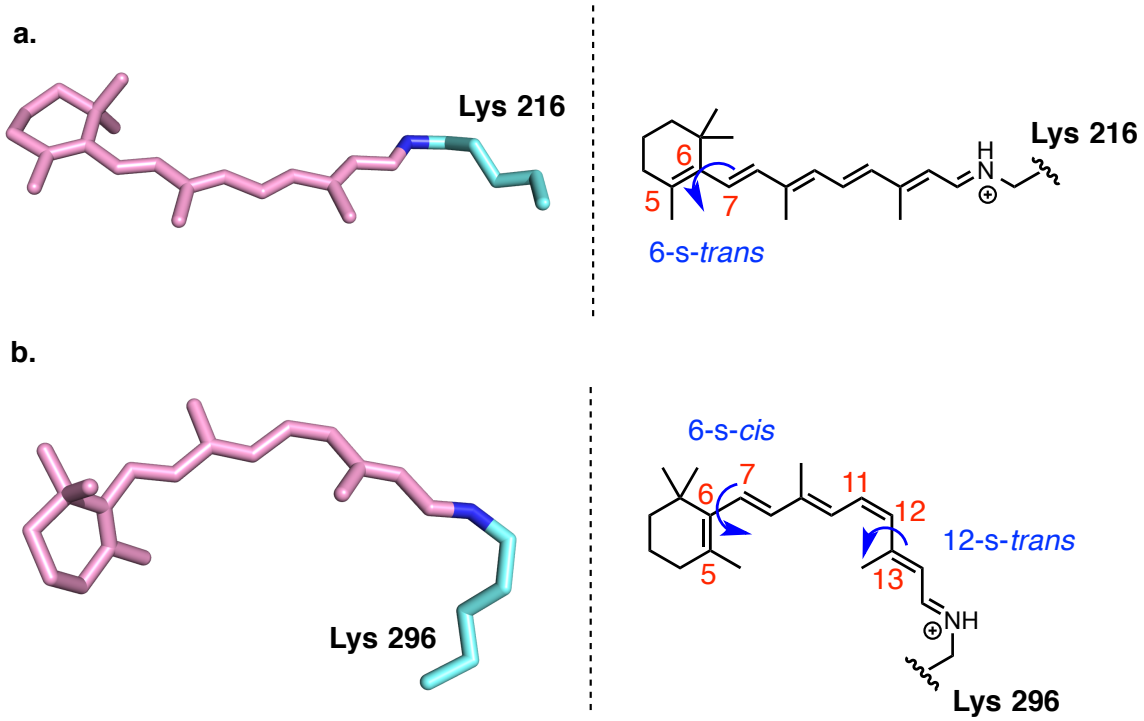


Figure II-3: Crystal structures of chromophores found in bacteriorhodopsin and bovine rhodopsin. **a.** The structure of all-*trans*-retinal bound to bacteriorhodopsin as a protonated Schiff base (PDB ID: 1FBB). Chromophore adopts a 6-s-*trans* conformation around the C6-C7 single bond. **b.** Crystal structure of bovine rhodopsin (PDB ID: 1F88) shows the conformation of the bound ligand, 11-*cis*-retinal, adopting a 6-s-*cis* and 12-s-*trans* conformation. Drawings on the right hand side depict the carbon numbering and highlight the torsions along the chromophore. Lysine residues are shown in cyan while the chromophores are indicated in pink for both **a** and **b**.

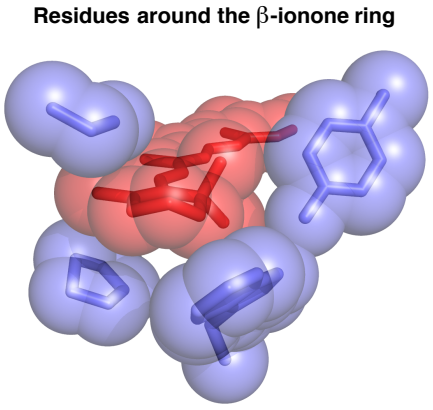
comparison with its solid state form.[65,66] Surprisingly, the 11-*cis*-retinylidene ligand adopts a 12-s-*trans* conformation in bovine rhodopsin, yielding virtually a planar geometry along the polyene chain (**Figure II-2 bottom** versus **Figure II-3b**). On the contrary, the ring-chain junction of the bound chromophore retains

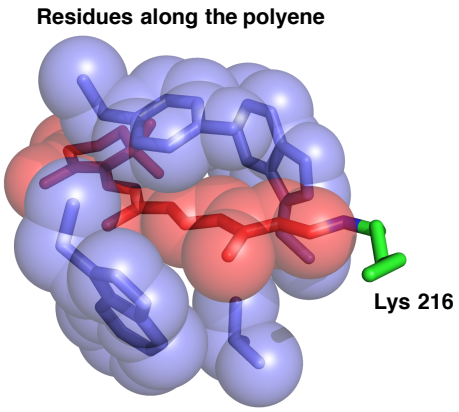
the 6-s-*cis* conformation, but in a severely distorted geometry (torsion angle ~ 59°). This highlights that both microbial opsins and visual rhodopsins retains the planarity of the polyene chain, however but they twist the β-ionone ring (with respect to the poylene backbone (about the C6-C7 single bond). Since microbial opsins favor the 6-s-*trans* conformation while visual rhodopsins produce a skewed 6-s-*cis* geometry, microbial opsins usually yield more red-shifted pigments with a better conjugation as compared to the visual rhodopsins. This was also proposed to be the possible distinction between the cone pigments (red, green, blue). It was suggested that the chromophore in red and green cone pigments is more planar, whereas the twist angle about the C6-C7 bond is larger in blue cones.[42,67] Unfortunately, the crystal structures of cone pigments have not been obtained yet to be able to dissect the conformational differentiations of the chromophores and interpret the relevant spectroscopic data.

One possible approach to understanding the conformational impact on spectral tuning of opsins is to modify the chromophore structure and systematically study the wavelength changes with respect to the modifications. Many artificial pigments based on synthetic retinal analogs have been extensively applied in an effort to elucidate the structure and function of bacteriorhodopsin and bovine rhodopsin. **Figure II-4** depicts a few examples of all-*trans*-retinal derivatives that were coupled with the bacteriorhodopsin in order to investigate the photophysical properties of the resulting artificial pigments.[25,68-82] Removal of the methyl groups from the β-ionone ring of the native

chromophore (**Figure II-4**, **A1**) yields artificial pigments absorbing at shorter wavelengths (see results for analogs **B1** to **E1**). This actually contradicts with the predictions in which chromophores having less steric hindrance at the ring-chain junction provide more planar conformations with a better π-electron delocalization. Indeed, these synthetic retinal aldehydes display red-shifted absorption maxima in solution in comparison with all-*trans*-retinal. However, binding of these analogs to the bacteriorhodopsin cause hypsochromic shifts in spectra, suggesting that the presence of the β-ionone ring along with the methyl substituents is crucial in order to restrict the degree of motion and full encapsulation of the ligand by the protein binding pocket (see the packed chromophore in **Figure II-4**). Namely, one could expect that chromophores devoid of the ring or its substituents exhibit a structural flexibility and a bond alteration in the binding cavity where more room is available for the reduced size of the chromophores. Perhaps, that is why the synthetic chromophore, **E1**, shows the largest blue shift as compared to other analogs (**B1** to **D1**) due to deplanarization of its last double bond from the rest of the poylene chain. Further support of this idea comes from the analogs in which the ring double bond is saturated. Analogs **F1**, **G1** (with reduced C5-C6 double bond), and **N1** (without the ring moiety) generate pigments, displaying similar absorption maxima. Since there is no olefinic double bond associated with a head group of the chromohore, the presence or the absence of the ring moiety along with its substituents does not have a role on wavelength regulation and the degree of freedom around the

ring-chain junction does not account for spectral tuning. In such pigments, only the conformation along the polyene backbone dictates the spectral changes. All





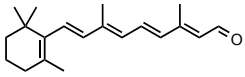
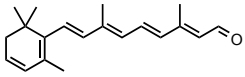
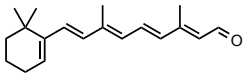
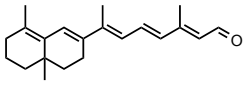
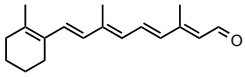
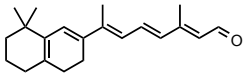
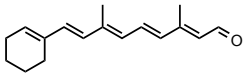
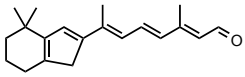
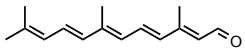
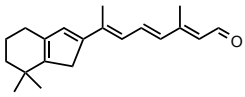
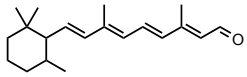
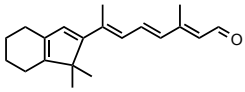
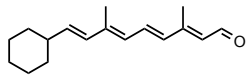
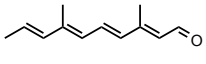
Chromophore	λ^{max} (nm)	Ref.	Chromophore	λ^{max} (nm)	Ref.
A1 	568	25,81	H1 	590	78
B1 	548	71,76	I1 	572	70
C1 	552	71	J1 	Two pigments 509 & 596	70,82
D1 	539	71,79	K1 	Two pigments 505 & 594	73
E1 	530	68,77,81	L1 	574	73
F1 	≈ 470	25,68,81	M1 	584	73
G1 	460	75	N1 	466	77

Figure II-4: Structures of synthetic retinal analogs studied in bacteriorhodopsin. The λ^{max} values corresponding to each bacteriorhodopsin/ligand are taken from literature reports. The space filling depiction of the crystal structure of bacteriorhodopsin (PDB ID: 1FBB) shows the residues (blue spheres) around the bound all-*trans*-retinylidene ligand (red sphere). The active site lysine residue is indicated in stick representation (green). For the retinal analog **F1**, the absorption maximum is averaged based on three different measurements reported in literature. For more analog studies, see references 69, 74 and 80.

three analogs (**F1**, **G1**, and **N1**) have a uniform olefinic chain and that is why it is not surprising to observe similar absorption regimes for the resulting artificial pigments. The synthesis of retinal derivatives with locked polyenes (structures are not shown here) has been also reported by Albeck *et al*.[83,84] It was shown that the artificial pigments formed from these retinal analogs have about the same spectral shifts in comparison with the all-*trans*-retinal, implying that there is no conformational restrictions along the olefinic chain of the chromophore in bacteriorhodopsin. This was then confirmed by the crystal structure of the bacterioprhodopsin in which the bound ligand is fully planarized within the binding pocket (**Figure II-3a** and **Figure II-4**).

As it is mentioned earlier, all-*trans*-retinal adopts a distorted 6-s-*cis* conformation about the C6-C7 single bond in both solid state (**Figure II-2**)[60,62] and in solution[85] with twist angles of 59° and ~ 30° respectively. Furthermore, the C6-C7 bond of retinal protonated Schiff base favors the twisted s-*cis* conformation in methanol solution,[86] whereas the protonated retinylidene iminium bound to bacteriorhodopsin is found to be in a 6-s-*trans* conformation without any torsions.[64] In order to understand the spectral contribution from conformational twists of retinal in the protein, ring locked retinal analogs were synthesized and spectroscopically analyzed both in solution and in the protein environment. By taking the assumption that these analogs are planar, both in solution and in protein bound states, the resulting red shift upon binding to a protein is due to the electrostatic interactions and not ring-chain planarization. Thus, the difference in

absorption of these analogs in comparison to the native chromophore, all-*trans*-retinal, could be interpreted as the influence of conformational changes on spectral tuning. Van der Steen and coworkers used the "ring locked" retinal analogs, **I1** and **J1** (**Figure II-4**), and calculated the contribution of the C6-C7 planarization to the opsin shift as 1300 cm^{-1} (**Table II-2**). On the other hand, Iwasa and his colleagues synthesized the 6-s-*cis* "fixed" retinals, which have a five membered ring fusion with varying positions of *gem*-dimethyl groups on the ring (**Figure II-4 K1**, **L1** and **M1**). Their study indicated that a C6-C7 bond planarization (skewed 6-s-*cis* → planar 6-s-*cis*) accounts for ~ 2330 cm^{-1} opsin shift (the number is averaged by all three analog measurements). This value is larger than the previously proposed value by Van der Steen (skewed 6-s-*cis* →

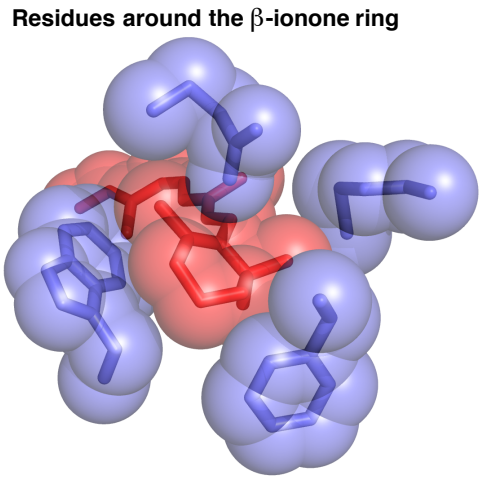
Table II-2. Absorption maxima of retinal analogs, their PSB's and artificial bacteriorhodopsin pigments.

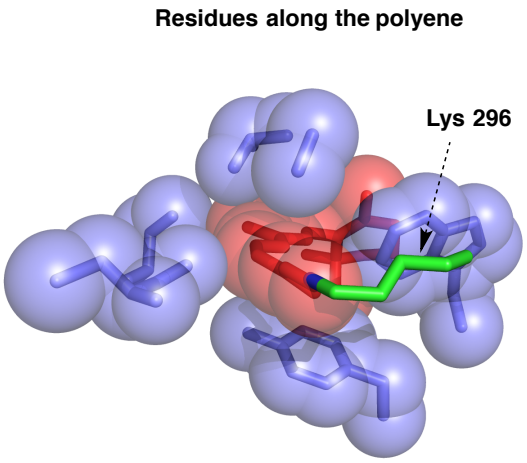
Retinal Analog	λ^{max} (nm) Aldehyde	λ^{max} (nm) PSB	λ^{max} (nm) in protein	Opsin Shift (cm^{-1})	ΔOpsin Shift (cm^{-1})
A1	380	440	568	5100	0
J1	400	465	572	3800	1300
I1	415	485	596	3800	1300
K1	425	502	594	3100	2000
L1	425	503	574	2500	2600
M1	425	504	584	2700	2400

Absorption maxima of aldehydes and protonated *n*-butylamine Schiff bases (PSB) were measured in methanol. Opsin shift is the difference in wavenumbers between the λ^{max} value of the PSB of the analog and that of the resulting artificial bacteriorhodopsin pigment. The contribution for planarization of the ring with the polyene to the total opsin shift (ΔOpsin Shift) is calculated by taking the difference of opsin shifts of all-*trans*-retinal (**A1**) and each retinal analog.

planar 6-s-*cis* or 6-s-*trans*, 1300 cm^{-1}) and could be considered as a more accurate estimation since the five-membered fixed analogs display better planarity than the six-membered locked forms.

Similar to artificial bacteriorhodopsin studies with synthetic retinal derivatives, studies with retinal analogs and bovine rhodopsin were instrumental to understand the structural changes that influence spectroscopic properties of the pigment.[57,72,84,87-99] Changes of the methyl substituents on the β-ionone ring of 11-*cis*-retinal (**Figure II-5**, **A2**) did not give rise to significant shifts in absorption maxima of the resultant pigments (see results for analogs; **B2**, **C2**, **F2**) but highlighted that the presence of these substituents is essential for a successful complexation of the ligand with the bovine rhodopsin (analog **E2** fails in pigment formation). Although synthetic retinals with modified β-ionone ring caused minimal changes in absorption maxima of artificial rhodopsins (± 3 nm on average), the isomeric analogs of these retinals coupled with bacteriorhodopsin yielded pronounced blue shifts (± 25 nm on average, **Figure II-4** see analogs **A1** to **E1**). The reason for the observed difference between these two proteins is due to the twist angle of the C6-C7 bond for their bound ligands. The crystal structure of bacteriorhodopsin shows that the C6-C7 torsion is almost zero for bound all-*trans*-retinal, proving a fully conjugated π-system along the chromophore. On the other hand, β-ionone ring of 11-*cis*-retinal is severely distorted in bovine rhodopsin, resulting in weak participation of the ring double bond with the delocalized π-electron system along the polyene (**Figure II-3**). As anticipated, the





Chromophore	λ^{max} (nm)	Ref.	Chromophore	λ^{max} (nm)	Ref.
A2	498	92-94,99	H2	517	57,90
B2	492	91	I2	498	72,96
C2	500	91	J2	485	84
D2	495	97	K2	490	87
E2	No pigment formation	91	L2	546	98
F2	496	97	M2	539	95,98
G2	460	88,89	N2	545	95

Figure II-5: Structures of synthetic retinal chromophores studied in bovine rhodopsin. The λ^{max} values are obtained from literature reports (as indicated in references). The space filling depiction of the crystal structure of bovine rhodopsin (PDB ID: 1F88) shows the residues (blue spheres) around the bound 11-*cis*-retinylidene ligand (red sphere). The active site lysine residue is indicated in stick representation (green).

structural changes on the ring moiety of the 11-*cis*-retinal does not cause significant spectral shifts for rhodopsin pigments. However, the same modifications for all-*trans*-retinal strongly influence the absorptive properties of the corresponding artificial bacteriorhodopsin complexes, because even the slight geometrical variations about the C6-C7 bond would be enough to destroy the fully conjugated π-system. In fact, the same phenomenon was observed with 5,6-dihydro retinal analogs. The coupling of 5,6-dihydro all-*trans*-retinal with bacteriorhodopsin leads to ~100 nm hypsochromic shift in the absorption spectrum (**Figure II-4**, **F1**), but this shift is about three fold less for bovine rhodopsin (**Figure II-5**, **G2**). Not surprisingly, 3,4-dehydro retinal analogs afforded red-shifted pigments in both bacteriorhodopsin and bovine opsin. Unexpectedly, the enhancement effect in bovine rhodopsin is similar to that seen in bacteriohodopsin. From these facts, it could be concluded that the sensitivity of absorption maxima of artificial pigments with respect to modifications on the β-ionone ring of the ligands is strongly dependent on the C6-C7 torsion. Less torsion around the C6-C7 single bond results in better conjugation in which the absorption maxima of the protein/ligand complex is more susceptible to structural changes of the ring.

For further investigation of the chromophore conformation around the ionone ring, different C6-C7 single bond locked bicyclic retinals were tested with bovine rhodopsin (**Figure II-5**, **L2**-**N2**). Opsin shifts of those artificial pigments were found to be smaller than that of rhodopsin/11-*cis*-retinal complex (**Table II-**

3). This was also observed in bacteriorhodopsin studies of Van der Steen and Iwasa, supporting that a part of the opsin shift originates from the coplanarization of the ligand in the protein binding site. However, neither form of the 11-*cis*-retinylidene is planar about the C6-C7 bond (in bovine crystal structure and in

Table II-3. Absorption maxima of retinal analogs, their PSB's and artificial bovine rhodopsin pigments.

Retinal Analog	λ^{max} (nm) Aldehyde	λ^{max} (nm) PSB	λ^{max} (nm) in protein	Opsin Shift (cm^{-1})	ΔOpsin Shift (cm^{-1})
A2	380	440	498	2650	0
L2	416	496	546	1850	800
M2	422	506	539	1210	1440
N2	405	495	545	1850	800

Absorption maxima of aldehydes and protonated *n*-butylamine Schiff bases (PSB) were measured in ethanol (PSB of **L2** in methanol). Opsin shift is the difference in wavenumbers between the λ^{max} value of the PSB of the analog and that of the resulting artificial bovine opsin pigment. The contribution for planarization of the ring with the polyene to the total opsin shift (ΔOpsin Shift) is calculated by taking the difference of opsin shifts of 11-*cis*-retinal (**A2**) and each retinal analog.

solution), thus it is hard to estimate the relative contribution to the opsin shift from of conformational events. Moreover, CD spectral measurements of bicyclic 11-*cis*-retinal analogs (**Figure II-5**, **L2**-**N2**) suggest that these ligands are still twisted around their C8-C9 single bond within the protein binding pocket.[95,98]

Finally, it is also important to clarify the steric interactions that originate from the 11*Z*-double bond of the chromophore in rhodopsin. The 11-*cis*-fixed retinal analog (**Figure II-5**, **I2**) was synthesized by Ito *et al*. and coupled with bovine rhodopsin. The resultant pigment absorbs at identical λ^{max} of the native rhodopsin, indicating that the conformation about the C12-C13 or C10-C11 single

bonds of the chromophore does not adopt a significant twist in the protein binding pocket.[96] Further support is provided by the artificial pigments derived from analogs of **J2** and **K2**. These pigments also yield opsin shifts close to that of the native rhodopsin, implying that the chromophore is not necessarily distorted along the polyene region. Indeed, the resolved crystal structure of bovine rhodopsin verified the latter assumptions. The 11-*cis*-retinylidene ligand was found in a linear conformation along the "polyene tail" (**Figure II-3b**) but slightly twisted about the C12-C13 single bond (torsion angle ~24°).

It should be noted that the absorption properties of synthetic retinal analogs are not only useful for artificial opsin studies, but also could be utilized directly for gas phase investigations in which the intrinsic properties of the chromophores are revealed in the absence of protein environment or solvent effects. Recently, the gas phase absorption spectrum of the protonated Schiff base of *n*-butyliminium of retinal was obtained, showing that the molecule has an intrinsic absorption shoulder at 610 nm.[100-102] This value is referred as an upper limit of the absorption of the chromophore where there are no interactions between the molecule and the solvents or any counteranions. The authors suggest a new definition of the opsin shift by using experimentally obtained gas phase absorption value (unbiased absorption reference). It was proposed that the opsin causes a blue shift in the absorption maximum of the protonated Schiff base of retinal via numerous interactions such as hydrogen bonding, dispersion interactions, Coulomb interactions and conformational changes. Furthermore,

Rajput *et al*. applied an improved experimental approach in order to obtain high-resolution gas phase absorption spectrum of retinal PSB.[103] **Figure II-6a** shows that the absorption spectrum of retinal PSB is a broad peak with a λ^{max} range from 530 nm to 610 nm. To assess the local structures of the recorded spectrum, the absorption profiles of two synthetic retinal chromophores were also measured. Both analogs of 5,6-dihydro and 1,1-didemethyl retinal protonated Schiff bases display clear peaks. As anticipated, the 5,6-dihydro analog absorbs at the blue edge of the spectrum with a λ^{max} of 525 nm, whereas the 1,1-didemethyl retinal PSB has a peak at 618 nm (**Figure II-6b**). This chromophore is more red-shifted because of its fully conjugated π-system in the 6-s-*trans* conformation. The overlay of experimentally obtained gas phase spectra of retinal PSB with these analogs suggest that the β-ionone ring of retinal PSB is able to explore different rotations about the C6-C7 single bond and spans all possible values from highly twisted to fully planar conformations. This implies that new reference wavelengths for the unbiased absorption of two retinal PSB conformers are with λ^{max} values of 530 nm (6-s-*cis*) and 610 nm (6-s-*trans*), respectively. Notably, both reports show the reference value of 610 nm as an upper limit of the retinal PSB, claiming that no retinal binding protein absorbs redder than this value. However, this statement is highly questionable, as our CRBPII and CRABPII rhodopsin mimics display highly red-shifted pigments (higher than 610 nm, up to 646 nm) upon forming a retinal PSB.[49,51]

The investigations of several synthetic retinal analogs in different environments (in solution, bound to proteins and gas phase) have clearly shown the importance of conformational changes for spectral regulation.

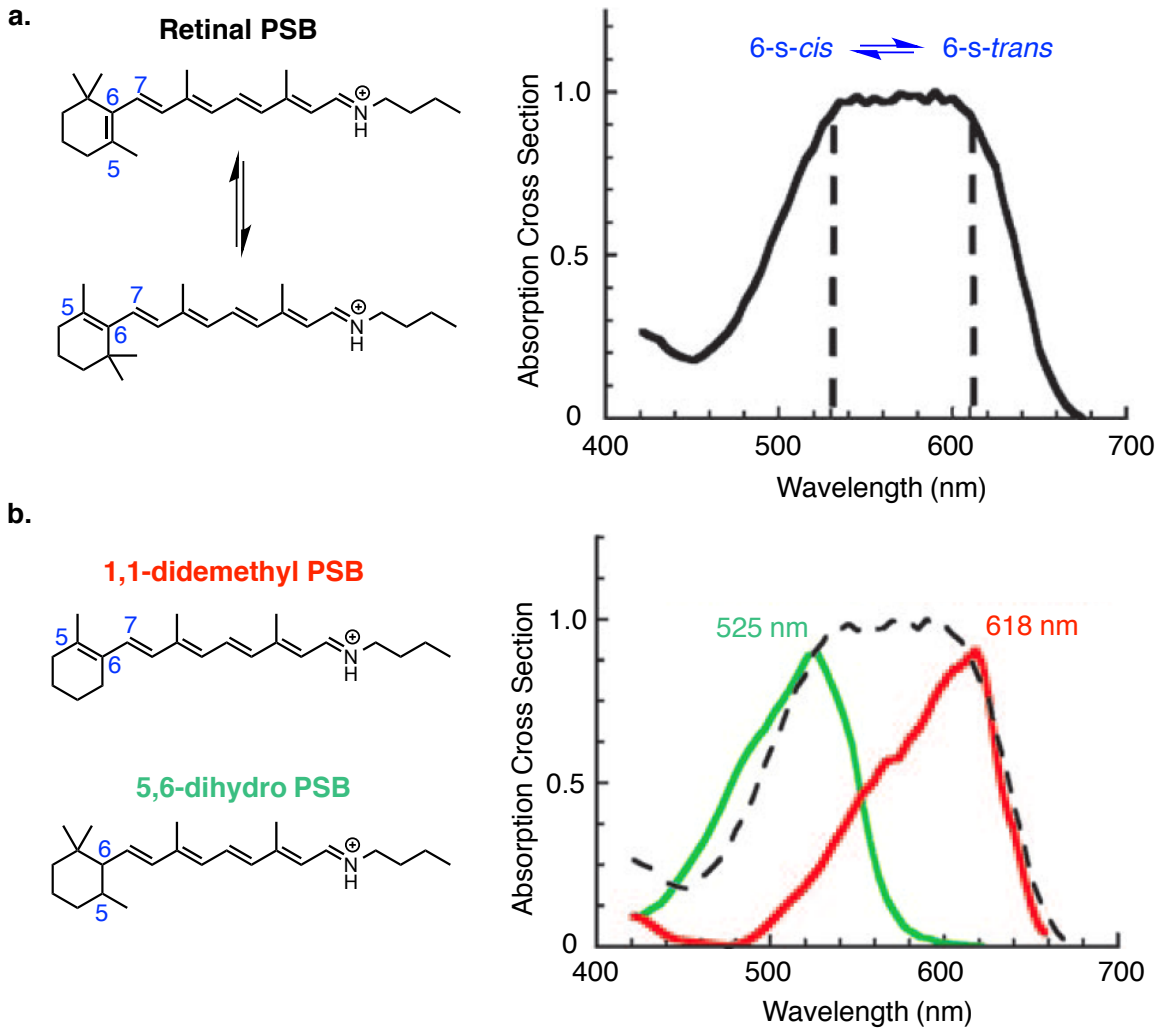


Figure II-6: Experimentally measured gas phase absorption cross sections of protonated Schiff bases (PSB) of retinal and retinal analogs.[103] **a.** PSB of all-*trans*-retinal with *n*-butyliminium of retinal has two conformations about the C6-C7 single bond. The recorded spectrum shows a broad peak extending from 530 nm to 610 nm. These values are taken as reference wavelengths for the planar 6-s-*trans* and twisted 6-s-*cis* conformations. **b.** The absorption spectra of two analogs were superimposed over the broad absorption of native retinal PSB (dashed line). The absorption cross sections of 1,1-didemethyl and 5,6-dihydro retinal protonated Schiff bases are shown in red and green, respectively.

Crystallographic data on microbial opsins and animal visual rhodopsins has greatly contributed to a better understanding of the physical interactions of these proteins with their native retinal ligands. However, structural information remains elusive for the opsins complexed with other retinal analogs. Also, the paucity of the methods to access high resolution structural data for rhodopsins limits research in this field. From the point of view of structural chemical biology, the next challenging frontier in the field would be to obtain a detailed insight on wavelength changes with respect to chromophore modifications. Thus, a suitable platform in which covalent modulations of the chromophoric molecules could be analyzed both spectroscopically and structurally within their interacting protein partners is a worthwhile endeavor. Our lab recently had the occasion to engineer a small cytosolic protein, hCRBPII into rhodopsin mimics. Our rationally designed hCRBPII rhodopsin mimics have enjoyed the great success in recapitulating wavelength regulation observed with retinal bound opsin proteins. These mimics suffice to serve as testing grounds for many retinal derivatives, aiding in the generation of new artificial pigments and, providing crystallographic data of diverse sets of protein-ligand complexes. In this study, we have focused on coupling of the reengineered hCRBPII rhodopsin mimics with different retinal analogs in order to systematically study the role of structural variations of these ligands on spectral tuning of the resultant pigments.

II.2 Previous Study In HCRBPII Rhodopsin Mimics

Our former lab member Dr. Wenjing Wang has demonstrated that all-*trans*-retinal results in colored pigments upon complexation with engineered hCRBPII rhodopsin mimics.[51] The iminium-based complexation was achieved from the reaction of retinal with an active site nucleophilic lysine residue. With the aid of the wild-type hCRBPII crystal structure, the position of Glu108 was chosen

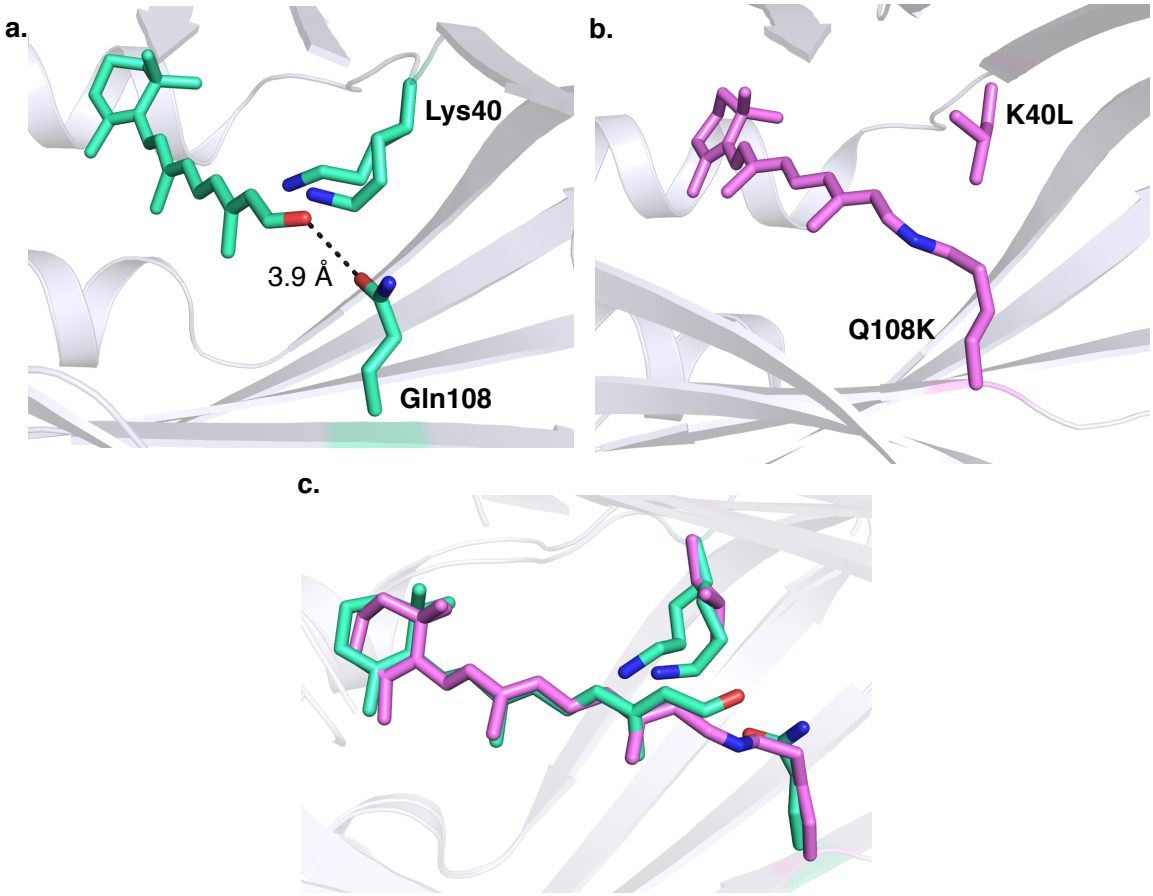
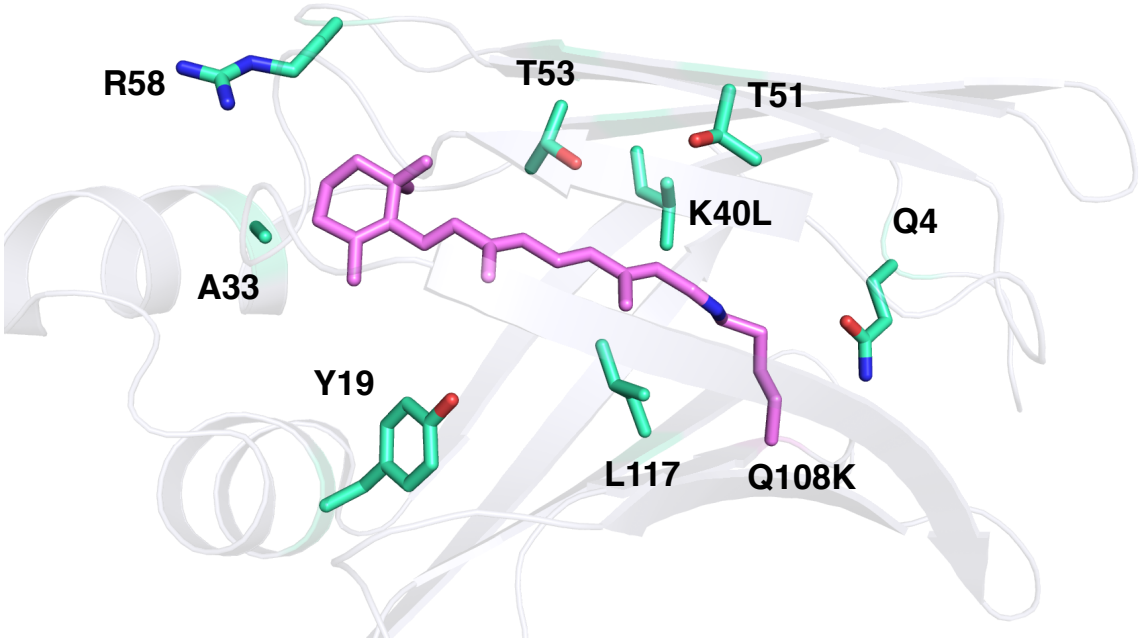


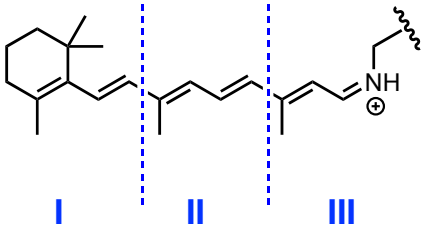
Figure II-7: Structure of wild-type hCRBPII and KL mutant bound to all-*trans*-retinal. **a.** Crystal structure of wild-type hCRBPII complexed with all-*trans*-retinal (PDB ID: 4QYP). The residues, Gln108 and K40L (in two conformations) are highlighted in stick model. **b.** Crystal structure of hCRBPII KL mutant bound to all-*trans*-retinal with a Schiff base linkage (PDB ID: 4EXZ). The installed nucleophilic lysine, Q108K, and the mutated K40L position are shown in sticks. **c.** The overlaid structures of **a** and **b** (wild-type hCRBPII/retinal - green, KL/retinal - pink)

as the site to install the nucleophilic lysine residue. A second mutation, K40L, was necessary in order to provide an ideal protein partner for retinal with an efficient protonation of the *in situ* formed Schiff base (**Figure II-7**). The resultant Q108K:K40L mutant (KL) showed high affinity towards all-*trans*-retinal and successfully formed a stable PSB, yielding a KL/retinal complex with a λ^{max} of 508 nm at physiological pH. The readily crystalizable nature of the hCRBPII protein was invaluable to obtain structural data on protein-ligand conjugates. Inspection of the crystal structure of KL/retinal complex highlighted the key residues around the bound chromophore that could be mutated into different amino acids to change the electrostatic environment of the protein binding pocket (**Figure II-8**). By dividing the chromophore environment into three segments, the modifications to amino acid residues in each zone were examined. At the onset, changes to zone I (ionone ring end) required installation of hydrophobic aromatic residues to embed the chromophore within the binding pocket and also to enclose the binding cavity from the aqueous environment. The mutations of A33W and R58W were crucial to reduce the solvent exposure of the bound chromophore for enhancing its sensitivity towards further changes in the binding site. Isolation of the protein interior from the bulk medium greatly promoted conjugation of the positive charge along the bound chromophore in an environment of low dielectric constant. On the other hand, changes to amino acids in zone II (middle of the polyene chain) did not cause significant shift in the absorption of the bound ligand. Several mutations were examined in place of

Thr53 and Tyr19, but the resultant pigments exhibited about 10 nm spectral shifts in their absorption maxima. Finally, modifications to zone III (PSB region) substantially affected the absorptive properties of the bound cationic chromophore. As anticipated, installation of negative polarities around the iminium yielded hypsochromically shifted spectra in pigments as a result of localized charge at PSB region, decreasing the π-electron conjugation. Therefore, mutations of T51D, L117E, and K40S resulted in significantly blue-shifted hCRBPII/retinal complexes (**Figure II-8**, **M1**, **M2** and **M3**). Bathochromic shift, on the other hand, was obtained by reversal of polarity on these residues. The isosteric replacement of Thr51 with a neutral hydrophobic residue, valine, led to a red shift in absorption of the bound retinal (**M5**, and other T51V containing mutants in the list). By removal of the polarity around the PSB region, the positive charge on nitrogen atom was destabilized, yielding an even distribution of the partial charge along the entire ligand. Similarly, disruption of a water network residing around the PSB, which was seen in all hCRBPII/retinal complexes, helped to remove any further polar environment in zone III to reach the most bathochromically shifted pigments. This water network originated from Thr1 to Gln4 and extended to the iminium nitrogen. By mutating the Gln4 residue, the effect of the water network in stabilization of the positive charge at iminium was abolished, resulting in highly conjugated chromophoric systems (**M9**, and other variants with Q4 mutations). The combination of the all critical point mutations yielded “super-red” shifted pigments absorbing over 600 nm in the visible



Segments Along The Chromophore



Entry	hCRBPII Mutant	λ^{max} (nm)
M1	Q108K:T51D:L117E	425
M2	Q108K:T51D	474
M3	Q108K:K40S	482
M4	Q108K:K40L	508
M5	Q108K:K40L:T51V	533
M6	Q108K:K40L:T51V:T53C	539
M7	Q108K:K40L:T51V:R59W	570
M8	Q108K:K40L:T51V:Y19W:R59W	577
M9	Q108K:K40L:T51V:Y19W:R59W:Q4W	593
M10	Q108K:K40L:T51V:T53C:Y19W:R59W:Q4W	613
M11	Q108K:K40L:T51V:T53C:Y19W:R59W:Q4R	622
M12	Q108K:K40L:T51V:T53C:Y19W:R59W:Q4R:A33W	644

Figure II-8: Crystal structure of KL/retinal complex (top) shows the residues close to different locations of the bound chromophore. The ligand is divided into three segments, indicated by blue dashed lines. The table lists hCRBPII variants with different mutations along the polyene and the corresponding absorption maximum.

spectrum. Obtaining crystal structures of such red-shifted pigments revealed that the ligand is fully encapsulated within the protein cavity and isolated from the bulk medium. Additionally, structures of the most red-shifted mutants showed almost coplanar geometry about the C6-C7 single bond, yielding maximally conjugated π-electron system. Therefore, extreme red shifts were accomplished by the uniform distribution of the charge along the bound ligand within an isolated protein environment.

In conclusion, rationally designed hCRBPII rhodopsin mimics enabled us to recapitulate the wavelength regulation observed with retinal bound opsin proteins. These mimics are applicable for systematically studying the correlation between structure and spectroscopic properties of protein-ligand complexes. Finally, it is important to emphasize that the availability of the crystal structures of engineered hCRBPII mutants is crucial for further optimization of protein-ligand interactions with desired spectroscopic characteristics.

II.3 Mechanistic Studies Of Wavelength Tuning In HCRBPII Rhodopsin Mimics With Different Retinal Analogs

Having several hCRBPII mutants in hand, we initiated a study to test diverse sets of retinoid ligands in order to evaluate their photophysical properties. Based on previously obtained crystal structures of the hCRBPII mutants with all-*trans*-retinal, we observed that the planarity along the polyene backbone of the chromophore is preserved in the protein cavity. However, different hCRBPII variants could lead to conformational variations about the C6-C7 bond. In the

present coarse-grained study, retinal analogs with modified ionone rings were screened to elucidate geometrical alterations as a function of observed wavelength of absorption. Commercially available all-*trans*-retinal and *trans*-3-dehydro-retinal were purchased from Sigma Aldrich and Toronto Research Chemicals, respectively. Synthetic procedures to obtain other retinal analogs used in this study are described in Chapter IV.13.2. In order to simplify the name of the protein-ligand complexes, each retinoid aldehyde was abbreviated as follows; i) all-*trans*-retinal, **RET**, ii) 3,4-dehydro-retinal, **DEHYDRO**, iii) 5-demethyl retinal, **5DEM**, iv) 1,1-didemethyl retinal, **GEMLESS**, v) cyclohexene retinal, **CYCHXE**, vi) cyclohexane retinal, **CYCHXA**.

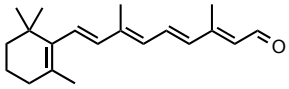
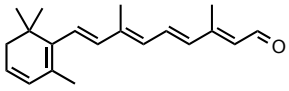
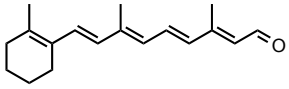
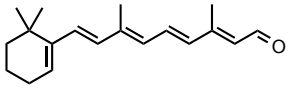
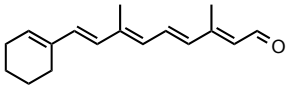
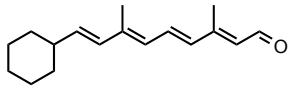
Before complexation of retinal aldehydes with hCRBPII rhodopsin mimics, spectroscopic properties of these compounds were analyzed in ethanol solution. **Table II-4** lists wavelengths for each chromophore in aldehyde, Schiff base, and the corresponding protonated Schiff base form. It is not surprising that the GEMLESS chromophore exhibits the most bathochromic shift. It was previously calculated that the energy difference between the optimized conformers (twisted 6-s-*cis* and planar 6-s-*trans*) is 3.8 Kcal/mol, in favor of the fully planar 6-s-*trans* conformer.[103] Thus, the coplanarity of the chromophore provides a strong conjugation along the chromophore, yielding a red-shifted spectrum. Analogs of DEHYDRO and CYCHXE shows similar absorptive properties. Although DEHYDRO chromophore has an additional double bond in the ring moiety, it adopts a twisted conformation to alleviate steric hindrance due to C5-CH_3/H8

interaction (same phenomenon is observed for retinal). Predictably, the bare cyclohexene ring does not suffer from steric strain and provides a better conjugated system. That is presumably why the twisted DEHYDRO analog and the planar CYCHXE chromophore have close λ^{max} values in solution. Substitution of the gem-dimethyl group on the cyclohexene ring at C1 possibly disturbs the planar geometry of the chromophore, thereby resulting in a blue shift in the absorption maximum of the 5DEM analog. However, the 5DEM chromophore still remains red-shifted as compared to retinal due to the absence of the unfavored C5-CH_3/H8 interaction. Lastly, removal of the ring double bond changes the λ^{max} of CYCHXE from 396 nm to 367 nm (wavelength difference is 29 nm), which is in agreement with tWoodward-Fieser prediction.

In order to elucidate the photophysical characteristics of the aforementioned retinal analogs in protein environments, we have followed previously applied systematic mutagenesis of hCRBPII. Our initial study of hCRBPII mutants with all-*trans*-retinal revealed that electrostatic changes close to the middle of the polyene chain does not have a significant impact on spectral tuning of the chromophore (~ 10 nm wavelength shifts). On the other hand, high sensitivity towards electrostatic modulations near the PSB region was seen for the all-*trans*-retinylidene chromophore. To examine behavior of retinal derivatives faced with similar electrostatic changes, a series of hCRBPII mutants were designed and coupled with retinal aldehydes. Our studies commenced with the

evaluation of a set of residues close to the middle of the polyene chain. Using the KL double mutant as a template, Thr53 was changed into neutral residues such as alanine, leucine, valine and serine. These changes foreseeably afforded pigments absorbing at similar wavelengths for each chromophore. The five hCRBPII mutants (varying residues at position 53) altered the spectra up to 10 nm with the ligand series (**Table II-5**). However, comparison between the pigments generated with all-*trans*-retinal and the analogs hinted at the order of conjugation for covalently bound chromophores. The differences in λ^{max} of

Table II-4. Absorption maxima of retinal analogs, their Schiff bases (SB) with *n*-butylamine and protonated forms (PSB).

Retinal Analog		ε (M.cm)$^{-1}$ Aldehyde	λ^{max} (nm) Aldehyde	λ^{max} (nm) SB	λ^{max} (nm) PSB
RET		48000	380	365	444
DEHYDRO		36500	398	376	463
GEMLESS		64000	404	379	478
5DEM		47800	388	368	453
CYCHXE		66800	396	374	466
CYCHXA		40900	367	349	428

Absorption maxima of aldehydes, SBs and PSBs were measured in ethanol. Protonation of SB was performed by addition of 37% aqueous HCl solution. **ε** is an extinction coefficient of the chromophore in ethanol.

retinal and the dehydro aldehyde in KL double mutant to KL:T53S (mutants from left to right in **Table II-5**) are 33, 34, 39, 37, and 32, nm respectively. The average enhancement with an additional double bond in the β-ionone ring seems to be 35 nm in these mutant series. Noteworthy, similar bathochromic shift is seen in protein/GEMLESS pigments. Removal of gem-dimethyl groups leads to the coplanar arrangement about the C6-C7 bond, resulting in maximal conjugation. Analysis of 5DEM hCRBPII complexes exhibits red-shifted spectra in comparison with the all-*trans*-retinal bound mutants but the resultant red shifts (~ 13 nm) remain smaller than the GEMLESS analog. This is perhaps due to the slightly deviated 5-desmehtyl ionone ring moiety in 6-s-*cis* conformation as opposed to a fully planar 6-s-*trans* geometry that is favored with the 1,1-didemethyl ring. These findings indicate that in the presence of the protein environment, modified retinal substrates are sufficiently well differentiated conformationally to allow for distinct absorption profiles.

Table II-5. Complexation of retinal and retinal analogs with hCRBPII mutants having Thr53 mutation.

Ligand	KL	KL:T53A	KL:T53L	KL:T53V	KL:T53S	Δλ
RET	508	512	496	501	504	0
DEHYDRO	541	546	535	538	536	35
GEMLESS	537	539	534	534	540	33
5DEM	518	520	517	517	515	13

Ligands (0.4 to 0.6 equivalents) were incubated with 20 μM of hCRBPII mutants in phosphate buffer saline (PBS) at room temperature. The absorption maximum of each protein-ligand complex is reported in nm. Δλ is calculated based on the average λ^{max} difference between retinal and retinal analogs. For example, the λ^{max} differences between retinal and dehydro retinal are 33, 34, 39, 37, and 32 nanometers for the mutants from left to right, respectively. The average of these values is 35 nm.

In pursuit of electrostatic modulations around the PSB region, Thr51 was replaced with a hydrophobic valine residue in the hope that decreasing the negative polarity near the iminium leads a red shift in absorption. Luckily, the resultant hCRBPII proteins with T51V mutation showed high binding affinity towards the retinoid ligands and formed stable colored pigments (**Table II-6**). In all of these instances, mutants carrying valine at position 51 afforded bathochromically shifted complexes, similar to the results with all-*trans*-retinal (KL versus KL:T51V mutant). In contrast to the spectra resulting from

Table II-6. The effect of T51V substitution on the absorption of hCRBPII-retinal analog complexes.

Ligand	Position 51	KL	KL:T53A	KL:T53L	KL:T53V	KL:T53S	Δλ
RET	T	508	512	496	501	504	29
	V	538	536	*~512*	*~519*	533	
DEHYDRO	T	541	546	535	538	536	35
	V	575	582	563	570	582	
GEMLESS	T	537	539	534	534	540	25
	V	557	564	557	560	569	
5DEM	T	518	520	517	517	515	25
	V	537	545	546	542	541	

Ligands (0.4 to 0.6 equivalents) were incubated with 20 μM of hCRBPII mutants in PBS at room temperature. The absorption maximum of each protein-ligand complex is reported in nm. Δλ is calculated based on the average λ^{max} difference before (black) and after T51V (red) mutation. The average Δλ value is calculated by excluding the KL:T51V:T53L and KL:T51V:T53V data (italic red) as mentioned in the text.

complexation with DEHYRDO, GEMLESS and 5DEM ligands, all-*trans*-retinal yielded an asymmetric absorption band when bound to KL:T51V:T53V and KL:T51V:T53L, suggesting an overlap of multiple peaks. Deviation from the

symmetry is more pronounced in case of T53L in which different rotameric forms of the leucine side chain could presumably causes changes to the absorption profile of the bound ligand. The enhancement induced by T51V mutation for each chromophore appeared to be consistent as the largest red shift is observed in the DEHYDRO bound pigments (average of 35 nm increase). Specifically, the installation of T51V into the KL:T53S mutant remarkably shifts the absorption by 46 nm for the DEHYDRO bound protein, whereas the same change only results in ~29 nm red shift in other three ligands, as shown in **Table II-6** (KL:T53S versus KL:T51V:T53S).

Table II-7. Comparison of absorption maxima of all-*trans*-retinal and its analogs bound to hCRBPII mutants.

Ligand	KL:T51V	KL:T51V:T53A	KL:T51V:T53S	Δλ
RET	538	536	533	0
DEHYDRO	575	582	582	44
GEMLESS	557	564	569	28
5DEM	537	545	541	5

Ligands (0.4 to 0.6 equivalents) were incubated with 20 μM of hCRBPII mutants in PBS at room temperature. The absorption maximum of each protein-ligand complex is reported in nm. Δλ is calculated based on the average λ^{max} difference between retinal and retinal analogs.

Further analysis of the KL:T51V:T53X (X=T,A,L,V, and S) hCRBPII series reveals that the increase of hydrophobicity around the PSB results in perturbation of conjugation in the bound ligands. As highlighted in **Table II-5**, the average enhancement of an additional double bond in the β-ionone ring is ~35 nm in those mutant series (RET versus DEHYDRO). Nevertheless, this effect is increased to 44 nm in the presence of T51V mutation (**Table II-7**, RET versus

DEHYDRO). In sharp contrast, the average λ^{max} difference between all-*trans*-retinal and the other two analogs, 5DEM and GEMLESS, was not increased in the presence of T51V mutation (Δλ values in **Table II-5** versus **Table II-7**). To illustrate, 5DEM displays ~13 nm red-shifted spectra in KL:T53X hCRBPII series (X=T,A,L,V, and S) as compared to all-*trans*-retinal, however the observed red shift is about half of that (~5 nm) when T51V is introduced. Likewise, the red shift accounting for the hCRBPII-KL:T53X/GEMLESS complexes as compared to all-*trans*-retinal bound equivalents decrease from 33 nm to 28 nm upon T51V replacement. This shows that the increase of hydrophobicity around the PSB (with T51V mutation) also influences the relative λ^{max} difference between the retinal and the retinal analogs.

While a substantial red shift was achieved by installation of T51V mutation, further reduction of the polar environment close to the PSB would require the disruption of the water network residing around this region. As described in our previous design of hCRBPII rhodopsin mimics, this water network was observed in many crystal structures and involves conserved sets of amino acid residues (**Figure II-9**). The first water molecule (W1 in **Figure II-9a**) in the network hydrogen bonds to the iminium nitrogen and to the carbonyl of Gln4, contributing to stabilization of the positive charge on the iminium. On the other hand, the second water molecule (W2) makes hydrogen bonds with the main chain carbonyls of Thr1 and Asp91 along with the amide nitrogen of Ile110. This

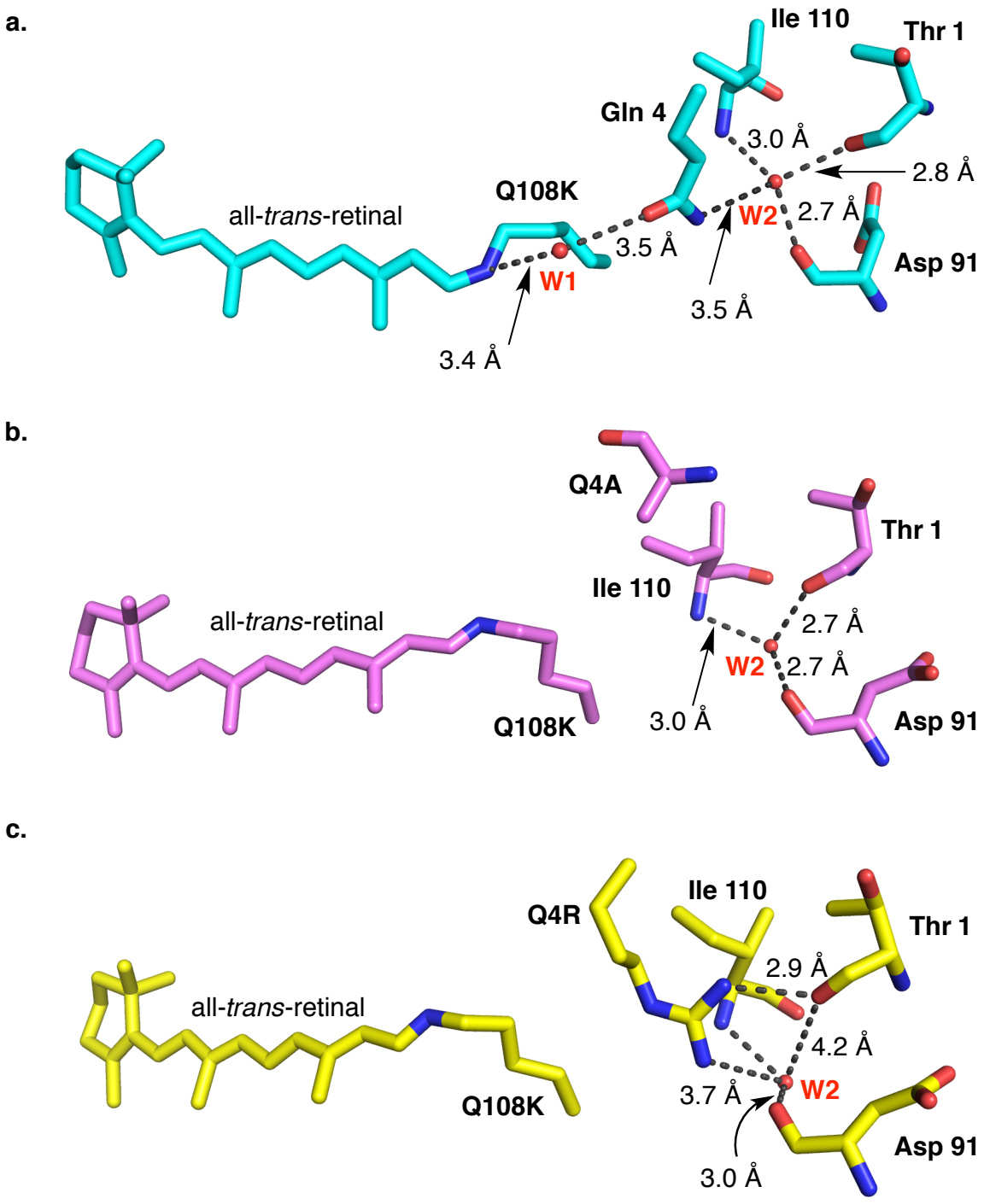


Figure II-9: A comparison of the interacting water networks in hCRBPII mutants. **a.** Crystal structure of KL:T51V:T53C:Y19W:R59W:T29L hCRBPII (cyan, PDB ID: 4EFG) with highlighted residues to show the water network extending to the iminium nitrogen. Water molecules are indicated as red spheres. **b.** Crystal structure of KL:T51V:T53C:Y19W:R59W:T29L:Q4A hCRBPII (pink, PDB ID: 4GKC). **c.** Crystal structure of KL:T51V:T53C:Y19W:R59W:T29L:Q4R hCRBPII (yellow, PDB ID: 4EEJ). The mutation of Gln4 results in the loss of an ordered water molecule (W1) in structures of **b** and **c**.

network, assisted by W2, orients the side chain of Gln4 side chain. Mutations of Gln4 (to Phe, Ala, Trp, Leu, Asn, and Thr) had a dramatic effect on the spectroscopic behavior of hCRBPII/retinal complexes, putatively as a result of in the loss of the ordered water molecule (W1) that destabilizes the protonation of iminium nitrogen. All of the Gln4 variants afforded red-shifted retinylidene bound protein complexes indicating an enhanced distribution of resonating charge along the polyene. Satisfyingly, obtaining crystallographic data provided detailed structural information about these mutants. Close inspection of the crystal structures of hCRBPII variants incorporating Q4A and Q4R replacements show that the absence of the W1 leads to a configurational change of the resultant iminium geometry. While hCRBPII complexes with all-*trans*-retinal favor the *cis*-iminium geometry in the presence of an ordered W1 molecule as shown in **Figure II-9a**, *trans*-imine geometry is adopted in all Gln4 modified mutants (**Figure II-9b** and **c**). Although, the effect of this configurational change is not clear on wavelength regulation, displacement of the Gln4 is crucial to improve the conjugation along the polyene.

The generality of the effect of Gln4 mutation was also investigated with synthetic retinal analogs bound to Q4F modified hCRBPII mutants. The resultant complexes yielded iminiums with low pK_a as compared to the parent mutants. As examples, the UV-vis spectra of KL:T51V:T53A:Q4F hCRBPII are illusrated in **Figure II-10** for each retinal analog. Although the amount of protonated Schiff base is lower as compared to Schiff base, acidification of the protein solution

facilitates the protonation of the imine nitrogen and increases the absorbance of the PSB peak. Likewise, other Q4F mutants (**Table II-8**) also fared poorly in formation of abundant PSB peak at physiological pH. Depressed pK_a of the iminium in such complexes is presumably the result of poor stabilization of the charge in the absence of the water network (W1 in **Figure II-9**). But despite the lowered pK_a, installation of Q4F remarkably induces a red shift in the resultant protein-ligand complexes. **Table II-8** lists the change in absorption as a result of the Q4F substitution for each chromophoric system. Protein complexes generated by the DEHYDRO ligand exhibit the most bathochromic shift in the presence of Phe4 similar to that seen with T51V mutated pigments (**Table II-6**). Obviously, both Gln4 and Thr51 residues, located around the iminium, strong

Table II-8. The effect of Q4F substitution on absorption of hCRBPII-retinal analog complexes.

Ligand	Position 4	KLV	KLV: T53A	KLV: T53L	KLV: T53V	KLV: T53S	Δλ
RET	Q	538	536	*~512*	*~519*	533	33
	F	569	571	567	567	567	
DEHYDRO	Q	575	582	563	570	582	42
	F	610	622	614	614	622	
GEMLESS	Q	557	564	557	560	569	26
	F	580	591	586	585	594	
5DEM	Q	537	545	546	542	541	28
	F	562	575	575	569	569	

Ligands (0.4 to 0.6 equivalents) were incubated with 20 μM of hCRBPII mutants in PBS at room temperature. The absorption maximum of each protein-ligand complex is reported in nm. KLV stands for the KL:T51V hCRBPII mutant. Δλ is calculated based on the average λ^{max} difference before (black) and after Q4F (red) mutation. The average Δλ value is calculated by excluding the KL:T51V:T53L and KL:T51V:T53V data (italic black) as mentioned in the text.

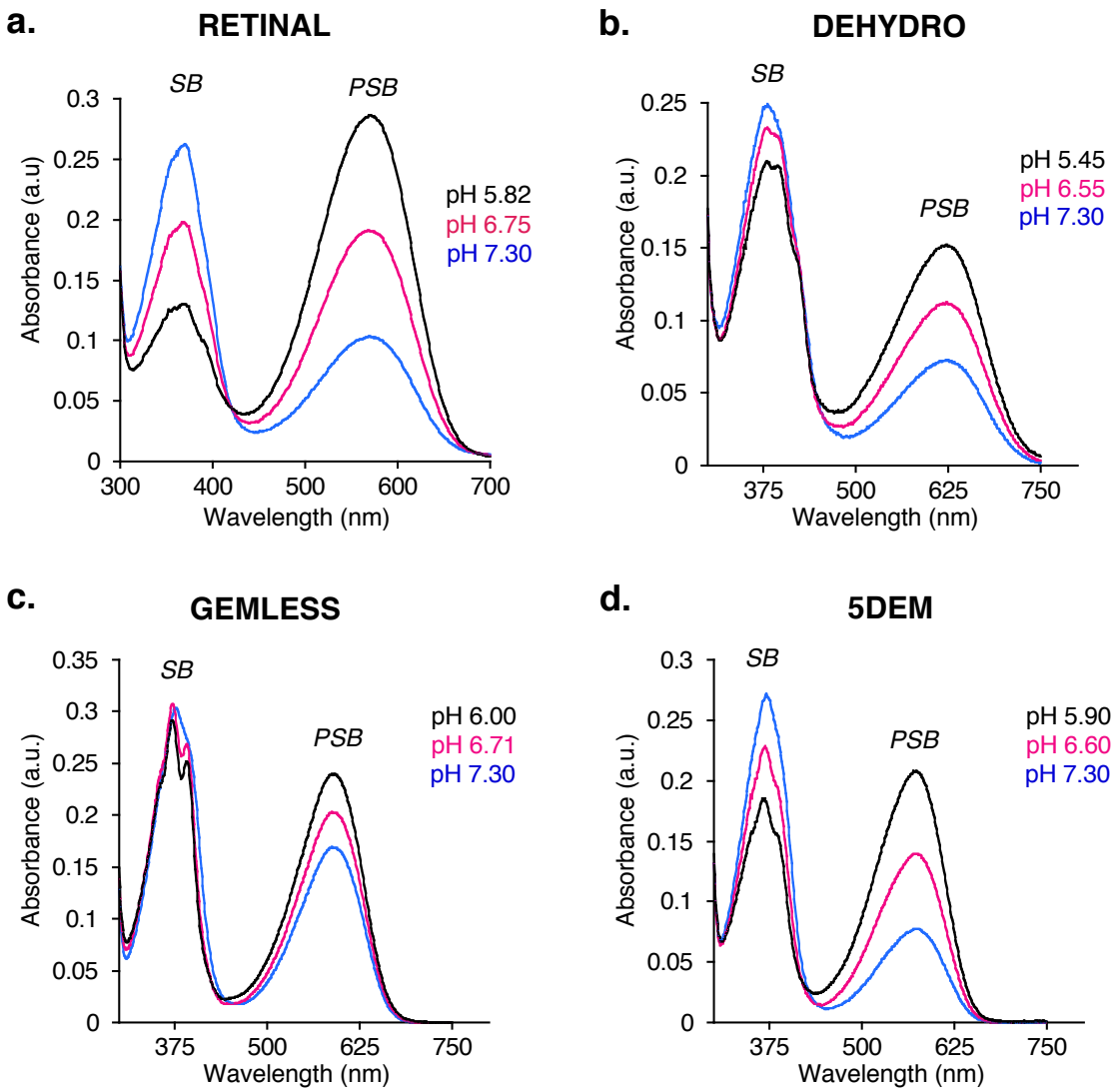


Figure II-10: UV-vis spectra of KL:T51V:T53A:Q4F hCRBPII complexed with retinal and its derivatives. The change in absorbance is monitored upon acidification of the protein-ligand solution (pH values are indicated for each spectrum). The absorption of SB and PSB peaks of the ligand bound protein are shown for all-*trans*-retinal, DEHYDRO, GEMLESS, and 5DEM in **a** to **d**, respectively.

effects on charge stabilization. Removal of these elements results in insufficient support to localize the positive charge on the iminium nitrogen and promotes charge distribution along the ligand. One would expect the largest bathochromic shift for chromophores in which the structure is more conjugated, thus allowing a

better propagation of the cationic charge. As such, it is not surprising that hCRBPII/DEHYDRO complexes, having a highly conjugated ligand, display the greatest sensitivity towards mutational changes. Additionally, comparison of the synthetic retinal analogs with all-*trans*-retinal provides further evidence for fact that the DEHYDRO ligand has an exceptionally high degree of charge delocalization, yielding the most red-shifted pigments amongst all the ligands (**Table II-9**). It should be also emphasized that **Table II-7** and **Table II-9** highlight the consistent order of λ^{max} differences between all-*trans*-retinal and its analogs. Seemingly, these sets of mutants (KLV:T53X and KLV:Q4F:T53X series) share similar binding modes for the associated ligands and only vary with respect to electrostatic environments.

Table II-9. Comparisons of absorption maxima of all-*trans*-retinal and its analogs in T51V and Q4F mutated hCRBPII variants.

Ligand	KLV: Q4F	KLV:T53A: Q4F	KLV:T53L: Q4F	KLV:T53V: Q4F	KLV:T53S: Q4F	Δλ
RET	569	571	567	567	567	0
DEHYDRO	610	622	614	614	622	48
GEMLESS	580	591	586	585	594	19
5DEM	562	575	575	569	569	2

Ligands (0.4 to 0.6 equivalents) were incubated with 20 μM of hCRBPII mutants in PBS at room temperature. The absorption maximum of each protein-ligand complex is reported in nm. Δλ is calculated based on the average λ^{max} difference between retinal and retinal analogs. KL:T51V hCRBPII is abbreviated as KLV.

II.4 Crystal Structures Of KL:T51V:T53S HCRBPII With Retinoids

In order to understand the structural elements that lead to the spectroscopic characteristics observed with our protein-ligand complexes, we

resorted to co-crystallize retinal and its derivatives with hCRBPII rhodopsin mimics. To this end, KL:T51V:T53S hCRBPII mutant yielded single crystals under optimized conditions with all four retinoid ligands. The crystal structure of KL:T51V:T53S/all-*trans*-retinal complex was obtained by Dr. Zahra Nossoni while the other sets of crystals with synthetic retinal analogs were collected collaboratively by Meisam Nosrati and Ipek Yapici. Close inspection of KL:T51V:T53S protein structure revealed a distinct reorganization of the T53S side chain with a nearby residue of Gln38 (**Figure II-11**). This arrangement has not been seen for other hCRBPII crystal structures, making the KL:T51V:T53S variant unique in its tertiary scaffold. In regular hCRBPII structures, a well-organized water mediated hydrogen bonding interaction between the Gln38 and the Gln128 side chains helps to position these amide residues in the protein cavity. As shown in **Figure II-11**, two glutamine side chains interact with each other via a water mediated hydrogen bond in the wild type hCRBPII. Likewise, KL:T53C and KL:T51V:T53C hCRBPII variants (Dr. Zahra Nossoni's data) exhibits the conserved interaction between Gln38 and Gln128 residues where two amide groups face to each other and form a bridge through the ordered water molecules (**Figure II-11c** and **11d**). Nonetheless, this favored interaction is suppressed when threonine is replaced with a serine residue at position 53. The newly installed Ser53 residue establishes tight hydrogen bonding interactions with the side chain carbonyl group of Gln38, making the amide moiety adopt a different trajectory as opposed to its original arrangement. Consequently, the

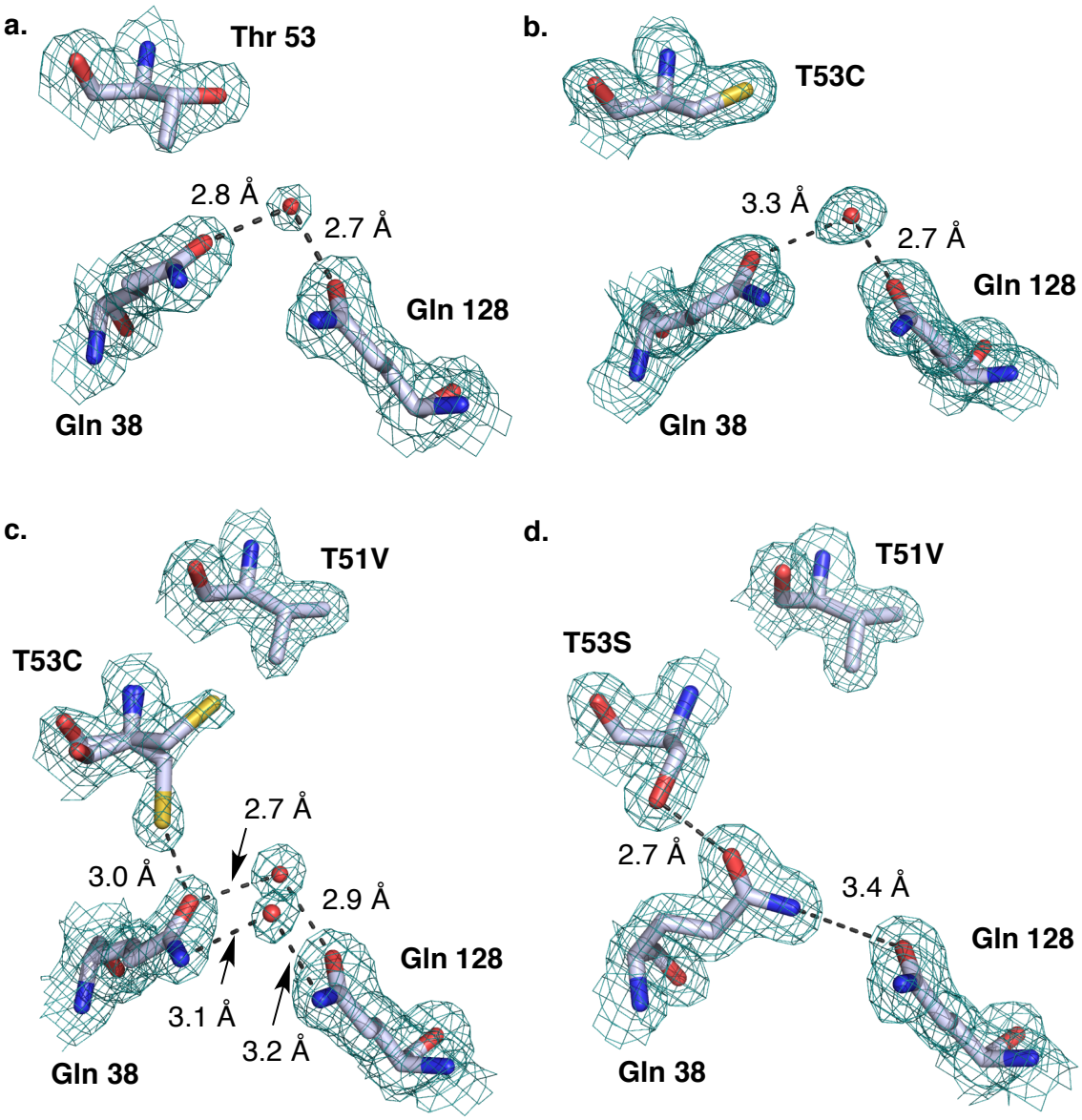


Figure II-11: Highlighted interactions of glutamine residues in hCRBPII. **a.** Crystal structure of wild type hCRBPII (PDB ID: 4QZT) showing Gln38, Gln128 and Thr53 residues. The 2Fo-Fc electron density map contoured at 1.0 sigma around the highlighted amino acids. The water molecule is indicated by the red sphere. **b.** Structure of KL:T53C depicts the conserved interaction between the mentioned glutamine residues. **c.** Crystal structure of KL:T51V:T53C reveals two conformations for the T53C side chain; one hydrogen bonds with the side chain carbonyl of Gln38. Two ordered water molecules assist the hydrogen bonding between the amide moieties of glutamine residues. **d.** Installation of the T53S mutation leads to a different organization of Gln38. The alcohol moiety and the carbonyl of the amide form a tight hydrogen bond and result in the expulsion of the water-mediated interaction with Gln128. Noticeably, the exact order of the side chain carbonyl and the amine moieties participating in hydrogen bonding is not certain.

water-mediated interaction between Gln38 and Gln128 is disturbed, resulting in loss of ordered water molecules as evident from the crystal structure.

Analysis of KL:T51V:T53S/retinoid complexes enabled us to identify structural features of significance for the bound ligands. Although the electron density of the protein scaffold is well defined for all KL:T51V:T53S complexes, it is unfortunate that poor densities are seen for the DEHYDRO and the GEMLESS ligands (**Figure II-12**). Despite clear evidence for the ligand occupying the binding pockets, we are unable to make judgments about the specific orientations of the ring moieties or the torsion angles for those ligand structures. Satisfyingly, all-*trans*-retinal and 5DEM chromophores display sufficient electron density for determining the conformations of their bound states. An overlay of retinal and 5DEM bound KL:T51V:T53S crystal structures revealed the flexibility of the binding pocket of the hCRBPII scaffold. As the tertiary folding features of the protein remain equivalent for both protein-ligand complexes, the mouth of the cavity consisting of two loops and the α-helicies displays considerable plasticity (see **Figure II-13a** red and blue locations). Further investigation into the binding site of the pigments highlights the distinct conformational preferences of the two ligands. Predictably, 5DEM chromophore adopts a 6-s-*cis* conformation about the C6-C7 single bond, having a slight torsion angle, $\psi_1 = 24^\circ$ (**Figure II-13b**). In sharp contrast, the same mutant yields a skewed 6-s-*cis* conformation for the bound all-*trans*-retinal with torsion angles departing significantly from planarity. As anticipated, the greater steric demand of the C5-CH_3 group in the retinal

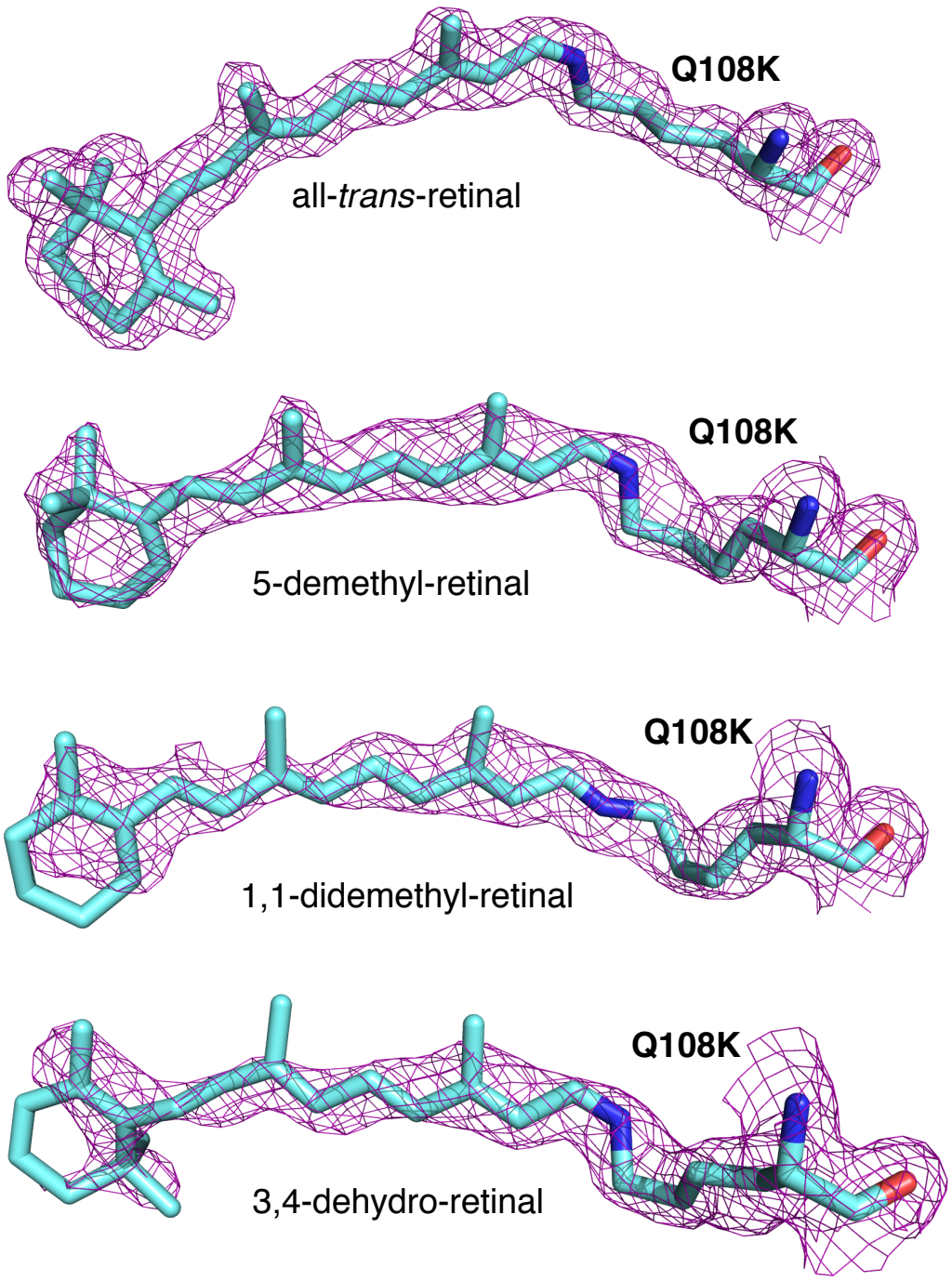


Figure II-12: Crystal structures of KL:T51V:T53S mutant with retinal and synthetic analogs. The 2Fo-Fc electron density map contoured at 1.0 sigma around the bound ligand and the active site Q108K residue. The resolutions of the crystal structures are 1.5 Å, 1.6 Å, 1.7 Å, and 1.6 Å from top to bottom structures, respectively.

structure creates unfavorable contact with the C8-H and induces a rotation of the

ionone ring about the C6-C7 single bond. The resultant twist of the C6-C7 bond

alleviates repulsive intramolecular contacts, and also causes a movement in the chromophore position within the binding site (**Figure II-13d**).

Structural analysis of the KL:T51V:T53S pigments suggests multiple possible mechanisms for the observed absorption characteristics. One important factor to consider is the electrostatic environment around the bound chromophore. The presence of polarizable groups or local polarities can lead to a change in the charge distribution across the polyene. In crystal structure of KL:T51V:T53S/5DEM complex, the ring moiety is tightly against Gln38 and Phe16 (**Figure II-13f**). The polar Gln38 side chain makes close contact with the C5 of the chromophore (3.4 Å away), and possibly stabilizes the delocalized positive charge on the ring double bond. However, the twisted ring arrangement of the all-*trans*-retinal increases the space between the Gln38 side chain and the chromophore (4.2 Å to C8 and C5 carbons), where a putative partial positive charge resonating from the iminium would be anticipated (**Figure II-13e**). This presumably results in a weak stabilization of the cationic chromophore and allows even distribution of the positive charge.

The PSB of retinal absorbs at 444 nm in solution while the subsequent pigment with KL:T51V:T53S exhibits a λ^{max} of 533 nm ($\Delta\lambda$ = 89 nm, opsin shift is 3761 cm^{-1}). On the other hand, the absorption maximum of the PSB of 5DEM chromophore is 453 nm and the resultant KL:T51V:T53S complex shows an absorption maximum at 541 nm ($\Delta\lambda$ = 88 nm, opsin shift is 3591 cm^{-1}). Another

key factor to take into account is the planarity of the chromophore π-system. Considering the fact that some part of an opsin shift is arising from the planarization, one would expect to see a smaller opsin shift for the 5DEM chromophore due to its higher planarity in solution as opposed to all-*trans*-retinal. However, the crystal structure of the KL:T51V:T53S/RET complex reveals that the double bond at the ionone ring is almost out of conjugation relative to the polyene chain of the retinal (**Figure II-13c**). Therefore, it is hard to consider the chromophore planarization as a contributing factor in the observed opsin shift for all-*trans*-retinal. Namely, the major component of the induced red shift in retinal absorption should be the changes in electrostatic environment around the ligand inside the protein cavity. Likewise, the observed opsin shift of 5DEM would be considered as the result of electrostatic interactions because the chromophore possesses a planar conformation both in solution and in protein cavity (**Figure II-13b**). Consequently, it is not surprising that the difference in the opsin shifts of 5DEM and retinal, bound to the same mutant, is small since they did not display substantial conformational changes from the solution to the protein binding pocket.

II.5 Elucidating The Effect Of Arg58 Mutation

As stated earlier, embedding and encapsulation of the ligand is necessary in order to reach super red-shifted protein/chromophore systems. Prior studies have shown that two key amino acids (Arg58 and Ala33), residing at the mouth of the binding cavity must be changed into hydrophobic aromatic residues in order

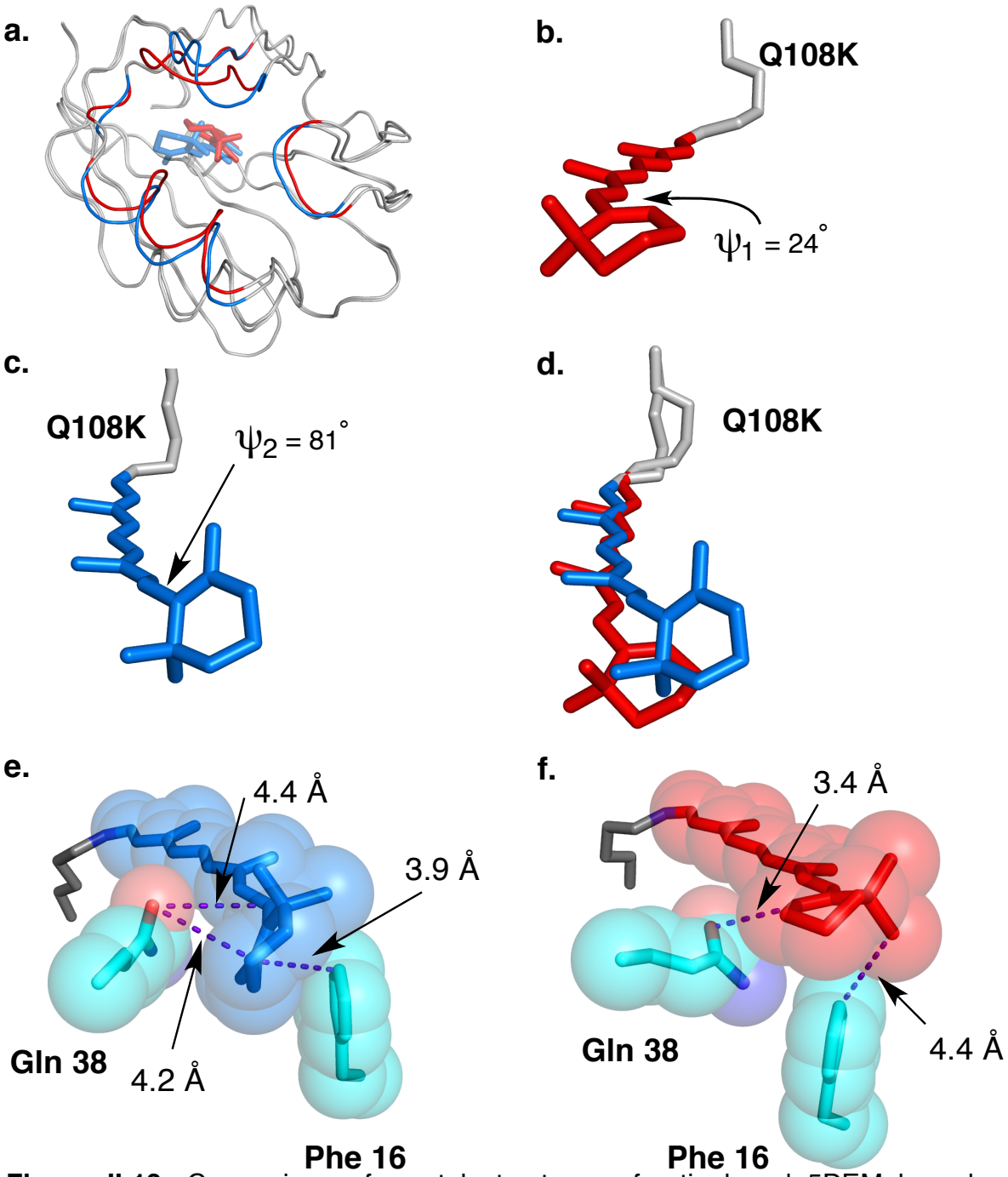


Figure II-13: Comparison of crystal structures of retinal and 5DEM bound KL:T51V:T53S hCRBPII complexes. **a.** Overlaid crystal structures indicate the rearranged (blue and red) versus conserved (gray) regions in the protein fold. **b.** Expanded view of protein cavity shows the planar geometry of the 5DEM chromophore. **c.** Skewed 6-s-cis conformation of the bound retinal is indicated in stick model. **d.** The structures of all-*trans*-retinal (blue) and 5DEM (red) ligands are overlaid, highlighting the geometrical differences of the chromophores. **e.** Space filling model of retinal (blue) and nearby residues of Gln38 and Phe16 depicts the packing interactions around the ionone ring. **f.** Same representation shown in **e** is illustrated for the 5DEM chromophore.

to isolate the bound chromophore from the bulk aqueous media. It was found that the combination of R58W and A33W mutations has the greatest impact on sequestering the binding pocket, resulting in larger red shifts in spectrum of the all-*trans*-retinal. Several crystal structures of hCRBPII variants that incorporate R58F, R58Y and R58W mutations were previously collected and analyzed with respect to the side chain motions of the installed aromatic residues and their ability to enclose the ligand binding site. Collected structural information revealed that the flexibility of the loop region allows different rotameric forms to of the Trp58 and Phe58 side chains. However, a conserved motif in terms of the side chain conformation was seen in many cases with the Tyr58 mutation (**Figure II-14b**). The phenolic group makes a water-mediated hydrogen bond with the side chain hydroxyl of the Thr29 and stabilizes the location of the aromatic Tyr residue (**Figure II-14c**). In hCBRPII structures with T29L, there is a direct hydrogen bond observed in between the phenolic group and the backbone carbonyl of Leu29. (**Figure II-14d**). Having this structural information in hand, we initiated our investigations with enclosed and sequestered binding pocket (R58Y mutants) with synthetic retinals. In this manner, little or no conformational issues would be anticipated for the R58Y replacement and direct electrostatic changes projected on the bound chromophores could be examined. At the outset, the effect of the R58Y mutation was tested on KL:T51V:T53X (X= T, A, L, V, and S) series of hCRBPII. Similar to previously obtained results from the hCRBPII/retinal complexes, the substitution of Arg58 with Tyr substantially shifts the spectra of

all-*trans*-retinylidene chromophore to longer wavelengths. As summarized in **Table II-10**, the absorption maxima of each protein-retinal complex is red-shifted ~37 nm as a result of the R58Y mutation. Notably, DEHYDRO bound hCRBPII pigments exhibit the same magnitude of red shifting in absorption, showing that both chromophores respond in a similar fashion to the change of Arg58. These findings fit well with the previously proposed hypothesis emphasizing that the

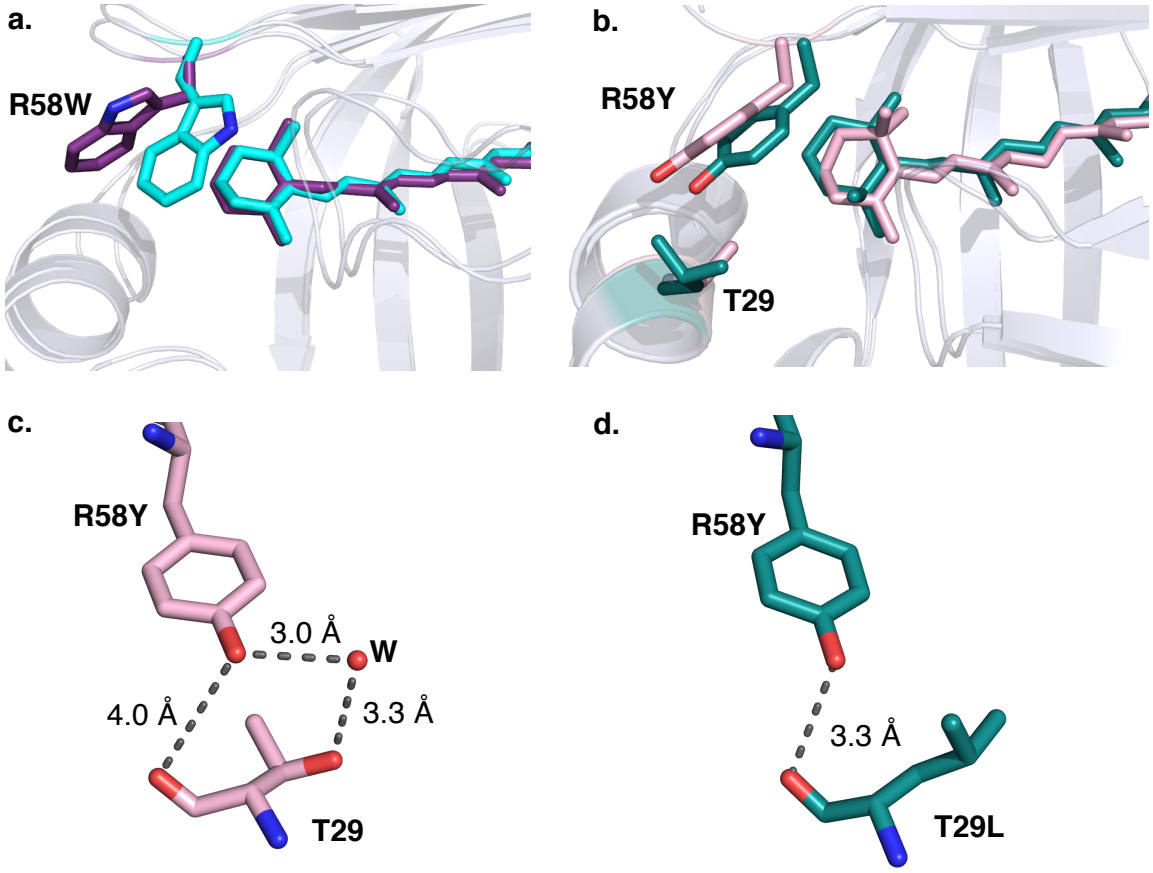


Figure II-14: Side chain orientations of Arg58 mutations. **a.** Illustration of different R58W rotomers in hCRBPII crystal structures. Overlaid structures show the alternative conformations of the R58W residue. **b.** Limited plasticity of the R58Y residue in different crystal structures of hCRBPII mutants, adopting a conserved orientation. **c.** Water mediated hydrogen bonding interactions between R58Y and Thr29 residues. **d.** The installation of T29L mutation removes the water mediated interactions seen in **c**, but gains a hydrogen bond with a backbone carbonyl of T29L.

embedded chromophore in an isolated protein binding pocket displays an enhanced charge distribution and a bathochromically shifted absorption profile. In sharp contrast, our new findings from the two other analogs, 5DEM and GEMLESS, are not consistent with the suggested mechanism. The enhancement, as a result of the R58Y substitution is small in GEMLESS bound proteins, while a subtle hypsochromic shift is seen for most cases with hCRBPII/5DEM complexes. If the major role of the aromatic residues placed instead of Arg58 is to enclose the binding pocket and create a buffer-isolated environment around the ligand, one could argue that the overall impact should be comparable for all bound chromophores regardless of their ring modifications. Nevertheless, upon inspection of the data, it becomes apparent that the R58Y mutation does not always leads a red shift in absorption of the bound ligands.

Table II-10. The effect of R58Y substitution on the absorption of hCRBPII-retinal analog complexes.

Ligand	Position 58	KLV	KLV: T53A	KLV: T53L	KLV: T53V	KLV: T53S	Δλ
RET	R	538	536	*~512*	*~519*	533	37
	Y	565	576	565	574	576	
DEHYDRO	R	575	582	563	570	582	36
	Y	603	615	605	612	616	
GEMLESS	R	557	564	557	560	569	7
	Y	565	572	564	568	575	
5DEM	R	537	545	546	542	541	-1
	Y	540	541	537	547	541	

Ligands (0.4 to 0.6 equivalents) were incubated with 20 μM of hCRBPII mutants in PBS at room temperature. The absorption maximum of each protein-ligand complex is reported in nm. KLV is abbreviated for KL:T51V hCRBPII mutant. Δλ is calculated based on the average λ^{max} difference before (black) and after R58Y (red) mutation. The average Δλ value is calculated by excluding the KL:T51V:T53L and KL:T51V:T53V data (italic black).

Further support for the latter statement has also come from the studies on hCRBPII mutants R58Y was tested in combination with Q4F mutation. In these protein series, the R58Y mutation yields red-shifted spectra for the DEHYDRO and the all-*trans*-retinal bound pigments (with almost same bathochromic shift), but fails to the same for the GEMLESS and 5DEM ligands. It is surprising that chromophores, which lack methyl substituents on the ring, show distinct absorptive properties than the ones having a canonical ionone ring in their structures.

In order to answer why the R58Y mutation has differential effects on the absorption of ligands as a function of the methyl substituents, several factors should be considered. To start with, both CYCHXE and CYCHXA were coupled with the hCRBPII mutants listed in **Table II-10** and **Table-11** in order to verify the role of encapsulation of the ligand in a closed cavity. Noticeably, these two

Table II-11. The effect of R58Y substitution on the absorption of hCRBPII-retinal analog complexes.

Ligand	Position 58	KLV: Q4F	KLVA: Q4F	KLVL: Q4F	KLVV: Q4F	KLVS: Q4F	Δλ
RET	R	569	571	567	567	567	22
	Y	579	595	592	591	592	
DEHYDRO	R	610	622	614	614	622	20
	Y	622	642	639	634	636	
GEMLESS	R	580	591	586	585	594	3
	Y	581	594	594	587	593	
5DEM	R	562	575	575	569	569	-4
	Y	559	570	566	568	567	

Ligands (0.4 to 0.6 equivalents) were incubated with 20 μM of hCRBPII mutants in PBS at room temperature. The absorption maximum of each protein-ligand complex is reported in nm. KLVX is abbreviated forKL:T51V:T53X hCRBPII mutants, where X can be T,A,L,V, and S. Δλ is calculated based on the average λ^{max} difference before (black) and after R58Y (red) mutation.

chromophores have no methyl groups on the ring and only differ in the number of double bonds. As both ligands share similar isosteric interactions, the subsequent effect of the R58Y mutation on the absorption spectra would be attributed to the direct consequence of the polarity changes in the binding pocket. Surprisingly, however, spectral analysis of CYCHXE and CYCHXA bound hCRBPII mutants revealed that there is not much wavelength modulation that could be achieved by the R58Y mutation (**Table II-12**). In most cases, the change in the absorption maximum stays within few nanometers, implying that the installation of aromatic residues at the mouth of the protein cavity does not serve as a key element to facilitate charge distribution across the bound ligand. Howbeit these results contradict prior work in which aromatic amino acids substituted for Arg58 bathochromically shift the absorption maxima of all-*trans*-retinal bound pigments. In like manner, hCRBPII/DEHYDRO complexes also exhibit red-shifted spectra as a result of the R58Y mutation (**Table II-10** and **Table II-11**). In these retinyl chromophores, a β-ionone ring has three methyl substituents that are influenced by the R58Y mutation in a manner that red shifts the wavelength. It is intriguing to consider factors that might dictate this unique interaction. There are certainly a large number of specific interaction that can be proposed in this context. One plausible model is that the wavelength is varied by the regulation of flexibility around the ring. The R58Y mutation could lead to protein structural changes that restrict motion of the ligand due to interactions between the three methyl groups in the β-ionone ring and the surrounding

residues. The afforded rigidity could orient the chromohore for better charge delocalization. Yet, chromophores lacking ring methyl substituents would be able to fall into the available space in the protein binding pocket, exploring structural freedom as a result of less steric constraints. Thus, presence of the large aromatic amino acids at Arg58 would presumably be less effective in rigidifying the more flexible chromophores, leading to subtle changes in wavelength. A second type of model can involve specific electrostatic interactions between the

Table II-12. Spectral analysis of CYCHXE and CYCHXA bound hCRBPII complexes.

Ligand	Position 58	KLV	KLVA	KLVL	KLVV	KLVS	Δλ
CYCHXE	R	543	546	546	544	548	2
	Y	544	552	540	547	552	
CYCHXA	R	494	496	494	493	496	-2
	Y	489	493	488	493	491	
		KLV: Q4F	**KLVA: Q4F**	**KLVL: Q4F**	**KLVV: Q4F**	**KLVS: Q4F**	
CYCHXE	R	565	570	572	569	571	-4
	Y	560	570	567	564	566	
CYCHXA	R	510	514	512	510	512	-10
	Y	499	505	500	504	501	

Ligands (0.4 equivalents) were incubated with 20 μM of hCRBPII mutants in PBS at room temperature. The absorption maximum of each protein-ligand complex is reported in nm. KLVX is abbreviated for KL:T51V:T53X hCRBPII mutants where X can be T,A,L,V, and S. Δλ is calculated based on the average λ^{max} difference before (black) and after R58Y (red) mutation.

chromophore and the side chain of R58Y. It is likely that the installation of R58Y residue in close proximity of the β-ionone ring creates strong hydrophobic-hydrophobic contacts between the three methyl groups and the aromatic ring. On the other hand, the chromophores having less methyl substituents perhaps are

not as tightly embedded in the protein as the ones with a canonical β-ionone ring. This packing organization could be a direct manifestation of the hydrophobic-hydrophobic interactions. Simply stated, our present study through site-directed mutagenesis highlights the significance of ring methyl substituents in wavelength regulation. However, at this point the major contributing factors (sterics or electrostatics) in wavelength modulation of R58Y-hCRBPII variants is not clear. Moreover, the inability to regulate the wavelength with chromophores that lack the methyl groups in the ring remains a mystery.

II.6 Structural Analysis On R58Y-hCRBPII Retinoid Pigments

Next, we crystallize a series of hCRBPII mutants that contained the R58Y mutation bound to the synthetic retinals. Our initial crystallization attempt was with KL:T51V:T53S:R58Y hCRBPII mutant that led to several crystals under previously established crystallization conditions. Subsequent refinement of the data obtained from these crystals did not show an ordered conformation for the bound ligands, mostly yielding poor electron density for the chromophores (data not shown). We then focused on crystallizing of KL:T51V:Y19W:R58Y hCRBPII with the retinal analogs since we already had good structural data with this mutant complexed with all-*trans*-retinal (obtained and solved by Dr. Zahra Nossoni and Meisam Nosrati). Using the KL:T51V:Y19W:R58Y mutant, we have successfully crystallized and solved the structures of three retinal analogs (DEHYDRO, GEMLESS, and 5DEM) that were previously unattainable. As discussed earlier (Section II.2), changes in Try19 do not have a significant

influence on wavelength regulation of the bound all-*trans*-retinylidene chromophore. This was also tested on modified retinals in presence and absence of the R58Y mutation (**Table II-13**). Predictably, substitution of Try19 with a Trp residue had little effect on absorption maxima of the pigments (~ 4 nm), which is consistent being in line with the previously obtained data.

Table II-13. The effect of R58Y substitution on the absorption of hCRBPII retinal analog complexes.

Ligand	Position 58	KLV	KLV:Y19W	Δλ
RET	R Y	538 565	538 565	27
DEHYDRO	R Y	575 603	577 604	28
GEMLESS	R Y	557 565	557 569	10
5DEM	R Y	537 540	539 544	-4
CYCHXE	R Y	543 544	540 545	3
CYCHXA	R Y	494 489	494 492	-4

Ligands (0.4 to 0.6 equivalents) were incubated with 20 μM of hCRBPII mutants in PBS at room temperature. The absorption maximum of each protein-ligand complex is reported in nm. KLV is abbreviated for KL:T51V hCRBPII mutant. Δλ is calculated based on the average λ^{max} difference before (black) and after R58Y (red) mutation.

Gratifyingly, all chromophore bound structures of KL:T51V:Y19W:R58Y hCRBPII protein provided a well-defined electron density map of the ligand (**Figure II-15**). Close inspection of the active sites revealed that there are a few noticeable differences between the RET/DEHYDRO complexes versus GEMLESS/5DEM bound pigments. Despite the fact that all four ligands occupy

the same binding space, retinal and DEHYDRO chromophores reside slightly closer to the mouth of the cavity, whereas GEMLESS and 5DEM analogs are located more towards the interior of the protein binding pocket (**Figure II-16a**). Indeed, the closer proximity of retinal and the DEHYDRO ligands to the protein mouth could be result of better hydrophobic-hydrophobic contact with the aromatic R58Y side chain as compared to the GEMLESS and the 5DEM chromophores. Furthermore, two additional aromatic residues, Phe16 and

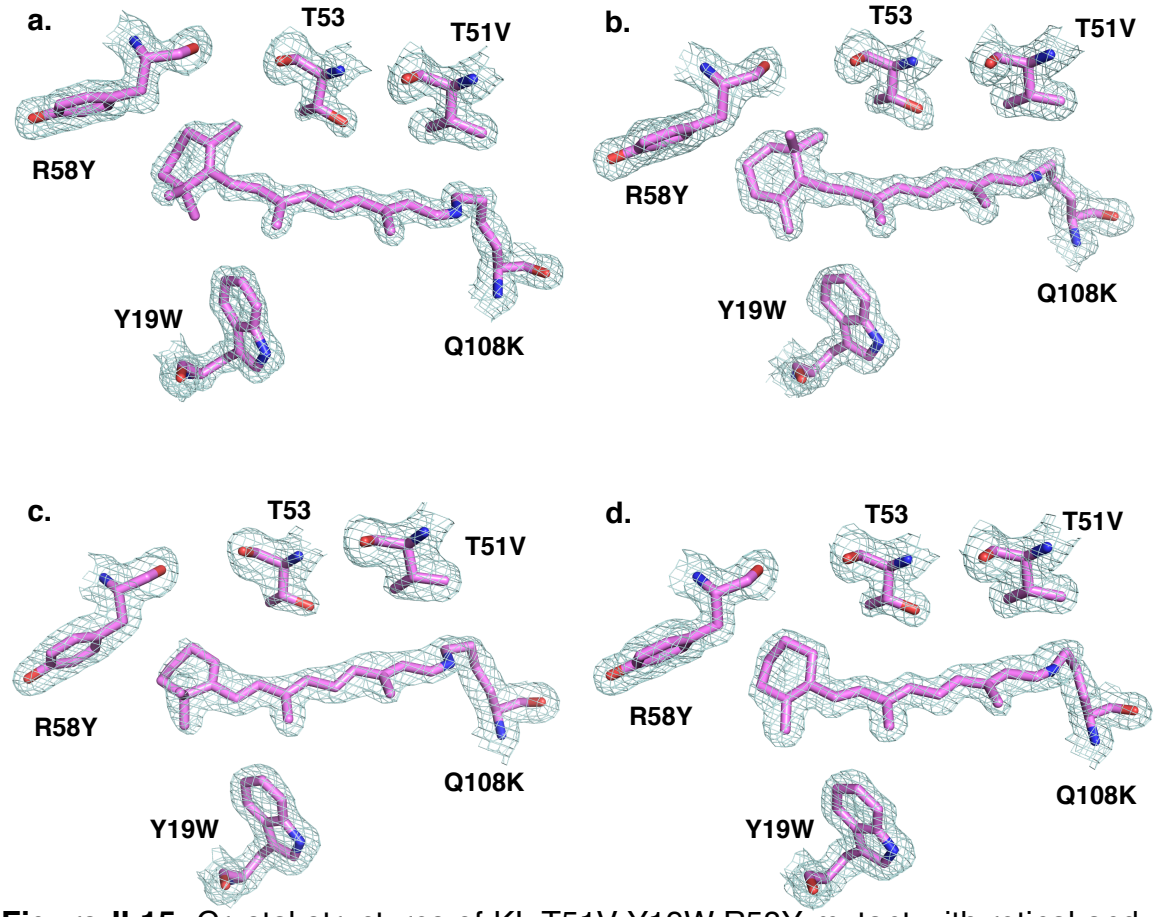


Figure II-15: Crystal structures of KL:T51V:Y19W:R58Y mutant with retinal and synthetic analogs. The 2Fo-Fc electron density map contoured at 1.0 sigma around the bound ligand and the highlighted residues. The resolutions of the crystal structures are 1.3 Å, 1.5 Å, 1.6 Å, and 1.5 Å for all-*trans*-retinal, DEHYDRO, 5DEM, and GEMLESS bound pigments (from **a** to **d**), respectively.

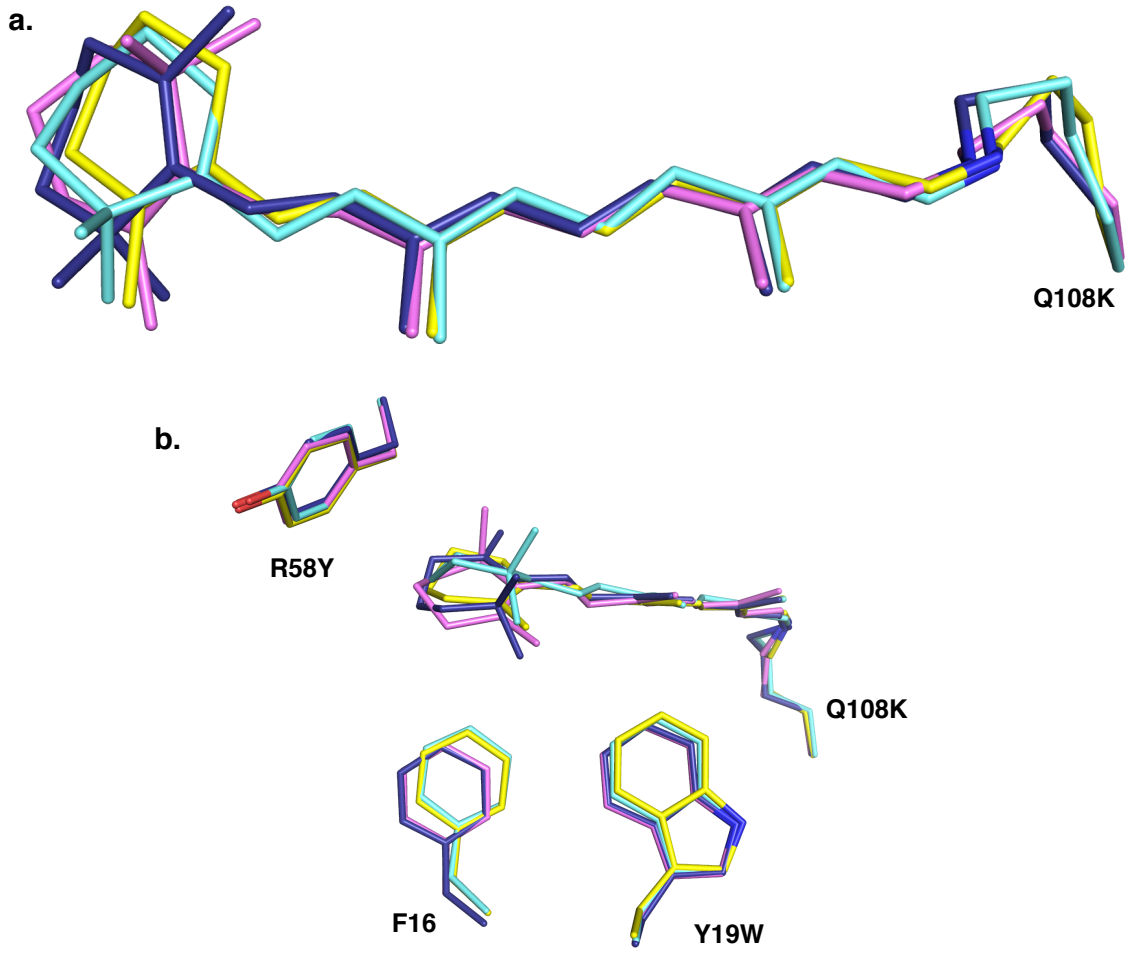


Figure II-16: Overlaid crystal structures of retinal chromophores. **a.** Stick model representations of RET, DEHYDRO, 5DEM, and GEMLESS ligands (indigo, magenta, cyan, and yellow respectively) in KL:T51V:Y19W:R58Y hCRBPII mutant. **b.** Side chain conformations of the aromatic residues around the bound ligand. Color assignments are the same as that indicated in **a**.

Trp19, in close proximity to the ring moieties of the bound ligands show little side chain movement. Side chain conformations of Phe16 and Trp19 residues are well aligned in the RET and DEHYDRO structures. Conversely, both residues display a sideways-tilted orientation in the other two crystal structures where the tilting is more pronounced for the Phe16 residue (**Figure II-16b**). Plausibly, these aromatic residues are important in providing tighter packing around the ring

moieties and also to provide optimal hydrophobic environments.

The data in **Table II-14** maps out the opsin shift as a function of ionone ring modifications. As anticipated, CYCHXA comprising of four double bonds and a bare cyclohexane ring exhibits the smallest opsin shift upon binding to hCRBPII mutant. Owing to the large dependence on the nature of the ring structure and the spectral shifts of the pigments studied, the opsin shift of CYCHXA serves as a reference point for the study of spectral contribution of modified ionone rings to the total opsin shift. Also, binding of CYCHXE with hCRBPII mutants red shifts the λ^{max} about 50 nm as compared to CYCHXA. The opsin shift for both ligands however, is similar. This is reasonable because both ligands share common ring substitutions, and when complexed with the same hCRBPII mutant, they are subjected to equivalent electrostatic environment.

In its complex with KL:T51V:Y19W:R58Y hCRBPII, all-*trans*-retinal displays an absorption band peaking at 565 nm, which is close to the λ^{max} of the GEMLESS bound complex. However, the same protein environment imposes much larger opsin shift on the retinal chromophore in comparison with the GEMLESS ligand. The dilemma of these findings is exemplified by the fact that the two chromophoric systems yield pigments absorbing closely, yet their calculated opsin shifts differ significantly. A possible explanation for this could be the planarization of the ring and the chain segments. If the GEMLESS chromophore exists in an identical conformation in free solution as in the protein

bound form, then the resulting opsin shift accounts for only electrostatic interactions between the protein and its ligand. On the other hand, conformation of retinal could be different in free versus bound state. In such a case the induced opsin shift would be a combination of electrostatic interactions and conformational changes. It has been shown previously that the all-*trans*-retinylidene iminium favors a twisted 6-s-*cis* conformation in solution.[86] In the protein bound state, an equivalent conformer for retinal was observed (**Figure II-**

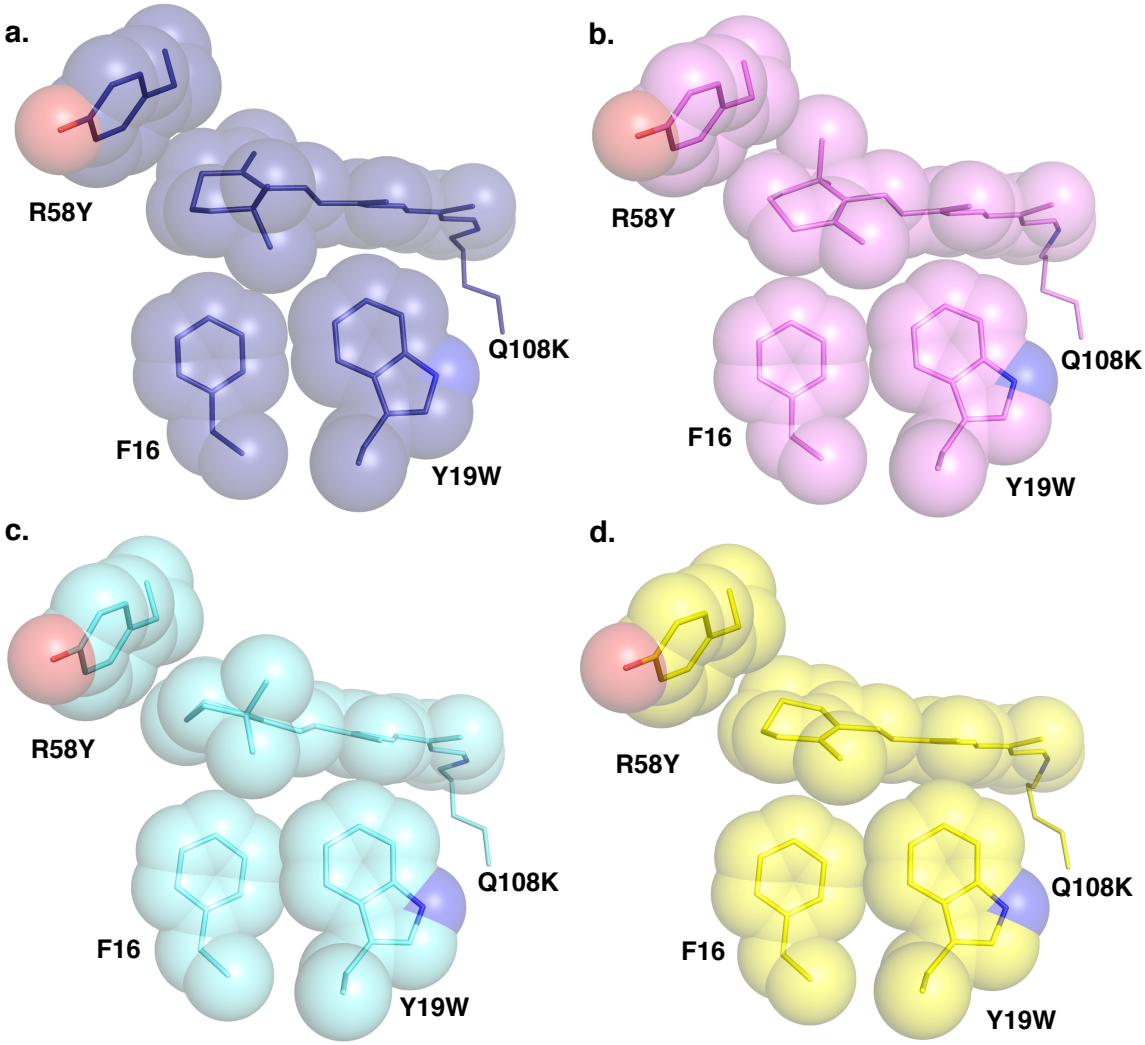


Figure II-17: Packing interactions are shown with space filling model of the chromophores and nearby aromatic residues for the structures of RET, DEHYDRO, 5DEM, and GEMLESS bound KL:T51V:Y19W:R58Y hCRBPII mutant in **a** to **d,** respectively.

18), revealing little to no change in the C6-C7 torsion. In this respect, planarization of the ring and polyene segments can be excluded from the calculated opsin shift for retinal. Likewise, the GEMLESS chromophore displays similar conformations in the protein binding pocket as in free solution (predicted by computational calculations) or in solid state crystal (Chapter IV, Figure IV-15). In all cases, the GEMLESS chromophore adopts the 6-s-*trans* conformation, leading to minimal conformational effects on the resulting opsin shift. From the latter experimental observations, it appears that the main contributor to the opsin shift for these chromophores is the result of electrostatic interactions between the

Table II-14. Opsin shifts of retinal chromophores in KL:T51V:Y19W:R58Y hCRBPII mutant.

Ligand	λ^{max} (nm) PSB	λ^{max} (nm) in protein	Opsin Shift (cm^{-1})	ΔOpsin Shift (cm^{-1})
RET	444	565	4823	1784
DEHYDRO	463	604	5042	2003
GEMLESS	478	569	3346	307
5DEM	453	544	3693	654
CYCHXE	466	545	3111	72
CYCHXA	428	492	3039	0

Absorption maxima of protonated *n*-butylamine Schiff bases (PSB) were measured in ethanol. Opsin shift is the difference in wavenumbers between the λ^{max} value of the PSB of the analog and that of the resulting hCRBPII pigment. The contribution of ring modifications to the total opsin shift (ΔOpsin Shift) is calculated by taking the difference of opsin shifts of hCRBPII/CYCHXA pigment and other pigments.

protein and its cognate ligand. This accentuates the fact that the relationship between the protein-ligand pair leading to a higher opsin shift is exceptionally fortified with the methyl substituents.

With regards to the ionone ring substituents, the trend in opsin shift can be summarized in **Table II-14**. It should be emphasized, however, that the observed trend is only apparent when aromatic residues are substituted for Arg58. Unfortunately, it is difficult to understand how the aromatic substitutions stimulate larger red shifts with chromophores that contain more methyl groups on the

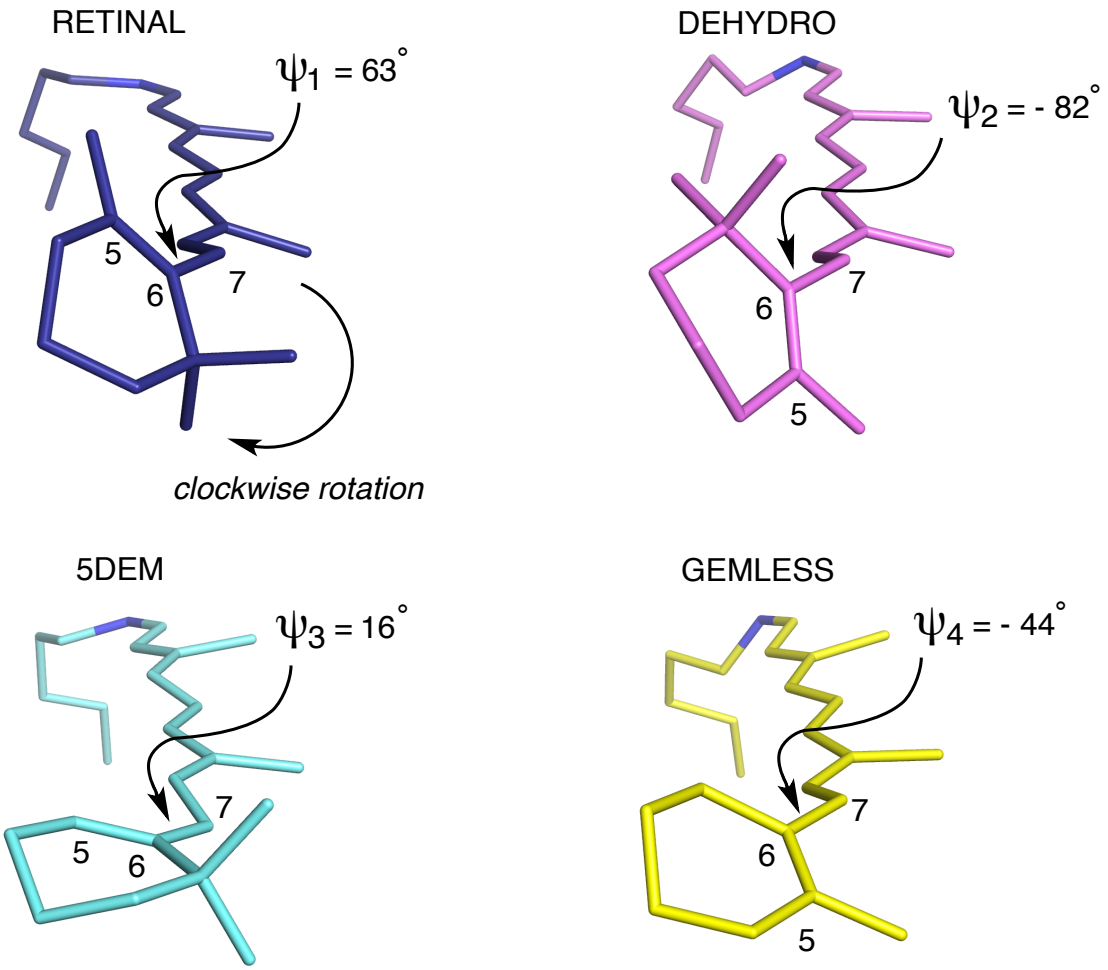


Figure II-18: Torsion about the C6-C7 bond observed with retinal and its analogs in KL:T51V:Y19W:R58Y hCRBPII crystals. The torsion angle (ψ) about a bond is defined as the angle made by the projection of the C5-C6 double bond with respect to the C7-C8 olefinic bond. When viewed down the C6-C7 single bond, positive angles (0 to + 180°) are for a clockwise rotation, as shown in thefigure, while negative angles (0 to -180°) are for a counterclockwise rotation. Zero angle is defined for an eclipsed position of the C5-C6 and C7-C6 double bonds in a 6-s-*cis* conformer.

ionone ring. Herein, our results have been essentially discussed in terms of reorientation of the ligands with the binding pocket, however, as pointed to previously, the influence of hydrophobic contacts around the chromophores could also be an important consideration.

II.7 Conclusion

The structure of retinal and related compounds is of interest because of the importance of these small molecules in molecular biology, particularly in the area of photosensitive opsin pigments. Several derivatives of retinal chromophore (naturally existing or synthetically modified analogs) have been studied with the hope to understand the physical and spectroscopic properties of the opsin proteins. In this coarse-grained study, we have concentrated on the latter aspect by utilizing our rationally designed hCRBPII rhodopsin mimics. With systematic approach to investigate protein-ligand interaction, we have shown that our mimics suffice to serve as testing grounds in which covalent modifications of the chromophoric molecules could be analyzed both spectroscopically and structurally within their interacting protein partners. We have evaluated diverse sets of hCRBPII-retinoid complexes with respect to their photophysical properties with the available high resolution crystal structures. Given the results reported here, we highlight the significance of ring methyl substituents in wavelength regulation not only as a handle for conformational issues, but also surprisingly as modulators for electrostatic interactions. From the experimental observations, it is clear that a single mutation, R58Y, was needed to extract differential response

from the modified chromophores, leading to the observed spectral shifts. These findings have led us to question our previous proposed mode of action for Arg58 mutants, where bathochromic shifts were attributed solely due encapsulating and sequestering of the chromophore within the protein binding pocket. We suggest that further insight by considering the electrostatic perturbations and steric consideration of the key factors.

REFERENCES

REFERENCES

1. Engel, A.; Gaub, H. E., "Structure and mechanics of membrane proteins" *Annu Rev Biochem* **2008**, *77*, 127.

2. Heijne, G. V., "Membrane-protein topology" *Nature Reviews* **2006**, *7*, 909.

3. Muller, D. J.; Wu, N.; Palczewski, K., "Vertebrate membrane proteins: Structure, function, and insights from biophysical approaches" *Pharmacol Rev* **2008**, *60*, 43.

4. Sachs, J. N.; Engelman, D. M., "Introduction to the membrane protein reviews: The interplay of structure, dynamics, and environment in membrane protein function" *Annu Rev Biochem* **2006**, *75*, 707.

5. Chitnis, P. R., "Photosystem I: Function and physiology" *Annu Rev Plant Phys* **2001**, *52*, 593.

6. Deamer, D. W. *Light Transducing Membranes Structure, Function, And Evolution* Academic Press, Inc. (London) Ltd., 1978.

7. Lodish, H.; Berk, A.; Zipursky, L.; Matsudaira, P.; Baltimore, D.; Darnell, J., "Molecular Cell Biology, 4th edition" *New York: W . H. Freeman* **2000**.

8. Nelson, N.; Yocum, C. F., "Structure and function of photosystems I and II" *Annu Rev Plant Biol* **2006**, *57*, 521.

9. Ernst, O. P.; Lodowski, D. T.; Elstner, M.; Hegemann, P.; Brown, L. S.; Kandori, H., "Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms" *Chem Rev* **2014**, *114*, 126.

10. Lamb, T. D.; Collin, S. P.; Pugh, E. N., "Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup" *Nat Rev Neurosci* **2007**, *8*, 960.

11. Porter, M. L.; Blasic, J. R.; Bok, M. J.; Cameron, E. G.; Pringle, T.; Cronin, T. W.; Robinson, P. R., "Shedding new light on opsin evolution" *P Roy Soc B-Biol Sci* **2012**, *279*, 3.

12. Shichida, Y.; Imai, H., "Visual pigment: G-protein-coupled receptor for light signals" *Cell Mol Life Sci* **1998**, *54*, 1299.

13. Shichida, Y.; Matsuyama, T., "Evolution of opsins and phototransduction" *Philos T R Soc B* **2009**, *364*, 2881.

14. Spudich, J. L.; Yang, C. S.; Jung, K. H.; Spudich, E. N., "Retinylidene proteins: Structures and functions from archaea to humans" *Annu Rev Cell Dev Bi* **2000**, *16*, 365.

15. Terakita, A., "The opsins" *Genome Biol* **2005**, *6*.

16. Terakita, A.; Koyanagi, M.; Shichida, Y., "Investigation of the non-visual pigment parapinopsins in lower vertebrates" *Zool Sci* **2005**, *22*, 1474.

17. Zhang, F.; Vierock, J.; Yizhar, O.; Fenno, L. E.; Tsunoda, S.; Kianianmomeni, A.; Prigge, M.; Berndt, A.; Cushman, J.; Polle, J.; Magnuson, J.; Hegemann, P.; Deisseroth, K., "The Microbial Opsin Family of Optogenetic Tools" *Cell* **2011**, *147*, 1446.

18. Fu, H. Y.; Lin, Y. C.; Chang, Y. N.; Tseng, H.; Huang, C. C.; Liu, K. C.; Huang, C. S.; Su, C. W.; Weng, R. R.; Lee, Y. Y.; Ng, W. V.; Yang, C. S., "A Novel Six-Rhodopsin System in a Single Archaeon" *J Bacteriol* **2010**, *192*, 5866.

19. Asenjo, A. B.; Rim, J.; Oprian, D. D., "Molecular Determinants of Human Red/Green Color Discrimination" *Neuron* **1994**, *12*, 1131.

20. Chan, T.; Lee, M.; Sakmar, T. P., "Introduction of Hydroxyl-Bearing Amino-Acids Causes Bathochromic Spectral Shifts in Rhodopsin - Amino-Acid Substitutions Responsible for Red-Green Color Pigment Spectral Tuning" *J Biol Chem* **1992**, *267*, 9478.

21. Janz, J. M.; Farrens, D. L., "Engineering a functional blue-wavelength-shifted rhodopsin mutant" *Biochemistry-Us* **2001**, *40*, 7219.

22. Kakitani, T.; Beppu, Y.; Yamada, A., "Color tuning mechanism of human red and green visual pigments" *Photochem Photobiol* **1999**, *70*, 686.

23. Lin, S. W.; Kochendoerfer, G. G.; Carroll, K. 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*, 24583.

24. Livnah, N.; Sheves, M., "Model Compounds Can Mimic Spectroscopic Properties of Bovine Rhodopsin" *J Am Chem Soc* **1993**, *115*, 351.

25. Nakanishi, K.; Baloghnair, V.; Arnaboldi, M.; Tsujimoto, K.; Honig, B., "An External Point-Charge Model for Bacteriorhodopsin to Account for Its Purple Color" *J Am Chem Soc* **1980**, *102*, 7945.

26. Neitz, M.; Neitz, J.; Jacobs, G. H., "Spectral Tuning of Pigments Underlying Red-Green Color-Vision" *Invest Ophth Vis Sci* **1991**, *32*, 1092.

27. Parry, J. W. L.; Poopalasundaram, S.; Bowmaker, J. K.; Hunt, D. M., "A novel amino acid substitution is responsible for spectral tuning in a rodent violet-sensitive visual pigment" *Biochemistry-Us* **2004**, *43*, 8014.

28. Sakmar, T. P.; Franke, R. R.; Khorana, H. G., "The Role of the Retinylidene Schiff-Base Counterion in Rhodopsin in Determining Wavelength Absorbency and Schiff-Base Pka" *P Natl Acad Sci USA* **1991**, *88*, 3079.

29. Spudich, J. L.; Mccain, D. A.; Nakanishi, K.; Okabe, M.; Shimizu, N.; Rodman, H.; Honig, B.; Bogomolni, R. A., "Chromophore Protein-Interaction in Bacterial Sensory Rhodopsin and Bacteriorhodopsin" *Biophys J* **1986**, *49*, 479.

30. Takahashi, T.; Yan, B.; Mazur, P.; Derguini, F.; Nakanishi, K.; Spudich, J. L., "Color Regulation in the Archaebacterial Phototaxis Receptor Phoborhodopsin (Sensory Rhodopsin-Ii)" *Biochemistry-Us* **1990**, *29*, 8467.

31. Wakakuwa, M.; Terakita, A.; Koyanagi, M.; Stavenga, D. G.; Shichida, Y.; Arikawa, K., "Evolution and Mechanism of Spectral Tuning of Blue-Absorbing Visual Pigments in Butterflies" *Plos One* **2010**, *5*.

32. Coto, P. B.; Strambi, A.; Ferre, N.; Olivucci, M., "The color of rhodopsins at the ab initio multiconfigurational perturbation theory resolution" *P Natl Acad Sci USA* **2006**, *103*, 17154.

33. Irving, C. S.; Byers, G. W.; Leermake.Pa, "Spectroscopic Model for Visual Pigments - Influence of Microenvironmental Polarizability" *Biochemistry-Us* **1970**, *9*, 858.

34. 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*, 953.

35. Rajamani, R.; Gao, J. L., "Combined QM/MM study of the opsin shift in bacteriorhodopsin" *J Comput Chem* **2002**, *23*, 96.

36. Rajamani, R.; Lin, Y. L.; Gao, J. L., "The Opsin Shift and Mechanism of Spectral Tuning in Rhodopsin" *J Comput Chem* **2011**, *32*, 854.

37. 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*, 3108.

38. Sekharan, S.; Altun, A.; Morokuma, K., "QM/MM Study of Dehydro and Dihydro beta-Ionone Retinal Analogues in Squid and Bovine Rhodopsins: Implications for Vision in Salamander Rhodopsin" *J Am Chem Soc* **2010**, *132*, 15856.

39. Sekharan, S.; Mooney, V. L.; Rivalta, I.; Kazmi, M. A.; Neitz, M.; Neitz, J.; Sakmar, T. P.; Yan, E. C. Y.; Batista, V. S., "Spectral Tuning of Ultraviolet Cone Pigments: An Interhelical Lock Mechanism" *J Am Chem Soc* **2013**, *135*, 19064.

40. Sekharan, S.; Sugihara, M.; Buss, V., "Origin of spectral tuning in rhodopsin - It is not the binding pocket" *Angew Chem Int Edit* **2007**, *46*, 269.

41. Sugihara, M.; Buss, V.; Entel, P.; Elstner, M.; Frauenheim, T., "11-cis-retinal protonated Schiff base: Influence of the protein environment on the geometry of the rhodopsin chromophore" *Biochemistry-Us* **2002**, *41*, 15259.

42. Trabanino, R. J.; Vaidehi, N.; Goddard, W. A., "Exploring the molecular mechanism for color distinction in humans" *J Phys Chem B* **2006**, *110*, 17230.

43. Wanko, M.; Hoffmann, M.; Strodel, P.; Koslowski, A.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M., "Calculating absorption shifts for retinal proteins: Computational challenges" *J Phys Chem B* **2005**, *109*, 3606.

44. Zhou, X. W.; Sundholm, D.; Wesolowski, T. A.; Kaila, V. R. I., "Spectral Tuning of Rhodopsin and Visual Cone Pigments" *J Am Chem Soc* **2014**, *136*, 2723.

45. Kochendoerfer, G. G.; Lin, S. W.; Sakmar, T. P.; Mathies, R. A., "How color visual pigments are tuned" *Trends Biochem Sci* **1999**, *24*, 300.

46. Nakanishi, K.; Berova, N.; Fishkin, N.; Fujioka, N., "Bioorganic studies on rhodopsins" *J Chin Chem Soc-Taip* **2002**, *49*, 443.

47. Nielsen, M. B., "Model systems for understanding absorption tuning by opsin proteins" *Chem Soc Rev* **2009**, *38*, 913.

48. Van der Horst, M. A.; Hellingwerf, K. J., "Photoreceptor proteins, "star actors of modern times": A review of the functional dynamics in the structure of representative members of six different photoreceptor families" *Accounts Chem Res* **2004**, *37*, 13.

49. Berbasova, T.; Nosrati, M.; Vasileiou, C.; Wang, W. J.; Sing, K.; Lee, S.; Yapici, I.; Geiger, J. H.; Borhan, B., "Rational Design of a Colorimetric pH Sensor from a Soluble Retinoic Acid Chaperone" *J Am Chem Soc* **2013**, *135*, 16111.

50. Sakmar, T. P., "Redder Than Red" *Science* **2012**, *338*, 1299.

51. Wang, W. J.; Nossoni, Z.; Berbasova, T.; Watson, C. T.; Yapici, I.; Lee, K. S. S.; Vasileiou, C.; Geiger, J. H.; Borhan, B., "Tuning the Electronic Absorption of Protein-Embedded All-trans-Retinal" *Science* **2012**, *338*, 1340.

52. Blatz, P. E.; Liebman, P. A., "Wavelength Regulation in Visual Pigments" *Exp Eye Res* **1973**, *17*, 573.

53. Honig, B.; Greenberg, A. D.; Dinur, U.; Ebrey, T. G., "Visual-Pigment Spectra - Implications of Protonation of Retinal Schiff-Base" *Biochemistry-Us* **1976**, *15*, 4593.

54. Tsin, A. T. C.; Liebman, P. A.; Beatty, D. D.; Drzymala, R., "Rod and Cone Visual Pigments in the Goldfish" *Vision Res* **1981**, *21*, 943.

55. Wald, G., "The Biochemistry of Vision" *Annu Rev Biochem* **1953**, *22*, 497.

56. Wald, G.; Brown, P. K.; Smith, P. H., "Iodopsin" *J Gen Physiol* **1955**, *38*, 623.

57. Wald, G., "Vision" *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **1953**, *12*, 606.

58. Honig, B.; Ebrey, T. G., "Structure and Spectra of Chromophore of Visual Pigments" *Annu Rev Biophys Bio* **1974**, *3*, 151.

59. Labhart, H., "Fe Theory Including an Elastic Sigma-Skeleton .1. Spectra and Bond Lengths in Long Polyenes" *J Chem Phys* **1957**, *27*, 957.

60. Gilardi, R.; Karle, I. L.; Karle, J.; Sperling, W., "Crystal Structure of Visual Chromophores, 11-Cis and All-Trans Retinal" *Nature* **1971**, *232*, 187.

61. Gilardi, R. D.; Karle, J.; Karle, I. L., "Crystal and Molecular Structure of 11-Cis-Retinal" *Acta Crystall B-Stru* **1972**, *B 28*, 2605.

62. Hamanaka, T.; Kakudo, M.; Ashida, T.; Mitsui, T., "Crystal-Structure of All-Trans Retinal" *Acta Crystall B-Stru* **1972**, *B 28*, 214.

63. Royant, A.; Nollert, P.; Edman, K.; Neutze, R.; Landau, E. M.; Pebay-Peyroula, E.; Navarro, J., "X-ray structure of sensory rhodopsin II at 2.1-angstrom resolution" *P Natl Acad Sci USA* **2001**, *98*, 10131.

64. Subramaniam, S.; Henderson, R., "Molecular mechanism of vectorial proton translocation by bacteriorhodopsin" *Nature* **2000**, *406*, 653.

65. Murakami, M.; Kouyama, T., "Crystal structure of squid rhodopsin" *Nature* **2008**, *453*, 363.

66. 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*, 739.

67. 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*, 6597.

68. Aharoni, A.; Khatchatouriants, A.; Manevitch, A.; Lewis, A.; Sheves, M., "Protein-beta-ionone ring interactions enhance the light-induced dipole of the chromophore in bacteriorhodopsin" *J Phys Chem B* **2003**, *107*, 6221.

69. Albeck, A.; Friedman, N.; Sheves, M.; Ottolenghi, M., "Factors Affecting the Absorption Maxima of Acidic Forms of Bacteriorhodopsin - a Study with Artificial Pigments" *Biophys J* **1989**, *56*, 1259.

70. Baselt, D. R.; Fodor, S. P. A.; Vandersteen, R.; Lugtenburg, J.; Bogomolni, R. A.; Mathies, R. A., "Halorhodopsin and Sensory Rhodopsin Contain a C-6-C-7 S-Trans Retinal Chromophore" *Biophys J* **1989**, *55*, 193.

71. Courtin, J. M. L.; Verhagen, L.; Biesheuvel, P. L.; Lugtenburg, J.; Vanderbend, R. L.; Vandam, K., "Bacteriorhodopsin - the Influence of the Cyclohexene-Ring Methyls" *Recl Trav Chim Pay B* **1987**, *106*, 112.

72. Fukada, Y.; Shichida, Y.; Yoshizawa, T.; Ito, M.; Kodama, A.; Tsukida, K., "Studies on Structure and Function of Rhodopsin by Use of Cyclopentatrienylidene 11-Cis-Locked-Rhodopsin" *Biochemistry-Us* **1984**, *23*, 5826.

73. Iwasa, T.; Ito, M.; Tokunaga, F., "A Novel Bacteriorhodopsin Analog with Conformationally 6-S-Cis Fixed Retinals" *Photochem Photobiol* **1992**, *56*, 921.

74. Koganov, E. S.; Hirshfeld, A.; Sheves, M., "Retinal beta-Ionone Ring-Salinixanthin Interactions in Xanthorhodopsin: A Study Using Artificial Pigments" *Biochemistry-Us* **2013**, *52*, 1290.

75. Mao, B.; Govindjee, R.; Ebrey, T. G.; Arnaboldi, M.; Baloghnair, V.; Nakanishi, K.; Crouch, R., "Photochemical and Functional-Properties of Bacteriorhodopsins Formed from 5,6-Dihydrodesmethylretinals and 5,6-Dihydrodesmethylretinals" *Biochemistry-Us* **1981**, *20*, 428.

76. Muradinszweykowska, M.; Vanamsterdam, L. J. P.; Rodenburg, L. J. M.; Lugtenburg, J.; Vanderbend, R. L.; Vandam, K., "(5-Demethyl)-Bacteriorhodopsin Analog - Its Formation and Light-Driven Proton Pump Action" *Febs Lett* **1983**, *154*, 180.

77. Rao, V. J.; Zingoni, J. P.; Crouch, R.; Denny, M.; Liu, R. S. H., "Isomers of 3,7,11-Trimethyldodeca-2,4,6,8,10-Pentaenal (a Linear Analog of Retinal) and Lower Homologs in Their Interaction with Bovine Opsin and Bacterioopsin" *Photochem Photobiol* **1985**, *41*, 171.

78. Schimz, A.; Sperling, W.; Ermann, P.; Bestmann, H. J.; Hildebrand, E., "Substitution of Retinal by Analogs in Retinal Pigments of Halobacterium-Halobium - Contribution of Bacteriorhodopsin and Halorhodopsin to Photosensory Activity" *Photochem Photobiol* **1983**, *38*, 417.

79. Sheves, M.; Friedman, N.; Rosenbach, V.; Ottolenghi, M., "Preparation of (1,1,5-Tri-Demethyl)Bacteriorhodopsin Pigment and Its Photocycle Study" *Febs Lett* **1984**, *166*, 245.

80. Singh, A. K.; Hota, P. K., "Development of bacteriorhodopsin analogues and studies of charge separated excited states in the photoprocesses of linear polyenes" *Photochem Photobiol* **2007**, *83*, 50.

81. Steinberg, G.; Friedman, N.; Sheves, M.; Ottolenghi, M., "Isomer Composition and Spectra of the Dark and Light Adapted Forms of Artificial Bacteriorhodopsins" *Photochem Photobiol* **1991**, *54*, 969.

82. Vandersteen, R.; Biesheuvel, P. L.; Mathies, R. A.; Lugtenburg, J., "Retinal Analogs with Locked 6-7 Conformations Show That Bacteriorhodopsin Requires the 6-S-Trans Conformation of the Chromophore" *J Am Chem Soc* **1986**, *108*, 6410.

83. Albeck, A.; Friedman, N.; Sheves, M.; Ottolenghi, M., "Role of Retinal Isomerizations and Rotations in the Photocycle of Bacteriorhodopsin" *J Am Chem Soc* **1986**, *108*, 4614.

84. Sheves, M.; Albeck, A.; Ottolenghi, M.; Boveegeurts, P. H. M.; Degrip, W. J.; Einterz, C. M.; Lewis, J. W.; Schaechter, L. E.; Kliger, D. S., "An Artificial Visual Pigment with Restricted C9-C-11 Motion Forms Normal Photolysis Intermediates" *J Am Chem Soc* **1986**, *108*, 6440.

85. Honig, B.; Hudson, B.; Sykes, B. D.; Karplus, M., "Ring Orientation in Beta-Ionone and Retinals (Theoretical/Nmr/Semi-Empirical Approach/Overhauser Effect/S-Cis Conformation)" *P Natl Acad Sci USA* **1971**, *68*, 1289.

86. 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" *J Am Chem Soc* **1992**, *114*, 2400.

87. Akita, H.; Tanis, S. P.; Adams, M.; Baloghnair, V.; Nakanishi, K., "Non-Bleachable Rhodopsins Retaining the Full Natural Chromophore" *J Am Chem Soc* **1980**, *102*, 6370.

88. Albeck, A.; Friedman, N.; Ottolenghi, M.; Sheves, M.; Einterz, C. M.; Hug, S. J.; Lewis, J. W.; Kliger, D. S., "Photolysis Intermediates of the Artificial Visual Pigment Cis-5,6-Dihydro-Isorhodopsin" *Biophys J* **1989**, *55*, 233.

89. Arnaboldi, M.; Motto, M. G.; Tsujimoto, K.; Baloghnair, V.; Nakanishi, K., "Hydro-Retinals and Hydro-Rhodopsins" *J Am Chem Soc* **1979**, *101*, 7082.

90. Azuma, M.; Azuma, K.; Kito, Y., "Circular-Dichroism of Visual Pigment Analogs Containing 3-Dehydroretinal and 5,6-Epoxy-3-Dehydroretinal as Chromophore" *Biochim Biophys Acta* **1973**, *295*, 520.

91. Dominguez, M.; Alvarez, R.; Perez, M.; Palczewski, K.; de Lera, A. R., "The role of the 11-cis-retinal ring methyl substituents in visual pigment formation" *Chembiochem* **2006**, *7*, 1815.

92. Hubbard, R., "Cis-Trans Isomers of Vitamin-a and Retinene in the Rhodopsin System" *Fed Proc* **1952**, *11*, 233.

93. Hubbard, R.; Wald, G., "Cis-Trans Isomers of Vitamin-a and Retinene in Vision" *Science* **1952**, *115*, 60.

94. Hubbard, R.; Wald, G., "Cis-Trans Isomers of Vitamin-a and Retinene in the Rhodopsin System" *J Gen Physiol* **1952**, *36*, 269.

95. Ito, M.; Katsuta, Y.; Imamoto, Y.; Shichida, Y.; Yoshizawa, T., "Conformational-Analysis of the Rhodopsin Chromophore Using Bicyclic Retinal Analogs" *Photochem Photobiol* **1992**, *56*, 915.

96. Ito, M.; Kodama, A.; Tsukida, K.; Fukada, Y.; Shichida, Y.; Yoshizawa, T., "A Novel Rhodopsin Analog Possessing the Cyclopentatrienylidene Structure as the 11-Cis-Locked and the Full Planar Chromophore" *Chem Pharm Bull* **1982**, *30*, 1913.

97. Karnaukhova, E.; Hu, S. H.; Boonyasai, R.; Tan, Q.; Nakanishi, K., "Bioactivity of visual pigments with sterically modified retinal analogs" *Bioorg Chem* **1999**, *27*, 372.

98. 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**, 2430.

99. Wald, G., "The photochemistry of vision" *Documenta Ophthalmologica* **1949**, *3*, 94.

100. 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*, 12347.

101. Nielsen, I. B.; Petersen, M. A.; Lammich, L.; Nielsen, M. B.; Andersen, L. H., "Absorption studies of neutral retinal Schiff base chromophores" *J Phys Chem A* **2006**, *110*, 12592.

102. Petersen, M. A.; Nielsen, I. B.; Kristensen, M. B.; Kadziola, A.; Lammich, L.; Andersen, L. H.; Nielsen, M. B., "Novel retinylidene iminium salts for defining opsin shifts: synthesis and intrinsic chromophoric properties" *Org Biomol Chem* **2006**, *4*, 1546.

103. 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 Edit* **2010**, *49*, 1790.

CHAPTER III: DE NOVO DESIGN OF A PHOTOACTIVATABLE PROTEIN SYSTEM BY USING A SYNTHETIC ANALOG OF AZOBENZENE

III.1 Photophysical Features Of Azo-chromophores

The synthesis and the development of azobenzene based compounds dates back to 1800s. Peter Griess, a German chemist, synthesized the first azo compound-Aniline yellow in 1858 and since then an enormous variety of azo compounds have been produced. The main structural feature of azo compounds is that they contain an azo moiety (-N=N-), which conjugates two aromatic ring systems. These compounds possess bright colors (particularly orange, yellow and red), making them one of the most used organic chromophores for coloring/dyeing industry.[1] Beside their robust and stable chemical structure, one of the most fascinating properties of these chromophores is the reversible photoisomerization of the azo bond.[2-4] The first discovery of the isomerization of the azo bond was reported by Hartley in 1937 where he carefully extracted and analyzed the *cis*-form of the azobenzene.[5]

The photoisomerization of azobenzene is readily induced by light excitation. The *trans* isomer shows an intense peak in the UV region (~ 320 nm), associated with the $\pi \rightarrow \pi^*$ transition, and a weak absorption band near 440 nm, corresponding to the $n \rightarrow \pi^*$ transition (**Figure III-1**). UV-irradiation of the *trans* isomer triggers the photoisomerization and results in increasing of the *cis*-isomer. This leads to a substantial decrease in the intensity of the UV absorption band since the *cis*-isomer exhibits a bent conformation with its phenyl rings (~55°),

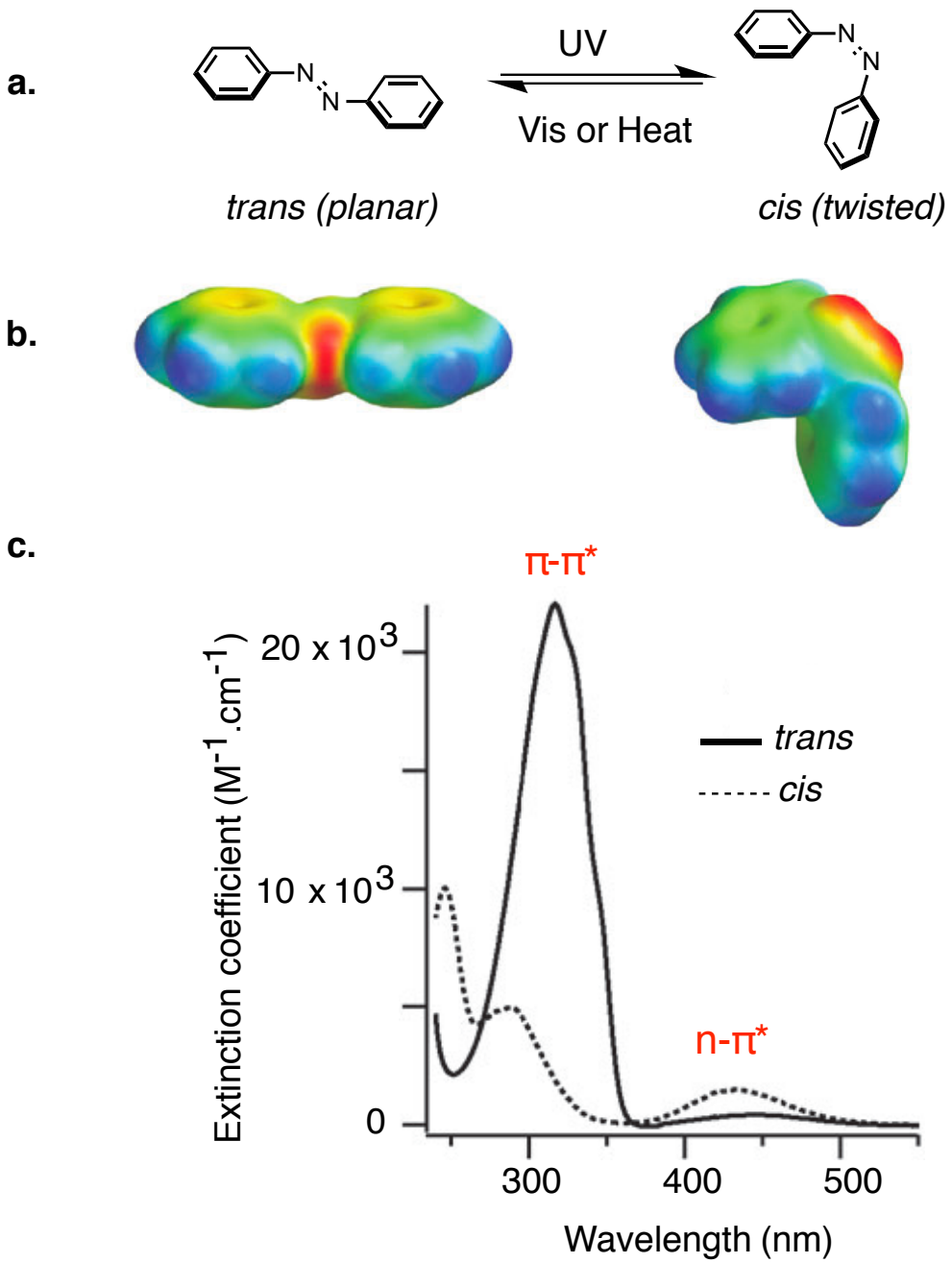


Figure III-1: Physical features of the azobenzene molecule. **a.** Chemical structures of *trans* and *cis* isomers of azobenzene. **b.** Electrostatic potential of azo isomers shown in spacefilling model (red-negative, blue-positive).[19] **c.** UV-vis spectra of *trans* and *cis* forms of azobenzene.

weakening the π orbital conjugation.[6] On the other hand, the *cis*-isomer has a stronger n → π* absorption peak as compared to that of seen in a *trans*-isomer.

Regeneration of the *trans*-isomer could be achieved by irradiation at 450 nm or thermal relaxation of the *cis*-azobenzene (*trans*-isomer is about 12 Kcal/mol more stable than the *cis*-isomer).[3,7] However, it should be noted that the absorption profiles of the two isomers substantially overlap and, thus, it is not possible to generate exclusively one isomer over the other by light irradiation. In other words, light irradiation creates a photostationary state in which *trans* isomerization to *cis* and *cis* isomerization to *trans* take place simultaneously. The quantum efficiency of the isomerization for the *trans* → *cis* conversion was calculated to be 0.40 (upon irradiation with UV light) whereas the reverse photoconversion, *cis* → *trans*, was found to be more effective with 0.53 quantum efficiency (when excited at 290 nm).[8] Since *trans*-isomer is energetically more stable than the *cis* form, the thermal relaxation of the *cis* → *trans* process effectively yields the *trans* isomer (>99.9%).[3,7]

The mechanism of *cis-trans* isomerization has attracted much attention for many years but it is still debated (**Figure III-2**).[4,9,10] Two possible mechanisms were postulated to explain the azo bond isomerization; i) rotation or torsion around the N-N bond with a disruption of the double bond,[11,12] and ii) via an inversion with a hybridized transition state.[13] However, a different isomerization pathway was proposed recently in which mixed mechanisms can take place. In this scenario, "concerted inversion", both -C-N=N- bonds bend in the same molecular plane to flatten the -CNNC- bond.[4,8,14,15] It is noteworthy that the isomerization processes occurs on a picosecond time scale but the kinetics of

thermal isomerization of *cis* to *trans* isomer are varied, depend on the structure, and can take from seconds to a few days.[16-19]

It is well known that different substituents on the phenyl rings alter the

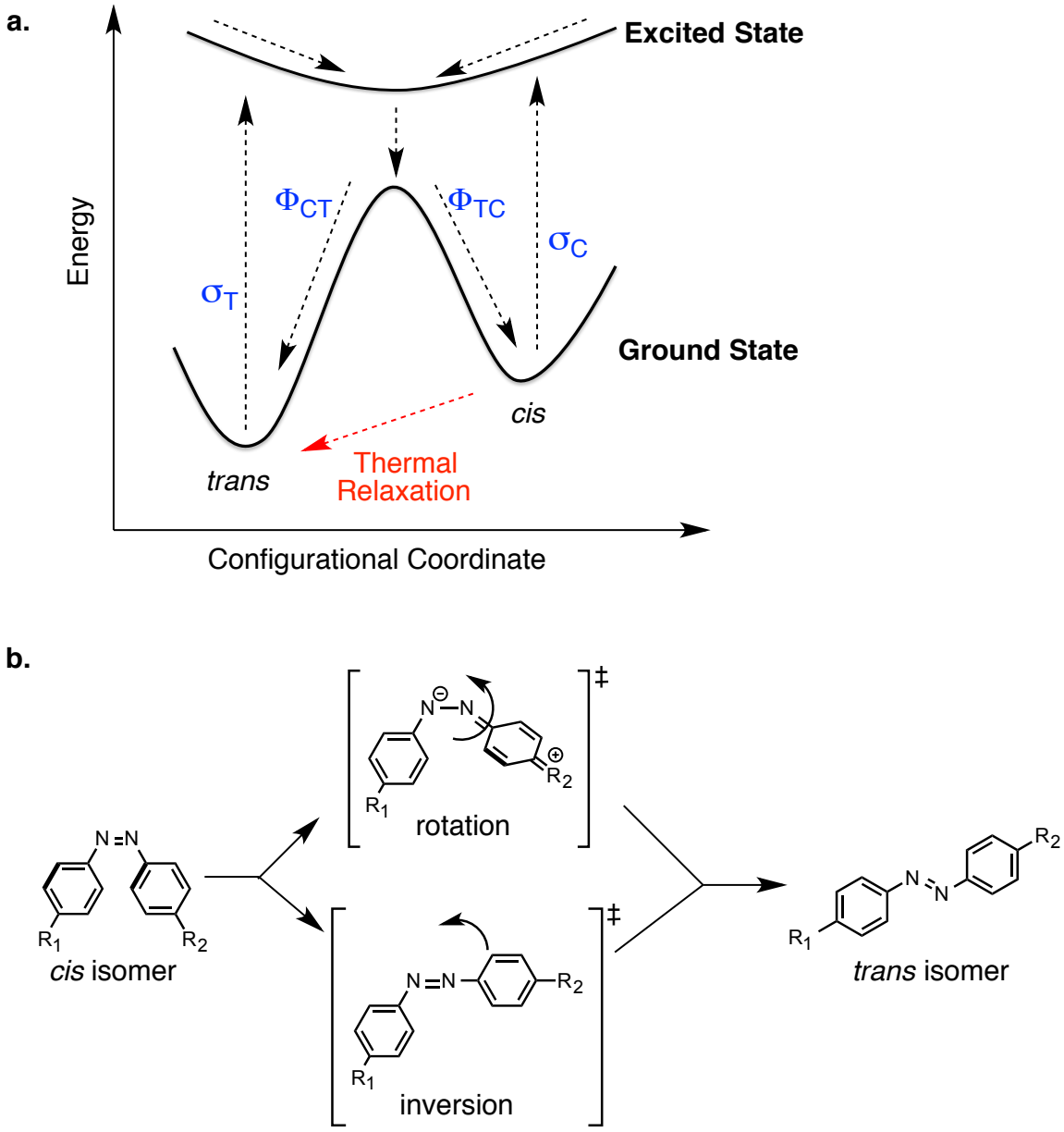


Figure III-2: Schematic representation of *cis-trans* isomerization. **a.** Energy diagram of azobenzene isomers. Molecular cross section for the light absorption is indicated by σ_T (*trans* isomer) and σ_C (*cis* isomer). Φ_{CT} (*cis* → *trans*), and Φ_{TC} (*trans* → *cis*) are the quantum efficiencies for each transition. **b.** Two proposed mechanisms for the *cis-trans* isomerization.

spectroscopic properties of the azo compounds. Depending on the electronic features of the substituents, azobenzene based photochromic systems can be classified in three categories: i) azobenzenes having no polar substitutions, ii) aminobenzenes with an electron donating group installed at the *para* position, and iii) pseudo-stilbenes with push-pull configuration (**Figure III-3a**).[20] The absorption profile of "azobenzenes" is similar to the unsubstituted azobenzene counterpart. However, installation of amino groups at the *ortho* and *para* positions induces a significant red shift in the spectrum. Moreover, push-pull azobenzenes (pseudo-stilbenes) can display further red shift, with absorption bands in the visible region.[3,16,18-22] **Figure III-3b** depicts some of the functionalized azobenzene derivatives with corresponding absorption maxima.[22-26]

The effect of substituents also influences the rate of thermal *cis* → *trans* relaxation. Even though the mechanism of isomerization still under investigation, it is clear that an increase of dipole character along the chromophore facilitates the thermal relaxation. In these systems, enhanced resonance over the molecule significantly increases the single bond character in the azo moiety, and thus lowers the activation barrier for thermal isomerization. Therefore, it is not surprising that push-pull azobenzene based photochromic systems have the shortest life-time for the *cis*-isomers. Additionally, there are studies showing that the rates could be solvent and pH dependent. In general, the increase of solvent

polarity and low pH regimes (protonation of the azo moiety) enhance thermal relaxation rates.[13,27-31]

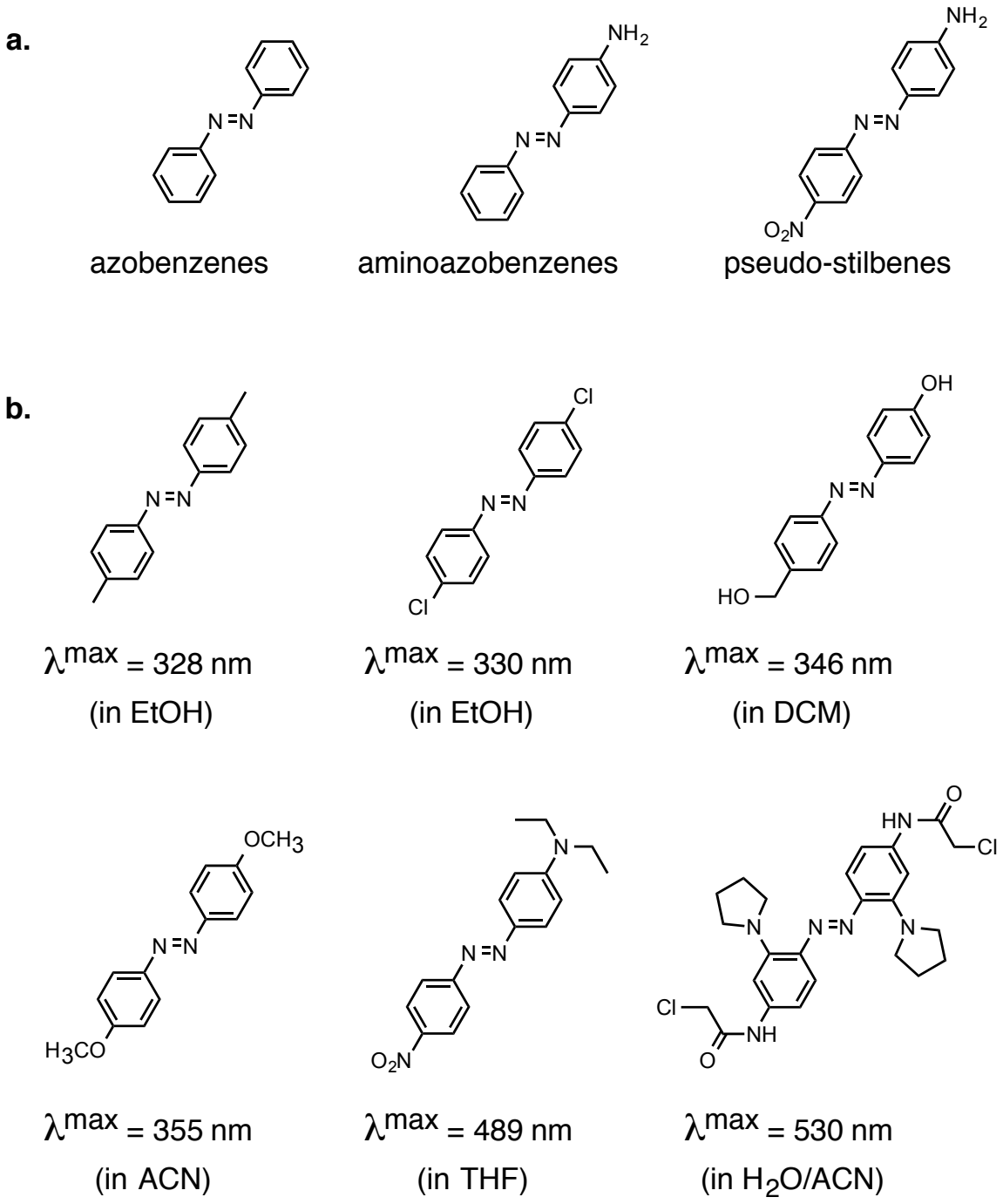
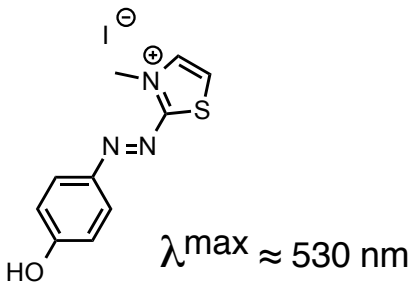


Figure III-3: Chemical stuructures of azobenzene derivatives. **a.** Types of azobenzene molecules based on their different substitutions. **b.** Few examples of modified azobenzenes with the corresponding absorption maxima.

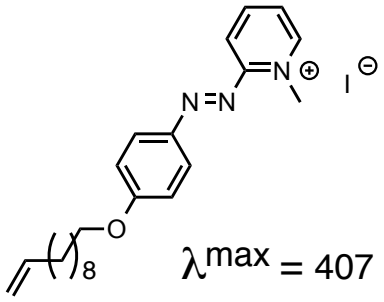
Recent work has shown that new azo compounds can be synthesized such that a variety of heterocyclic moieties is attached to the azo bond. These compounds exhibit remarkable photostability and different *cis* → *trans* conversions rates. The latter property of the azo switch is highly important for its utility and application. In cases of fast photoswitching systems, it is desirable to have a rapid thermal relaxation (*cis* → *trans*) rates to regenerate the initial state of the system without the need for a second irradiation. Recently, Garcia-Amoros and coworkers synthesized a series of photoactive azoderivatives that display



$\lambda^{max} \approx 530$ nm

$\tau = 55$ ns

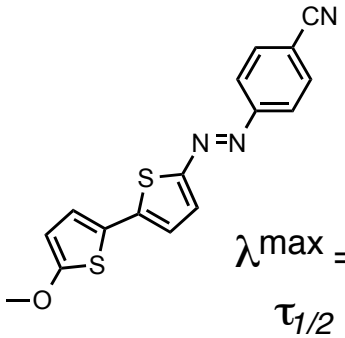
benzothiazolium azophenolic salt (***A***)



$\lambda^{max} = 407$ nm

$\tau = 570$ μs

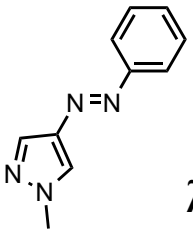
azopyridinium salt (***B***)



$\lambda^{max} = 495$ nm

$\tau_{1/2} = 33$ s

2,2'-bithiophene azo dye (***C***)



$\lambda^{max} = 417$ nm

$\tau_{1/2} \approx 1000$ days

arylazopyrazole (***D***)

Figure III-4: Chemical structures and photophysical properties of azoheteroarenes. Each λ^{max} value shows the n → π* transition of the azo dye. Thermal relaxation time and the half life-time for *cis* → *trans* isomerization are indicated by τ and $\tau_{1/2}$, respectively.

ultrafast photoswitching kinetics of isomerization.[32] Compound **A**, a benzothiazolium azophenolic salt, can be irradiated with a green laser at 532 nm to trigger a *trans* → *cis* conversion in 5 ns. Furthermore, without a need for a second illumination, the photogenerated *cis* isomer thermally switches back to the stable *trans* form with a life-time of 55 ns, highlighting that a full switching cycle could be completed on a nanosecond timescale. These benzothiazolium salts are the fastest reported azo switches so far. Another class of azo salts was reported for its fast *cis* → *trans* isomerization kinetics. Azopyridinium salts contain a pyridinic cation, which is used as an electron withdrawing group to generate a strong push-pull system (compound **B** in **Figure III-4**).[33] This feature allows a large red shift in the π → π* transition, absorbing around 410 nm, which overlaps with the n → π* electronic transition. The resultant red shift is attributed to the strong charge transfer from the alkoxy group to the positively charged nitrogen atom of the pyridinium ring. This strong electron transfer also promotes a partial breaking of the -N=N- bond, thereby lowering the energy barrier for the rotation around the azo moiety to recover the more stable *trans* isomer. It is also important to mention that the incorporation of heterocycles not only affects the kinetics of isomerization processes, but also induces different spectroscopic properties. To illustrate, compound **C** exhibits a photochromic behavior in which isomerization of *cis* → *trans* (or vice versa) results in a color change.[34]

Although short life-times of the *cis* isomer could be beneficial for rapid and complete resetting of the switch, one might prefer to maintain the on/off state of

the photoswitch over a longer time frame. In this scenario, a long-lived *cis* state requires a slow *cis* → *trans* thermal relaxation. Typical life-times for the *cis* isomers of azobenzenes (**Figure III-3a**, no polar substitutions) are on the orders of hours. Howbeit, this is not considered long for compounds to be treated as a stable two state system. Therefore, considerable effort has been directed at the development of new azo chromophores for elongating the life-time of the *cis* isomer. Nishimura *et al*. found that *ortho*-alkylation of aromatic rings slows down the *cis* → *trans* thermal isomerization. This was attributed to the increase in steric demand between the *ortho* substituents and the orbitals of nitrogen atoms undergoing inversion. It was also proposed that *ortho* groups could form either a hydrophobic cage around the azo moiety to stabilize the *cis* structure or create a local non-polar environment to destabilize the polar transition state during the *cis* → *trans* isomerization.[35,36] Beside incorporations of *ortho* substituents, recent reports showed that five membered arylazopyrazoles offer substantial thermal stability of the *cis* isomers, thereby providing fully controllable photoswitching systems.[37,38] Compound **D** (**Figure III-4**) is converted into its *cis* isomer with a high efficiency (>98%) when excited with a broad band 330-400 nm light source. The resultant *cis* species possesses a thermal half-life of near 1000 days, which is considered to be one of the slowest azo switches reported to date. In sharp contrast to classical azobenzenes, these arylazopyrazoles show faster thermal relaxation times in the presence of *ortho* substituents. This raises questions about the proposed mechanisms for the thermal isomerization of azoarenes.[39-41]

Lastly, numerous macrocyclic azobenzene compounds have been also synthesized. These molecules have high *cis* stability and elongated life-times (up to six years), as a result of conformational strain experience during the isomerization process.[42-45]

Azobenzene switches have been intensively investigated ever since the discovery of their photoswitching properties. Their photochemistry has not only illuminated several fundamental aspects of photochemistry, but also offered practical applicability in different fields of science and industry. These switches have been integrated in a wide range of applications such as optical-switches, surface patterning and artificial bio-mimics (summarized by number of excellent reviews).[16,19,20,46-51] Specifically, tools based on azobenzene photochromic systems have become widely appreciated in molecular biology. Since light can be localized in time and space, it is a powerful and ideal tool to trigger a variety of cellular processes. The union of robust and tunable photoisomerization of azo switches with the high temporal and spatial precision of light is a promising avenue for development of artificial biomolecules with light-controlled function. In this report, we aimed at designing a protein system that couples with an azobenzene-based chromophore to respond to light stimulation. Such photo-controlled protein-small molecule hybrids could be useful in studying protein-protein interactions and generating a bio-machinery for affinity-based protein separation methodology.

III.2 Design Of A Photoactivatable Protein System

Protein engineering is a broad area in which the design of new proteins with specialized functions could be achieved via many different engineering methods. Rational design (site-directed mutagenesis) and evolutionary systems (random mutagenesis and selection) are examples of effective protein engineering approaches to develop new proteins or enzymes with desired properties. Inspired by several applicable protein engineering approaches, our research group has initiated the design of photoactivatable protein systems by using small cytosolic lipid binding proteins and azobenzene-based photochromic molecules as the protein-binding partners. Our goal mainly focuses on finding a protein-chromophore complex that can respond to different wavelengths of light such that the interaction between the protein and its chromophore could be controlled. As illustrated in **Figure III-5a**, we are interested in finding protein scaffolds that show high binding affinity to the *trans* isomer of azobenzene moiety but disfavor the binding of the *cis* isomer. The interaction between the protein and its ligand would be controlled by light irradiation as *trans* azobenzene readily isomerizes to the *cis* form upon illumination. Such a photo-controlled protein-small molecule hybrid could be useful in developing a photoswitchable protein affinity purification system (**Figure III-5b**). In this purification method, the protein of interest could be fused to the photoactivable protein tag and purified over a solid support. First, the solid surface is chemically modified to immobilize azo molecules (*trans* isomer at initial stage) and then treated with a mixture of

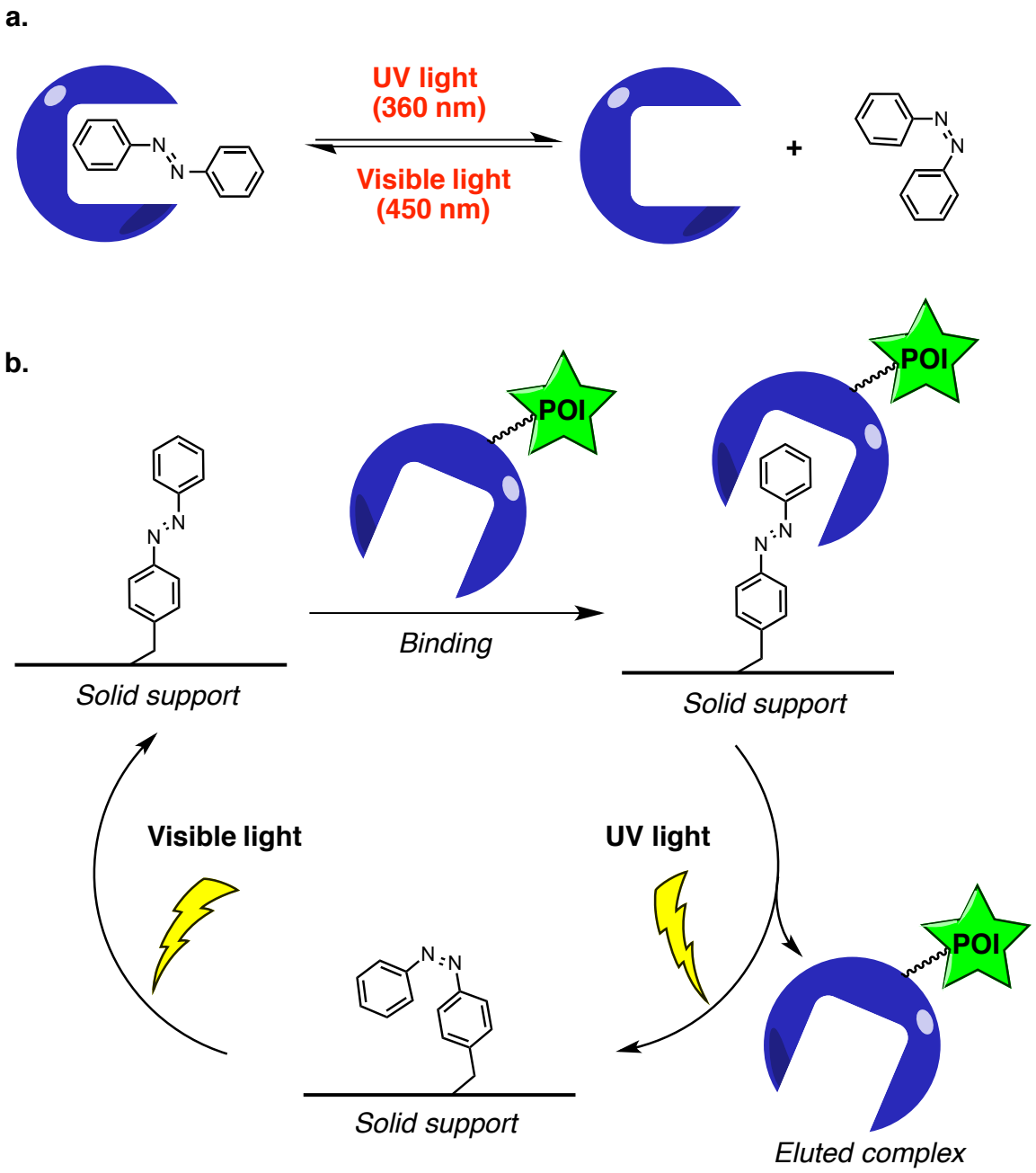


Figure III-5: Azobenzene-based protein photoswitches. **a.** General scheme for a photoswitchable protein-chromophore system. **b.** Photoswitchable affinity purification by tagging the protein of interest (POI, green star) with the photoactivatable protein tag.

proteins. As protein of interest stays intact on the solid surface with the help of the protein tag (having a high binding affinity to the *trans* azo molecule), the other undesired proteins will be washed out. Once the solid surface is UV irradiated,

trans → *cis* isomerization ensues and results in the elution of the protein tag together with the protein of interest. Lastly, the *cis* isomer of the azo molecules could be converted back to the *trans* form by either visible light irradiation or through thermal relaxation, regenerating the solid support for purification runs. Noticeably, the eluted protein of interest can be cleaved from its tag by a number of imbedded protease sequences. This purification technique allows isolation of proteins without the need for chemical eluents and time consuming clean up procedures such as dialysis or gel filtration.

III.2.1 Previous Design Work For Photoswitchable Proteins

It is critical to choose an appropriate protein platform to design an effective and robust photoactivatable protein system. As described earlier, Cellular Retinoic Acid Binding Protein II (CRABPII) and Cellular Retinol Binding Protein (CRBPII) have remarkable tolerance to mutations with a robust β-barrel topology that provides a large binding cavity. These proteins provide an ideal framework for applications in protein re-design, particularly engineering of rhodopsin mimics. Therefore, we have taken an approach to engineer these two protein systems in a manner to endow them with light-controlled function. At the outset, we have chosen CRABPII as a suitable protein platform (project initiated by Dr. Wenjing Wang). It was previously shown that CRABPII could accommodate a variety of carboxylic acid derivatives within its binding pocket (**Figure III-6**).[52] Crystal structures reveal that the key interactions between the carboxylic acid moiety of the ligands and the P2-myelin motif (Arg-Try-Arg triad)[53] are conserved in all the

structures. This data suggests that an azobenzene switch with a benzoic acid moiety could be a good ligand candidate for the CRABPII scaffold. To that end, we synthesized several derivatives of azobenzene-based carboxylic acids (see an example in **Figure III-6c**) in order to evaluate their binding affinity with CRABPII. Initial measurements showed that azo-carboxylic acid derivatives display moderate binding affinity to wild type CRABPII, with micro molar range of dissociation constants. Although the affinity of CRABPII to azo derivatives was

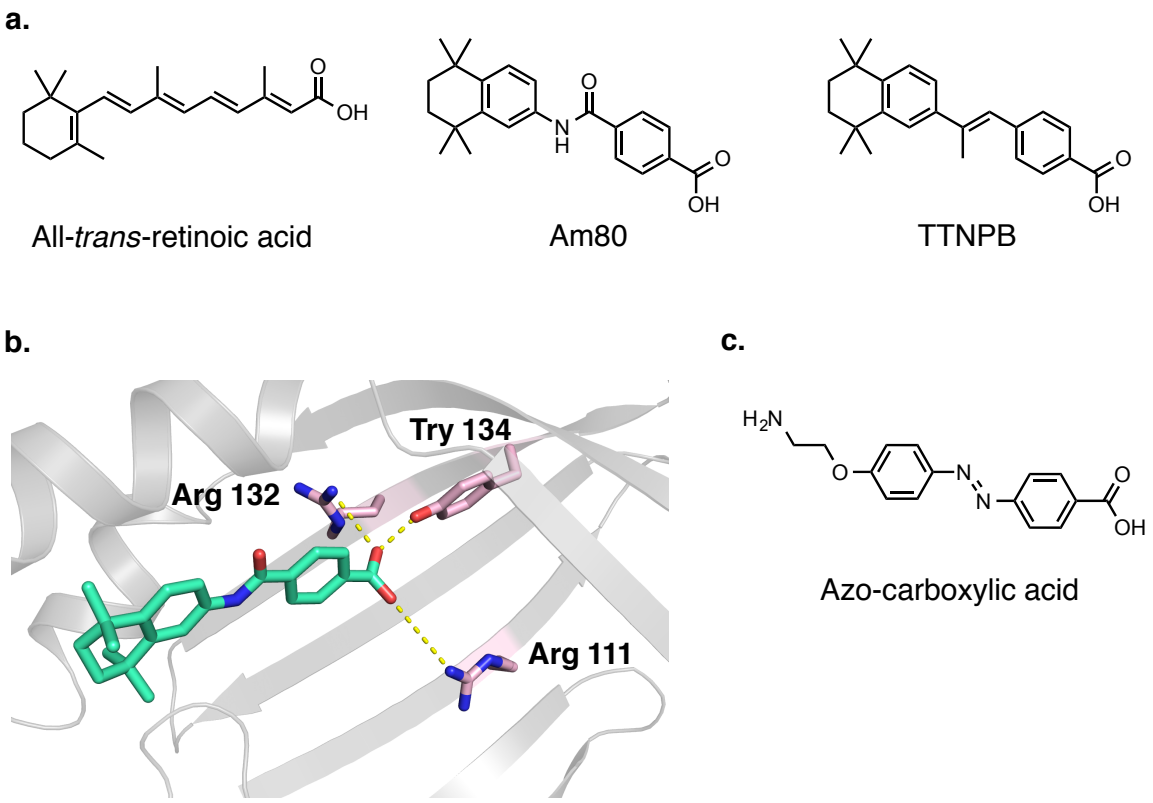


Figure III-6: Different substrate structures for a CRABPII. **a.** Structures of native ligand (all-*trans*-retinoic acid) and synthetic ligands (retinobenzoic acids). **b.** Crystal structure of Am80 (green) bound CRABPII (PDB ID: 2CBR). The highlighted residues (pink) are the important amino acids for the protein-ligand interactions. **c.** Structure of the designed azo chromophore for a photoswitchable system.

not poor, we proceeded with further affinity maturations via randomized CRABPII libraries utilizing phage display technology. Disappointingly, we had several issues during the biopanning of the constructed phage libraries, and we were unsuccessful in obtaining consistent and reproducible results. Furthermore, the failure of CRABPII-merocyanine **1** system *in vivo* studies (discussed in Chapter I) led to doubts of its feasibility and/or stability as a scaffold for phage display screening. As a result, we charged the direction of the project and decided to continue with hCRBPII as a protein platform.

III.2.2 Rational Design Of HCRBPII Into A Photoswitchable Protein

Disappointing results with engineering a photoswitchable CRABPII system via phage display led us to take another approach with a different protein. As mentioned earlier, hCRBPII has the same structural robustness and adaptable binding site similar to the CRABPII. Additionally, it exhibits consistency in its characteristic properties (expression, folding and recognition of its ligand) both *in vitro* and *in vivo* conditions. In this respect, it is reasonable to choose hCRBPII as a candidate for protein engineering.

Despite the fact that phage display is a widely used technique in the field of protein engineering, we faced several optimization problems. Also, we severely suffered from contamination of the library with the wild type phage. As a result, we decided to apply a rational design strategy in the hope that an optimum protein-chromophore interaction could be established. Similar to our previous approach with CRABPII, we structurally designed the azobenzene-based

chromophore that would show some structural similarity to the native substrate of the wild type hCRBPII. Inspection of the crystal structure of holo hCRBPII was invaluable to see the necessary elements of interaction between the protein and its cognitive ligand. It is important to mention that the hydrophobicity of the hCRBPII binding pocket is one of the most critical driving forces for its ligand as all-*trans*-retinol is a small hydrophobic molecule. Furthermore, there are two critical residues, Gln108 and Lys40, assisting all-*trans*-retinol to interact with the protein via hydrogen bonding (**Figure III-7**). These residues form strong hydrogen bonding interactions with the hydroxyl moiety of the ligand in which the ligand adopts a favorable conformation within the protein binding cavity.

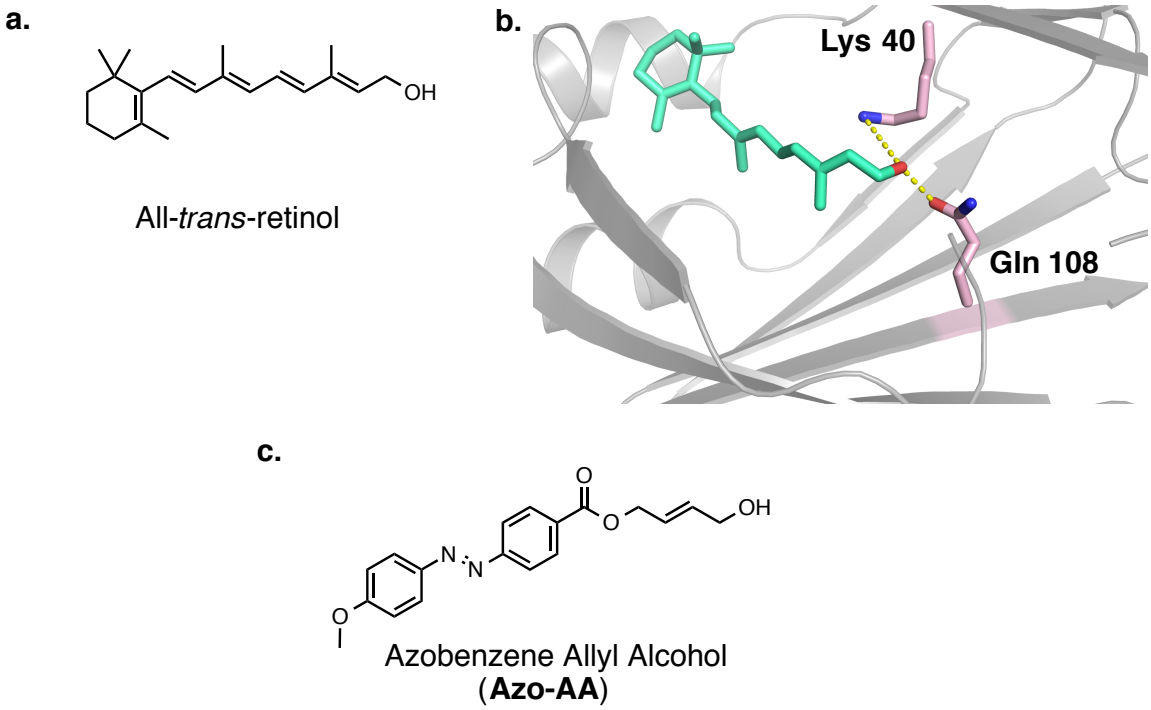


Figure III-7: Structural data of hCRBPII and its ligands. **a.** Structure of the all-*trans*-retinol, native substrate of hCRBPII. **b.** Crystal structure of all-*trans*-retinol (green) bound wild type hCRBPII (PDB ID: 4QZT). The highlighted residues (pink) are the important amino acids for protein-ligand interactions. **c.** Structure of the designed **Azo-AA** chromophore for a photoswitchable system.

In the light of the available structural data, we considered an azobenzene derivative appended with an allylic alcohol as a reasonable mimic of the native chromophore to initiate a rational design strategy. Moreover, the azobenzene core provides an ideal hydrophobic segment for driving the association of the ligand with its protein partner. We were able to synthesize the designed azobenzene allyl alcohol ligand (**Azo-AA**) in a quick fashion and good yield (see details in Chapter IV.13.6). The UV-vis spectrum of **Azo-AA** shows two characteristics absorption bands for both isomers. The thermodynamically stable *trans* isomer exhibits an intense peak corresponding to the $\pi \rightarrow \pi^*$ transition with a λ^{max} of 360 nm and a small peak at around 420 nm associated with the $n \rightarrow \pi^*$ transition (**Figure III-8**). Upon UV light exposure, **Azo-AA** readily isomerizes to the *cis* form and displays a more pronounced absorption peak of the $n \rightarrow \pi^*$ transition. Next, we evaluated the dark isomerization of the *cis* product. After populating the *cis* isomer of **Azo-AA** by UV light, the solution was kept in the dark and scanned by UV-vis spectrometer every 1.5 h. As anticipated, the *cis* product shows thermal relaxation, yielding the thermodynamically favored *trans* form in solution (60% of the *cis* species generates the *trans* product over 12 h).

III.3 Crystal Structure Of Azo-AA Bound Wild Type HCRBPII

With an efficient azobenzene photochromic switch in hand, an effort was made to generate stable protein-chromophore complex through a rational design. However, a solid rational design approach requires tentative structural information about the protein of interest with the target chromophore. In this

respect, we first focused on co-crystallization of **Azo-AA** with wild type hCRBPII. Our initial affinity measurement revealed that **Azo-AA** has a a K_d of 1.42 μM with wild type hCRBPII (**Figure III-9**). Subsequently, freshly expressed wild type

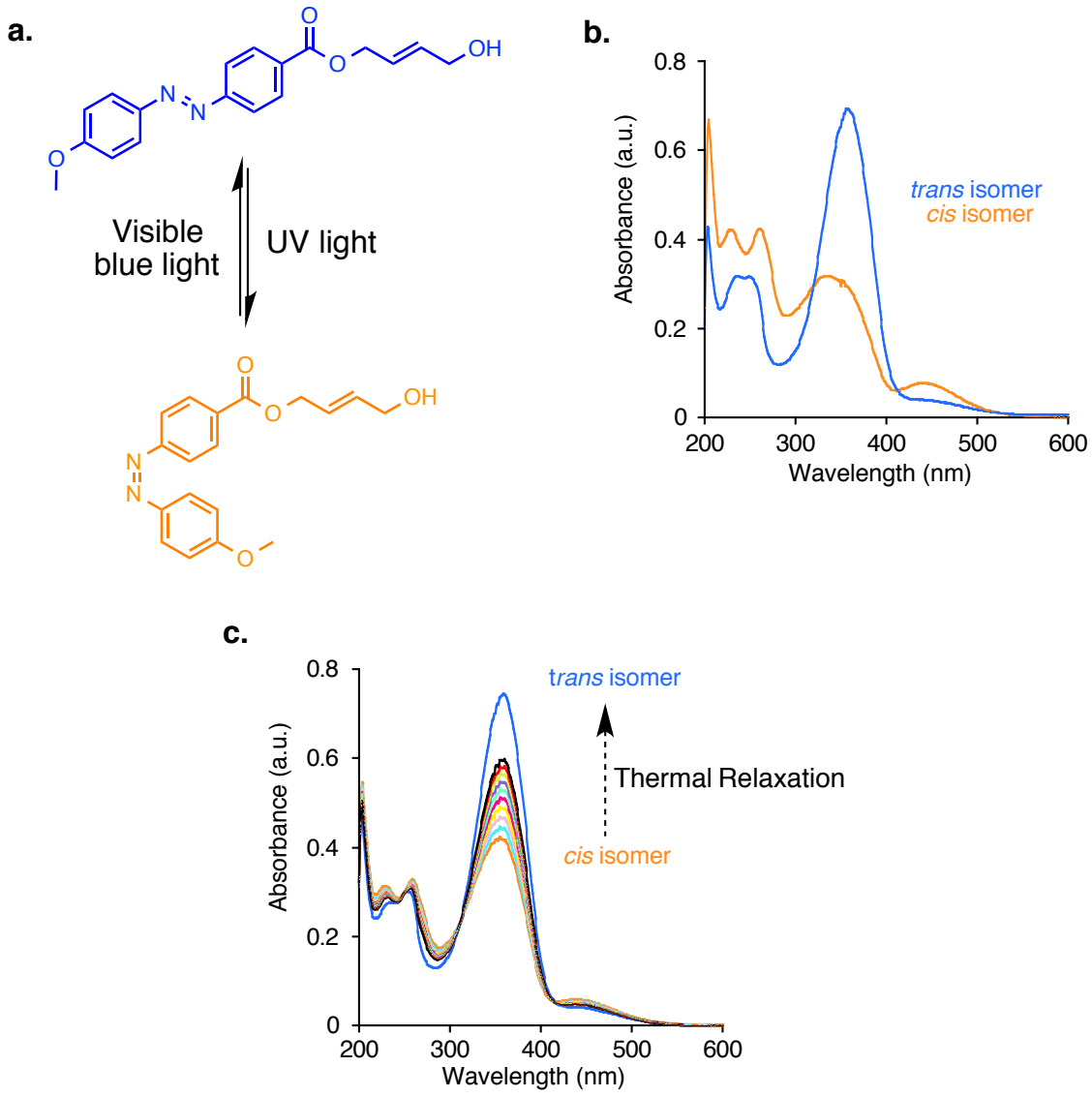


Figure III-8: Characterization of **Azo-AA**. **a.** *Cis-trans* isomerization of **Azo-AA**. **b.** UV-vis spectra of **Azo-AA** isomers in ethanol. **c.** Thermal *cis* → *trans* relaxation monitored by UV-vis. *Trans* isomer (blue absorption) was irradiated with UV light to yield the *cis* isomer (orange absorption) and then kept at dark. Time laps between each scan is 1.5 h. Black spectrum shows the final measurement after 12 h.

protein (~12 mg/mL) was incubated with an excess amount of **Azo-AA** (4 equivalents) at room temperature over 12 h and subjected to hanging drop crystallization method. Luckily, we were able to obtain several single hCRBPII/**Azo-AA** crystals (**Figure IV-10c**) and collect diffraction data. Meisam Nosrati refined the crystallographic data in which four protein molecules per asymmetric unit were found by molecular replacement. Although not all the

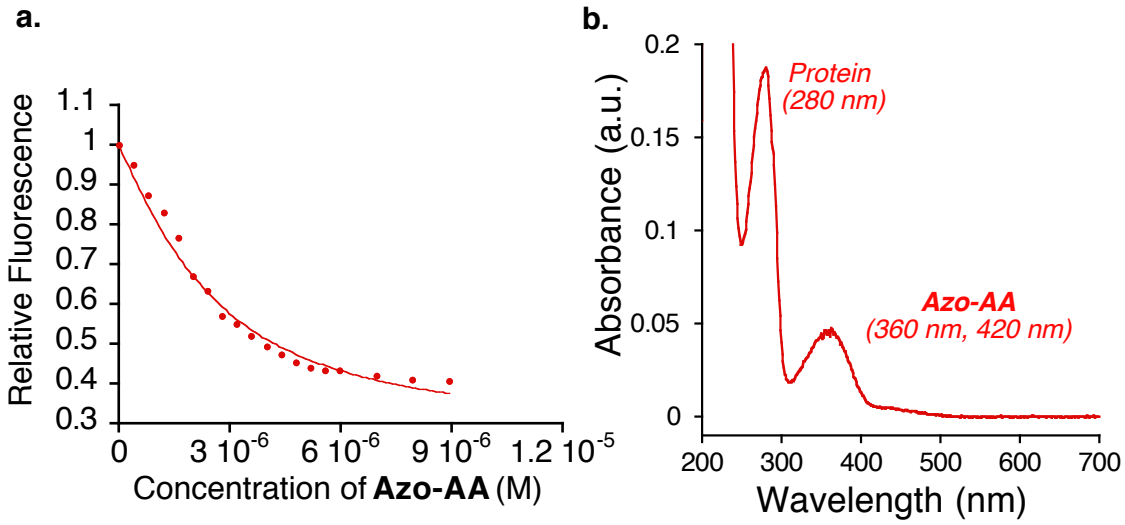


Figure III-9: Binding analysis of **Azo-AA** to wild type hCRBPII. **a.** Fluorescence quenching assay of hCRBPII with **Azo-AA**. **b.** UV-vis spectrum of hCRBPII/**Azo-AA** complex after overnight incubation.

chains displayed an occupied protein cavity, chain C provided a reasonable electron density for the **Azo-AA** in the ligand binding site. Analysis of the crystal structure of the complex demonstrated that the *trans* isomer of the **Azo-AA** chromophore is fully encapsulated by the protein, similar to that seen with the native substrate bound (all-*trans*-retinol) (**Figure III-10**). The second phenyl ring of the **Azo-AA**, which is substituted with the methoxide group, occupied the same location as the β-ionone ring of retinol (see overlaid structures in **Figure III-10c**).

Furthermore, the allylic alcohol moiety of the azo chromophore adopts a conformation that enables it to interact with Gln108 and Lys40 residues much like that observed retinol bound structure (**Figure III-7b** and **Figure III-10c**). We also

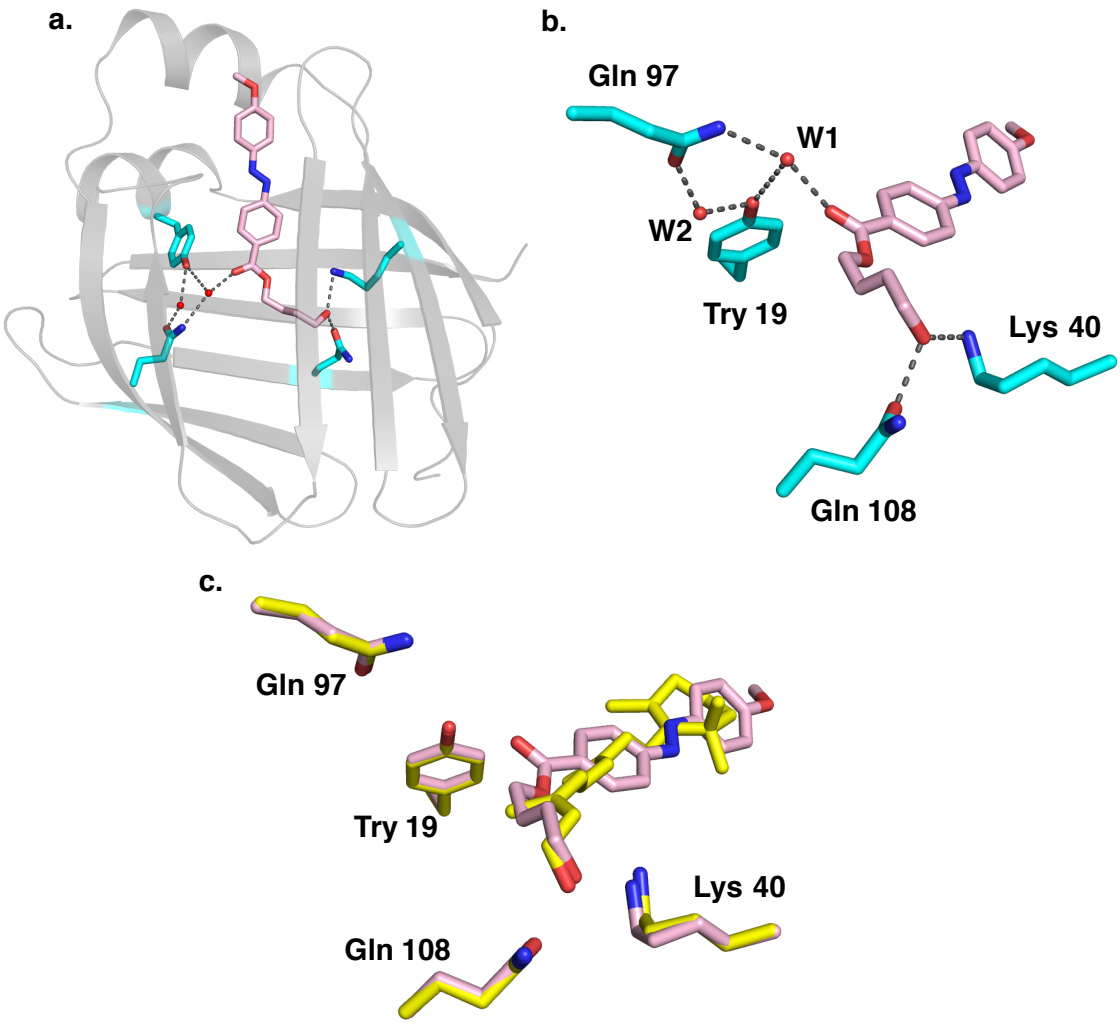


Figure III-10: Crystal structure of the wild type hCRBPII complexed with **Azo-AA**. **a.** The tertiary structure of hCRBPII protein (gray cartoon) bound to **Azo-AA** (pink). The important residues are highlighted (cyan sticks) inside the binding cavity. **b.** Expanded view of the ligand binding site. Hyrogen bonding interactions (water-mediated or direct) are indicated by dash lines between the side chains of the amino acid residues (cyan) and the bound chromophore (pink). The water molecules are shown as red spheres (**W1** and **W2**). **c.** Overlaid structure of the wild type hCRBPII/all-*trans*-retinol (yellow, PDB ID: 4QZT) and hCRBPII/**Azo-AA** (pink) complexes.

noticed that the carbonyl of the ester group participates in water-mediated hydrogen bonding interactions with the binding cavity residues. The small water network that includes the side chains of Gln97 and Try19 further fortifies the interaction between the azo chromophore and the protein. However, the binding affinity of **Azo-AA** with hCRBPII protein is in the micro molar range, which is not considered as a tight protein-ligand pair. Therefore, it is desirable to apply a rational design strategy, with available crystallographic data, to maximize the interaction of the protein with its ligand partner.

III.4 Structure Guided Mutagenesis In HCRBPII

In pursuit of developing the hCRBPII/**Azo-AA** system as a robust photoactivatable protein tag, several point mutations were applied to increase its affinity for the **Azo-AA** ligand. At the outset, seven amino acid residues residing at different zones of the chromophore were targeted (**Figure III-11**). As discussed above, the alcohol and ester moieties of the **Azo-AA** are important to form hydrogen bonding interactions with the surrounding residues. Switching hydrophobic interactions to hydrogen bonding interactions could be a good strategy to introduce new hydrogen bonds between the chromophore and the protein, providing a strong source of specificity. Therefore, we focused on identifying hydrophobic residues in close proximity to the bound ligand. We noticed that two hydrophobic residues, Leu117 and Leu119 could be good candidates for installing polar amino acids, which can act as hydrogen donors or acceptors. From this perspective, the mutations of L117E and L119E were tested

to study the change in binding affinity of the resultant protein variants. Nevertheless, both L117E and L119E mutations resulted in destabilizing the complex formation, increasing the K_d from 1.42 μM (wild type hCRBPII) to 2.15 μM and 4.54 μM, respectively. Despite being within hydrogen bonding distance of the chromophore, these mutations might be destabilizing due to the unfavorable packing.

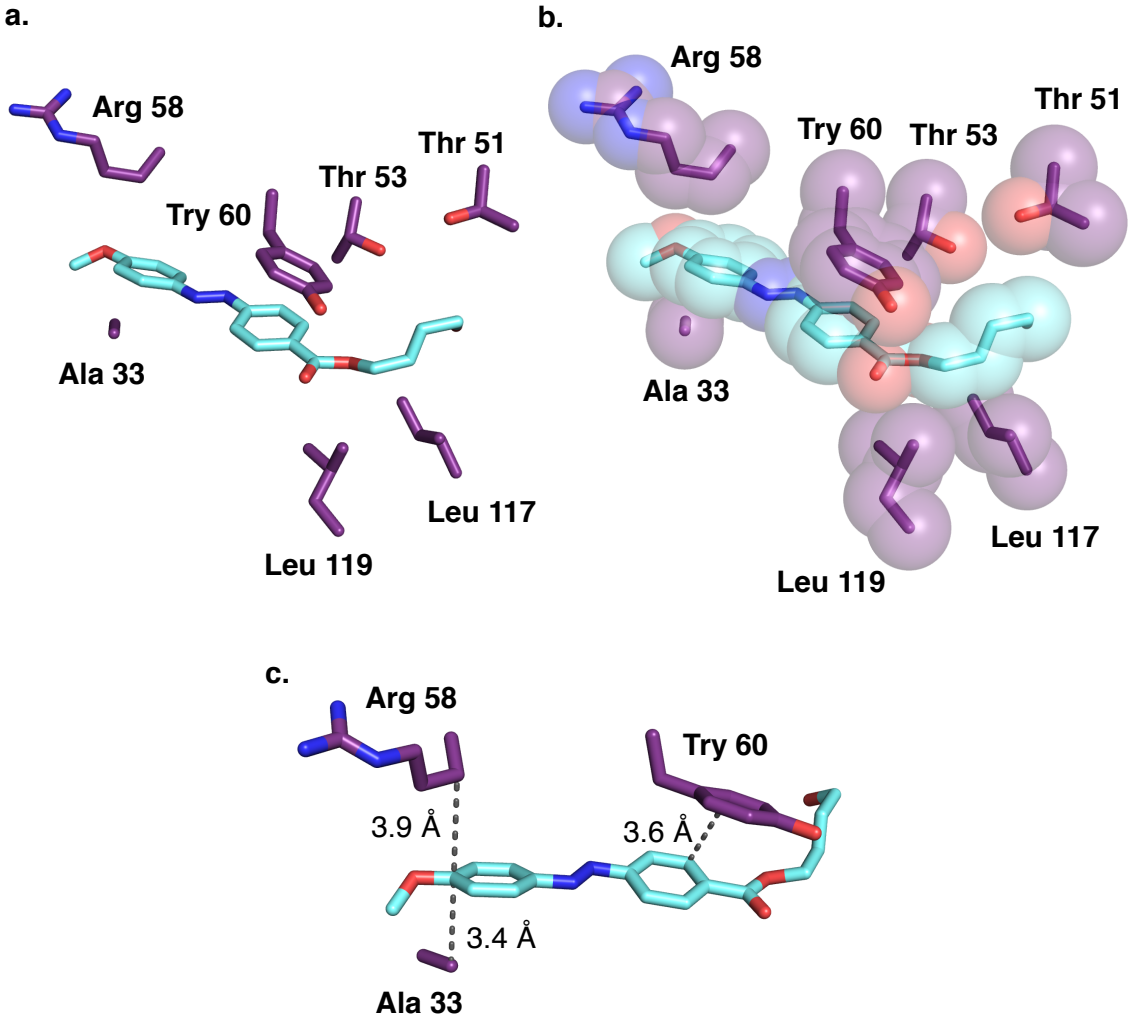


Figure III-11: Illustration of mutated residues discussed in the text. **a.** Stick model of residues (purple) surrounding the **Azo-AA** ligand (cyan). **b.** The space filling depiction of **a**. **c.** Residues in close proximity to the phenyl rings of the bound chromophore.

The initial failure of the design strategy relying on creating buried hydrogen bonding interactions led us to reverse our approach. We aimed to make use of polar to non-polar swaps in order to provide a suitable hydrophobic binding pocket for the accommodated chromophore. Our previous study in wavelength regulation of retinal-based ligands indicated that replacing threonine residues with hydrophobic amino acids could strongly change the polarity of the binding pocket, resulting in a substantial shift in the absorption spectrum of bound retinylidene chromophores. Consequently, two threonine residues, Thr51 and Thr53, in close proximity to the **Azo-AA** were identified and mutated to non-polar residues. For both T51V and T53A mutations, we observed an increase in binding affinity (hCRBPII-T51V K_d = 0.71 μM, hCRBPII-T53A K_d = 1.09 μM). Although the change in binding affinity was not substantial, these results were encouraging with respect to a single point mutation at the binding cavity.

To further tune the polarity of the binding site and to find more favorable contacts between the protein and the bound ligand, a series of hydrophobic aromatic substitutions were made. The hope was to promote π-stacking interactions between aromatic amino acids and the bound chromophore and/or contribute to the hydrophobicity of the protein interior. To start with, Tyr60 residing 3.6 Å away from the first aromatic ring of the **Azo-AA** was replaced with a Trp residue in the hope that a large surface of the Trp might result in a strong π-stacking interaction with the ligand (**Figure III-11c**). However, Y60W weakened binding affinity by ~6-fold in comparison to the wild type protein. Next, the

residues close to the second phenyl ring (methoxide substituted) of the azo chromophore were targeted. Similar to Y60W, A33W mutation did not stabilize the protein-chromophore interaction, resulting in a significant loss in binding affinity (hCRBPII-A33W $K_d \approx 16$ μM). In sharp contrast to the aforementioned Y60W and A33W results, R58W provided a better protein-ligand pair with a K_d value of 0.79 μM. It is interesting that the R58W mutation was critically important in our previous wavelength regulation studies. As discussed earlier (Chapter II), this residue is highly effective in red shifting the absorption profiles of retinylidene ligands, particularly all-*trans*-retinal and 3-dehydro-retinal, by sequestering the protein cavity and forming strong hydrophobic interactions with the chromophore head groups. It is worth mentioning that similar factors could play a role in the increased binding affinity.

III.4.1 Addition Of Affinity Enhancing Mutations In HCRBPII-R58W Scaffold

Because the addition of R58W mutation was a promising head for creating a stable protein-ligand pair, we tested if we could make further optimizations by including additional mutations to the hCRBPII-R58W scaffold. Similar to the Arg58 residue, Leu77 is another amino acid that resides at the mouth of the binding cavity, making close contact with the **Azo-AA** ligand (**Figure III-12a**). It seemed reasonable that introducing an aromatic side chain at position 77 could give similar results obtained with R58W. Therefore, we constructed two mutants, L77F and L77W, of the hCRBPII-R58W variant, expecting an increase in binding

specificity. However, both double mutants failed in providing higher affinities than the parent R58W construct (**Table III-1**, entries **3** and **4**). Perhaps, installation of large aromatic residues results in unfavorable steric congestion.

Another well-known strategy for redesigning binding specificities is to modify the complementary shape of the ligand-protein interface via the "bump and hole" strategy.[54,55] In this design, accommodation of the target ligand can be modulated by pruning back or filling space with neighboring residues. To illustrate, if the ligand is too large to fit within the binding site of the wild type protein, then creating a hole is an appropriate approach to enlarge the ligand binding cavity. On the other hand, smaller ligands can be tightly packet within the active site of the protein by filling the space (creating bumps) with larger amino acid substitutions. Since **Azo-AA** occupies larger surface area in comparison with the natural substrate of hCRBPII, carving the binding cavity could provide a

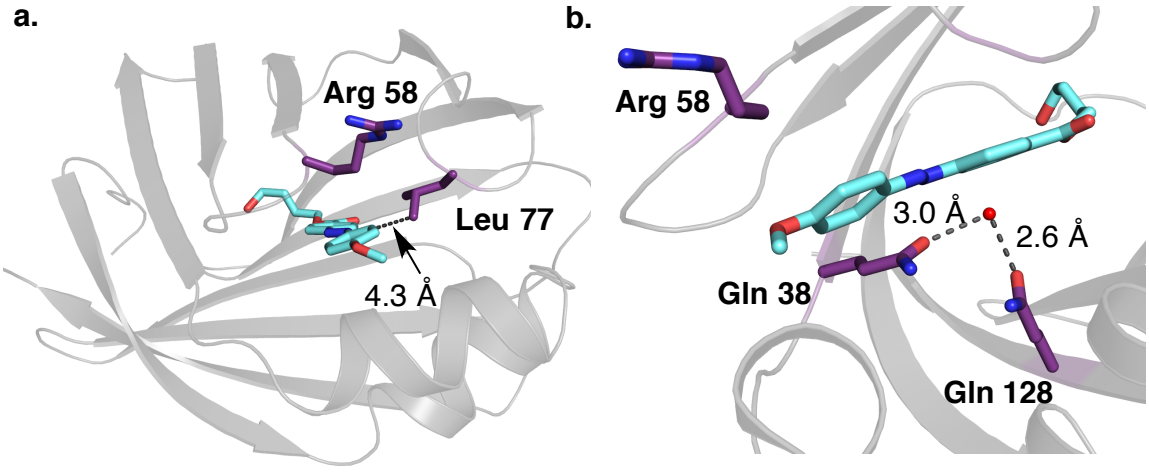


Figure III-12: Proposed sites for mutagenesis in hCRBPII/**Azo-AA** crystal structure. **a.** The location of Arg58 and Leu77 residues on loops connecting the β-strands. **b.** Expanded view of the binding site to highlight the polar Gln residues. The water-mediated hydrogen bonding interaction is shown by dash lines. The water molecule is represented as red sphere.

favorable ligand-protein interface. Applying this approach, we focused on pruning back residues around the ligand interface to create an open space that can allow deeper binding of the chromophore. At the outset, we screened a series hCRBPII variants in which the Leu residues at positions of 117 and 119 were replaced with smaller amino acid side chains such as Val and Ser (**Table III-1**, entries **6-9**). Once again it was observed that polar replacements at these positions are not favorable and dramatically destabilize complex formation. However, non-polar to non-polar swaps by smaller residues slightly improved the chromophore association with the protein. Following with the same approach, the residues found in the middle zone of the bound ligand were also substituted with non-polar

Table III-1. Binding affinity measurements of hCRBPII mutants to **Azo-AA**.

Entry	hCRBPII Variants	K_d Values
1	Wild Type	1.42 μM
2	R58W	0.79 μM
3	R58W:L77F	3.67 μM
4	R58W:L77W	2.23 μM
5	R58W:F16W	p.b.
6	R58W:L117S	p.b.
7	R58W:L117V	0.11 μM
8	R58W:L117S:L119S	p.b.
9	R58W:L117V:L119V	1.64 μM
10	R58W:Q128L	0.65 μM
11	R58W:Q128V	0.51 μM
12	R58W:Q38L	0.94 μM
13	R58W:Q38V	99.16 nM
14	R58W:Q38V:Q128V	4.74 μM
15	R58W:Q38V:L117V	0.58 μM
16	R58W:Q38V:L117V:L119V	1.68 μM
17	R58W:Q38V:T51A:T53A	0.61 μM
18	R58W:Q38V:T51G:T53S	4.45 μM

p.b. = poor binding

groups. We exchanged two Gln residues residing near the azobenzene core with hydrophobic amino acids, Leu and Val (**Figure III-12b**). It was exciting that we obtained one of the best experimental results with these substitutions. Particularly, the combination of R58W and Q38V resulted in a large increase in binding affinity, providing an 8-fold improvement (**Table III-1**, entries **2** versus **13**). This double mutant lowered the K_d value from 1.42 μM (wild type) to ~100 nM, providing a stabilized protein-ligand complex. Having a stronger binder in hand, we further investigated the additive effects of other point mutations that resulted in enhanced binding specificity. The combinations of numerous point mutations were tested, including Leu117, Thr51, and Thr53 together with the aforementioned R58W:Q38V double mutation. Unfortunately, we did not observe any improvement in the binding affinity (**Table III-1**, entries **14** to **18**).

III.5 Conclusion And Future Work

Creating an optimal protein-ligand pair is challenging because there are several interactions that require optimization simultaneously to dramatically improve the specificity of the protein to its cognitive ligand. These interactions such as salt bridges, hydrogen bonds, Van der Waals contacts, hydrophobic-hydrophobic and dipole interactions do not only change the binding energies but also determine the geometry and the orientation of the ligand within the binding site. This means that optimizing one interaction could conflict with optimizing others and thus could be unpredictable. Herein, we demonstrate our preliminary results in design of a photoactivatable protein switch by using an azobenzene-

based chromophoric system. Despite screening a small set of protein variants, we were able to obtain a single protein mutant that shows favorable interactions with the synthetic azo ligand. Amongst the applied strategies, we observed that the swap of hydrophobic for polar interactions gave better results in enhanced binding affinity. However, we are well aware of the fact that more mutants should be screened to find tight binders for the *trans* isomer, while showing poor affinity for the *cis* form. A future project could apply different strategies for identifying more diverse sets of protein mutants with the desired characteristics. Genetic screening is one approach to test large libraries of protein mutants (via installation of random mutations) but requires intense experimental work. Alternatively, structure-based molecular modeling has become a widely used method to predict and also design protein-protein or protein-ligand interfaces. The computational approach can allow quick and accurate predictions and avoid overwhelming experimental protocols. Therefore, a computational design method could be a proper choice for future endeavors, especially with the readily crystallizable nature of the hCRBPII scaffold.

.

REFERENCES

REFERENCES

1. Venkataraman, K. *The Chemistry Of Synthetic Dyes*; New York: Academic Press, 1956.

2. Griffith.J, "Selected Aspects of Photochemistry .2. Photochemistry of Azobenzene and Its Derivatives" *Chem Soc Rev* **1972**, *1*, 481.

3. Rau, H., "Spectroscopic Properties of Organic Azo-Compounds" *Angew Chem Int Edit* **1973**, *12*, 224.

4. Tiberio, G.; Muccioli, L.; Berardi, R.; Zannoni, C., "How Does the Trans-Cis Photoisomerization of Azobenzene Take Place in Organic Solvents?" *Chemphyschem* **2010**, *11*, 1018.

5. Hartley, G. S., "The cis-form of azobenzene" *Nature* **1937**, *140*, 281.

6. Fliegl, H.; Kohn, A.; Hattig, C.; Ahlrichs, R., "Ab initio calculation of the vibrational and electronic spectra of trans- and cis-azobenzene" *J Am Chem Soc* **2003**, *125*, 9821.

7. Dias, A. R.; Dapiedade, M. E. M.; Simoes, J. A. M.; Simoni, J. A.; Teixeira, C.; Diogo, H. P.; Yang, M. Y.; Pilcher, G., "Enthalpies of Formation of Cis-Azobenzene and Trans-Azobenzene" *J Chem Thermodyn* **1992**, *24*, 439.

8. Conti, I.; Garavelli, M.; Orlandi, G., "The different photoisomerization efficiency of azobenzene in the lowest n pi* and,pi pi* singlets: The role of a phantom state" *J Am Chem Soc* **2008**, *130*, 5216.

9. Rau, H.; Luddecke, E., "On the Rotation-Inversion Controversy on Photo-Isomerization of Azobenzenes - Experimental Proof of Inversion" *J Am Chem Soc* **1982**, *104*, 1616.

10. Rau, H.; Luddecke, E.; Nitsch, H.; Patzelt, H., "The Mechanism of Photo-Isomerization of Azobenzenes" *B Soc Chim Belg* **1982**, *91*, 475.

11. Binenboy.J; Burcat, A.; Lifshitz, A.; Shamir, J., "Energy Barrier of Cis-Trans Isomerization of Difluorodiazine" *J Am Chem Soc* **1966**, *88*, 5039.

12. Talaty, E. R.; Fargo, J. C., "Thermal Cis-Trans-Isomerization of Substituted Azobenzenes - a Correction of Literature" *Chem Commun* **1967**, 65.

13. Lefevre, R. J. W.; Northcott, J., "The Effects of Substituents and Solvents on the Cis-]Trans Change of Azobenzene" *J Chem Soc* **1953**, 867.

14. Diau, E. W. G., "A new trans-to-cis photoisomerization mechanism of azobenzene on the S-1(n,pi*) surface" *J Phys Chem A* **2004**, *108*, 950.

15. Diau, E. W. G.; Chang, C. W., "Resolution of the inversion-rotation controversy of azobenzene in solution via femtosecond fluorescence anisotropy on S1 excitation." *Abstr Pap Am Chem S* **2004**, *228*, U287.

16. Garcia-Amoros, J.; Velasco, D., "Recent advances towards azobenzene-based light-driven real-time information-transmitting materials" *Beilstein J Org Chem* **2012**, *8*, 1003.

17. Garcia-Amoros, J.; Diaz-Lobo, M.; Nonell, S.; Velasco, D., "Fastest Thermal Isomerization of an Azobenzene for Nanosecond Photoswitching Applications under Physiological Conditions" *Angewandte Chemie-International Edition* **2012**, *51*, 12820.

18. Beharry, A. A.; Sadovski, O.; Woolley, G. A., "Azobenzene Photoswitching without Ultraviolet Light" *J Am Chem Soc* **2011**, *133*, 19684.

19. Beharry, A. A.; Woolley, G. A., "Azobenzene photoswitches for biomolecules" *Chem Soc Rev* **2011**, *40*, 4422.

20. Yager, K. G.; Barrett, C. J., "Novel photo-switching using azobenzene functional materials" *J Photoch Photobio A* **2006**, *182*, 250.

21. Beharry, A. A.; Woolley, G. A., "Photoswitch Design" *Neuromethods* **2011**, *55*, 171.

22. Goulet-Hanssens, A.; Corkery, T. C.; Priimagi, A.; Barrett, C. J., "Effect of head group size on the photoswitching applications of azobenzene Disperse Red 1 analogues" *J Mater Chem C* **2014**, *2*, 7505.

23. Patel, S. A.; Sinha, S.; Mishra, A. N.; Kamath, B. V.; Ram, R. N., "Olefin epoxidation catalysed by Mn(II) Schiff base complex in heterogenised-homogeneous systems" *J Mol Catal a-Chem* **2003**, *192*, 53.

24. Wegner, H. A., "Azobenzenes in a New LightuSwitching In Vivo" *Angewandte Chemie-International Edition* **2012**, *51*, 4787.

25. Garcia-Amoros, J.; Martinez, M.; Finkelmann, H.; Velasco, D., "Kinetico-Mechanistic Study of the Thermal Cis-to-Trans Isomerization of 4,4 '-Dialkoxyazoderivatives in Nematic Liquid Crystals" *J Phys Chem B* **2010**, *114*, 1287.

26. Sierocki, P.; Maas, H.; Dragut, P.; Richardt, G.; Vogtle, F.; Cola, L.; Brouwer, F. A. M.; Zink, J. I., "Photoisomerization of azobenzene derivatives in nanostructured silica" *J Phys Chem B* **2006**, *110*, 24390.

27. Brode, W. R.; Gould, J. H.; Wyman, G. M., "The Relation between the Absorption Spectra and the Chemical Constitution of Dyes .25. Phototropism and Cis-Trans Isomerism in Aromatic Azo Compounds" *J Am Chem Soc* **1952**, *74*, 4641.

28. Dunn, N. J.; Humphries, W. H.; Offenbacher, A. R.; King, T. L.; Gray, J. A., "pH-Dependent cis -> trans Isomerization Rates for Azobenzene Dyes in Aqueous Solution" *J Phys Chem A* **2009**, *113*, 13144.

29. Sawicki, E., "Physical Properties of the Aminoazobenzene Dyes .8. Absorption Spectra in Acid Solution" *J Org Chem* **1957**, *22*, 1084.

30. Schanze, K. S.; Mattox, T. F.; Whitten, D. G., "Solvent Effects Upon the Thermal Cis-Trans Isomerization and Charge-Transfer Absorption of "4-(Diethylamino)-4'-Nitroazobenzene" *J Org Chem* **1983**, *48*, 2808.

31. Jaffe, H. H.; Yeh, S. J.; Gardner, R. W., "The Electronic Spectra of Azobenzene Derivatives and Their Conjugate Acids" *J Mol Spectrosc* **1958**, *2*, 120.

32. Garcia-Amoros, J.; Bucinskas, A.; Reig, M.; Nonell, S.; Velasco, D., "Fastest molecular photochromic switches based on nanosecond isomerizing benzothiazolium azophenolic salts" *J Mater Chem C* **2014**, *2*, 474.

33. Garcia-Amoros, J.; Massad, W. A.; Nonell, S.; Velasco, D., "Fast Isomerizing Methyl Iodide Azopyridinium Salts For Molecular Switches" *Organic Letters* **2010**, *12*, 3514.

34. Coelho, P. J.; Carvalho, L. M.; Moura, J. C. V. P.; Raposo, M. M. M., "Novel photochromic 2,2 '-bithiophene azo dyes" *Dyes Pigments* **2009**, *82*, 130.

35. Nishimura, N.; Sueyoshi, T.; Yamanaka, H.; Imai, E.; Yamamoto, S.; Hasegawa, S., "Thermal Cis-to-Trans Isomerization of Substituted Azobenzenes .2. Substituent and Solvent Effects" *B Chem Soc Jpn* **1976**, *49*, 1381.

36. Bunce, N. J.; Ferguson, G.; Forber, C. L.; Stachnyk, G. J., "Sterically Hindered Azobenzenes - Isolation of Cis Isomers and Kinetics of Thermal Cis-] Trans Isomerization" *J Org Chem* **1987**, *52*, 394.

37. Weston, C. E.; Richardson, R. D.; Haycock, P. R.; White, A. J. P.; Fuchter, M. J., "Arylazopyrazoles: Azoheteroarene Photoswitches Offering

Quantitative Isomerization and Long Thermal Half-Lives" *J Am Chem Soc* **2014**, *136*, 11878.

38. Wendler, T.; Schutt, C.; Nather, C.; Herges, R., "Photoswitchable Azoheterocycles via Coupling of Lithiated Imidazoles with Benzenediazonium Salts" *J Org Chem* **2012**, *77*, 3284.

39. Cembran, A.; Bernardi, F.; Garavelli, M.; Gagliardi, L.; Orlandi, G., "On the mechanism of the cis-trans isomerization in the lowest electronic states of azobenzene: S-0, S-1, and T-1" *J Am Chem Soc* **2004**, *126*, 3234.

40. Dokic, J.; Gothe, M.; Wirth, J.; Peters, M. V.; Schwarz, J.; Hecht, S.; Saalfrank, P., "Quantum Chemical Investigation of Thermal Cis-to-Trans Isomerization of Azobenzene Derivatives: Substituent Effects, Solvent Effects, and Comparison to Experimental Data" *J Phys Chem A* **2009**, *113*, 6763.

41. Ikegami, T.; Kurita, N.; Sekino, H.; Ishikawa, Y., "Mechanism of cis-to-trans isomerization of azobenzene: Direct MD study" *J Phys Chem A* **2003**, *107*, 4555.

42. Norikane, Y.; Kitamoto, K.; Tamaoki, N., "Novel crystal structure, Cis-Trans isomerization, and host property of meta-substituted macrocyclic azobenzenes with the shortest linkers" *J Org Chem* **2003**, *68*, 8291.

43. Rau, H.; Rottger, D., "Photochromic Azobenzenes Which Are Stable in the Trans and Cis Forms" *Mol Cryst Liq Crys A* **1994**, *246*, 143.

44. Rottger, D.; Rau, H., "Photochemistry of azobenzenophanes with three-membered bridges" *J Photoch Photobio A* **1996**, *101*, 205.

45. Nagamani, S. A.; Norikane, Y.; Tamaoki, N., "Photoinduced hinge-like molecular motion: Studies on xanthene-based cyclic azobenzene dimers" *J Org Chem* **2005**, *70*, 9304.

46. Beharry, A. A.; Sadovski, O.; Wong, L.; Tropepe, V.; Woolley, G. A., "Azobenzene Photo-switches for Cellular Applications" *Biopolymers* **2011**, *96*, 431.

47. Emoto, A.; Uchida, E.; Fukuda, T., "Optical and Physical Applications of Photocontrollable Materials: Azobenzene-Containing and Liquid Crystalline Polymers" *Polymers-Basel* **2012**, *4*, 150.

48. Mahimwalla, Z.; Yager, K. G.; Mamiya, J.; Shishido, A.; Priimagi, A.; Barrett, C. J., "Azobenzene photomechanics: prospects and potential applications" *Polym Bull* **2012**, *69*, 967.

49. Merino, E.; Ribagorda, M., "Control over molecular motion using the cis-trans photoisomerization of the azo group" *Beilstein J Org Chem* **2012**, *8*, 1071.

50. Russew, M. M.; Hecht, S., "Photoswitches: From Molecules to Materials" *Adv Mater* **2010**, *22*, 3348.

51. Zatsepin, T. S.; Abrosimova, L. A.; Monakhova, M. V.; Hien, L. T.; Pingoud, A.; Kubareva, E. A.; Oretskaya, T. S., "Design of photocontrolled biomolecules based on azobenzene derivatives" *Russ Chem Rev+* **2013**, *82*, 942.

52. Chaudhuri, B. N.; Kleywegt, G. J.; Broutin-L'Hermite, I.; Bergfors, T.; Senn, H.; Le Motte, P.; Partouche, O.; Jones, T. A., "Structures of cellular retinoic acid binding proteins I and II in complex with synthetic retinoids" *Acta Crystallogr D* **1999**, *55*, 1850.

53. Jones, T. A.; Bergfors, T.; Sedzik, J.; Unge, T., "The 3-Dimensional Structure of P2 Myelin Protein" *Embo J* **1988**, *7*, 1597.

54. Belshaw, P. J.; Schoepfer, J. G.; Liu, K. Q.; Morrison, K. L.; Schreiber, S. L., "Rational Design of Orthogonal Receptor-Ligand Combinations" *Angew Chem Int Edit* **1995**, *34*, 2129.

55. Belshaw, P. J.; Schoepfer, J. G.; Liu, K.; Morrison, K. L.; Schreiber, S. L., "Rational Design of New Receptor-Ligand Pairs Related to Cyclophilin Cyclosporine" *Abstr Pap Am Chem S* **1995**, *210*, 137.

CHAPTER IV: MATERIALS AND METHODS

IV.1 Mutagenesis Of HCRBPII And CRABPII

Mutagenesis of all CRABPII and hCRBPII proteins was performed with the PET17b-hCRABPII and PET17b-hCRBPII vector, as described below. The polymerase chain reaction (PCR) was done according to specified PCR conditions (**Table IV-1**). The primers used for the mutations were ordered from Integrated DNA Technologies, having a range of melting point between 52 °C to 66 °C (depending on primer sequence). PCR products were subjected to *Dpn* I (New England BioLabs®) digestion in order to destroy the original template DNA. *Dpn* I restriction enzyme (12 units) was added to 30 μL of PCR sample and the reaction mixture was incubated at 37 °C for 1 h. The resulting mixture was transformed to XL-1 Blue super (Novagen®) competent cells for DNA amplification and purification.

XL-1 Blue competent cells (100 μL) thawed on ice and 7 μL of PCR product was added to the cells. After incubation on ice for 30 min, the cells were heat-shocked at 42 °C for 45 seconds and were gently spread on LB agar plate containing ampicillin (100 μg/mL) and tetracycline (7.5 μg/mL) and incubated at 37 °C for 12 to 16 h.

A single colony was picked from the agar plate and inoculated in 15 mL of LB solution containing ampicillin (100 μg/mL) and tetracycline (7.5 μg/mL). The

cell culture was grown at 37 °C for 12 to 16 h and the cells were harvested via centrifucation at 13000 rpm for 1 min. DNA isolation was performed by using Promega Wizard® Plus SV Miniprep (A1330) DNA purification kit according to the manufacturers' directions, with the exception of using 50 μL of Nuclease-Free water for DNA elution instead of the recommended 100 μL. DNA concentration was measured by using the Thermo Scientific NanoDrop™ 1000 Spectrophotometer. The average DNA concentration in a 50 μL solution was 120

Table IV-1. PCR condition and protocol.

Total Reaction Volume	50 μL
Template (DNA plasmid)	100 ng (x μL)
Primer forward	20 pmol (y μL)
Primer reverse	20 pmol (z μL)
dNTP	1 μL
DMSO	5 μL
10× pfu buffer	5 μL
Pfu Turbo (DNApolymerase)	1 μL
DI water	38-(x+y+z) μL

PCR program		
1×	94 °C	3 min
20×	94 °C	20 sec
	temperature 3-5 °C lower than primer melting temperature	50 sec
	72 °C	4 min
1×	72 °C	8 min
1×	10 °C	5 min

ng/μL. A sample containing at least 700 ng of DNA was transferred into another eppendorf tube and sequenced by The Research Technology Support Facility at Michigan State University. A primer corresponding to the T7 promoter was used for sequencing.

IV.2 Protein Expression And Purification Of HCRBPII And HCRABPII

The target gene (100 ng of DNA for 100 μL of cells) was transformed into BL21(DE3) pLysS (Invitrogen™) *E. coli* competent cells for protein expression. A protocol similar to that described for transformation of PCR products was used. The only difference was that after heat-shock at 42 °C, the cell solution was spread directly onto an LB plate with ampicillin (100 μg/mL) and chloroamphenicol (170 μg/mL) antibiotics.

In order to grow a cell culture, a single colony was used from the agar plate to inoculate 1 L of modified Luria Broth* with ampicillin (100 μg/mL) and chloroamphenicol (170 μg/mL). The modified-LB solution contains: Tryptone (10 g), yeast extract (8 g), NaCl (5 g) and 1 L of dH_2O (sterilized by autoclave). The cell culture was grown at 37 °C while shaking for 6 to 9 h. The overexpression was induced by the addition of 1 mL of 1 M IPTG solution into 1 L cell culture (overall concentration 1.0 mM. The culture was shaken at 23 °C for 12 to 18 h at 220 rpm. In order to scale up protein expression, a second IPTG induction was

applied by addition of 0.5 mL of 1 M IPTG solution into the same 1 L cell culture after 12 h together with 200 mL of autoclaved modified-LB solution.

The cells, from the culture shaken at 23 °C for up to 18 h, were harvested at 4 °C by centrifugation for 12 min at 5000 rpm. The supernatant was discarded and the cells were resuspended in 50 mL of Tris buffer (10 mM Tris•HCl, pH=8.0), lysed by sonication (Power 60%, 1 min x 3), treated with DNAase (300 units/ 50 mL suspension) and $MgCl_2$ (0.12 mmol/ 50 mL suspension) and centrifuged at 16 °C (5000 rpm, 20 min).

The protein was purified by ion exchange chromatography using Q Sepharose™, Fast Flow resin (column diameter: ~4 cm; height: ~10 cm) at 4 °C. The FastQ column was pre-washed twice with 100 mL of a solution containing 2 M NaCl and 10 mM Tris•HCl, pH=8.0, and then equilibrated with 200 mL of Tris buffer (10 mM Tris•HCl, pH=8.0). The cell supernatant was then loaded onto the column, the column was washed with 40 mL of Tris buffer, and the protein was eluted with 45 mL of the elution buffer (10 mM Tris•HCl, 200 mM NaCl, pH=8.0). The eluent was desalted using an ultrafiltration cell under nitrogen pressure (~20 psi) equipped with a 10,000 MW cut-off membrane (ultrafiltration membrane, regenerated cellulose, filter code: YM10, diameter 63.5 mm, NMWL: 10,000). The protein was first concentrated to ~25 mL and then diluted to 160 mL with Tris buffer (10 mM Tris•HCl, pH=8.0). The resulting diluted solution was then re-

concentrated to ~25 mL. Further purification was performed with Fast Protein Liquid Chromatography (Biologic Duo Flow, Biorad) as described in **Table IV-2**. Most of the CRABPII and hCRBPII variants were collected between 4% to 15% elution.

Table IV-2. FPLC protocol.

	Description		Parameters
1.	Isocratic Flow	pH=8.1, 0% B	10.00 mL, 3.00 mL/min
2.	Load Sample	Dynamic Loop	'*Volume of protein*' mL, 2.0 mL/min
3.	Isocratic Flow	pH=8.1, 0% B	10.00 mL, 3.0 mL/min
4.	Linear Gradient	pH=8.1, 0→4% B	10.00 mL, 3.00 mL/min
5.	Isocratic Flow	pH=8.1, 4% B	20.00 mL, 3.00 mL/min
6.	Linear Gradient	pH=8.1, 4→8% B	10.00 mL, 3.00 mL/min
7.	Linear Gradient	pH=8.1, 8% B	20.00 mL, 3.00 mL/min
8.	Linear Gradient	pH=8.1, 8→15% B	10.00 mL, 3.00 mL/min
9.	Linear Gradient	pH=8.1, 15% B	15.00 mL, 3.00 mL/min
10.	Isocratic Flow	pH=8.1, 15→75% B	10.00 mL, 3.00 mL/min
11.	Isocratic Flow	pH=8.1, 100% B	20.00 mL, 3.00 mL/min
12.	Isocratic Flow	pH=8.1, 0% B	30.00 mL, 3.00 mL/min
13.	End protocol		

IV.3 Protein Characterization

Spectroscopic characterizations of purified proteins were carried out using a Cary 300 UV-Vis spectrophotometer and a Cary 5000 UV-Vis-NIR spectrophotometer (Agilent Varian Instruments). Fluorescent experiments were performed with a Fluorolog 3 spectrafluorimeter (Horiba Scientific).

IV.3.1 Extinction Coefficient Determination And Yield Calculation

The extinction coefficients of the proteins, measured at 280 nm, were determined according to the method described by Gill and Vonhippel.[1] The

theoretical extinction coefficient (ε^{Theo}) is calculated based on the following equation:

$$\varepsilon^{Theo} = 5690x + 1280y + 120z$$

where x, y and z are the number of Trp, Tyr and Cys residues, respectively, in the protein. The coefficients for x, y, and z are the extinction coefficients of these residues determined previously ($\varepsilon^{Trp} = 5690\ M^{-1}\ cm^{-1}$, $\varepsilon^{Tyr} = 1280\ M^{-1}\ cm^{-1}$, $\varepsilon^{Cys} = 120\ M^{-1}\ cm^{-1}$). The concentration of the protein was measured in 6 M-guanidine HCl (denaturing solution) solution following Beer's Law:

$$Abs^{280\ Denatured} = b.c.\varepsilon^{Theo}$$

where b is the cuvette path length and c is the concentration of the protein. The absorbance value ($Abs^{280\ Native}$) at the same concentration of protein was measured in phosphate buffered saline, PBS, (0.24 g KH_2PO_4, 1.45 g Na_2HPO_4, 0.8 g NaCl, 0.2 g KCl, pH=7.3, total volume 1 L) under identical settings. In this way, the experimental extinction coefficient (ε^{Exp}) of the protein can be derived:

$$\varepsilon^{Exp} = (Abs^{280\ Native} / Abs^{280\ Denatured}).\varepsilon^{Theo}$$

The yield of protein expression was calculated by measuring the absorbance value in PBS at 280 nm. The protein concentration was diluted by half (500 μL of

native buffer, PBS, mixed with 500 μL of protein solution) and the following calculations were applied:

$$c = 2(Abs^{280\ Native} / \varepsilon^{Exp}) \text{ and yield (g)} = c.V.M_W$$

where c is the concentration of protein solution (mol/L), V is the volume of protein solution (L) and M_W is the molecular weight of the protein (g/mol).

Table IV-3. Extinction coefficient of the hCRBPII mutants.

hCRBPII Variants	ε	hCRBPII Variants	ε
Q108K:K40L (KL)	27,800	WILD TYPE (WT)	28,900
KL:T53A	27,260	L117E	27,640
KL:T53V	27,900	L119E	27,450
KL:T53L	27,990	R58W	33,320
KL:T53S	27,200	R58W:L117S	34,770
KL:T51V (KLV)	27,800	R58W:L117V	31,410
KLV:T53A	27,520	R58W:L117S:L119S	33,770
KLV:T53V	24,850	R58W:L117V:L119V	33,930
KLV:T53L	25,640	T53A	28,090
KLV:T53S	27,200	T51V	28,260
KLV:Q4F	28,060	A33S	28,720
KLV:T53A:Q4F	27,440	A33W	35,670
KLV:T53V:Q4F	30,260	Y60W (monomer)	36,030
KLV:T53L:Q4F	27,800	R58W:L77F	29,450
KLV:T53S:Q4F	27,710	R58W:L77W	32,170
KLV:R58Y	28,800	R58W:L77A	33,930
KLV:T53A:R58Y	26,800	R58W:F16W	40,740
KLV:T53V:R58Y	32,570	R58W:Q128L	31,980
KLV:T53L:R58Y	30,790	R58W:Q128V	31,180
KLV:T53S:R58Y	29,780	R58W:Q38L	33,540
KLV:R58Y:Q4F	28,890	R58W:Q38V	34,200
KLV:T53A:R58Y:Q4F	27,530	R58W:Q38V:Q128L	27,190
KLV:T53V:R58Y:Q4F	28,070	R58W:Q38V:Q128V	38,470
KLV:T53L:R58Y:Q4F	27,990	R58W:Q38V:L117V:L119V	43,370
KLV:T53S:R58Y:Q4F	27,230	R58W:Q38V:T51A:T53A	41,630
KLV:Y19W	32,500	R58W:Q38V:T51G:T53S	35,640
KLV:Y19W:R58Y	32,800	R58W:Q38V:L117V	28,200

ε = extinction coefficient of the protein at 280 nm in units of $(M.cm)^{-1}$.

Table IV-4. Extinction coefficient of the CRABPII mutants.

CRABPII Variants	ε (M.cm)$^{-1}$
R132K:R111L (KL)	20,072
R132K:R111L:L121E (KLE)	17,400
R132K:R111L:L121D (KLD)	20,450
KL:R59W	25,980
KL:R59Y	22,140
KLD:R59W	25,990
KLE:R59W	24,500
KLE:R59H	20,200
KLE:R59Y	21,950
KLE:R59F	20,100
KLE:R59L	19,650
KLE:V76W	24,840
KLE:S37W	23,530
KLE:I31W	23,440
KLE:L28W	23,350
KLE:P39W	25,980
KLE:M123W	25,150
KLE:A32W	24,140
KLE:V76W:S37W	28,900
KLE:V76W:L28W	30,840
KLE:R59W:I31W	31,250
KLE:R59W:L28W	27,190
KLE:R59W:M123W	31,960
KLE:R59W:A32W	29,250
KLE:R59W:L28W:I31W:A32W	39,700
KL:L121Y	20,860
KL:L121Y:R59W	26,130
KL:Y134F	20,530
KLE:L19R	20,320
KLE:L19R:R59W	25,270
KLE:T56L	19,950
KLE:T56W	20,570
KLE:R59W:F15W	24,420
KLE:T56W:F15W	28,290
KLE:R59W:T56W	30,860
KLE:R59W:T56L	25,030
KLE:R59W:T56W:F15W	30,340
KLE:R59W:A32W:I31W	41,600
KLE:M123W:A32W	30,700

Mutants below the red line were not involved in Chapter I discussions.

The typical yield for CRABPII and hCRBPII mutants is ~40 mg/L of modified Luria Broth.

IV.3.2 Acid Dissociation Constant Measurements Of Protonated Schiff Bases

For pK_a (acid dissociation constant) measurements, protein (20 μM in PBS) was incubated with ligand (0.3 to 0.5 equivalents) at room temperature till full PSB formation. The UV spectrum and the pH of the solution were recorded after addition of each portion of 5 M NaOH solution. The λ^{max} of the protein-ligand complex versus pH was plotted (see examples in **Figure IV-1**). A polynomial fit of the data (3^{rd} power) was applied and the pK_a value was calculated when the second derivative of the equation was set to zero.

$$y = M_3X^3 + M_2X^2 + M_1X + M_o$$

$$y' = 3M_3X^2 + 2M_2X + M_1$$

$$y'' = 6M_3X + 2M_2$$

If $y''=0$ then pK_a $(X) = - [M_2 / 3M_3]$

IV.4 Denaturation Studies Of CRABPII Mutants

We carried denaturation studies only on CRABPII proteins in order to study the protein specificity towards the merocyanine ligand. CRABPII mutants (4 μM) were incubated with 0.2 equivalents of merocyanine aldehyde **1** in PBS at room

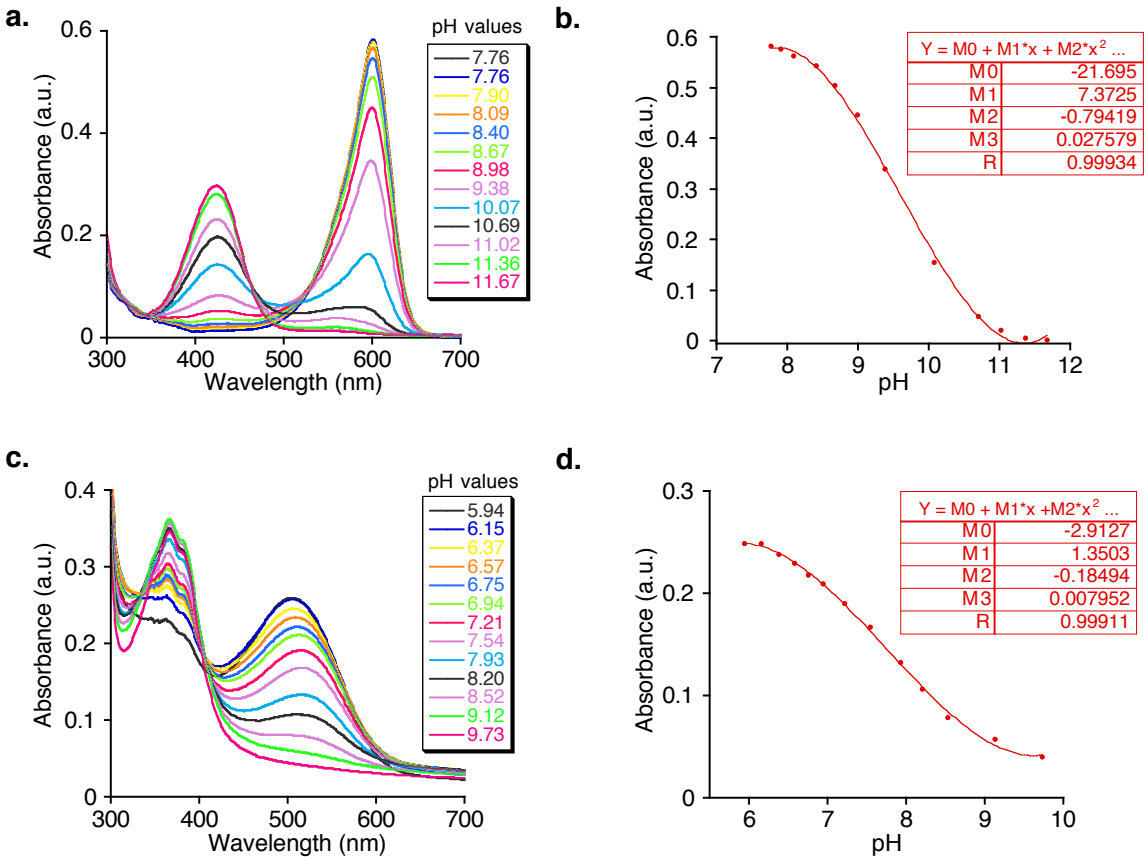


Figure IV-1: Derivation of pK_a values of protein-ligand protonated Schiff bases. **a.** Base titration of CRABPII-KL/merocyanine complex. The pH values are indicated in the box on the right. **b.** Absorbance of KL/merocyanine complex at λ^{max} (600 nm) as a function of pH with an apparent pK_a of 9.60. **c.** Base titration of hCRBPII-KL:T53C/5DEM complex. **d.** The plot of λ^{max} (521 nm) as a function of pH results in a pK_a value of 7.80.

temperature. Once the CRABPII/merocyanine complex formation was complete, 5 μL of 0.2 M SDS solution was added to the protein complex. UV-vis spectra were taken before and after SDS treatment. Since denatured proteins no longer encapsulate the covalently bound ligand (lose tertiary structure), the maximum absorption of the resulting water exposed ligand shifts from the wavelength native to the complex to 584 nm (**Figure I-21a** and **b**). The latter observed λ^{max}

corresponds exactly to that observed for merocyanine-iminium in PBS buffer with the 20 μM of BSA (**Figure I-21c** and **d**).

IV.5 Quantum Yield Measurements and Brightness Calculations of CRABPII/Merocyanine Complexes

All quantum yields (Φ) were determined by comparing the integrated area of the corrected emission spectra for each CRABPII/merocyanine complex with the corresponding integrated area obtained from a solution of fluorescent standards Oxazine-170 (purchased from across organics, lot# A0098689) or Oxazine-1 (purchased from Exciton, lot# 26424). Integrals at various concentrations were plotted against the absorbance obtained at the wavelength corresponding to the excitation wavelength and the slope of the curve obtained for protein-merocyanine complexes was compared to the slope of the curves found for reference fluorophores (**Figure IV-2**). We kept all the absorbance values at excitation wavelength under 0.1 absorbance unit. All samples were excited at 565 nm and the emission was collected from 575 nm to 800 nm. Temperature was set to 22 $^{\circ}$C and kept constant with a temperature controller during measurements. The absolute quantum yields for reference fluorophores were taken to be 0.579 for Oxazine-170 and 0.141 for Oxazine-1 in ethanol.[2] Quantum yield calculations were corrected for the refractive index difference between ethanol (η=1.36) and PBS (η=1.33). To calculate the quantum yield of CRABPII/merocyanine complexes, we applied the following equation:

$$\Phi_X = \Phi_{ST} (Grad^X / Grad^{ST}).(\eta^X / \eta^{ST})^2$$

Where the subscripts ST and X denote standard and test respectively, Φ is the fluorescence quantum yield, *Grad* is the gradient from the plot of integrated fluorescence intensity versus absorbance, and η the refractive index of the solvent.

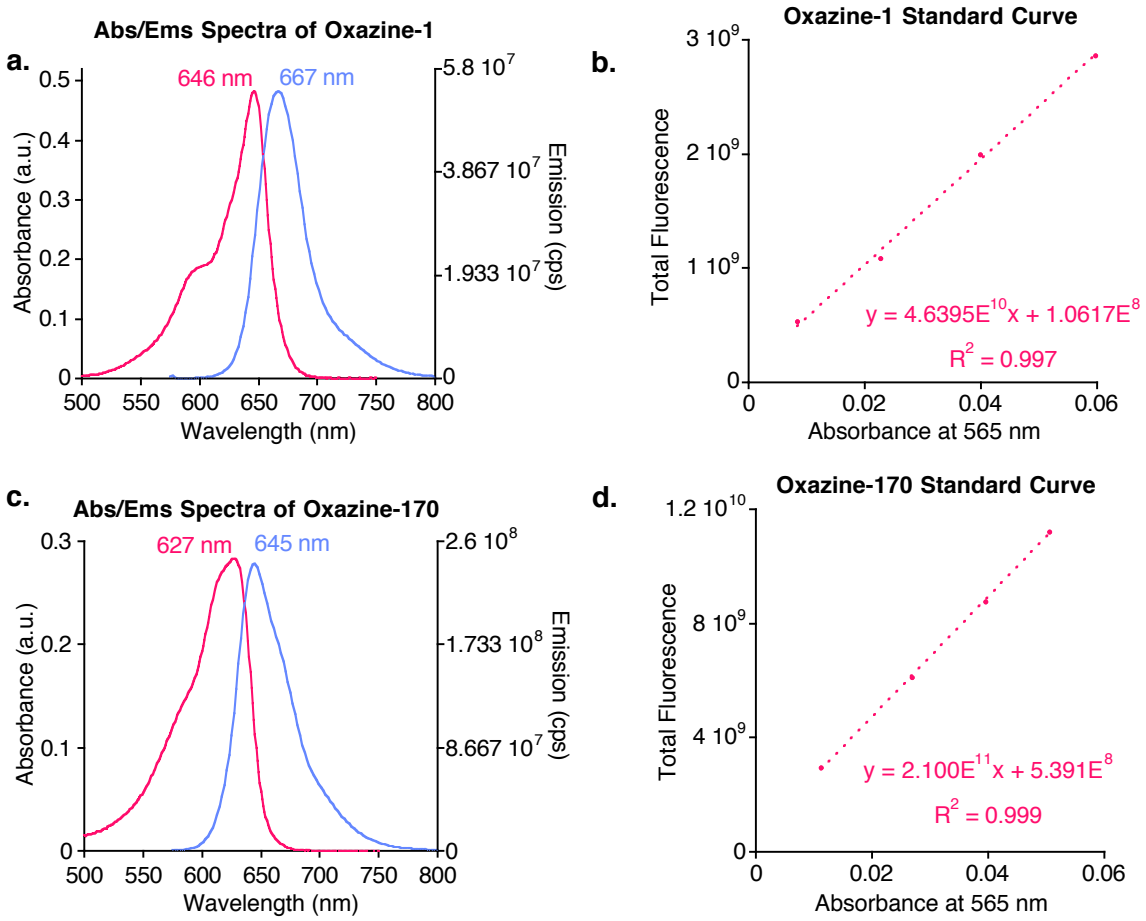


Figure IV-2: Spectroscopic characterization of Oxazine-1 and Oxazine-170. **a.** Absorption (pink) and emission (blue) spectra of Oxazine-1 in ethanol. **b.** The plot shows the total fluorescence emission (integrated area from 575 nm to 800 nm, excitation at 565 nm) as a function of absorption at 565 nm. Slope of the line is correlated to the quantum efficiency as described in the text. **c.** Absorption (pink) and emission (blue) spectra of Oxazine-170 in ethanol. **d.** Same analysis of **b** for the Oxazine-170 under identical settings.

The extinction coefficient of merocyanine aldehyde **1** was measured by UV-vis spectroscopy and found to be 77,000 $M^{-1}cm^{-1}$. Protein solutions (10 μM in PBS) were incubated with merocyanine aldehyde (0.8 mM stock solution in EtOH) to a final concentration of 2 μM. The solution was kept at room temperature for 1.5 h by which time maximum PSB absorbance is achieved. The absorbance value of each CRABPII/merocyanine complex was recorded at the corresponding λ^{max} and the extinction coefficient was determined via application of the Beer's law:

Table IV-5. Brightness calculations for CRABPII mutants complexed with merocyanine **1**.

Fluorescent Protein	Φ(%)	ε (M.cm)$^{-1}$	Brightness (% mKate2)	Brightness (% EGFP)
mKate2	40	62,500	100	74
mRFP	27	50,000	54	40
R132K:R111L (KL)	18	111,700	80	60
KL:L121E (KLE)	33	93,200	123	91
KL:L121D (KLD)	30	97,500	117	87
KL:R59W	23	140,100	129	95
KLE:R59W	39	169,800	265	196
KLE:V76W	33	77,600	102	76
KLE:S37W	30	96,500	116	87
KLE:L28W	31	58,600	73	54
KLE:A32W	31	104,300	129	96
KLE:V76W:L28W	32	67,000	86	63
KLE:R59W:L28W	38	158,400	241	178
KLE:R59W:A32W	39	142,900	223	165

Φ = quantum yield, **ε** = extinction coefficient of the fluorescent proteins and CRABPII/**1** complexes.

$$Abs^{CRABPII/Mero} = b.c.\varepsilon^{CRABPII/Mero}$$

$$\varepsilon^{CRABPII/Mero} = (Abs^{CRABPII/Mero} / 2 \times 10^{-6}) M^{-1}cm^{-1}$$

where b is the cuvette path length (1 cm), c is the concentration of CRABPII/merocyanine complex (2 μM), and $Abs^{CRABPII/Mero}$ is the absorbance value of PSB at λ^{max} of CRABPII/merocyanine complex.

Brightness was calculated as the product of $\varepsilon^{CRABPII/Mero}$ and the fluorescence quantum yield (Φ) of each CRABPII/merocyanine pigment and was compared to known fluorescent proteins such as mKate2, mRFP, mCardinal2 and EGFP (**Table I-2** and **Table IV-5**).[3,4]

IV.6 Kinetic Measurements

IV.6.1 Determination Of Relative Rates Of CRABPII Mutants Binding With Merocyanine Aldehyde 1

Kinetic measurements were performed at 23 °C, with the aid of a temperature controller for the UV-vis spectrophotometer. The protein (20 μM in PBS) was mixed with merocyanine aldehyde **1** (0.3 equivalents) and the increase in absorbance of each CRABPII mutant at their respective λ^{max} was recorded at 0.5 second intervals for 80 min to 180 min. The data was fit to an exponential rise (pseudo first order) with KaleidaGraph 4.1.3 using the following equation:

$$A = A_0 (1- \exp(-kt)) + c$$

where A is the absorbance value at each recorded time-point, A_0 is the final absorbance value after the complex formation is complete, k is the rate constant,

t is time after addition, *A* is the recorded absorbance, and *c* is a free constant, which accounts for the time delay from merocyanine **1** addition to the point when recording was started. This equation was rewritten in KaleidaGraph in the following format;

$$y = m1(1-\exp(-m2.m0))+m3$$

where the calculated value of m2 is the rate constant (see **Figure IV-3** for examples). The rate constant of CRABPII KL mutant was taken as a reference (rate = 1) to calculate the relative rates of other CRABPII series. The data is

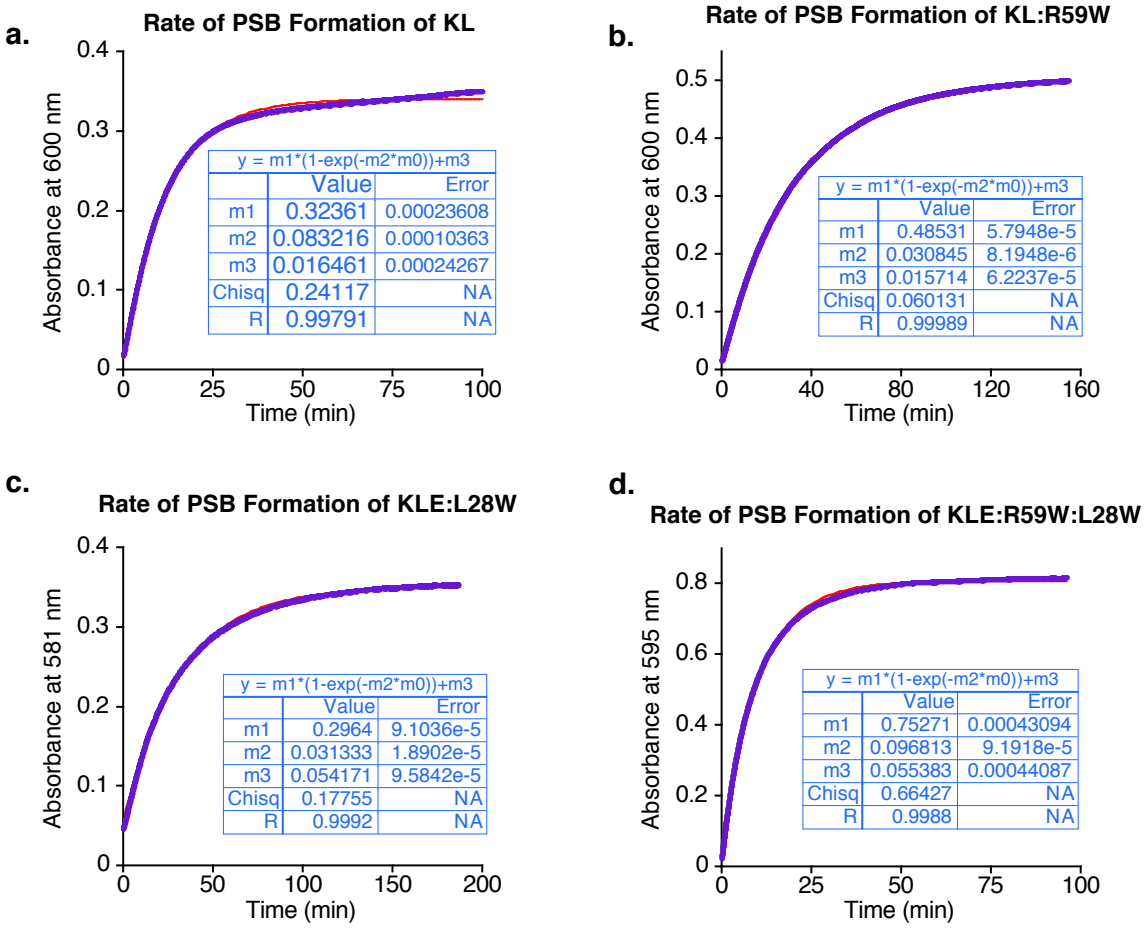


Figure IV-3: Kinetic data of CRABPII mutants binding with merocyanine aldehyde **1**. Increase in PSB absorbances was monitored and pseudo first order fit was applied to deduce the rate constants for each mutants (**a** to **d**) as described in the text.

shown in Chapter I, **Table I-3**.

IV.6.2 Determination Of Second-Order Rate Constant for KLE:R59W Binding with Merocyanine Aldehyde 1

Measurement of the second order rate constant for the complexation of KLE:R59W with merocyanine **1** followed previously established protocols.[5,6] Fluorescence measurements were performed on a FluoroLog 3 instrument, equipped with a temperature controller (Horiba Jobin Yvon). KLE:R59W (100 nM) was reacted with an excess amount of merocyanine **1** in PBS buffer (37 °C), at different concentrations (1 to 5 μM), and the increase of fluorescence intensity was monitored for each reaction. The fluorescence intensity of the samples was collected at 0.5 second intervals, with excitation at 565 nm (1 nm slit width) and emission at 616 nm (12 nm slit width). The raw fluorescence intensity data were converted into complexation fraction by following equation:

$$[\text{complexation fraction}] = (F_t - F_0) / (F_{max} - F_0)$$

where F_t, F_0, and F_{max} are the observed, initial, and maximum fluorescence intensities, respectively. Since the reaction conditions lead to pseudo-first order kinetics [merocyanine**1**] >> [protein], the data at each merocyanine concentration was fit to a first-order kinetic equation as follows:

$$[\text{complexation fraction}] = 1 - \exp(-k^{obs}t)$$

The latter equation was rewritten in KaleidaGraph as shown below:

$$y = m1(1-\exp(-m2m0))+m3$$

where the calculated value of m2 is the pseudo-first-order rate constant (k^{obs}).

The calculated k^{obs} for each reaction was plotted vs. the merocyanine concentration, leading to a linear fit (**Figure I-19**). The resultant slope yields the second-order rate constant (k_2) for the complex formation:

$$k^{obs} = k_2 [\text{merocyanine } \mathbf{1}]$$

IV.6.3 Determination Of Half Life of Complexation of KLE:R59W with Merocyanine Aldehyde 1

Half-life binding measurements ($t_{1/2}$) were determined under stoichiometric conditions (5 μM KLE:R59W and merocyanine) in PBS buffer at 37 °C. The increase in fluorescence intensity, indicative of binding and PSB formation, was recorded at 0.5 second intervals, with excitation at 565 nm (1 nm slit width) and emission at 616 nm (12 nm slit width). The data was fit to a second-order rate equation derived as shown in **Figure IV-4**. In this figure, **C** is directly proportional to fluorescence intensity (cps), plotted vs. time (s). The maximum fluorescence intensity is adjusted to 5 μM, which is the final concentration of the complex, in order to fit the data shown below (**Figure IV-5**). From the plot ($t_{1/2}$) was calculated for a second-order reaction where $t_{1/2} = 1/k_2$.[concentration]. Therefore, $t_{1/2} = 1/5160.8[5\text{x}10^{-6}] = 39$ seconds.

A + **B** --> **C** **A** = CRABPII, **B** = Merocyanine**1**, **C** = Complex

$$\frac{d[\mathbf{C}]}{dt} = k_2[\mathbf{A}][\mathbf{B}]$$

$[\mathbf{A_0}] = [\mathbf{B_0}] = 5x10^{-6}$ M initial concentrations of reactants

as the reaction continues, both **A** and **B** get consumed simultaneously in order to form a complex, **C**

$$[\mathbf{A}]=[\mathbf{B}] =[5x10^{-6} - \mathbf{C}]$$

$$\Downarrow$$

$$\frac{d[\mathbf{C}]}{dt} = k_2[5x10^{-6} - \mathbf{C}]^2$$

$$\Downarrow$$

$$\frac{d[\mathbf{C}]}{[5x10^{-6} - \mathbf{C}]^2} = k_2\,dt$$

$$\Downarrow$$

$$\int \frac{d[\mathbf{C}]}{[5x10^{-6} - \mathbf{C}]^2} = \int k_2\,dt$$

$$\Downarrow$$

$$\frac{1}{[5x10^{-6} - \mathbf{C}]} = k_2 t + m \text{ where m is a constant}$$

$$\Downarrow$$

$$\mathbf{C} = 5x10^{-6} - \frac{1}{[k_2 t + m]}$$

$$\Downarrow$$

The equation is rewriten in KaleidaGraph;

$$\mathbf{y} = m_3 - \frac{1}{[m_2.m_0 + m_1]}$$

where $m_3 = 5x10^{-6}$, $m_2 = k_2$, $m_0 = t$ and $m_1 = 1$ which is a constant

Figure IV-4: Second-order rate equation applied for kinetic studies.

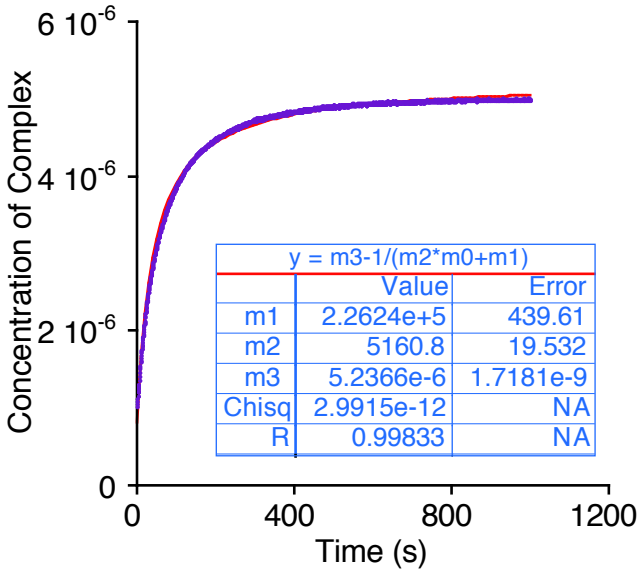


Figure IV-5: Calculation of $t_{1/2}$ from the second-order reaction rate. The plot shows the second-order rate constant, m_2, for PSB formation of CRABPII-KLE:R59W mutant with merocyanine aldehyde **1**.

IV.6.4 Rate Comparison For Iminium Formation Between KLE:R59W Versus BSA

Combination of the high affinity binding of hydrophobic, retinal-like molecules, and the active site Lys residue engineered within the CRABPII mutants, leads to a higher rate of iminium formation with the target proteins as compared to non-specific binding with other proteins. This is demonstrated below (**Figure IV-6** and also **Figure I-22c**), where upon incubation of merocyanine **1** (5 μM) with BSA (5 μM) in comparison to KLE:R59W (5 μM) in PBS (37 °C), the latter achieves 7% of the fluorescence output as compared to the former within the first minute of the reaction (excitation, 565 nm, 1 nm slit

width; emission 616 nm, 12 nm slit width, data was recorded at 0.5 second intervals). During the same time period (1 min), KLE:R59W has reached over 66% of its total fluorescence output, clearly indicating the higher level of specificity for binding merocyanine **1**. The higher specificity for binding KLE:R59W is maintained even at much higher BSA concentration (10-fold excess) as illustrated in **Figure IV-6**.

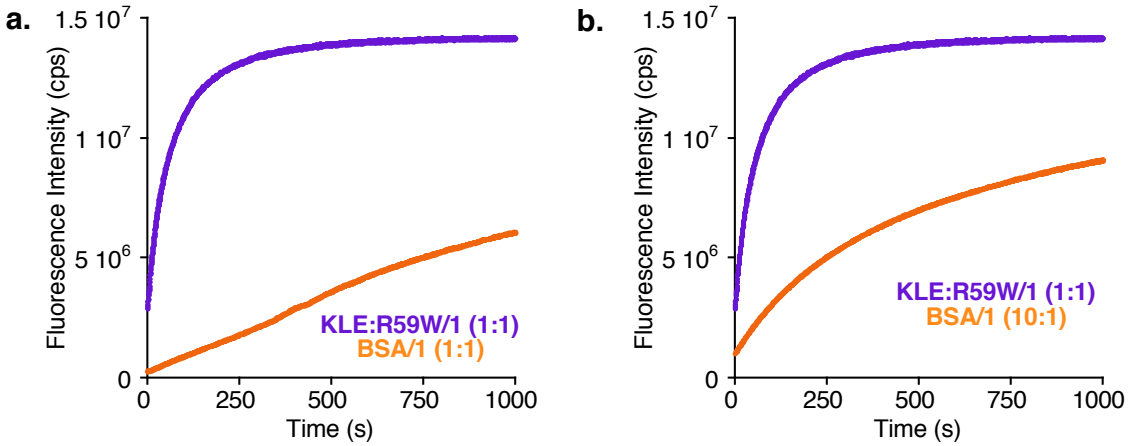


Figure IV-6: Comparison for rate of iminium formation between KLE:R59W and BSA with merocyanine **1**. **a.** Increase of fluorescence intensity was followed upon addition of the merocyanine ligand into the KLE:R59W mutant (purple plot) with an one to one ratio mixture. The orange plot shows the fluorescence intensity increase for the BSA experiment under identical settings. **b.** The rate of iminium formation was compared with the BSA/merocyanine mixture when 10-fold excess of BSA was used.

IV.7 Live Cell imaging In Bacteria

BL21(DE3) pLysS (Invitrogen™) were transformed with 100 ng of plasmid DNA expressing CRABPII variants under the T7 promoter. Cells were plated, grown overnight and single colonies were used for inoculation of 2 mL of Luria Broth for 6 h containing ampicillin and chloroamphenicol (ampicilin 100 μg/mL

and chloroamphenicol 170 μg/mL). Expression was induced with IPTG (final concentration 1 mM) and the culture was shaken at room temperature overnight. For control experiment, 5 mL of competent BL21(DE3) pLysS cells were inoculated in 2 mL Luria Broth containing chloroamphenicol (170 μg/mL) and the culture was shaken at 37 °C for 6 h.

The cultures were heated to 37 °C for 10 min and merocyanine aldehyde, at a final concentration of 1 mM, was added to the pre-warmed cultures. The cells were immediately harvested by centrifugation at 5,000 rpm for 1 min and the supernatant was discarded. The pellets were washed three times with PBS (preheated to 37 °C) followed by 1 min centrifugation. After the last wash the cells were resuspended in PBS and plated on a glass coverslip. In order to stop bacteria from floating, the coverslip was heated to 40 °C for 1 min and covered with no.1 cover glass. Confocal imaging was performed using a Zeiss 510 Meta FCS inverted microscope with a 63x oil immersed objective. The sample was placed with the cover glass facing objective and imaged with a 594 nm laser and a 615 nm long pass filter corresponding to the CRABPII/merocyanine complexes absorbance and emission. Kalman averaging 8 was applied during imaging. All the images have pseudo color. **Figure I-24** depicts images from *E. coli* cells incubated with merocyanine **1**.

IV.7.1 Live Cell Imaging Of Basal Expression Of CRABPII Mutant In Bacteria

In the previous example expression of the CRABPII mutants in *E. coli* was induced with the strong T7 promoter, leading to high expression yields of the target protein. In order to illustrate the feasibility of visualizing low expressed CRABPII protein, we resorted to investigate basal, uninduced, *E. coli* systems that have low yields of the CRABPII mutants as a result of leaky expression.

BL21(DE3) pLysS were transformed with 100 ng of plasmid DNA expressing CRABPII KLE:R59W variants under the T7 promoter. Cells were plated, grown overnight and single colonies were used for inoculation of 10 mL of Luria Broth for 6 h containing ampicillin and chloroamphenicol (ampicilin 100 μg/mL and chloroamphenicol 170 μg/mL). The culture was cooled down to room temperature and shaken for additional 8 h. Same protocol was followed for control experiment in which 5 μL of competent BL21(DE3) pLysS cells were inoculated in 10 mL Luria Broth containing chloroamphenicol (170 μg/mL).

Merocyanine aldehyde, at a final concentration of 1 μM, was added to 1 mL of grown cultures and the cells were immediately harvested by centrifugation at 5,000 rpm for 1 min and the supernatant was discarded. The pellets were washed three times with PBS (preheated to 37 °C) followed by 1 min centrifugation. The cells were resuspended in PBS and a small sample was plated on a glass coverslip. The sample was covered with no.1 cover glass and heated to 40 °C for a minute to fix the bacterial cells. Confocal imaging was performed as described in the previous section. The lack of observable

fluorescence in **Figure IV-7** is again illustrative of insignificant background signal arising from nonspecific binding. On the other hand, bright red fluorescence is observed for basal expression of KLE:R59W.

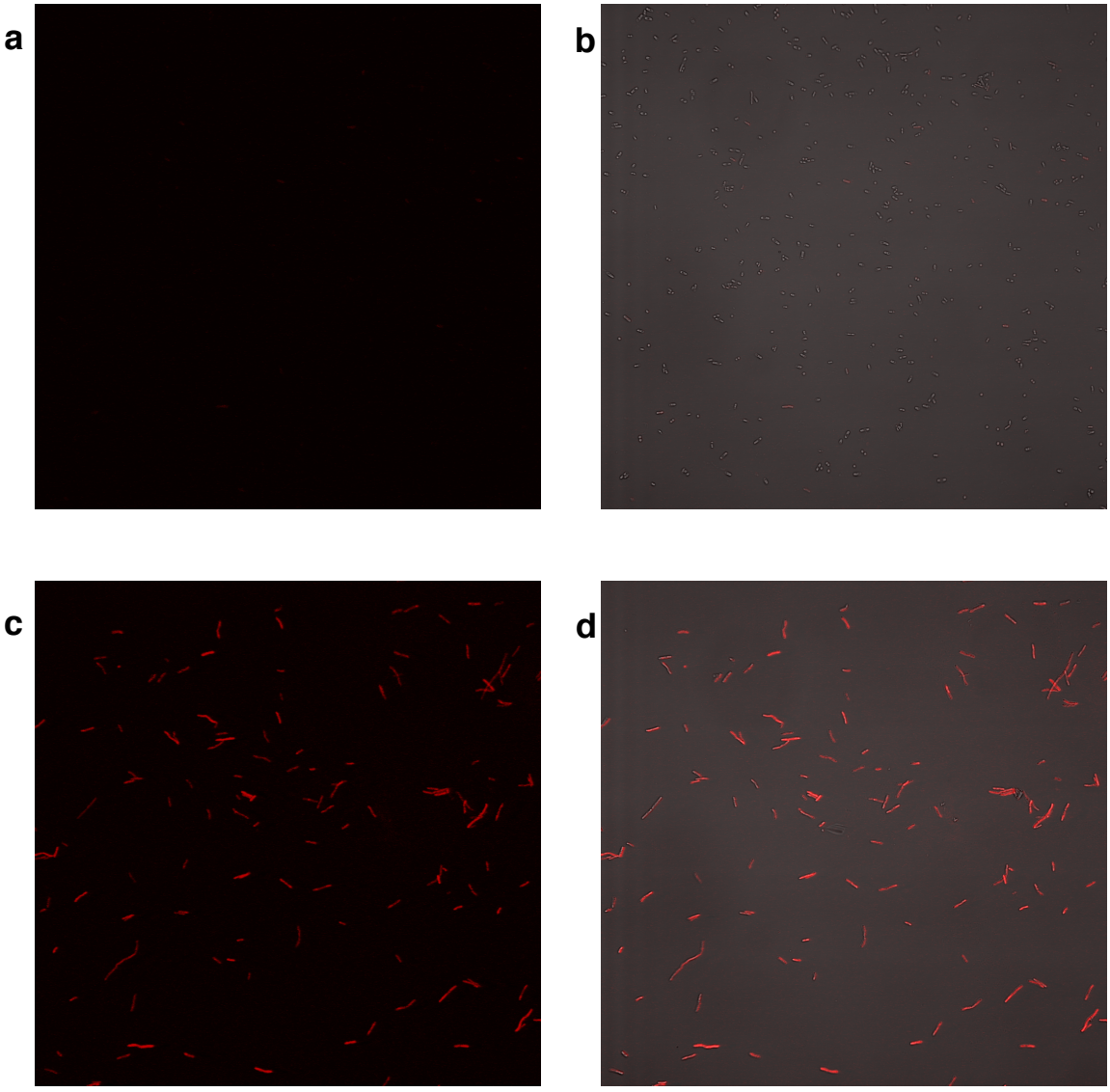


Figure IV-7: Imaging KLE:R59W-merocyanine fluorescence in *E. coli* cells. **a.** Control panel shows no fluorescence upon merocyanine incubation with non-transformed BL21 *E. coli* cells. **b.** Overlay of control fluorescence and bright field image. **c.** Red fluorescence of KLE:R59W-CRABPII mutant complexed with merocyanine **1** at 594 nm excitation. **d.** Overlay of fluorescence and bright field images of KLE:R59W/**1** complex in *E. coli*.

IV.8 Photobleaching Experiment Of CRABPII/Merocyanine Adducts In Bacteria

Laser scanning confocal microscopy (LSCM) photobleaching experiments were conducted with BL21(DE3) pLysS cells after expressing CRABPII mutants and labeling with merocyanine. Cells expressing mRFP with no additional labeling were used for control experiment. The glass coverslips were prepared as described above. Photobleaching was performed using a 63x oil immersion objective (Zeiss 510 Meta FCS). Laser line HeNe594 was adjusted to 100% of output power (2.0 mW). Detector gain and amplifier offset were brought to maximum (1250 and 0.5, respectively). Selected frame size had 1024x1024 pixels and 5x digital zoom. The number of collected frames for each experiment was 200. A single cell having approximately the same dimensions and intensity under the fixed instrument settings was chosen for photobleaching assays. Fluorescence using the 594 nm laser was recorded with a 488/594 main dichroic mirror, 545 nm dichroic beam splitter and 615 nm longpass barrier filter. To produce comparable bleaching curves, we simply scale the raw time coordinates to normalize the intensity to 100 photons/sec of initial emission. The results are illustrated in **Figure I-25**.

IV.9 Expression and Characterization Of Monomeric Red Fluorescent Protein 1

We cloned and expressed monomeric red fluorescent protein 1 (mRFP1) as a control of our quantum yield measurements.

The plasmid mRFP-Rab5 containing the mRFP1 sequence was obtained from Addgene (plasmid #14437).[7] This plasmid was used as a template for PCR amplification with the following primers:

Forward-*NdeI*- (5'-CCG GTC GCC **CAT ATG** GCC TCC TCC-3')

Reverse-*XhoI*- (5'-GAG ATC TGA GTA CTT **CTC GAG** GGC GCC GGT G-3')

The amplified DNA fragment was digested with *Nde*I and *Xho*I (New England Biolabs), and cloned into PET22b expression vector (Novagen), resulting in the plasmid PET22b-mRFP1-His6. The mRFP1 gene was sequenced by the first tier/DMSO sequencing method and the results were identical to those reported by Campbell *et al.*[8] The expression of mRFP1 was carried with the same procedures as described above for CRABPII mutants.

IV.9.1 Protein Isolation And Purification Of MRFP1

The protein isolation and purification were performed as described previously, with modifications.[9] Cells, harvested from culture, were centrifuged at 4 $^{\circ}$C for 12 min at 5,000 rpm. The supernatant was discarded and the bacterial pellet was resuspended in 50 mM KH_2PO_4, pH 8.0, and 0.02% sodium azide (buffer A). The resulting suspension was ultrasonicated (Power 60%, 1 min x 3) and centrifuged at 5000 rpm for 20 min. The following steps were carried under red light to prevent photobleaching of the protein.

The supernatant was applied to a Ni^{+2}-chelating Sepharose Fast Flow column and incubated for 15 min at 4 $^{\circ}$C. The beads were washed with 40 mL of

buffer A. The protein was eluted with 50 mL of elution buffer containing 250 mM of imidazole, 20 mM of Tris, pH=8.0. Buffer exchange was performed by an ultrafiltration cell under the nitrogen pressure (~20 psi) equipped with a 10,000 MW cutoff filter (regenerated cellulose, filter code: YM10, diameter 63.5 mm, NMWL: 10,000). The protein was first concentrated to ~25 mL and then diluted to 160 mL with Tris buffer (10 mM Tris•HCl, pH=8.0). The resulting dilute solution was concentrated back to ~25 mL. Protein was stored at 4 °C in a sterilized tube covered with aluminum foil.

IV.9.2 Quantum Yield Determination Of MRFP1

Quantum yield of mRFP1 was measured under identical conditions with CRABPII variants. As described above, the excitation wavelength was set to 565 nm and the emission spectrum was collected between 575 nm to 800 nm. The

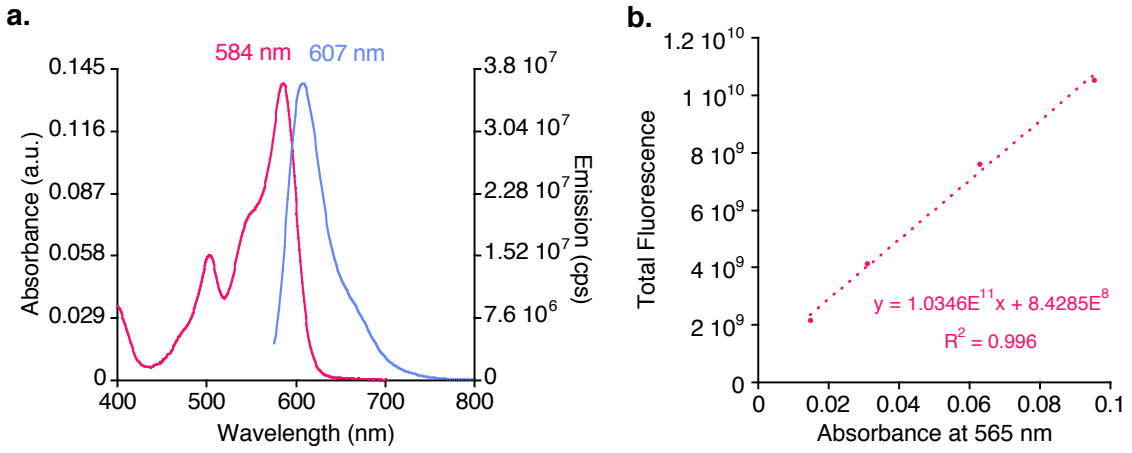


Figure IV-8: Spectroscopic characterization of mRFP1. **a.** Absorption (pink) and emission (blue) spectra of mRFP1 in PBS. **b.** The plot shows the total fluorescence emission (integrated area from 575 nm to 800 nm, excitation at 565 nm) as a function of absorption at 565 nm. Slope of the line is correlated to the quantum efficiency.

quantum yield was calculated as 27.2% based on comparison with standards, which is in agreement with published data (**Figure IV-8**).[8]

IV.10 Cloning Of CRABPII Gene Into Mammalian Expression Vectors

In order to carry out live cell imaging in eukaryotic cells, we cloned CRABPII gene in different mammalian expression vectors, under the cytomegalovirus (CMV) promoter. The construction of EGFP-CRABPII (KLE:R59W:A32W) fusion was performed in the pEGFP-C2 vector which was a generous gift from Prof. William Henry (Associate Chairperson, MSU, Biochemistry and Molecular Biology). The CRABPII gene was amplified and cloned after the EGFP gene by using the forward and reverse primers as indicated below:

Forward-*XhoI*- (5'-CC**C TCG AG**G CCA AAC TTC TCT GGC AAC TGG AAA ATC-3')

Reverse-*HindIII*- (5'-GCA GAA TTC G**AA GCT T**CA CTC TCG GAC GTA GAC TTT GGT GCA CAC-3')

The second construct that tested in eukaryotic cell imaging was achieved by cloning the CRABPII gene into the pECFP-N1 (so called ECFP-GalT) vector (Addgene plasmid #11937).[10] *XhoI* and *BamHI* restriction sites were chosen for cutting out the GalT segment and inserting the CRABPII gene (KLE:R59W:A32W mutant) upstream of the ECFP. The following primers were used to amplify and clone the CRABPII gene:

Forward-*XhoI*- (5'- GGA GAT ATT **CTC GAG** CTC GCC ACC ATG CCA AAC TTC TCT GGC AAC TGG AAA ATC-3')

Reverse-*BamHI*- (5'- GTC ATT T**GG ATC C**CG TCC CTC TCG GAC GTA GAC CTT GGT-3')

In this construct, CRABPII-ECFP fusion was expressed in mammalian cells and the bright fluorescence of ECFP was visualized via confocal microscopy (**Figure I-26b**). In order to generate non-tagged CRBAPII protein, we introduced a stop codon (**TGA**) at the end of CRABPII, preventing the translation of the ECFP protein. This was done by a PCR method as described above (**Table IV-1**), but setting the elongation (extension) time for 5.5 min at 72 $^{\circ}$C. The primers used for the PCR method are shown below:

Forward primer (5'-C GTC CGA GAG **TGA** CGG GAT CCA CCG-3')

Reverse primer (5'-CGG TGG ATC CCG **TCA** CTC TCG GAC G-3')

Lastly, we utilized the pFlag-CVM2 vector (obtained from Addgene but currently not available) to carry out further imaging experiments in mammalian cells. The CRABPII-KLE:R59W mutant gene amplified with the corresponding primers (illustrated below) and the resultant cDNA was cloned downstream of the Flag epitope sequence into the *HindIII* and *EcoRI* sites of the pFlag CVM2 vector.

Forward-*HindIII*- (5'-GAT ATA **AAG CTT** CCA AAC TTC TCT GGC-3')

Reverse-*EcoRI*- (5'-CT GCA **GAA TTC** TCA CTC TCG GAC G-3')

All the plasmids with the designed constructs were sequenced by the MSU gene sequencing facility. The corresponding sequencing primers (45 pmol) were

added to the DNA samples before submission; the sequence of the primers is shown below:

Sequencing of pFlag-CRABPII gene (5'-GGG AGG TCT ATA TAA GCA GAG CTC G-3')

Sequencing of genes under CVM promoter (5'-GGT CTA TAT AAG CAG AGC TGG TTT AG -3')

All the imaging studies were carried by Dr. Tetyana Berbasova and Ms. Elizabeth Santos. Detailed protocols were provided in Dr. Tetyana Berbasova's dissertation.

IV.10.1 General Cloning Protocol

All PCR amplifications of the target genes were performed according to specified PCR conditions (**Table IV-1**), save that the extension times were changed based on the size of the targets gene (1 min per 1000 bp) instead of the 4 min extension time. To ensure high yield of the amplified gene (insert), we set up four or five PCR tubes. The resultant PCR products were combined and subjected to agarose gel electrophoresis (1% agarose, 2 μg/mL of ethidium bromide, 150V, 15 min). The gene fragment was visualized under UV hand lamp and cut out for further purification. Wizard® SV Gel and PCR Clean-Up System (Promega) was used to purify and isolate the amplified fragment by following the kit instructions. It is important to mention that better quality and yield of DNA was obtained when Zymo-Spin™ IC spin columns used.

The purified insert and the cloning vector were digested with proper restriction enzymes. The restriction enzymes were supplied from (New England BioLabs®) along with the CutSmart buffer. It is important to note that longer incubation times in 1.5X CutSmart buffer can increase the efficiency of the digestion process. The resultant digested samples were purified by agarose gel electrophoresis by following the same protocol described above. The final purified cut insert and vector fragments were ligated via T4 DNA Ligase (New England BioLabs®). It is suggested to set up different insert to vector ratios in a order to optimize the ligation reaction. The best results were usually obtained in 3:1 ratio (insert:vector) of ligation solution in a 1.5X ligation buffer at room temperature for 2 h. Ligated products were transformed to XL-1 Blue super (Novagen®) competent cells for DNA amplification and purification (see **Section IV-1** for further DNA isolation protocols).

IV.11 Determination Of Dissociation Constant For HCRBPII Mutants With Azobenzene Allyl Alcohol

The affinity of hCRBPII mutants for azobenzene allyl alcohol (**Azo-AA**) was measured via fluorescence quenching assay which was previously established by Wang and Cogan *et al.*[11,12] Briefly, 20 μM of hCRBPII sample was prepared in 3 mL of PBS solution and transferred into a clean fluorescence cuvette. The concentration of the stock **Azo-AA** (in absolute ethanol) was determined by UV-vis, using the molar extinction coefficient of 29,530 $M^{-1}cm^{-1}$.

The concentration of the stock solution was prepared to be around 1 mM in order to minimize the addition of a high volume of ethanol solution during the fluorescence titration. The protein sample was excited at 280 nm (slit width of 1 nm) and the fluorescence emission was recorded from 345 nm to 355 nm (slit width of 12 nm). This was repeated three times until a stable emission intensity at 350 nm was reached. **Azo-AA** was added to the cuvette at varying equivalents from the stock solution and fluorescence intensity at 350 nm was recorded for each chromophore addition at 23 °C. The titration was complete when there was no observable quenching of fluorescence upon addition of the chromophore. The data were analyzed by nonlinear square fit of the equation shown in **Figure IV-9**.

IV.12 Crystallization of HCRBPII And CRABPII Mutants

Except the azobenzene allyl alcohol, the light sensitive polyenes were dissolved in ethanol, stored at -80 °C, and protected from light in a covered vial. The stock solution of azobenzene allyl alcohol was freshly prepared right after crystallization from the solid compound and kept at 4 °C. All the protein/ligand complex formation was achieved via addition of the ligand solution (up to 4 equivalents) to the concentrated protein solution (~ 20 mg/mL for CRABPII, ~ 12 mg/mL for hCRBPII). It is important to keep the ethanolic ligand solution less than 10% in the concentrated protein solution in order to prevent protein precipitation. For the crystallization of CRABPII variants, the protein/merocyanine **1** sample was incubated at room temprature in the dark for 1 h to reach complete PSB

formation. However, hCRBPII mutants were incubated with their cognitive ligands between 5 h to 12 h before setting up crystal boxes. All crystals were grown by

a.

$$\frac{F}{F_f} = 1 + \left(\frac{F_b}{F_f} - 1\right).\left(\frac{P_t + A_t + K_d - \sqrt{(P_t + A_t + K_d)^2 - 4P_tA_t}}{2P_t}\right)$$

where F is observed fluorescence,

F_b is the fluorescence of the bound hCRBPII,

F_f is the fluorescence of the free hCRBPII,

P_t is the total hCRBPII concentration,

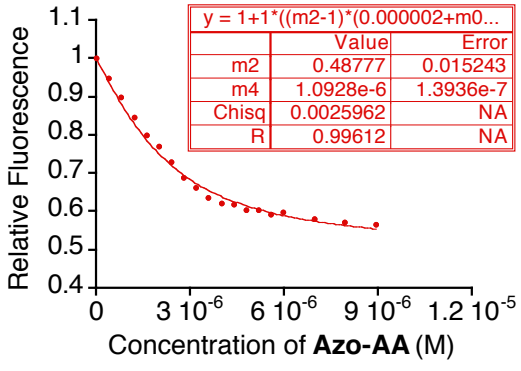
A_t is the total azobenzene allyl alcohol concentration (this is variable),

K_d is the dissociation constant.

The equation is rewriten in KaleidaGraph;

$$y = 1 + \left(m2 - 1\right).\left(\frac{2x10^{-5} + m0 + m4 - \sqrt{(2x10^{-5} + m0 + m4)^2 - 8x10^{-5}.m0}}{4x10^{-5}}\right)$$

b.



c.

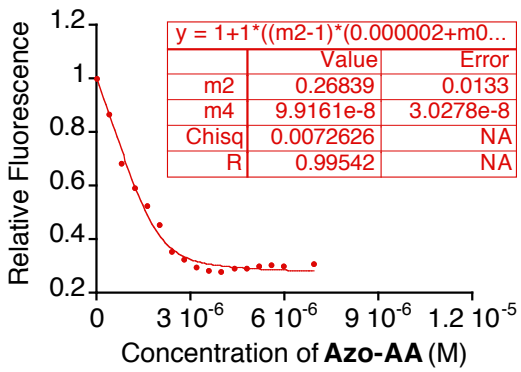


Figure IV-9: Analysis of affinity measurements of the hCRBPII mutants for **Azo-AA**. **a.** The nonlinear square fit equation to calculate K_d values. **b.** The curve fit of relative fluorescence versus concentration of **Azo-AA** for the T53A hCRBPII mutant (m4 represents the K_d value). **c.** Another example of affinity measurement between **Azo-AA** and R58W:Q38V hCRBPII mutant.

hanging drop vapor diffussion (at 4 °C for CRABPII/merocyanine **1** complexes, 23 °C for hCRBPII/ligand complexes) with the drop containing an equal mixture of crystallization and complex solution. Crystallizations were set up under red light, and the boxes were kept in the dark for the polyene chromophores. Crystals appear within 3-4 days and grow to their maximum size in two weeks (**Figure IV-10**). In case of hCRBPII/**Azo-AA** complex, crystals appeared within 12 h and were collected immediately. Crystals were briefly soaked in a cryoprotectant solution containing the mother liquor and 20% glycerol and flash frozen in liquid nitrogen under red light.

Diffraction data were collected at the Advanced Photon Source (APS)

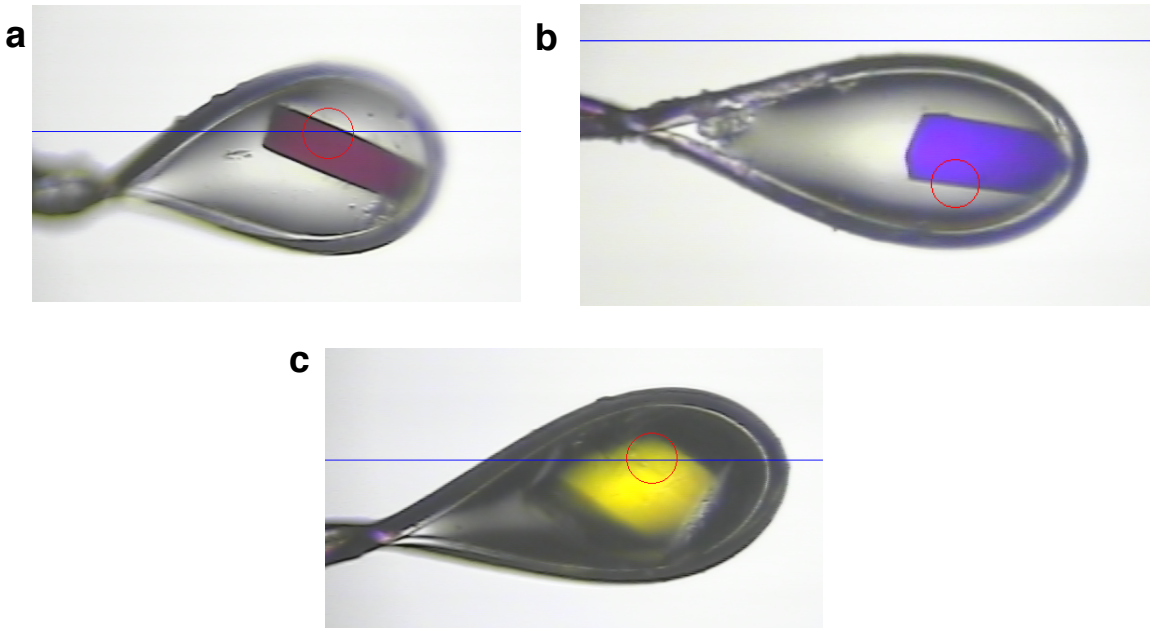


Figure IV-10: Photos of different ptotein/ligand crystals under the beamline. **a.** Single crystal of a hCRBPII mutant (KL:T51V:T53S) complexed with GEMLESS chromophore. **b.** CRABPII-KLE/merocyanine **1** crystal. **c.** Crystal photo of hCRBPII/**Azo-AA** crystal.

(Argonne IL) at the LSCAT (sector 21) using 1.00 Å wavelength radiation at 100 K. Data reduction and scaling were performed using the HKL2000 program package.[13] The structures were determined using the Molecular Replacement program in the CCP4 program suite, and refined using the CCP4 and PHENIX program packages.[14-16] Several cycles of refinement, model adjustment and placement of ordered water molecules were performed for each structure. All model building and placement of water molecules were preformed using COOT (0.6.1).[17] The chromophore was manually fitted in the electron density near the end of the refinement.

Table IV-6. CRABPII/merocyanine **1** crystallization box conditions.

500μL PEG 200μL Salt-**1** 200μL H_2O 100μL BTP-**1**	500μL PEG 200μL Salt-**2** 200μL H_2O 100μL BTP-**2**	500μL PEG 83μL Tac-5 100μL BTP-**1** 317μL H_2O	500μL PEG 83μL Tac-6 100μL BTP-**2** 317μL H_2O	600μL PEG 100μL BTP-**2** 100μL Salt-**3** 200μL H_2O
600μL PEG 100μL BTP-**1** 100μL Salt-**3** 200μL H_2O	500μL PEG 100μL BTP-**2** 100μL Salt-**3** 300μL H_2O	500μL PEG 100μL BTP-**1** 100μL Salt-**3** 300μL H_2O	500μL PEG 100μL BTP-**1** 100μL Salt-**3** 300μL H_2O	500μL PEG 100μL BTP-**3** 100μL Salt-**3** 300μL H_2O

PEG = 40% Poly(ethylene glycol) 3350

Salt-**1** = 1M Sodium Malonate pH=5.5
Salt-**2** = 1M Sodium Malonate pH=6.5
Salt-**3** = 2M Sodium Fluoride

Tac-5 = 24% (v/v) Tacsimate pH=5.0
Tac-6 = 24% (v/v) Tacsimate pH=6.0

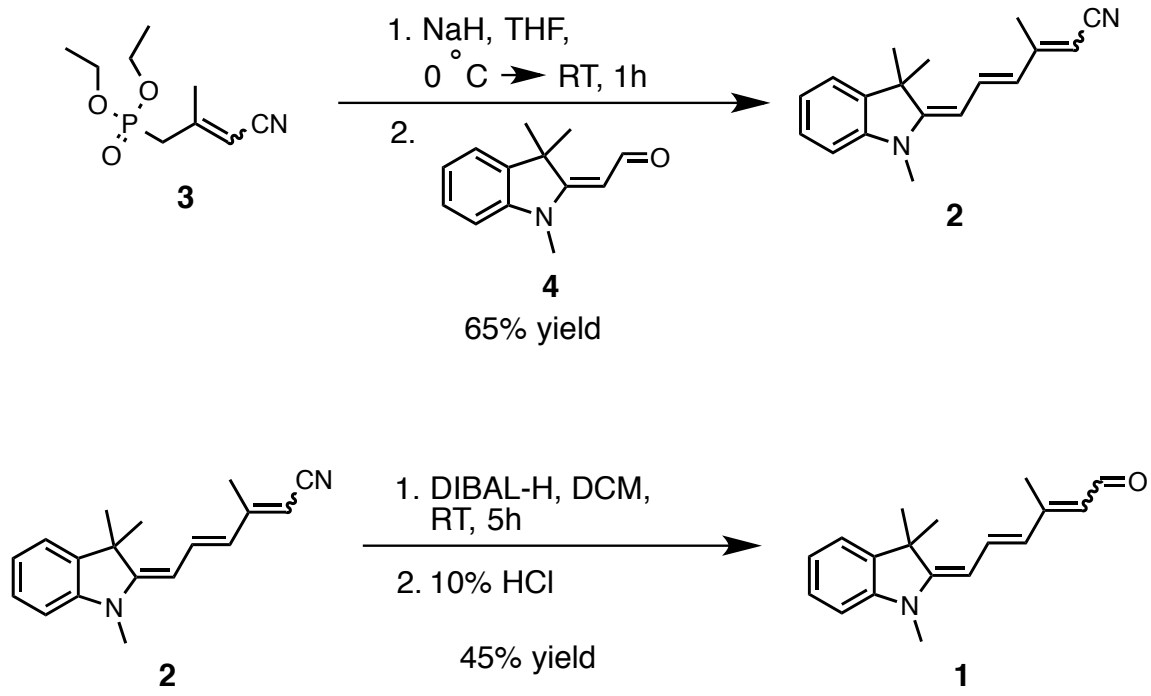
BTP-**1** = 1M Bis-Tris Propane pH= 5.5
BTP-**2** = 1M Bis-Tris Propane pH= 6.5
BTP-**3** = 1M Bis-Tris Propane pH= 7.0

Although CRABPII and hCRBPII share high structural similarity, crystallization of these proteins requires different conditions. The hCRBPII crystals were obtained in conditions that were previously established by Prof. James Geiger. The best hCRBPII crystals were grown by mixing 1 μL of reservoir solution (30-35% PEG 4000, 0.1 M sodium acetate [pH 4.6-4.8], and 0.1 M ammonium acetate) with 1 μL of hCRBPII/ligand complex. On the other hand, CRABPII/merocyanine **1** complexes were crystallized in conditions as outlined in **Table IV-6**.

IV.13 Experimental Procedures For Synthesis Of Chromophores

IV.13.1 Synthesis Of Merocyanine Aldehyde 1

The synthesis of ethyl phosphonate, compound **3** and Fisher's base



Scheme IV-1: Synthesis of merocyanine aldehyde **1**.

aldehyde, compound **4** were described previously.[18-20] Both compound **3** and **4** can also be purchased from Sigma Aldrich (cas number: 87549-50-6 and 84-83-3, respectively). All the syntheses shown in **Scheme IV-1** were performed under red light and the products were stored at -20 °C, under an atmosphere of nitrogen gas.

*Synthesis of compound **2***: NaH (60% by weight, 0.65 g suspension containing 0.39 g NaH, 1.62 mmol) was placed in a dry 100 mL round bottom flask. Hexane (~15 mL) was added and the mixture was stirred briefly, the solid particles were allowed to settle. The solvent-oil solution was removed with a glass pipette. Washing was performed twice. The round bottom flask was then capped and purged with nitrogen. The washed NaH was resuspended in anhydrous THF (35 mL). The stabilized nitrile ylide **3** (0.3 g, 1.39 mmol) was added to the NaH/THF suspension at 0 °C with vigorous stirring during 5 min. The reaction mixture was left stirring in the ice-bath for 10 min and then slowly warmed to room temperature (RT) for an hour. Fisher's base aldehyde, **4**, (0.23 g, 1.16 mmol) was added dropwise during 5 min. Once the addition of the aldehyde was completed, the reaction was kept under nitrogen atmosphere and was left stirring for 8 h.

Work up: Brine solution was added to reaction mixture and the resulting aqueous layer was extracted three times with diethyl ether. Organic layers were combined and washed with deionized water. The combined organic phase was

dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure to yield an oily mixture. Purification was performed with a 0.1% triethylamine basified silica gel column with 20% ethyl acetate/hexane mixture to yield the conjugated nitrile **2** as a mixture of two isomers (0.2 g overall yield, 65%).

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.56 (*s*, 6 H), 2.19 (*d*, *J* = 0.5 Hz, 3 H), 3.12 (*s*, 3 H), 4.97 (*s*, 1 H), 5.32 (*d*, *J* = 13.0 Hz, 1 H), 6.04 (*d*, *J* = 14.0 Hz, 1 H), 6.63 (*d*, *J* = 7.9 Hz, 1 H), 6.86 (*td*, ^1J = 7.5 Hz, ^2J = 0.9 Hz, 1 H), 7.12 (*dd*, ^1J = 7.3 Hz, ^2J = 1.0 Hz, 1 H), 7.15 (*dd*, ^1J = 14.5 Hz, ^2J = 12.0 Hz, 1 H), δ 7.17 (*td*, ^1J = 7.7 Hz, ^2J = 1.2 Hz, 1 H).

^{13}C NMR (125 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 16.7, 28.5, 29.1, 45.9, 91.3, 95.5, 106.3, 119.6, 120.2, 121.6, 123.2, 127.9, 133.1, 138.9, 144.7, 157.5, 160.8.

*Synthesis of merocyanine aldehyde **1***: Compound **2** (0.15 g, 0.57 mmol) was dissolved in dry dichloromethane (25 mL) in a 200 mL round bottom flask. A 1 M solution of DIBAL-H (2.27 mL, 2.27 mmol) in hexane was added over 15 min and kept stirring for 5 h. The reaction was quenched with 10% HCl and the pH of the organic layer was monitored not to go below pH 4.0. The resultant aqueous and organic layers were vigorously stirred for 5 min at room temperature.

Work up: The organic and aqueous layers were separated and the aqueous layer was extracted with dichloromethane, twice. The combined organic layers were dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure to yield a dark oily mixture. Further purification was performed by 0.1% triethylamine basified silica gel with a stepwise elution column chromatography. The solvent polarity was slowly increased from 5% ethyl acetate/hexane mixture to 20% ethyl acetate/hexane mixture during chromatography. The fractions containing merocyanine aldehyde were collected and the solvent was removed under reduced pressure to yield merocyanine aldehyde, **1**, as a red solid (68.3 mg, 45% yield). ^{1}H NMR analysis indicates a 3:1 ratio of *E:Z* olefinic isomers (double bond directly conjugated to the aldehyde).

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.56 (*s*, 6 H), 2.3 (*s*, 3 H), 3.15 (*s*, 3 H), 5.43 (*d*, *J* = 12.0 Hz, 1 H), 5.89 (*d*, *J* = 8.5 Hz, 1 H), 6.11 (*d*, *J* = 15.0 Hz, 1 H), 6.66 (*d*, *J* = 8.0 Hz, 1 H), 6.88 (*t*, *J* = 7.5 Hz, 1 H), 7.14 (*d*, *J* = 7.5 Hz, 1 H), 7.18, (*t*, *J* = 7.5 Hz, 1 H), 7.38 (*dd*, ^1J = 14.0 Hz, ^2J = 12.0 Hz, 1 H), 10.02 (*d*, *J* = 8.0 Hz, 1 H).

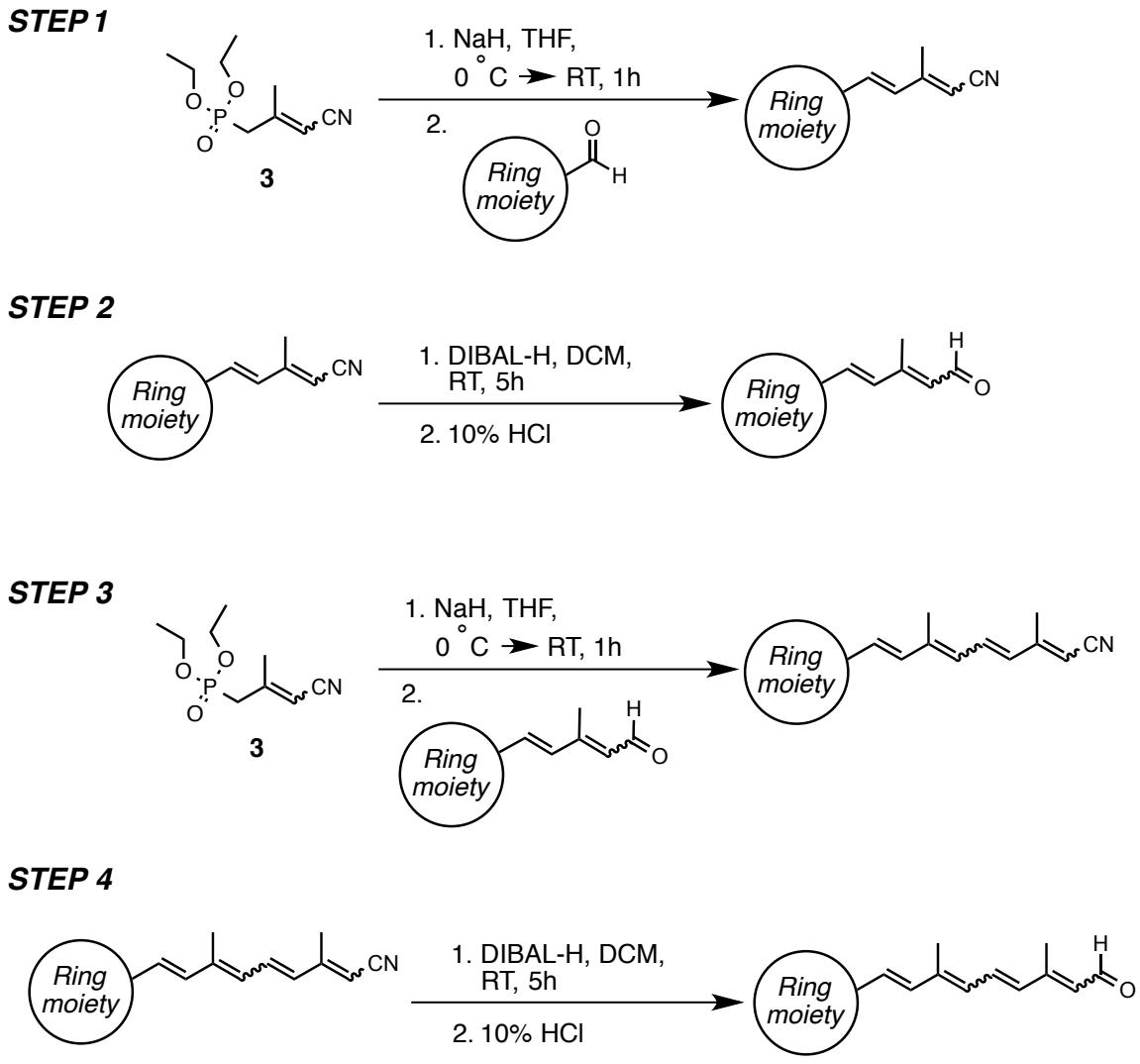
^{13}C NMR (125 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 13.2, 28.5, 29.2, 46.0, 96.4, 106.5, 120.4, 121.6, 125.5, 126.2, 127.9, 133.7, 138.8, 144.6, 156.4, 161.2, 190.6.

HRMS analysis (ESI): Calculated for $C_{18}H_{22}NO$: 268.1701; Found: 268.1694

$[M+H]^+$. UV-vis spectrum of merocyanine aldehyde **1** is illustrated in Chapter I, **Figure I-11** (extinction coefficient: 77,000 $M^{-1}cm^{-1}$ at 492 nm).

IV.13.2 Synthesis Of Artificial Retinal Compounds

All the retinal analog syntheses were achieved by following four consecutive steps shown in **Scheme IV-2.** All the steps were performed under red light and the products were stored at -80 °C, under an atmosphere of argon



Scheme IV-2: Four-step synthesis of artificial retinal analogs.

gas. Ring moieties with the appended aldehyde were pre-synthesized for the 5DEM and GEMLESS chromophores prior to **Scheme IV-2** protocol. However, the head groups for the CYCHXE and CYCHXA ligands were purchased from Sigma Aldrich. The cas numbers for 1-cyclohexene-1-carboxaldehyde and cyclohexanecarboxaldehyde are 1192-88-7 and 2043-61-0, respectively.

General procedure for Step 1 and Step 3: NaH (60% by weight, 1.4 equivalents) was placed in a dry 100 mL round bottom flask. Hexane (~10 to 15 mL) was added and the mixture was stirred briefly, the solid particles were allowed to settle. The solvent-oil solution was removed with a glass pipette. Washing was performed twice. The round bottom flask was then capped and purged with nitrogen. The washed NaH was resuspended in anhydrous THF (20 to 40 mL). The stabilized nitrile ylide **3** (1.2 equivalents) was added to the NaH/THF suspension at 0 $^{\circ}$C with vigorous stirring during 5 min. The reaction mixture was left stirring in an ice-bath for 10 min and then slowly warmed to room temperature (RT) for an hour. The aldehyde moiety (1 equivalent) was added dropwise during 5 min. Once the addition of the aldehyde was completed, the reaction was kept under nitrogen atmosphere and was left stirring for 6 to 8 h.

Work up: Brine solution was added to reaction mixture and the resulting aqueous layer was extracted three times with diethyl ether. Organic layers were combined and washed with deionized water. The combined organic phase was dried over anhydrous Na_2SO_4. The solvent was removed under reduced

pressure to yield an oily mixture. Purification was performed with a 0.1% triethylamine basified silica gel column with 15% ethyl acetate/hexane mixture. The fractions containing the final conjugated nitrile were collected and the solvent was removed under reduced pressure. The yields obtained from these reactions range from 50% to 90% in which step 3 usually results in lower yield in comparison with step 1.

General procedure for Step 2 and Step 4: Conjugated nitrile (1 equivalent) was dissolved in dry dichloromethane (20 to 25 mL) in a 200 mL round bottom flask. A 1 M solution of DIBAL-H (3.5 equivalents) in hexane was added over 15 min to the reaction solution and then, it was kept stirring for additional 5 to 6 h at room temperature. The reaction was quenched with 10% HCl solution. Usually quenching of excess DIBAL-H results in particle formation in the reaction mixture but it is important to dissolve these solid particles in order to achieve higher yields and have an easier purification. This can be achieved by the addition of a little excess of HCl solution, save that the pH of the organic layer was monitored, with a pH paper, not to go below pH 4.0. The resultant aqueous and organic layers were vigorously stirred for 5 min at room temperature.

Work up: The organic and aqueous layers were separated and the aqueous layer was extracted with dichloromethane, twice. The combined organic layers were dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure and further purification was performed with a 0.1% triethylamine basified silica gel column chromatography. The solvent polarity was

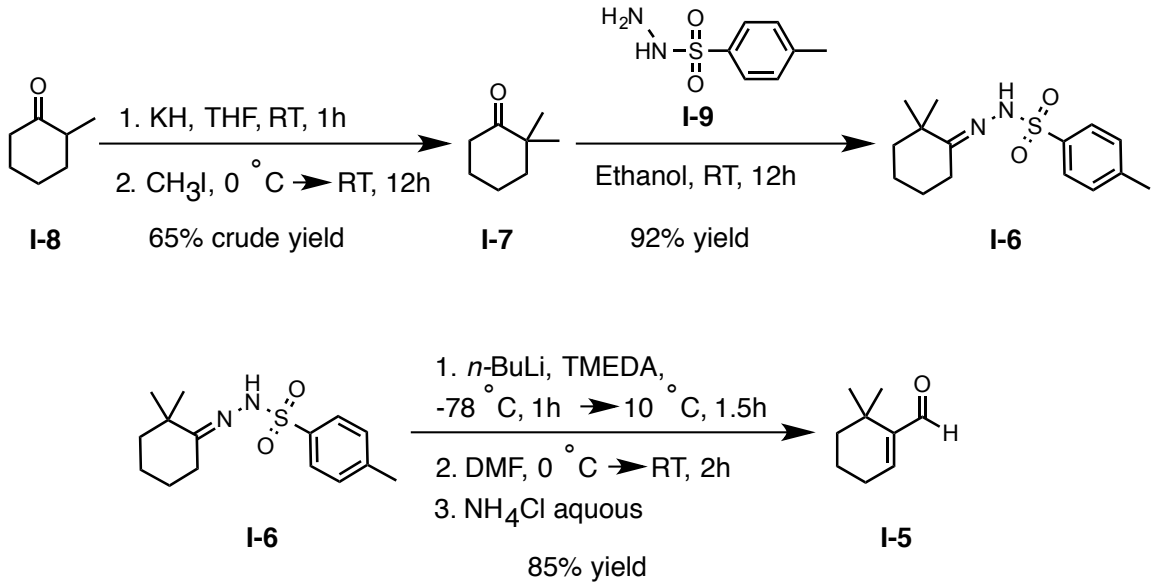
slowly increased from 5% ethyl acetate/hexane mixture to 20% ethyl acetate/hexane mixture during chromatography. The fractions containing the conjugated aldeyhde were collected and the solvent was removed under reduced pressure. The yields obtained from these reactions range from 60% to 85%.

All proton and carbon NMR data and other spectroscopic details of the synthesized compound will be described in Section IV.13.5.

IV.13.3 Synthesis Of 5DEM Chromophore Ring Moiety

The synthesis of the 5DEM ring moiety, compound **I-5**, was achieved by following the protocol which was reported by Kurth and Beard (**Scheme IV-3**).[21] Synthesis of compound **I-5** was also reported by other groups with different routes.[22,23]

*Synthesis of 2,2-dimethylcyclohexanone, **I-7***: Compound **I-7** is a commercially available molecule but conversion of 2-methylcyclohexanone (**I-8**) to 2,2-dimethylcyclohexanone (**I-7**) was carried out according to published



Scheme IV-3: Synthesis of 5DEM chromophore ring moiety (compound **I-5**).

procedures.[24]

Potassium hyride was obtained as a dispersion in mineral oil containing 30% KH by weight. KH/mineral oil suspension (7.93 g, 0.059 mol) was placed in a dry 200 mL round bottom flask. Mixture was allowed to settle and the oil was removed with the glass pipette. Rinsing with a hexane solution was used for further cleaning. Hexane (~ 20 mL) was added and the mixture was stirred briefly, then allowed to settle. The solvent-oil solution was removed with a glass pipette. Washing was performed three times. A round bottom flask was then capped and purged with Nitrogen. KH solid was resuspended in distilled THF (40 mL). The 2-methylcyclohexanone (6 mL, 0.049 mol) was added to the KH/THF suspension at room temperature with vigorous stirring during 20 min. The reaction was left stirring at room temperature for 1 h (monitored by TLC). The reaction was cooled down 0 $^{\circ}$C in an ice bath and methyl iodide (3.39 mL, 0.054 mol) was added dropwise during 15 min. Once the addition of methyl iodide was complete, the reaction was brought to room temperature and left stirring overnight.

Work up: The reaction mixture was filtered through a celite pad eluting with diethyl ether. The organic layer was washed with deionized water and dried over anhydrous sodium sulfate. Removal of solvent by rotovap gave an oily crude compound **I-7**, which includes 87% of 2,2-dimethylcyclohexanone as a thermodynamic product (65% crude yield).

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 1.09 (*s*, 6 H), 1.62 - 1.82 (*m*, 6 H), 2.37 (*t*, *J* = 6.5 Hz, 2 H).

Synthesis of hydrazone, ***I-6***: *p*-Tolylsulfonyl hydrazine **I-9** (1.17 g, 6.28 mmol) and 2,2-dimethylcyclohexanone (**I-7**, 1.5 mL from crude) were added to a small round bottom flask and stirred together under nitrogen atmosphere. Ethanol (5 mL) was added in one portion. The mixture was gently heated until the solids were completely dissolved. The solution was allowed to cool to room temperature, and left stirring for overnight.

Work up: Ethanol was evaporated and the white precipitate was suspended in cold hexane (50 mL). Solid particles crashed out and were allowed to settle in the fridge for 1 hour. The solvent was pipetted out and the white hydrazone crystals (1.7 g, 5.77 mmol) were dried under vacuum (92% yield according to hydrazine).

The above synthetic procedure was taken from Kurth *et al*.[21]

^{1}H NMR (500 MHz, $CDCl_3$) δ [ppm] = 1.02 (*s*, 6 H), 1.44 (*t*, *J* = 5.5 Hz, 2 H), 1.55 (*m*, 4 H), 2.18 (t, *J* = 6.0 Hz, 2 H), 2.41 (*s*, 3 H), ~ 7.0 (*broad s*, N-**H**), 7.27 (*d*, *J* = 8.0 Hz, 2 H), 7.82 (*d*, *J* = 8.0 Hz, 2H).

^{13}C NMR (125 MHz, $CDCl_3$) δ [ppm] = 21.3, 21.6, 23.0, 26.1, 26.8, 39.2, 40.8, 128.2, 129.2, 129.2, 129.9, 135.2, 143.7, 166.7.

Melting point: 125 - 130 $^{\circ}$C

Synthesis of 5DEM ring moiety, ***I-5***: An ovendried, two-necked 100-mL round-bottom flask was charged with dry TMEDA (11 mL, 0.0734 mol), *n*-butyllithium in hexanes (16.3 mL, 1.126 M, 0.0184 mol) and was cooled to -78 °C. The solid hydrazine **I-6** (1.5 g, 5.1 mmol) was added over a period of 20 min, and the resulting mixture was stirred at -78 °C for an additional 1 h. The red solution was warmed to 10 °C and kept stirred during 1.5 h at this temperature. The solution was cooled to 0 °C and dry dimethylformamide (4 mL, 0.051 mol) was added over 10 min. The solution was stirred for 2 h at room temperature.

Work up: The reaction mixture was poured into water and then ether was added. The organic layer was washed with 5% aqueous copper sulfate solution and brine (2 times) and dried (Na_2SO_4). The solvent was removed under reduced pressure to yield a yellow oil. The product was purified using silica-gel column with 10% ethyl acetate/hexane, yielding 0.6 g of product (4.35 mmol, 85% yield). It is important to take the resultant aldehyde to the next step as soon as possible due to its instability even at low temperatures (fast oxidation to a carboxylic acid).

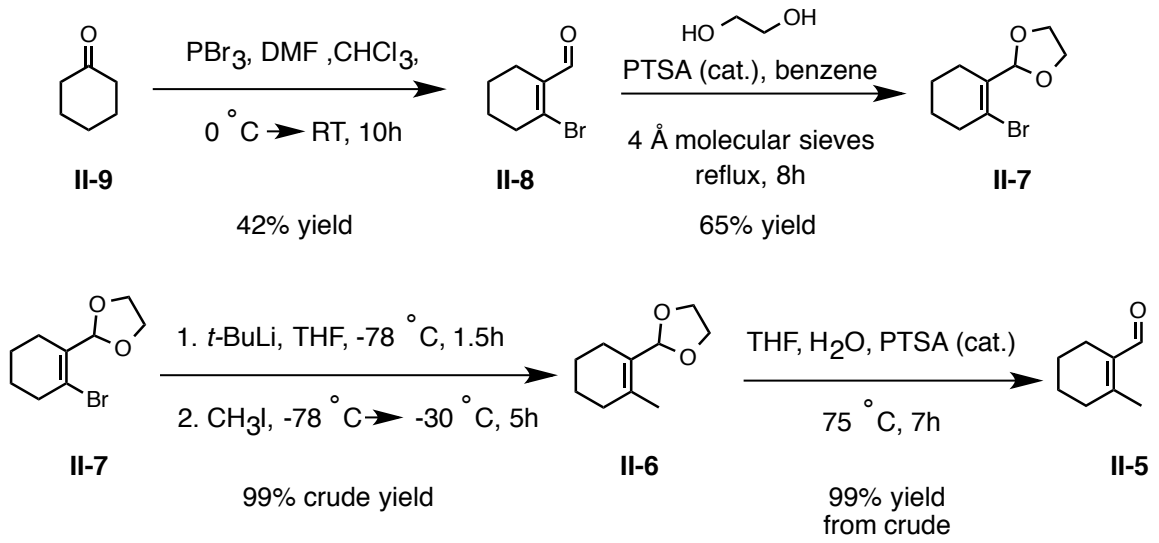
^{1}H NMR (500 MHz, $CDCl_3$) δ [ppm] = 1.18 (*s*, 6 H), 1.43 - 1.47 (*m*, 2 H), 1.58 - 1.64 (*m*, 2 H), 2.25-2.31 (*m*, 2 H), 6.67 (*t*, *J* = 3.8 Hz, 1 H), 9.31 (*s*, 1H).

^{13}C NMR (125 MHz, $CDCl_3$) δ [ppm] = 18.3, 27.1, 27.3, 32.4, 39.6, 148.6, 153.0,

194.8.

IV.13.4 Synthesis Of GEMLESS Chromophore Ring Moiety

The synthesis of GEMLESS chromophore ring moiety (compound **II-5**) was carried out according to the synthetic scheme shown below.



Scheme IV-4: Synthesis of GEMLESS chromophore ring moiety (compound **II-5**).

*Synthesis of 2-bromo,1-cyclohexene carbaldehyde, **II-8**:* Chloroform (50 mL, 2 days incubated with activated 4Å molecular sieves) and distilled DMF (11.7 mL, 152.7 mmol) were added to an oven dried 500 mL of round bottom flask and the mixture was cooled down to 0 °C in an ice bath. PBr_3 (13.8 mL, 137.6 mmol) was added dropwise to the reaction mixture over 15 min. The resulting solution (whitish suspension) was stirred for 1 h before the addition of cyclohexanone (**II-9**) (5 g, 50.9 mmol). The reaction mixture turned into a bright red clear solution and was stirred overnight at room temperature.

Work up: The reaction was cooled down and neutralized with solid $NaHCO_3$ on ice bath to a pH between 6.0-7.0. The resultant orange mixture was transferred to a 1 L extraction funnel and was washed with water (250 mL) and dichloromethane (250 mL). The water layer was extracted twice with dichloromethane. The combined organic layers were washed with brine solution (300 mL), dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. The residue was purified by elution through a short silica column with 10% ethyl acetate/hexane to afford 2-bromo,1-cyclohexene carbaldehyde as an oil (4 g, 21.16 mmol, 42% yield).

The above synthetic procedure was taken from Lin *et al.*[25]

^{1}H NMR (500 MHz, $CDCl_3$) δ [ppm] = 1.64 - 1.69 (*m*, 2 H), 1.72 - 1.77 (*m*, 2 H), 2.24 - 2.28 (*m*, 2 H), 2.71 - 2.74 (*m*, 2 H), 10.00 (*s*, 1H).

^{13}C-NMR (125 MHz, $CDCl_3$) δ [ppm] = 21.1, 24.2, 24.9, 38.8, 135.3, 143.5, 193.7.

*Synthesis of compound **II-7**:* 2-Bromocyclohex-1-ene-1-carbaldehyde **II-8** (4 g, 0.021 mol), ethylene glycol (3.3 g, 0.053 mol) and catalytic amount of *para*-toluene sulfonic acid, PTSA, (200 mg, 1.06 mmol) were added to benzene (15 mL). Activated molecular sieves were added to the solution. Using a Dean Stark apparatus, the reaction mixture was refluxed for 8 h.

Work up: After being cooled to room temperature, the mixture was filtered over celite and washed with dichloromethane. The organic phase was wash with brine, dried over anhydrous Na_2SO_4, and concentrated under reduced pressure. The residue was purified by elution through a short silica column with 10% ethyl acetate/hexane to afford the protected aldeyhde (3.207 g, 0.013 mol, 65% yield). Note that, this product is not stable. It can be stored at -80 °C for only couple of days.

^{1}H NMR (500 MHz, $CDCl_3$) δ [ppm] = 1.62 - 1.71 (*m*, 4 H), 2.13 - 2.16 (*m*, 2 H), 2.29 - 2.52 (*m*, 2 H), 3.88 - 4.02 (*m*, 4 H), 5.73 (*s*, 1H).

*Synthesis of compound **II-6**:* Acetal-protected aldehyde, **II-7**, (3.207 g, 0.013 mol) was dissolved in dry THF (35 mL), and the resulting solution was cooled down to -78 °C. *tert*-Butyl lithium in pentane, (16.8 mL, 1.7 M, 0.0286 mol) was added dropwise to the reaction mixture. The reaction mixture was kept at -78 °C for 1.5 h while stirring. Formation of the organolithium intermediate was observed as white particles. Iodomethane (CH_3I) (3.3 mL, 7.381 g, 0.052 mol) was added dropwise to the reaction. The reaction was warmed to -30 °C and kept at this temperature for 5 h.

Work up: The reaction was quenched with the addition of brine and extracted with diethyl ether (twice). The organic layer was dried over anhydrous

Na_2SO_4, and concentrated under reduced pressure. Compound **II-6** was not purified further and the crude product (2.3 g) was carried out to the next step.

^{1}H NMR (500 MHz, $CDCl_3$) δ [ppm] = 1.56 - 1.58 (*m*, 4 H), 1.71 (*s*, 3 H), 1.94 - 1.98 (*m*, 2 H), 1.99 -2.03 (*m*, 2 H), 3.86 - 3.99 (*m*, 4 H), 5.62 (*s*, 1H).

^{13}C-NMR (125 MHz, $CDCl_3$) δ [ppm] = 18.6, 21.9, 22.5, 22.8, 32.5, 65.2, 101.5, 126.9, 135.4.

Synthesis of GEMLESS ring moiety, ***II-5***: The crude **II-6** from the previous step (2.3 g) was dissolved in THF (25 mL). Catalytic amount of *para*-toluene sulfonic acid, PTSA, (100 mg, 0.52 mmol) together with water (10 mL) was added into the reaction mixture and was stirred at 75 °C for 7 h.

Work up: The reaction was cooled down to room temperature and was extracted with brine and dichloromethane. The final aldehyde was obtained from the organic layer, dried over Na_2SO_4, and concentrated under reduced pressure. The oily product was carried out to next step without further purification.

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 1.56 - 1.62 (*m*, 4H), 2.11 (*s*, 3H), 2.14 - 2.21 (*m*, 4H), 10.13 (*s*, 1H).

^{13}C-NMR (125 MHz, $CDCl_3$) δ [ppm] = 18.3, 21.7, 22.0, 22.2, 34.2, 133.8, 156.1, 191.0.

The analytical data are in agreement with the previously reported in literature.[26,27]

IV.13.5 Spectroscopic Data Of Polyenic Compounds

As illustrated in **Scheme IV-2,** all the retinal analogs were synthesized via four consecutive steps. Below, spectroscopic data for the polyenic compounds (shown in **Figure IV-11**) are provided below.

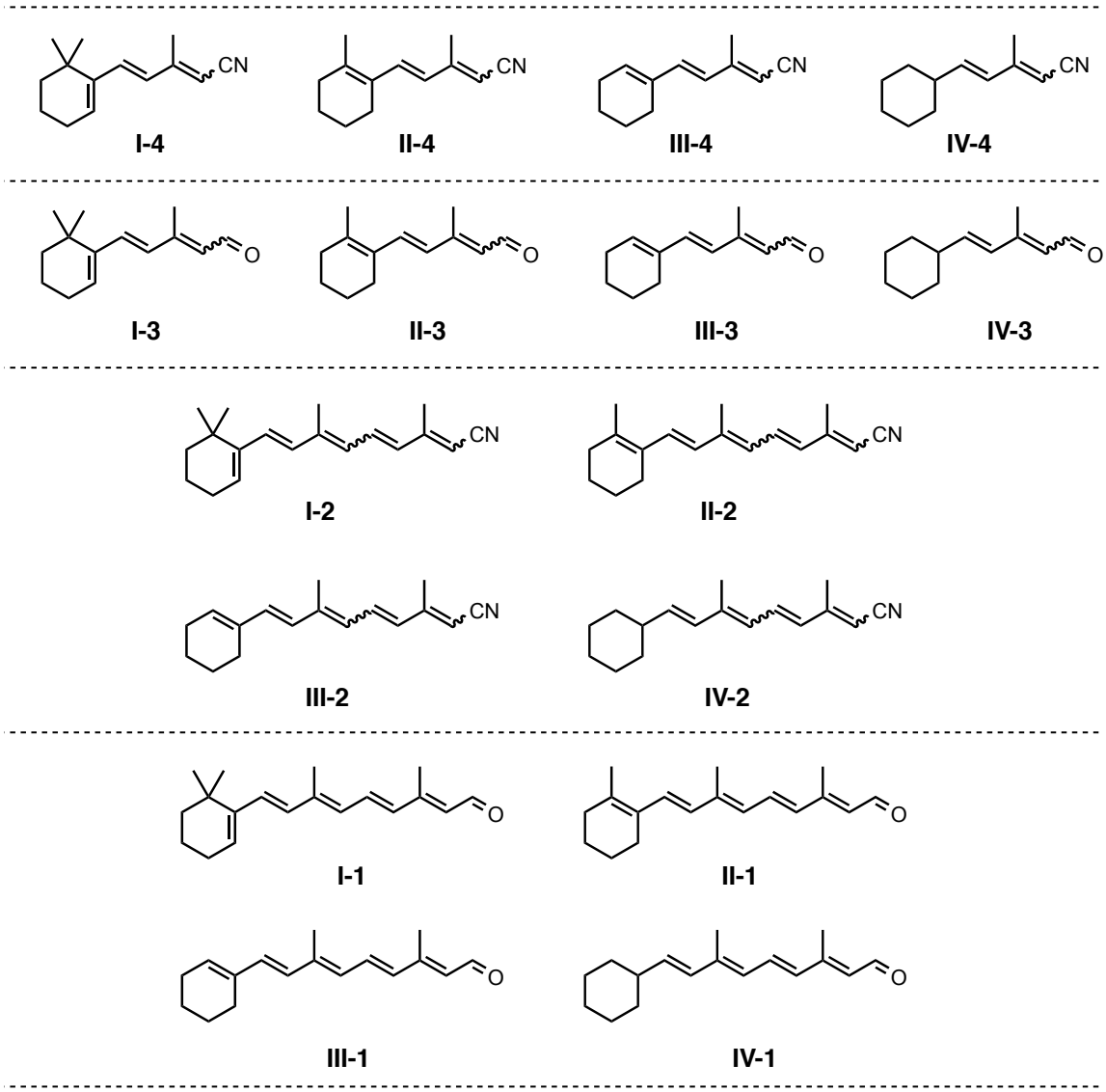


Figure IV-11: Structures of olefinic compounds discussed in the text.

*Compound **I-4***: This compound was obtained as an oily yellow product with a distinct citrus fruit smell and showed a mixture of isomers at C10 (double

bond directly conjugated to the nitrile). The all-*trans* isomer was the major product with a ratio of 3:1. ^{1}H NMR data is reported individually for the *E* and *Z* isomers while the carbon NMR peaks are reported together.

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.05 (*s*, 6 H), 1.45 - 1.49 (*m*, 2 H), 1.55 - 1.61 (*m*, 2 H), 2.05 - 2.12 (*m*, 2 H), 2.15 (*d*, *J* = 1.2 Hz, 3 H), 5.15 (*s*, 1H), 5.94 (*t*, *J* = 4.2 Hz, 1 H), 6.38 (*d*, *J* = 16.0 Hz, 1 H), 6.52 (*d*, *J*= 16.0 Hz, 1 H).

^{1}H NMR (500 MHz, $CDCl_3$) for 10-*Z* isomer δ [ppm] = 1.07 (*s*, 6 H), 1.45 - 1.50 (*m*, 2 H), 1.55 - 1.63 (*m*, 2 H), 2.05 - 2.11 (*m*, 2 H), 2.15 (*d*, *J* = 0.9 Hz, 3 H), 5.05 (*s*, 1H), 6.04 (*t*, *J* = 4.2 Hz, 1 H), 6.56 (*d, J* = 16.5 Hz, 1 H), 6.91 (*d*, *J* = 15.5 Hz, 1 H).

^{13}C NMR (125 MHz, $CDCl_3$) for mixture of isomers δ [ppm] = 14.11, 16.62, 18.79, 19.41, 22.68, 26.53, 28.45, 28.48, 29.35, 29.68, 30.29, 31.92, 33.38, 33.41, 39.19, 39.24, 94.93, 96.36, 118.10, 125.07, 127.47, 127.68, 128.17, 136.52, 137.13, 143.13, 156.81, 157.33.

*Compound **II-4***: This compound was obtained as an oily needle shaped yellow crystals. Similar to compound **I-4**, a mixture of isomers was isolated after purification with a 5:1 ratio of *E:Z* olefinic isomers (double bond directly conjugated to the nitrile). Dr. Richard Staples was able to extract a single crystal to provide crystallographic X-ray data of the all-*trans* isomer (**Figure IV-12**).

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.57 - 1.67 (*m*, 4 H), 1.84 (*s*, 3 H), 2.10 - 2.14 (*m*, 4 H), 2.18 (*s*, 3 H), 5.15 (*s*, 1H), 6.18 (*d*, *J* = 16.5 Hz, 1 H), 7.04 (*d*, *J* = 16.0 Hz, 1 H).

^{1}H NMR (500 MHz, $CDCl_3$) for 10-*Z* isomer δ [ppm] = 1.57 - 1.69 (*m*, 4 H), 2.03 (*s*, 3 H), 2.11 - 2.14 (*m*, 4 H), 2.15 (*s,* 3 H), 5.04 (*s*, 1H), 6.69 (*d*, *J* = 16.0 Hz, 1 H), 7.09 (*d*, *J* = 15.5 Hz, 1 H).

^{13}C NMR (125 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 16.65, 19.66, 22.51, 22.58, 25.29, 33.49, 95.84, 118.32, 124.93, 127.63, 134.43, 140.47, 157.87.

^{13}C NMR (125 MHz, $CDCl_3$) for 10-*Z* isomer δ [ppm] = 16.62, 19.48, 22.17, 22.55, 25.35, 33.58, 94.50, 122.33, 124.68, 128.04, 134.90, 139.69, 157.26.

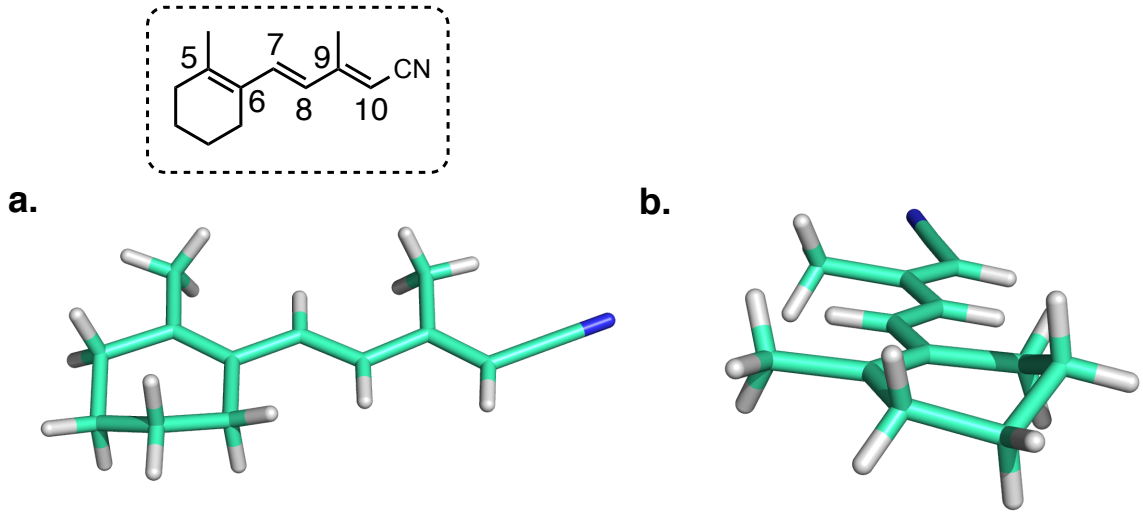


Figure IV-12: X-ray crystal structure of compound **II-4**. **a.** The chemical structure is shown in the dashed box and the carbons are numbered as indicated. A single crystal of the all-*trans* isomer in 6-s-*trans* conformation is represented (hydrogens-white, carbons-green, nitrogen-blue). **b.** The polyene chain displays a planar geometry while the cyclohexene ring is slightly puckered.

*Compound **III-4***: Compound **II-4** and compound **III-4** exhibits similar physical properties. Compound **III-4** was also obtained as oily needle shaped yellow crystals as a mixture of isomers (3:1 ratio of *E:Z* olefinic isomers). The crystallographic X-ray data of the all-*trans* isomer is shown in **Figure IV-13**.

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.57 - 1.62 (*m*, 2 H), 1.64 - 1.70 (*m*, 2 H), 2.11 - 2.14 (*m*, 2 H), 2.15 (*s*, 3 H), 2.16 - 2.2 (*m*, 2 H), 5.15 (*s*, 1H), 6.00 (*t*, *J* = 4.0 Hz, 1 H), 6.14 (*d*, *J* = 16.5 Hz, 1 H), 6.52 (*d*, *J* = 16.0 Hz, 1 H).

^{1}H NMR (500 MHz, $CDCl_3$) for 10-*Z* isomer δ [ppm] = 1.57 - 1.62 (*m*, 2 H), 1.64 - 1.70 (*m*, 2 H), 2.11 - 2.14 (*m*, 2 H), 2.00 (*d*, *J* = 1.5 Hz, 3 H), 2.16 - 2.2 (*m*, 2 H), 5.05 (*s*, 1H), 6.00 (*t*, *J* = 4.0 Hz, 1 H), 6.55 (*d*, *J* = 16.0 Hz, 1 H), 6.65 (*d*, *J* = 16.0

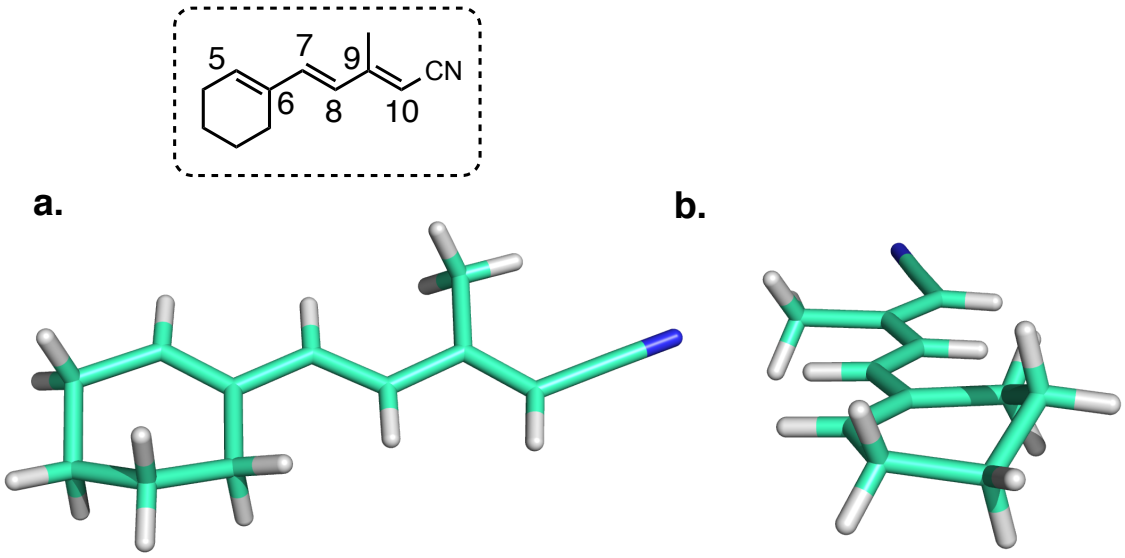


Figure IV-13: X-ray crystal structure of compound **III-4**. **a.** The chemical structure is shown in the dashed box and the carbons are numbered as indicated. A single crystal of the all-*trans* isomer in 6-s-*trans* conformation is represented (hydrogens-white, carbons-green, nitrogen-blue). **b.** The polyene chain displays a planar geometry while the cyclohexene ring is slightly puckered.

Hz, 1 H).

^{13}C NMR (125 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 16.59, 19.42, 22.13, 24.26, 26.35, 96.18, 118.17, 124.62, 135.31, 136.05, 139.68, 157.42.

^{13}C NMR (125 MHz, $CDCl_3$) for 10-*Z* isomer δ [ppm] = 16.59, 19.42, 22.16, 24.31, 26.41, 33.58, 94.87, 117.29, 122.03, 135.71, 136.20, 140.25, 156.84.

*Compound **IV-4***: Compound **IV-4** was obtained as a colorless oil after purification. The ratio of the *E*:*Z* isomers was determined to be 2:1 based on proton NMR. Due to overlapping peaks, we were not able to assign the individual NMR signals for each isomer.

^{1}H NMR (500 MHz, $CDCl_3$) for mixture of isomers δ [ppm] = 1.06 - 1.19 (*m*), 1.23 - 1.31 (*m*), 1.64 - 1.67 (*m*), 1.70 -1.75 (*m*), 1.96 (*s*, CH_3-*Z* isomer), 2.11 (*s*, CH_3-*E* isomer), 5.04 (*s*, C10-H, *Z* isomer), 5.10 (*s*, C10-H, *E* isomer), 5.99 - 6.09 (overlapped doublets for C7-C8 protons), 6.60 (*d*, *J* = 15.5 Hz, C7-H or C8-H, *Z* isomer).

^{13}C NMR (125 MHz, $CDCl_3$) for mixture of isomers δ [ppm] = 16.68, 19.46, 25.77, 25.91, 32.39, 41.11, 41.26, 94.75, 96.06, 117.11, 117.94, 125.60, 128.09, 145.01, 145.77, 156.99, 157.54.

*Compound **I-3***: This compound was obtained as a dark oily material, showing a mixture of isomers with a 3:1 ratio of *E:Z* olefinic compounds (double bond directly conjugated to the aldehyde, C10). The E/Z stereochemistry of the

C9-C10 double bond was verified by nOE and for the rest of the products it was assumed that the same selectivity was maintained.

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.08 (*s*, 6 H), 1.47 - 1.50 (*m*, 2 H), 1.57 - 1.62 (*m*, 2 H), 2.09 - 2.12 (*m*, 2 H), 2.26 (*s*, 3 H), 5.94 (*d*, *J* = 8.5 Hz, 1H), 6.00 (*t*, *J* = 4.0 Hz, 1 H), 6.46 (*d*, *J* = 16.0 Hz, 1 H), 6.70 (*d*, *J* = 15.5 Hz, 1 H), 10.09 (*d*, *J* = 8.5 Hz, 1 H).

^{1}H NMR (500 MHz, $CDCl_3$) for 10-*Z* isomer δ [ppm] = 1.08 (*s*, 6 H), 1.47 - 1.50 (*m*, 2 H), 1.57 - 1.62 (*m*, 2 H), 2.08 (*s*, 3 H), 2.09 - 2.12 (*m*, 2 H), 5.82 (*d*, *J* = 8.0 Hz, 1H), 6.00 (*t*, *J* = 4.0 Hz, 1 H), 6.60 (*d*, *J* = 15.5 Hz, 1 H), 7.36 (*d*, *J*= 16 Hz, 1 H), 10.2 (*d*, *J* = 8.0 Hz, 1 H).

^{13}C NMR (125 MHz, $CDCl_3$) for mixture of isomers δ [ppm] = 13.07, 18.83, 26.56, 26.59, 28.52, 28.53, 30.30, 33.46, 39.23, 127.66, 127.69, 128.93, 130.64, 143.42, 155.13, 190.11, 191.30.

*Compound **II-3***: Compound **II-3** was purified as a dark oily mixture containing 12:1 ratio of *E*:*Z* isomers. As a result of the high *E*:*Z* isomeric ratio, the abundance of ^{13}C for the 10-*Z* isomer is low and not detectable.

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.58 - 1.66 (*m*, 4 H), 1.86 (*s*, 3 H), 2.11 - 2.18 (*m*, 4 H), 2.39 (*s,* 3 H), 5.96 (*d*, *J* = 8.0 Hz, 1 H), 6.25 (*d*, *J* = 16.0 Hz, 1 H), 7.24 (*d*, *J* = 15.5 Hz, 1 H), 10.07 (*d*, *J* = 7.5 Hz, 1 H).

^{1}H NMR (500 MHz, $CDCl_3$) for 10-*Z* isomer δ [ppm] = 1.58 - 1.66 (*m*, 4 H), 1.86

(*s*, 3 H), 2.11 - 2.18 (*m*, 4 H), 2.26 (*s,* 3 H), 5.81 (*d*, *J* = 8.0 Hz, 1 H), 6.70 (*d*, *J* = 15.5 Hz, 1 H), 7.25 (*d*, *J* = 15.5 Hz, 1 H), 10.18 (*d*, *J* = 8.0 Hz, 1 H).

^{13}C NMR (125 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 13.04, 19.63, 22.44, 22.51, 25.26, 33.48, 127.79, 128.05, 128.56, 134.33, 140.53, 155.93, 191.07.

Compound ***III-3***: This compound was obtained as a oily solid material with an *E*:*Z* ratio of 3:1.

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.57 - 1.62 (*m*, 2 H), 1.66 - 1.71 (*m*, 2 H), 2.16 - 2.20 (*m*, 4 H), 2.26 (*d*, *J* = 1.0 Hz, 3 H), 5.94 (*d*, *J* = 8.0 Hz, 1 H), 6.03 (*t*, *J* = 3.0 Hz, 1 H), 6.21 (*d*, *J* = 15.5 Hz, 1 H), 6.70 (*d*, *J* = 16.0 Hz, 1 H), 10.08 (*d*, *J* = 8.0 Hz, 1 H).

^{1}H NMR (500 MHz, $CDCl_3$) for 10-*Z* isomer δ [ppm] = 1.57 - 1.62 (*m*, 2 H), 1.66 - 1.71 (*m*, 2 H), 2.08 (*d*, *J* = 1.5 Hz, 3 H), 2.16 - 2.20 (*m*, 4 H), 5.82 (*d*, *J* = 8.0 Hz, 1 H), 6.03 (*t*, *J* = 3.0 Hz, 1 H), 6.60 (*d*, *J* = 15.5 Hz, 1 H), 7.12 (*d*, *J* = 15.5 Hz, 1 H), 10.18 (*d*, *J* = 8.0 Hz, 1 H).

^{13}C NMR (125 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 13.03, 21.26, 22.18, 24.34, 26.43, 127.65, 128.97, 135.81, 136.08, 139.66, 155.38, 191.18.

^{13}C NMR (125 MHz, $CDCl_3$) for 10-*Z* isomer δ [ppm] = only following distinct peaks are noticeable; 119.50, 127.56, 136.25, 140.69, 155.32, 190.06.

Compound ***IV-3***: This compound was obtained as an oily solid material with an *E*:*Z* ratio of 2.3:1.0. Most of the ^{1}H NMR signals overlap with each other;

therefore, it is not possible to report the individual NMR data for each isomer.

^{1}H NMR (500 MHz, $CDCl_3$) for mixture of isomers δ [ppm] = 1.08 - 1.32 (*m*), 1.65 - 1.69 (*m*), 1.73 - 1.76 (*m*), 2.03 (*d*, *J* = 1.5 Hz, CH_3-*Z* isomer), 2.08 - 2.14 (*m,* 1H), 2.22 (*d*, *J* = 1.5 Hz, CH_3-*E* isomer), 5.79 (*d*, *J* = 8.0 Hz, C10-H, *Z* isomer), 5.87 (*d*, *J* = 8.5 Hz, C10-H, *E* isomer), 6.07 - 6.22 (overlapped doublets for C7-C8 protons), 7.01 (*d*, *J* = 16.0 Hz, C7-H or C8-H, *Z* isomer), 10.08 (*d*, *J* = 8.5 Hz, C10-H, *E* isomer), 10.15 (*d*, *J* = 8.5 Hz, C10-H, *Z* isomer).

^{13}C NMR (125 MHz, $CDCl_3$) for mixture of isomers δ [ppm] = 13.07, 21.42, 25.78, 25.80, 25.95, 32.44, 41.41, 41.65, 122.93, 127.50, 128.54, 131.02, 145.02, 145.08, 155.31, 155.49, 190.29, 191.49.

All elongated nitriles, compounds **I-2**, **II-2**, **III-2** and **IV-2**, were isolated as a mixture of different isomeric forms. The reactions were monitored by thin layer chromatography (TLC) to reach completion. Additionally, these compounds have distinct citrus fruit smell, which is different than the starting aldehyde. Since NMR data is complicated due to the presence of different isomeric forms, the spectroscopic information of these compounds is not reported here. More detailed information for the following products is provided below.

*Compound **I-1** (5DEM retinal aldehyde)*: All-*trans* isomer of 5DEM aldehyde was purified as an oily needle shaped yellow crystals. Unfortunately, crystals were too small to obtain X-ray crystallographic data.

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.07 (*s*, 6H), 1.46 - 1.49 (*m*, 2H), 1.56 - 1.61 (*m*, 2H), 1.98 (*s*, 3H), 2.08 (*q*, *J* = 5.5 Hz, 2H), 2.30 (*s*, 3H), 5.89 (*t*, *J* = 4.2 Hz, 1H), 5.95 (*d*, *J* = 8.5 Hz, 1H), 6.20 (*d*, *J* = 11.0 Hz, 1H), 6.34 (*d*, *J* = 15.5 Hz, 1H), 6.35 (*d*, *J* = 15.0 Hz, 1H), 6.46 (*d*, *J* = 16.0 Hz, 1H), 7.1 (*dd*, *J* = 11.0, 15.0 Hz, 1H), 10.08 (*d*, *J* = 8.5 Hz, 1H).

^{13}C NMR (125 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 13.10, 13.15, 18.99, 26.54, 28.60, 30.30, 33.52, 39.37, 124.61, 129.03, 129.80, 130.63, 132.00, 132.51, 134.55, 141.32, 143.91, 154.79, 191.11.

The UV-vis spectrum of the final 5DEM aldehyde was taken in ethanol (**Figure IV-14**). The free aldehyde absorbs maximally at 388 nm while the following Schiff base formed with *n*-butylamine results in a blue shift, showing a

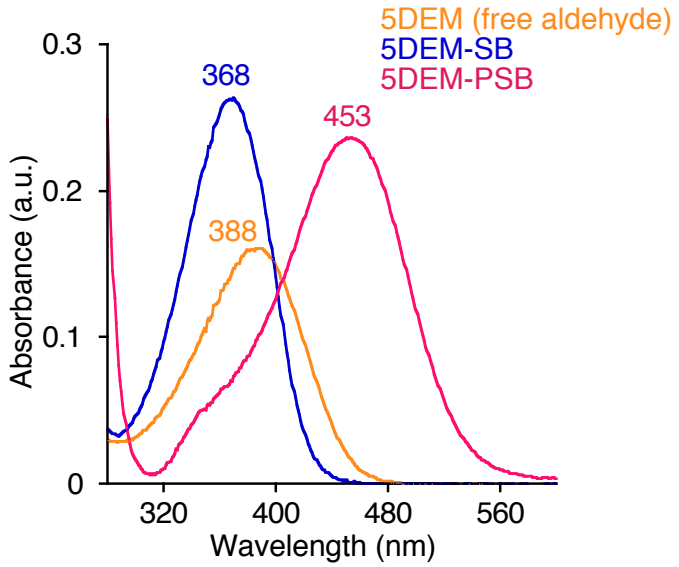


Figure IV-14: UV-vis spectra of compound **I-1** (5DEM retinal aldehyde) in ethanol.

λ^{max} value of 368 nm. Acidification of the Schiff base with 37% HCl solution yields the iminium (λ^{max} of 5DEM-PSB = 453 nm).

*Compound **II-1** (GEMLESS retinal aldehyde)*: All-*trans* isomer of GEMLESS aldehyde was purified as an oily needle shaped yellow crystal. X-ray crystallographic data was obtained by Dr. Richard Staples, revealing that the ring-polyene junction also favors a 6-s-*trans* conformation in the solid state, which was calculated to be the energetically more stable conformer in solution (**Figure IV-15**).[28] The UV-vis spectrum of the GEMLESS aldehyde is also shown in **Figure IV-15**.

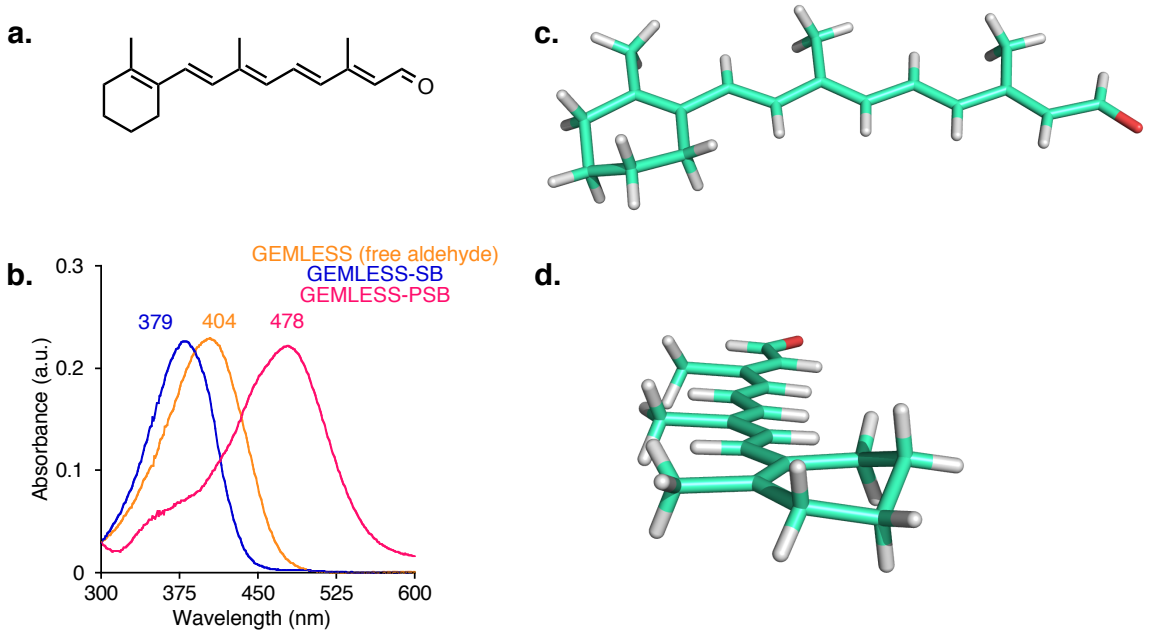


Figure IV-15: Characterization of compound **II-1** (GEMLESS retinal aldehyde). **a.** The structure of the GEMLESS retinal aldehyde. **b.** UV-vis spectrum of the GEMLESS chromophore in ethanol (orange-free aldehyde, blue-Schiff Base, pink-protonated Schiff Base). **c.** Crystal structure of the GEMLESS chromophore is represented by a stick model (hydrogens-white, carbons-green, oxygen-red). **d.** Side view of the chromophore showing the planar polyene backbone and the 6-s-*trans* ring-chain junction.

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.56 - 1.61 (*m*, 2H), 1.62 - 1.67 (*m*, 2H), 1.83 (*s*, 3H), 2.02 (*s,* 3H), 2.09 - 2.13 (*m*, 2H), 2.16 - 2.20 (*m*, 2H), 2.30 (*d, J* = 1.0 Hz, 3H), 5.95 (*d*, *J* = 8.0 Hz, 1H), 6.23 (*d*, *J* = 11.5 Hz, 1H), 6.27 (*d*, *J* = 16.0 Hz, 1H), 6.34 (*d*, *J* = 15.5 Hz, 1H), 6.88 (*d*, *J* = 16.0 Hz, 1H), 7.12 (*dd*, *J* = 15.0 Hz, 11.5 Hz, 1H), 10.08 (*d*, *J* = 8.5 Hz, 1H).

^{13}C NMR (125 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 13.11, 13.16, 19.62, 22.75, 22.83, 25.48, 33.43, 128.28, 128.91, 129.26, 129.49, 129.79, 132.65, 134.26, 136.58, 141.92, 154.84, 191.06.

*Compound **III-1** (CYCHXE retinal aldehyde)*: Similar to compound **II-1**, CYCHXE was also obtained as an oily needle shaped yellow crystals. **Figure IV-16** depicts the solid state crystal structure of the CYCHXE retinal aldehyde.

^{1}H NMR (500 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 1.57 - 1.62 (*m*, 2H), 1.66 - 1.70 (*m*, 2H), 1.99 (*s*, 3H), 2.15 - 2.19 (*m*, 4H), 2.30 (*d, J* = 0.5 Hz, 3H), 5.88 (*d*, *J* = 4.5 Hz, 1H), 5.95 (*d*, *J* = 8.0 Hz, 1H), 6.21 (*d*, *J* = 11.5 Hz, 1H), 6.23 (*d*, *J* = 15.5 Hz, 1H), 6.34 (*d*, *J* = 15.5 Hz, 1H), 6.39 (*d*, *J* = 16.0 Hz, 1H), 7.10 (*dd*, *J* = 14.5 Hz, 11.0 Hz, 1H), 10.08 (*d*, *J* = 8.0 Hz, 1H).

^{13}C NMR (125 MHz, $CDCl_3$) for all-*trans* isomer δ [ppm] = 13.08, 13.10, 22.41, 24.50, 26.30, 128.99, 129.15, 129.93, 132.33, 132.54, 134.36, 134.39, 136.08, 141.35, 154.80, 191.08.

*Compound **IV-1** (CYCHXA retinal aldehyde)*: All-*trans* isomer of the

compound **IV-1** was purified as oily needle shaped yellow crystals.

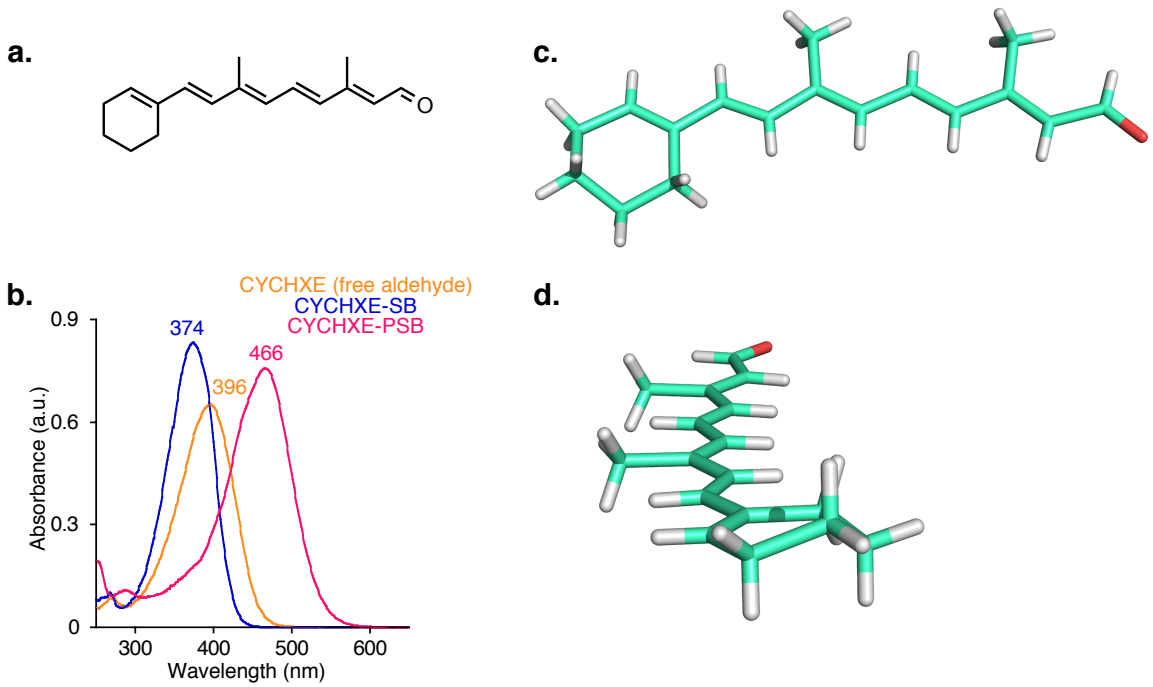


Figure IV-16: Characterization of compound **III-1** (CYCHXE retinal aldehyde). **a.** The structure of the CYCHXE retinal aldehyde. **b.** UV-vis spectrum of CYCHXE in ethanol (orange-free aldehyde, blue-Schiff Base, pink-protonated Schiff Base). **c.** Crystal structure of CYCHXE is represented by a stick model (hydrogens-white, carbons-green, oxygen-red). **d.** Side view of the chromophore showing the planar polyene backbone and the 6-s-*trans* ring-chain junction.

^{1}H NMR (500 MHz, $CDCl_3$) for mixture of isomers δ [ppm] = 1.08 - 1.28 (*m*, 5H), 1.63 - 1.73 (*m*, 5H), 1.93 (*s*, 3H), 2.02 - 2.08 (*m,* 1H), 2.29 (*s*, 3H), 5.80 (*dd*, *J* = 15.5 Hz, 6.5 Hz, 1H), 5.94 (*d*, *J* = 8.5 Hz, 1H), 6.10 (*d*, *J* = 16.0 Hz, 1H), 6.11 (*d*, *J* = 12.0 Hz, 1H), 6.32 (*d*, *J* = 15.0 Hz, 1H), 7.08 (*dd*, *J* = 15.0 Hz, 11.0 Hz, 1H), 10.08 (*d*, *J* = 8.5 Hz, 1H).

^{13}C NMR (125 MHz, $CDCl_3$) for mixture of isomers δ [ppm] = 13.09, 13.18, 25.97, 26.08, 32.93, 41.31, 128.73, 128.96, 131.89, 132.57, 134.29, 139.13,

141.19, 154.88, 191.15.

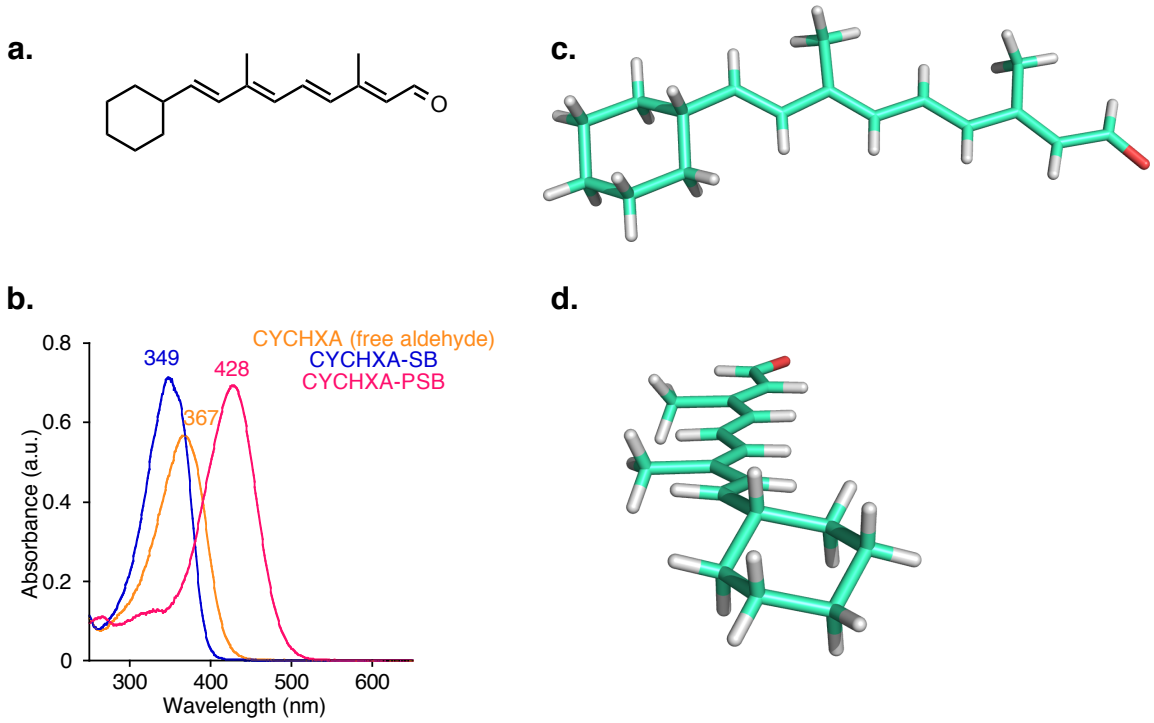
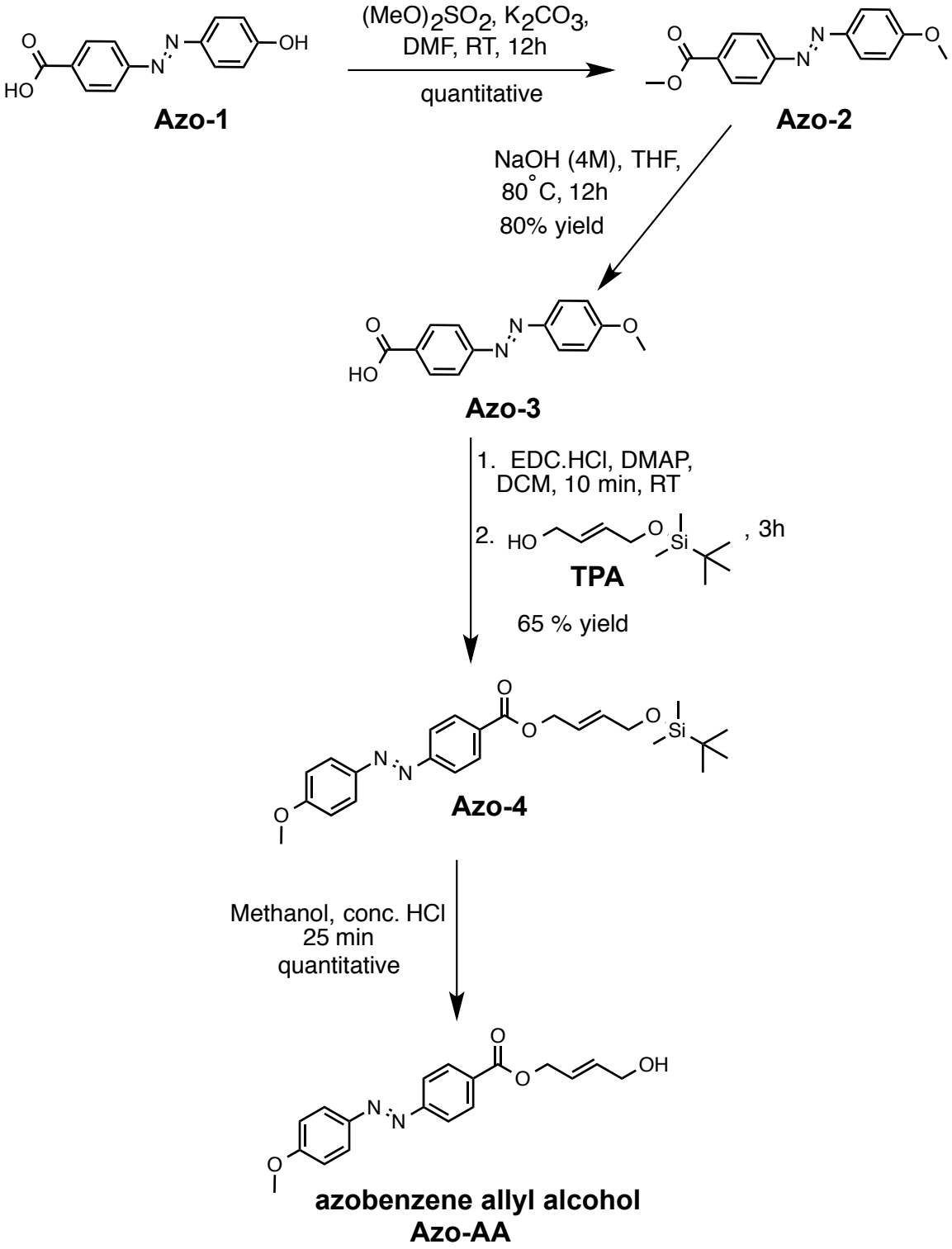


Figure IV-17: Characterization of compound **IV-1** (CYCHXA retinal aldehyde). **a.** The structure of the CYCHXA retinal aldehyde. **b.** UV-vis spectrum of CYCHXA in ethanol (orange-free aldehyde, blue-Schiff Base, pink-protonated Schiff Base). **c.** Crystal structure of CYCHXA is represented by a stick model (hydrogens-white, carbons-green, oxygen-red). **d.** Side view of the chromophore showing the planar polyene backbone and the chair form of a cyclohexyl ring.

IV.13.6 Synthesis Of Azobenzene Allyl Alcohol

The synthesis of azobenzene allyl alcohol (**Azo-AA**) was carried out as illustrated in **Scheme IV-5**.

*Synthesis of compound **Azo-2**:* The synthesis of **Azo-1** was described previously by Dr. Wenjing Wang in her dissertation. **Azo-1** (500 mg, 2.06 mmol) was dissolved in dry DMF (20 mL) together with 2.5 equivalents of K_2CO_3 (0.71 g, 5.16 mmol). The resulting mixture was vigorously stirred at room temperature



Scheme IV-5: Synthetic protocol of azobenzene allyl alcohol (**Azo-AA**).

for 20 min under inert atmosphere. Dimethyl sulfate, $(MeO)_2SO_2$, (0.78 g, 6.19 mmol) was added, to the solution and was stirred at room temperature. The

reaction was monitored by TLC until it was complete (12 h). The reaction required addition of another portions of K_2CO_3 (2.5 equivalents) and $(MeO)_2SO_2$, (2.5 equivalents) to achieve complete dimethylation. Otherwise, undesired mono-methylated side product (methyl ester with free phenol) was seen in the reaction mixture.

Work up: The reaction mixture was quenched with 1 M of aqueous HCl solution, extracted into dichloromethane, washed with water and dried over anhydrous Na_2SO_4. The mixture was transferred into a round bottom flask and the solvent was removed under vacuum to give a dark orange solid. The crude solid product was not purified further and carried to next step (quantitative crude yield). The NMR data were in agreement with the previous report.[29]

Retention Factor, Rf, (30% ethyl acetate/hexane) = 0.69

Melting point: 165 - 169 $^{\circ}$C

^{1}H NMR (500 MHz, DMSO-d^{6}) δ [ppm] = 3.88 (*s*, 3H), 3.89 (*s*, 3H), 7.16 (*d*, *J* = 9.0 Hz, 2H), 7.94 (*d*, *J* = 8.0 Hz, 2H), 7.95 (*d*, *J* = 8.5 Hz, 2H), 8.14 (*d*, *J* = 8.5 Hz, 2H).

^{13}C-NMR (125 MHz, DMSO-d^{6}) δ [ppm] = 52.39, 55.76, 114.79, 122.39, 125.08, 130.50, 130.88, 146.22, 154.62, 162.70, 165.68.

*Synthesis of compound **Azo-3**:* Hydrolysis of **Azo-2** (250 mg, 0.92 mmol) was performed in a mixture of aqueous NaOH solution (4 M) and THF. The reaction mixture was refluxed at 80 °C for 12 h.

Work up: The reaction mixture was quenched with 1 M of aqueous HCl solution, extracted into dichloromethane, washed with water and dried over anhydrous Na_2SO_4. The organic solvent was removed under vacuum to yield a solid orange residue. The residue was loaded on a silica column, packed with 30% ethyl acetate/hexane solvent system. After 50 mL of 30% ethyl acetate/hexane, the eluent was changed to dichloromethane and then gradually to dichloromethane:methanol mixture (1:1). The fractions containing the **Azo-3** were combined and concentrated under reduced pressure to yield an orange solid (80% yield).

Rf (30% ethyl acetate/hexane) = 0.25 (smeared spot)

Melting point: 245 - 250 °C

^{1}H NMR (500 MHz, DMSO-d^6) δ [ppm] = 3.88 (*s*, 3H), 7.16 (*d*, *J* = 8.5 Hz, 2H), 7.904 (*d*, *J* = 8.0 Hz, 2H), 7.94 (*d*, *J* = 9.0 Hz, 2H), 8.11 (*d*, *J* = 8.5 Hz, 2H), 13.19 (*broad s*, 1H)

^{13}C-NMR (125 MHz, DMSO-d^6) δ [ppm] = 55.75, 114.76, 122.17, 124.98, 130.54, 132.150, 146.23, 154.44, 162.55, 166.76.

*Synthesis of compound **Azo-4**:* **TPA** *tert*-Butyldimethylsilyl protected

alcohol was previously synthesized by Mr. Calvin Grant. The esterification of **TPA** and **Azo-3** was achieved via EDC.HCl coupling based on the protocol, which was established in Prof. Tepe's lab at MSU (procedure was adapted from Dr. Nicole Hewlett and Mr. Jake Ludwig). **Azo-3** (80 mg, 0.31 mmol) was added to dichloromethane, mixed with 1.5 equivalents of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (71 mg, 0.46 mmol), and 0.3 equivalents of 4-dimethylaminopyridine (11 mg, 0.09 mmol) and stirred for 10 min at room temperature. When all the solids were homogeneously dissolved in DCM, **TPA** (125 mg, 0.62 mmol) was added to the reaction mixture and vigorously stirred under inert atmosphere. Reaction was monitored by TLC until it was complete (3 h).

Work up: The reaction was quenched with saturated sodium bicarbonate solution, and extraction was performed with DCM twice. The combined organic layers were dried over anhydrous Na_2SO_4, and concentrated under reduced pressure. Following elution through a short silica column (20% ethyl acetate/hexane) the **Azo-4** was isolated as a bright orange solid (65% yield).

Rf (30% ethyl acetate/hexane) = 0.65

^{1}H NMR (500 MHz, DMSO-d^6) δ [ppm] = 0.05 (*s*, 6H), 0.087 (*s*, 9H), 3.88 (*s*, 3H), 4.19 (*m*, 2H), 4.84 (*dd*, *J* = 6.0 Hz, *J* =1.0 Hz, 2H), 5.87 - 5.93 (*m*, 1H), 5.98 (*ddd, J* = 15.5 Hz, *J* = 4.5 Hz, *J* = 4.5 Hz, 1H), 7.16 (*d, J* = 9.0 Hz, 2H), 7.94 (*d, J* = 8.5 Hz, 2H), 7.95 (*d, J* = 9.0 Hz, 2H), 8.15 (*d, J* = 9.0 Hz, 2H).

^{13}C-NMR (125 MHz, DMSO-d^{6}) δ [ppm] = -5.32, -5.28, 25.79, 55.75, 62.25, 64.79, 114.78, 122.42, 123.19, 125.09, 130.50, 130.91, 134.21, 146.22, 154.66, 162.69, 164.89.

Synthesis of ***azobenzene allyl alcohol (Azo-AA)****:* **Azo-4** (85 mg, 0.19 mmol) was dissolved in methanol (10 mL) and the mixture stirred at room temperature for a couple minutes. Addition of 3 drops of concentrated HCl to the solution facilitates the deprotection of the *tert*-butyldimethylsilyl group, TBDMS, in 25 min.

Work up: Methanol was removed under reduced pressure and the resulting residue was dissolved in dichloromethane, washed with brine, and dried over anhydrous Na_2SO_4. Bright orange solid product was obtained upon removal of solvent under reduced pressure with in quantitative yield (62.9 g).

Rf (30% ethyl acetate/hexane) = 0.23

Melting point: 109 - 113 °C

^{1}H NMR (500 MHz, DMSO-d^{6}) δ [ppm] = 3.88 (*s*, 3H), 3.99 (*m*, 2H), 4.84 (*dd*, *J* = 6.0 Hz, *J* =1.0 Hz, 2H), 5.84 - 5.90 (*m*, 1H), 5.99 (*ddd, J* = 15.0 Hz, *J* = 4.5 Hz, *J* = 4.5 Hz, 1H), 7.17 (*d, J* = 9.0 Hz, 2H), 7.94 (*d, J* = 8.5 Hz, 2H), 7.95 (*d, J* = 9.5 Hz, 2H), 8.15 (*d, J* = 8.5 Hz, 2H).

^{13}C-NMR (125 MHz, DMSO-d^{6}) δ [ppm] = 55.77, 60.58, 65.01, 114.79, 122.44, 122.74, 125.09, 130.53, 130.96, 135.71, 146.23, 154.65, 162.70, 164.92.

REFERENCES

REFERENCES

1. Gill, S. C.; Vonhippel, P. H., "Calculation of Protein Extinction Coefficients from Amino-Acid Sequence Data" *Anal Biochem* **1989**, *182*, 319.

2. Rurack, K.; Spieles, M., "Fluorescence Quantum Yields of a Series of Red and Near-Infrared Dyes Emitting at 600-1000 nm" *Anal Chem* **2011**, *83*, 1232.

3. Chu, J.; Haynes, R. D.; Corbel, S. Y.; Li, P. P.; Gonzalez-Gonzalez, E.; Burg, J. S.; Ataie, N. J.; Lam, A. J.; Cranfill, P. J.; Baird, M. A.; Davidson, M. W.; Ng, H. L.; Garcia, K. C.; Contag, C. H.; Shen, K.; Blau, H. M.; Lin, M. Z., "Non-invasive intravital imaging of cellular differentiation with a bright red-excitable fluorescent protein" *Nat Methods* **2014**, *11*, 572.

4. Shcherbo, D.; Murphy, C. S.; Ermakova, G. V.; Solovieva, E. A.; Chepurnykh, T. V.; Shcheglov, A. S.; Verkhusha, V. V.; Pletnev, V. Z.; Hazelwood, K. L.; Roche, P. M.; Lukyanov, S.; Zaraisky, A. G.; Davidson, M. W.; Chudakov, D. M., "Far-red fluorescent tags for protein imaging in living tissues" *Biochem J* **2009**, *418*, 567.

5. Hori, Y.; Norinobu, T.; Sato, M.; Arita, K.; Shirakawa, M.; Kikuchi, K., "Development of Fluorogenic Probes for Quick No-Wash Live-Cell Imaging of Intracellular Proteins" *J Am Chem Soc* **2013**, *135*, 12360.

6. Liu, T. K.; Hsieh, P. Y.; Zhuang, Y. D.; Hsia, C. Y.; Huang, C. L.; Lai, H. P.; Lin, H. S.; Chen, I. C.; Hsu, H. Y.; Tan, K. T., "A Rapid SNAP-Tag Fluorogenic Probe Based on an Environment-Sensitive Fluorophore for No-Wash Live Cell Imaging" *Acs Chem Biol* **2014**, *9*, 2359.

7. Vonderheit, A.; Helenius, A., "Rab7 associates with early endosomes to mediate sorting and transport of semliki forest virus to late endosomes" *Plos Biol* **2005**, *3*, 1225.

8. Campbell, R. E.; Tour, O.; Palmer, A. E.; Steinbach, P. A.; Baird, G. S.; Zacharias, D. A.; Tsien, R. Y., "A monomeric red fluorescent protein" *P Natl Acad Sci USA* **2002**, *99*, 7877.

9. Fischer, M.; Haase, I.; Simmeth, E.; Gerisch, G.; Muller-Taubenberger, A., "A brilliant monomeric red fluorescent protein to visualize cytoskeleton dynamics in Dicyostelium" *Febs Lett* **2004**, *577*, 227.

10. Cole, N. B.; Smith, C. L.; Sciaky, N.; Terasaki, M.; Edidin, M.; LippincottSchwartz, J., "Diffusional mobility of Golgi proteins in membranes of living cells" *Science* **1996**, *273*, 797.

11. Cogan, U.; Kopelman, M.; Mokady, S.; Shinitzky, M., "Binding Affinities of Retinol and Related Compounds to Retinol Binding-Proteins" *Eur J Biochem* **1976**, *65*, 71.

12. Wang, L. C.; Li, Y.; Yan, H. G., "Structure-function relationships of cellular retinoic acid-binding proteins - Quantitative analysis of the ligand binding properties of the wild-type proteins and site-directed mutants" *J Biol Chem* **1997**, *272*, 1541.

13. Otwinowski, Z.; Minor, W., "Processing of X-ray diffraction data collected in oscillation mode" *Method Enzymol* **1997**, *276*, 307.

14. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., "PHENIX: a comprehensive Python-based system for macromolecular structure solution" *Acta Crystallogr D* **2010**, *66*, 213.

15. Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A., "REFMAC5 for the refinement of macromolecular crystal structures" *Acta Crystallogr D* **2011**, *67*, 355.

16. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., "Overview of the CCP4 suite and current developments" *Acta Crystallogr D* **2011**, *67*, 235.

17. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., "Features and development of Coot" *Acta Crystallogr D* **2010**, *66*, 486.

18. Meguellati, K.; Spichty, M.; Ladame, S., "Reversible Synthesis and Characterization of Dynamic Imino Analogues of Trimethine and Pentamethine Cyanine Dyes" *Org Lett* **2009**, *11*, 1123.

19. Sovdat, T.; Bassolino, G.; Liebel, M.; Schnedermann, C.; Fletcher, S. P.; Kukura, P., "Backbone Modification of Retinal Induces Protein-like Excited State Dynamics in Solution" *J Am Chem Soc* **2012**, *134*, 8318.

20. Wang, J.; Cao, W. F.; Su, J. H.; Tian, H.; Huang, Y. H.; Sun, Z. R., "Syntheses and nonlinear absorption of novel unsymmetrical cyanines" *Dyes Pigments* **2003**, *57*, 171.

21. Kurth, M. J.; Beard, R. L., "Beta-Hydroxy Esters - (E)-Versus (Z)-Enolate Geometry in Dianionic Claisen Rearrangements" *J Org Chem* **1988**, *53*, 4085.

22. Laplace, D. R.; Verbraeken, B.; Van Hecke, K.; Winne, J. M., "Total Synthesis of (plus /-)-Frondosin B and (+/-)-5-epi-Liphagal by Using a Concise (4+3) Cycloaddition Approach" *Chem-Eur J* **2014**, *20*, 253.

23. Spijkerassink, M. B.; Winkel, C.; Baldwin, G. S.; Lugtenburg, J., "5-Demethylretinal and Its 5-H-2, 7-H-2 and 5,7-(H-2)-H-2 Isotopomers - Synthesis, Photochemistry and Spectroscopy" *Recl Trav Chim Pay B* **1988**, *107*, 125.

24. Negishi, E. I.; Chatterjee, S., "Organoboron Chemistry .54. Highly Regioselective Generation of Thermodynamic Enolates and Their Direct Characterization by Nmr" *Tetrahedron Lett* **1983**, *24*, 1341.

25. Lin, M. Y.; Das, A.; Liu, R. S., "Metal-catalyzed cycloisomerization of enyne functionalities via a 1,3-alkylidene migration" *J Am Chem Soc* **2006**, *128*, 9340.

26. Klahn, P.; Duschek, A.; Liebert, C.; Kirsch, S. F., "Total Synthesis of (+)-Cyperolone" *Org Lett* **2012**, *14*, 1250.

27. Lifchits, O.; Mahlau, M.; Reisinger, C. M.; Lee, A.; Fares, C.; Polyak, I.; Gopakumar, G.; Thiel, W.; List, B., "The Cinchona Primary Amine-Catalyzed Asymmetric Epoxidation and Hydroperoxidation of alpha,beta-Unsaturated Carbonyl Compounds with Hydrogen Peroxide" *J Am Chem Soc* **2013**, *135*, 6677.

28. 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 Edit* **2010**, *49*, 1790.

29. Davey, M. H.; Lee, V. Y.; Miller, R. D.; Marks, T. J., "Synthesis of aryl nitroso derivatives by tert-butyl hypochlorite oxidation in homogeneous media. Intermediates for the preparation of high-hyperpolarizability chromophore skeletons" *J Org Chem* **1999**, *64*, 4976.