ELUCIDATION OF ILBP FAMILY FOLDING PATHWAY AND STUDY OF REENGINEERING THEM AS FLUORESCENT PROTEIN TAGS VIA STRUCTURAL ANALYSIS

By

Zahra Assar

A DISSERTATION

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

Chemistry - Doctor of Philosophy

2018

ABSTRACT

ELUCIDATION OF ILBP FAMILY FOLDING PATHWAY AND STUDY OF REENGINEERING THEM AS FLUORESCENT PROTEIN TAGS VIA STRUCTURAL ANALYSIS

By

Zahra Assar

The intracellular lipid binding proteins (iLBP) family are found in the cells of mammals, birds, fish, amphibians and reptiles. They function to shuttle large insoluble hydrophobic molecules, including retinal, various long chain fatty acids and etc. throughout the cell around the cytosol and nucleus. The combination of their small size and relatively large binding pocket make them suitable templates in a variety of protein design applications, including the study of an innovative class of fluorescent proteins. To pursue our goals, we used human Cellular Retinol Binding Protein II (hCRBPII). We were the first group to achieve the structure of an *all-trans*-retinal and the first bonafide structure of retinol-bound hCRBPII.

In the course of these studies, we have discovered hCRBPII is surprisingly capable of folding as domain swapped dimer (DS), with single mutations able to shift the folding product from monomer to dimer. Structural analysis of both wild type and multiple mutant DS dimers led us to remarkable hypotheses regarding mechanism of this phenomenon, which is different from the previous studies on this family. We proposed that the N-terminal and C-terminal halves of hCRBPII are capable of at least partially folding independently, to form "open monomer." The dimer/monomer ratio depends on the relative rates of dimerization of the open monomers, versus closing of the two halves together to form the "closed monomer". In addition, by comparing structures of holo hCRBPII DS dimer variants, we identified an extremely large change in the relative orientation of the two domains upon ligand binding in dimers. This

suggests the possibility that iLBP domain swapped dimers could be allosterically regulated forms of these proteins, at least in some cases. Fatty acid binding protein 5 (FABP5), another member of iLBPs, has also been reported to forms a very similar DS dimer, which makes it likely that other family members could also form DS dimers, and have physiological importance for some members of the family.

As mentioned, a new class of fluorogenic proteins was created by binding fluorophore aldehydes in the binding pocket of hCRBPII via protonated Schiff base (PSB) formation. In this new system, emission of the designed solvatochromic fluorophore is flexible based on the polarity of the environment; therefore multicolor probes can be developed. More importantly, absorption/emission wavelengths can be tuned therefor; nonspecific labeling and background fluorescence can be reduced. By now, the absorption maxima are tuned from 501nm to 705nm and emission maxima from 613 nm to 744 nm. Covering both the red and far-red fluorescence wavelength regimes.

This Dissertation is lovingly dedicated to my grandmother, Maman joon.

ACKNOWLEDGMENTS

I am very much grateful to my PhD advisor, Professor James H. Geiger, for his support and guidance throughout my graduate work. He taught me how to think and design strategies as a scientist. He also trained me to think independently and to be more patient in science because “Good things do not come easy. The roak is lined with pitfalls”. *Desi Arnaz.*

Jim always supported me and my thoughts about my projects. Not only he is a great advisor, also he is very friendly approachable person. I am so glad and honored to have him as my advisor. Thank you Jim, for letting me to grow, even start my professional career before I defend and thank you for everything.

I would like to thank my second reader and committee member, Professor Babak Borhan. I still remember August 2012 when I first came to US, I was not sure about my graduate research field but Babak guided me very well. Also during biomeetings, He always pointed important discussions and suggested critical experiments which they all helped me to understand my research better. Whenever I needed help he was always available for me. Thank you Babak for everything.

I would like to appreciate Dr. William Henry for being on my committee and reading my dissertation. I received some great advice from him regarding my research. In addition, thank you to Professor Gary Blanchard for also reading my dissertation.

I would like to thank Dr. Zahra (Rafida) Nossoni and Dr. Wenjing Wang for opening a path for me to pursue my PhD. I am happy that I had Rafida right when I joined the Geiger lab. She

helped me in my first year a lot and she was always available especially when I needed help in data collection. She is a great sister-in-law and like a sister that I have never had.

I also would like to thank Dr. Chrysoula Vasileiou for her help. She always was available and helpful to me with her ideas and great discussions in our biomeetings. Thank You Chrysoula.

I am heartily thankful to our awesome group members, all current and former members in Geiger's group for being friendly and helpful to me. Specifically, thank you Dr. Remie Fawaz, Dr. Camille Watson and Dr. Meisam Nosrati, I still remember my first Argonne trip and how much Meisam helped me during my first time data collection. Thank you Hadi Nayebi and Alireza Ghanbarpour for being great lab members, I always enjoyed our scientific discussions and chats. Also thank you Nona Ehyaei for always cheering me up with your energy and helping me whenever I could not drive to East Lansing. Thank you Lindsey Gilbert and Janice Chiou, Kevin Kramer and Colin McCornack. I would like to thank the bio people in Professor Borhan's group Dr. Tetyana Berbasova, Dr. Ipek Yapici, Wei Sheng and especially Dr. Elizabeth Santos. Liz was the best roommate to live and great scientist to work with.

I would like to thank LS-CAT staff at the Argonne National Laboratory for all of their help during data collection. In addition, I am thankful to my new scientific family of Cayman Chemical Company, Dr. Adam Stein and Melissa Holt. I am very happy to work with them as a member of structural biology team.

I also appreciate my dear friends who supported me on this path.

Finally, I am grateful to my beloved family. My parents, brother and my husband Dr. Farid Nossoni, without them, their love and support, it would not be possible for me to finish this stage.

Farid always taught me to be positive. I am thankful to have him in my life, for inspiring me and supporting me to pursue my path.

And again I would like to say thank you to all of the amazing people who supported and helped me.

‘

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO SYMBOLS AND ABBREVIATIONS

Å	Angstrom
PDB ID	Protein Data Bank Identifier
PSB	Protonated Schiff Base
SB	Schiff Base
iLBP	intracellular Lipid Binding Proteins
hCRBPII	human Cellular Retinol Binding ProteinII
hCRABPII	human Cellular Retinoic Acid Binding ProteinII
hFABP4	human Fatty Acid Binding Protein4
hFABP5	human Fatty Acid Binding Protein5
σ	Sigma
UV	Ultra Violet
WT	Wild-Type
IPTG	Isopropyl &-D-1-thiogalactopyranoside
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
Vis	Visible
Kd	Dissociation constant
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
KDa	Kilo Dalton

E. Coli's	Escherichia coli
min	minute
S	Second
ε	Extinction coefficient
QY	Quantum Yield
λ_{max}	Maximum wavelength
λ_{ex}	Excitation wavelength
λ_{em}	Emission Wavelength
mM	Milimolar
µM	Micromolar
nM	Nanomolar
mol	mole
mmol	millimole
mg	milligram
ml	milliliter
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
rpm	rotation per minute
°C	degrees of centigrade
K	degrees of kelvin
pH	Logarithmic scale of hydrogen ion activity
NaOH	Sodium hydroxide

Equiv	equivalent
RMSD	root mean square deviation
Ala, A	Alanine
Arg, R	Argannine
Asn, N	Aspargine
Asp, D	Aspartate
Cys, C	Cysteine
Gln, Q	Glutamine
Glu, E	Glutamate
His, H	Histitidine
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

DS	Domain Swapping
LB	Luria Broth
R-factor	reliability factor
RT	Room temperature
RMSD	root-mean-square deviation

Chapter I: Structural Studies of Human Cellular Retinol Binding Protein II (hCRBPII) bound to Retinol and Retinal.

I-1 Introduction

Vitamin A (retinol) is an essential micronutrient that plays a key role in vision, cell growth, differentiation metabolism etc. [5] Because of its low solubility in aqueous medium, retinol is bound by specific binding proteins in body fluids and within the cell for its transport, bioavailability and stability. Retinol is converted to retinal, which is the essential molecule of vision, and also to retinoic acid, which is essential in cell differentiation. [6]

In different tissues, different binding proteins play important roles in transport, storage and metabolism of vitamin A, such as Retinol Binding Proteins (RBP), Cellular Retinoic Acid Binding Proteins (CRABP), Cellular Retinol Binding Protein (CRBP), Cellular Retinaldehyde Binding Protein (CRALBP) and Interphotoreceptor Retinol Binding Protein (IRBP). In all those proteins, the hydrophobic ligand and protein have non-covalent interactions.[7, 8]

CRBPs belong to the family of intracellular lipid binding proteins (iLBP's). iLBPs are a sub-family of the large family of calycins that, in addition to the iLBP's, include the avidins and lipocallins. All the Calycins have a beta-sandwich structure in common and all are chaperones for their hydrophobic ligands, which bind in their relatively large binding cavities, though sequence homology between the subfamilies is less than 10%. [6, 9-13] ILBPs are relatively small proteins (126-140 amino acid) found in the cytosols of fish, amphibians, reptiles, birds and mammals. They responsible for shuttling insoluble hydrophobic molecules, including retinal, retinoic acid, various long chain fatty acids, eicosinoids, cholesterol and hemes, throughout the

cell.[6, 10-13] Their structure consists of a ten-stranded β-barrel, and two short α-helices that cover the binding pocket, with the ligand buried deeply within the binding site. Basically, two short α helixes cover the binding pocket like a lid, isolating the ligand from bulk solution and the hydrophobic ligand is deeply buried inside the binding pocket. [14]**(Figure I-1a)** This family also includes the Cellular Retinoic Acid Binding Proteins (CRABP), and the Fatty Acid Binding Proteins (FABP) with quite similar structures.[8]

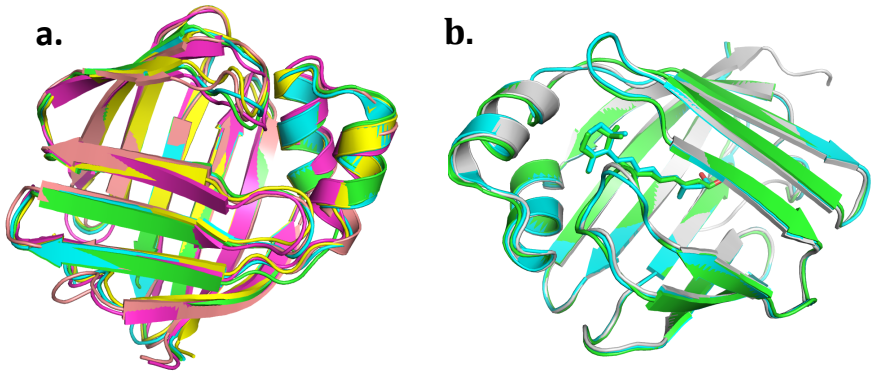


Figure I-1: a. An overlay of the structures of hCRBPII (green and cyan), hCRBPIII (1GGL, salmon), hCRBPIV (1LPJ, yellow) and hCRBPI (1KGL, pink) **b.** An overlay of the structures of apo hCRBPII (2RCQ, gray), wt hCRBPII bound to all-*trans*-retinal (4QYP, green) and to all-*trans*-retinol (4QYN, cyan).

There are four identified isoforms for CRBP (CRBPI, II, III and IV), among them only CRBPI and II have been shown to be involved in the metabolism and transport of intracellular retinol so far. HCRBPI is found predominantly in the liver and kidney while hCRBPII is more prevalent in the small intestine. [5, 15, 16]It was reported that hCRBPII regulates retinoid metabolism in intestine and facilitates the reduction of retinal to retinol and the subsequent esterification of retinal to retinyl ester. HCRBPIII mRNA is expressed in kidney and liver thus suggesting an impotent role as an intracellular mediator of retinol metabolism in these tissues. HCRBPIV

mRNA is expressed primarily in heart, and transverse colon. However, physiological roles of newly identified hCRBPIII and IV are still unknown.[8, 17]

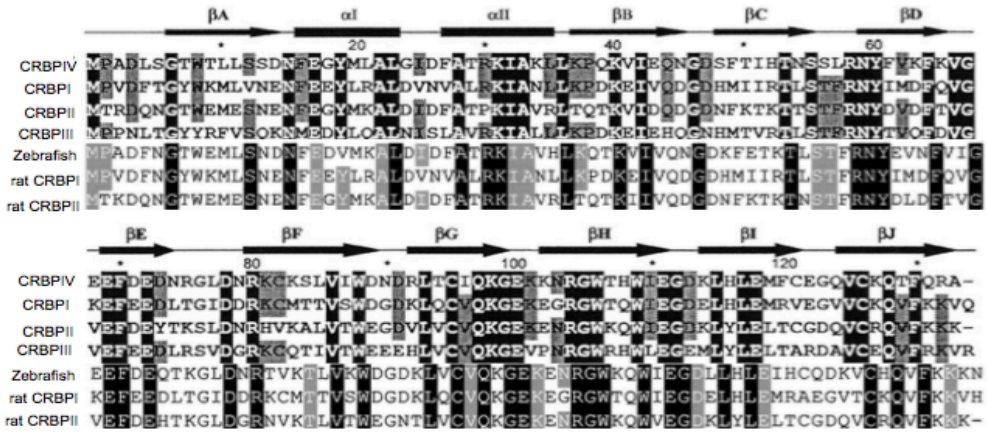


Figure I-2: The multiple sequence alignment between human, rat CRBPs and zebra fish CRBP.

CRBPI and II share 56% sequence identity and both proteins can bind all-*trans*-retinol, all *trans*-retinal and 13-*cis*-retinol, while neither bind 9-*cis*-retinol though CRBPII is found to be the more selective for all-*trans*-retinoids. [18-20] **(Figure I-2)** As the first crystal structures of CRBP, we can point to the apo and all-*trans*-retinol bound structure of rat CRBPII, which was published in 1993.[21-23] Comparison of the two structures showed that no significant differences were observed upon ligand binding. Later on, the crystal structures of human CRBPII and a zebra fish CRBP, both bound with all-*trans*-retinol, were also reported. [21, 22]However, in both cases the electron density of the end of the chromophore was not clear. The ligand was modeled as a mixture of retinol, and a retinol derivative whereas the C13-C14 bond is a single bond. (2RCT, 1KQW)

In our lab and in collaboration with prof. Borhan's group, we were the first group to achieve the structure of an all-*trans*-retinal bound CRBP (hCRBPII) by Dr. Nossoni. We also

were the first to show that a retinoid derivative virtually identical to that seen in previously published CRBPII structures can be generated by X-ray radiation damage, and this process is dependent on the wavelength of the X-ray radiation. In addition, we finally determined the first bonafide structure of retinol-bound hCRBPII.[6] **(Figure I-1b)**

I-2 The affinity of hCRBPII for retinol and retinal

The affinity of rat CRBPII for both retinol and retinal has been measured using both a fluorescence quenching assay and Fluorine NMR. The most recent study showed rat CRBPII binds to retinol and retinal with similar affinity (a K_d of about 10 nM).[20, 24, 25] As far as we knew, no affinity measurements for hCRBPII had been reported, and then, we decided to measure its affinity to both retinal and retinol using a Tryptophan fluorescence-quenching assay (Described in Experimental Section). While the binding of retinal was significantly weaker compared with that most recently reported for the rat protein, 130±10 nM VS 10nM in rat CRBPII, **Figure I-3d**)[24], hCRBPII showed significantly higher affinity for retinol, with a lower limit of 2.2±7.9 nM (**Figure I-3a**). The binding constant appeared to be too high for accurate measurement using our fluorescence assay. This seems to indicate a potential difference in the interaction of rat versus human CRBP for these ligands. The ligand is stabilized within binding pocket, allowing it to preserve its yellow colour at room temperature and in light for more than three days, while in ethanol solution a quarter of the chromophore is degraded after 4 hours when exposed to air.

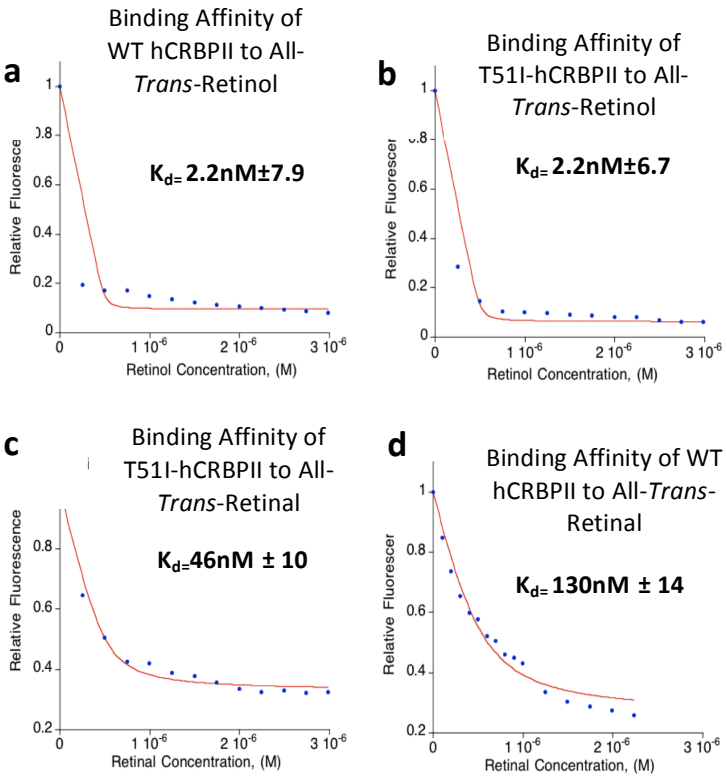


Figure I-3: Ligand binding measurements using the fluorescence quenching of tryptophan (excitation at 280 nm, emission at 350 nm). **a.** Titration of WT-hCRBPII with all-*trans*-retinol. **b.** Titration of hCRBPII T51I mutant with all-*trans*-retinol. **c.** Titration of hCRBPII T51I mutant with all-*trans*-retinal. **d.** Titration of WT-hCRBPII with all-*trans*-retinal.

I-3 Structure of the retinal-bound hCRBPII complex

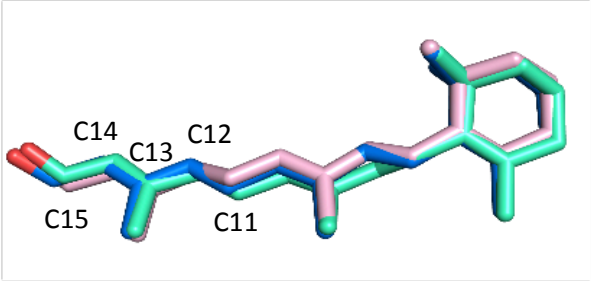
Retinal-bound hCRBPII crystallizes in the P1 space group with four molecules per asymmetric unit (Described in experimental section, **Table I-1**).[(6)] Electron density corresponding to the entire ligand was very clear in only one of the four chains (chain B). In chain C, density for retinal is well defined for all atoms but C15 and the hydroxyl group. In molecule A, electron density is relatively well-defined for the carbonyl group and C11-C15, but is weaker for the rest of the molecule. Scattered, uninterruptable electron density is seen in

molecule D. The overall structure of the protein is little changed either in the apo state, or when bound to either retinal or retinol. **(Figure I-1a, 1b)** The retinal is essentially identical in all three molecules where the electron density is clear **(Figure I-4a)**.

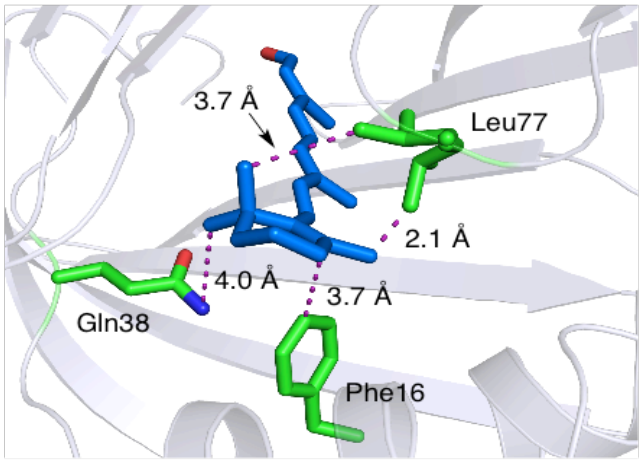
The β-ionone ring adopts the 6-s-*trans* conformation in both ligands. Two hydrophobic residues, Phe16 and Leu77, surround the ionone ring of the ligand, locking the chromophore in position **(Figure I-4b, 4c).** The distances between the C5 methyl group of retinal and Leu77 and Phe16 are about 3.7Å and 3.8Å respectively. (**Figure I-4b)**

Also, the carbonyl group of retinal makes a water mediated interaction with Gln108 and Lys40. The side chain of Gln108 is fixed in position by a small water network, including the main chain carbonyls of Thr1 and Asp91, and the side chains of Gln4, and Gln108 (**Figure I-4a**). This network is unique to hCRBPII and is not seen in zebra fish CRBP, rat CRBPII and hCRBPI (PDB entry 1KGL), hCRBPIII (PDB entry 1GGL) and hCRBPIV (PDB entry 1LPJ). *(14)*

a.



b.



c.

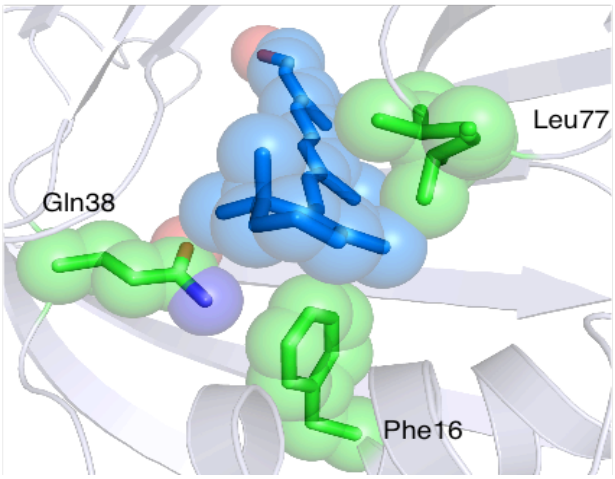


Figure I-4: Retinal binding in hCRBPII. **a.** An overlay of retinal from mol A (cyan), mol B (blue) and Mol C (pink), showing the similarity in binding in the three. **b**. Stick representation of the interaction between Phe16, Leu77 (both with green carbons, N blue and O, red) and the retinal ionone ring (blue carbons, Mol B). **c.** Space filling depiction of b.

In this case, water must act as hydrogen bond donor to the two main-chain carbonyl groups (Thr1 and Asp91), only the lone pairs on the oxygen of water are left to have hydrogen bonding to the amide nitrogen of Gln4, defining its orientation. Therefore it is the carbonyl oxygen of Gln4 that makes a hydrogen bond, necessarily with the amide of Gln108, defining its orientation as well. The orientation of Gln108, with the carbonyl group pointing toward the carbonyl of the retinal, abolishing the direct hydrogen bond to the retinal carbonyl oxygen and necessitates the water-mediated hydrogen bond. The hydrogen bond between this water and the retinal carbonyl is 2.4Å, which is an ideal distance for a strong low-barrier hydrogen bond (**Figure I-5 and Figure I-9c**). This water makes an additional hydrogen bond with the ε-amino group of Lys40, and is fixed in space by two hydrogen bonds in the binding pocket.

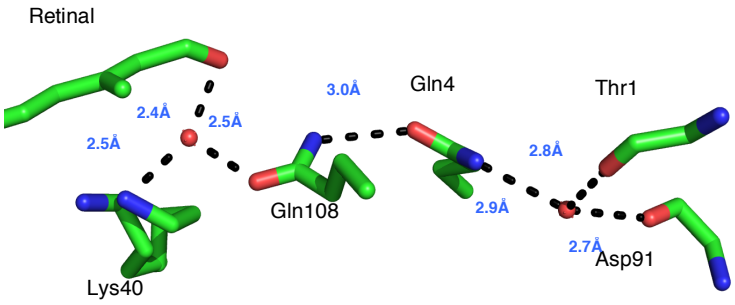


Figure I-5: The water mediated interaction of the retinal hydroxyl group in hCRBPII. Atoms colored by type, hydrogen bonded distances are shown.

I-4 Wavelength-dependent damage of the retinol in hCRBPII

As previously mentioned, the structures of the CRBPII-retinol complex from rat, human and zebra fish have already been determined. But the presence of the retinoid derivative that

also occupies the binding site in human and zebra fish structures made our interpretation complicated. [21, 22]

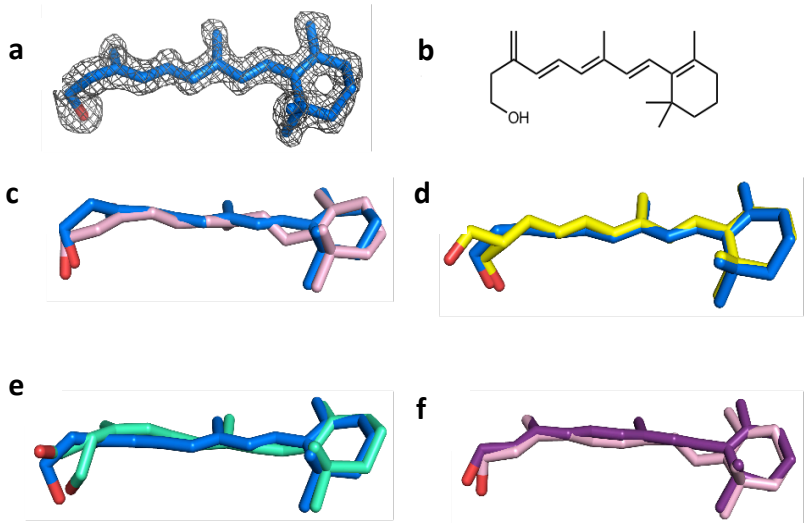


Figure I-6: **a**. The contoured 2Fo-Fc electron density map at 1.0 σ for rearranged retinol from data collected at 11 KeV/the highest flux (4QYN, blue). **b.** The chemical structure of a possible retinoid derivative consistent with the crystallographic data. **c.** The overlay of bonafide retinol structure (chain C) from data collected with an attenuated X-ray beam at an energy of 7 KeV (4QZT, Pink) and the structure of rearraged retinoid derivative (chain B) from data collected at 11 KeV and the highest flux (4QYN, blue). **d.** The structures of the noncanonical retinoid (refined in two conformations) previously seen in the human CRBPII (2RCT, yellow), and our current rearranged structure (4QYN, blue) were overlaid. **e.** The structures of the canonical and noncanonical retinoid previously seen in the zebrafish CRBP (1KQW, green), and our current rearranged structure (4QYN, blue) were overlaid. **f.** The overlaid structures of the canonical retinol obtained in the rat CRBPII (1OPB, purple) and our current bonafide all-trans-retinol structure in human CRBPII (4QZT, pink).

This other retinoid species, postulated to be a degradation product from the original retinol, is characterized by a torsion angle about the C13-C14 bond that is around 60°, inconsistent with the *trans* double bond between these two carbons expected for retinol. **(Figure I-6b, 9a and 9b)** The hydroxyl group of this species makes hydrogen bonds to both

Gln108 and Lys40, essentially occupying the same position as the water molecule responsible for the water mediated interaction seen in our hCRBPII-retinal complex. In rat CRBPII-retinol complex shows only retinol in the binding pocket, with the hydroxyl group in a distinctly different position to that seen in the other structures, allowing it to make hydrogen bonds to both the amide oxygen and nitrogen atoms of Gln108. **(Figure I-7c and 9c)** Since we were not confident on the previous hCRBPII-retinol complexes, the structure of the hCRBPII-retinol complex was re-determined in a new P1 crystal form with 2 molecules per asymmetric unit in our group. (Table II-1)

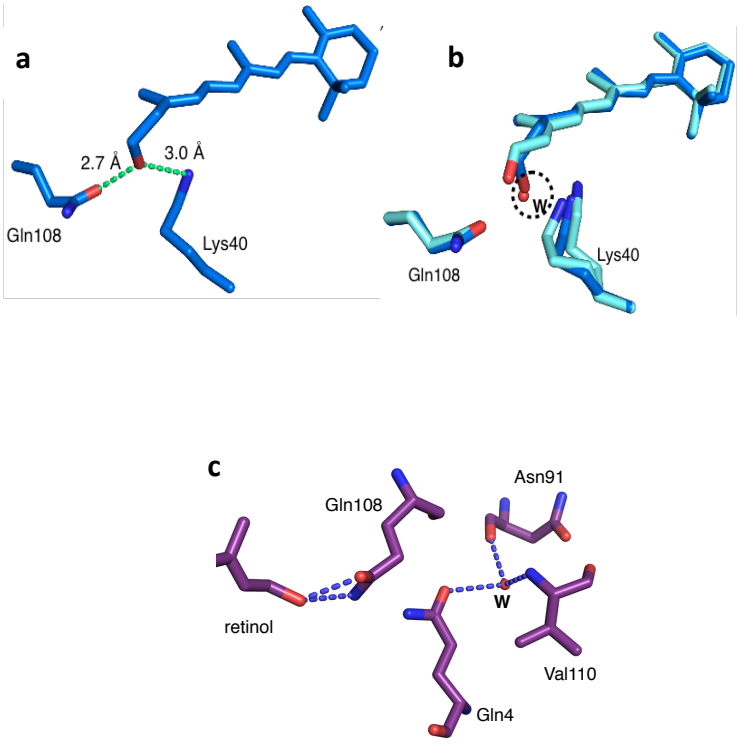


Figure I-7: **a.** The rearranged retinol from data collected at 11 KeV and the highest flux (4QYN, blue) and the interaction of alcohol moiety with the neighboring residues of Gln108 and Lys40. **b.** The structures of the retinal bound hCRBPII (4QYP, cyan, chain B), and our current rearranged retinol structure (4QYN, blue) were overlaid. The ordered water molecule (**W**) seen in all-*trans*-retinal crystal structure, is occupied by the hydroxyl group of the noncanonical retinoid (indicated in dashed circle). Similar case in retinol structure from data collected at 7 KeV and an attenuated beam has been seen. **c**. Interactions around Gln108 of retinol bound rat CRBPII (1OPB, purple).

Interestingly, The same retinoid derivative was seen in our structure. **(Figure I-6a, 6d and 6e)** The hydroxyl of the retinoid derivative makes hydrogen bonds to both Gln108 and Lys40, and occupies a position that is essentially identical to that of the water in the retinal complex **(Figure I-7).** No evidence for retinol is seen in our structures. Though the exact identity of the retinoid is unclear (**Figure I-6b).** NMR spectrum of the retinol sample was taken by Dr. Yapici to determine the purity of the retinol. The NMR showed the retinol to be quite pure with

no significant extraneous peaks. Further, Dr. Yapici performed a series of analytical HPLC of the sample, which they also showed only a single peak consistent with retinol. **(Figure I-8).** Since our crystals were all grown at a relatively low pH of 4.6, we investigated the possibility that the low pH of the crystallization condition could lead to the retinoid derivative by incubating retinol, both alone and in the presence of hCRBPII overnight. In both cases a decomposition product was identified by HPLC. However, this product was less polar as evident from the HPLC retention times, and is inconsistent with a hydroxyl-group containing compound. Then we considered that the low pH product was a dehydration product, similar to that previously seen, but was not the compound seen in the structure **(Figure I-a, 6b).**[(26)] We then grew over 30 crystals of the complex, and Dr. Yapici dissolved them in phosphate buffered saline extracted the retinoids by hexane and analysed them by HPLC and UV/Vis spectroscopy.

all of the other data sets, the human, zebrafish, and both of our hCRBPII complex structures, were collected at synchrotron sources using about 1Å (11 KeV) wavelength radiation. **(Figure I-6, 9)**[(23)]

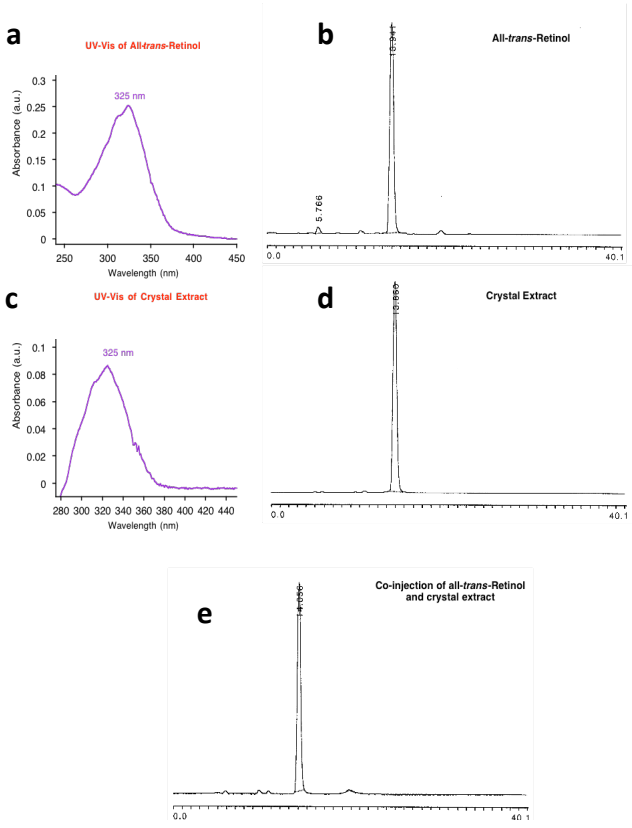


Figure I-8: This experiment performed by Dr. Yapici **a.** UV-vis spectrum of all-*trans*-retinol. **b.** The HPLC trace of all-*trans*-retinol. **c.** UV-vis spectrum of the retinol extracted from the crystals and purified by HPLC. **d.** HPLC chromatogram of the retinol extracted from crystals. **e.** HPLC chromatogram of bonafide all-*trans*-retinol and the retinol extracted from crystals. Showing that they are the same species.

This showed only a single peak that co-eluted precisely with retinol, indicating that the crystals do in fact contain only retinol. (**Figure I-8**) It therefore appears that the retinoid derivative seen in these structures is due to X-ray induced damage during data collection. Consistent with this possibility, It was observed that the yellow colour of the crystal changed to dark orange after its exposure to the X-ray beam and this colour is maintained to the end of

data collection.[27-29] It is interesting to note that for the rat CRBPII retinol complex, the only bonafide retinol-bound CRBPII structure, data was collected using a Cu *K*α home X-ray source,

In an effort to finally produce a bonafide hCRBPII retinol complex, we collected a variety of data sets at the synchrotron. Since most X-ray damage of protein crystals are thought to be dosage related. We first collected data at 11 KeV but with the beam attenuated to less than 20% intensity. The electron density was still consistent with only the rearranged retinol in the active site (data not shown). Then we investigated the effect of X-ray energy by collecting data at 7 KeV, 8 KeV and 9 KeV, all using a beam attenuated to 20% of the original, and an additional data set at 11 KeV. Dr Spencer Anderson from Argonne National Laboratory helped us in collecting the attenuated data at multiple wavelengths. Surprisingly, the retinoid electron density in the 7 KeV data set clearly showed an un-rearranged all-*trans*-retinol **(Figure I-9c).** All the other data sets showed some mixture of probably retinol and the rearranged product. Together, these data show that the rearranged product is a result of both dosage and wavelength-dependent X-ray damage, and that data collection at lower energy results in the first structure of the bonafide hCRBPII-retinol complex. **(Figure I-9d, 10, 1b)**[6]

I-5 The structural difference between retinol-bound hCRBPII and retinal-bound.

The difference in binding of retinal versus retinol lies in the orientation of the hydroxyl group (**Figure I-10c, 10d).** In the hCRBPII retinol complex the hydroxyl is positioned to make hydrogen bonds with Lys40 and Gln108, making hydrogen bonds with both carbonyl oxygen and nitrogen atoms of Gln108 simultaneously (an interaction that would be impossible for retinal, which cannot act as hydrogen bond donor to the amide carbonyl oxygen). This results in

a different orientation around C15, making room for the water molecule between Gln108 and Lys40 and the creation of the hydrogen bond to this water molecule.

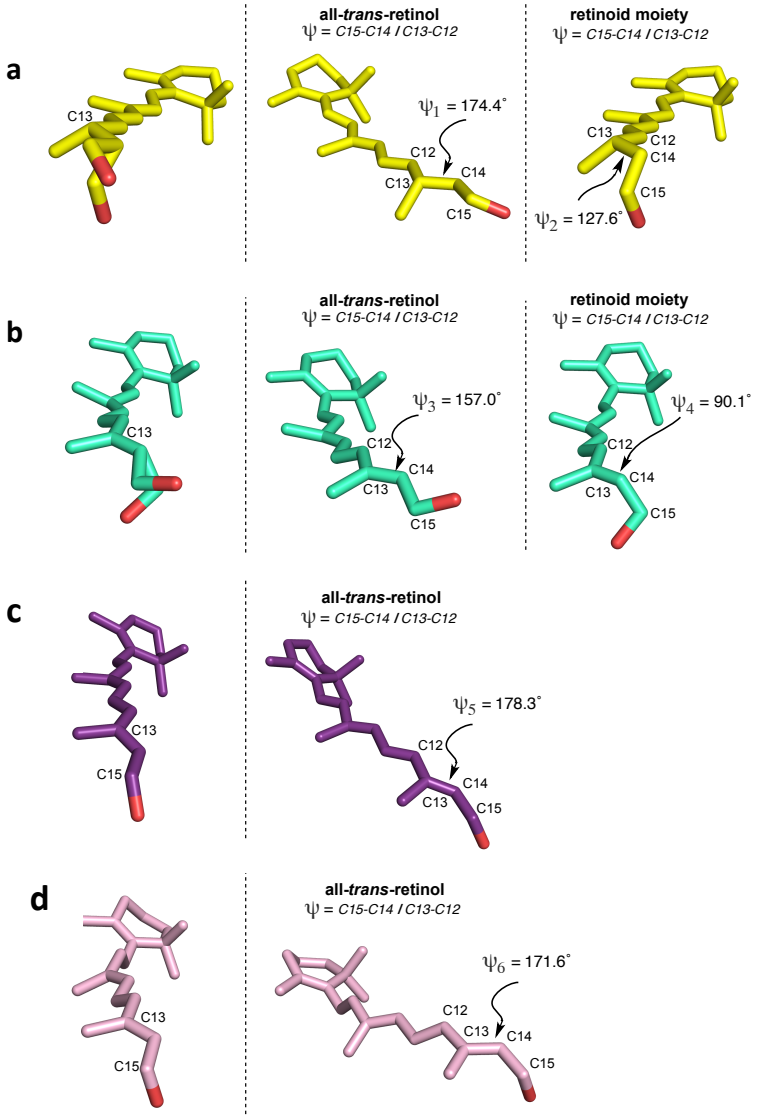


Figure I-9: A comparison of the retinol and retinoid moieties found in the crystal structures of human, zebra fish and rat CRBPII. **a.** The retinol (left) and retinoid (right) moieties found in the previously published hCRBPII (2RCT, yellow). The value of 128° torsion angle (ψ) about the C13-C14 bond is inconsistent with a double bond (180°). **b.** The retinol and retinoid moieties seen in the crystal structure of zebrafish CRBP (1KQW, green). The torsion angle about the C13-C14 bond is also inconsistent with a double bond. **c.** The retinol structure in the rat CRBPII (1OPB, purple). **d.** The retinol found in the X-ray structure of hCRBPII using data collected at an energy of 7 KeV and an attenuated beam (4QZT, pink) with an approximate 180° torsion angle around C13-C14 bond.

By overlaying these two structures, it is clear that the position of the retinol hydroxyl group in hCRBPII would in fact clash with this water molecule. The difference in affinity for retinol versus retinal may be explained by their different interactions with hCRBPII. However, the physiological relevance of this difference in binding constant is unclear.

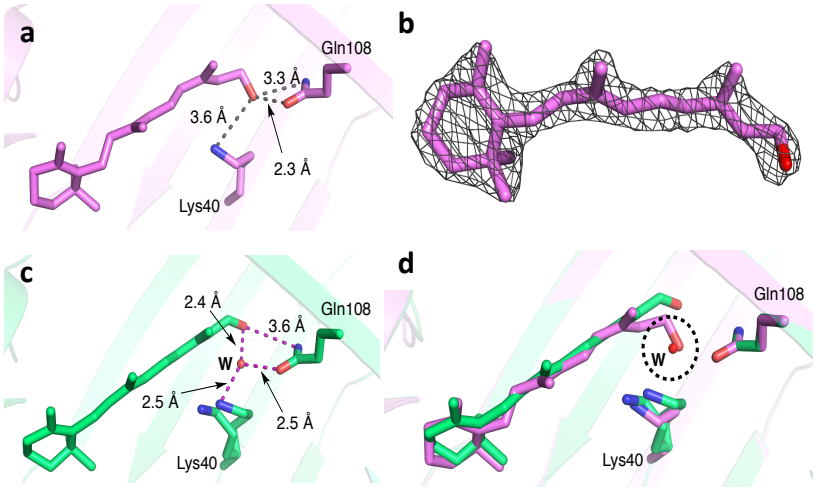


Figure I-10: Retinal versus retinol binding in hCRBPII. **a.** The structure of the retinol-bound WT hCRBPII**. b.** The 2Fo-Fc electron density map contoured at 1.0σ around retinol of 7Kev data. **c.** The structure of all-*trans*-retinal bound to WT hCRBPII. **d.** The structures of all-*trans*-retinal (green) and retinol (magenta). The position of an ordered water molecule (**W**) occupied by the hydroxyl group of retinol is indicated in dashed circle.

I-6 Ligand Binding in CRBPI versus CRBPII

In rat, CRBPI and II have similar binding affinities for all-*trans*-retinol. [(20)] An overlay of the crystal structures of retinal-bound and retinol-bound hCRBPII and the structure of hCRBPI (PDB entry 1KGL) and rat CRBPI (PDB entry 1CRB) illustrate that in hCRBPII the ligand is more than 1Å deeper in the binding pocket.[(14)] The identity of the amino acid at position 51 appears

to be the deciding factor for ligand position **(Figure I-11).** Position 51 is Isoleucine in CRBPI (both rat and human) while it is threonine in hCRBPII. By comparing the structures of human and rat CRBPI (PDB entry 1CRB)[30], It can be seen that retinal cannot bind hCRBPI in the same way that it does in hCRBPII. This is due to a steric clash between the side chain of isoleucine and the retinal carbonyl group.

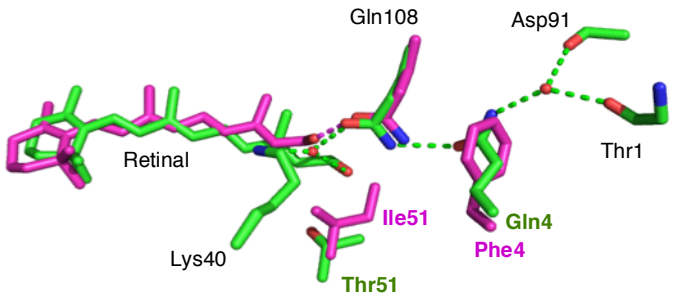


Figure I-11: The structures of retinol-bound rat CRBPI (magenta carbons and hydrogen bonds) and retinal-bound hCRBPII (green carbons and hydrogen-bonds) are overlayed. Amino acid labels are colored similarly when different in the two structures. Note the change in the water network introduced by Phe3.

The combination of the addition of the larger isoleucine at position 51, and the altered position of the ligand leave no room for a water molecule between Gln108 and Lys40. In both CRBPI structures a hydrogen bond is formed between Gln108 and the retinol hydroxyl group **(Figure I-11).** Though there is no structure of a CRBPI bound to retinal, it is clear that a hydrogen bond to Gln108 can only be made if the amide nitrogen is pointing toward the retinal carbonyl, because it must be the hydrogen donor.

This is opposite to the orientation of Gln108 in hCRBPII, where the orientation of Gln108 is determined by Gln4 and a water network as described previously. However, in CRBPI, there is

a phenylalanine at position 4 (Phe3 in hCRBPI) instead of glutamine, which may allow Gln108 to rotate its sidechain, such that the amide nitrogen is positioned to make a hydrogen bond with the aldehyde carbonyl, allowing retinal to bind CRBPI very similarly to that of retinol, resulting in the similar affinity seen for the two ligands in CRBPII.

In an attempt to understand the role of the residue at position 51, the hCRBPII-T51I mutant was produced and its binding affinity for retinol and retinal measured. The dissociation constants of this mutant for both all-*trans*-retinol and all-*trans*-retinal were similar to that of the wild type protein (**Figure I-3b, 3d**) indicating that position 51 is not responsible for the differences in affinity in the two proteins. Unfortunately attempts at crystallization of hCRBPII-T51I bound to all-*trans*-retinal were not successful. However, crystals of the retinol-bound CRBPII-T51V complex were obtained and its structure determined. However, this structure, collected with an intense X-ray beam with an energy of 11 KeV, also showed the rearranged retinol product in the active site, making it difficult to determine the effect of a hydrophobic residue at position 51 on bonafide retinol binding.

I-7 Conclusion

Based on the crystal structure of holo wt hCRBPII with all-*trans*-retinal, there are at least three distinct binding modes for retinoids in CRBPI and II. In the human and rat CRBPII retinol-bound structures, retinol is able to make hydrogen bonds with both the carbonyl and amide nitrogen of the side chain of Gln108. This mode of interaction is not feasible in retinal, due to the lack of a hydrogen bond donor, resulting in a conformational change of the carbonyl leading to a water-mediated interaction between the retinal carbonyl and Lys40 and Gln108. This is consistent with the higher binding affinity for retinol relative to retinal in hCRBPII. Since retinol

is able to make three hydrogen bonds to Gln108 and Lys40. Also, the presence of Ile51 causes a translation of the ligand in the CRBPI binding pocket, resulting in only a single hydrogen bond being made between Gln108 and the retinol hydroxyl group.

I-8 Experimental

I-8-1 Material and Method: Protein Expression and Purification

The hCRBPII gene was purchased from ATCC and cloned in pET17b vector, with using the NdeI and XhoI restriction enzyme in the N- and C terminus respectively. 1 µL of the resulting plasmid were transformed in 50 µL of *E.coli* DH5α (Novagen®) competent cells. The cells were incubated for 30 min in ice and heat shocked at 42 °C for 28 seconds, then 450 µL of Luria-Bertani broth (LB) was added and the cells were incubated at 37°C for an hour. The resulting mixture were spread on an LB agar plate with (ampicillin 100 µg/mL) and incubated at 37°C for 16 hours. A single colony was picked from the plate and inoculated in 5 mL of LB media contains 100 µg/mL ampicillin. The cell culture was grown over night (12-16 hours) at 37°C, then the media were centrifuged at 15000 rpm for 1 min. DNA extraction and isolation from the cell pellet was done according to the manufacturers instructions (Promega Wizard and SV Miniprep (A1330) DNA purification kit). The target gene was transformed into BL21(DE3) (Invitrogen**TM**) *E.coli* competent cells for protein expression. A single colony was picked and inoculated in 25 mL of LB, containing 100 µg/mL ampicillin at 37°C for 12-16 hours. This media was then transferred into 1L of LB media with ampicillin (100 mg/L) to grow large scale and incubated at 37°C until OD_{600} reached 0.5-1.0. The protein expression was induced with 1 mM isopropyl β-D-1 thiogalactopyranoside (IPTG) (purchased from Gold Biotechnology) overnight at 25°C. The

cells were harvested by centrifugation at 5000 rpm for 20 min. The harvested cells were resuspended in Tris buffer (10 mM Tris, 10 mM NaCl pH 8.0, 50mL). The suspended cells were lysed by sonication and the lysed cells were centrifuged at 4°C (14,000 rpm, 20 min).

The supernatant was loaded on a Fast Q anion exchange resin (purchased from GE healthcare), which was equilibrated with buffer A (10 mM Tris, 10 mM NaCl pH 8.0). Then after the binding, the resin was washed three times with 50 mL of buffer A. The protein was eluted with 40 mL of elution buffer (10 mM Tris, 100 mM NaCl pH 8.0). The eluted protein was desalted using a Centriprep® centrifugal filter (10kDa cutoff) at 2500 rpm. Protein was then purified on a 15Q anion exchange column using a BioLogic DuoFlow system for the second step of purification. The purity of the protein was determined with SDS-PAGE.

I-8-2 Kd Determination via Fluorescence Quenching Assay

The dissociation constant measurement, (Kd) for both all-*trans*-retinol and all-*trans*-retinal (purchased from Sigma) with hCRBPII was determined by fluorescence quenching assay. All samples were stored in glass containers, as many plastic containers will leach fluorescent impurities into the sample. Salinized glassware was used to avoid loss of protein and a change in protein concentration.

The cuvette was allowed to sit with 3 mL of a 0.01% gelatin containing PBS buffer (4 mM NaH2PO4, 16 mM Na2HPO4, 150 mM NaCl, pH 7.3) for 30 minutes. The clean cuvette was rinsed once with distilled water, and the protein solution was added (3 mL, 20 μM). The sample was excited at 280 nm with an excitation slit width of 1.5 nm. The fluorescence was measured at the peak maximum (345 nm). This was repeated three times until a stable emission intensity at 350 nm was reached. Retinal was added to the cuvette in varying amounts from a 1.5 mM

stock solution in ethanol maintained in the dark. Care was taken to ensure that the EtOH volume remained below 2%. The titration was complete when there was no observable quenching of fluorescence upon addition of the chromophore. This was plotted as concentration of chromophore versus relative fluorescence intensity. The data were analyzed by nonlinear square fit of the equation. The Kd for each hCRBPII was determined according to the method previously described.[(31)]

I-8-3 Crystallization and Structure Determination

The pure WT hCRBPII protein was concentrated to 5-10 mg/mL. The complexes of protein with all-*trans*-retinol were prepared by adding 3-4 equivalents of retinal and retinol solution (30 mM retinol in ethanol). The final concentration of ethanol in protein solution was kept below 10% V/V. The mixtures of protein and ligands were incubated at room temperature in Black LiteSafe™ Microcentrifuge Tubes for 2 hours. The crystal of the complex of protein and ligands were prepared at room temperature by hanging drop vapor diffusion in Limbro plates wrapped in aluminum foil to prevent light-initiated degradation of the retinal and retinol. The best crystals were grown using Hampton Research Screens (the best crystals growing in 30-35% PEG 4000 (Sigma-Aldrich), 0.1 M sodium acetate (Columbus Chemical Industry) pH 4.6-4.8 and 0.1 M ammonium acetate (J. T. Baker) with 1 µL of hCRBPII-ligand complex and 1 µL of mother liquor). The crystals appeared after 3 days and reached their maximum size in one week. The crystals were flash frozen in liquid nitrogen using a cryo-protectant solution (30% PEG 4000, 0.1 M sodium acetate pH 4.6-4.8, 0.1 M ammonium acetate, 20% glycerol) and stored in a liquid nitrogen dewar.

All of the diffraction data were collected at beamline 21-ID-D, LS-CAT (Argonne National Laboratory, Advanced Photon Source, Chicago, IL) using a MAR300 detector and 1.00Å wavelength radiation at 100K. The diffraction data were indexed using the HKL2000 software package.[32] The structure was solved by molecular replacement using the MOLREP program implemented in the ccp4 programing package using human CRBPII (Protein data bank, accession code 2RCQ[22]) as model. Initial electron density maps and structure refinement was performed using REFMAC5 in the CCP4 suite.[33, 34] [35]All rebuilding and placement of ordered water molecules was done manually using COOT.[36] The chromophore was created using j-ligand and manually fitted in the electron density at the final step of the refinement.

Table I-1: Data-collection and refinement statistics.

	wt hCRBPII-retinol	wt hCRBPII-retinol-7KeV	wt hCRBPII-retinol-11KeV
Space group	$P2_1$	P1	P1
a(Å)	34.65	36.41	36.48
b(Å)	75.14	54.18	54.14
c (Å)	54.65	68.44	68.30
α(°)	90.00	107.72	107.64
β(°)	100.78	97.19	96.94
γ(°)	90.00	103.59	103.71
Resolution (Å)	50.00-1.19(1.21-1.19)[a]	29.79-1.89(1.94-1.89)[a]	34.661-1.496 (1.51-1.50)[a]
Total reflection	306503	76284	358399
Unique Reflection	87302	34646	76284
Completeness (%)	99.8(84.0)[a]	92.36(86.0)[a]	95.20(88.0)[a]
Molecules per Asymmetric Unit	2	4	4
Average I/σ	20.19(2.04)[a]	12.97(6.23)[a]	26.45(2.41)[a]
R_{merge}(%)	9.1(43.4)[a]	7.8(32.5)[a]	4.2(59.8)[a]
Mosaicity (°)	0.486	0.58	0.58

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

Table I-1 (cont'd)

Refinement statistics R_{work}/ R_{free} (%)			
Bond Length (Å)	0.0297	0.008	0.007
Angle (°)	2.381	1.079	1.066
B average for main chain	9.148	34.628	23.69
B average for water and side chain	16.410	45.016	39.549
Number of water molecules	507	331	468
Ramachandran plot	97.10	96.04	96.20
Most favored (%)	2.07	3.21	2.28
Allowed (%)	0.83	2.12	2.49
Outliers (%)			
PDB#	4QYN	4QZT	4QZU

REFERENCES

REFERENCES

1. Ong, D. E. (1984) A NOVEL RETINOL-BINDING PROTEIN FROM RAT - PURIFICATION AND PARTIAL CHARACTERIZATION, *Journal of Biological Chemistry 259*, 1476-1482.

2. Nossoni, Z., Assar, Z., Yapici, I., Nosrati, M., Wang, W., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal, *Acta crystallographica. Section D, Biological crystallography 70*, 3226-3232.

3. Veerkamp, J. H., Vankuppevelt, T., Maatman, R., and Prinsen, C. F. M. (1993) STRUCTURAL AND FUNCTIONAL-ASPECTS OF CYTOSOLIC FATTY-ACID-BINDING PROTEINS, *Prostaglandins Leukotrienes and Essential Fatty Acids 49*, 887-906.

4. Banaszak, L., Winter, N., Xu, Z. H., Bernlohr, D. A., Cowan, S., and Jones, T. A. (1994) LIPID-BINDING PROTEINS - A FAMILY OF FATTY-ACID AND RETINOID TRANSPORT PROTEINS, *Advances in Protein Chemistry, Vol 45 45*, 89-151.

5. Vaezeslami, S., Mathes, E., Vasilelou, C., Borhan, B., and Geiger, J. H. (2006) The structure of apo-wild-type cellular retinoic acid binding protein II at 1.4 A and its relationship to ligand binding and nuclear translocation, *Journal of Molecular Biology 363*, 687-701.

6. Haunerland, N. H., and Spener, F. (2004) Fatty acid-binding proteins--insights from genetic manipulations, *Progress in lipid research 43*, 328-349.

7. Lucke, C., Huang, S., Rademacher, M., and Ruterjans, H. (2002) New insights into intracellular lipid binding proteins: The role of buried water, *Protein science : a publication of the Protein Society 11*, 2382-2392.

8. Schaap, F. G., van der Vusse, G. J., and Glatz, J. F. (2002) Evolution of the family of intracellular lipid binding proteins in vertebrates, *Molecular and cellular biochemistry 239*, 69-77.

9. Smathers, R. L., and Petersen, D. R. (2011) The human fatty acid-binding protein family: evolutionary divergences and functions, *Human genomics 5*, 170-191.

10. Franzoni, L., Lucke, C., Perez, C., Cavazzini, D., Rademacher, M., Ludwig, C., Spisni, A., Rossi, G. L., and Ruterjans, H. (2002) Structure and backbone dynamics of apo- and holo-cellular retinol-binding protein in solution, *Journal of Biological Chemistry 277*, 21983-21997.

11. Crow, J. A., and Ong, D. E. (1985) CELL-SPECIFIC IMMUNOHISTOCHEMICAL LOCALIZATION OF A CELLULAR RETINOL-BINDING PROTEIN (TYPE-2) IN THE SMALL-INTESTINE OF RAT, *Proceedings of the National Academy of Sciences of the United States of America 82*, 4707-4711.

12. Vogel, S., Mendelsohn, C. L., Mertz, J. R., Piantedosi, R., Waldburger, C., Gottesman, M. E., and Blaner, W. S. (2001) Characterization of a new member of the fatty acid-binding protein family that binds all-trans-retinol, *Journal of Biological Chemistry 276*, 1353-1360.

13. Folli, C., Calderone, V., Ottonello, S., Bolchi, A., Zanotti, G., Stoppini, M., and Berni, R. (2001) Identification, retinoid binding, and x-ray analysis of a human retinol-binding protein, *Proceedings of the National Academy of Sciences of the United States of America 98*, 3710-3715.

14. Li, E., Demmer, L. A., Sweetser, D. A., Ong, D. E., and Gordon, J. I. (1986) ISOLATION OF A CLONED CDNA-ENCODING AN ABUNDANT INTESTINAL RETINOL BINDING-PROTEIN, *Clinical Research 34*, A685-A685.

15. Li, E., and Norris, A. W. (1996) Structure/function of cytoplasmic vitamin A-binding proteins, *Annual Review of Nutrition 16*, 205-234.

16. Li, E., Qian, S. J., Winter, N. S., Davignon, A., Levin, M. S., and Gordon, J. I. (1991) FLUORINE NUCLEAR-MAGNETIC-RESONANCE ANALYSIS OF THE LIGAND-BINDING PROPERTIES OF 2 HOMOLOGOUS RAT CELLULAR RETINOL-BINDING PROTEINS EXPRESSED IN ESCHERICHIA-COLI, *Journal of Biological Chemistry 266*, 3622-3629.

17. Calderone, V., Folli, C., Marchesani, A., Berni, R., and Zanotti, G. (2002) Identification and structural analysis of a zebrafish apo and holo cellular retinol-binding protein, *Journal of Molecular Biology 321*, 527-535.

18. Tarter, M., Capaldi, S., Carrizo, M. E., Ambrosi, E., Perduca, M., and Monaco, H. L. (2008) Crystal structure of human cellular retinol-binding protein 11 to 1.2 angstrom resolution, *Proteins-Structure Function and Bioinformatics 70*, 1626-1630.

19. Winter, N. S., Bratt, J. M., and Banaszak, L. J. (1993) CRYSTAL-STRUCTURES OF HOLO AND APO-CELLULAR RETINOL-BINDING PROTEIN-II, *Journal of Molecular Biology 230*, 1247-1259.

20. Kane, M. A., Bright, F. V., and Napoli, J. L. (2011) Binding affinities of CRBPI and CRBPII for 9-cis-retinoids, *Biochimica Et Biophysica Acta-General Subjects 1810*, 514-518.

21. Malpeli, G., Stoppini, M., Zapponi, M. C., Folli, C., and Berni, R. (1995) INTERACTIONS WITH RETINOL AND RETINOIDS OF BOVINE CELLULAR RETINOL-BINDING PROTEIN, *European Journal of Biochemistry 229*, 486-493.

22. McBee, J. K., Kuksa, V., Alvarez, R., de Lera, A. R., Prezhdo, O., Haeseleer, F., Sokal, I., and Palczewski, K. (2000) Isomerization of all-trans-retinol to cis-retinols in bovine retinal pigment epithelial cells: Dependence on the specificity of retinoid-binding proteins, *Biochemistry 39*, 11370-11380.

23. Borshchevskiy, V. I., Round, E. S., Popov, A. N., Bueldt, G., and Gordeliy, V. I. (2011) X-ray-Radiation-Induced Changes in Bacteriorhodopsin Structure, *Journal of Molecular Biology 409*, 813-825.

24. Matsui, Y., Sakai, K., Murakami, M., Shiro, Y., Adachi, S., Okumura, H., and Kouyama, T. (2002) Specific damage induced by X-ray radiation and structural changes in the primary photoreaction of bacteriorhodopsin, *Journal of Molecular Biology 324*, 469-481.

25. Schreckenbach, T., Walckhoff, B., and Oesterhelt, D. (1977) STUDIES ON RETINAL-PROTEIN INTERACTION IN BACTERIORHODOPSIN, *European Journal of Biochemistry 76*, 499-511.

26. Cowan, S. W., Newcomer, M. E., and Jones, T. A. (1993) CRYSTALLOGRAPHIC STUDIES ON A FAMILY OF CELLULAR LIPOPHILIC TRANSPORT PROTEINS - REFINEMENT OF P2 MYELIN PROTEIN AND THE STRUCTURE DETERMINATION AND REFINEMENT OF CELLULAR RETINOL-BINDING PROTEIN IN COMPLEX WITH ALL-TRANS-RETINOL, *Journal of Molecular Biology 230*, 1225-1246.

27. Vasileiou, C., Vaezeslami, S., Crist, R. M., Rabago-Smith, M., Geiger, J. H., and Borhan, B. (2007) Protein design: reengineering cellular retinoic acid binding protein II into a rhodopsin protein mimic, *J Am Chem Soc 129*, 6140-6148.

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

29. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *Journal of Applied Crystallography 30*, 1022-1025.

30. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP, *Acta Crystallographica Section D-Biological Crystallography 66*, 22-25.

31. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallographica Section D-Biological Crystallography 67*, 355-367.

Chapter II: Domain swapping in hCRBPII

II-1 Introduction

Protein oligomers have evolved because of their advantages over their monomers. These advantages include the possibility of allosteric control, larger binding surfaces, higher concentration of active sites, new active sites at subunit interfaces and higher stability. However, the mechanisms for the evolution and assembly of oligomeric species during protein synthesis remain unclear. Different mechanisms have been proposed for the evolution of protein oligomers, three-dimensional (3D) domain swapping is among them. [37-42] Three-dimensional (3D) domain swapping (DS) is a process in which two or more monomer protein molecules exchange their identical structural elements to form dimers or higher-order oligomers **(Figure II-1).** [37-42]

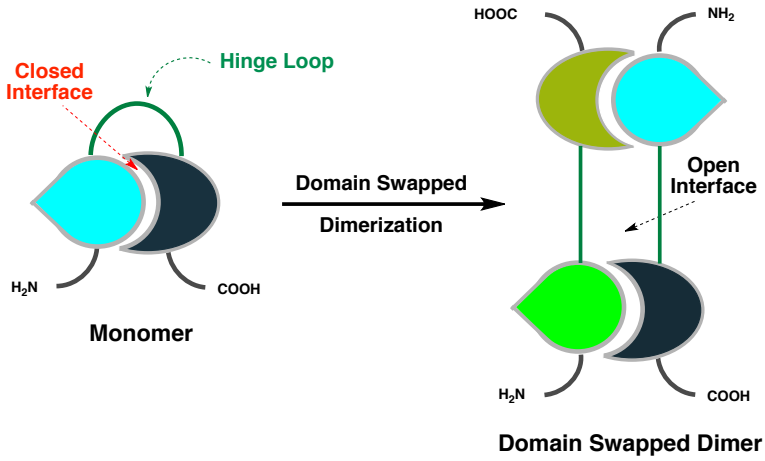


Figure II-1: Cartoon illustration of 3D domain swapping. Identical structural units swap between 2 monomers, followed by creating the open interface between them.

The region that differs markedly between monomer and oligomer, called the "switch region" or hinge region" is relatively small. In most cases it consists of a loop in the monomer that becomes straightened in the DS oligomer. In many cases it is this region that seems to be the most important determinant for the occurrence of domain swapping. It was first discovered by Eisenberg and coworkers in 1994 to account for the dimerization mechanism of diphtheria toxin (DT) **(Figure II-2)** and was later shown to be responsible for the dimerization of lyophilized Bovine pancreatic ribonuclease (Rnase A, which can have either N-terminal or C-terminal swapped dimers). So far only about 40 distinctly different proteins have been observed to form domain swapped dimers. *(43-45)*

Diphtheria toxin (DT) is a 533-residues protein toxin secreted from a bacterium that causes diphtheria. DT has three domains: a catalytic domain (C), a transmembrane domain (T) and a receptor-binding domain (R) *(37-42)*. The dimeric form of DT does not form spontaneously even at higher concentrations. However, DT can be induced to form dimers by decreasing pH, which converts monomeric DT into open monomers, which form a DT dimer at high concentration by swapping the globular domain (R domain) *(37-42)* The structure of monomer DT revealed the mechanism by which low pH could trigger changes in monomer DT and thereby form an open monomer. In the monomer, the interdomain interface between the R domain and C domain is charged, with nine basic and three acidic residues on the R domain interface and seven acidic residues on the C domain interface. There are three salt bridges stabilizing the interface at neutral pH. The decrease in pH causes protonation of those acidic residues and disruption of the salt bridges. Furthermore, buried positive charges in the interface favor the open monomer formation. During the dimerization of DT, higher order oligomers were

observed by size exclusion HPLC. These oligomers include trimers, tetramers and pentamers [37-42]. In most cases the exchanged region is located at the N- or C-terminus of the protein, though the exchanged part is found in the middle of the sequence in a few instances. [46-49]

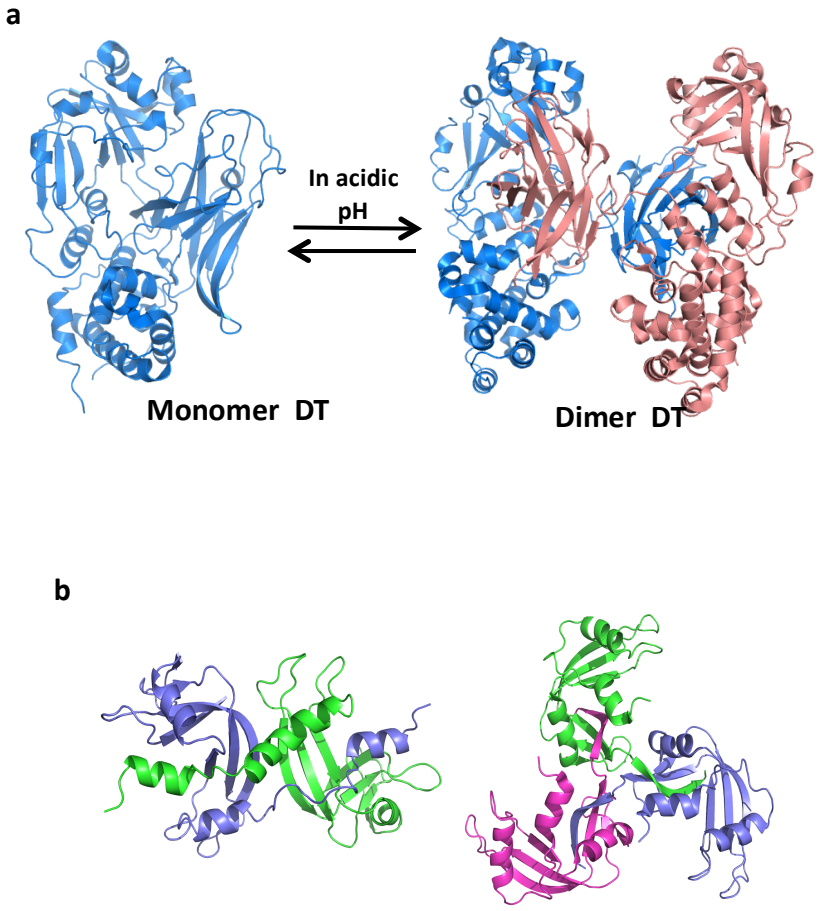


Figure II-2: a. Left: DT monomer (PDB ID 4OW6). Right: DT dimer (4AE1). The two subunits are blue and pink. In acidic pH equilibrium goes toward DS dimer formation. **b.** Left structure of DS dimer of RnaseA (PDB 1A2W) and cyclic DS trimes (PDB 5SRA).

Domain swapping, combined with gene duplication, has been postulated to be responsible for the evolution of a number of larger protein domains from smaller fragments[50]**.** Domain swapping can also lead to aggregation, amyloid and fibril formation, which lead to

protein malfunction[51]**.** Protein misfolding and aggregation is the cause of a number of maladies including Parkinson's disease, Alzheimer disease, diabetes, Huntington's disease and many others[52-58]. Triggers of domain-swapping include lyophilization, non-physiological pH, unusually high protein concentration [59, 60] [43, 61] temperature, mutation and even ligand binding[62, 63]**.**

The classic example of domain swapping is bovine pancreatic ribonuclease A (RNase A), which, the DS of the N-terminal exchange, was proposed in 1962 by Crestfield, Stein, and Moore to describe its behavior under acidic conditions. Followed by the first X-ray structures for a domain swapped dimer in the late nineties. In RNase case, either the N-terminal helix or the C-terminal strand can domain swap. The swapped unit and its oligomeric state can vary significantly and RNase A illustrates the possibilities of domain swapping in the trimer both N- and C-terminal units are swapped and a circular arrangement succeeds **(Figure II-2b).** Another example is Cyanovirin-N (CV-N), which is a 101 amino acid cyanobacterial lectin. CV-N inactivates HIV and is a general virucidal agent against other enveloped viruses. The original solution structure was a monomer, whereas the subsequently determined X-ray structures were domain swapped dimers, which each domain comprises a triple-stranded β-sheet with a β-hairpin packed on top. A helical linker is located in the middle of the sequence. It was shown that the DS dimer is a kinetically trapped folding intermediate at high protein concentrations that convert into the slightly more stable monomer form at physiological (>30^{O}C) temperature.[1d, 10] At room temperature or below, however, the dimer lifetime is sufficiently long for structural characterization in solution. The fact that both monomeric and domain-swapped dimeric CV-N coexist in solution under identical conditions indicates that the free energies of folding for both

quaternary states must be comparable and the kinetic barrier between the monomer and dimer has to be significant. This can be altered by mutation of residues in the hinge region. For instance, changing the proline in the hinge region to glycine resulted in a substantial stabilization of this monomeric P51G mutant by >5 kcal/mol compared to wild-type. A S52P mutant yielded predominantly dimeric protein due to destabilization of the monomer. More and more domain swapped protein structures are becoming elucidated, and, for several cases, growing evidence supports that the dimer or multimer is an active, biological important structure. Regardless of whether domain swapping is a specific mechanism for regulation *in vivo*, it is becoming clear that domain swapping is a means by which stable oligomers can be generated under evolutionary force. By protein expression or refolding at high concentrations, suggesting that high concentrations of folding intermediates may be involved in the process.[64, 65] So far no case has been observed where both the monomer and dimer of a wild-type protein are expressed without rapid, equilibrating monomer/dimer exchange. Domain-Swapping can also be caused by mutation. For example a single site mutation in protein L from Peptostreptococcus Magnus leads to 3D domain swapping [66]**.**

As discussed in the previous chapter, iLBPs are relatively small (126-140 amino acid, predominantly β sheet proteins) cytosolic proteins. Structures of a number of iLBPs have been determined and all share this monomeric fold. **(Figure II-1)** Numerous studies of the folding mechanism of members of the iLBP family have been conducted, specially focused on human Cellular Retinoic Acid Binding Protein I (hCRABPI). [67-71] In an elegant series of experiments, using a variety of techniques, the Gierasch lab has concluded that CRABPI undergoes early beta barrel collapse.

They used CRAPI with its simple architecture to study structural basis of β sheet proteins formation. Regarding this purpose, three only Tryptophan residues, with structurally distinct locations as three probes, were mutated. Folding of these mutants was examined using stopped-flow fluorescence and circular dichroism. As a result within 10 ms, the rapid hydrophobic collapse occurs and stopped-flow circular dichroism shows significant secondary structure content and it (chain topology) gets developed by 100 ms. Then followed by the specific packing of the β-sheet sidechains and formation of the native hydrogen-bond network (T = 1 s). These are hallmarks of the pathway, thus preventing off pathway aggregation and potential amyloid formation in this protein.[72]

Investigations into the folding mechanism of other members of the family have also been conducted, all of which are consistent with a relatively rugged energy landscape, indicating a number of possible stable or meta-stable intermediates, a situation that is generally characteristic of proteins that are largely beta sheet in structure.[72]

Also in chapter II was mentioned that CRBPs belong to the iLBP family. [6, 10-13] The combination of small size and relatively large binding cavity made these family members ideal templates for use in protein design applications.

The Geiger and Borhan labs have used these templates for a wide-ranging protein design applications for more than a decade. First both hCRABPII and hCRBPII were modified to be Rhodopsin mimics that bind and react with retinal to form the protonated Schiff base (PSB) which is found in the binding cavity of all rhodopsins.[31, 73] Then the environment around the bound chromophore was modified to study the mechanisms of absorption wavelength tuning that give rise to color vision, and the light selectivity of all the other rhodopsins, which is

required for their function.[74-76] Rhodopsin mimics that photo-isomerize both in solution and in the crystal were created to allow the detailed study of isomerization processes in rhodopsins.[77, 78] In our collaboration, we have also created a new class of fluorescent proteins by binding fluorophores in the binding pocket that also react to form PSB's and allow regulation of the emission wavelength. [79] (Some of this work is described further in chapter IV) To perform all the above-mentioned projects, we had to make, express and evaluate hundreds of mutants of these proteins. During the preparation of mutants for these studies, Dr. Wang and Dr. Nossoni for the fist time observed protein bands that eluted separately on ion-exchange chromatography, but nonetheless, were the same species on SDS-PAGE. This led us to the discovery of a set of hCRBPII protein mutants capable of domain swapped dimerization. It was a surprise to discover that some hCRBPII mutants predominantly form dimers via domain swapping instead of the monomeric form. [45] **(Figure II-3)**

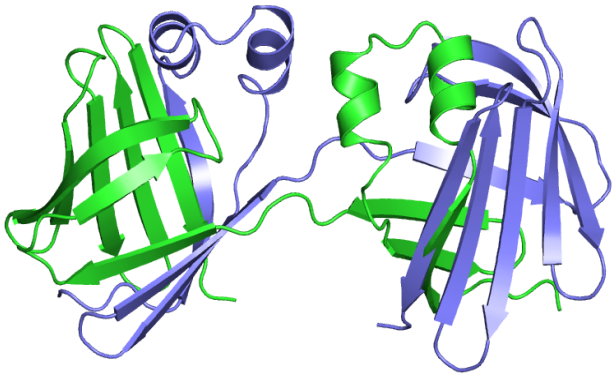


Figure II-3: The hCRBPII DS dimer.

We were the first group to report domain swapped structures of hCRBPII. [45] The extent of domain swapping is large (almost 50% of the protein sequence) and occurs during routine

protein bacterial over-expression at 16-30 °C. HCRBPII domain swapped dimerization is favored by high expression levels in E. coli and high concentrations in vitro. Most importantly protein stability studies show the domain swapped dimer to be a less stable, kinetically trapped species (depending on the mutant). The resultant dimer structure represents a fold not identified in the SCOP or CATH protein fold databases. Together, our data supports the hypothesis that folding of hCRBPII occurs via a structurally ordered "open monomer," with the relative orientation of the N- and C- terminal structural units determining the propensity for dimerization.[2c] The existence of this dimer for hCRBPII suggests a folding pathway distinct from that characterized for hCRABPI, indicating that evolutionary related members of the same structural family may follow different folding pathways.

Later on, we determined structures of holo hCRBPII domain swapped dimer variants. Together, these structures show an extremely large and reproducible change in the relative orientation of the two domains of the dimer upon ligand binding, inviting the possibility that iLBP domain swapped dimers could be allosterically-regulated forms of these proteins, at least in some cases. Ideally, the domain swapped dimer may represent a new, biologically relevant structural fold for the iLBP protein family, though it remains to be seen if they form in vivo, and have distinct functionality relative to their parent, monomeric forms.

II-2 Origin of dimerization in hCRBPII

As it was mentioned above, we developed hCRBPII as rhodopsin mimics to study the mechanism of wavelength tuning. To this purpose, the binding site was redesigned to bind all-*trans* retinal and form a PSB via the Q108K:K40L (KL) double mutant by Dr. Wang previously [(80)].

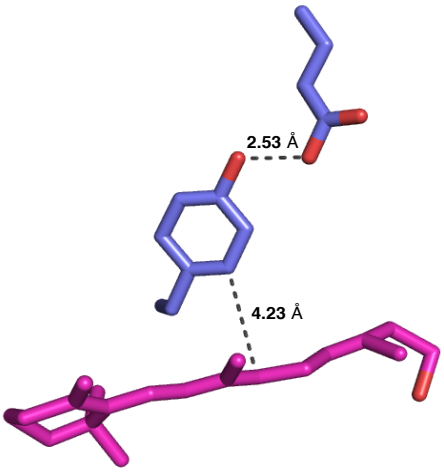


Figure II-4: Trajectory of Tyr 60, Glu 72 (both shown as stick representations, colored blue) in WT hCRBPII with all-*trans*-retinal.

The residues situated around the chromophore were then selected for systematic mutation in an effort to understand the mechanism of wavelength tuning. Tyr60, conserved only in CRBPs, is located in beta-strand C and approximately 4Å from the center of the polyene chain of all-*trans*-retinal. **(Figure II-4)** [17, 80-82].

The mutant Q108K:K40L:Y60W (KLY60W) was generated by Dr. Wang, expressed and purified as previously described in the hopes of producing a π-π interaction with the chromophore that could potentially delocalize the positive charge of the PSB along the chromophore, leading to shifted spectra.[80] In its purification process, Source-Q ion exchange chromatography (IEX) unexpectedly gave two KLY60W hCRBPII-containing protein peak fractions, one eluting at 80mM NaCl, where hCRBPII mutants typically elute, and a second eluting at 160 mM NaCl [45] **(Figure II-5a)**. The UV-Vis spectra of the all-*trans*-retinylidene PSB complex with protein from each of the fractions gave distinct spectra. The 80 mM and 160 mM protein fractions, when bound to retinal, showed λ_{max} of 496nm and 514nm, respectively **(Figure II-5b, 5c).** At lower expression temperature (25°C) significantly more of the 80 mM

fraction was produced while a higher expression temperature (30°C) gave similar amounts of each species **(Table II-1)**

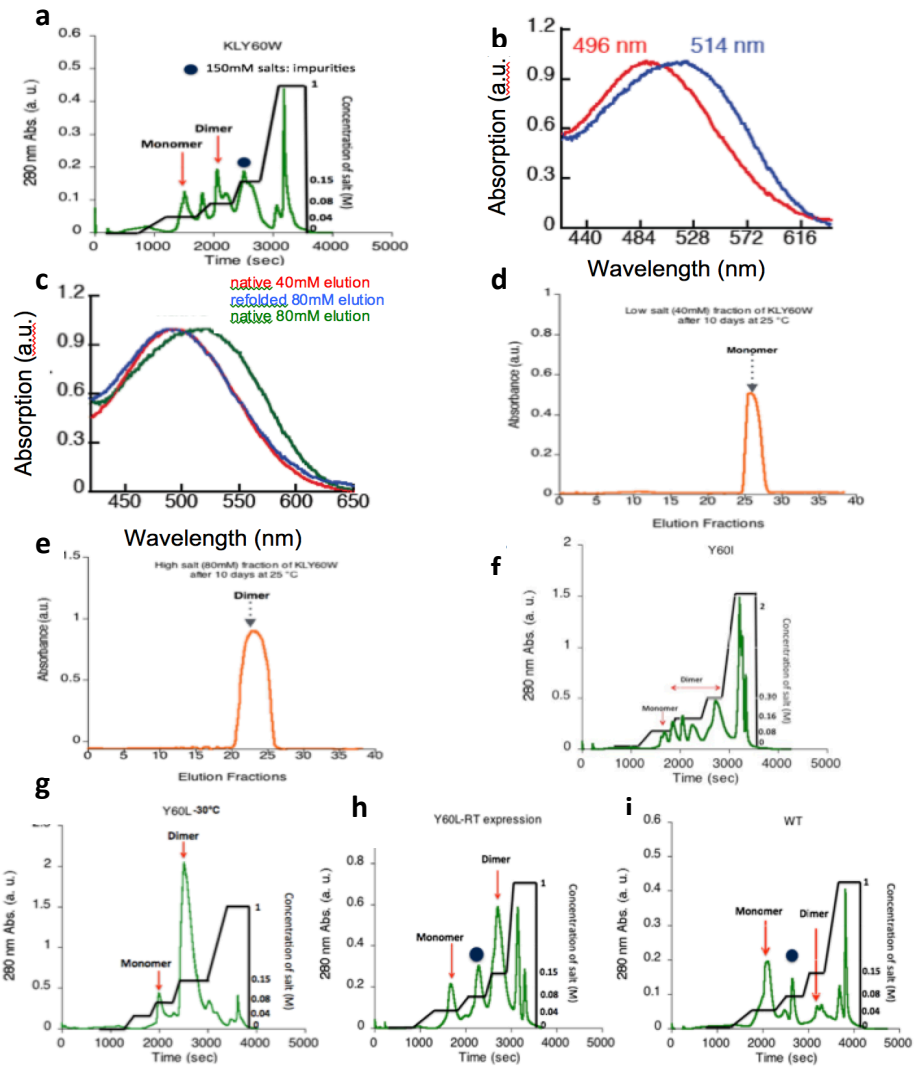


Figure II-5: a. IEX chromatogram of the KLY60W, monitored at 280nm. **b.** The UV-Vis spectra of the 80mM salt elution (red) and the 160mM salt elution (blue) of KLY60W incubated with all-*trans*-retinal. **c.** The UV-Vis spectra of the refolded 80mM eluent (low salt, blue) after denaturation and refolding at 0.03 mM concentration and the native 160mM (high salt, green) and native 80mM elution (red), all bound to all-*trans*-retinal. **d**. The size exclusion chromatogram (Superdex 200 16/75 column) for dimer (the 160mM elution) KLY60W, after maintaining the protein at room temperature for several days, showing only the monomer size peak. **e.** Gel filtration chromatogram of monomer (the 80mM elution) of KLY60W, showing only the dimer fractions. **f-i.** Ion exchange chromatographs of WT and various hCRBPII mutants, monitored at 280 nm. **f.** Y60I at 30°C. **g.** Y60L at 30°C. **h.** Y60L at 25°C. **i.** WT hCRBPII expressed at 30 °C.

Size-exclusion chromatography (SEC) showed that the 80 mM protein eluted as a 15kD monomer, and the 160 mM fraction eluted as a 30kD dimer. No equilibrium between the monomer and dimer forms was observed, even after incubation for 10 days at room temperature **(Figure II-5d and 5e).** Each species has been tested to achieve more information about their stability. Their melting curves showed the dimer to be less stable than the monomer (denaturing at 69°C, explained in **II-7-4 section, Figure II-6c and 31**), however the monomer was stable to 95°C **(Figure II-6f and II-31).** The 160 mM salt fraction was denatured with 8M urea and refolded at low concentration (0.03 mM) **(Table II-1**). The UV-Vis spectrum of the retinal PSB-bound refolded monomer was identical to the originally isolated monomer, indicating the refolded proteins to be structurally similar **(Figure II-5c).**

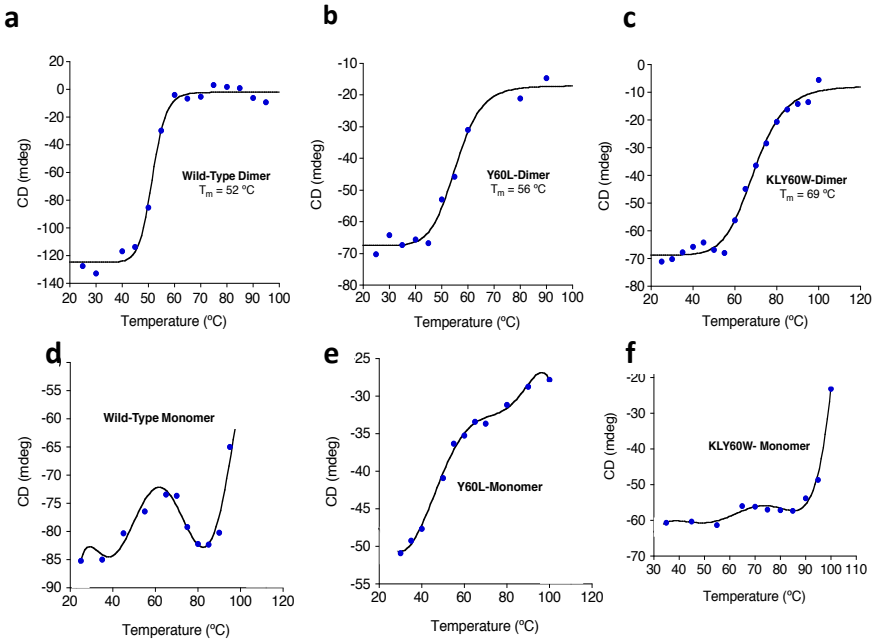


Figure II-6: Thermal melting curves: **a.** WT hCRBPII dimer (θ measured at 220 nm), T_m = 52 °C; **b.** Y60L dimer (θ measured at 227 nm), T_m = 56 °C; **c.** KLY60W dimer (θ measured at 220 nm), T_m = 69 °C. **d.** WT hCRBPII monomer (θ measured at 220 nm); **e.** Y60L monomer (θ measured at 215 nm); **f.** KLY60W monomer (θ measured at 220 nm). Note that most of the secondary structure is preserved in both the WT and KLY60W monomers, even at the highest temperature achieved (CD intensities of-65mdeg and -50mdeg respectively). Thus melting curves can only give a lower limit for the T_M, consistent with the view that the monomers are more stable than the dimers. The apparent lack of a T_m for the monomers might suggest the presence of thermally induced folding intermediates. Clearly the unfolding process is not monolithic as that observed for the dimeric species. This lends further support to the suggestion that the open monomer, which is presumably obtained from melting of the dimer, has a different folding trajectory (and consequently a different unfolding trajectory) as compared to the monomeric species.

II-3 Dimer formation of other hCRBPII variants

At first, we hypothesized mutation of the residue 60 in the KLY60W series was most likely responsible for the formation of the dimeric species since many hCRBPII proteins mutated at the 40 and 108 positions have been purified in monomeric form in our lab previously[(80)]. To this purpose, we made several single mutants at this position **(Table II-2).**

Y60W mutant yielded the same monomer/dimer ratios as KLY60W, confirming the importance of position 60. [45] In fact most mutations resulted in increased dimer formation, including Y60F, Y60L and Y60I, however the Y60H mutant produced no dimeric protein. **(Table II-2 and Figure II-5f-h)** Together, this data emphasizez the fact that position 60 mutations can significantly favor the formation of dimeric species. Position 60 lies near the N-terminal end of strand 4, and is not a position conserved in iLBP family members, though it is a Tyr in most retinol binding proteins.[81]

Table II-1: The Effect of expression temperature and *in vitro* refolding protein concentration on the dimer/monomer (D/M) ratio.

Protein	D/M (%)[a] at 25 °C	D/M (%)[a] at 30 °C	*In vitro* refolding[b]	Percent of D/M (%)[c]
WT-hCRBPII	0/100	10/90	0.03 mM 0.13 mM 0.3 mM	0/100 0/100 40/60
Q108K:K40L: Y60W	30/70	60/40	0.03 mM 0.6 mM	0/100 30/70
Y60L	80/20	90/10	0.03 mM 0.13 mM 0.4 mM	20/80 40/60 70/30

a. Mass ratio of dimer to monomer; b. Starting concentration; c. Mass ratio, after *in vitro* refolding.

We decided to perform several *in vitro* refolding experiments on the WT and several of the mutants **(Table II-1 and Figure II-7)**, to better understand the nature of the dimerization in this protein.

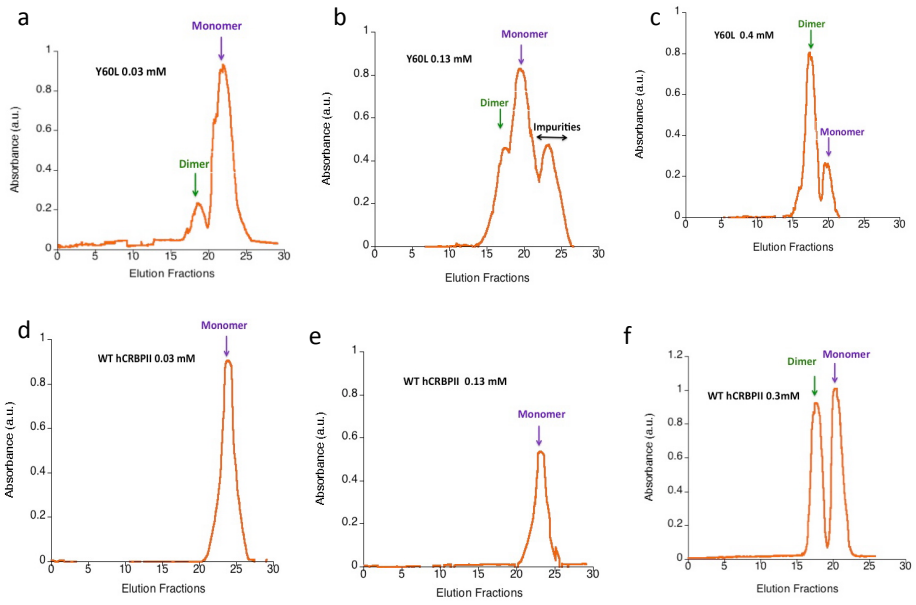


Figure II-7: Size exclusion chromatography after protein denaturation and refolding, all monitored at 280 nm. **a.** Y60L mutant refolded at 0.03 mM concentration. **b.** Y60L refolded at 0.13 mM concentration. **c.** Y60L refolded at 0.4 mM concentration. **d.** WT hCRBPII refolded at 0.03 mM concentration. **e.** WT hCRBPII refolded at 0.13 mM concentration. **f.** WT hCRBPII refolded at 0.3 mM concentration.

In each case, dimer formation was favored by higher refolding concentrations while monomer formation was favored at lower refolding concentrations. This correlated with our *E. coli*-expression results, where expression at higher temperatures, which leads to higher levels of expression and therefore higher concentrations of pre-folded proteins, (30 °C versus 25 °C, **Table II-1 and Figure II-7**). Melting experiments also showed the dimer to be less stable than the monomer for two of the three variants (WT, KLY60W, **Figure II-6a, 6c, 6d and 6f**), indicating the dimer to be the kinetic product (the melting curve for the Y60L monomer was not readily interpretable, **Figure II-6b and 6e**). [(45)]

Table II-2: Monomer/dimer ratio of various residue 60 mutants. During bacterial expression at 30 °C

Mutations	Dimer/Monomer mass ratio (%)
WT	10/90 (no dimer at RT)
Y60L	90/10
S55W	20/80
S55W:Y60L	60/40
Y60I	80/20
Y60Q	20/80
Y60D	30/70
Y60T	No dimer
Y60W	50/50
Q108K:K40L:Y60F	40/60
Q108K:K40L:Y60W	60/40
Δ56:WT	20/80
F57G:Y60L	70/30
T56P-Y60L	70/30
Δ56:Y60L	90/10
E72A	50/50
Q108K:K40L:Y60H	No dimer

Overall, these experiments are consistent with the dimer/monomer ratio being kinetically controlled, with higher concentrations favoring dimer formation. [45] In the meanwhile, we were investigating the possibility that WT hCRBPII could also form a dimeric species by overexpressing hCRBPII at the dimer favoring induction temperature of 30 °C, resulting in both monomeric and dimeric protein **(Figure II-5i).** To our knowledge, this was the first example of a domain swapped system where both monomer and dimer of the wild-type species are expressed without monomer/dimer exchange **(Figure II-8)**. [45]

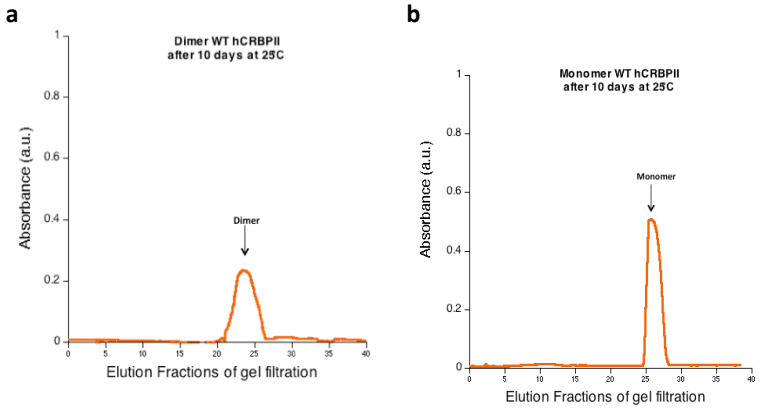


Figure II-8: a. The size exclusion chromatogram for dimer of WT hCRBPII, after maintaining the protein at room temperature for several days, showing only the dimer species. **b.** Gel filtration chromatogram of monomer of WT, showing only the monomer protein.

Binding assays also show that this dimer has a similar affinity (in nM range, compared to monomer) for retinol and retinal as well **(Figure II-9,** method explained in II-3-2 section).

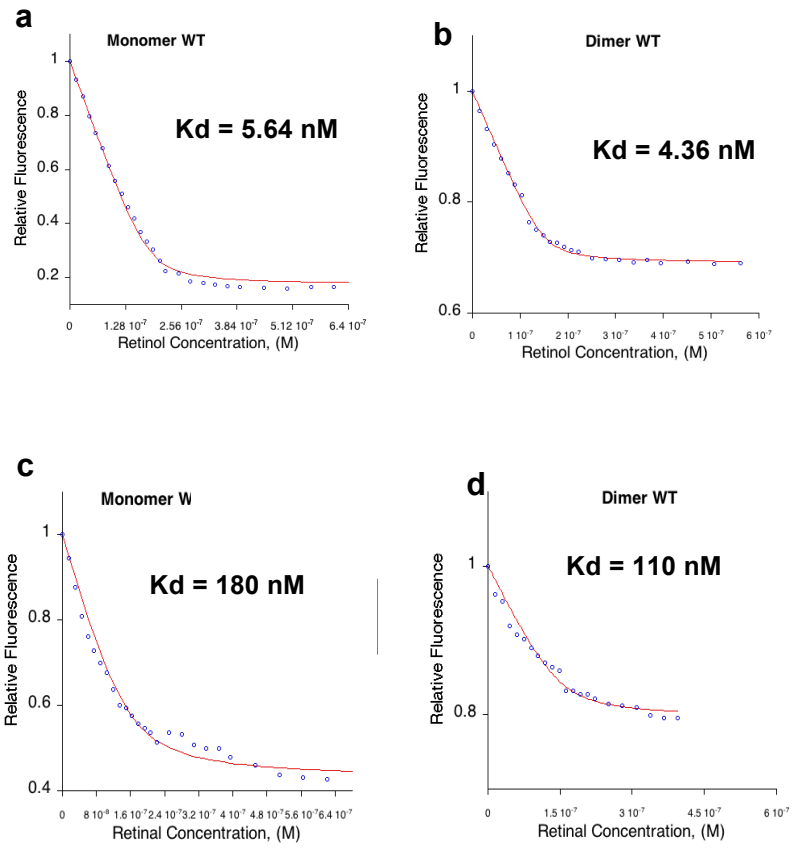


Figure II-9: Ligand binding, monitored by Tryptophan fluorescence quenching of monomeric and dimeric hCRBPII. **a.** Monomer WT hCRBPII with all-*trans*-retinol. **b.** Dimer WT hCRBPII with all-*trans*-retinol. **c.** Monomer WT hCRBPII with all-*trans*-retinal. **d.** Dimer WT hCRBPII with all-*trans*-retinal.

II-4 Structural Studies

II-4-1 Structural analysis reveals an extensive domain swapped dimer

We were able to crystallize and determine several dimer structures to understand the nature of this new form of hCRBPII. [(45)] **(Figure II-3)** In this attempt, structures of dimer mutants (KLY60W, Y60W and Y60L) and wild type (WT) were obtained respectively **(Table II-7).** The same extensive, domain swapped dimer was observed for all four of these structures **(Figures II-3, 10,**

11, 12a and 12b). The domain swapping is extensive, involving residue 1-56, which includes 3 beta strands and two helices **(Figure II-3, 12a, 12b).**

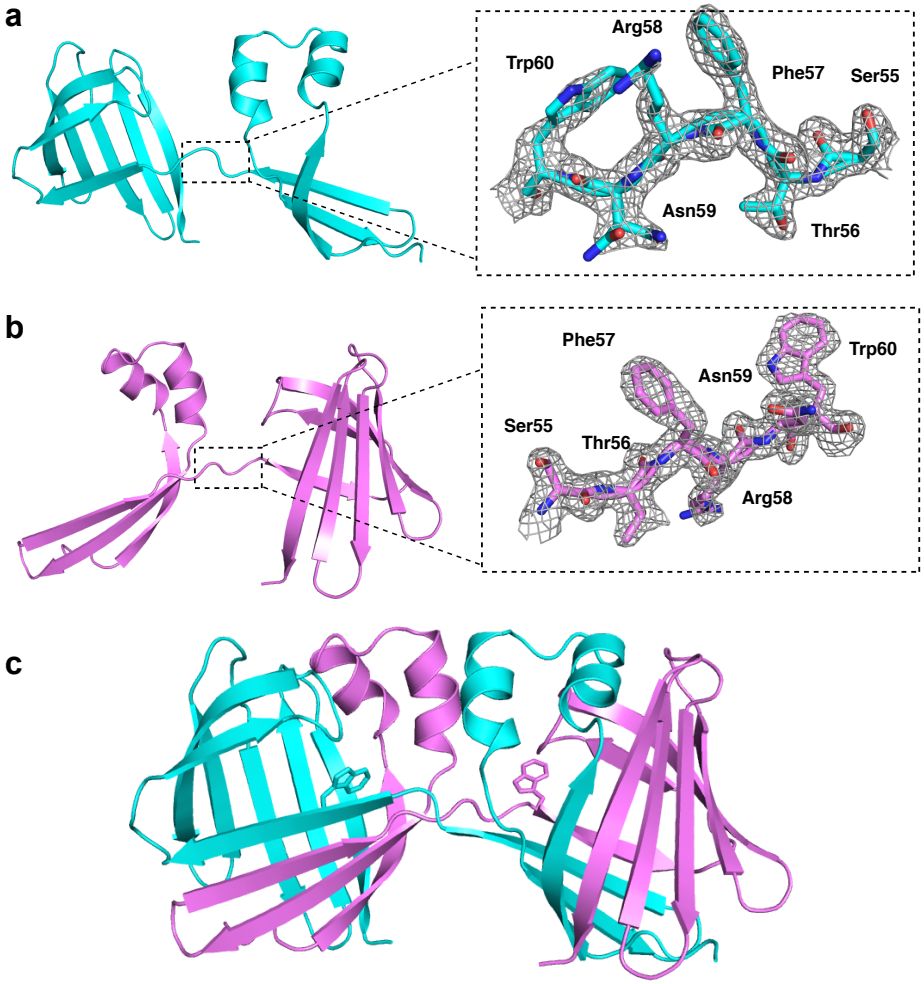


Figure II-10: Crystal structure of dimer KLY60W mutant and its monmer components **a.** Chain A of KLY60W (where only the Thr56 psi angle differs from the closed monomer form) with the electron density, contoured at 1.0 σ, of the hinge loop region encompassing amino acids 55-60. Atoms colored by type, N, blue, O, red, C, cyan. **b.** Chain B, showing the electron density, contoured at 1.0 σ, of the hinge loop region encompassing amino acids 55-60. Atoms colored by type, N, blue, O, red, C, cyan. **c.** The complete structure of the KLY60W dimer: In chain A (cyan) Trp60 is pointing inside the binding pocket, while in chain B (pink) Trp60 is pointing toward the solvent.

This corresponds to an 83% extent of swapping as defined in the domain swapping 3D knowledge database(http://caps.ncbs.res.in/3dswap/index.html).[83] Each of the two domains in the dimer are highly similar to that of the monomer hCRBPII structure with RMSD ranging from 0.382- 0.397 between monomeric hCRBPII and a single domain of the dimer **(Figure II-11)**. These domain swapped dimer structures have a large rotation about the psi angle of Thr 56 in common.

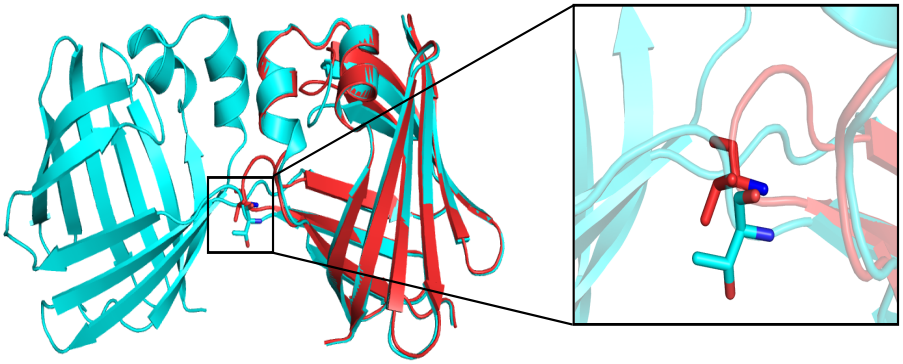


Figure II-11: An overlay of the dimeric (cyan) and monomeric (red) KLY60W mutant structures. Res56 for both structures is shown in sticks. The hinge loop region in the monomer and dimer is encircled, showing the dramatic difference in the two structures.

While the psi angle of Thr56 is about -12° in monomeric hCRBPII, it ranges from 132° – 169° in the domain swapped dimers **(Figure II-12a)** This 150°+ rotation about the psi angle of Thr56 represents the hinge motion required for domain swapping **(Figure II-11).** Though mutation at residue 60 substantially affects the monomer/dimer ratio, it is not in the hinge loop region, but is instead 3 amino acids away, near the middle of β-strand 4th. In fact substantial deviations in torsion angles from the monomer structure can be seen in residues 55 – 60 (from the hinge loop region till even the first 3 amino acids of the 4th beta sheet).

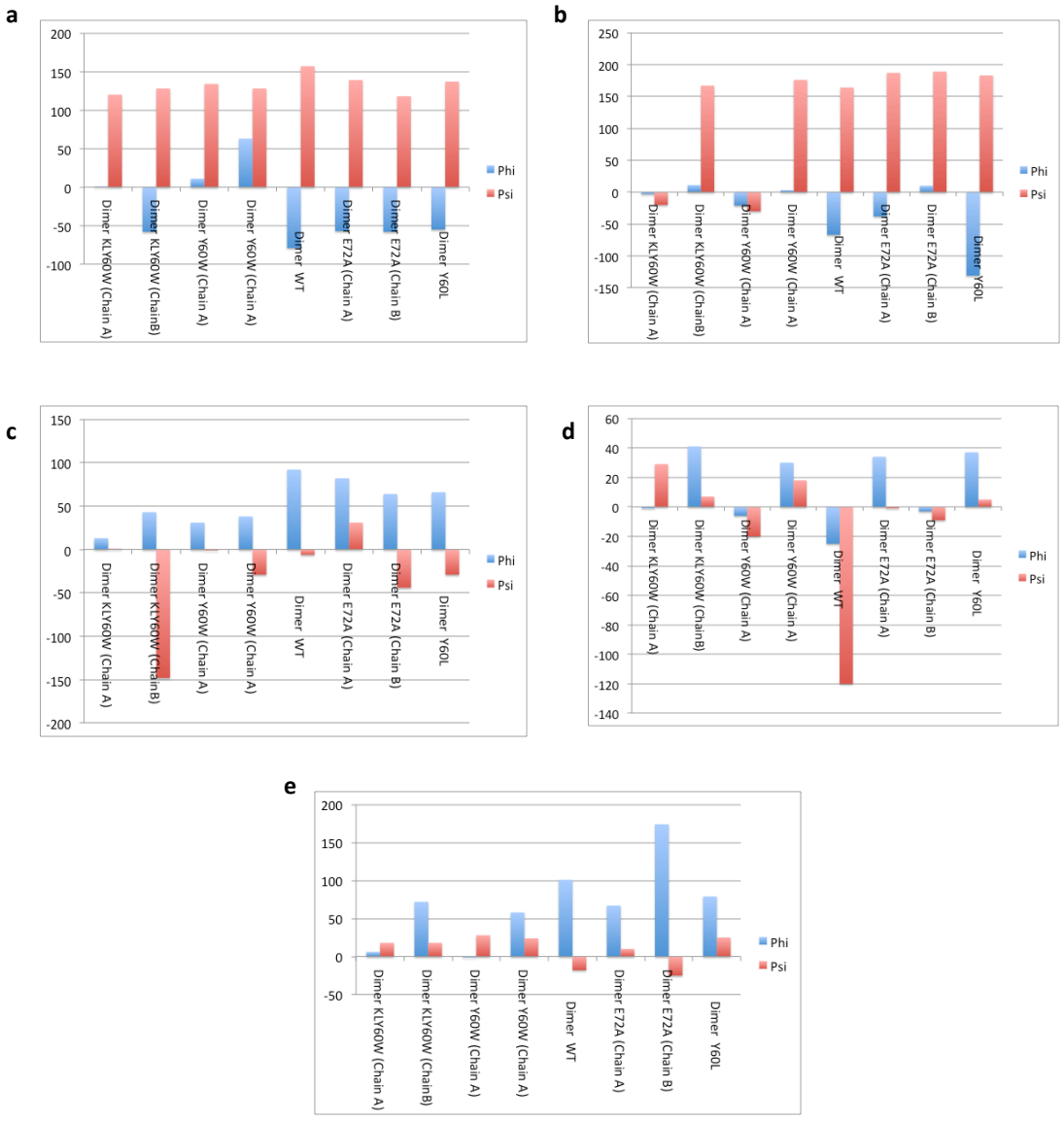


Figure II-12: a-e: Torsion angle differences between monomer and dimer differences of Phi angle are in blue and psi angle differences are in red in each case. **a.** Thr56. **b.** Phe57. **c.** Arg58. **d.** Asn59. **e.** Res60.

The domain swapping orients the opening of the two binding sites to face one another, creating a continuous and large internal cavity that stretches from one side of the dimer to the other (nearly 40 Å in length, (please refer to section II-6). A search of the various structural fold databases (SCOP 1.75, CATH 3.5) reveals no other protein of similar architecture.[(84)]

II-4-2 Symmetry VS asymmetry

Although all four-dimer structures shared the same domain swapped interface, there are substantial differences in their structures. As crystallographers and structural biologists, we believed that the structural details of the DS dimers are key to understanding the mechanism of domain swapping.

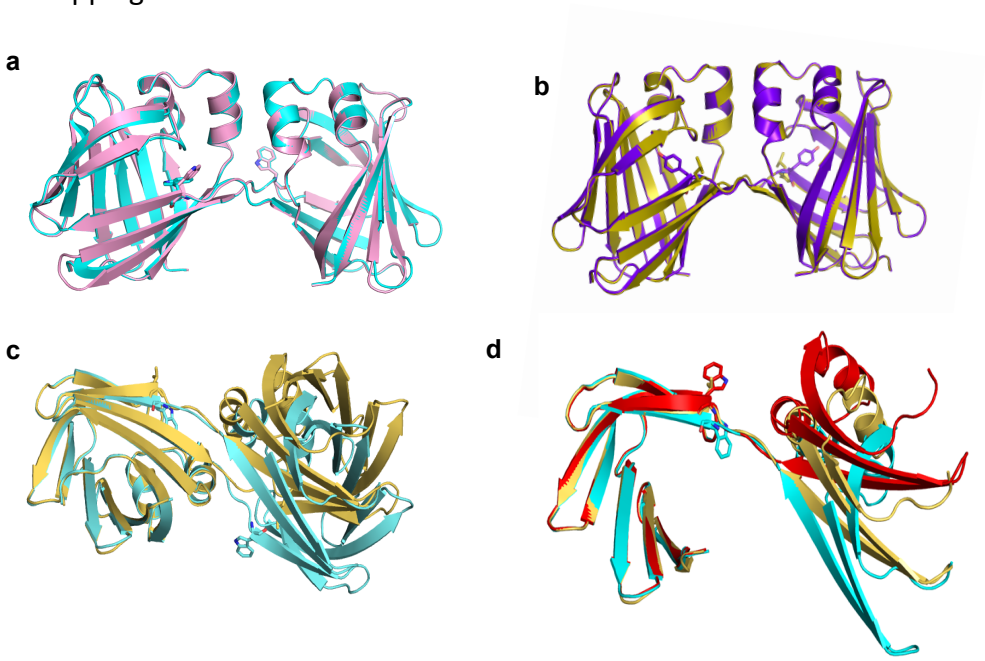


Figure II-13: The differences in the asymmetric and symmetric structures of hCRBPII dimers. Res60 in all structures is shown in sticks **a.** The asymmetric structures: The overlay structure of domain swapped dimer KLY60W (cyan) and Y60W (pink). In chain A dimer KLY60W, Trp60 is pointing inside the binding pocket. **b.** The symmetric structures. The overlay of WT hCRBPII (purple) and Y60L (yellow) domain swapped dimer. **c.** Comparing asymmetric (dimer KLY60W in cyan) and symmetric dimer (dimer Y60L in yellow) by looking down the two-fold axis and it shows the different position of helices in N-terminus. **d.** One chain of dimer Y60L (yellow) and both chains of asymmetric dimer Y60W (chain A in cyan and chain B in red) are overlaid at their C-terminus.

We found that hCRBPII DS dimers can be classified as either symmetric (WT hCRBPII, Y60L and other mutants described in II-6 section), where the two chains which make the dimer are essentially identical, or asymmetric (KLY60W, Y60W and holo DS dimers described in II-6), where the two chains are distinct from one another **(Figure II-13).** [45] As shown in **Figure II-13b**

the two symmetric structures (WT and Y60L) are similar, with the relative orientation of the two domains in the dimer essentially identical. However, there are significant differences in the torsion angles around residue 60, with an associated repositioning of the side chain. Tyr60 is buried inside of the protein, making a hydrogen bond with Glu72 in both the WT monomer and dimer, while Leu60 is flipped out of the inside of the protein and solvent exposed in the Y60L dimer **(Figure II-14a).**

This "flipped out" conformation is the cause of the residue 59 phi angle large deviation, relative to its monomer value (about -121° in the monomer versus -80° in the Y60L dimer) **(Figure II-14c)** In contrast, the phi torsion angle of Asn59 in the WT dimer is -142° (similar to its monomer value), which leads to the "flipped in" conformation (**Figures II-14a and 14b).** To balance the resulting change in trajectory of the main chain, the Asn59 psi angle is radically rotated relative to all of the other structures (-142° in the WT dimer versus 161° in the monomer and 165 in the dimer Y60L, see **Figures II-b),** resulting in a substantial Ramachandran outlier at Asn59. Nonetheless it results in the main chains of the two symmetric structures once again closely tracking each other **(Figure II-14a).**

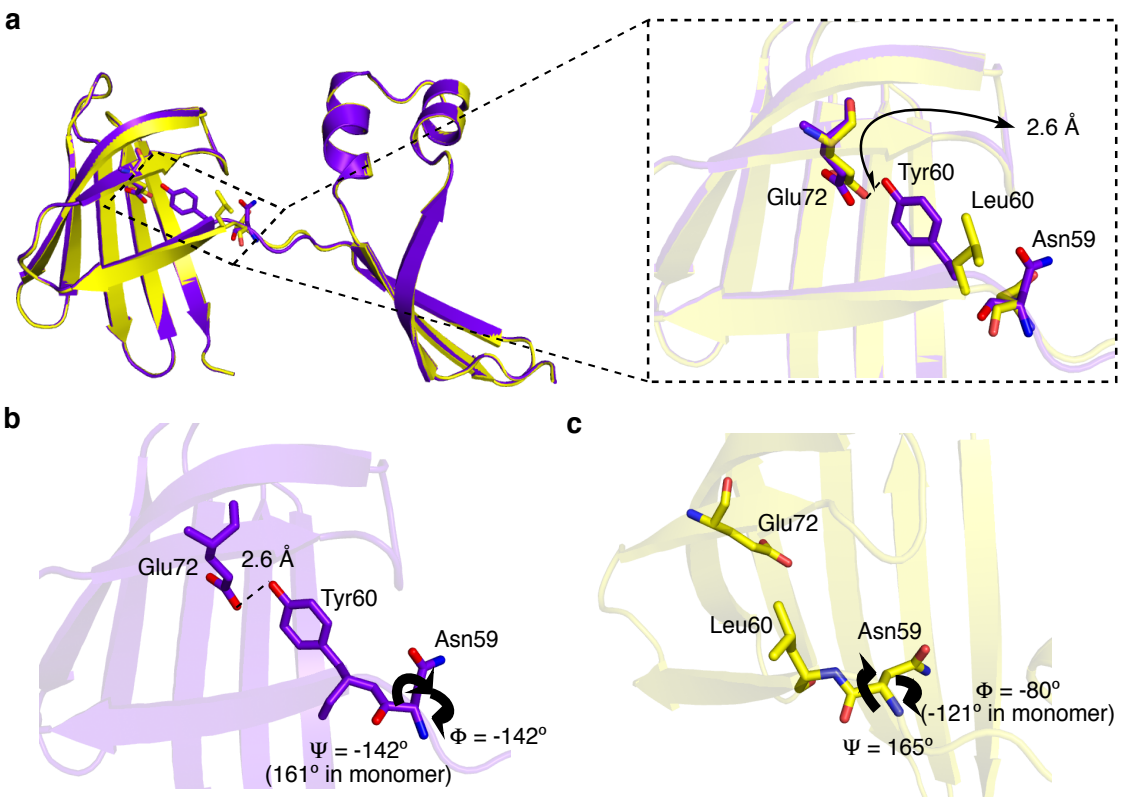


Figure II-14. Torsion angle deviations outside the hinge-loop region define the relative orientation of the two domains of the dimer, with the highest deviation seen in Asn59. **a.** The WT (purple) and Y60L (yellow) dimers are overlaid. Inset, the hydrogen bond between Tyr60 and Glu72 compared to the "flipped out" conformation of Leu60. **b.** The critical Tyr60 and Asn59 region in the WT hCRBPII dimer, showing the key phi/psi angles. **c.** The same region in the Y60L dimer, showing the "flipped out" Leu60 and the key phi/psi angles. Comparison of the two shows how the large difference in the phi angle is compensated for in the WT N59 psi angle, keeping the main chain of the two on a similar trajectory.

In contrast to the symmetric dimers, asymmetric dimer (i.e. Y60W and KLY60W) structures are similar to one another and asymmetric **(Figures II-13a).** However, there is a significant deviation in the position of the N-terminal region when the C-terminal domains are overlaid **(Figures II-13c and 13d).** Basically, in asymmetric dimers, there is a large deviation in the relative orientation of the N- and C-terminal regions in the two-polypeptide chains (subunits) that make the dimer. **(Figure II-13d)** [45]

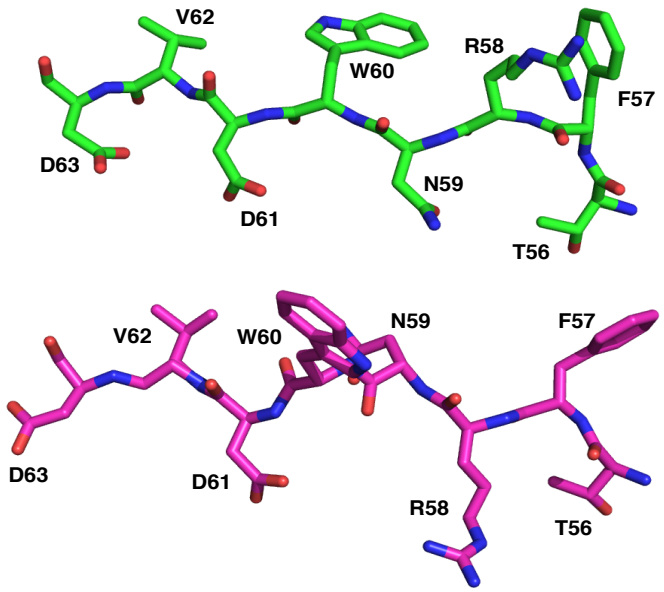


Figure II-15: Residues 56-63 in the chains A (green, top) and B (magenta, bottom) of Y60W hCRBPII.

This positional deviation is as large as 14 Å between the two conformations of the DS dimer. This represents a radical change in the relative orientation of the two domains relative to the symmetric dimer.

The source of this large deviation can be seen in the torsion angles of the residues between 56-60 **(Figures II-12).** Where in chain A of both KLY60W and Y60W dimers, all of the torsion angles, except the psi angle of Thr56, are similar to those of a monomer. In chain B a number of other torsion angles, namely the psi angles of residues Thr56 and Phe57, and the phi angle of Phe57, deviate from their monomer values **(Figures II-12a and 12b).** Presumably, these differences in torsion angles are required to orient the two halves of the chain properly for dimer formation.

Interestingly, the A subunit of the KLY60W and Y60W mutants is the most similar to the monomer. It has phi/psi angles almost identical to those of the monomer, except the psi angle of hinge residue T56, which is rotated by almost 180 degrees in all the DS dimers. Phe57 and Arg58 side chains are on the same side of the strand, similar to the conformation seen in the loop connecting the two strands in the monomer **(Figure II-12c and 15).** However, the phi/psi angles of the B subunit had other significant differences with monomer, especially the phi angle of residue 59 **(Figure II-12d),** which resulted in the flipping of Trp60 from inside the binding cavity to the outside and Asn59 from outside to inside of the cavity (relative to the monomer) **(Figure II-15).** This places both residues 60 and 61 outside the binding cavity, which is necessary to re-phase the strand (explained in the next section, II-4-2). The asymmetry of the Y60W mutant dimers tells us that the "canonical" DS dimer, where the only torsion angle change is in the hinge residue, is not possible because the relative orientations of the N- and C-terminal regions of each domains are incompatible with dimer formation; thus subunit B must adjust itself, via the torsion angles along the connecting strand, to adopt a conformation consistent with dimer formation to subunit A.[45] **(Figure II-16)**

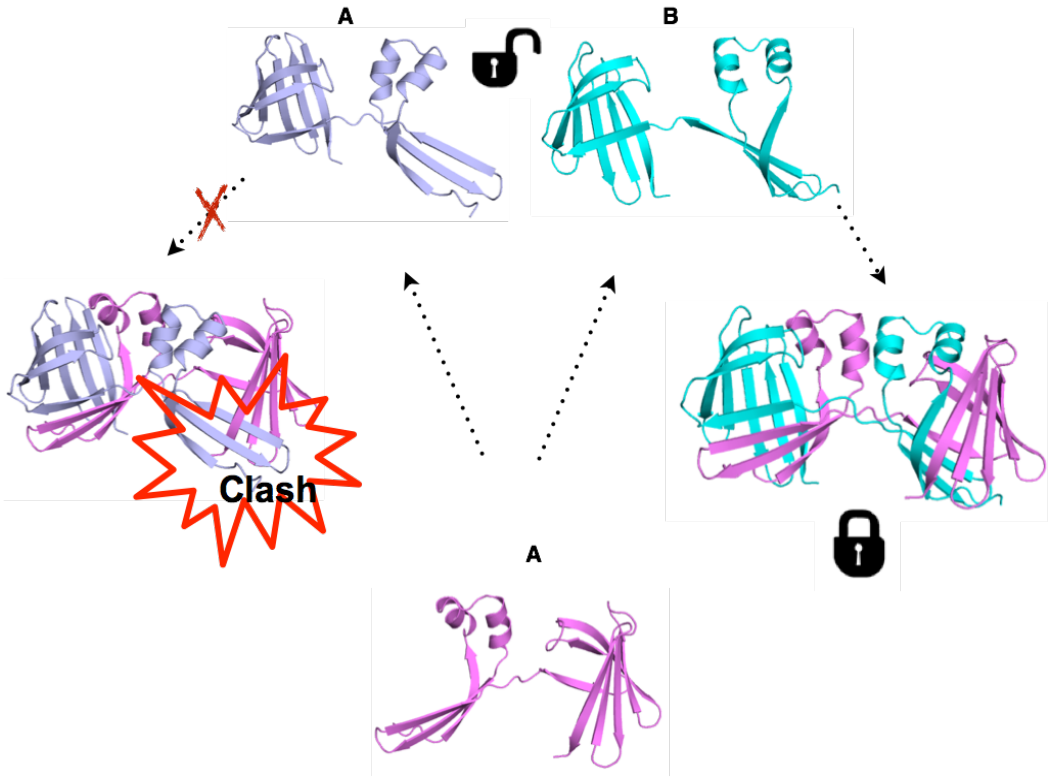


Figure II-16: Overview of dimer formation in asymmetric structures: Two chains of Y60W dimer are labeled in the picture. The same chain (molecule A and/or B) in this asymmetric dimers cannot form dimers since they clash. However, the interaction of different chain with each other (molecule A and B) could lead to proper DS dimer formation with the different orientation.

II-4-3 Phase Relationship

In all cases domain swapping is the result of the straightened connecting loop between beta strands 3 and 4 of the monomer, resulting in this loop and beta strands 3rd and 4th becoming a single beta strand stretching the length of the dimer **(Figure II-3).**

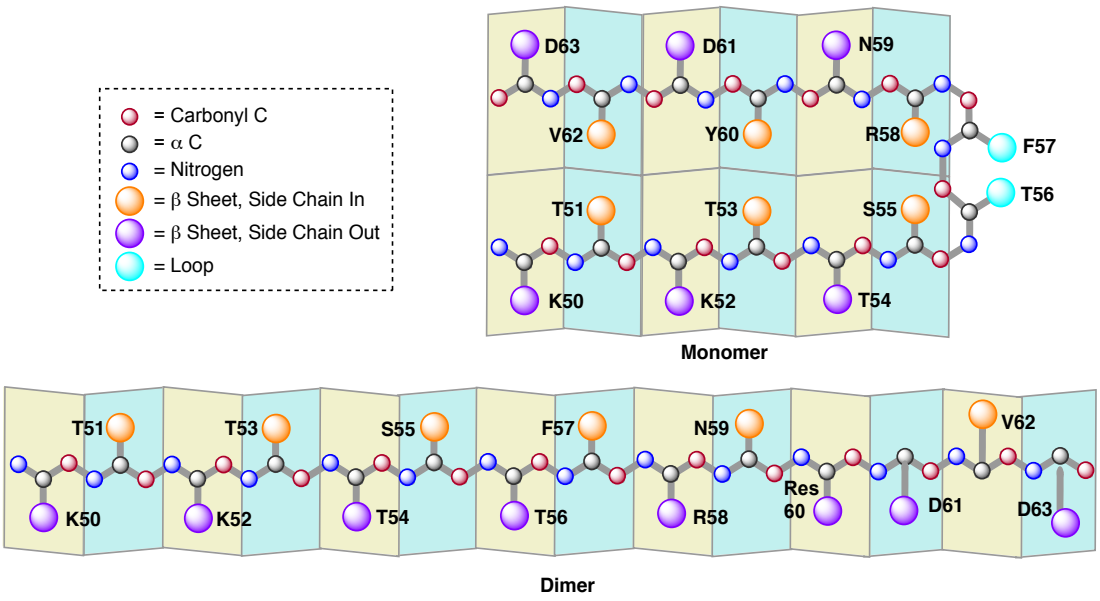


Figure II-17: Schematic representation of "phase relationship" in DS dimerization. Top: monomer, bottom: dimer with residue 60 as mutant.

In order to have a DS dimer, with a perfect stretched beta sheet in that region, eventually we will have "phase problem" because, in monomer while the *odd* side chains of beta strand 3 face inside the binding pocket, the even side chains of beta strand 4 face inside. With an even number of residues in the loop, this means that the single beta strand formed in the DS dimer must, at some point, re-phase the strand, which would then put the even side chains of the C-terminal half of the strand "in phase" with the odd numbered side chains of the first half.

The way that mutants (i.e. Y60L and Y60W) solve the phase problem might be the key point in the domain swapping mechanism. Conformations of residues Y60 and D61 are important to consider as well. In Y60L, Y60W (chain B), KLY60W (chain B) mutant domain swapped dimers of hCRBPII, these two residues are pointed toward the solvent, which re-phase the strand in order to fix the phase problem **(figure II-17).** However, in wild type hCRBPII, which gave us mostly monomer, residue Y60 is inside the binding pocket. This may suggest one of the

significant points in the mechanism for dimerization in hCRBPII.

II-5 possible mechanism for domain swapping in hCRBPII

With the existence of the WT domain swapped dimer and the fact that there is no Interconversion between monomer and dimer, indicating each as a unique fold of this protein. Therefore, we proposed a mechanism for the hCRBPII folding pathway, different from what has been seen for other iLBPs earlier[(72)]. We suggested that the two halves of hCRBPII (the N-terminal and C-terminal halves) are capable of at least partially folding independently, initially as an extended "open monomer." The dimer/monomer ratio would then depend on the relative rates of dimerization of the open monomers, versus rotation of the n- and c-termini together to form the "closed monomer".

Our data shows that three things are required for domain swapping in this system, the large hinge motion at Thr56, subsequent proper orientation of the n- and c-termini to adopt a conformation consistent with dimer formation and re-phasing of the connecting strand with the last two strongly correlated. **(Figure II-18)**

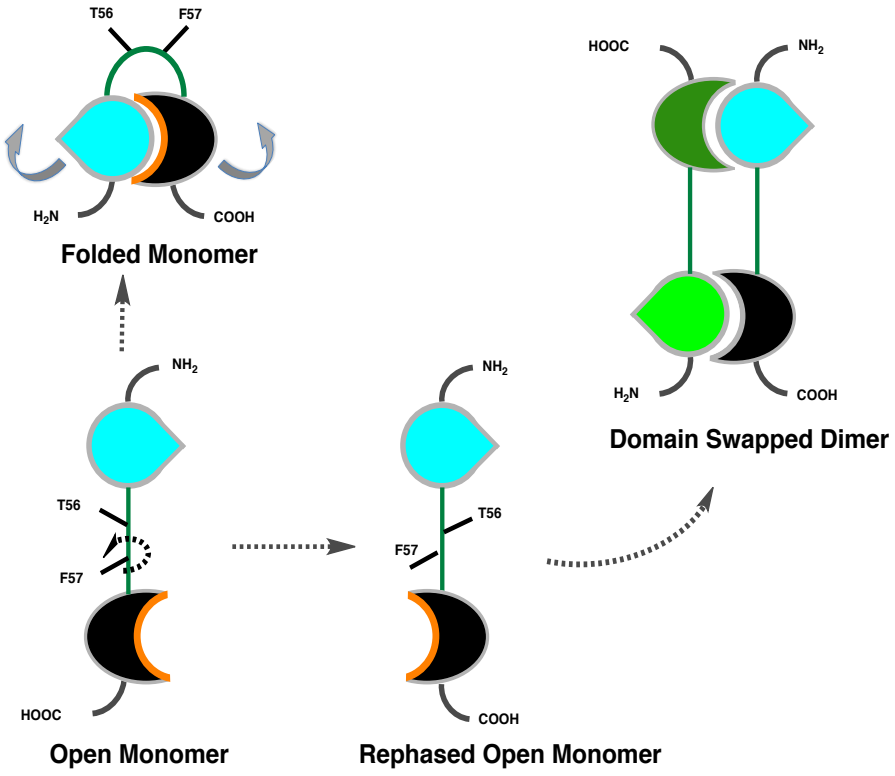


Figure II-18: DS dimerization requires: 1. Rotation about Thr56 psi. 2. Orientation of the two halves to accommodate dimerization. 3. Rephasing of the connecting strand.

As one would expect, this is a rate-governed process, thus factors such as temperature and concentration should affect the outcome. In all the symmetric dimers thus far structurally characterized virtually the same dimer is seen, where both subunits of the dimer have very similar relative orientations of their n-and c-termini, with Asn59 flipped inside the binding cavity, which is "in phase" with the n-terminal side of the connecting beta strand. **(Figure II-21)** However, reaching this symmetric dimer arrangement is accomplished in different ways. Flipping out of the residue at position 60, in both the Y60L symmetric dimer and molecule B of both asymmetric dimers **(Figures II-14a and 15)** is one low energy pathway for a "dimer friendly" orientation. In contrast WT hCRBPII is "spring loaded" against this conformation by the

hydrogen bond made between Glu72 and Tyr60, which serves to hold Tyr60 in the "flipped in" position **(Figures II-14b).**

Formation of the WT symmetric dimer then requires an almost 90° rotation of the Asn59 psi angle resulting in a substantial Ramachandran outlier **(Figure II-12d).** ("Twist against the spring" situation). This gave rise to the small quantities of WT dimer and its relative instability. Tyr at position 60 is resistant to flipping out due to the hydrogen bond with Glu72, which "spring loaded" it against dimerization, **(Figure II-19)** and therefore requiring an unfavorable psi angle in Asn59 to achieve the relative orientation required for dimerization **(Figure II- 12d)**.

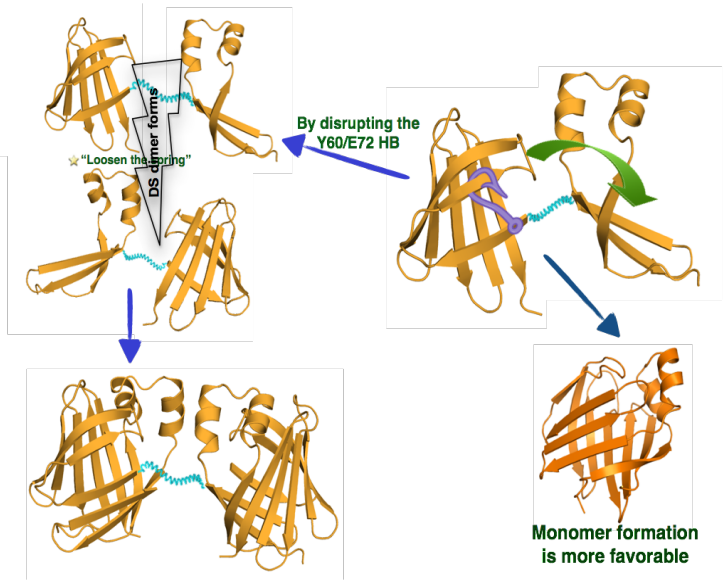


Figure II-19: DS dimer vs. monomer formation in WT hCRBP II.

In none of the other DS dimer structures we see Ramachandran outliers in the connecting strand region, indicating since Tyr60 is inside the binding pocket and making its hydrogen bond with Glu72 results in a relatively high energetic penalty for dimerization, due

both subunits to the need to rephase the sidechains on the connecting strand, and to allow the n- and c-termini to adopt a conformation consistent with dimerization.

Based on this hypothesis mutation of E72 to a residue incompetent for hydrogen bonding (a small hydrophobic residue) will "loosen the spring" resulting in an increase in dimer formation. **(Figure II-19)**

II-5-1 Dimerization and structure of the E72A mutant

As a first step to confirm and test the abovementioned hypothesis, the E72A mutant was created, expressed and purified. As predicted this mutation led to a dramatic increase in dimerization relative to WT hCRBPII. (E72A mutant produced 50% dimer vs. only10% in WT) **(Table II-9, Figure II-20).** The E72A mutant was then crystallized and its structure determined.

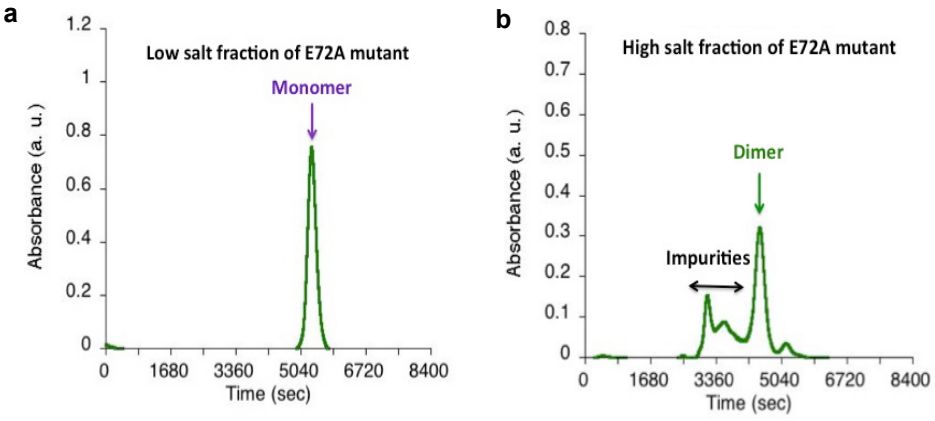


Figure II-20: Size exclusion chromatography of low salt and high salt (80 mM and 160 mM NaCl,) E72A hCRBPII respectively, expressed at 25°C (since 30°C expressed produced virtually no soluble protein).

Although a crystallographic two-fold axis does not relate the two monomers in this dimer, the relative orientation of the N- and C-terminal halves of the protein are essentially identical to that of the symmetric dimers **(Figure II-21C).** There are deviations in its torsion

angles at position 57 (phi angles of -85° versus -133) and most notably, position 60 (phi angles of 35° versus -72° and psi angles of 92° versus 127) **(Figure II-12b and e).** The result is that Tyr 60 is in the "flipped out" conformation, similar to that seen in both the Y60L dimer and no Ramachandran outliers in the connecting strand residues **(Figure II-21a),** thus inducing the flexibility between the n- and c-terminal domains (compare to WT, **Figure II-21b**), favoring dimer formation, and suggesting that the hydrogen bond between Tyr60 and Glu72 is there to repress dimer formation and consistent with the hypothesis that the interaction between Y60 and E72 in the open monomer folding intermediate effectively holds the two domains in a conformation unfavorable to dimerization, promoting the formation of the physiologically relevant monomeric species. This suggests that the n- and c-terminal domains are capable of at least partially folding in isolation, and would represent the "foldons"[(85-89)] for hCRBPII. **(Figure II-18 and 19)** Note that Tyr60 is not conserved in the iLBP family, and is found only in CRBP's.

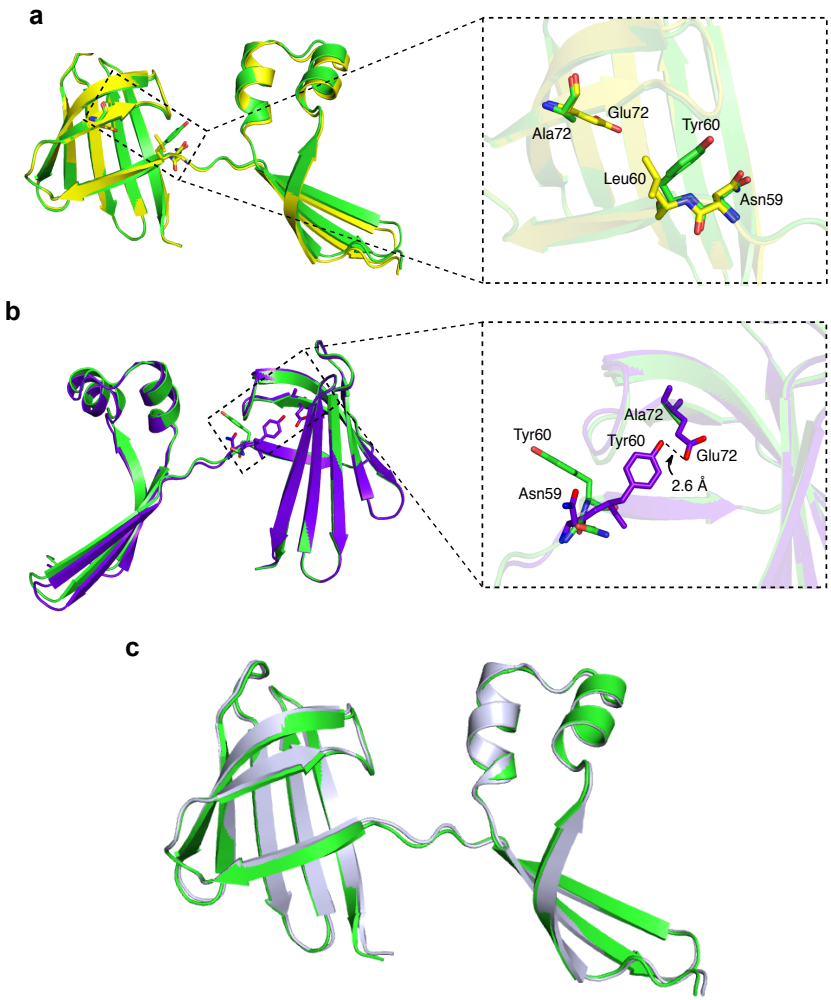


Figure II-21: a. Overlay of chain A of the E72A dimer (green) and the Y60L dimer (yellow). Note that both show residue 60 pointing toward the solvent. **b.** Overlay of the E72A chain A (blue) and WT (green) dimers. Inset shows that without the Glu 72 hydrogen bond donor, residue 60 is free to flip out, leading to a more relaxed dimer structure. **c.** Two chains of the E72A dimer (green and grey) are overlaid, showing their strong similarity.

II-5-2 Study the folding pathway

As mentioned above, we proposed that the N and C terminus of the protein might fold independently (at least partially) to form an open monomer folding intermediate. Then, the ratio of monomer/dimer would depend on the relative rates of dimerization of open monomers, versus rotation of the N- and C-termini together to form the “closed monomer”. It

was seen before that this is concentration dependent **(Table II-1)** and would depend on the energetics of proper orientation for dimer formation.

In this Scenario, the DS dimer would then represent the kineticly trapped folding intermediate on the native folding pathway of hCRBPII. One of the early steps to detect and study the folding intermediates is to monitor the protein at different denaturant concentrations, with the assumption that the structures formed at intermediate denaturant concentrations will reflect the intermediates formed in the folding pathway.[90, 91] To this end, we tried to probe the structure of WT-hCRBPII over a range of denaturant concentrations (Gd-HCl in this case) using both Circular dichroism spectroscopy and tryptophan fluorescence spectroscopy.[1] **(Figure II-22)** We initially were interested to plot the change in ellipticity versus denaturant, **(Figure II-22c)** to compare it to a two-state ($N \rightleftarrows U$) and three state ($N \rightleftarrows (1/n)I_n \rightleftarrows U$ or even higher order model) which would suggest the presence of intermediates (excluded the two-state model) in the folding pathway[92]. It should be also considered that the two halves of the protein may have significantly different stabilities. After performing these series of experiments, the helix/strand ratio was calculated from the CD spectra at each individual denaturant concentration and compared to the secondary structure of the native and predicted intermediate (which almost consists of the same percentages of helix/strands as native) **(Figure II-19 and** Table II-8**).**[1] Also the concentration dependence may suggest that dimer association occurs early in the protein-folding pathway. Based on the result from our CD spectra and calculated percentages of helix and beta sheet at different Gd.Hcl concentrations, helixes may be more stable compare to beta sheets. From 1M of denaturant concentration, the percentages of beta sheets decrease to 30% while for helixes, even at 2M still there are intensities relative

to α-helix signal and after 4M it drops to about 20% of the native (Table II-8), proposing that N-terminus maybe (partially) fold first. These calculations are all considered with the respect of 220nm regarding beta sheets and 222nm/208nm for helixes.

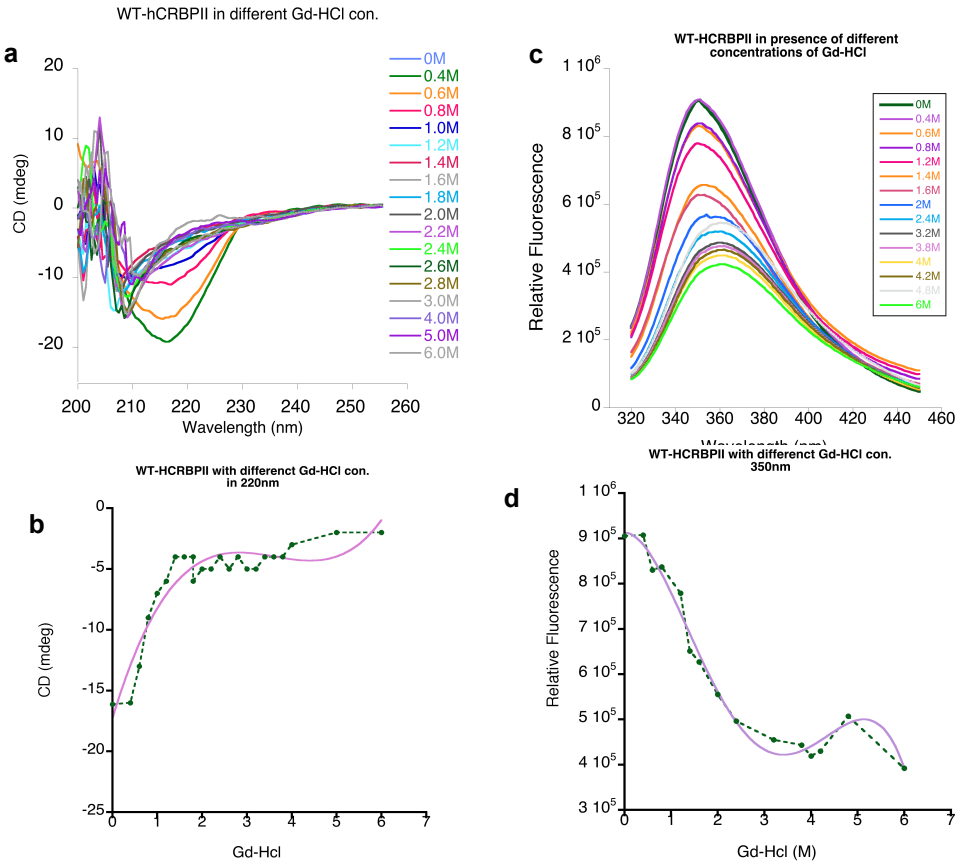


Figure II-22: Equilibrium Unfolding Experiment for wild-type CRBPII in presence of different concentration of Gd-HCl, **a.-b**. monitoring by Circular Dichroism (CD). **c-d.** Tryptophan fluorescence spectroscopy, right, performed with WT hCRBPII, to study the presence an intermediate in the folding pathway. Purple solid line represents the polynomial fit (3rd degree) in b and d. Both b and d spectra are in correlate with previous studies. *(1)*

Also, based on the previous studies on analysis of three-state protein unfolding data, our CD spectra of unfolding supports more three-state model than two state. [(1)] Even, Prof. Lapidus have tried to fit our data into two-state model and it did not fit correctly. However further analysis, especially with three-state model fit, need to be done in order to test this hypothesis. Same set of experiments can be done with the dimer favoring mutants (i.e Y60L), and conduct them at various monomer-favoring and dimer-favoring concentrations. Observing a similar intermediate in the folding of these mutants confirms our mechanism even though the folding product is different and it would allow us manipulating the folding pathway to control the folding product for this protein, and potentially other members of the iLBP family as well.

Completing mentioned calculations requires more recourses period, however during the course of my graduate research, were unable to complete the mentioned task. Though we are optimistic this project will carry on

II-5-3 HCRBPII folding route VS other iLBP members

A number of folding studies have been conducted on a few other iLBP family members, namely hCRBPI and human cellular retinoic acid binding protein I (hCRABPI). In the latter case, a series of experiments have indicated that the folding pathway of hCRABPI involves an early barrel closure in the folding pathway, even before substantial secondary structure is evident[(67-71)].

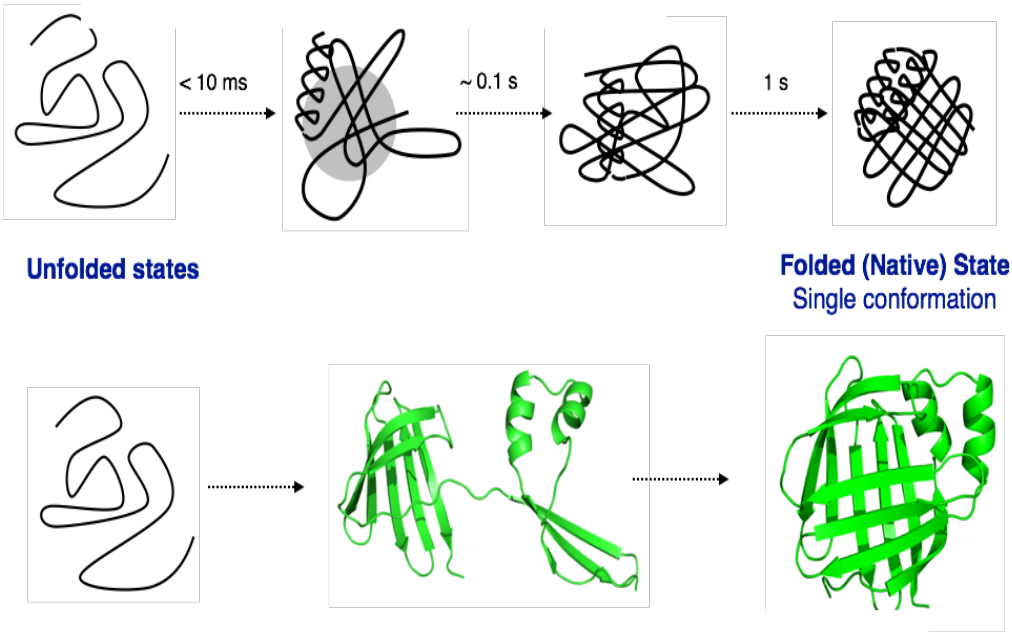


Figure II-23: overview picture of hCRABPI folding pathway compare to hCRBPII.

This prevented the protein from forming intermediates that were prone to aggregation and potentially fibril formation. This mechanism is clearly distinct from what we proposed for hCRBPII. However, there is no guarantee that all members of a structural family would follow the same folding pathway. Recently, the crystal structure of a FABP5 DS dimer was reported. The swapped region was almost identical to that of hCRBPII.[(93)] (Refer to next Chapter, IV) This shows that other family members ((with only 36% sequence identity between FABP5 and hCRBPII) may also have a propensity to form the same DS dimer. The identification of a second, DS dimer iLBP family member lead us to the idea *that a subset of iLBP family members* may indeed be DS dimers in their physiologically relevant forms.

II-6 Conformational change driven by ligand binding in hCRBPII

It was previously mentioned that the cellular retinol binding proteins (CRBP's) are involved in the trafficking of both retinol and retinal within the cell. Binding studies of hCRBPII DS dimers

have demonstrated that DS dimers are competent for ligand binding.[45] **(Figure II-9)** Knowing these, in continues collaboration between our lab (prof. Geiger group members) and Prof. Borhan group; we successfully expressed, purified and determined structures of multiple variants of holo hCRBPII domain swapped dimer mutants (up to 90% dimer formation) bound Protonated Schiff Base, PSB, to retinal as one of its natural ligand and in the absence of ligand in that series. [74-76]c (**Figure II- 25, II-28 and II-29)** Q108K:T51D holo structure obtained by Dr. Nosrati and apo form obtained by Alireza Ghanbarpour. **(Table II- 3)**

We found that ligand binding in dimers leads to a reproducible and noticeable conformational change. Among those, all of the apo structures reveal a symmetric dimer very similar to the other symmetric dimers previously mentioned in section II-4-2 **(Figure II-24),** though some of them might differ in their crystal packing. Therefore, the data indicate the symmetric dimer to be the most common and robust form of DS dimer structure, with the asymmetric dimer a special case caused by the Y60W mutation.

Table II-3: Summary of DS dimer crystal structures of all hCRBPII mutants

Mutants	Holo/Apo	# molecules per asymmetric unit	Symmetric/ asymmetric
Q108K:T51D	Apo	1	*Symmetric*
Q108K:T51D	Holo	12	*Asymmetric*
Q108K:K40D	Holo	12	*Asymmetric*

On the other hand, structures of the retinylidene-bound dimers reveal different relative orientation of the two domains of the dimer in them when compared to the apo symmetric dimers (**Figure II-25)**.

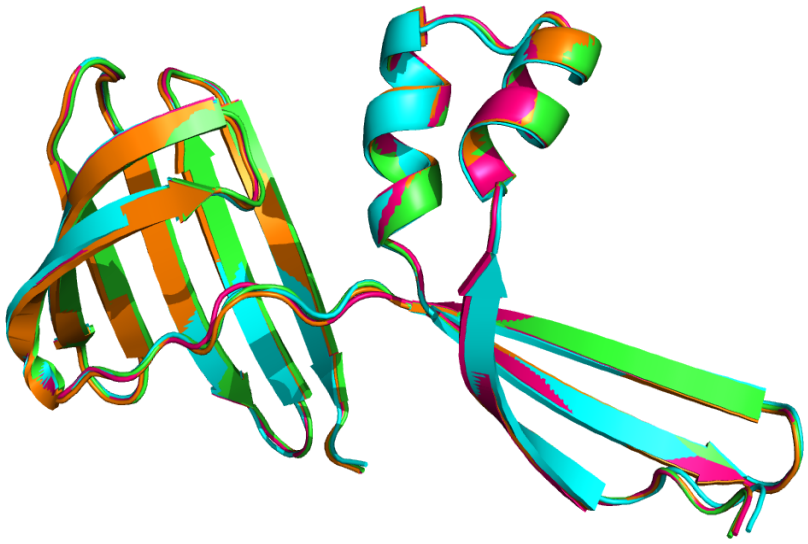


Figure II-24: A comparison of four symmetric dimers, Y60L (orange), WT (red), E72A (cyan) and Q108K,K40D (green).

Further, most of the holo dimers crystallize with between 4 and 12 monomers in the asymmetric unit, all of which exhibit small differences in relative orientation of the two domains that make up the DS dimer. This indicates that, in contrast to the apo DS dimer, which seems to exist as a relatively rigid structure, holo dimers have more flexibility along their

strand. However, nearly all dimer molecules of holo structures (i.e. 6 pairs in Q108K:K40D) are pretty similar to one another. **(Figure II-28a-f)**

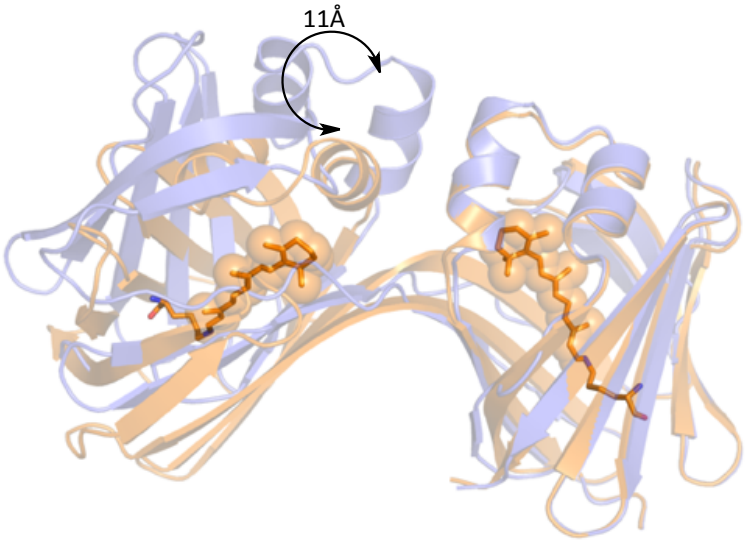


Figure II-25: Overlay of holo Q108K:K40D (chains F and A, both shown in orange) and apo Q108K:T51D (shown in blue) showing the large motion of helix 1 upon ligand binding (about 15 Å). Bound retinal molecules in holo structures are shown as transparent spherical models. Note, complete dimer in the apo structures was generated by crystallographic two-fold symmetry operation.

The major cause of this large relative deviation in two domains is related to the position Thr59 and Tyr60. In all of the mentioned apo symmetric dimers, the Asn59 side chain is flipped into the binding cavity and the residue 60 is flipped out of the binding cavity (except the WT-hCRBPII dimer has Y60 flipped in form).[(94)] This is in contrast to every hCRBPII monomer (with over 40 structures so far determined), where Asn59 is always outside the binding pocket. In holo DS dimer structures the steric bulk of the ligand forces Asn59 to rotate out of the binding pocket, resulting in the rotation of the entire domain. (**Figure II-26a and b),** and Tyr60 is found inside the binding pocket.

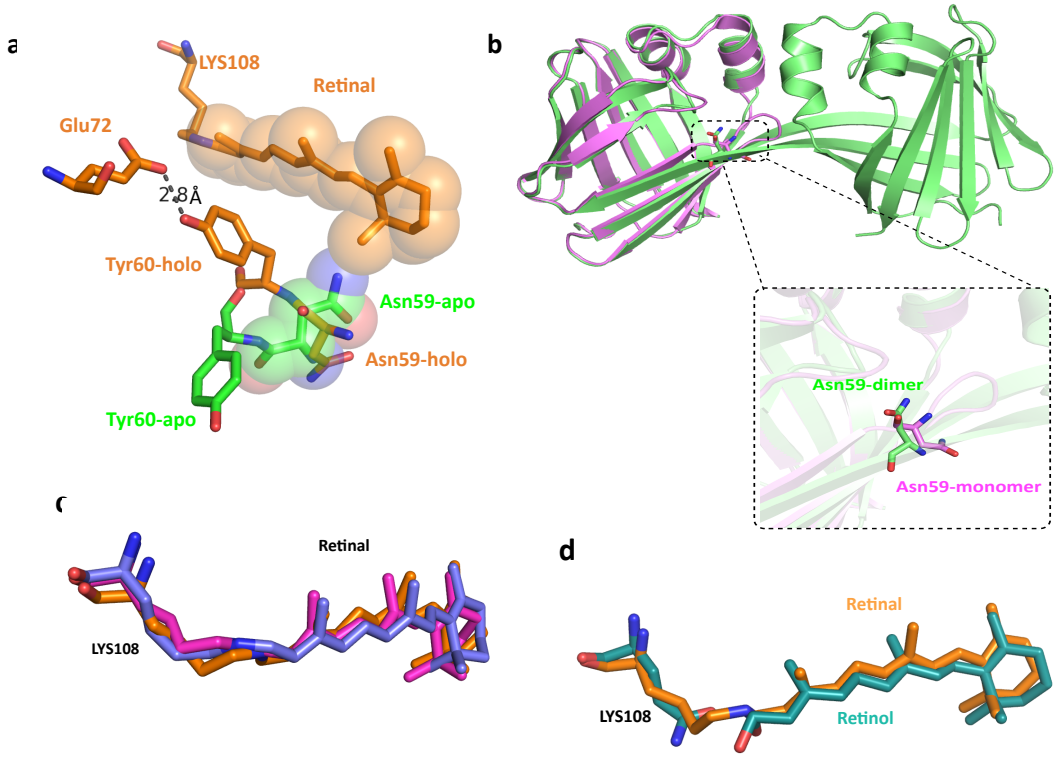


Figure II-26: Ligand binding leads Asn59 to flip out of the binding pocket due to steric effect. **a**. Overlay of Q108K:T51D-apo (green, obtained by ALireza) and Q108K:K40D-holo (orange). Allosteric conformational change between apo and holo dimers is caused by the orientation of the Tyr 60 and Asn 59 sidechains. Asn59 would sterically clash with the bound retinylidene, which leads to the flipped-out conformation of Asn59 and flipped-in conformation of Tyr60. **b.** Overlay of WT apo monomer (pink, PDB code: 2RCQ) and WT apo hCRBPII dimer (green, PDB code: 4ZH9). The "flipped in" Asn59 conformation is required to adopt the orientation to form the symmetric DS dimer. **c.** The similar trajectory of retinal in monomer and dimer structures of hCRBPII mutants. Q108K:K40L monomer (purple, PDB code: 4RUU) Q108K:K40D dimer (orange) and Q108K:K40L:T51DD dimer (pink, obtained by Dr.Nosrati) which are all bound to retinal via a covalent bond to Lys108. **d.** Overlay of WT monomer hCRBPII bound to retinol (green, PDB code: 4QZT) with retinal-bound Q108K:K40D dimer (orange).

As shown **(Figure II-26c and d),** the trajectory of the retinylidene through the dimer binding pocket is very similar to that seen in both retinol-bound and retinal PSB-bound monomers, indicating that dimers bind ligands very similar to that of the monomer structures, whether the ligand is a retinol or retinal.

Interestingly, other than the bound retinyldine mutants, the conformational change of the two domains can be observed for WT-hCRBPII DS dimer bound with retinal (obtained by Alireza Ghanbarpour) as shown in **Figure II-27**. Probably due to the low resolution of this structure (about 3.4 Å) the electron density of the ligand is weak. However, a strong hydrogen bond between Tyr60 and Glu72 is apparently required for the conformational change (similar to that seen in apo WT-hCRBPII, PDB ID 4ZH9) in all holo DS dimers **(Figure II-26a and 28g-l) T**o test this new hypothesis around Tyr60, Q108K:K40L:51F:Y60A mutant structure was determined by Alireza Ghanbarpour.and showed Thr 59 position pointing into the binding pocket.(Data not shown)

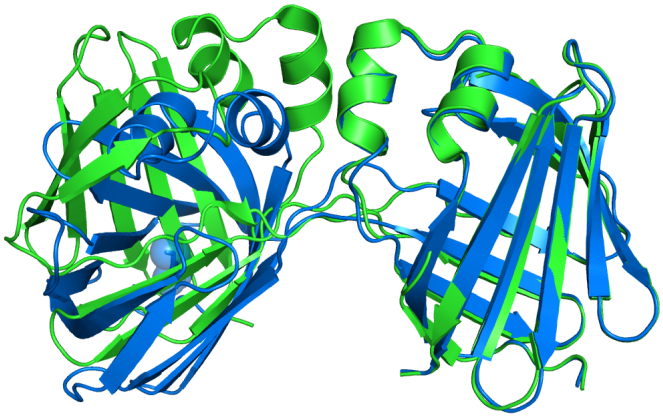


Figure II-27: Overlay of WT-hCRBPII (holo) DS dimer shown by dark blue and WT-hCRBPII (apo) (PDB code:4ZH9) shown by green.

This reveals the essential role that Tyr 60, and presumably Asn59, play in the mechanism of allosteric conformational change seen in hCRBPII domain swapped dimers. During my graduate work, majority of mutants and analysis were done on Tyr60 however, more analysis needs to be carried on Thr59 in the future.

Finally, this observation shows DS dimers could potentially be allosterically regulated during ligand binding, in a way that is not available to iLBP monomeric structures, which invariably lead to only modest changes in conformation[95]. In fact allosteric regulation via ligand binding has been suggested to be important for the function of a number of iLBP family members (FABP4 and CRABPI and etc).[96] [97, 98]

Meanwhile, the identification of several mutational "hot spots" and the apparent ease at which the dominant folding product can be dramatically altered in hCRBPII have led us to ask whether the physiologically relevant structure for iLBP family members might actually be the DS dimer. As the other major goals of this project was to use the information gained regarding the mechanism and determinants of DS dimer formation to identify other family members of such a subset. The DS dimer would certainly have distinct and important characteristics *in vivo*, including its ability to carry two ligands simultaneously, which will be described in more detail in the next chapter.

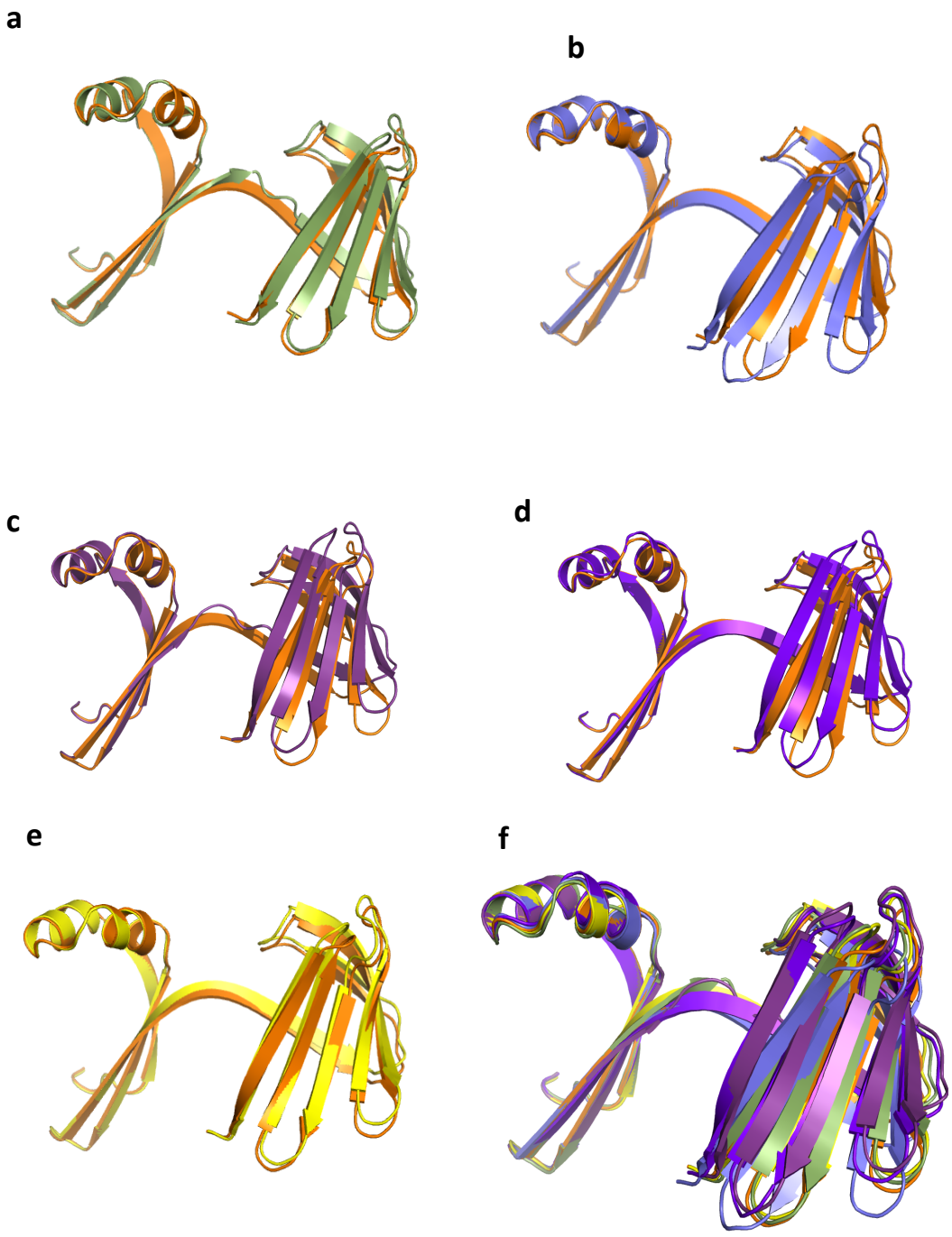
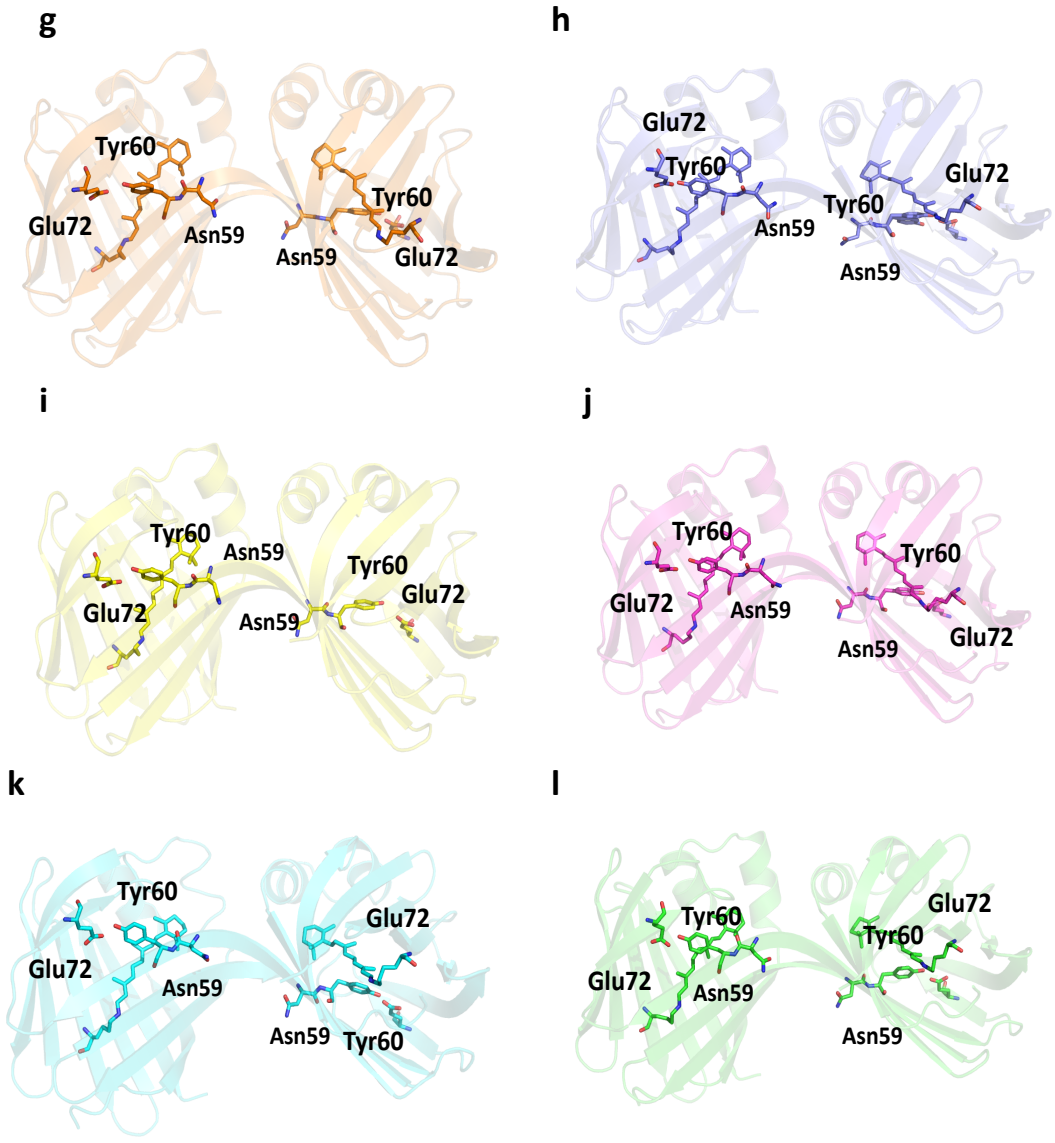


Figure II-28: Overlay of different chains in holo Q108K:K40D DS dimer. **a.** Chains A shown by orange and I shown by green. **b.** Chains A shown by orange and L shown by blue. **c.** Chain A shown by orange and D shown by pale purple **d.** Chains A shown by orange and G shown by purple**. e.** Overlay of chains A shown by orange and B shown by yellow. **f.** Overlay of chains A shown by green and G shown by cyan. g-l: E72…Y60 Hydrogen boning motifs in Q108K:K40D holo DS dimer: **g.** Chains A and F **h**. Chains D and H. **i.** Chain L and J. **j.** Chains B and E **k.** Chains K and I. **l.** Chains C and G

Figure II-28 (cont'd)



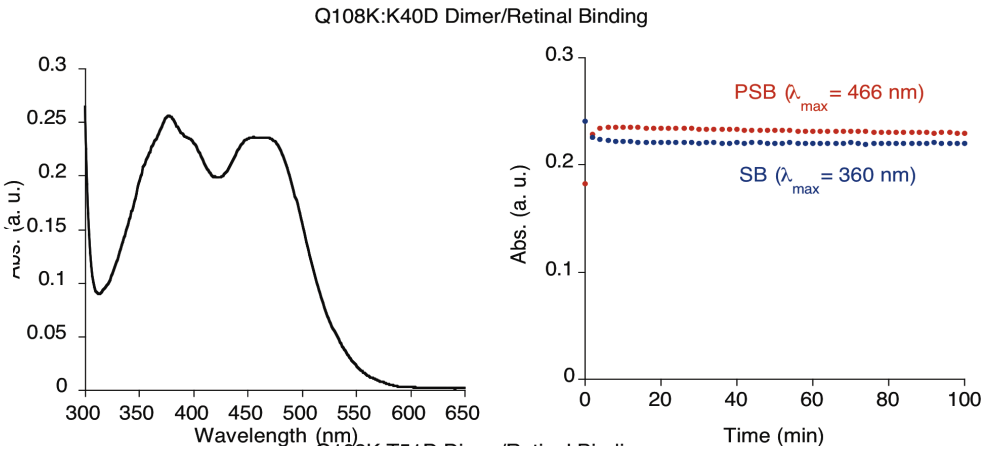


Figure II-29: UV data spectrum for different dimer mutants with *all-trans* retinal was carried by Dr. Santos. Left, UV spectra represents the kinetically formed product between the hCRBPII dimer and retinal. Right, this panel depicts the formation of both SB and PSB as a function of time.

II-7 Experimental

II-7-1 Material and Method: Site-Directed Mutagenesis

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

Table II-4: PCR protocol for hCRABPII and hCRBPII mutagenesis.

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

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

The PCR product (50µL) was transformed into DH5alpha cells (50µL) and grown on Luria-Bertani (LB)-agar plates supplemented with Ampicillin (75 g/mL) for 15h. A single colony was inoculated in 15 mL LB medium containing 100mg/mL ampicillin and grown at 37°C while shaking, for 12 hours. DNA purification was performed using a QIAGEN Miniprep DNA purification kit. The average isolated plasmid DNA concentration was 120ng/ l in 30 µL in solution. The DNA sequence was verified by the MSU gene sequencing facility using T7 primer.

Primers

Y60W
Forward: 5’- CACATTCCGCAACTGGGATGTGGATTTCAC-3’
Reverse: 5’- GTGAAATCCACATCCCAGTTGCGGAATGTG-3’

Y60F
Forward: 5’- CACATTCCGCAACTTTGATGTGGATTTCAC -3’
Reverse: 5’- GTGAAATCCACATCAAAGTTGCGGAATGTG-3’

Y60L
Forward: 5’- TTCCGCAACCTTGATGTGGAT-3’
Reverse: 5’- ATCCACATCAAGGTTGCGGAA-3’

E72A
Forward: 5’- AGAGTTTGACGCGTACACAAAGAG- 3’
Reverse: 5’- ACTCCAACAGTGAAATCC- 3’

Y60I
Forward: 5’- CATTCCGCAACATTGATGTGGATTTC-3’ Reverse: 5’-
GAAATCCACATCAATGTTGCGGAATG -3’

Y60T
Forward: 5’-CATTCCGCAACACTGATGTGGATTTC-3’
Reverse: 5’-GAAATCCACATCAGTGTTGCGGAATG-3’

Y60E
Forward: 5’-TTCCGCAACGAGGATGTGGATTTC-3’ Reverse: 5’-GAAATCCACATCCTCGTTGCGGAA -
3’

Y60D
Forward: 5'- CATTCCGCAACGATGATGTGGATTTC 3'
Reverse: 5'-GAAATCCACATCATCGTTGCGGAATG -3'

Y60Q
Forward: 5'-CATTCCGCAACCAGGATGTGGATTTC-3'
Reverse: 5'-GAAATCCACATCCTGGTTGCGGAATG-3'

Y60A
Forward: 5’-CATTCCGCAACGCTGATGTGGATTTC -3’ Reverse: 5’-
GAAATCCACATCAGCGTTGCGGAATG-3’

T56 DELETION

Forward: 5’ CTTCAAGACAAAAACCACTAGCTTCCGCAACTGGGATGTGGAT -3’
Reverse: 5’- ATCCACATCCCAGTTGAGGAAGCTAGTGGTTTTTGTCTTGAA-3’

T56 Deleteion-Y60L
Forward: 5’- AAGACAAAAACCACTAGCTTCCGCAACCTTGATGTG-3’
Reverse: 5’-CACATCAAGGTTGCGGAAGCTAGTGGTTTTTGTCTT-3’

T56P
Forward: 5’- ACCACTAGCCCATTCCGCAAC-3’
Reverse: 5’- GTTGCGGAATGGGCTAGTGGT-3’

F57G
Forward: 5’- ACCACTAGCACAGGCCGCAACCTTGAT-3’ Reverse: 5’-
ATCAAGGTTGCGGCCTGTGCTAGTGGT-3’

S55W
Forward: 5'-GACAAAAACCACTTGGACATTCCGC-3' Reverse: 5'-

GCGGAATGTCCAAGTGGTTTTTGTC-3'

II-7-2 Protein Expression and Purification of DS dimer mutants

The protein expression and purification were performed as previously mentioned in II-3-1 with the following modifications: To increase the dimer/monomer ratio, the induction temperature was increased from room temperature to 30^{o}C for overnight growth. In the first step of purification (Q Sepharose Fast Flow resin (GE Health Sciences)) the bound protein was eluted with 10 mM Tris, 150 mM NaCl, pH = 8.0, the pure fractions desalted by dialysis overnight against 10 mM Tris pH=8.0 buffer, and loaded on a second anion exchange column (15Q, GE Health Sciences, BioLogic DuoFlow system) by a gradient run **(Table II-5).** The low and high salt fractions were separately concentrated at 4°C to 5-8 mg/mL and further purified by size exclusion chromatography (Superdex 200 16/75 column, GE Health Sciences, packed in-

house) in a buffer containing 10 mM Tris, 150 mM NaCl, pH = 8.0.

Table II-5: Anion Exchange purification protocol for hCRBPII. The buffer used for above protocol is 50 mM Tris, pH is adjusted automatically. The proteins elute between 4% -8 % 2M NaCl.

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

II-7-3 Protein Refolding

The monomeric protein fraction was denatured by addition of urea to 8 M at 4 °C and refolded by dialysis against 10 mM Tris pH = 8.0 with a Slide-A-Lyzer® Dialysis Cassette (3,500 MWCO cut off, Thermo Fisher Sci.). The renatured protein was concentrated and analyzed by size exclusion chromatography. [45] **(Table II-1)**

II-7-4 Thermal Melting Curves for monomer and dimer mutants.

10 μM protein (in 50 mM phosphate buffer, pH = 7.5) was heated and monitored by CD spectroscopy using a JASCO (J-810) Spectropolarimeter connected to a NESLAB RTE-111 bath circulator. The temperature was increased by 5°C for every measurement and the CD was measured. The curves were fitted based on the Levenberg-Marquardt formula using OriginLab (OriginLab, Northampton, MA, **Figure II-6 and 31)** *(99, 100)*

Final concentration of protein for folding studies was 15 μM in folding studies experiments to see a resalable signal between 220-230nm regarding beta sheets however for fluorescence studies were at 9μM. (Section II-5-2) Excitation wavelength was at 280nm and emission wavelength between 320-450nm. **(Figure II-22)** All of those mentioned experiments were done in Prof. Borhan and Prof. Lapidus (physics department) labs.

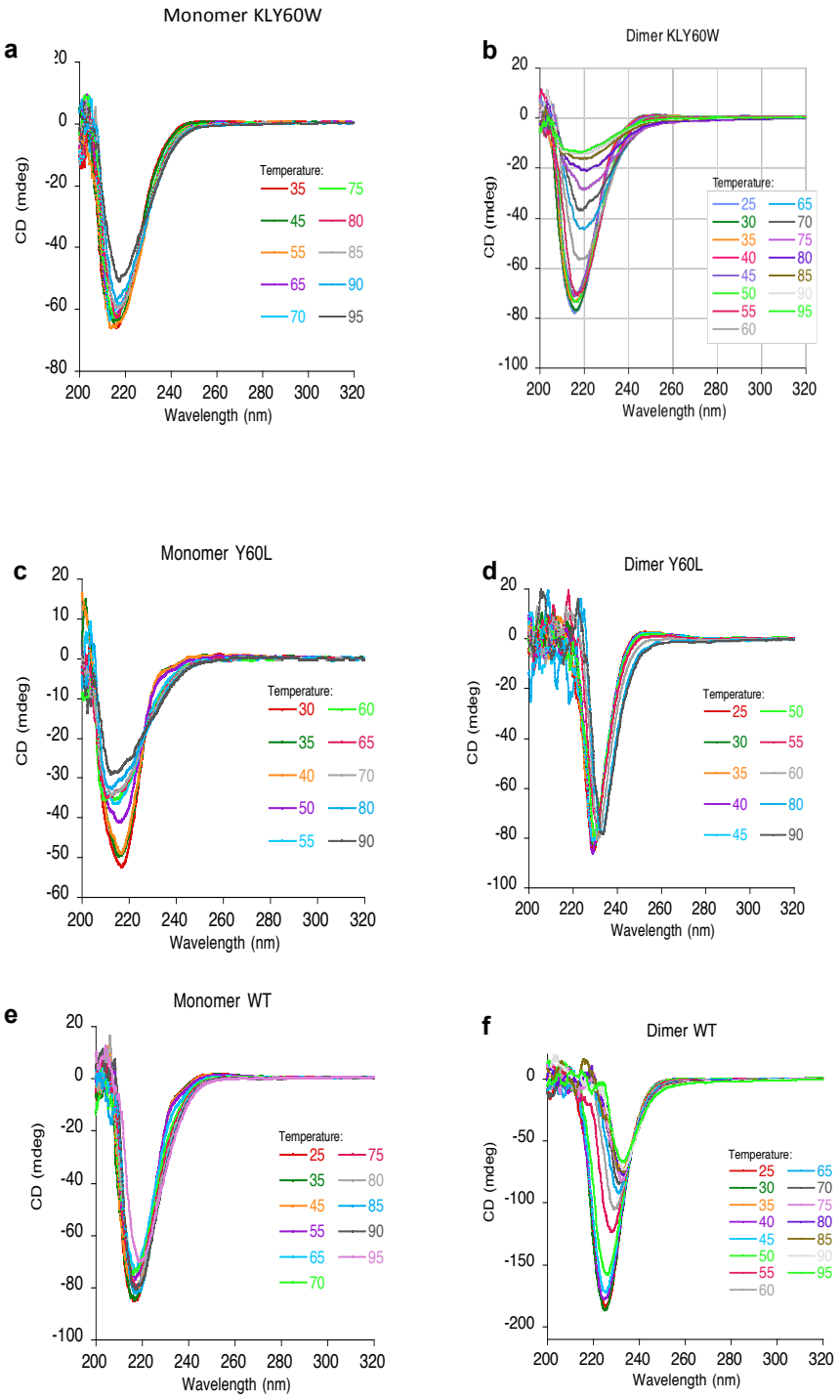


Figure II-31: Circular Dichroism spectra of various hCRBPII mutants as a function of temperature. **a.** Monomer KLY60W. **b.** Dimer KLY60W. **c.** Monomer Y60L. **d.** Dimer Y60L. **e.** Monomer WT hCRBPII. **f.** Dimer WT hCRBPII. Note, most of the secondary structure is preserved in both the monomers, even at the highest temperature achieved. Thus melting curves can only give a lower limit for the T_M, consistent with the view that the monomers are more stable than the dimers.

Table II-6: Calculated percentage of helix/strand for each denaturant (GD.HCl) concentration during unfolding experiment. 220nm corresponds to beta sheets. 222nm/208nm corresponds to α-helixes

Concentration of denaturant	mdeg at 220nm	% of mdeg at 220nm	mdeg at 222nm/ 208nm	% of mdeg at 222nm/208nm
0M	-20	100	-12/-14	100%
0.6	-17	85	-11/-11	100/78
0.8	-14	70	-8/-9	66/64
1.0	-7	35	-6/-10	50/71
1.2	-6	30	-5/-12	41/85
1.4	-4	20	-3/-9	25/64
2.0	-4	20	-3/-11	25/78
2.8	-5	25	-4/-14	30/100
4.0	-4	20	-3/-7	25/50
6	-2	10	-2/-4	16/26

II-7-5 K_d determination

The dissociation constant (K_d) of both monomer and dimer species with all-*trans*-retinal and all-*trans*-retinol were determined by fluorescence titration as previously explained in chapter I.

II-7-6 Extinction Coefficient Determination

The absorption extinction coefficients for the various CRBPII mutants were determined according to the method described by Gill and von Hippel [(101)].

II-7-7 Crystallization, data collection and refinement

Each mutant protein was concentrated to between 4 to 6 mg/mL in a buffer containing 10 mM Tris, 100 mM NaCl, pH = 8.0. Crystals were grown using the hanging drop vapor diffusion method using 1 µL of protein solution and 1 µL of crystallization solution in the drop and 1 mL of crystallization solution in the reservoir. All of the crystallization trials with retinal

were wrapped in aluminum foil to protect the light sensitive chromophore. The best crystals grew using a crystallization solution of 25-30% PEG4000, 0.1 M $NaCH_3CO_2Na$ pH 4.2, 0.1 M $NH_4CH_3CO_2$. The crystals were soaked in a cryoprotectant solution (30% PEG4000, 0.1 M sodium acetate pH = 4.0, 0.1 M ammonium acetate with 15% glycerol), flash frozen in loops in liquid nitrogen and stored in a liquid nitrogen Dewar prior to data collection.

Diffraction data were collected at the Advanced Photon Source (APS) (Argonne National Laboratory IL) LS-CAT, (sector 21-ID-D,F,G) using a MAR300 detector and 1.00Å wavelength radiation at 100K. The initial diffraction data were indexed, processed and scaled using the HKL2000 software package.[32] The structures were solved by molecular replacement using MOLREP in CCP4 or PHASER in PHENIX and hCRBPII (PDB entry 2RCQ) as a search model.[22] The initial electron density map was produced by REFMAC5 in the CCP4 package or Phaser-MR in PHENIX.[33, 102, 103] Model rebuilding, placement of water molecules etc. were done using COOT.[36] The structures were refined using either the CCP4 and PHENIX program packages. The chromophore was built into the KLY60W mutant structures using the JLIGAND program to generate restraints.[104, 105] Crystallographic, atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org.

Table II-7: X-ray crystallographic data and refinement statistics for monomeric and dimeric hCRBPII mutants.

	Q108K:K40L:Y60W (monomer)	Y60L (monomer)
Space group	P1	P1
a(Å)	30.08	30.27
b(Å)	36.05	36.38
c (Å)	63.92	126.10
α(°)	90.94	85.43
β(°)	91.99	85.18
δ(°)	113.93	66.65
Molecules per Asymmetric Unit	2	4
Total reflection	124534	268817
Unique Reflection	41751	77790
Completeness (%)	92.8(68.0) [a]	94.7(94.6)[a]
Average I/σ	50.30(3.58) [a]	43.37(1.64)[a]
R_{merge}(%)	4.1(29.3)[a]	3.3(77.6)[a]
Resolution (Å) (Last Shell)	21.30-1.47(1.52-1.50)[a]	50-1.50(1.53-1.50)[a]
R_{work}/ R_{free} (%)	19.86/25.06	22.72/24.74
RMSD from ideal value		
Bond Length (Å)	0.025	0.0109
Bond Angle (°)	2.36	1.22
Average B factor	19.98	30.86
Number of water molecules	228	459
PDB ID	4ZJ0	5DG4

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

Table II-7 (cont'd)

	Q108K: K40L:Y60W (dimer)	Y60W (dimer)	Y60L (dimer)	E72A (dimer)	WT (dimer)	Q108K: K40D (dimer)
Space group	$P2_12_12$	$P2_12_12$	$P2_12_1$	$P2_1$	$P2_12$	$P2_12_12_1$
a(Å)	62.77	62.62	63.60	62.	64.0	65.47
b(Å)	109.22	109.77	60.42	72.	60.5	73.60
c (Å)	36.39	36.32	36.77	60.	36.8	351.60
α(°)	90.00	90.00	90.00	90.	90.0	90.00
β(°)	90.00	90.00	90.00	90.	90.0	90.00
δ(°)	90.00	90.00	90.00	90.	90.0	90.00
Molecules per Asymmetric	2	2	1	2	1	12
Total reflection	288797	345311	30108	3807	3389	2363956
Unique	20048	55151	21293	2724	4455	53735
Completeness	99.9(10	98.6(10	100(9	100(	100(	99.80
Average I/σ	78.16(6	66.13(7	82.03	75.1	34.6	26.2(2.7
R_{merge}(%)	4.7(46.5)[a]	5.2(31.7)[a]	4.1(69.5)[a]	5.5(84.7)[a]	5.9(73.1)[a]	3.5(30.0)[a]
Resolution (Å) (Last Shell)	50.00-1.91(1.94-1.91)[a]	50.00-1.70(1.73-1.70)[a]	50.00-1.55(1.58-1.55)[a]	50-1.77(1.8-1.77)[a]	50.00-2.66(2.71-2.66)[a]	19.92-2.60(2.64-2.60)[a]
R_{work}/ R_{free} (%)	18.00/21.84	18.17/23.94	19.91/24.57	17.94/22.41	17.57/26.28	17.66/28.05

Table II-7 (cont'd)

RMSD from ideal values					
Bond Length (Å)	0.008	0.007	0.013	0.002	0.008
Bond Angle (°)	1.11	1.03	1.28	0.51	1.24
Average B factor	28.71	34.31	34.17	36.71	55.85
Number of water molecules	352	147	462	50	241
PDB ID	4ZR2	4ZH6	5DPQ	4ZH9	5U6G

REFERENCES

REFERENCES

1. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping - A mechanism for oligomer assembly, *Protein Science 4*, 2455-2468.

2. Crestfield, A. M., Stein, W. H., and Moore, S. (1962) On the aggregation of bovine pancreatic ribonuclease, *Archives of biochemistry and biophysics Suppl 1*, 217-222.

3. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: As domains continue to swap, *Protein Science 11*, 1285-1299.

4. Liu, Y. S., Gotte, G., Libonati, M., and Eisenberg, D. (2002) Structures of the two 3D domain-swapped RNase A trimers, *Protein Science 11*, 371-380.

5. Liu, Y. S., Gotte, G., Libonati, M., and Eisenberg, D. S. (2001) Crystal structure of an Rnase A dimer displaying a new type of 3D domain swapping, *Nucleic Acid Database*.

6. Liu, Y. S., Hart, P. J., Schlunegger, M. P., and Eisenberg, D. (1998) The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-angstrom resolution, *Proceedings of the National Academy of Sciences of the United States of America 95*, 3437-3442.

7. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: as domains continue to swap, *Protein science : a publication of the Protein Society 11*, 1285-1299.

8. Koharudin, L. M., Liu, L., and Gronenborn, A. M. (2013) Different 3D domain-swapped oligomeric cyanovirin-N structures suggest trapped folding intermediates, *Proc Natl Acad Sci U S A 110*, 7702-7707.

9. Assar, Z., Nossoni, Z., Wang, W., Santos, E. M., Kramer, K., McCornack, C., Vasileiou, C., Borhan, B., and Geiger, J. H. (2016) Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates, *Structure 24*, 1590-1598.

10. Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., Blundell, T. L., and Slingsby, C. (1990) X-ray analysis of beta-B2-crystallin and evolution of oligomeric lens proteins, *Nature 347*, 776-780.

11. Hakansson, M., Svensson, A., Fast, J., and Linse, S. (2001) An extended hydrophobic core induces EF-hand swapping, *Protein Science 10*, 927-933.

12. Rousseau, F., Schymkowitz, J., and Itzhaki, L. S. (2012) Implications of 3D domain swapping for protein folding, misfolding and function, In *Protein Dimerization and Oligomerization in Biology* (Mattews, J. M., Ed.), pp 137-152.

13. Yang, F., Bewley, C. A., Louis, J. M., Gustafson, K. R., Boyd, M. R., Gronenborn, A. M., Clore, G. M., and Wlodawer, A. (1999) Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping, *Journal of Molecular Biology 288*, 403-412.

14. Szilagyi, A., Zhang, Y., and Zavodszky, P. (2012) Intra-Chain 3D Segment Swapping Spawns the Evolution of New Multidomain Protein Architectures, *Journal of Molecular Biology* 14. Szilagyi, A., Zhang, Y., and Zavodszky, P. (2012) Intra-Chain 3D Segment Swapping Spawns the Evolution of New Multidomain Protein Architectures, *Journal of Molecular Biology 415*, 221-235.

15. Stroud, J. C., Liu, C., Teng, P. K., and Eisenberg, D. (2012) Toxic fibrillar oligomers of amyloid-beta have cross-beta structure, *Proceedings of the National Academy of Sciences of the United States of America 109*, 7717-7722.

16. Cleary, J. P., Walsh, D. M., Hofmeister, J. J., Shankar, G. M., Kuskowski, M. A., Selkoe, D. J., and Ashe, K. H. (2005) Natural oligomers of the amyloid-protein specifically disrupt cognitive function, *Nature Neuroscience 8*, 79-84.

16. Stroud, J. C., Liu, C., Teng, P. K., and Eisenberg, D. (2012) Toxic fibrillar oligomers of amyloid-beta have cross-beta structure, *Proceedings of the National Academy of Sciences of the United States of America 109*, 7717-7722.

17. Cleary, J. P., Walsh, D. M., Hofmeister, J. J., Shankar, G. M., Kuskowski, M. A., Selkoe, D. J., and Ashe, K. H. (2005) Natural oligomers of the amyloid-protein specifically disrupt cognitive function, *Nature Neuroscience 8*, 79-84.

18. Fandrich, M., Fletcher, M. A., and Dobson, C. M. (2001) Amyloid fibrils from muscle myoglobin - Even an ordinary globular protein can assume a rogue guise if conditions are right, *Nature 410*, 165-166.

19. Gronenborn, A. M. (2009) Protein acrobatics in pairs - dimerization via domain swapping, *Current Opinion in Structural Biology 19*, 39-49.

20. Schlunegger, M. P., Bennett, M. J., and Eisenberg, D. (1997) Oligomer formation by 3D domain swapping: A model for protein assembly and misassembly, In *Advances in Protein Chemistry, Vol 50: Protein Misassembly* (Wetzel, R., Ed.), pp 61-122.

21. Selkoe, D. J. (2004) Folding proteins in fatal ways (vol 426, pg 900, 2003), *Nature 428*, 445-445.

22. Shtilerman, M. D., Ding, T. T., and Lansbury, P. T. (2002) Molecular crowding accelerates fibrillization of alpha-synuclein: Could an increase in the cytoplasmic protein concentration induce Parkinson's disease?, *Biochemistry 41*, 3855-3860.

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

24. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Refined structure of dimeric diphtheria toxin at 2.0 A resolution, *Protein science : a publication of the Protein Society 3*, 1444-1463.

25. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D., and Zhang, K. Y. (2001) Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus, *Structure 9*, 1017-1027.

26. Ellis, R. J. (2001) Macromolecular crowding: an important but neglected aspect of the intracellular environment, *Curr Opin Struct Biol 11*, 114-119.

27. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Rlefined strucutre of dimeric diphtheria-toxin at 2.0-angstrom resolution, *Protein Science 3*, 1444-1463.

28. Carroll, S. F., Barbieri, J. T., and Collier, R. J. (1986) Dimeric form of diphtheria-toxin-purification and characterization, *Biochemistry 25*, 2425-2430.

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

30. Minton, A. P. (2000) Implications of macromolecular crowding for protein assembly, *Current Opinion in Structural Biology 10*, 34-39.

31. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D., and Zhang, K. Y. J. (2001) Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus, *Structure 9*, 1017-1027.

32. Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014) Energy landscapes of functional proteins are inherently risky, *Nature chemical biology 10*, 884-891.

33. Ignatova, Z., and Gierasch, L. M. (2004) Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling, *Proc Natl Acad Sci U S A 101*, 523-528.

34. Ignatova, Z., and Gierasch, L. M. (2006) Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant, *Proc Natl Acad Sci U S A 103*, 13357-13361.

35. Ropson, I. J., Boyer, J. A., Schaeffer, B. A., and Dalessio, P. M. (2009) Comparison of the folding mechanism of highly homologous proteins in the lipid-binding protein family, *Proteins: Structure, Function, and Bioinformatics 75*, 799-806.

36. Burns, L. L., and Ropson, I. J. (2001) Folding of intracellular retinol and retinoic acid binding proteins, *Proteins: Structure, Function, and Bioinformatics 43*, 292-302.

37. Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014) Energy landscapes of functional proteins are inherently risky, *Nat. Chem. Biol. 10*, 884-891.

38. Haunerland, N. H., and Spener, F. (2004) Fatty acid-binding proteins--insights from genetic manipulations, *Progress in lipid research 43*, 328-349.

39. Lucke, C., Huang, S., Rademacher, M., and Ruterjans, H. (2002) New insights into intracellular lipid binding proteins: The role of buried water, *Protein science : a publication of the Protein Society 11*, 2382-2392.

40. Nossoni, Z., Assar, Z., Yapici, I., Nosrati, M., Wang, W., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal, *Acta crystallographica. Section D, Biological crystallography 70*, 3226-3232.

41. Schaap, F. G., van der Vusse, G. J., and Glatz, J. F. (2002) Evolution of the family of intracellular lipid binding proteins in vertebrates, *Molecular and cellular biochemistry 239*, 69-77.

42. Smathers, R. L., and Petersen, D. R. (2011) The human fatty acid-binding protein family: evolutionary divergences and functions, *Human genomics 5*, 170-191.

43. Crist, R. M., Vasileiou, C., Rabago-Smith, M., Geiger, J. H., and Borhan, B. (2006) Engineering a Rhodopsin Protein Mimic, *Journal of the American Chemical Society 128*, 4522-4523.

44. Vasileiou, C., Vaezeslami, S., Crist, R. M., Rabago-Smith, M., Geiger, J. H., and Borhan, B. (2007) Protein design: reengineering cellular retinoic acid binding protein II into a rhodopsin protein mimic, *J Am Chem Soc 129*, 6140-6148.

45. Berbasova, T., Nosrati, M., Vasileiou, C., Wang, W., Lee, K. S., Yapici, I., Geiger, J. H., and Borhan, B. (2013) Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone, *J Am Chem Soc 135*, 16111-16119.

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

47. Wang, W., Nossoni, Z., Berbasova, T., Watson, C. T., Yapici, I., Lee, K. S., Vasileiou, C., Geiger, J. H., and Borhan, B. (2012) Tuning the electronic absorption of protein-embedded all-trans-retinal, *Science 338*, 1340-1343.

48. Nosrati, M., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. H. (2016) A Photoisomerizing Rhodopsin Mimic Observed at Atomic Resolution, *J Am Chem Soc 138*, 8802-8808.

49. Berbasova, T., Santos, E. M., Nosrati, M., Vasileiou, C., Geiger, J. H., and Borhan, B. (2016) Light-Activated Reversible Imine Isomerization: Towards a Photochromic Protein Switch, *Chembiochem : a European journal of chemical biology 17*, 407-414.

50. Yapici, I., Lee, K. S., Berbasova, T., Nosrati, M., Jia, X., Vasileiou, C., Wang, W., Santos, E. M., Geiger, J. H., and Borhan, B. (2015) "Turn-on" protein fluorescence: in situ formation of cyanine dyes, *J Am Chem Soc 137*, 1073-1080.

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

52. Folli, C., Calderone, V., Ottonello, S., Bolchi, A., Zanotti, G., Stoppini, M., and Berni, R. (2001) Identification, retinoid binding, and x-ray analysis of a human retinol-binding protein, *Proceedings of the National Academy of Sciences of the United States of America 98*, 3710-3715.

53. Gunasekaran, K., Hagler, A. T., and Gierasch, L. M. (2004) Sequence and structural analysis of cellular retinoic acid-binding proteins reveals a network of conserved hydrophobic interactions, *Proteins-Structure Function and Genetics 54*, 179-194.

54. Nossoni, Z., Ass ar, Z., Yapici, I., Nosrati, M., Wang, W., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal, *Acta Crystallogr. Sect. D. Biol. Crystallogr. 70*, 3226-3232.

55. Shameer, K., Shingate, P. N., Manjunath, S. C. P., Karthika, M., Pugalenthi, G., and Sowdhamini, R. (2011) 3DSwap: curated knowledgebase of proteins involved in 3D domain swapping, *Database-the Journal of Biological Databases and Curation*.

56. Sillitoe, I., Cuff, A. L., Dessailly, B. H., Dawson, N. L., Furnham, N., Lee, D., Lees, J. G., Lewis, T. E., Studer, R. A., Rentzsch, R., Yeats, C., Thornton, J. M., and Orengo, C. A. (2013) New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures, *Nucleic Acids Research 41*, D490-D498.

57. Englander, S. W., and Mayne, L. (2014) The nature of protein folding pathways, *Proc Natl Acad Sci U S A 111*, 15873-15880.

58. Englander, S. W., Mayne, L., and Krishna, M. M. (2007) Protein folding and misfolding: mechanism and principles, *Quarterly reviews of biophysics 40*, 287-326.

59. Krishna, M. M., and Englander, S. W. (2007) A unified mechanism for protein folding: predetermined pathways with optional errors, *Protein science : a publication of the Protein Society 16*, 449-464.

60. Mayne, L., and Englander, S. W. (2000) Two-state vs. multistate protein unfolding studied by optical melting and hydrogen exchange, *Protein science : a publication of the Protein Society 9*, 1873-1877.

61. Roder, H., Elove, G. A., and Englander, S. W. (1988) Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR, *Nature 335*, 700-704.

62. Martins, N. F., and Santoro, M. M. (1999) Partially folded intermediates during trypsinogen denaturation, *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas 32*, 673-682.

63. Medina, E., Cordova, C., Villalobos, P., Reyes, J., Komives, E. A., Ramirez-Sarmiento, C. A., and Babul, J. (2016) Three-Dimensional Domain Swapping Changes the Folding Mechanism of the Forkhead Domain of FoxP1, *Biophysical journal 110*, 2349-2360.

64. Zwanzig, R. (1997) Two-state models of protein folding kinetics, *Proceedings of the National Academy of Sciences of the United States of America 94*, 148-150.

65. Sanson, B., Wang, T., Sun, J., Wang, L., Kaczocha, M., Ojima, I., Deutsch, D., and Li, H. (2014) Crystallographic study of FABP5 as an intracellular endocannabinoid transporter, *Acta crystallographica. Section D, Biological crystallography 70*, 290-298.

66. Assar, Z., Nossoni, Z., Wang, W., Santos, Elizabeth M., Kramer, K., McCornack, C., Vasileiou, C., Borhan, B., and Geiger, James H. Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates, *Structure*.

67. Silvaroli, J. A., Arne, J. M., Chelstowska, S., Kiser, P. D., Banerjee, S., and Golczak, M. (2016) Ligand Binding Induces Conformational Changes in Human Cellular Retinol-binding Protein 1 (CRBP1) Revealed by Atomic Resolution Crystal Structures, *The Journal of biological chemistry 291*, 8528-8540.

68. Gillilan, R. E., Ayers, S. D., and Noy, N. (2007) Structural basis for activation of fatty acid-binding protein 4, *J Mol Biol 372*, 1246-1260.

69. Sessler, R. J., and Noy, N. (2005) A ligand-activated nuclear localization signal in cellular retinoic acid binding protein-II, *Molecular cell 18*, 343-353.

70. Armstrong, E. H., Goswami, D., Griffin, P. R., Noy, N., and Ortlund, E. A. (2014) Structural Basis for Ligand Regulation of the Fatty Acid-binding Protein 5, Peroxisome Proliferator-

activated Receptor β/δ (FABP5-PPARβ/δ) Signaling Pathway, *The Journal of biological chemistry 289*, 14941-14954.

71. Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, *Nature Protocols 1*, 2527-2535.

72. Ramos, C. R. R., Oyama, S., Jr., Sforca, M. L., Pertinhez, T. A., Ho, P. L., and Spisni, A. (2008) 2POA: Schistosoma mansoni Sm14 Fatty Acid-Binding Protein: improvement of protein stability by substitution of the single Cys62 residue, *Worldwide Protein Data Bank*.

73. Gill, S. C., and Vonhippel, P. H. (1989) CALCULATION OF PROTEIN EXTINCTION COEFFICIENTS FROM AMINO-ACID SEQUENCE DATA, *Analytical Biochemistry 182*, 319-326.

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

75. Tarter, M., Capaldi, S., Carrizo, M. E., Ambrosi, E., Perduca, M., and Monaco, H. L. (2008) Crystal structure of human cellular retinol-binding protein 11 to 1.2 angstrom resolution, *Proteins-Structure Function and Bioinformatics 70*, 1626-1630.

76. 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., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallographica Section D-Biological Crystallography 66*, 213-221.

77. Bailey, S. (1994) THE CCP4 SUITE - PROGRAMS FOR PROTEIN CRYSTALLOGRAPHY, *Acta Crystallographica Section D-Biological Crystallography 50*, 760-763.

78. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *Journal of Applied Crystallography 30*, 1022-1025.

79. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallographica Section D-Biological Crystallography 66*, 486-501.

80. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L.-W., Jain, S., Kapral, G. J., Kunstleve, R. W. G., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2011) The Phenix software for automated determination of macromolecular structures, *Methods 55*, 94-106.

81. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallographica Section D-Biological Crystallography 53*, 240-255.

71. Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, *Nature Protocols 1*, 2527-2535.

72. Ramos, C. R. R., Oyama, S., Jr., Sforca, M. L., Pertinhez, T. A., Ho, P. L., and Spisni, A. (2008) 2POA: Schistosoma mansoni Sm14 Fatty Acid-Binding Protein: improvement of protein stability by substitution of the single Cys62 residue, *Worldwide Protein Data Bank*.

73. Gill, S. C., and Vonhippel, P. H. (1989) CALCULATION OF PROTEIN EXTINCTION COEFFICIENTS FROM AMINO-ACID SEQUENCE DATA, *Analytical Biochemistry 182*, 319-326.

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

75. Tarter, M., Capaldi, S., Carrizo, M. E., Ambrosi, E., Perduca, M., and Monaco, H. L. (2008) Crystal structure of human cellular retinol-binding protein 11 to 1.2 angstrom resolution, *Proteins-Structure Function and Bioinformatics 70*, 1626-1630.

76. 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., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallographica Section D-Biological Crystallography 66*, 213-221.

77. Bailey, S. (1994) THE CCP4 SUITE - PROGRAMS FOR PROTEIN CRYSTALLOGRAPHY, *Acta Crystallographica Section D-Biological Crystallography 50*, 760-763.

78. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *Journal of Applied Crystallography 30*, 1022-1025.

79. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallographica Section D-Biological Crystallography 66*, 486-501.

80. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L.-W., Jain, S., Kapral, G. J., Kunstleve, R. W. G., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H.

(2011) The Phenix software for automated determination of macromolecular structures, *Methods 55*, 94-106.

81. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallographica Section D-Biological Crystallography 53*, 240-255.

Chapter III: Can we predict and control the folding product of proteins through their amino acid sequence (at least in iLBP family)?

III-1 Introduction

As we discussed in the previous chapter, intracellular lipid binding proteins (iLBP) are a family of relatively small, soluble proteins. They chaperone lypophilic, insoluble molecules through the cytosol of virtually all animal tissues.[106, 107] Fatty acid binding proteins (FABPs) are members of this family[107]. Recently it was shown that the inhibition of intracellular endocannabinoid carriers such as FABPs may provide an alternative strategy to modulate endocannabinoid inactivation and anti-inflammatory effects. [93] There is also evidence of the direct transfer of their lypophilic cargo to the regulatory transcription factors, including the retinoic acid receptor (RAR) in the case of retinoic acid and the peroxisome proliferator-activated receptor (PPAR) in the case of Linoleic acid. [108] Many other recent reports have suggested a mechanism of allosteric regulation for some iLBP family members, namely FABP4 and CRABPII. [96] [97, 98] Specifically, it has been suggested in the case of both FABP4 and CRABPII that ligand binding leads to conformational change that results in translocation to the nucleus, delivering either retinoic acid or fatty acids to their respective nuclear receptors. [96] [97, 98]

However, structures of apo and holo iLBP monomers show only subtle changes in structure upon binding. Based on our results, discussed in section III-6, we suggest the DS dimer structure to be exquisitely sensitive to ligand binding, undergoing reproducible and dramatic conformational change upon ligand association. This begs the question of whether the DS dimer may form *in vivo* for some members of the iLBP family, and play important physiological roles, especially related to allosteric regulation.

The ultimate goal of our research is to have the tools necessary to predict the folding products of iLBP family members based on their sequences. This will require understanding of the effect of sequence insertions, significant changes in amino acid identity at certain positions, etc. This would allow us to control and predict the folding pathway of iLBP family members to achieve the desired folding product, either monomer or DS dimer. The accomplishment of this goal will allow us to predict from sequence the likely folding outcome for members of the family, and will also allow us to control at will the folding product via mutagenesis

III-2 Investigation of potential domain swapping in other iLBP family members: Case Study 1, human FABP4

III-2-1 History of FABP4

As previously mentioned, gene-activating lipophilic compounds are carried from the cytosol to the nucleus by FABP. FABP4 (adipocyte FABP), and FABP5 (epidermal FABP) specifically deliver to Peroxisome-proliferators-activated receptors (PPARγ, and PPARβ/δ repectively) followed by enhancing the transcriptional activity of the receptor.

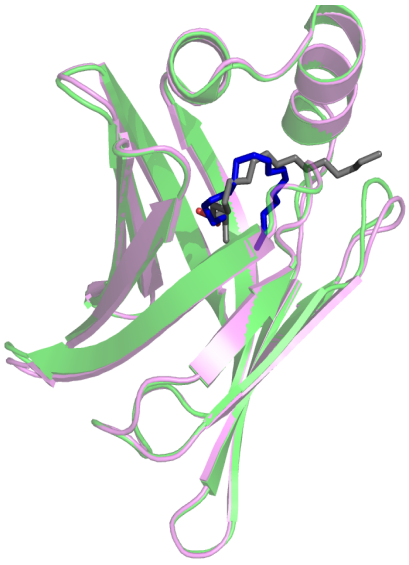


Figure III-1: Overlaid of FABP4 (2q9s, green) -Linoleic acid (blue), FABP4 (2ans, pink)-oleic acid (grey) shows the exit of the oleic acid from the binding pocket.

The activated receptor is involved in regulation of adipocyte differentiation, insulin sensitivity etc. [96] [97, 98] FABP4 binds to activating (linoleic acid) and non-activating compounds (oleic acid). Recent studies showed activated and non-activated conformations of the protein have different homodimeric configurations through crystal packing. Non-activating ligands exit from the portal (lid), preventing binding pocket closure. **(Figure III-1)** Crystallographic symmetry yields a homodimer in which opposing helix-turn-helix subdomains form the principal protein–protein contact surface followed by masking its NLS site. [44]

III-2-2 Investigation on hFABP4 bacterial expression

Initially we picked this protein as the first case study. We thought the "homodimeric configuration" due to crystal packing might be the incorrectly refined DS dimer structure. subsequently we found out about the presence of dimer species in our size exclusion chromatography profile. (**Figure III-2**) However it was due to a disulfide bond linkage between Cys in the N-terminus of FABP4 and it was confirmed by addition of denaturant, 2mM DTT, in our purification buffers and running an SDS-gel with and without fresh denaturant reagent, which we only saw one species as monomer size whenever DTT was present. Even while performing denaturation/ renaturation experiment (explained in section III-3), only monomer was obtained at higher concentrations of material (0.3mM). (Data not shown)

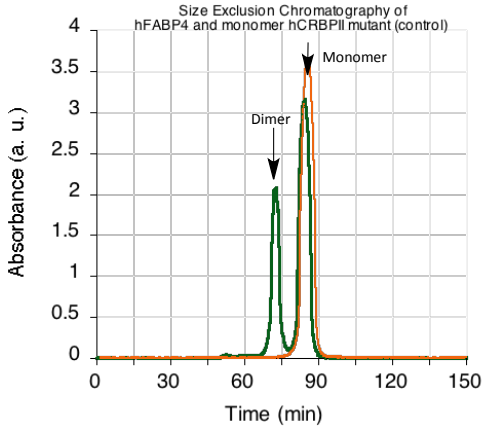


Figure III-2: Size Exclusion Chromatogram of human FABP4, which was expressed in *E.coli* (at 30°C as dimer favoring temperature, explained in section III-3, green). SEC profile of monomer hCRBPII mutant (orange) as our size control.

III-3 Investigation of potential domain swapping in other iLBP family members: Case Study 1, human FABP5

III-3-1 History of hFABP5

Fatty acid binding protein 5 (FABP5) is found mostly in epidermal cells, but also in other tissues, such as brain, liver, kidney, lung, and adipose tissue.

Recently it has been shown that FABP5 is an intracellular carrier of anandamide (AEA, the best characterized endocannabinoid which activates cannabinoid receptors in the central nervous system and peripheral tissues, i.e. CB1) from plasma membrane to intracellular fatty acid amide hydrolase (FAAH) for hydrolysis, since the period of endocannabinoid signaling, by endocannabinoid, is regulated through 'on demand' biosynthesis and prompt catabolism. [(93)]

(**Figure III-3)**

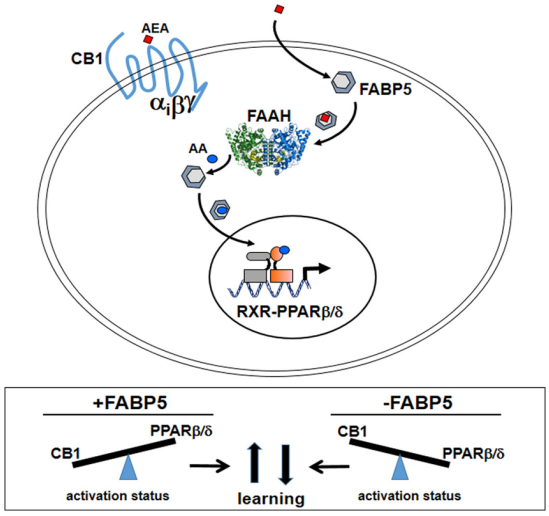


Figure III-3: Model for the role of-FABP5 and its involvement in regulating.

Also it has been shown that the selective FABP5 inhibitor, BMS-309403, reduces AEA inactivation in cells and increases endocannabinoid levels, produceing beneficial anti-inflammatory and antinociceptive effects.[109, 110] FABP5 has been discovered to form a domain swapped dimer, with hinge residues of Ser59, Thr60, Val61 (equivalent of Ser55, Thr56, Phe57 in hCRBPII) (**Figure III-4 and III-5b**).[93] Existence of domain swapping in FABP5 was very interesting for us, since this protein is in the same family as hCRBPII.

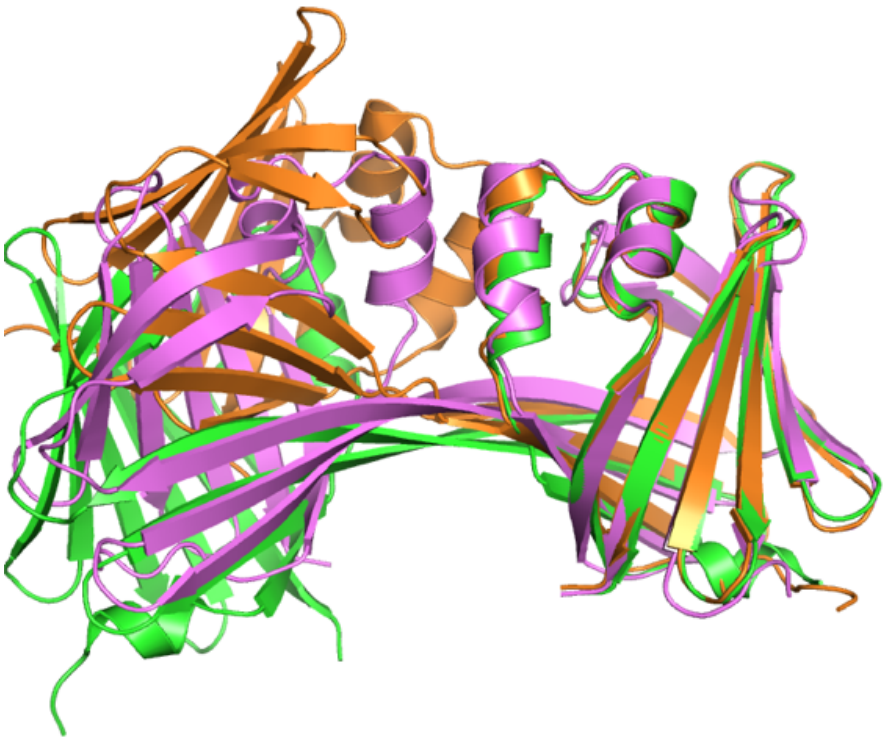


Figure III-4: Overlaid of DS dimers of holo hFABP5 and hCRBPII. HCRBPII (purple, PDB ID 4ZH9), BMS-hFABP5 (green), AEA-hFABP5 (orange). There is a big different orientation of two domains upon ligand binding.

Two hFABP5 DS dimer structures have thus far been determined. Both are ligand-bound, but the two ligands are distinct, one being the endocannabinoid anandamide, an actIIIator and the other BMS-309403, a high-affinity inhibitor. [93] A large change in the relatIIIe orientation of the two domains is seen in the two ligand-bound DS dimer structures, presumably due to the very different binding modes of the two ligands. **(Figure III-5)** It is therefore possible that different ligands can result in distinct protein conformations, which could then provide distinct, allosterically regulated function, to the protein. Although validation of such a mechanism requires further exploration of the iLBP family members *in vIIIo*, the fact that the wild-type sequence of a second member of the iLBP family also domain swaps in a very

similar way to that of hCRBPII, suggests that a significant subset of iLBP family members may also form DS dimers.

Therefore, We have conducted mutagenesis and structural elucidation on human Fatty Acid Binding Protein 5 (hFABP5) to analyze its domain swapping pathway and with the future goal of predicting potential sequences of FABP's (or other iLBPs members) to produce DS dimers as well. For instance, if we think that the presence of hydrophobic residues in positions that are only inside the binding pocket in the DS dimer lead to significantly increased DS dimerization, then it needs to be investigated among other iLBP family members whose sequences have such residues. Several such sequences can already be identified **(Figure III-11)** which other members of our group will investigate on them from now on. Eventually this will lead to a re-evaluation of the structural biology of the iLBP family.

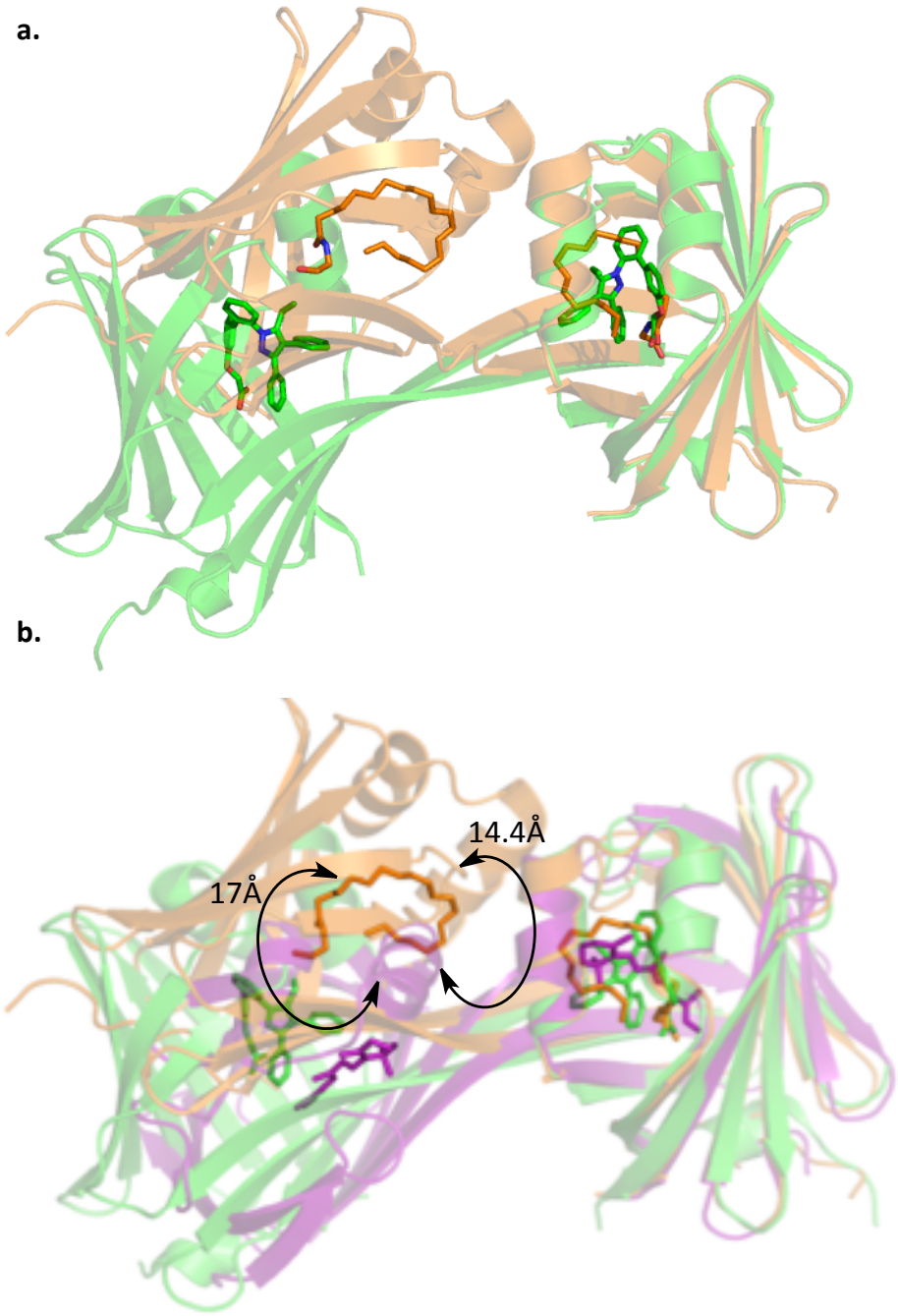


Figure III-5: a. Overlaid of holo DS dimers of BMS-hFABP5 (green), AEA-hFABP5 (orange). Dramatic change is seen in the relative orientation of two domains in these DS dimers, which is induced by different ligands. **b.** Overlaid of DS dimers of BMS-hFABP5 (green), AEA-hFABP5 (orange) and holo Q108K:K40D DS dimer of hCRBPII. AEA, BMS and retinal are shown as sticks. Between structures of dimer hCRBPII and AEA-hFABP5, there is 14Å deviation while for BMS-hFABP5 dimer there is a larger deviation of 17Å.

III-3-2 Investigate domain swapping in human fatty acid binding protein 5

In order to investigate domain swapping in hFABP5, we have decided to first study the domain swapping of apo protein (no addition of fatty acids). Surprisingly, during its purification, the chromatogram of the SEC of the elution fractions of post Ni-NTA purification (since it is a his-tagged protein) showed different results each time for each batch of expression. Most of the time, we achieved a large monomer peak and very small peaks for the dimer and higher oligomers. These observations were the same even at different expression temperatures (20°C and 30°C and 37°C), which was different than hCRBPII. However, one time we were able to get a significant amount of dimer (35%) after purification, which was confirmed by SDS-PAGE. (**FigureIII-6a and 6b,** The reason for getting different results for expression is still not clear for us. Further, we have found no evidence of monomer/dimer exchange over 10 days at room temperature, though it was reported that monomer/dimer exchange might be occurring much more rapidly. *(93)*

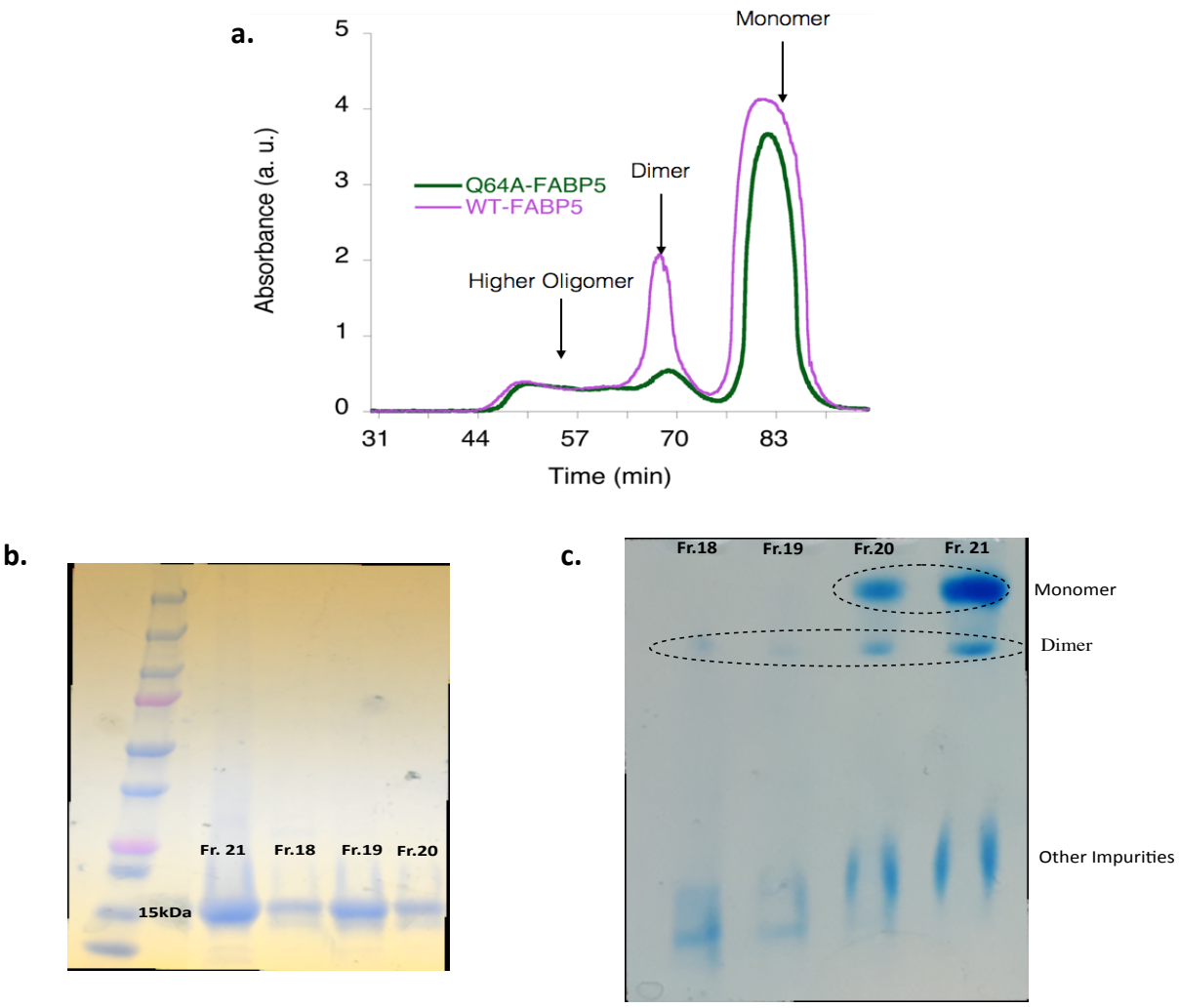


Figure III-6: a. SEC of WT and Q4A mutant FABP5 purification by SEC after E. coli expression. **b.** SDS-PAGE of SEC fractions 18-21 of WT-hFABP5. **c.** Native gel of the same SEC fractions. It shows that in the dimer peak, both monomer and dimer species are present.

However, these studies were conducted after de-lipidation treatment (we did not treat it with lipidex), which could potentially decrease the barrier between monomer and dimer. Basically delipidation process is performed by lipidex resin and it is recommended to remove all the non-specific lipids (which in this case, they may come from *E.coli* cells). Followed by crystallization trials of the fractions we thought to be dimer. We were able to crystallize and determine the structure in $P2_12_12_1$ space group **(Table III-1).** However it was not the DS dimer

stricter as we expected it, instead it was two monomers per asymmetric unit. Where in the crystal packing, two monomers meeting right at the loop of connecting 3rd and 4th beta strand (hinge loop in DS dimer structures) where the swapping should occur, but the density for this part is not very good to build the DS dimer structure. **(Figure III-7, Table III-1)** This led us to believe that the corresponded dimer fractions may be a mix of dimer and monomer. This was then confirmed by native-gel electrophoresis (**Figure III-6c**). Though the Dimer species is present in all the fractions, it is only relatively pure in Fraction 18. Pooling all the fractions from 18 to 20 gave a mixture of monomer and dimer. This is due to the fact that monomer and dimer do not separate well on our gel filtration column.

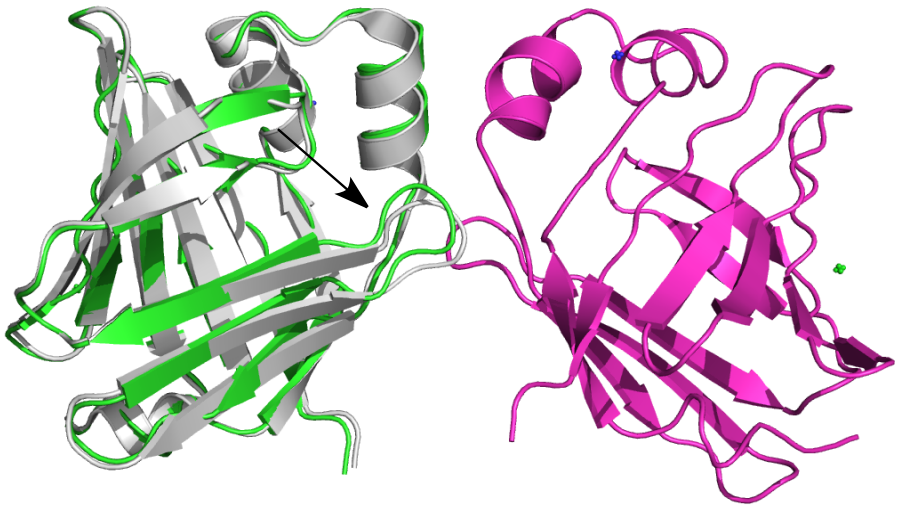


Figure III-7: Overlaid of our current monomer structure (chain A in green and chain B in pink) with hFABP5 monomer (grey, PDB ID 4lkp)**.** There is a different conformation between their 3rd and 4th beta sheet (hinge loop region in DS dimer), which is pointed by arrow.

Nevertheless, The structure of our monomer has different conformation compare to the other monomer structures reported between their 3rd and 4th beta sheet (hinge loop region in DS dimer structures)[(98)] **(Figure III-7)** All of these raise the hypothesis that due to the existence

of both monomer and dimer species in the solution, these monomer crystals maybe were nucleated by dimer species , therefore we cannot build these structure models as DS dimers.

After that, we expressed hFABP5 multiple times and one time we could get dimer peak again. This time knowing that only fr. 18 of the SEC run contains dimer species, we separately pooled and were able to crystallize it (performed by Nona Ehyai). As we expected, we observed electron density for the continuous hinge loop region and the structure was determined in $P6_322$ space group as symmetric domain swapped dimer. **(Figure III-8a)** Although the crystal packing of this structure and the previously mentioned monomer structure are comparable, there is a large deviation in their relative orientation of the two molecules in those structures **(Figure III-8b)** similar to the **Figure III-5.** Therefor, this may give some evidence that the previous pseudo monomer structure was nucleated by a dimer. However, there is an extra electron density in one of the dimer chains binding pocket, which seems to be the size of a fatty acid. Since we never added any potential ligand to the protein used in crystallization, we suspect that the ligand is a fatty acid from the E. coli the protein was expressed in. Interestingly, this DS dimer structure share very similar relative orientation compare to the DS dimer structure of hFABP5-AEA. (**Figure III-8c**).

As previously mentioned, the structure of apo hFABP5 domain swapped dimer is not known and is presumably the product of dimerization during folding. Our group actively will try to obtain the apo structure of DS dimer hFABP5 and to use it to design a set of mutagenesis experiments to probe the determinants of the monomer/dimer ratio in FABP5 as described in chapter III for hCRBPII. A similar strategy will be used to probe the phase relationship in FABP5 as described for hCRBPII as well. **(Figure III-11)** In order to achieve this structure and

understand dimer formation for this protein, the reasons behind the different expression levels of domain swapped dimers need to be investigated, since the expression process is not the same every time.

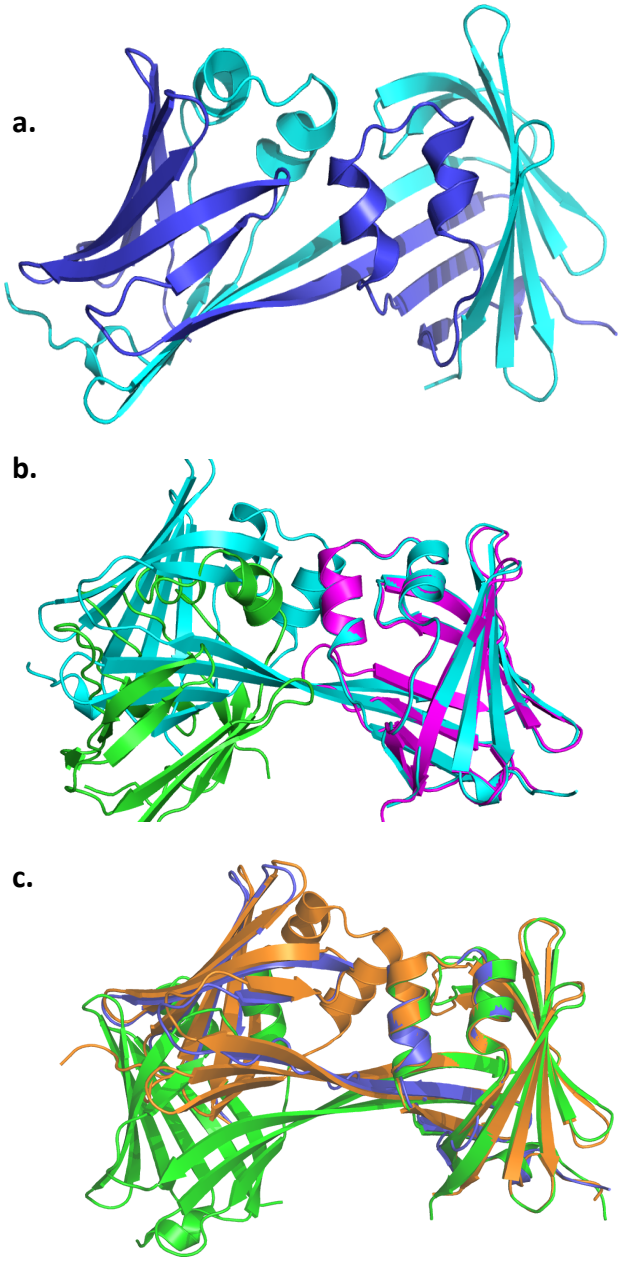


Figure III_8: a. Crystal structure of DS dimer of hFABP5 of our trial. Chain A in blue, Chain B in cyan. **b.** Overlaid of DS hFABP5 structure (obtained by Nona, cyan**)** with abovementioned pseudo monomer hFABP5 (chain A in pink and chain B in green)**c.** Overlay of our determined DS hFABP5 structure (blue**)** with DS dimer of BMS-hFABP5 (green, 4AZM), and AEA-hFABP5 (orange, 4AZR). This dimer structure is very similar to the AEA- structure and both have different orientation of their domains compared with the BMS-structure.

III-3-3 Study domain swapping occurrence by mutational analysis of human FABP5

We were looking for mutations that could increase the ratio of dimerization of hFABP5. In order to do this, we compared the structure of the monomer and dimer of hFABP5. In the monomer, Thr64 (equivalent of Y60 in hCRBPII, **Figure III-11)** points in, toward the binding pocket and makes a water-mediated hydrogen bond with Thr57, (Note: res 59 to 61 are in the hinge loop region). **(Figure III-9b)** However, in the dimer, Thr64 points outside and there is a new hydrogen bond between of one chain to Thr57 of the other chain of the dimer **(Figure III-9a).** We reasoned that since both Gln65 and Thr64 switche sides from outside to inside the binding pocket in the monomer and DS dimer, respectively, this position likely to be a "mutational hot spot" for DS dimerization

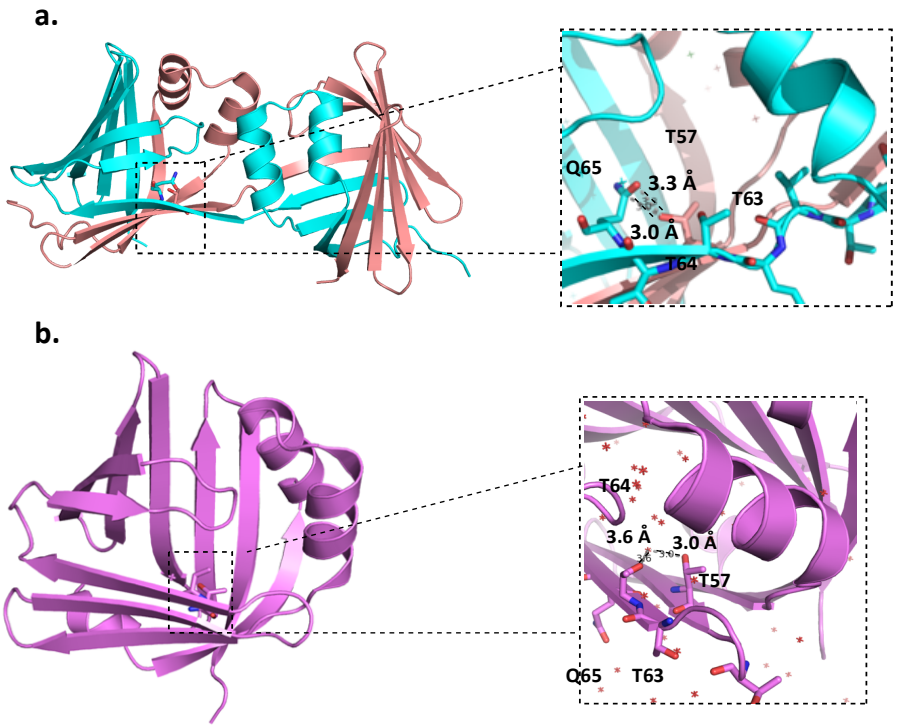


Figure III-9: a. In DS dimer of hFABP5 (Chain A in salmon, chain B in cyan, PDB 4AZR), there is a hydrogen bond between Q65 and T57. **b**. While in monomer structure (purple, PDB ID 4LKP) Gln65 is pointing out of toward the solvent and T64 makes a hydrogen bond to Thr57 via water network. Res. 57, 64 and 65 are represented in sticks. Water molecules are in red.

Therefore, we thought that by removing this hydrogen bond between Gln65 with Thr57 in the dimer, the formation of the monomer could be more feasible. We mutated Gln to Ala, which is a small hydrophobic residue. This mutation yielded almost no dimer formation compare to the WT-hFABP5 which confirmed by native gel electrophoreses **(Figure IV-6a)** Although, for hFABP5 we cannot say exact ratio of monomer/dimer like hCRBPII since we could not get dimer dependently than monomer. We could get small crystals of this mutant but they did not diffract. Regarding this scenario, Q65 was mutated to W and M as well, **(Figure III-10a)** which in both mutations, monomer of hFABP5 was only expressed (though Q65M made higher order oligomers in non reducing environment). We could get only 2 crystals of the Q65W mutant, unfortunately they diffracted poorly, to about 5Å.

We generated more mutations of hFABP5. However, no mutation has been discovered yet that increased the ratio of dimer to monomer significantly.

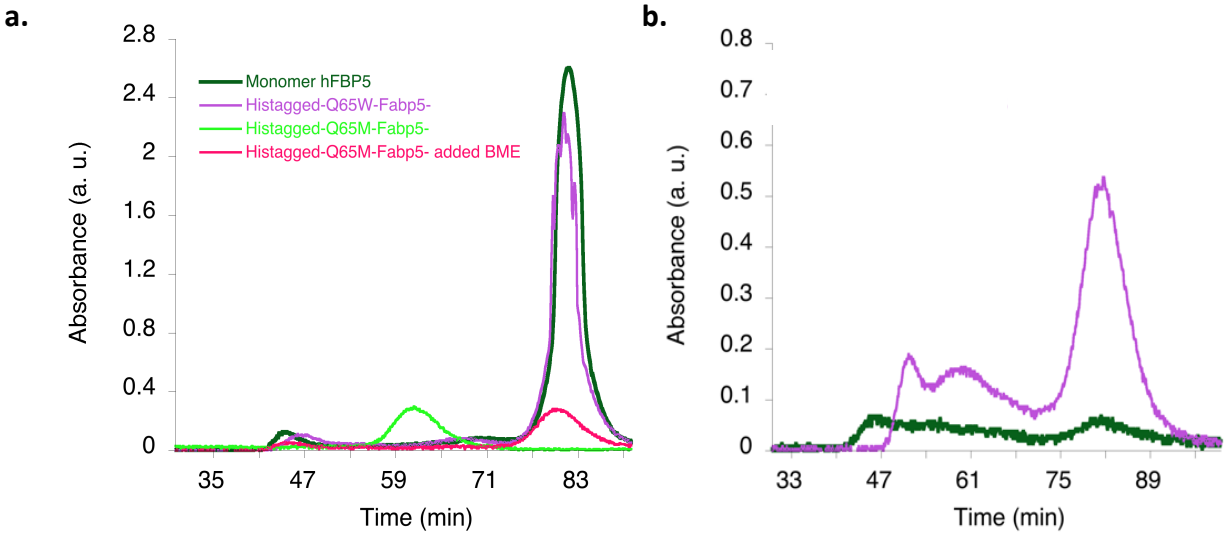


Figure III-10: a. Size exclusion chromatogram of Gln65 mutants of hFABP5 during purification by SEC. **b.** Size exclusion chromatogram of Thr64 mutants of hFABP5 (Thr63E pink, T63L green) during purification by SEC.

We applied the phase relationship for monomer and dimer hFABP5 (rephrasing occurs at Phe 66. Refer to section II-4-3 for more details, **Figure III-10b and III-11**). For example, the Thr 64 side chain is going from inside to outside of the binding cavity in monomer to dimer, it could be mutated to hydrophilic residues to increase the dimer formation. To test this hypothesis, we first mutated Thr64 to Leu (small hydrophobic residue to mimic the Y60L mutant of hCRBPII). It did not express well and the yield was low. Then T64E was mutated and expressed which it was mostly monomer and higher oligomers. **(Figure III-10b)** Other members of our group are carrying on the crystallization trial for these mutants.

As our other mutation candidates, Thr63 can be chosen to some residues, which are more hydrophobic. As it mentioned before, hinge loop residues involved in hFABP5 DS dimer structures are: Ser59, Thr60, Val61 (the same as Ser55, Thr56 and phe57 in hCRBPII). Therefore, we hypothesized that by mutating Thr60 to a rigid amino acid (i.e. Pro), it will make it harder for the loop to close back and form monomer then increases the dimer formation. To this end we mutated and expressed T60P, however our hypothesis seems not to be right and it also only formed monomer. But this could be because Pro is not hydrophilic and since this residue is pointed toward the solvent in dimer, it needed to be mutated to more hydrophilic residues instead. Other residues, which we found them important, are Lys62, Thr63 and Phe66 because they have different conformation in dimer from monomer. It should be considered that we are not able to identify the propensity of the dimer the same way as hCRBPII (which descried in chapter II) since we cannot separate dimer vs. monomer in purification steps but we only can distinguish the mutants that generate only one of the species: dimer or monomer.

Together, this data will probe the importance of establishing the phase relationship

required for dimerization and our group still is focusing on this aim.

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

	FABP5	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69
	FABP5 sequence	**N**	**L**	**T**	**I**	**K**	**T**	**E**	**S**	**T**	**V**	**K**	**T**	**T**	**Q**	**F**	**S**	**C**	**T**
Monomer	WT-FABP5	O	I	O	I	O	I	O	I	L	L	I	O	I	O	I	O	I	O
Symmetric Dimers	WT(AZO)	O	I	O	I	O	I	O	I	O	I	**O**	**I**	**O**	**I**	I	O	I	O
	WT(AZR)	O	I	O	I	O	I	O	I	O	I	**O**	**I**	**O**	**I**	I	O	I	O

I: Inside; **O**: Outside; **S**: Neither Inside or Outside, but Solvent exposed; **L**: loop

Figure III-11: Phase relationship in iLPBP family.

III-4 Experimental

III-4-1 Material and Method: Site-Directed Mutagenesis

The hFABP5 and hFABP4 plasmids described were ordered from IDT®. Then we cloned them into pET28a (Novagen) between BamHI and HindIII cut sites and it yielded N-terminal his-tagged (cleavable by thrombin). We used these constructs for mutagenesis as well. (Detailed protocol explained in section III-7).

Synthesized amino acid sequence of hFABP5 (N-terminal Histagged-thrombin):

HHHHHHSSGLVPRGSHMGSMGSHMATVQQLEGRWRLVDSKGFDEYMKELGVGIALRKMGAMAKPDCII

TCDGKNLTIKTESTVKTTQFSCTLGEKFEETTADGRKTQTVCNFTDGALVQHQEWDGKESTITRKLKDGKLVV

ECVMNNVTCTRIYEKVE

Primers

T64A-hFABP5
Forward: 5**'**- CGTGAAGACGGCGCAGTTTTCAT-3**'**
Reverse: 5**'**- GTCGATTCAGTTTTAATAGTTAAGGTTCTTA-3**'**

T60P-hFABP5
Forward: 5**'**- AACTGAATCGCCCGTGAAGACGA -3**'**
Reverse: 5**'**- TTAATAGTTAAGTTCTTACCATCGC-3**'**

T64E-hFABP5
Forward: 5'- CGTGAAGACGGAGCAGTTTTCATG-3'
Reverse: 5'- GTCGATTCAGTTTTAATAGTTAAG-3'

T64L-hFABP5
Forward: 5'- CGTGGAGACGCTGCAGTTTTCATG- 3'
Reverse: 5'- GTCGATTCAGTTTTAATAGTTAAG- 3'

Q65W-hFABP5
Forward: 5'- GAAGACGACGTGGTTTTCATGCAC-3'
Reverse: 5'-ACGGTCGATTCAGTTTTAATAG -3'

Q65M-hFABP5
Forward: 5'-GAAGACGATGTTTTCATGCAC-3'
Reverse: 5'-ACGGTCGATTCAGTTTTAATAG-3'

Q65W-hFABP5
Forward: 5'- GAAGACGACGTGGTTTTCATGCAC-3'
Reverse: 5'-ACGGTCGATTCAGTTTTAATAG -3'

Q65M-hFABP5
Forward: 5'-GAAGACGATGTTTTCATGCAC-3'
Reverse: 5'-ACGGTCGATTCAGTTTTAATAG-3'

III-4-2 Protein Expression and Purification of hFABP4 and hFABP5

Proteins were expressed in Escherichia coli BL21 (DE3) cells using the T7 expression system. Cells were grown in LB media and were induced at OD_{600} of 0.7, by adding IPTG to a final concentration of 0.4 mM. After 20 h incubation at 20°C, the cells were collected by centrifugation at 5000rpm at 4°C for 30 min. Cells were lysed in 30 ml ice-cold column buffer: 20mM Tris pH 8.5 and 250 mM NaCl (20mM Tris pH 7.4, 500mM NaCl and 5mM imidazole in the case of hFABP4), by sonication on ice. Then followed by 30 min centrifugation at 10,000rpm at 4°C. The supernatant was loaded onto a Ni–NTA column (GE healthcare) and batched bind for one hour at 4°C. The column was washed with 10CV buffer containing 20 mM imidazole. The proteins were then eluted with column buffer containing 200 mM imidazole. In case of delipidation, protein were added to Lipidex-5000 resin and incubated for 1 hour at 37°C and collected the flow through. (sigma) The affinity-purified proteins were incubated with thrombin (GE Healthcare Life Sciences) at 10 units per milligram of protein at 4°C overnight in order to remove the 6 His tagged. The cleaved protein solutions were loaded onto an Ni–NTA columns and the unbound fractions of cleaved materials collected and then concentrated. The affinity-purified samples were concentrated to about 2mL and loaded onto our SEC column (Superdex S75 16/600 HiLoad column, GE healthcare) equilibrated with 1CV PBS pH 8.5. The peak fractions were collected. (In the case of hFABP5, only few times delipidated by incubation with Lipidex- 5000, Sigma at 37° C for 1 h with occasional mixing followed by second run of SEC). Almost all mentioned (except T64L, which didn't express well) mutations, purified proteins were concentrated to approximately 10mg /mL for crystallization purpose.

III-4-3 Crystallization, data collection and refinement

The concentrated un-cleaved human FABP5 apo was crystallized in 25% PEG 3350, 100 mM bis-tris pH 6.0. For details on data collection, processing the data and structure refinement please refer to section **II-7-8**

Table III-1: Crystallographic data of hFABP5 (pseudo monomer).

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

REFERENCES

REFERENCES

1. Ong, D. E. (1984) A NOVEL RETINOL-BINDING PROTEIN FROM RAT - PURIFICATION AND PARTIAL CHARACTERIZATION, *Journal of Biological Chemistry 259*, 1476-1482.

2. Harder, M. E., Deinzer, M. L., Leid, M. E., and Schimerlik, M. I. (2004) Global analysis of three-state protein unfolding data, *Protein Science 13*, 2207-2222.

3. Nossoni, Z., Assar, Z., Yapici, I., Nosrati, M., Wang, W., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal, *Acta crystallographica. Section D, Biological crystallography 70*, 3226-3232.

4. Veerkamp, J. H., Vankuppevelt, T., Maatman, R., and Prinsen, C. F. M. (1993) STRUCTURAL AND FUNCTIONAL-ASPECTS OF CYTOSOLIC FATTY-ACID-BINDING PROTEINS, *Prostaglandins Leukotrienes and Essential Fatty Acids 49*, 887-906.

5. Banaszak, L., Winter, N., Xu, Z. H., Bernlohr, D. A., Cowan, S., and Jones, T. A. (1994) LIPID-BINDING PROTEINS - A FAMILY OF FATTY-ACID AND RETINOID TRANSPORT PROTEINS, *Advances in Protein Chemistry, Vol 45 45*, 89-151.

6. Vaezeslami, S., Mathes, E., Vasilelou, C., Borhan, B., and Geiger, J. H. (2006) The structure of apo-wild-type cellular retinoic acid binding protein II at 1.4 A and its relationship to ligand binding and nuclear translocation, *Journal of Molecular Biology 363*, 687-701.

7. Haunerland, N. H., and Spener, F. (2004) Fatty acid-binding proteins--insights from genetic manipulations, *Progress in lipid research 43*, 328-349.

8. Lucke, C., Huang, S., Rademacher, M., and Ruterjans, H. (2002) New insights into intracellular lipid binding proteins: The role of buried water, *Protein science : a publication of the Protein Society 11*, 2382-2392.

9. Schaap, F. G., van der Vusse, G. J., and Glatz, J. F. (2002) Evolution of the family of intracellular lipid binding proteins in vertebrates, *Molecular and cellular biochemistry 239*, 69-77.

10. Smathers, R. L., and Petersen, D. R. (2011) The human fatty acid-binding protein family: evolutionary divergences and functions, *Human genomics 5*, 170-191.

11. Franzoni, L., Lucke, C., Perez, C., Cavazzini, D., Rademacher, M., Ludwig, C., Spisni, A., Rossi, G. L., and Ruterjans, H. (2002) Structure and backbone dynamics of apo- and holo-

cellular retinol-binding protein in solution, *Journal of Biological Chemistry 277*, 21983-21997.

12. Crow, J. A., and Ong, D. E. (1985) CELL-SPECIFIC IMMUNOHISTOCHEMICAL LOCALIZATION OF A CELLULAR RETINOL-BINDING PROTEIN (TYPE-2) IN THE SMALL-INTESTINE OF RAT, *Proceedings of the National Academy of Sciences of the United States of America 82*, 4707-4711.

13. Vogel, S., Mendelsohn, C. L., Mertz, J. R., Piantedosi, R., Waldburger, C., Gottesman, M. E., and Blaner, W. S. (2001) Characterization of a new member of the fatty acid-binding protein family that binds all-trans-retinol, *Journal of Biological Chemistry 276*, 1353-1360.

14. Folli, C., Calderone, V., Ottonello, S., Bolchi, A., Zanotti, G., Stoppini, M., and Berni, R. (2001) Identification, retinoid binding, and x-ray analysis of a human retinol-binding protein, *Proceedings of the National Academy of Sciences of the United States of America 98*, 3710-3715.

15. Li, E., Demmer, L. A., Sweetser, D. A., Ong, D. E., and Gordon, J. I. (1986) ISOLATION OF A CLONED CDNA-ENCODING AN ABUNDANT INTESTINAL RETINOL BINDING-PROTEIN, *Clinical Research 34*, A685-A685.

16. Li, E., and Norris, A. W. (1996) Structure/function of cytoplasmic vitamin A-binding proteins, *Annual Review of Nutrition 16*, 205-234.

17. Li, E., Qian, S. J., Winter, N. S., Davignon, A., Levin, M. S., and Gordon, J. I. (1991) FLUORINE NUCLEAR-MAGNETIC-RESONANCE ANALYSIS OF THE LIGAND-BINDING PROPERTIES OF 2 HOMOLOGOUS RAT CELLULAR RETINOL-BINDING PROTEINS EXPRESSED IN ESCHERICHIA-COLI, *Journal of Biological Chemistry 266*, 3622-3629.

18. Calderone, V., Folli, C., Marchesani, A., Berni, R., and Zanotti, G. (2002) Identification and structural analysis of a zebrafish apo and holo cellular retinol-binding protein, *Journal of Molecular Biology 321*, 527-535.

19. Tarter, M., Capaldi, S., Carrizo, M. E., Ambrosi, E., Perduca, M., and Monaco, H. L. (2008) Crystal structure of human cellular retinol-binding protein 11 to 1.2 angstrom resolution, *Proteins-Structure Function and Bioinformatics 70*, 1626-1630.

20. Winter, N. S., Bratt, J. M., and Banaszak, L. J. (1993) CRYSTAL-STRUCTURES OF HOLO AND APO-CELLULAR RETINOL-BINDING PROTEIN-II, *Journal of Molecular Biology 230*, 1247-1259.

21. Kane, M. A., Bright, F. V., and Napoli, J. L. (2011) Binding affinities of CRBPI and CRBPII for 9-cis-retinoids, *Biochimica Et Biophysica Acta-General Subjects 1810*, 514-518.

22. Malpeli, G., Stoppini, M., Zapponi, M. C., Folli, C., and Berni, R. (1995) INTERACTIONS WITH RETINOL AND RETINOIDS OF BOVINE CELLULAR RETINOL-BINDING PROTEIN, *European Journal of Biochemistry 229*, 486-493.

23. McBee, J. K., Kuksa, V., Alvarez, R., de Lera, A. R., Prezhdo, O., Haeseleer, F., Sokal, I., and Palczewski, K. (2000) Isomerization of all-trans-retinol to cis-retinols in bovine retinal pigment epithelial cells: Dependence on the specificity of retinoid-binding proteins, *Biochemistry 39*, 11370-11380.

24. Borshchevskiy, V. I., Round, E. S., Popov, A. N., Bueldt, G., and Gordeliy, V. I. (2011) X-ray-Radiation-Induced Changes in Bacteriorhodopsin Structure, *Journal of Molecular Biology 409*, 813-825.

25. Matsui, Y., Sakai, K., Murakami, M., Shiro, Y., Adachi, S., Okumura, H., and Kouyama, T. (2002) Specific damage induced by X-ray radiation and structural changes in the primary photoreaction of bacteriorhodopsin, *Journal of Molecular Biology 324*, 469-481.

26. Schreckenbach, T., Walckhoff, B., and Oesterhelt, D. (1977) STUDIES ON RETINAL-PROTEIN INTERACTION IN BACTERIORHODOPSIN, *European Journal of Biochemistry 76*, 499-511.

27. Cowan, S. W., Newcomer, M. E., and Jones, T. A. (1993) CRYSTALLOGRAPHIC STUDIES ON A FAMILY OF CELLULAR LIPOPHILIC TRANSPORT PROTEINS - REFINEMENT OF P2 MYELIN PROTEIN AND THE STRUCTURE DETERMINATION AND REFINEMENT OF CELLULAR RETINOL-BINDING PROTEIN IN COMPLEX WITH ALL-TRANS-RETINOL, *Journal of Molecular Biology 230*, 1225-1246.

28. Vasileiou, C., Vaezeslami, S., Crist, R. M., Rabago-Smith, M., Geiger, J. H., and Borhan, B. (2007) Protein design: reengineering cellular retinoic acid binding protein II into a rhodopsin protein mimic, *J Am Chem Soc 129*, 6140-6148.

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

30. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *Journal of Applied Crystallography 30*, 1022-1025.

31. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP, *Acta Crystallographica Section D-Biological Crystallography 66*, 22-25.

32. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallographica Section D-Biological Crystallography 67*, 355-367.

33. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallographica Section D-Biological Crystallography 66*, 486-501.

34. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping - A mechanism for oligomer assembly, *Protein Science 4*, 2455-2468.

35. Crestfield, A. M., Stein, W. H., and Moore, S. (1962) On the aggregation of bovine pancreatic ribonuclease, *Archives of biochemistry and biophysics Suppl 1*, 217-222.

36. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: As domains continue to swap, *Protein Science 11*, 1285-1299.

37. Liu, Y. S., Gotte, G., Libonati, M., and Eisenberg, D. (2002) Structures of the two 3D domain-swapped RNase A trimers, *Protein Science 11*, 371-380.

38. Liu, Y. S., Gotte, G., Libonati, M., and Eisenberg, D. S. (2001) Crystal structure of an Rnase A dimer displaying a new type of 3D domain swapping, *Nucleic Acid Database*.

39. Liu, Y. S., Hart, P. J., Schlunegger, M. P., and Eisenberg, D. (1998) The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-angstrom resolution, *Proceedings of the National Academy of Sciences of the United States of America 95*, 3437-3442.

40. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: as domains continue to swap, *Protein science : a publication of the Protein Society 11*, 1285-1299.

41. Koharudin, L. M., Liu, L., and Gronenborn, A. M. (2013) Different 3D domain-swapped oligomeric cyanovirin-N structures suggest trapped folding intermediates, *Proc Natl Acad Sci U S A 110*, 7702-7707.

42. Assar, Z., Nossoni, Z., Wang, W., Santos, E. M., Kramer, K., McCornack, C., Vasileiou, C., Borhan, B., and Geiger, J. H. (2016) Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates, *Structure 24*, 1590-1598.

43. Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., Blundell, T. L., and Slingsby, C. (1990) X-ray analysis of beta-B2-crystallin and evolution of oligomeric lens proteins, *Nature 347*, 776-780.

44. Hakansson, M., Svensson, A., Fast, J., and Linse, S. (2001) An extended hydrophobic core induces EF-hand swapping, *Protein Science 10*, 927-933.

45. Rousseau, F., Schymkowitz, J., and Itzhaki, L. S. (2012) Implications of 3D domain swapping for protein folding, misfolding and function, In *Protein Dimerization and Oligomerization in Biology* (Mattews, J. M., Ed.), pp 137-152.

46. Yang, F., Bewley, C. A., Louis, J. M., Gustafson, K. R., Boyd, M. R., Gronenborn, A. M., Clore, G. M., and Wlodawer, A. (1999) Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping, *Journal of Molecular Biology 288*, 403-412.

47. Szilagyi, A., Zhang, Y., and Zavodszky, P. (2012) Intra-Chain 3D Segment Swapping Spawns the Evolution of New Multidomain Protein Architectures, *Journal of Molecular Biology 415*, 221-235.

48. Stroud, J. C., Liu, C., Teng, P. K., and Eisenberg, D. (2012) Toxic fibrillar oligomers of amyloid-beta have cross-beta structure, *Proceedings of the National Academy of Sciences of the United States of America 109*, 7717-7722.

49. Cleary, J. P., Walsh, D. M., Hofmeister, J. J., Shankar, G. M., Kuskowski, M. A., Selkoe, D. J., and Ashe, K. H. (2005) Natural oligomers of the amyloid-protein specifically disrupt cognitive function, *Nature Neuroscience 8*, 79-84.

50. Fandrich, M., Fletcher, M. A., and Dobson, C. M. (2001) Amyloid fibrils from muscle myoglobin - Even an ordinary globular protein can assume a rogue guise if conditions are right, *Nature 410*, 165-166.

51. Gronenborn, A. M. (2009) Protein acrobatics in pairs - dimerization via domain swapping, *Current Opinion in Structural Biology 19*, 39-49.

52. Schlunegger, M. P., Bennett, M. J., and Eisenberg, D. (1997) Oligomer formation by 3D domain swapping: A model for protein assembly and misassembly, In *Advances in Protein Chemistry, Vol 50: Protein Misassembly* (Wetzel, R., Ed.), pp 61-122.

53. Selkoe, D. J. (2004) Folding proteins in fatal ways (vol 426, pg 900, 2003), *Nature 428*, 445-445.

54. Shtilerman, M. D., Ding, T. T., and Lansbury, P. T. (2002) Molecular crowding accelerates fibrillization of alpha-synuclein: Could an increase in the cytoplasmic protein concentration induce Parkinson's disease?, *Biochemistry 41*, 3855-3860.

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

56. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Refined structure of dimeric diphtheria toxin at 2.0 A resolution, *Protein science : a publication of the Protein Society 3*, 1444-1463.

57. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D., and Zhang, K. Y. (2001) Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus, *Structure 9*, 1017-1027.

58. Ellis, R. J. (2001) Macromolecular crowding: an important but neglected aspect of the intracellular environment, *Curr Opin Struct Biol 11*, 114-119.

59. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Rlefined strucutre of dimeric diphtheria-toxin at 2.0-angstrom resolution, *Protein Science 3*, 1444-1463.

60. Carroll, S. F., Barbieri, J. T., and Collier, R. J. (1986) Dimeric form of diphtheria-toxin-purification and characterization, *Biochemistry 25*, 2425-2430.

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

62. Minton, A. P. (2000) Implications of macromolecular crowding for protein assembly, *Current Opinion in Structural Biology 10*, 34-39.

63. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D., and Zhang, K. Y. J. (2001) Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus, *Structure 9*, 1017-1027.

64. Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014) Energy landscapes of functional proteins are inherently risky, *Nature chemical biology 10*, 884-891.

65. Ignatova, Z., and Gierasch, L. M. (2004) Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling, *Proc Natl Acad Sci U S A 101*, 523-528.

66. Ignatova, Z., and Gierasch, L. M. (2006) Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant, *Proc Natl Acad Sci U S A 103*, 13357-13361.

67. Ropson, I. J., Boyer, J. A., Schaeffer, B. A., and Dalessio, P. M. (2009) Comparison of the folding mechanism of highly homologous proteins in the lipid-binding protein family, *Proteins: Structure, Function, and Bioinformatics 75*, 799-806.

68. Burns, L. L., and Ropson, I. J. (2001) Folding of intracellular retinol and retinoic acid binding proteins, *Proteins: Structure, Function, and Bioinformatics 43*, 292-302.

69. Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014) Energy landscapes of functional proteins are inherently risky, *Nat. Chem. Biol. 10*, 884-891.

70. Crist, R. M., Vasileiou, C., Rabago-Smith, M., Geiger, J. H., and Borhan, B. (2006) Engineering a Rhodopsin Protein Mimic, *Journal of the American Chemical Society 128*, 4522-4523.

71. Berbasova, T., Nosrati, M., Vasileiou, C., Wang, W., Lee, K. S., Yapici, I., Geiger, J. H., and Borhan, B. (2013) Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone, *J Am Chem Soc 135*, 16111-16119.

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

73. Wang, W., Nossoni, Z., Berbasova, T., Watson, C. T., Yapici, I., Lee, K. S., Vasileiou, C., Geiger, J. H., and Borhan, B. (2012) Tuning the electronic absorption of protein-embedded all-trans-retinal, *Science 338*, 1340-1343.

74. Nosrati, M., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. H. (2016) A Photoisomerizing Rhodopsin Mimic Observed at Atomic Resolution, *J Am Chem Soc 138*, 8802-8808.

75. Berbasova, T., Santos, E. M., Nosrati, M., Vasileiou, C., Geiger, J. H., and Borhan, B. (2016) Light-Activated Reversible Imine Isomerization: Towards a Photochromic Protein Switch, *Chembiochem : a European journal of chemical biology 17*, 407-414.

76. Yapici, I., Lee, K. S., Berbasova, T., Nosrati, M., Jia, X., Vasileiou, C., Wang, W., Santos, E. M., Geiger, J. H., and Borhan, B. (2015) "Turn-on" protein fluorescence: in situ formation of cyanine dyes, *J Am Chem Soc 137*, 1073-1080.

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

78. Gunasekaran, K., Hagler, A. T., and Gierasch, L. M. (2004) Sequence and structural analysis of cellular retinoic acid-binding proteins reveals a network of conserved hydrophobic interactions, *Proteins-Structure Function and Genetics 54*, 179-194.

79. Nossoni, Z., Assar, Z., Yapici, I., Nosrati, M., Wang, W., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal, *Acta Crystallogr. Sect. D. Biol. Crystallogr. 70*, 3226-3232.

80. Shameer, K., Shingate, P. N., Manjunath, S. C. P., Karthika, M., Pugalenthi, G., and Sowdhamini, R. (2011) 3DSwap: curated knowledgebase of proteins involved in 3D domain swapping, *Database-the Journal of Biological Databases and Curation*.

81. Sillitoe, I., Cuff, A. L., Dessailly, B. H., Dawson, N. L., Furnham, N., Lee, D., Lees, J. G., Lewis, T. E., Studer, R. A., Rentzsch, R., Yeats, C., Thornton, J. M., and Orengo, C. A. (2013) New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures, *Nucleic Acids Research 41*, D490-D498.

82. Englander, S. W., and Mayne, L. (2014) The nature of protein folding pathways, *Proc Natl Acad Sci U S A 111*, 15873-15880.

83. Englander, S. W., Mayne, L., and Krishna, M. M. (2007) Protein folding and misfolding: mechanism and principles, *Quarterly reviews of biophysics 40*, 287-326.

84. Krishna, M. M., and Englander, S. W. (2007) A unified mechanism for protein folding: predetermined pathways with optional errors, *Protein science : a publication of the Protein Society 16*, 449-464.

85. Mayne, L., and Englander, S. W. (2000) Two-state vs. multistate protein unfolding studied by optical melting and hydrogen exchange, *Protein science : a publication of the Protein Society 9*, 1873-1877.

86. Roder, H., Elove, G. A., and Englander, S. W. (1988) Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR, *Nature 335*, 700-704.

87. Martins, N. F., and Santoro, M. M. (1999) Partially folded intermediates during trypsinogen denaturation, *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas 32*, 673-682.

88. Medina, E., Cordova, C., Villalobos, P., Reyes, J., Komives, E. A., Ramirez-Sarmiento, C. A., and Babul, J. (2016) Three-Dimensional Domain Swapping Changes the Folding Mechanism of the Forkhead Domain of FoxP1, *Biophysical journal 110*, 2349-2360.

89. Zwanzig, R. (1997) Two-state models of protein folding kinetics, *Proceedings of the National Academy of Sciences of the United States of America 94*, 148-150.

90. Sanson, B., Wang, T., Sun, J., Wang, L., Kaczocha, M., Ojima, I., Deutsch, D., and Li, H. (2014) Crystallographic study of FABP5 as an intracellular endocannabinoid transporter, *Acta crystallographica. Section D, Biological crystallography 70*, 290-298.

91. Assar, Z., Nossoni, Z., Wang, W., Santos, Elizabeth M., Kramer, K., McCornack, C., Vasileiou, C., Borhan, B., and Geiger, James H. Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates, *Structure*.

92. Silvaroli, J. A., Arne, J. M., Chelstowska, S., Kiser, P. D., Banerjee, S., and Golczak, M. (2016) Ligand Binding Induces Conformational Changes in Human Cellular Retinol-

binding Protein 1 (CRBP1) Revealed by Atomic Resolution Crystal Structures, *The Journal of biological chemistry 291*, 8528-8540.

93. Gillilan, R. E., Ayers, S. D., and Noy, N. (2007) Structural basis for activation of fatty acid-binding protein 4, *J Mol Biol 372*, 1246-1260.

94. Sessler, R. J., and Noy, N. (2005) A ligand-activated nuclear localization signal in cellular retinoic acid binding protein-II, *Molecular cell 18*, 343-353.

95. Armstrong, E. H., Goswami, D., Griffin, P. R., Noy, N., and Ortlund, E. A. (2014) Structural Basis for Ligand Regulation of the Fatty Acid-binding Protein 5, Peroxisome Proliferator-activated Receptor β/δ (FABP5-PPARβ/δ) Signaling Pathway, *The Journal of biological chemistry 289*, 14941-14954.

96. Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, *Nature Protocols 1*, 2527-2535.

97. Ramos, C. R. R., Oyama, S., Jr., Sforca, M. L., Pertinhez, T. A., Ho, P. L., and Spisni, A. (2008) 2POA: Schistosoma mansoni Sm14 Fatty Acid-Binding Protein: improvement of protein stability by substitution of the single Cys62 residue, *Worldwide Protein Data Bank*.

98. Gill, S. C., and Vonhippel, P. H. (1989) CALCULATION OF PROTEIN EXTINCTION COEFFICIENTS FROM AMINO-ACID SEQUENCE DATA, *Analytical Biochemistry 182*, 319-326.

99. 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., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallographica Section D-Biological Crystallography 66*, 213-221.

100. Bailey, S. (1994) THE CCP4 SUITE - PROGRAMS FOR PROTEIN CRYSTALLOGRAPHY, *Acta Crystallographica Section D-Biological Crystallography 50*, 760-763.

101. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L.-W., Jain, S., Kapral, G. J., Kunstleve, R. W. G., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2011) The Phenix software for automated determination of macromolecular structures, *Methods 55*, 94-106.

102. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallographica Section D-Biological Crystallography 53*, 240-255.

103. Ong, D. E., and Page, D. L. (1987) Cellular retinol-binding protein (type two) is abundant in human small intestine, *Journal of Lipid Research 28*, 739-745.

104. Haunerland, N. H., and Spener, F. (2004) Fatty acid-binding proteins – insights from genetic manipulations, *Progress in Lipid Research 43*, 328-349.

105. Gillilan, R. E., Ayers, S. D., and Noy, N. (2007) Structural Basis for Activation of Fatty Acid-binding Protein 4, *Journal of Molecular Biology 372*, 1246-1260.

106. Kaczocha, M., Glaser, S. T., and Deutsch, D. G. (2009) Identification of intracellular carriers for the endocannabinoid anandamide, *Proceedings of the National Academy of Sciences of the United States of America 106*, 6375-6380.

Chapter IV: STRUCTURAL INSIGHTS INTO EMISSION REGULATION OF HCRBBPII-BOUND SOLVATOCHROMIC FLUOROPHORE

IV-1 Introduction

As one part of the collaboration (described in chapter II) between Prof. Borhan and Prof. Geiger lab, hCRBPII was reengineered so that it can bind to the fluorophores to make a new system with the overreaching goal of using these systems as fluorescent protein tags over the entire visible spectrum. [(75)]

This new class of fluorogenic proteins was created by binding fluorophore aldehydes in the binding pocket of hCRBPII via protonated Schiff base (PSB) formation, which as a result is sensitive to pH. [(111)] Numerous applications can be intended such as application in multi-color imaging and *in vivo* pH sensing.[(112)] One of the other benefits of this system is that no oxygen is required (unlike green fluorescent protein, GFP) and therefore it could find potential in obligate anaerobes. Dr. Santos in prof. Borhan's group has synthesized a new solvatochromic fluorophore, ThioFlour, and developed this fluorogenic system. **(Figure IV-1c)**

Solvatochromism describes the alteration in wavelength (and intensity), of an absorption and/or emission band that comes from the polarity change of its medium.[(113)], [(114)] Since specificity is crucial in engineering a practical and efficient fluorogenic system, solvatochromic fluorophores were chosen to couple with the hCRBPII variants. Also previous studies with a merocyanine aldehyde showed non-specific background fluorescence in imaging of live cells (arising from the iminium formation with nonspecific proteins), due to the static excitation and emission profiles of these fluorophore bound hCRBPII complexes. [(111)]

In this new system, emission of solvatochromic fluorophores is flexible based on the polarity of their environment. Therefore multicolor probes can be developed. More importantly, due to possibility of the hCRBPII/fluorophore emission tuning, we can reduce nonspecific labeling. This would result in a probe that yields little background fluorescence.

IV-2 Rational behind the parent template: Q108K:K40L:T51V:T53S

Most solvatochromic dyes show relatively low fluorescence quantum yields in polar solvents and high fluorescence quantum yields in apolar solvents.[114],[115] ThioFlour spectroscopic data in various solvents suggests the same result. (Performed by Dr. Santos, Data not shown) It also shows no fluorescent in PBS solutions (can be considered as no fluorescence being observed of unbound probe during *in vivo* applications). This could be beneficial in live cell imaging. Her studies also show that ThioFluor-PSB is much less solvatochromic than the ThioFluor aldehyde, indicating that PSB emission does not follow the same solvatochromic trend and some other factors, such as hydrogen bonding, may affect the emission wavelength of ThioFlour-PSB.[116]

To peruse the study of ThioFluor bound protein interactions, Dr. Santos generated a large number of hCRBPII variants by point mutation. **Table IV-1** shows only the ones to being discussed in this chapter. For this purpose the polarity can be removed in two ways: 1) Removal of the more polar residues in the binding pocket nearest the chromophore and 2) Removal of the direct water molecules interacting with the iminium region. Removal of these negative dipoles in the PSB region should encourage positive charge delocalization.[76]

As previously mentioned in Chapter II, hCRBPII has a large hydrophobic binding pocket, with few polar residues.[76] Mutagenesis was first performed on T51, T53, Q4, Q38 and R58

residues (**Figure IV-1**), mutating them to similarly sized hydrophobic residues in the parent double mutant Q108K:K40L (KL) form. KL was used. Q108K couldn't be used due to its precipitation in acidic pH. [76] Additionally, we only discuss the monomer spectroscopic properties here. Since T51V is capable of monomer formation of hCRBPII variants, it was thus retained in most mutants tested.

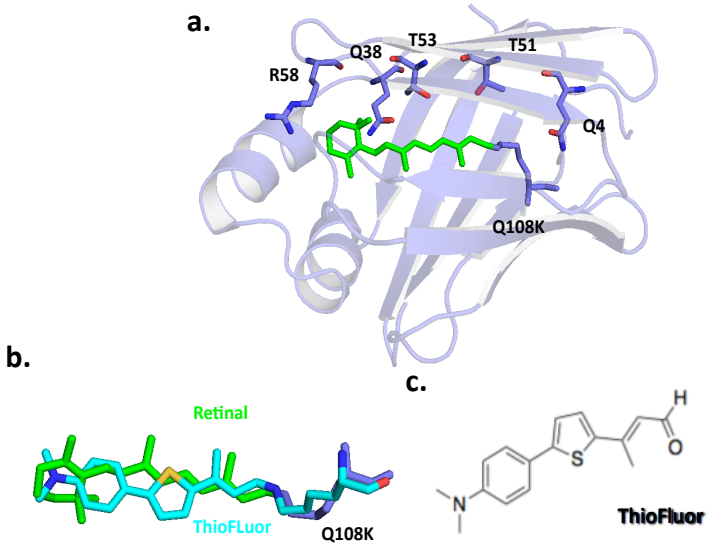


Figure IV-1: a. Polar residues around the ligand mutated in order to remove polarity from the binding pocket. Coordinates obtained from PDB 4EXZ (hCRBPII-Q108K:K40L/retinal). **b.** Both retinal and ThioFluor bind and lay in the same region of the binding ppocket. (the Thioflour coordinates are from Q108K:K40L: T51V:T53S: R58Y structure, will be discussed later in the IV-3 section. **c.** The chemical Structure of the ThioFluor , which was synthesized by Dr. Santos

T53 is located near the middle of the polyene in the crystal structure of KL/retinal, as shown in **Figure IV-2a** making a hydrogen bond to T51 through a water molecule. In order to alter its polarity, T53 was replaced with various residues including alanine and serine in the parent template KL.

Interestingly, all these substitutions led to at least a slight red shift in absorption. From KL the change in absorption was 5 nm, 7 nm (3nm and 8nm in their emission wavelength) for T53A and T53S respectively (**Table IV-1**). As it was mentioned before, T51 and T53 interact via a water-mediated hydrogen bond in the crystal structure (of KL bound retinal, **Figure IV-2a**). By mutating T51 to valine, this interaction would be removed, leading to a red shift in wavelength. T53S shows a similar effect. There is a water network between Q38 and Q128 (**Figure IV-2b)**. T53S could potentially form a hydrogen bond with Q38, which would disrupt the conserved water-mediated interaction between Q38 and Q128 by ejection of water, and thus change the polarity, near the ThioFluor.

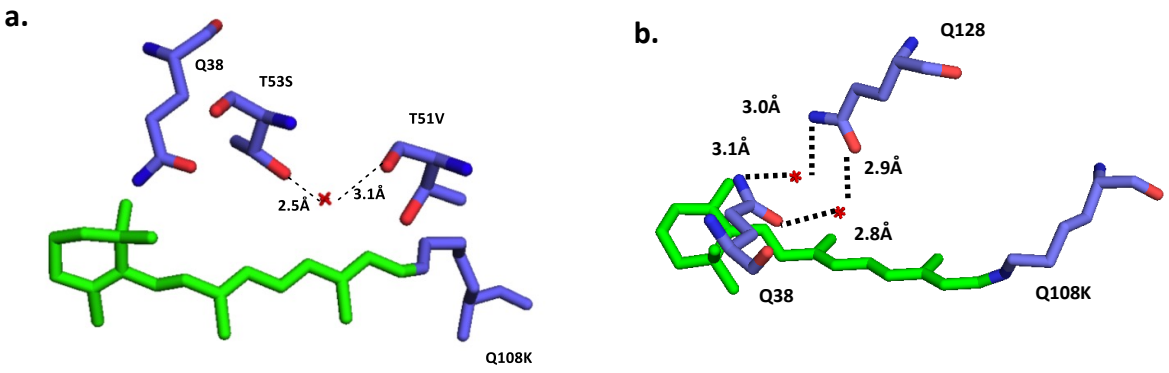


Figure IV-2: a. Water network hydrogen bonding between T51 and T53. **b.** Water mediated hydrogen bonding between Q38 and Q128. (PDB 4EXZ, hCRBPII-Q108K:K40L/retinal, Retinal is represented in green sticks).

To test this hypothesis, Dr. Santos generated Q108K:K40L:T53A:R58F. This mimics the T53S mutation and elimination of the water mediated interaction between Q38 and Q128 leading to removing polarity in the pocket. In case of residue 58 (R58), it was mutated to aromatic residues, tyrosine and phenylalanine in KL and both led to a bathochromic shift in absorption

11nm and 17nm and emission 7 and 4nm with ThioFluor respectively (**Table IV-1**). This result agrees with previous studies that showed that isolating the retinal chromophore in the binding pocket of hCRBPII resulted in a bathochromic shift due to more effective shielding of the binding pocket. [75, 76, 112]

It should be noted that residues closer to the iminium, Q4 and T51, led to a larger red-shift than T53 and Q38 residues farther away. Therefore, the template Q108K:K40L:T51V:T53S was used for further protein engineering by Dr. Santos. However Q4 mutations were not included, due to prior studies showing that the pK_a of hCRBPII/retinal complexes are severely low when Q4 was mutated. [76]

Table IV-1: Spectroscopic change of hCRBPII variants bound ThioFluor

Entry	Mutant	λ_{abs}	λ_{em}	φ
1	Q108K:K40L	580	674	nd
2	Q108K:K40L:T51V	608	690	0.12
3	Q108K:K40L:T53A	585	677	0.11
4	Q108K:K40L:T53S	587	682	0.08
5	Q108K:K40L: T51V:T53S	627	697	0.15
6	Q108K:K40L:Q38F	588	680	0.13
7	Q108K:K40L: T51V:T53S: R58H	643	705	0.14
8	Q108K:K40L: T51V:T53S: R58W	623	697	0.08
9	Q108K:K40L: T51V:T53S: R58Y	622	691	0.11
10	Q108K:K40L: T53A: R58F	589	681	0.12
11	Q108K:K40L: T51V:T53S: R58W:Y19W	653	719	0.10
12	Q108K:K40L: T51V:T53S: R58W: Y19W: A33W	666	724	0.14
13	Q108K:K40L: T51V:T53S: R58W:Y19W:L117E	558	673	0.31
14	Q108K:K40L: T51V:T53S: R58W:Y19W: A33W: L117E	539	628	0.28

IV-3 Attempts to isolate the binding cavity

Due to previous studies conducted by Dr. Wenjing Wang and Dr. Nossoni, an aromatic residue at position R58 leads to a bathochromic shift in absorption when retinal is used as a ligand, due to more effective shielding of the binding poc ket. [(75, 76, 112) (117)]Therefore R58F, R58Y, R58W and R58H mutations were introduced into the parent mutant

Q108K:K40L:T51V:T53S by Dr. Santos in this project. It should be noted that, unlike retinal, ThioFluor contains heteroatoms that can interact with surrounding residues through hydrogen bonding, hydrophobic interactions, pi-pi stacking, pi-cation interactions, ionic interactions, etc. Crystal Structure of Q108K:K40L: T51V:T53S: R58W/ThioFLuor was obtioned by Dr. Nosrati. Based on its structure, R58W sits directly on top of the *N,N*-dimethyl amino moiety of the bound ThioFluor. As shown (**Figure IV-3)**, the distance between the tryptophan and nitrogen is 3.7 Å, well within van der Waals' contact.

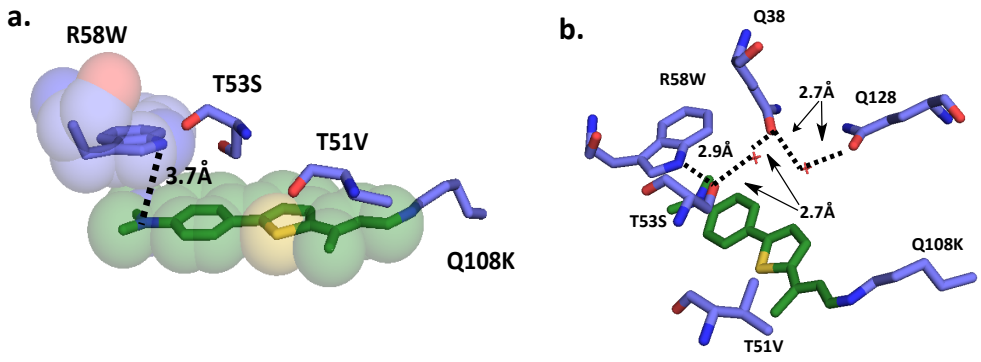


Figure IV-3: a. Crystal structure of Q108K:K40L:T51V:T53S:R58W/ThioFluor and space-filling representation obtained by Dr. Nosrati. R58W Is packed with the *N,N*-dimethyl amino moiety of the ThioFlour tail. **b.** The water mediated hydrogen bonding between T53S and R58W is highlighted.

In that series R58H mutation caused 16nm in absorption and 8nm in emission from Q108K:K40L:T51V:T53S. Unfortunately, crystallization of Q108K:K40L:T51V:T53S:R58H was not successful.

In the R58Y mutation series, tyrosine is flipped out of the pocket, providing no stacking interaction with the ligand (which yielded a much smaller wavelength alteration than Q108K:K40L:T51V:T53S.) Alternatively, the R58Y participates in a water-mediated hydrogen

bond with the hydroxyl side chain of T29 (**Figure IV-4**). As the ligand is not locked in place by the R58 mutation, this presumably causes the translocation and rotation of the chromophore.

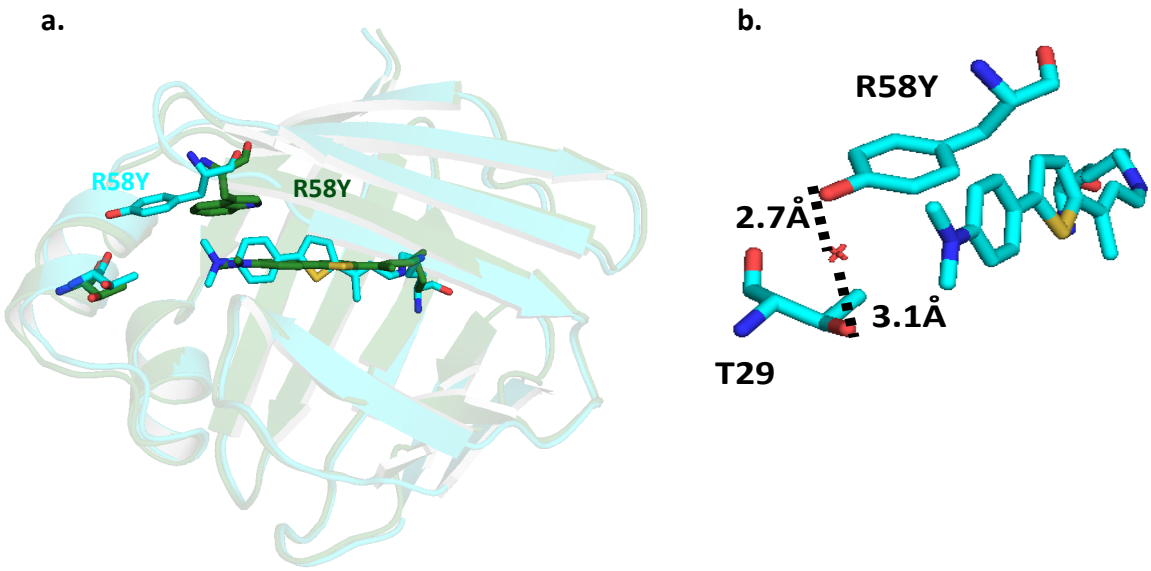


Figure IV-4: a. Overlay of Crystal structures of Q108K:K40L:T51V:T53S:R58W/ThioFluor (cyan) with Q108K:K40L:T51V:T53S:R58Y/ThioFluor (green). **b.** It shows Water mediated hydrogen bonding between R58Y and T29.

Between Q108K:K40L:T51V:T53S:R58W/ThioFluor and Q108K:K40L:T51V:T53S:R58Y/ThioFluor, ligand trajectories are perpendicular , both have cis iminium bonds. However R58Y makes a hydrogen bonding network through a water molecule and Thr29. On the other hand, Trp 58 packed with the tail of ligand and, makes a pi cation interaction with the ligand, therefore causing red shift emission for R58W.

Most remarkably, upon overlay of the structure with both Q108K:K40L:T51V:T53S:R58W/ThioFluor and Q108K:K40L:T51V:T53S:R58Y/ThioFluor, the flexibility of the loop at the entrance of the binding pocket can be seen (**Figure IV-4**). Although

the crystal structure of Q108K:K40L:T51V:T53S:R58F/ThioFluor was not obtained, the structure of the similar variant Q108K:K40L:T53A:R58F with ThioFluor was solved**. (Figure IV-5)**

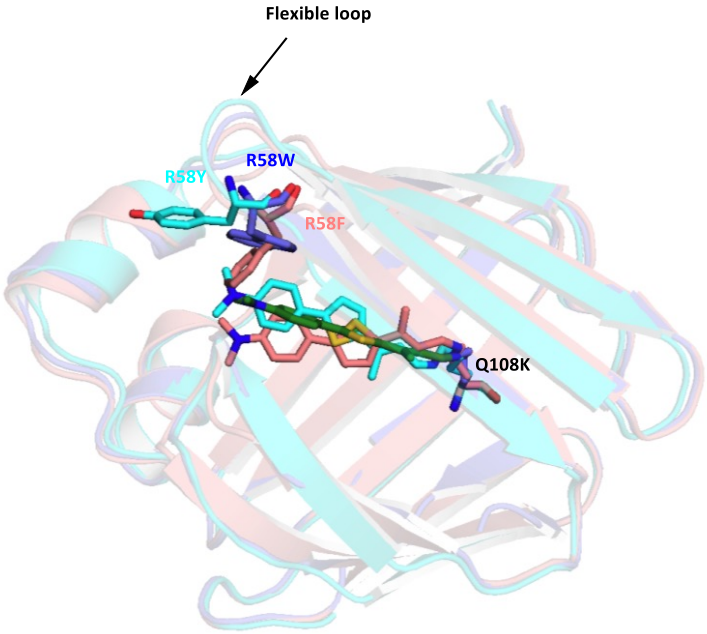


Figure IV-5: Overlay of Crystal structures of Q108K:K40L:T51V:T53S:R58W/ThioFluor (cyan) with Q108K:K40L:T51V:T53S:R58Y/ThioFluor (green) and Q108K:K40L:T53A:R58F/ThioFluor (salmon).

As observed, with an R58F mutation, the flexible loop allows movement such that phenylalanine is within van der Waals contact of the ligand. A slight bathochromic shift of 4 nm in absorption is observed for Q108K:K40L:T53A:R58F from Q108K:K40L:T53A, suggesting again that R58F does not successfully sequester the binding cavity**. (Table IV-1 and Figure IV-5)** By comparing Q108K:K40L:T51V:T53S:R58Y/ThioFluor and, Q108K:K40L:T53A:R58F/ThioFluor structures, a 180° difference in ligand trajectory has been revealed. In R58F due to the position of Phe 58, the benzene ring of the ligand has moved lower compared to R58Y**.(Figure IV-5)** Further crystallographic analysis would be required in order to determine the effect of other

Arg58 mutations as they are shown to dramatically alter the chromophore orientation.

IV-4 Extensive packing of the ligand and removal of water leads to bathochromic shift in absorption and emission.

Based on the structure of Q108K:K40L:T51V:T53S:R58W/ThioFluor, there is a conserved water network from the phenol of Tyr 19 to the thiophene sulfur of ThioFluor (**Figure IV-6)** As previously shown, removal of polarity yields to red-shift wavelength. Y19W mutation (less likely to provide hydrogen bonding, conducted by Dr. Santos) would result in removal of this water network. A red shift in absorption of 30 nm is observed with the introduction of Y19W to Q108K:K40L:T51V:T53S:R58W (**Table IV-1**).

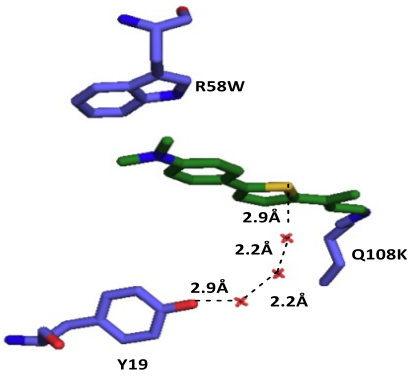


Figure IV-6: The water mediated hydrogen bonding between Tyr19 to the thiophene sulfur of ThioFluor highlighted in the Crystal structure of Q108K:K40L:T51V:T53S:R58W/ThioFluor obtained by Dr. Nosrati.

To confirm this hypothesis, the crystal structure of Q108K:K40L:T51V:T53S:R58W:Y19W/ThioFluor was determined, and as expected shows that the water network was indeed abolished, abrogating the water-mediated interation with the fluorophore **(Figure IV-7).** Moreover, the hydroxyl of Ser 53 hydrogen bonds with both the

indole side chain of Trp 58 and the amide moiety of Gln 38 through a water mediated hydrogen bond (**Figure IV-3b)**. These interactions presumably dictate the chromophore's orientation. Between Q108K:K40L:T51V:T53S:R58W and Q108K:K40L:T51V:T53S:R58W:Y19W, the only difference is the Y19W mutation, which removes the water network and makes a pi stacking interaction with the ligand, encapsulating the fluorophore and inducing a vast red shift in emission. (**Figure IV-7 and 9)**

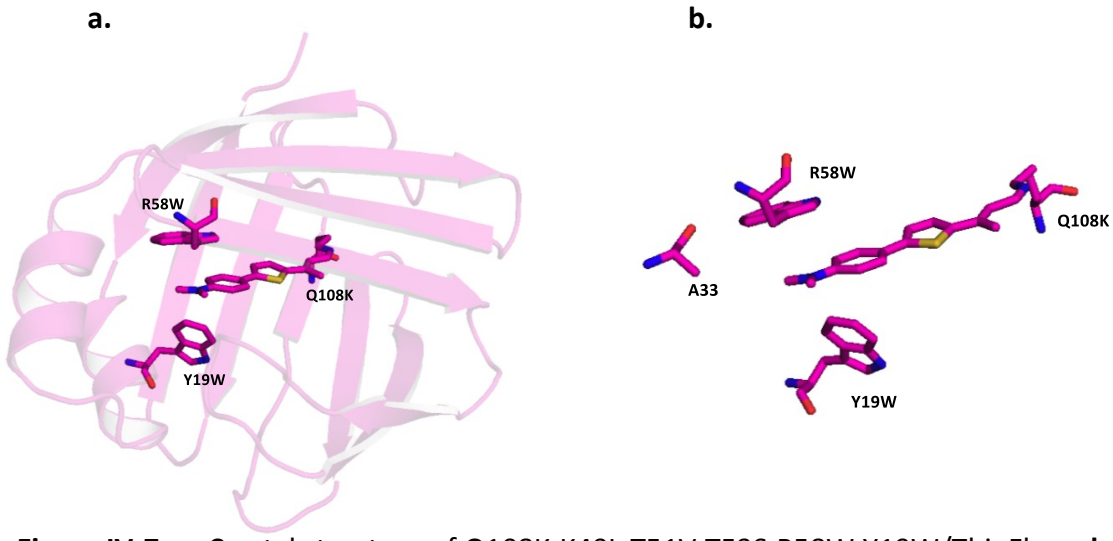


Figure IV-7: a. Crystal structure of Q108K:K40L:T51V:T53S:R58W:Y19W/ThioFluor. **b.** Water network is abolished in this structure. ThioFluor and residues around it are represented in sticks.

The last attempt in this series was to encapsulate the binding cavity, via the introduction of one more bulky tryptophan residue at A33, Q108K:K40L:T51V:T53S:R58W:Y19W (**Figure IV-7b**). Unfortunately, crystallization of Q108K:K40L:T51V:T53S:R58W:Y19W:A33W was not successful. Presumably, A33W adequately covers the binding cavity. (See **Figure IV-8** for a similar mutant) While efforts to crystallize Q108K:K40L:T51V:T53S:R58W:Y19W:A33W/ThioFluor were not successful, a crystal structure of

Q108K:K40L:T51V:T53S:R58W:Y19W:A33W:L117E/ThioFluor was achieved by Alireza Ghanbarpour. **(Figure IV-8 and 9)** It shows that A33W closes the lid of the binding pocket. Combination of R58W, Y19W and A33W encapsulates the binding cavity.

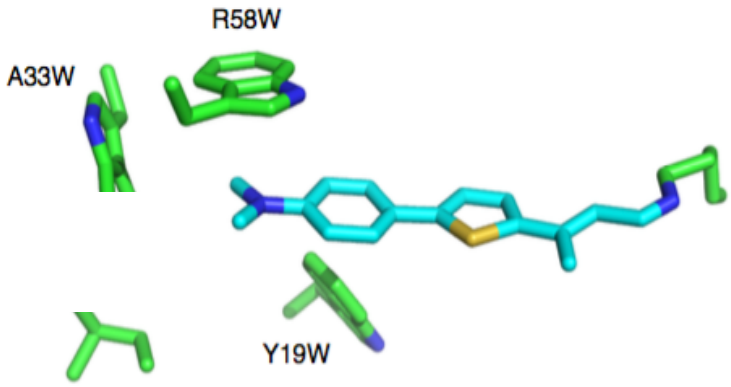


Figure IV-8: Crystal structure of Q108K:K40L:T51V:T53S:R58W:Y19W:A33W:L117E/ThioFluor obtained by Alireza. ThioFluor in cyan and residues around it in green are shown in sticks.

IV-5 Localization of charge leads to a blue shift in absorption and emission

Previouosly with hCRBPII/retinal, Dr. Wang and Dr. Nossoni demonstrated that an even distribution of electrostatic potential across the polyene led to a bathochromic shift in absorption, and by localizing the cation on the iminium nitrogen it led to a blue shift. Similarly, we sought to introduce negatively polarized residues near the iminium. **(Figure IV-9)**

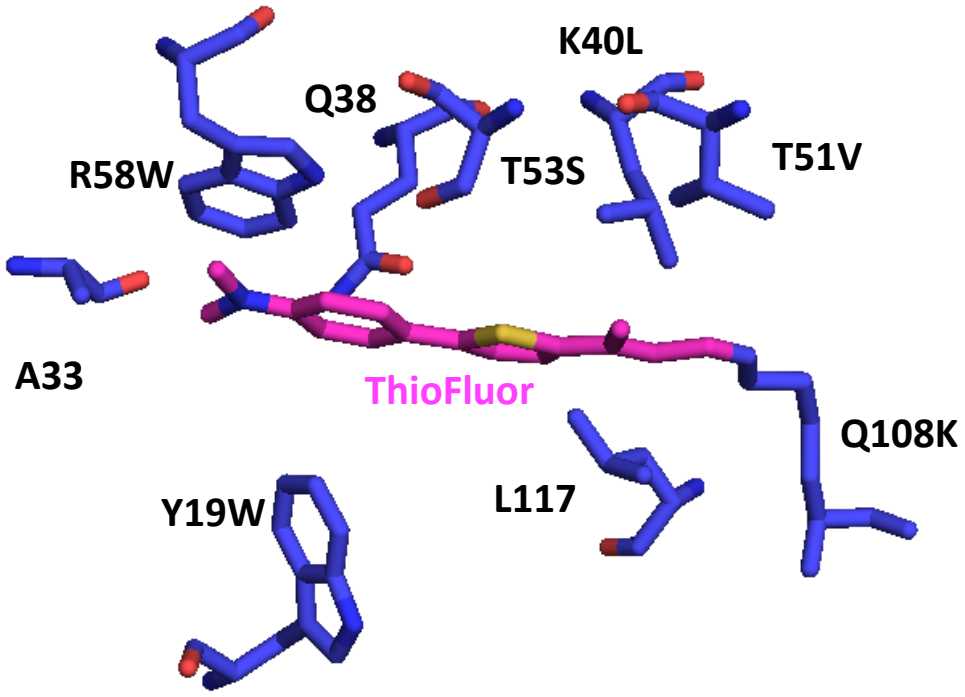


Figure IV-9: Mutated residues which discussed in this chapter (show in blue) in an attempt to encapsulate the binding cavity. Crystal structure is of Q108K:K40L:T51V:T53S:R58W:Y19W/ThioFluor.

To this end, Leu117 was mutated to glutamic acid in Q108K:K40L:T51V:T53S:R58W:Y19W and Q108K:K40L:T51V:T53S:R58W:Y19W:A33W, both of which when bound with ThioFluor producing a clearly observable red-shifted iminium. (Performed by Dr. Santos)

Between Structures of Q108K:K40L:T51V:T53S:R58W:Y19W and Q108K:K40L:T51V:T53S:R58W:Y19W:L117E they have cis and trans iminium bond respectively. L11E mutant makes hydrogen bonds with iminnium bond and localizes the positive charge of iminium. It also gives the imine trans over the cis isomer. Additionally, the entire ligand moved

in the horizontal direction as well. This explains why the carboxylate blue shifts 100nm and 46nm for absorption and emission respectively. **(Figure IV-14)**

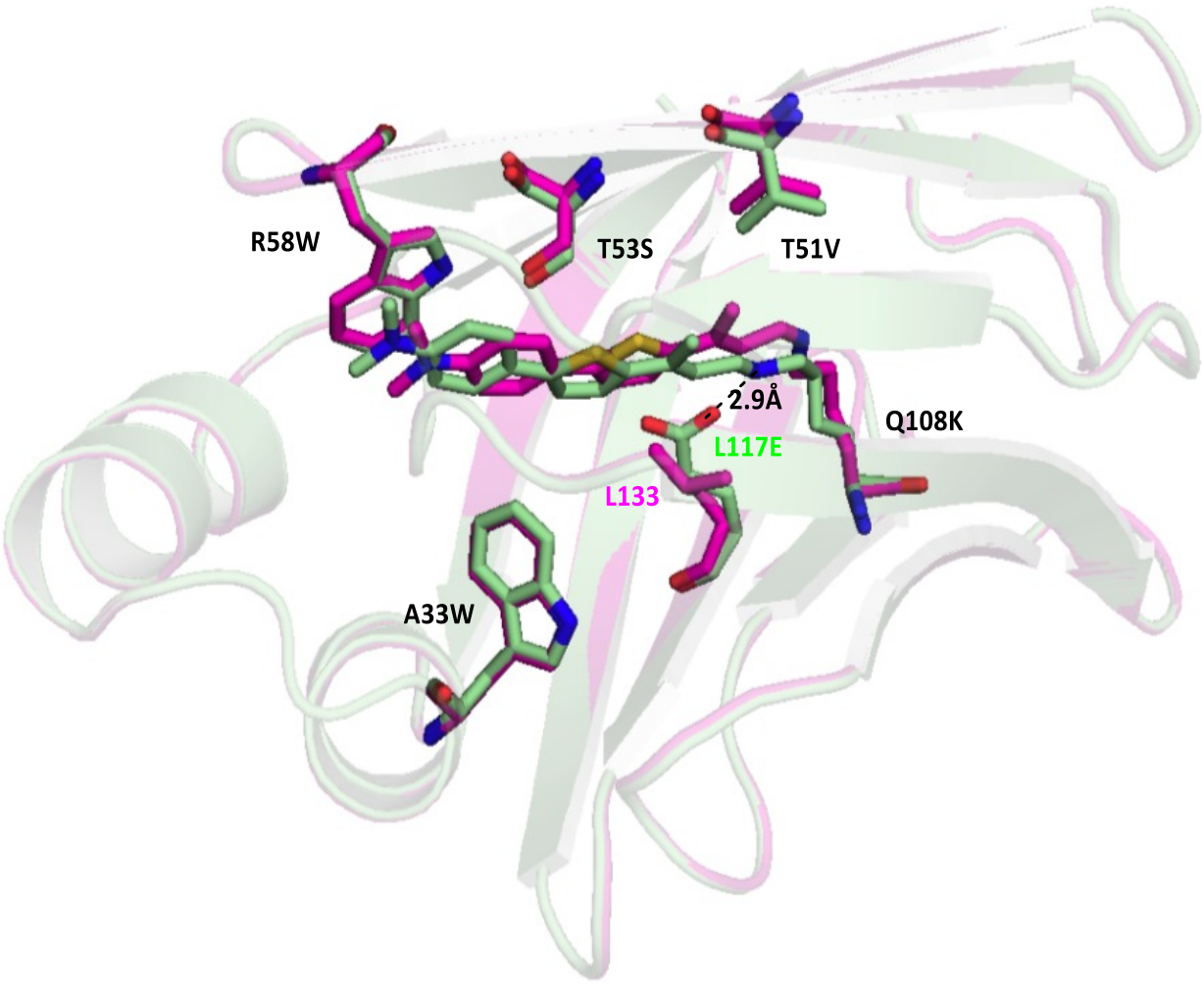


Figure IV-10: Crystal structure of Q108K:K40L:T51V:T53S:R58W:Y19W:A33W:L117E/ThioFluor. ThioFluor and highlighted residues are represented as sticks.

As other variants in these series, we can point out Q108K:K40L:T53A:R58F which lead to a more intense iminium species **(Figure IV-7).** As it was mentioned earlier in this chapter, our overreaching goal is to develop fluorescent proteins that span the entire visible spectra. By now

in this project, the absorption maxima are tuned from 501nm to 705nm and emission maxima from 613 nm to 744 nm. Covering both the red and far-red fluorescence wavelength regimes.

This fluorogenic system even was tested in mammalian cells for live cell imaging by Dr. Santos.[(111)] The data shows that ThioFluor is cell permeable and displays no background fluorescence when excited in the red, even when long incubation periods with ThioFluor are required, unlike merocyanine aldehyde. (Data not shown)

IV-6 Material and Methods

Designed hCRBPII mutants were obtained Dr. Santos. Mutants were expressed and purified according to the protocol described in Chapter I. The purified monomers were then used for crystallization trials.

Each mutant protein was concentrated to 8mg/mL in a buffer containing 10 mM Tris, 100 mM NaCl, pH = 8.0. Proteins were incubated with 4 equivalent of ThioFluor ligand (stock concentration of 0.014M in DMSO) in room temperature overnight. Crystals were grown using the hanging drop vapor diffusion method using 1 µL of protein solution and 1 µL of crystallization solution in the drop. The best crystals grew using a crystallization solution of 25-40% PEG4000, 0.1 M $NaCH_3CO_2Na$ pH 4.5 0.1 M $NH_4CH_3CO_2$. The crystals were soaked in a cryoprotectant solution (same crystallization solution with 15% glycerol), flash frozen in loops in liquid nitrogen and stored in a liquid nitrogen Dewar prior to data collection.

Data collection and solving the structures for the crystallized mutants (Q108K:K40L:T53A:R58F, Q108K:K40L:T51V:T53S:R58Y, Q108K:K40L:T51V:T53S:R58W:Y19W and Q108K:K40L:T51V:T53S:R58W:Y19W:L117E, **Table IV-2**) were performed according to the

method described in Chapter I. The chromophore using the JLIGAND program to generate restraints.[104, 105]

Table IV-2: X-ray crystallographic data and refinement statistics for monomeric hCRBPII mutants.

	Q108K:K40L:T53A:R58F	Q108K:K40L:T51A: T53S:R58Y	Q108K:K40L:T51A: T53S:R58W:Y19W	Q108K:K40L:T51A: T53S:R58W:Y19W:L117E
Space group	P1	P1	P1	P1
a(Å)	36.51	32.29	29.32	29.40
b(Å)	55.16	35.25	36.43	36.39
c (Å)	68.87	64.87	64.08	64.27
α(°)	109.94	92.51	89.74	89.84
β(°)	98.49	90.42	89.36	89.66
δ(°)	101.41	111.72	66.38	66.15
Molecules per Asymmetric Unit	4	2	2	2
Total reflection	597246	268817	544327	623264
Unique Reflection	82805	47165	72294	69070
Completeness (%)	95.71(91.0) [a]	79.63(94.6)[a]	93.53(90.0)[a]	94.8(89.0)[a]
Average I/σ	32.9(2.01) [a]	43.37(1.64)[a]	44.63(6.9)[a]	31.4(1.47)[a]
Resolution (Å) (Last Shell)	27.4-1.47(1.51-1.41)[a]	29.2-1.34(1.37-1.34)[a]	29.60-1.220(1.25-1.29)[a]	29.55-1.24(1.27-1.23)[a]
R_{work}/ R_{free} (%)	23.3/26.3	23.8/28.2	17.9/18.3	17.3/18.5
RMSD from ideal value				
Bond Length (Å)	0.008	0.013	0.01	0.007
Bond Angle (°)	1.249	1.56	1.58	1.30
Average B factor	15.89	20.58	12.07	13.66
Number of water molecules	513	284	321	211

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

REFERENCES

REFERENCES

1. Ong, D. E. (1984) A NOVEL RETINOL-BINDING PROTEIN FROM RAT - PURIFICATION AND PARTIAL CHARACTERIZATION, *Journal of Biological Chemistry 259*, 1476-1482.

2. Harder, M. E., Deinzer, M. L., Leid, M. E., and Schimerlik, M. I. (2004) Global analysis of three-state protein unfolding data, *Protein Science 13*, 2207-2222.

3. Nossoni, Z., Assar, Z., Yapici, I., Nosrati, M., Wang, W., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal, *Acta crystallographica. Section D, Biological crystallography 70*, 3226-3232.

4. Veerkamp, J. H., Vankuppevelt, T., Maatman, R., and Prinsen, C. F. M. (1993) STRUCTURAL AND FUNCTIONAL-ASPECTS OF CYTOSOLIC FATTY-ACID-BINDING PROTEINS, *Prostaglandins Leukotrienes and Essential Fatty Acids 49*, 887-906.

5. Banaszak, L., Winter, N., Xu, Z. H., Bernlohr, D. A., Cowan, S., and Jones, T. A. (1994) LIPID-BINDING PROTEINS - A FAMILY OF FATTY-ACID AND RETINOID TRANSPORT PROTEINS, *Advances in Protein Chemistry, Vol 45 45*, 89-151.

6. Vaezeslami, S., Mathes, E., Vasilelou, C., Borhan, B., and Geiger, J. H. (2006) The structure of apo-wild-type cellular retinoic acid binding protein II at 1.4 A and its relationship to ligand binding and nuclear translocation, *Journal of Molecular Biology 363*, 687-701.

7. Haunerland, N. H., and Spener, F. (2004) Fatty acid-binding proteins--insights from genetic manipulations, *Progress in lipid research 43*, 328-349.

8. Lucke, C., Huang, S., Rademacher, M., and Ruterjans, H. (2002) New insights into intracellular lipid binding proteins: The role of buried water, *Protein science : a publication of the Protein Society 11*, 2382-2392.

9. Schaap, F. G., van der Vusse, G. J., and Glatz, J. F. (2002) Evolution of the family of intracellular lipid binding proteins in vertebrates, *Molecular and cellular biochemistry 239*, 69-77.

10. Smathers, R. L., and Petersen, D. R. (2011) The human fatty acid-binding protein family: evolutionary divergences and functions, *Human genomics 5*, 170-191.

11. Franzoni, L., Lucke, C., Perez, C., Cavazzini, D., Rademacher, M., Ludwig, C., Spisni, A., Rossi, G. L., and Ruterjans, H. (2002) Structure and backbone dynamics of apo- and holo-

cellular retinol-binding protein in solution, *Journal of Biological Chemistry 277*, 21983-21997.

12. Crow, J. A., and Ong, D. E. (1985) CELL-SPECIFIC IMMUNOHISTOCHEMICAL LOCALIZATION OF A CELLULAR RETINOL-BINDING PROTEIN (TYPE-2) IN THE SMALL-INTESTINE OF RAT, *Proceedings of the National Academy of Sciences of the United States of America 82*, 4707-4711.

13. Vogel, S., Mendelsohn, C. L., Mertz, J. R., Piantedosi, R., Waldburger, C., Gottesman, M. E., and Blaner, W. S. (2001) Characterization of a new member of the fatty acid-binding protein family that binds all-trans-retinol, *Journal of Biological Chemistry 276*, 1353-1360.

14. Folli, C., Calderone, V., Ottonello, S., Bolchi, A., Zanotti, G., Stoppini, M., and Berni, R. (2001) Identification, retinoid binding, and x-ray analysis of a human retinol-binding protein, *Proceedings of the National Academy of Sciences of the United States of America 98*, 3710-3715.

15. Li, E., Demmer, L. A., Sweetser, D. A., Ong, D. E., and Gordon, J. I. (1986) ISOLATION OF A CLONED CDNA-ENCODING AN ABUNDANT INTESTINAL RETINOL BINDING-PROTEIN, *Clinical Research 34*, A685-A685.

16. Li, E., and Norris, A. W. (1996) Structure/function of cytoplasmic vitamin A-binding proteins, *Annual Review of Nutrition 16*, 205-234.

17. Li, E., Qian, S. J., Winter, N. S., Davignon, A., Levin, M. S., and Gordon, J. I. (1991) FLUORINE NUCLEAR-MAGNETIC-RESONANCE ANALYSIS OF THE LIGAND-BINDING PROPERTIES OF 2 HOMOLOGOUS RAT CELLULAR RETINOL-BINDING PROTEINS EXPRESSED IN ESCHERICHIA-COLI, *Journal of Biological Chemistry 266*, 3622-3629.

18. Calderone, V., Folli, C., Marchesani, A., Berni, R., and Zanotti, G. (2002) Identification and structural analysis of a zebrafish apo and holo cellular retinol-binding protein, *Journal of Molecular Biology 321*, 527-535.

19. Tarter, M., Capaldi, S., Carrizo, M. E., Ambrosi, E., Perduca, M., and Monaco, H. L. (2008) Crystal structure of human cellular retinol-binding protein 11 to 1.2 angstrom resolution, *Proteins-Structure Function and Bioinformatics 70*, 1626-1630.

20. Winter, N. S., Bratt, J. M., and Banaszak, L. J. (1993) CRYSTAL-STRUCTURES OF HOLO AND APO-CELLULAR RETINOL-BINDING PROTEIN-II, *Journal of Molecular Biology 230*, 1247-1259.

21. Kane, M. A., Bright, F. V., and Napoli, J. L. (2011) Binding affinities of CRBPI and CRBPII for 9-cis-retinoids, *Biochimica Et Biophysica Acta-General Subjects 1810*, 514-518.

22. Malpeli, G., Stoppini, M., Zapponi, M. C., Folli, C., and Berni, R. (1995) INTERACTIONS WITH RETINOL AND RETINOIDS OF BOVINE CELLULAR RETINOL-BINDING PROTEIN, *European Journal of Biochemistry 229*, 486-493.

23. McBee, J. K., Kuksa, V., Alvarez, R., de Lera, A. R., Prezhdo, O., Haeseleer, F., Sokal, I., and Palczewski, K. (2000) Isomerization of all-trans-retinol to cis-retinols in bovine retinal pigment epithelial cells: Dependence on the specificity of retinoid-binding proteins, *Biochemistry 39*, 11370-11380.

24. Borshchevskiy, V. I., Round, E. S., Popov, A. N., Bueldt, G., and Gordeliy, V. I. (2011) X-ray-Radiation-Induced Changes in Bacteriorhodopsin Structure, *Journal of Molecular Biology 409*, 813-825.

25. Matsui, Y., Sakai, K., Murakami, M., Shiro, Y., Adachi, S., Okumura, H., and Kouyama, T. (2002) Specific damage induced by X-ray radiation and structural changes in the primary photoreaction of bacteriorhodopsin, *Journal of Molecular Biology 324*, 469-481.

26. Schreckenbach, T., Walckhoff, B., and Oesterhelt, D. (1977) STUDIES ON RETINAL-PROTEIN INTERACTION IN BACTERIORHODOPSIN, *European Journal of Biochemistry 76*, 499-511.

27. Cowan, S. W., Newcomer, M. E., and Jones, T. A. (1993) CRYSTALLOGRAPHIC STUDIES ON A FAMILY OF CELLULAR LIPOPHILIC TRANSPORT PROTEINS - REFINEMENT OF P2 MYELIN PROTEIN AND THE STRUCTURE DETERMINATION AND REFINEMENT OF CELLULAR RETINOL-BINDING PROTEIN IN COMPLEX WITH ALL-TRANS-RETINOL, *Journal of Molecular Biology 230*, 1225-1246.

28. Vasileiou, C., Vaezeslami, S., Crist, R. M., Rabago-Smith, M., Geiger, J. H., and Borhan, B. (2007) Protein design: reengineering cellular retinoic acid binding protein II into a rhodopsin protein mimic, *J Am Chem Soc 129*, 6140-6148.

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

30. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *Journal of Applied Crystallography 30*, 1022-1025.

31. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP, *Acta Crystallographica Section D-Biological Crystallography 66*, 22-25.

32. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallographica Section D-Biological Crystallography 67*, 355-367.

33. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallographica Section D-Biological Crystallography 66*, 486-501.

34. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping - A mechanism for oligomer assembly, *Protein Science 4*, 2455-2468.

35. Crestfield, A. M., Stein, W. H., and Moore, S. (1962) On the aggregation of bovine pancreatic ribonuclease, *Archives of biochemistry and biophysics Suppl 1*, 217-222.

36. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: As domains continue to swap, *Protein Science 11*, 1285-1299.

37. Liu, Y. S., Gotte, G., Libonati, M., and Eisenberg, D. (2002) Structures of the two 3D domain-swapped RNase A trimers, *Protein Science 11*, 371-380.

38. Liu, Y. S., Gotte, G., Libonati, M., and Eisenberg, D. S. (2001) Crystal structure of an Rnase A dimer displaying a new type of 3D domain swapping, *Nucleic Acid Database*.

39. Liu, Y. S., Hart, P. J., Schlunegger, M. P., and Eisenberg, D. (1998) The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-angstrom resolution, *Proceedings of the National Academy of Sciences of the United States of America 95*, 3437-3442.

40. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: as domains continue to swap, *Protein science : a publication of the Protein Society 11*, 1285-1299.

41. Koharudin, L. M., Liu, L., and Gronenborn, A. M. (2013) Different 3D domain-swapped oligomeric cyanovirin-N structures suggest trapped folding intermediates, *Proc Natl Acad Sci U S A 110*, 7702-7707.

42. Assar, Z., Nossoni, Z., Wang, W., Santos, E. M., Kramer, K., McCornack, C., Vasileiou, C., Borhan, B., and Geiger, J. H. (2016) Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates, *Structure 24*, 1590-1598.

43. Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., Blundell, T. L., and Slingsby, C. (1990) X-ray analysis of beta-B2-crystallin and evolution of oligomeric lens proteins, *Nature 347*, 776-780.

44. Hakansson, M., Svensson, A., Fast, J., and Linse, S. (2001) An extended hydrophobic core induces EF-hand swapping, *Protein Science 10*, 927-933.

45. Rousseau, F., Schymkowitz, J., and Itzhaki, L. S. (2012) Implications of 3D domain swapping for protein folding, misfolding and function, In *Protein Dimerization and Oligomerization in Biology* (Mattews, J. M., Ed.), pp 137-152.

46. Yang, F., Bewley, C. A., Louis, J. M., Gustafson, K. R., Boyd, M. R., Gronenborn, A. M., Clore, G. M., and Wlodawer, A. (1999) Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping, *Journal of Molecular Biology 288*, 403-412.

47. Szilagyi, A., Zhang, Y., and Zavodszky, P. (2012) Intra-Chain 3D Segment Swapping Spawns the Evolution of New Multidomain Protein Architectures, *Journal of Molecular Biology 415*, 221-235.

48. Stroud, J. C., Liu, C., Teng, P. K., and Eisenberg, D. (2012) Toxic fibrillar oligomers of amyloid-beta have cross-beta structure, *Proceedings of the National Academy of Sciences of the United States of America 109*, 7717-7722.

49. Cleary, J. P., Walsh, D. M., Hofmeister, J. J., Shankar, G. M., Kuskowski, M. A., Selkoe, D. J., and Ashe, K. H. (2005) Natural oligomers of the amyloid-protein specifically disrupt cognitive function, *Nature Neuroscience 8*, 79-84.

50. Fandrich, M., Fletcher, M. A., and Dobson, C. M. (2001) Amyloid fibrils from muscle myoglobin - Even an ordinary globular protein can assume a rogue guise if conditions are right, *Nature 410*, 165-166.

51. Gronenborn, A. M. (2009) Protein acrobatics in pairs - dimerization via domain swapping, *Current Opinion in Structural Biology 19*, 39-49.

52. Schlunegger, M. P., Bennett, M. J., and Eisenberg, D. (1997) Oligomer formation by 3D domain swapping: A model for protein assembly and misassembly, In *Advances in Protein Chemistry, Vol 50: Protein Misassembly* (Wetzel, R., Ed.), pp 61-122.

53. Selkoe, D. J. (2004) Folding proteins in fatal ways (vol 426, pg 900, 2003), *Nature 428*, 445-445.

54. Shtilerman, M. D., Ding, T. T., and Lansbury, P. T. (2002) Molecular crowding accelerates fibrillization of alpha-synuclein: Could an increase in the cytoplasmic protein concentration induce Parkinson's disease?, *Biochemistry 41*, 3855-3860.

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

56. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Refined structure of dimeric diphtheria toxin at 2.0 A resolution, *Protein science : a publication of the Protein Society 3*, 1444-1463.

57. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D., and Zhang, K. Y. (2001) Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus, *Structure 9*, 1017-1027.

58. Ellis, R. J. (2001) Macromolecular crowding: an important but neglected aspect of the intracellular environment, *Curr Opin Struct Biol 11*, 114-119.

59. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Rlefined strucutre of dimeric diphtheria-toxin at 2.0-angstrom resolution, *Protein Science 3*, 1444-1463.

60. Carroll, S. F., Barbieri, J. T., and Collier, R. J. (1986) Dimeric form of diphtheria-toxin-purification and characterization, *Biochemistry 25*, 2425-2430.

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

62. Minton, A. P. (2000) Implications of macromolecular crowding for protein assembly, *Current Opinion in Structural Biology 10*, 34-39.

63. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D., and Zhang, K. Y. J. (2001) Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus, *Structure 9*, 1017-1027.

64. Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014) Energy landscapes of functional proteins are inherently risky, *Nature chemical biology 10*, 884-891.

65. Ignatova, Z., and Gierasch, L. M. (2004) Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling, *Proc Natl Acad Sci U S A 101*, 523-528.

66. Ignatova, Z., and Gierasch, L. M. (2006) Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant, *Proc Natl Acad Sci U S A 103*, 13357-13361.

67. Ropson, I. J., Boyer, J. A., Schaeffer, B. A., and Dalessio, P. M. (2009) Comparison of the folding mechanism of highly homologous proteins in the lipid-binding protein family, *Proteins: Structure, Function, and Bioinformatics 75*, 799-806.

68. Burns, L. L., and Ropson, I. J. (2001) Folding of intracellular retinol and retinoic acid binding proteins, *Proteins: Structure, Function, and Bioinformatics 43*, 292-302.

69. Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014) Energy landscapes of functional proteins are inherently risky, *Nat. Chem. Biol. 10*, 884-891.

70. Crist, R. M., Vasileiou, C., Rabago-Smith, M., Geiger, J. H., and Borhan, B. (2006) Engineering a Rhodopsin Protein Mimic, *Journal of the American Chemical Society 128*, 4522-4523.

71. Berbasova, T., Nosrati, M., Vasileiou, C., Wang, W., Lee, K. S., Yapici, I., Geiger, J. H., and Borhan, B. (2013) Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone, *J Am Chem Soc 135*, 16111-16119.

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

73. Wang, W., Nossoni, Z., Berbasova, T., Watson, C. T., Yapici, I., Lee, K. S., Vasileiou, C., Geiger, J. H., and Borhan, B. (2012) Tuning the electronic absorption of protein-embedded all-trans-retinal, *Science 338*, 1340-1343.

74. Nosrati, M., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. H. (2016) A Photoisomerizing Rhodopsin Mimic Observed at Atomic Resolution, *J Am Chem Soc 138*, 8802-8808.

75. Berbasova, T., Santos, E. M., Nosrati, M., Vasileiou, C., Geiger, J. H., and Borhan, B. (2016) Light-Activated Reversible Imine Isomerization: Towards a Photochromic Protein Switch, *Chembiochem : a European journal of chemical biology 17*, 407-414.

76. Yapici, I., Lee, K. S., Berbasova, T., Nosrati, M., Jia, X., Vasileiou, C., Wang, W., Santos, E. M., Geiger, J. H., and Borhan, B. (2015) "Turn-on" protein fluorescence: in situ formation of cyanine dyes, *J Am Chem Soc 137*, 1073-1080.

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

78. Gunasekaran, K., Hagler, A. T., and Gierasch, L. M. (2004) Sequence and structural analysis of cellular retinoic acid-binding proteins reveals a network of conserved hydrophobic interactions, *Proteins-Structure Function and Genetics 54*, 179-194.

79. Nossoni, Z., Assar, Z., Yapici, I., Nosrati, M., Wang, W., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal, *Acta Crystallogr. Sect. D. Biol. Crystallogr. 70*, 3226-3232.

80. Shameer, K., Shingate, P. N., Manjunath, S. C. P., Karthika, M., Pugalenthi, G., and Sowdhamini, R. (2011) 3DSwap: curated knowledgebase of proteins involved in 3D domain swapping, *Database-the Journal of Biological Databases and Curation*.

81. Sillitoe, I., Cuff, A. L., Dessailly, B. H., Dawson, N. L., Furnham, N., Lee, D., Lees, J. G., Lewis, T. E., Studer, R. A., Rentzsch, R., Yeats, C., Thornton, J. M., and Orengo, C. A. (2013) New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures, *Nucleic Acids Research 41*, D490-D498.

82. Englander, S. W., and Mayne, L. (2014) The nature of protein folding pathways, *Proc Natl Acad Sci U S A 111*, 15873-15880.

83. Englander, S. W., Mayne, L., and Krishna, M. M. (2007) Protein folding and misfolding: mechanism and principles, *Quarterly reviews of biophysics 40*, 287-326.

84. Krishna, M. M., and Englander, S. W. (2007) A unified mechanism for protein folding: predetermined pathways with optional errors, *Protein science : a publication of the Protein Society 16*, 449-464.

85. Mayne, L., and Englander, S. W. (2000) Two-state vs. multistate protein unfolding studied by optical melting and hydrogen exchange, *Protein science : a publication of the Protein Society 9*, 1873-1877.

86. Roder, H., Elove, G. A., and Englander, S. W. (1988) Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR, *Nature 335*, 700-704.

87. Martins, N. F., and Santoro, M. M. (1999) Partially folded intermediates during trypsinogen denaturation, *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas 32*, 673-682.

88. Medina, E., Cordova, C., Villalobos, P., Reyes, J., Komives, E. A., Ramirez-Sarmiento, C. A., and Babul, J. (2016) Three-Dimensional Domain Swapping Changes the Folding Mechanism of the Forkhead Domain of FoxP1, *Biophysical journal 110*, 2349-2360.

89. Zwanzig, R. (1997) Two-state models of protein folding kinetics, *Proceedings of the National Academy of Sciences of the United States of America 94*, 148-150.

90. Sanson, B., Wang, T., Sun, J., Wang, L., Kaczocha, M., Ojima, I., Deutsch, D., and Li, H. (2014) Crystallographic study of FABP5 as an intracellular endocannabinoid transporter, *Acta crystallographica. Section D, Biological crystallography 70*, 290-298.

91. Assar, Z., Nossoni, Z., Wang, W., Santos, Elizabeth M., Kramer, K., McCornack, C., Vasileiou, C., Borhan, B., and Geiger, James H. Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates, *Structure*.

92. Silvaroli, J. A., Arne, J. M., Chelstowska, S., Kiser, P. D., Banerjee, S., and Golczak, M. (2016) Ligand Binding Induces Conformational Changes in Human Cellular Retinol-

binding Protein 1 (CRBP1) Revealed by Atomic Resolution Crystal Structures, *The Journal of biological chemistry 291*, 8528-8540.

93. Gillilan, R. E., Ayers, S. D., and Noy, N. (2007) Structural basis for activation of fatty acid-binding protein 4, *J Mol Biol 372*, 1246-1260.

94. Sessler, R. J., and Noy, N. (2005) A ligand-activated nuclear localization signal in cellular retinoic acid binding protein-II, *Molecular cell 18*, 343-353.

95. Armstrong, E. H., Goswami, D., Griffin, P. R., Noy, N., and Ortlund, E. A. (2014) Structural Basis for Ligand Regulation of the Fatty Acid-binding Protein 5, Peroxisome Proliferator-activated Receptor β/δ (FABP5-PPARβ/δ) Signaling Pathway, *The Journal of biological chemistry 289*, 14941-14954.

96. Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, *Nature Protocols 1*, 2527-2535.

97. Ramos, C. R. R., Oyama, S., Jr., Sforca, M. L., Pertinhez, T. A., Ho, P. L., and Spisni, A. (2008) 2POA: Schistosoma mansoni Sm14 Fatty Acid-Binding Protein: improvement of protein stability by substitution of the single Cys62 residue, *Worldwide Protein Data Bank*.

98. Gill, S. C., and Vonhippel, P. H. (1989) CALCULATION OF PROTEIN EXTINCTION COEFFICIENTS FROM AMINO-ACID SEQUENCE DATA, *Analytical Biochemistry 182*, 319-326.

99. 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., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallographica Section D-Biological Crystallography 66*, 213-221.

100. Bailey, S. (1994) THE CCP4 SUITE - PROGRAMS FOR PROTEIN CRYSTALLOGRAPHY, *Acta Crystallographica Section D-Biological Crystallography 50*, 760-763.

101. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L.-W., Jain, S., Kapral, G. J., Kunstleve, R. W. G., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2011) The Phenix software for automated determination of macromolecular structures, *Methods 55*, 94-106.

102. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallographica Section D-Biological Crystallography 53*, 240-255.
103. Ong, D. E., and Page, D. L. (1987) Cellular retinol-binding protein (type two) is abundant in human small intestine, *Journal of Lipid Research 28*, 739-745.

104. Haunerland, N. H., and Spener, F. (2004) Fatty acid-binding proteins – insights from genetic manipulations, *Progress in Lipid Research 43*, 328-349.

105. Gillilan, R. E., Ayers, S. D., and Noy, N. (2007) Structural Basis for Activation of Fatty Acid-binding Protein 4, *Journal of Molecular Biology 372*, 1246-1260.

106. Kaczocha, M., Glaser, S. T., and Deutsch, D. G. (2009) Identification of intracellular carriers for the endocannabinoid anandamide, *Proceedings of the National Academy of Sciences of the United States of America 106*, 6375-6380.

107. Sulsky, R., Magnin, D. R., Huanga, Y., Simpkins, L., Taunk, P., Patel, M., Zhu, Y., Stouch, T. R., Bassolino-Klimas, D., Parker, R., Harrity, T., Stoffel, R., Taylor, D. S., Lavoie, T. B., Kish, K., Jacobson, B. L., Sheriff, S., Adam, L. P., Ewing, W. R., and Robl, J. A. (2007) Potent and selective biphenyl azole inhibitors of adipocyte fatty acid binding pirotein (aFABP), *Bioorganic & Medicinal Chemistry Letters 17*, 3511-3515.

108. Yapici, I., Lee, K. S. S., Berbasova, T., Nosrati, M., Jia, X. F., Vasileiou, C., Wang, W. J., Santos, E. M., Geiger, J. H., and Borhan, B. (2015) "Turn-On" Protein Fluorescence: In Situ Formation of Cyanine Dyes, *Journal of the American Chemical Society 137*, 1073-1080.

109. Berbasova, T., Nosrati, M., Vasileiou, C., Wang, W., Lee, K. S., Yapici, I., Geiger, J. H., and Borhan, B. (2013) Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone, *Journal of the American Chemical Society 135*, 16111-16119.

110. Abraham, M. H., Grellier, P. L., Abboud, J. L. M., Doherty, R. M., and Taft, R. W. (1988) Solvent Effects in Organic-Chemistry - Recent Developments, *Canadian Journal of Chemistry-Revue Canadienne De Chimie 66*, 2673-2686.

111. Ando, Y., Homma, Y., Hiruta, Y., Citterio, D., and Suzuki, K. (2009) Structural characteristics and optical properties of a series of solvatochromic fluorescent dyes displaying long-wavelength emission, *Dyes and Pigments 83*, 198-206.

112. Carlotti, B., Flamini, R., Kikas, I., Mazzucato, U., and Spalletti, A. (2012) Intramolecular charge transfer, solvatochromism and hyperpolarizability of compounds bearing ethenylene or ethynylene bridges, *Chemical Physics 407*, 9-19.

113. Kamlet, M. J., and Taft, R. W. (1976) Solvatochromic Comparison Method 1. Beta-Scale of Solvent Hydrogen-Bond Acceptor (Hba) Basicities, *Journal of the American Chemical Society 98*, 377-383.

114. Yutani, K., Ogasahara, K., Tsujita, T., and Sugino, Y. (1987) Dependence of Conformational Stability on Hydrophobicity of the Amino-Acid Residue in a Series of Variant Proteins Substituted at a Unique Position of Tryptophan Synthase Alpha-Subunit, *Proceedings of the National Academy of Sciences of the United States of America 84*, 4441-4444.

109. Berbasova, T., Nosrati, M., Vasileiou, C., Wang, W., Lee, K. S., Yapici, I., Geiger, J. H., and Borhan, B. (2013) Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone, *Journal of the American Chemical Society 135*, 16111-16119.

110. Abraham, M. H., Grellier, P. L., Abboud, J. L. M., Doherty, R. M., and Taft, R. W. (1988) Solvent Effects in Organic-Chemistry - Recent Developments, *Canadian Journal of Chemistry-Revue Canadienne De Chimie 66*, 2673-2686.

111. Ando, Y., Homma, Y., Hiruta, Y., Citterio, D., and Suzuki, K. (2009) Structural characteristics and optical properties of a series of solvatochromic fluorescent dyes displaying long-wavelength emission, *Dyes and Pigments 83*, 198-206.

112. Carlotti, B., Flamini, R., Kikas, I., Mazzucato, U., and Spalletti, A. (2012) Intramolecular charge transfer, solvatochromism and hyperpolarizability of compounds bearing ethenylene or ethynylene bridges, *Chemical Physics 407*, 9-19.
113. Kamlet, M. J., and Taft, R. W. (1976) Solvatochromic Comparison Method 1. Beta-Scale of Solvent Hydrogen-Bond Acceptor (Hba) Basicities, *Journal of the American Chemical Society 98*, 377-383.

114. Yutani, K., Ogasahara, K., Tsujita, T., and Sugino, Y. (1987) Dependence of Conformational Stability on Hydrophobicity of the Amino-Acid Residue in a Series of Variant Proteins Substituted at a Unique Position of Tryptophan Synthase Alpha-Subunit, *Proceedings of the National Academy of Sciences of the United States of America 84*, 4441-4444.

109. Berbasova, T., Nosrati, M., Vasileiou, C., Wang, W., Lee, K. S., Yapici, I., Geiger, J. H., and Borhan, B. (2013) Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone, *Journal of the American Chemical Society 135*, 16111-16119.

110. Abraham, M. H., Grellier, P. L., Abboud, J. L. M., Doherty, R. M., and Taft, R. W. (1988) Solvent Effects in Organic-Chemistry - Recent Developments, *Canadian Journal of Chemistry-Revue Canadienne De Chimie 66*, 2673-2686.

111. Ando, Y., Homma, Y., Hiruta, Y., Citterio, D., and Suzuki, K. (2009) Structural characteristics and optical properties of a series of solvatochromic fluorescent dyes displaying long-wavelength emission, *Dyes and Pigments 83*, 198-206.

112. Carlotti, B., Flamini, R., Kikas, I., Mazzucato, U., and Spalletti, A. (2012) Intramolecular charge transfer, solvatochromism and hyperpolarizability of compounds bearing ethenylene or ethynylene bridges, *Chemical Physics 407*, 9-19.

113. Kamlet, M. J., and Taft, R. W. (1976) Solvatochromic Comparison Method 1. Beta-Scale of Solvent Hydrogen-Bond Acceptor (Hba) Basicities, *Journal of the American Chemical Society 98*, 377-383.

114. Yutani, K., Ogasahara, K., Tsujita, T., and Sugino, Y. (1987) Dependence of Conformational Stability on Hydrophobicity of the Amino-Acid Residue in a Series of Variant Proteins Substituted at a Unique Position of Tryptophan Synthase Alpha-Subunit, *Proceedings of the National Academy of Sciences of the United States of America 84*, 4441-4444.

Chapter V: Mutational Studies on Starch Branching Enzyme I (RBEI) to understand its Surface Binding Mechanism.

V-1 introduction

Starch is the world's second most abundant polysaccharide. More than 1 billion tons of starch is consumed annually, as it provides the majority of caloric intake for humans and animals. It is also a renewable source of transportation fuels. It is therefore critical to understand the structure, function and specificity of the enzymes responsible for starch biosynthesis and modification[(118)]. Plants use exclusively branched linear amylose as a way to store excess glucose. The stored glucose is utilized when the demand for energy increases or the plant is unable to produce enough energy to maintain e functions. The polymer is stored either in special organelles called amyloplasts, found in tubers, seeds and the roots of plants, or in chlorplasts in the leaf.

In vivo, starch synthase is responsible synthesizing the linear glucan polymer using ADP-glucose building blocks. The final step in the starch biosynthetic pathway is branching enzyme catalyzes the branching of the polysaccharide by cleaving the previously formed linear α-1,4 glycosidic linkage by SN2 attack of the carboxylate nucleophile to form an enzyme/substrate covalent intermediat, followed by a second S_N2 reaction by the acceptor nucleophile to reattaching the chain via an α-1,6 bond that serves as a branching point on the newly formed amylopectin[(119)]. These polysaccharides preserved until energy stores are low, at which time they are enzymatically cleaved from the main chain and then easily metabolized by the organism. Fungi, humans, bacteria and other animals store excess glucose as glycogen (a similarly branched polysaccharide) in an analogous fashion.

Branching Enzymes (BEs) are classified as belonging to glycoside hydrolase (GH) family 13 according to the CAZy classification (http://www.cazy.org/) based on similarity in the amino acid sequence, typified by α-amylase and related enzymes, such as isoamylase, cyclodextrin glycosyltransferase (CGTase) and pullulanase. Members of the family consist of a catalytic central $(\beta/\alpha)_8$ barrel domain that contains the active site residues. [120, 121]

To date, starch branching enzymes (SBEs) have been characterized from multiple sources including maize, [122-131] rice, [132-137] wheat, [138-141] barley, [142-144] potato, [145, 146] cassava, [147] pea and soybean. [148-150] In rice endosperm, there are three isoforms of the branching enzymes, BEI, BEIIa, and BEIIb. BEI and BEIIb are typically specifically expressed in endosperm while BEIIa is ubiquitous in every tissue and it is been shown that BEIIa is required for leaf starch synthesis [133, 137] suggesting a different biochemical role. Studies on the properties of rice BE isozymes [137] supports a previous observation stating that the branching of the amylopectin amorphous lamellae is catalyzed by BEI while the branches residing on the border between the amorphous and crystalline lamellae are almost exclusively catalyzed by BEIIb. [151] Both BEI and BEIIb have different substrate preferences.

BEI plays a role in transferring less branched, longer chains with a degree of polymerization (DP) > 37,but most preferably those with degree of polymerization DP 6-15 to link multiple clusters of amylopectin [5f], while BEIIb specifically transfers short chains with mostly DP of 6 and 7. (Figure V-1) In contrast, the BEIIa-deficient mutant exhibited no significant change in the amylopectin chain. [152] The physiological role of BEIIa might be to support at least partially the function of BEI and BEIIb.[151], [5f]

The structure of truncated Oryza Sativa L (Asian rice) branching enzyme (RBEI) by

Noghuchi et al. revealed the catalytic residues and a few glucan surface binding sites (where the protein interacts and binds with amylose) were observed to play an important role on how BEI recognizes polysaccharides containing α-1,4 as well as α-1,6 linkages. [2]

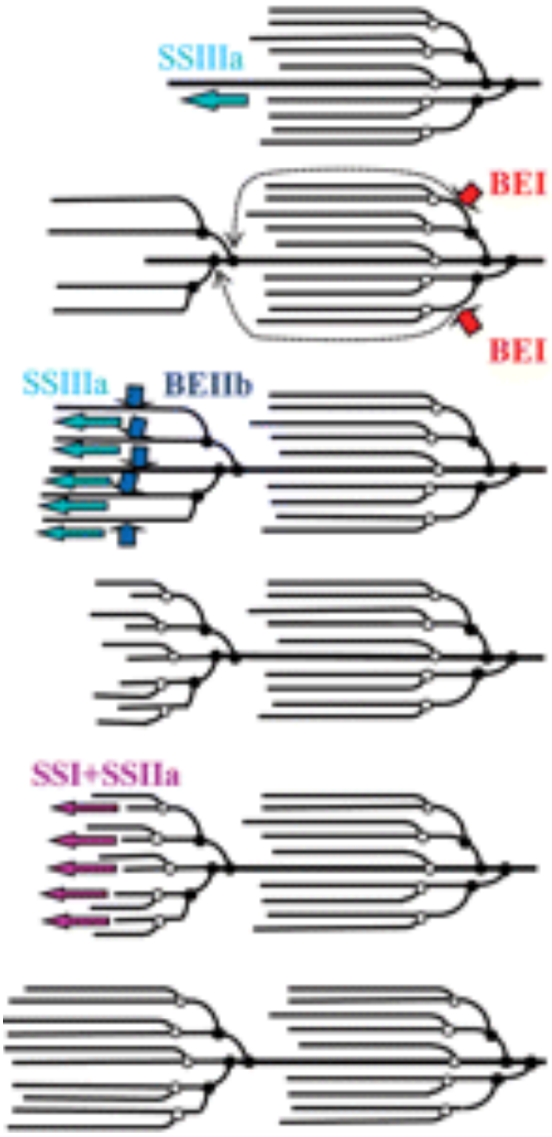


Figure V-1: Schematic view of the BE isozymes and SS isozymes in the amylopectin cluster.

V-2 Structure of the both N and C terminus truncated RBEI:

The crystal structure of the both N and C terminus truncated RBEI first has been determined by Noghuchi et al. in 2011. Also, Dr. Remie Fawaz (a former Geiger lab member)

has obtained another truncated structure of this enzyme recently with dodecaose (M12). (FigureV-2)

BEI is 820-residues. The N-terminal leader sequence (65 residues), which is responsible for transport into the amyloplast, is truncated in the structure and leads to mature BEI. A mature BEI consists of 755 residues and four domains: a central $(\beta/\alpha)_8$ catalytic domain (the GH13 module, residues 161– 587), the N-terminal carbohydrate-binding module 48 domain (CBM48; residues 59–160), the N-terminal helices (residues 1-58) and an α-amylase C- domain (residues 588–702) (FigureI-2). Asp344 and Glu399, the essential catalytic residues, are the nucleophile (base) and proton donor respectively. [2, 153]

Crystal structures of both rice mature BEI and $BEI_{\Delta C}$ (residues 702-755 are truncated) were obtained before. [118] In our group Dr. Fawaz was able to obtain well diffracting crystals of malto-dodecaose (M12) bound $BEI_{\Delta C}$ at 2.35 Å.

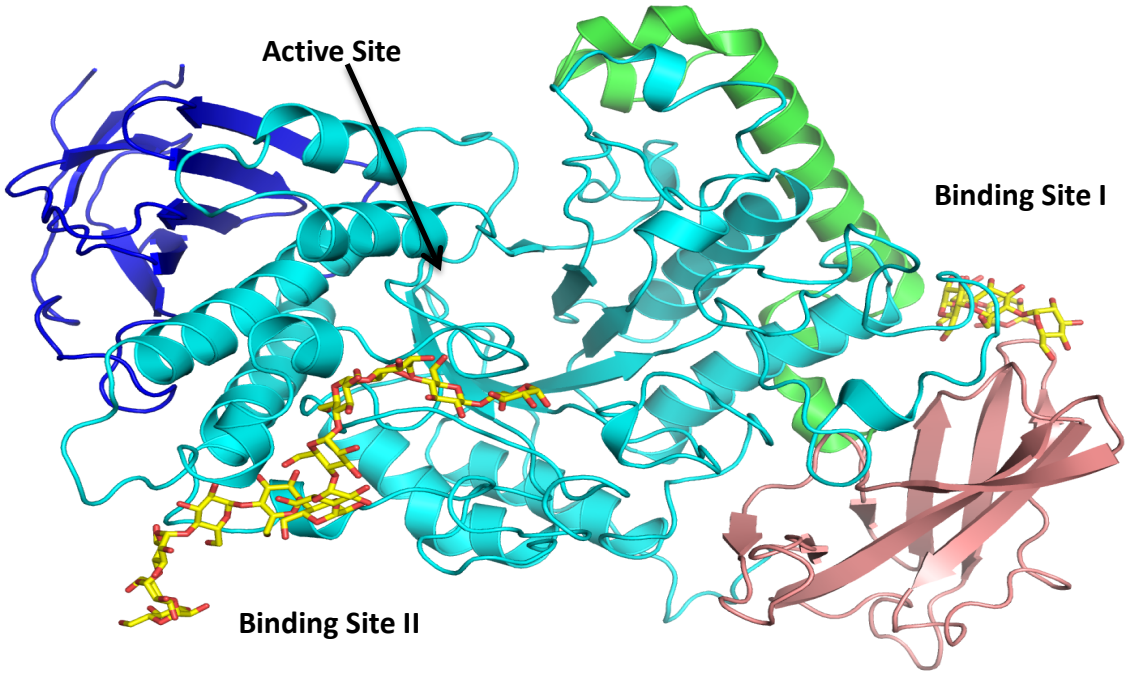


Figure V-2: Overall structure of rice Branching Enzyme I in complex with maltododecaose (M12). The N-terminal domain is shown in green, CBM48, pink, center catalytic domain, cyan, and the C-terminal α-amylase C domain in blue. Carbohydrates are represented in sticks and colored by atom type, carbons in yellow and oxygens in red. One oligosaccharide, M12, binds exclusively into the catalytic domain and hangs over the catalytic groove without reaching inside (site II), while the five glucose units visible for the second molecules (site I), between three domains: the N-terminal, the carbohydrate binding module, and the catalytic domain.

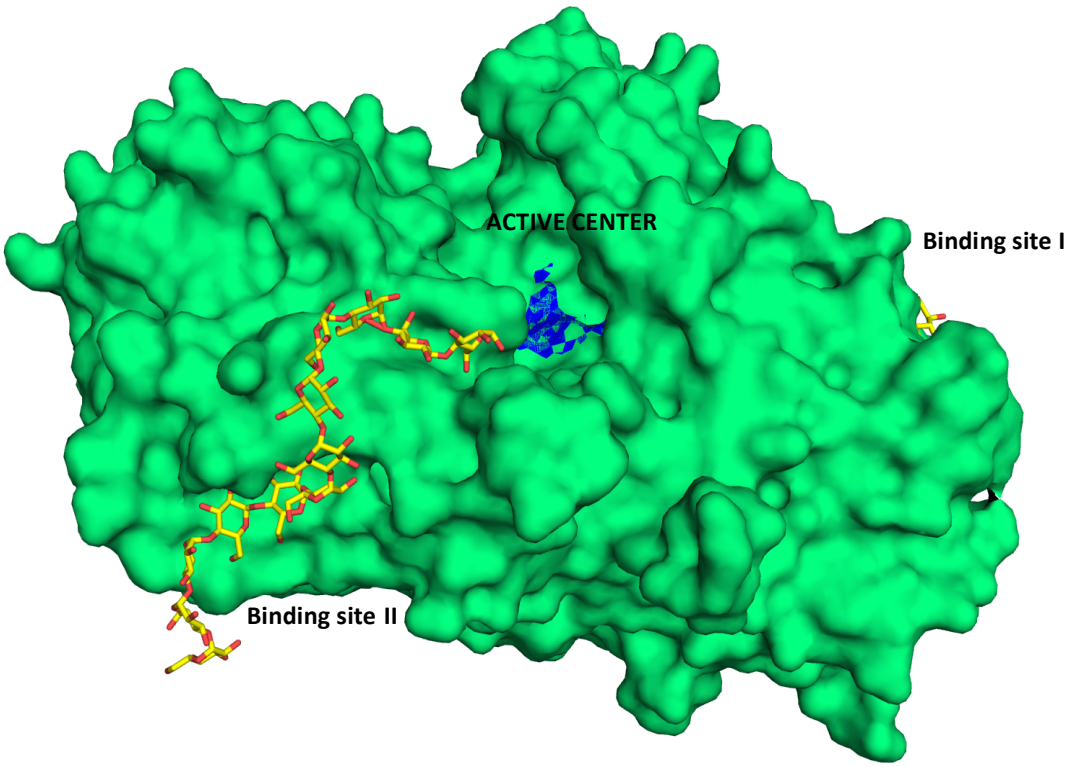


Figure V-3: Surface depiction of RBEI in complex with M12. At the center of this groove, residues involved in catalysis are shown in blue. These catalytic residues, Y235, D270, H275, R342, D344, E399, H467 and D468 according to RBEI sequence numbering, [2] were predicted based on biochemical and structural data of α-amylase. [3, 4]

V-2-1 Binding Site I: Carbohydrate Binding Module (CBM)

This binding site has been observed in both the truncated structures of RBEI, with maltopentaose (M5) and maltododecaose (M12). Based on the crystal structure, this site bridges the CBM48, N terminal helical and catalytic domains. In both structures, five units of glucose are visible in this site I. The saccharides make several hydrogen-bond interactions with the residues in that side along with the water networks. (Figure V-4).

The current rice Branching Enzyme I structure confirms once again the necessity of the CBM48 domain for carbohydrate binding. (TableV-1).

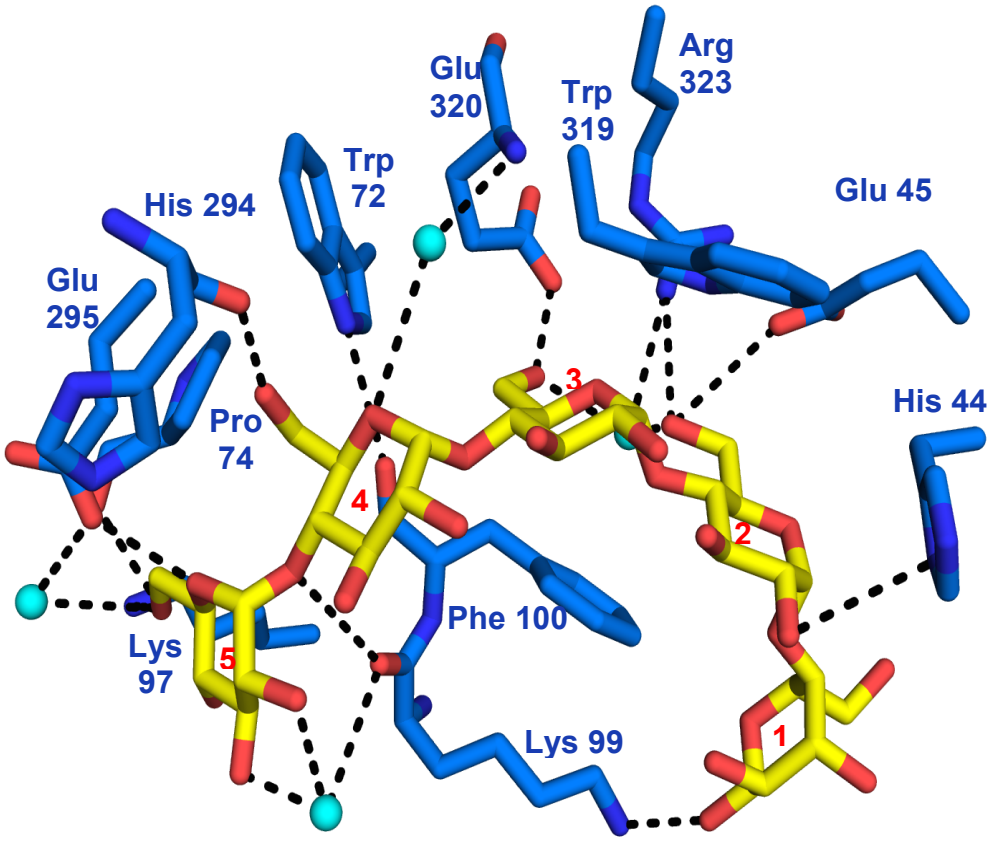


Figure V-4: Binding site I: detailed interactions between the oligosaccharide and RBEI. The protein atoms are colored by type: C in blue marine, O in red and N in dark blue. M12 atoms are also colored by type with C in yellow and O in red. Glucose units are numbered in red. Hydrogen bonds are shown in dotted black lines. Water molecules interacting in this site are represented in spheres and colored in cyan.

V-2-2 New Observed Binding Site II

The structure of M12-bound RBEI reveals a second M12 molecule bound. The second molecule of oligosaccharide and close residues around it, are divided into two binding sites. Binding site II are starting away from the active site. (Figure V-2)

Almost all glucose units in this carbohydrate make interactions (specially hydrogen binding) to residues of the enzyme. (Figure V-6 and Table V-1)

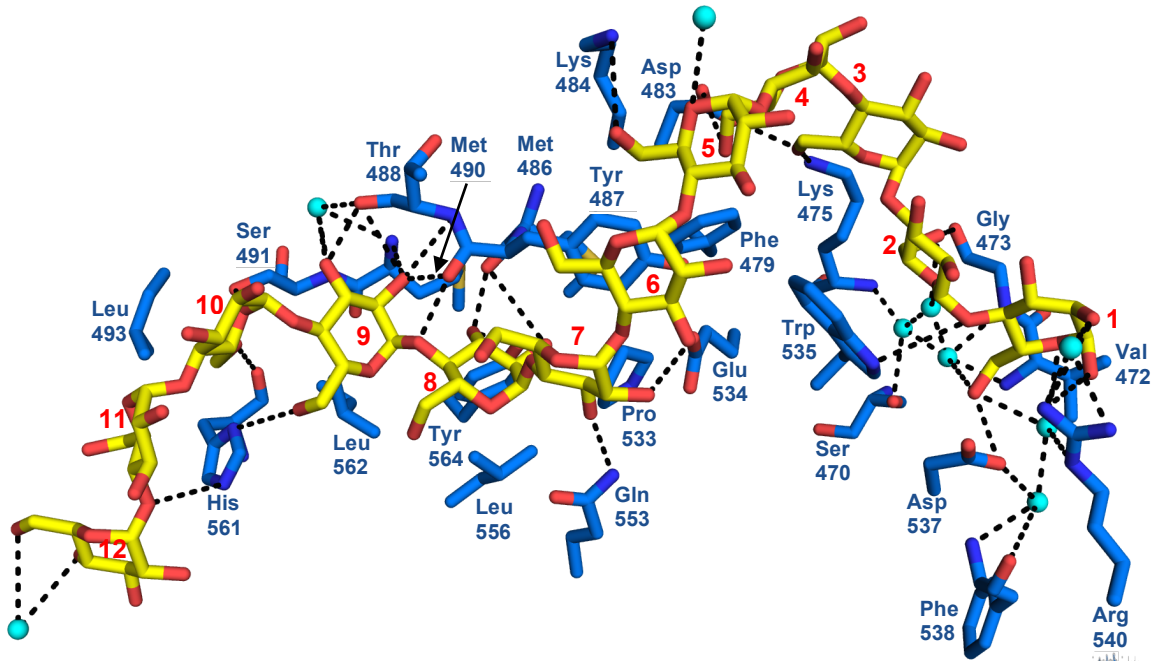


Figure V-5: Binding site II, detailed interactions between the oligosaccharide and RBEI. Numbers for glucose units, and the protein atoms are colored. Hydrogen bonds and water molecules are also represented. Hydrogen bonds are shown in dotted black lines.

V-3 Mutational studies on the observed binding sites and active site of RBEI

Several essential binding sites and conserved amino acid residues are now revealed in RBEI. Therefore in order to understand the effect of each residue on the protein activity and its involvement in determination of chain length transferred, it becomes important to perform a series of mutations starting with point mutations, and then after comparing results carrying double mutants of RBEI.

Table V-1: Residues involved in the interactions in Binding sites I, II and III of RBEI-M12.

Binding site	Residues involved in binding
I	H44, E45, W72, P74, K97, K99, F100, W319, H294, E295, E320, R323
II	H561, S491, Y487, T488, M490, Y564, M486, Q553, E534, L493, L562, L556, P533, K484, D483, K475, G473, S470, D537, F538, R540, F479, W535, V472

V-3-1 Effect of Active Site Mutation:

As previously explained in section V-2, Asp344 is the nucleophile in the catalytic process. (Figure V-3, Figure V-6) to confirm this the D344A mutant was made to study its effect on the enzyme activity and function. The goal is to confirm its' role as nucleophile, and to use the in-active enzyme in crystallization experiments, because the inactive enzyme may bind a glucan in the active site without degradation.

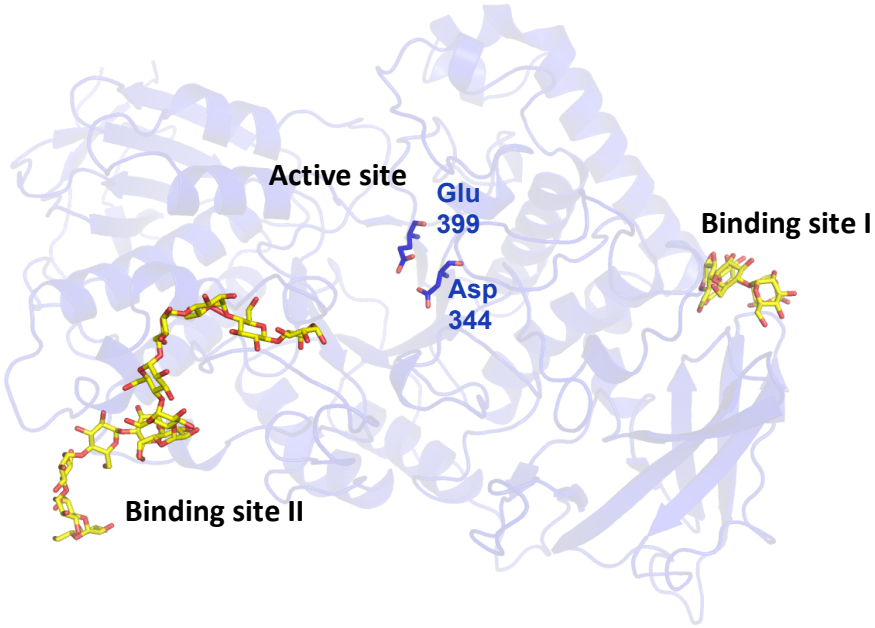


Figure V-6: Overall structure of RBEI. Active site residues are shown in blue sticks. Two oligosaccharide molecules are in yellow sticks. Oxygen atoms are in red. So far, we have not identified any glucose units close and around the active site yet.

The iodine-activity assay was carried out on purified mutated enzyme. Please refer to section V-4-1 and V-4-2 for full details on the purification step and the assay protocol. The absorbance of the glucan–iodine complex was almost constant during this time, which indicates the deactivation of enzyme due to the mutation in the active site. (Figure V-7) Specific acitivy of D344A was 0.02 µmol/min^{-1}mg^{-1} (U/mg) which was only 0.7% active compared to the activity of wild-type (WT) truncated RBEI. Data confirmed that we knocked out the enzyme activity by mutating its active site as we predicited.

Crystallization attempts with this mutant have so far not yielded well-diffracting crystals.

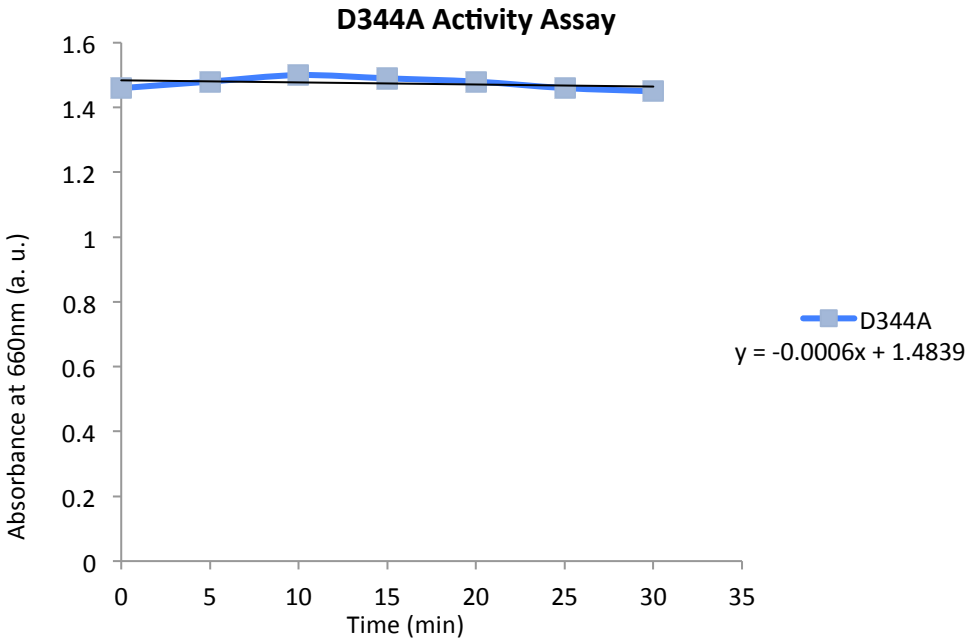


Figure V-7: Absorbance versus Time for the D344A mutation of truncated RBEI.The graph is the average for three trials. Slope/Protein indicates the activity (U/mg) of the enzyme.

V-3-1 Effect of E534A Mutation:

In binding site II of RBEI-M12, Glu 534 directly interacts with Glucose7 via a hydrogen bond to O2 and makes a water-mediated interaction with O3 (Figure V-5, Figure V-8).

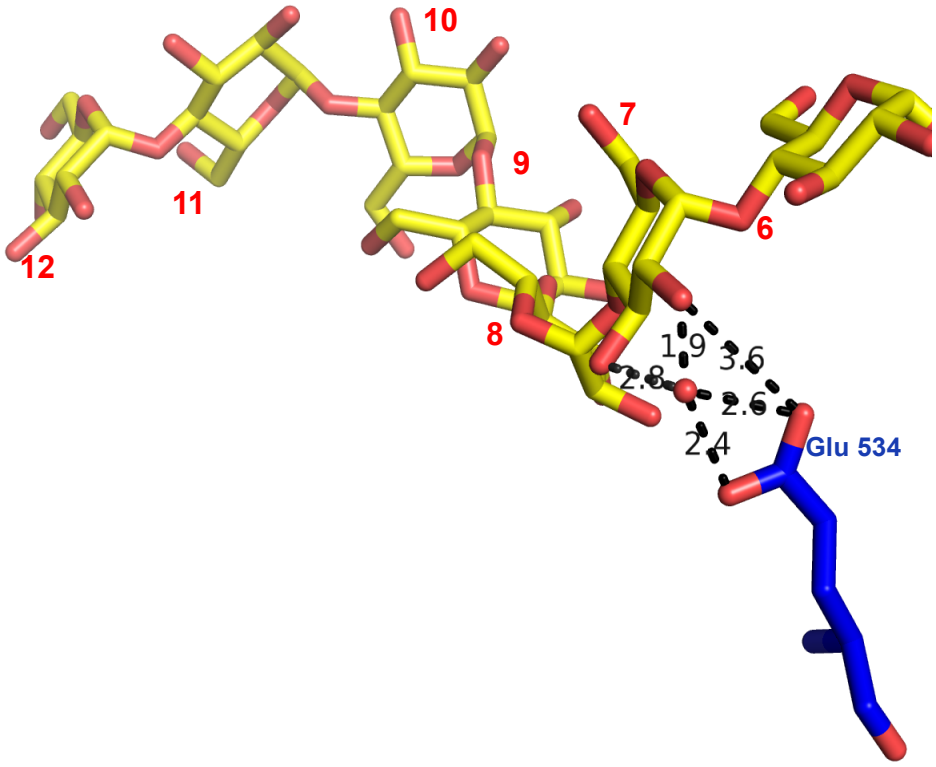


Figure V-8: Detailed interactions between the Glu7 in binding site II and residue E534 of RBEI. All interactions between them are in between 2-3.5 Å. Numbers for glucose units, and the protein atoms are colored. Hydrogen bonds and water molecules are also represented. Hydrogen bonds are shown in dotted black lines.

The E534A mutation was made and assayed to test the importance of this interaction, which is located far from the active site. However the iodine-activity assay result surprised us and E534A mutant was 83% active compare to the WT enzyme. (Figure V-9) Based on our result, Glu 534 in Site II, is more likely to involve binding enzyme to its substrate than directing oligosaccharide into the active site (because mutation of this residue, cannot deactivate the enzyme). This opens up the idea that maybe residues around that site are most likely to involve directing transferring chains (donor) and acceptor chains into the active site. To confirm our hypothesis regarding binding site II, a list of series of mutations were proposed (Table V-2). The

mutation's DNA were prepared but they need to be expressed and tested for their activity and crystallized in the future.

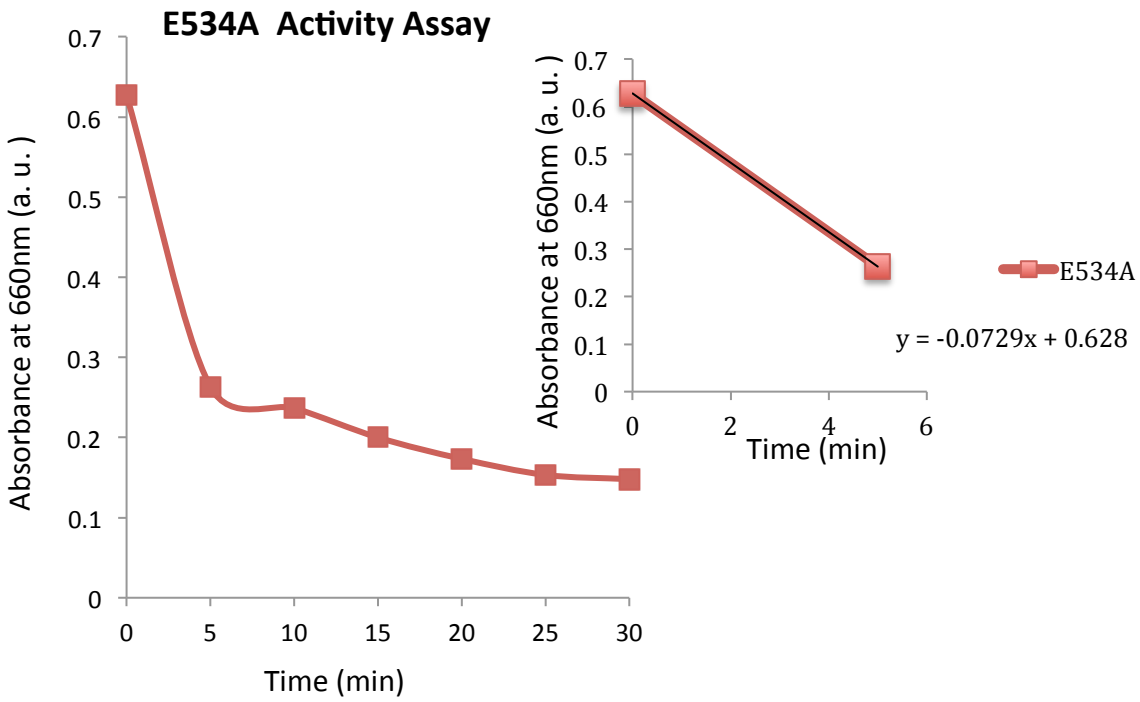


Figure V-9: Absorbance versus Time for the E534A mutation of truncated RBEI: The graph is the average for three trials. The decrease in absorbance is linear for the first 5 min then reaches a plateau due to the high activity of the enzyme.

Table V-2:
More suggested mutations to study for residues interacting with M12 in binding site II of RBEI.

Residue Number	Residue	Type of interaction to glucose units	Mutation
561	His	Hydrophilic	Ala
562	Leu	Hydrophobic	Phe
564	Tyr	Hydrophilic	Phe
487	Tyr	Hydrophilic	Ala

V-4 Experimental

V-4-1 Material and Method

Full-length RBEI (Oryza Sativa, Clone ID: 114619 from Japan) was cloned into a modified pet28a vector (Novagen) by Dr. Fawaz. It contained N-terminus His_6-tag and SUMO site right before the N-terminus of RBEI (the first 65 residues were excluded). There is also a tobacco etch virus (TEV) recognition site for cleavage in the C-terminus of the RBEI sequence (between res 694-695). To mutate the WT DNA, following primers were ordered and site-directed mutagenesis was performed to mutate and result in pure DNA. This method was performed by pfuTutrbo DNA polymerase (agilent) and temperature circle in PCR instrument. (Table V-3)

PCR Protocol for RBEI mutagenesis:

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

Tm stands for the melting temperature of the primer for each mutation and it is mentioned on each ordered primer from IDT.

PCR Primers for D344A mutant:
Forward: 5’- GGCTTCCGATTTGCTGGGGTTACGTCA -3’
Reverse: 5’- TGATGTAACCCCAGCAAATCGGAAGCC -3’

PCR Primers for E534A mutant:
Forward: 5’- TTTGGCCATCCAGCCTGGATTGACTTT-3’
Reverse: 5’- AAAGTCCAATCCAGGCTGGATGGCCAAA-3’

PCR Primers for D344A mutant:
Forward: 5’- GGCTTCCGATTTGCTGGGGTTACGTCA -3’
Reverse: 5’- TGATGTAACCCCAGCAAATCGGAAGCC -3’

PCR Primers for E534A mutant:
Forward: 5’- TTTGGCCATCCAGCCTGGATTGACTTT-3’
Reverse: 5’- AAAGTCCAATCCAGGCTGGATGGCCAAA-3’

PCR Primers for H561A mutant:
Forward: 5’- CGACACTGATGCCCTTCGATACA-3’
Reverse: 5’- TGTATCGAAGGGCATCAGTGTCG-3’

PCR Primers for D344A mutant:
Forward: 5’- GGCTTCCGATTTGCTGGGGTTACGTCA -3’
Reverse: 5’- TGATGTAACCCCAGCAAATCGGAAGCC -3’

PCR Primers for E534A mutant:
Forward: 5’- TTTGGCCATCCAGCCTGGATTGACTTT-3’
Reverse: 5’- AAAGTCCAATCCAGGCTGGATGGCCAAA-3’

PCR Primers for D344A mutant:
Forward: 5’- GGCTTCCGATTTGCTGGGGTTACGTCA -3’
Reverse: 5’- TGATGTAACCCCAGCAAATCGGAAGCC -3’

PCR Primers for E534A mutant:
Forward: 5’- TTTGGCCATCCAGCCTGGATTGACTTT-3’
Reverse: 5’- AAAGTCCAATCCAGGCTGGATGGCCAAA-3’

PCR Primers for H561A mutant:
Forward: 5’- CGACACTGATGCCCTTCGATACA-3’
Reverse: 5’- TGTATCGAAGGGCATCAGTGTCG-3’

PCR Primers for H561A mutant:
Forward: 5’- CGACACTGATGCCCTTCGATACA-3’
Reverse: 5’- TGTATCGAAGGGCATCAGTGTCG-3’

PCR Primers for L562F mutant:
Forward: 5’- TACAATTCCTTGGCCATCAAGAGAA-3’
Reverse: 5’- TTCTCTTGATGGCCAAGGAAATGTA-3’

PCR Primers for Y564F mutant:
Forward: 5’- GCATTCATATACTTGAATCGAAGGTGATCAG-3’
Reverse: 5’- CTGATCACCTTCGATTCAAGTATATGAATGC-3’

PCR Primers for H561A mutant:
Forward: 5’- CGACACTGATGCCCTTCGATACA-3’
Reverse: 5’- TGTATCGAAGGGCATCAGTGTCG-3’

PCR Primers for L562F mutant:
Forward: 5’- TACAATTCCTTGGCCATCAAGAGAA-3’
Reverse: 5’- TTCTCTTGATGGCCAAGGAAATGTA-3’

PCR Primers for Y564F mutant:
Forward: 5’- GCATTCATATACTTGAATCGAAGGTGATCAG-3’
Reverse: 5’- CTGATCACCTTCGATTCAAGTATATGAATGC-3’

Protein over-expression: The desired DNA of RBEI was then transformed into BL21(DE3) *E.coli cells.* The cells were grown in LB media, after the OD_{600} reached too 0.5-0.6 induced with 0.5mM IPTG and over expressed at 25°C for 5 hours before centrifugation and freezing at -20°C. The frozen cells (obtained from 6L) were re-suspended and supplemented with one table of protease inhibitor tablet in lysis buffer (5mL/g of pellelts) of (50mM Tris, pH=8.0, 100mM NaCl, 1mM BME, 10mM imidazole). The re-suspended cells were lysed by sonication (three 1 min with 1 min relaxation between each) and centrifuged for 20 min, 13000 rpm. The supernatant was subjected to Ni-NTA affinity resin (from Qiagen). Resin washed with the same buffer plus 30 mM imidazole until it reached to the baseline. Next, Sumo protease added on the column before eluting the protein and it was collected in different fractions. 30 µL of sample is collected for running polyacrylamide gel electrophoresis (SDS-PAGE) at each step.

The C-terminus cut of RBEI was carried by addition of the TEV protease (Concentration 2 mg/mL) and dialyzed against buffer 50mM Tris-HCL pH 8.0, 0.5mM EDTA, 1mM DTT for overnight at 4°C. Protein was concentrated using centrifugal concentrator with 10000 MW cutoffs and purified on a Superdex 75 16/600 GE column. Fractions with right size of truncated RBEI (80kDa) were analyzed by 12% SDS PAGE. Pure protein then was concentrated by centrifugal concentrator to (3-5 mg/ml), as determined by Bio-Rad protein assay at 600nm.

V-4-2 Activity Assay

We assayed the enzymatic activity for the truncated RBEI through iodine-staining assay. Branching enzymes have been tested for their activity since the 1970s [(131, 154, 155)]. The most commonly used assay, was developed by Boyer, C. and Preiss, J. in the Biochemistry department of Michigan State University (MSU). This assay is based on the decrease in

absorption of a glucan-iodine complex. [154]. The absorbance of the glucan–iodine complex is decreased by the branching of the substrate (amylose from Sigma Scientific) with branching enzyme. There are 3 trials for this assay simultaneously (to minimize variables in substrates or instrument and therefore obtained three data sets per protein) all contain 80μL of amylose stock solution (0.5ml 10%NaOH, 50mg amylose, 2ml H_2O), 6 μL HCL and 820 μL H_2O. Solutions are set to pH 8.0. Tubes stabilize at 30°C for about 10 minutes. The iodine reagent was made daily from 2.9 mL of stock solution (0.26 g of I2 and 2.6g of KI in 10mL of water). Branching enzyme I (30 μg) is added to each tube of amylose and an aliquot of 50 µL of reaction mixture is withdrawn to add to the iodine solution every 5 minutes over a 30 minute-period. The absorbance was determined at 660 nm for the reaction mixtures containing amylose. We monitor the change in absorbance (y-axis) at a wavelength of 660nm in 5 minute intervals (X-axis) over a 30-minute total reaction time. The greater decrease in absorbance, the greater the activity.

$$\text{Activity (U / mg)} = \frac{\text{Absorbance / min}}{\text{Protein Amount (mg)}} = \frac{\text{slope}}{0.03\,\text{mg}}$$

$$\text{Percent Activity} = \frac{\text{Protein Activity (U/mg)}}{\text{WT Activity (U/mg)}} \times 100$$

In the case of E534A Mutant, the plot of absorbance versus time was linear for the first 5 min, but then reached a plateau. Therefore, we used the linear part of the graph to calculate the slope. One unit of activity is defined as a decrease in absorbance of 1.0 absorbance unit per min at 660 nm [154] and is measured in U/mg of protein. [154, 156]

V-4-3 Crystallization of RBEI mutants

Each mutant protein was concentrated to 3-5mg/mL. Initial screens were done on the mutant's proteins. After optimizing different conditions for crystals to grow, the best condition was 28% PEG 8000, 550mM NaOAC, 0.1 M $(CH_3)_2AsO_2Na \cdot 3H_2O$ pH 6.9. Few small crystals formed after a month. However they still were small and did not diffract.

REFERENCES

REFERENCES

1. Harder, M. E., Deinzer, M. L., Leid, M. E., and Schimerlik, M. I. (2004) Global analysis of three-state protein unfolding data, *Protein Science 13*, 2207-2222.

2. Noguchi, J., Chaen, K., Vu, N. T., Akasaka, T., Shimada, H., Nakashima, T., Nishi, A., Satoh, H., Omori, T., Kakuta, Y., and Kimura, M. (2011) Crystal structure of the branching enzyme I (BEI) from *Oryza sativa* L with implication sfor catalysis and susbtrate binding, *Glycobiology 21*, 1108-1116.

3. Brzozowski, A. M., and Davies, G. J. (1997) Structure of the Aspergillus oryzae Alpha-Amylase Complexed with the Inhibitor Acarbose at 2.0 Å Resolution, *Biochemistry 36*, 10837-10845.

4. Uitdehaag, J. C. M., van Alebeek, G. J., Van Der Veen, B., Dijkhuizen, L., and Dijkstra, B. W. (2000) Structures of Maltohexaose and Maltoheptaose Bound at the Donor Sites of Cyclodextrin Glycosyltransferase Give Insight into the Mechanisms of Transglycosylation Activity and Cyclodextrin Size Specificity, *Biochemistry 39*, 7772-7780.

5. Ong, D. E. (1984) A NOVEL RETINOL-BINDING PROTEIN FROM RAT - PURIFICATION AND PARTIAL CHARACTERIZATION, *Journal of Biological Chemistry 259*, 1476-1482.

6. Nossoni, Z., Assar, Z., Yapici, I., Nosrati, M., Wang, W., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal, *Acta crystallographica. Section D, Biological crystallography 70*, 3226-3232.

7. Veerkamp, J. H., Vankuppevelt, T., Maatman, R., and Prinsen, C. F. M. (1993) STRUCTURAL AND FUNCTIONAL-ASPECTS OF CYTOSOLIC FATTY-ACID-BINDING PROTEINS, *Prostaglandins Leukotrienes and Essential Fatty Acids 49*, 887-906.

8. Banaszak, L., Winter, N., Xu, Z. H., Bernlohr, D. A., Cowan, S., and Jones, T. A. (1994) LIPID-BINDING PROTEINS - A FAMILY OF FATTY-ACID AND RETINOID TRANSPORT PROTEINS, *Advances in Protein Chemistry, Vol 45 45*, 89-151.

9. Vaezeslami, S., Mathes, E., Vasilelou, C., Borhan, B., and Geiger, J. H. (2006) The structure of apo-wild-type cellular retinoic acid binding protein II at 1.4 A and its relationship to ligand binding and nuclear translocation, *Journal of Molecular Biology 363*, 687-701.

10. Haunerland, N. H., and Spener, F. (2004) Fatty acid-binding proteins--insights from genetic manipulations, *Progress in lipid research 43*, 328-349.

11. Lucke, C., Huang, S., Rademacher, M., and Ruterjans, H. (2002) New insights into intracellular lipid binding proteins: The role of buried water, *Protein science : a publication of the Protein Society 11*, 2382-2392.

12. Schaap, F. G., van der Vusse, G. J., and Glatz, J. F. (2002) Evolution of the family of intracellular lipid binding proteins in vertebrates, *Molecular and cellular biochemistry 239*, 69-77.

13. Smathers, R. L., and Petersen, D. R. (2011) The human fatty acid-binding protein family: evolutionary divergences and functions, *Human genomics 5*, 170-191.

14. Franzoni, L., Lucke, C., Perez, C., Cavazzini, D., Rademacher, M., Ludwig, C., Spisni, A., Rossi, G. L., and Ruterjans, H. (2002) Structure and backbone dynamics of apo- and holo-cellular retinol-binding protein in solution, *Journal of Biological Chemistry 277*, 21983-21997.

15. Crow, J. A., and Ong, D. E. (1985) CELL-SPECIFIC IMMUNOHISTOCHEMICAL LOCALIZATION OF A CELLULAR RETINOL-BINDING PROTEIN (TYPE-2) IN THE SMALL-INTESTINE OF RAT, *Proceedings of the National Academy of Sciences of the United States of America 82*, 4707-4711.

16. Vogel, S., Mendelsohn, C. L., Mertz, J. R., Piantedosi, R., Waldburger, C., Gottesman, M. E., and Blaner, W. S. (2001) Characterization of a new member of the fatty acid-binding protein family that binds all-trans-retinol, *Journal of Biological Chemistry 276*, 1353-1360.

17. Folli, C., Calderone, V., Ottonello, S., Bolchi, A., Zanotti, G., Stoppini, M., and Berni, R. (2001) Identification, retinoid binding, and x-ray analysis of a human retinol-binding protein, *Proceedings of the National Academy of Sciences of the United States of America 98*, 3710-3715.

18. Li, E., Demmer, L. A., Sweetser, D. A., Ong, D. E., and Gordon, J. I. (1986) ISOLATION OF A CLONED CDNA-ENCODING AN ABUNDANT INTESTINAL RETINOL BINDING-PROTEIN, *Clinical Research 34*, A685-A685.

19. Li, E., and Norris, A. W. (1996) Structure/function of cytoplasmic vitamin A-binding proteins, *Annual Review of Nutrition 16*, 205-234.

20. Li, E., Qian, S. J., Winter, N. S., Davignon, A., Levin, M. S., and Gordon, J. I. (1991) FLUORINE NUCLEAR-MAGNETIC-RESONANCE ANALYSIS OF THE LIGAND-BINDING PROPERTIES OF 2 HOMOLOGOUS RAT CELLULAR RETINOL-BINDING PROTEINS EXPRESSED IN ESCHERICHIA-COLI, *Journal of Biological Chemistry 266*, 3622-3629.

21. Calderone, V., Folli, C., Marchesani, A., Berni, R., and Zanotti, G. (2002) Identification and structural analysis of a zebrafish apo and holo cellular retinol-binding protein, *Journal of Molecular Biology 321*, 527-535.

22. Tarter, M., Capaldi, S., Carrizo, M. E., Ambrosi, E., Perduca, M., and Monaco, H. L. (2008) Crystal structure of human cellular retinol-binding protein 11 to 1.2 angstrom resolution, *Proteins-Structure Function and Bioinformatics 70*, 1626-1630.

23. Winter, N. S., Bratt, J. M., and Banaszak, L. J. (1993) CRYSTAL-STRUCTURES OF HOLO AND APO-CELLULAR RETINOL-BINDING PROTEIN-II, *Journal of Molecular Biology 230*, 1247-1259.

24. Kane, M. A., Bright, F. V., and Napoli, J. L. (2011) Binding affinities of CRBPI and CRBPII for 9-cis-retinoids, *Biochimica Et Biophysica Acta-General Subjects 1810*, 514-518.

25. Malpeli, G., Stoppini, M., Zapponi, M. C., Folli, C., and Berni, R. (1995) INTERACTIONS WITH RETINOL AND RETINOIDS OF BOVINE CELLULAR RETINOL-BINDING PROTEIN, *European Journal of Biochemistry 229*, 486-493.

26. McBee, J. K., Kuksa, V., Alvarez, R., de Lera, A. R., Prezhdo, O., Haeseleer, F., Sokal, I., and Palczewski, K. (2000) Isomerization of all-trans-retinol to cis-retinols in bovine retinal pigment epithelial cells: Dependence on the specificity of retinoid-binding proteins, *Biochemistry 39*, 11370-11380.

27. Borshchevskiy, V. I., Round, E. S., Popov, A. N., Bueldt, G., and Gordeliy, V. I. (2011) X-ray-Radiation-Induced Changes in Bacteriorhodopsin Structure, *Journal of Molecular Biology 409*, 813-825.

28. Matsui, Y., Sakai, K., Murakami, M., Shiro, Y., Adachi, S., Okumura, H., and Kouyama, T. (2002) Specific damage induced by X-ray radiation and structural changes in the primary photoreaction of bacteriorhodopsin, *Journal of Molecular Biology 324*, 469-481.

29. Schreckenbach, T., Walckhoff, B., and Oesterhelt, D. (1977) STUDIES ON RETINAL-PROTEIN INTERACTION IN BACTERIORHODOPSIN, *European Journal of Biochemistry 76*, 499-511.

30. Cowan, S. W., Newcomer, M. E., and Jones, T. A. (1993) CRYSTALLOGRAPHIC STUDIES ON A FAMILY OF CELLULAR LIPOPHILIC TRANSPORT PROTEINS - REFINEMENT OF P2 MYELIN PROTEIN AND THE STRUCTURE DETERMINATION AND REFINEMENT OF CELLULAR RETINOL-BINDING PROTEIN IN COMPLEX WITH ALL-TRANS-RETINOL, *Journal of Molecular Biology 230*, 1225-1246.

31. Vasileiou, C., Vaezeslami, S., Crist, R. M., Rabago-Smith, M., Geiger, J. H., and Borhan, B. (2007) Protein design: reengineering cellular retinoic acid binding protein II into a rhodopsin protein mimic, *J Am Chem Soc 129*, 6140-6148.

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

33. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *Journal of Applied Crystallography 30*, 1022-1025.

34. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP, *Acta Crystallographica Section D-Biological Crystallography 66*, 22-25.

35. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallographica Section D-Biological Crystallography 67*, 355-367.

36. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallographica Section D-Biological Crystallography 66*, 486-501.

37. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping - A mechanism for oligomer assembly, *Protein Science 4*, 2455-2468.

38. Crestfield, A. M., Stein, W. H., and Moore, S. (1962) On the aggregation of bovine pancreatic ribonuclease, *Archives of biochemistry and biophysics Suppl 1*, 217-222.

39. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: As domains continue to swap, *Protein Science 11*, 1285-1299.

40. Liu, Y. S., Gotte, G., Libonati, M., and Eisenberg, D. (2002) Structures of the two 3D domain-swapped RNase A trimers, *Protein Science 11*, 371-380.

41. Liu, Y. S., Gotte, G., Libonati, M., and Eisenberg, D. S. (2001) Crystal structure of an Rnase A dimer displaying a new type of 3D domain swapping, *Nucleic Acid Database*.

42. Liu, Y. S., Hart, P. J., Schlunegger, M. P., and Eisenberg, D. (1998) The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-angstrom resolution, *Proceedings of the National Academy of Sciences of the United States of America 95*, 3437-3442.

43. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: as domains continue to swap, *Protein science : a publication of the Protein Society 11*, 1285-1299.

44. Koharudin, L. M., Liu, L., and Gronenborn, A. M. (2013) Different 3D domain-swapped oligomeric cyanovirin-N structures suggest trapped folding intermediates, *Proc Natl Acad Sci U S A 110*, 7702-7707.

45. Assar, Z., Nossoni, Z., Wang, W., Santos, E. M., Kramer, K., McCornack, C., Vasileiou, C., Borhan, B., and Geiger, J. H. (2016) Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates, *Structure 24*, 1590-1598.

46. Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., Blundell, T. L., and Slingsby, C. (1990) X-ray analysis of beta-B2-crystallin and evolution of oligomeric lens proteins, *Nature 347*, 776-780.

47. Hakansson, M., Svensson, A., Fast, J., and Linse, S. (2001) An extended hydrophobic core induces EF-hand swapping, *Protein Science 10*, 927-933.

48. Rousseau, F., Schymkowitz, J., and Itzhaki, L. S. (2012) Implications of 3D domain swapping for protein folding, misfolding and function, In *Protein Dimerization and Oligomerization in Biology* (Mattews, J. M., Ed.), pp 137-152.

49. Yang, F., Bewley, C. A., Louis, J. M., Gustafson, K. R., Boyd, M. R., Gronenborn, A. M., Clore, G. M., and Wlodawer, A. (1999) Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping, *Journal of Molecular Biology 288*, 403-412.

50. Szilagyi, A., Zhang, Y., and Zavodszky, P. (2012) Intra-Chain 3D Segment Swapping Spawns the Evolution of New Multidomain Protein Architectures, *Journal of Molecular Biology 415*, 221-235.

51. Stroud, J. C., Liu, C., Teng, P. K., and Eisenberg, D. (2012) Toxic fibrillar oligomers of amyloid-beta have cross-beta structure, *Proceedings of the National Academy of Sciences of the United States of America 109*, 7717-7722.

52. Cleary, J. P., Walsh, D. M., Hofmeister, J. J., Shankar, G. M., Kuskowski, M. A., Selkoe, D. J., and Ashe, K. H. (2005) Natural oligomers of the amyloid-protein specifically disrupt cognitive function, *Nature Neuroscience 8*, 79-84.

53. Fandrich, M., Fletcher, M. A., and Dobson, C. M. (2001) Amyloid fibrils from muscle myoglobin - Even an ordinary globular protein can assume a rogue guise if conditions are right, *Nature 410*, 165-166.

54. Gronenborn, A. M. (2009) Protein acrobatics in pairs - dimerization via domain swapping, *Current Opinion in Structural Biology 19*, 39-49.

55. Schlunegger, M. P., Bennett, M. J., and Eisenberg, D. (1997) Oligomer formation by 3D domain swapping: A model for protein assembly and misassembly, In *Advances in Protein Chemistry, Vol 50: Protein Misassembly* (Wetzel, R., Ed.), pp 61-122.

56. Selkoe, D. J. (2004) Folding proteins in fatal ways (vol 426, pg 900, 2003), *Nature 428*, 445-445.

57. Shtilerman, M. D., Ding, T. T., and Lansbury, P. T. (2002) Molecular crowding accelerates fibrillization of alpha-synuclein: Could an increase in the cytoplasmic protein concentration induce Parkinson's disease?, *Biochemistry 41*, 3855-3860.

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

59. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Refined structure of dimeric diphtheria toxin at 2.0 A resolution, *Protein science : a publication of the Protein Society 3*, 1444-1463.

60. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D., and Zhang, K. Y. (2001) Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus, *Structure 9*, 1017-1027.

61. Ellis, R. J. (2001) Macromolecular crowding: an important but neglected aspect of the intracellular environment, *Curr Opin Struct Biol 11*, 114-119.

62. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Rlefined strucutre of dimeric diphtheria-toxin at 2.0-angstrom resolution, *Protein Science 3*, 1444-1463.

63. Carroll, S. F., Barbieri, J. T., and Collier, R. J. (1986) Dimeric form of diphtheria-toxin-purification and characterization, *Biochemistry 25*, 2425-2430.

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

65. Minton, A. P. (2000) Implications of macromolecular crowding for protein assembly, *Current Opinion in Structural Biology 10*, 34-39.

66. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D., and Zhang, K. Y. J. (2001) Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus, *Structure 9*, 1017-1027.

67. Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014) Energy landscapes of functional proteins are inherently risky, *Nature chemical biology 10*, 884-891.

68. Ignatova, Z., and Gierasch, L. M. (2004) Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling, *Proc Natl Acad Sci U S A 101*, 523-528.

69. Ignatova, Z., and Gierasch, L. M. (2006) Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant, *Proc Natl Acad Sci U S A 103*, 13357-13361.

70. Ropson, I. J., Boyer, J. A., Schaeffer, B. A., and Dalessio, P. M. (2009) Comparison of the folding mechanism of highly homologous proteins in the lipid-binding protein family, *Proteins: Structure, Function, and Bioinformatics 75*, 799-806.

71. Burns, L. L., and Ropson, I. J. (2001) Folding of intracellular retinol and retinoic acid binding proteins, *Proteins: Structure, Function, and Bioinformatics 43*, 292-302.

72. Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014) Energy landscapes of functional proteins are inherently risky, *Nat. Chem. Biol. 10*, 884-891.

73. Crist, R. M., Vasileiou, C., Rabago-Smith, M., Geiger, J. H., and Borhan, B. (2006) Engineering a Rhodopsin Protein Mimic, *Journal of the American Chemical Society 128*, 4522-4523.

74. Berbasova, T., Nosrati, M., Vasileiou, C., Wang, W., Lee, K. S., Yapici, I., Geiger, J. H., and Borhan, B. (2013) Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone, *J Am Chem Soc 135*, 16111-16119.

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

76. Wang, W., Nossoni, Z., Berbasova, T., Watson, C. T., Yapici, I., Lee, K. S., Vasileiou, C., Geiger, J. H., and Borhan, B. (2012) Tuning the electronic absorption of protein-embedded all-trans-retinal, *Science 338*, 1340-1343.

77. Nosrati, M., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. H. (2016) A Photoisomerizing Rhodopsin Mimic Observed at Atomic Resolution, *J Am Chem Soc 138*, 8802-8808.

78. Berbasova, T., Santos, E. M., Nosrati, M., Vasileiou, C., Geiger, J. H., and Borhan, B. (2016) Light-Activated Reversible Imine Isomerization: Towards a Photochromic Protein Switch, *Chembiochem : a European journal of chemical biology 17*, 407-414.

79. Yapici, I., Lee, K. S., Berbasova, T., Nosrati, M., Jia, X., Vasileiou, C., Wang, W., Santos, E. M., Geiger, J. H., and Borhan, B. (2015) "Turn-on" protein fluorescence: in situ formation of cyanine dyes, *J Am Chem Soc 137*, 1073-1080.

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

81. Gunasekaran, K., Hagler, A. T., and Gierasch, L. M. (2004) Sequence and structural analysis of cellular retinoic acid-binding proteins reveals a network of conserved hydrophobic interactions, *Proteins-Structure Function and Genetics 54*, 179-194.

82. Nossoni, Z., Assar, Z., Yapici, I., Nosrati, M., Wang, W., Berbasova, T., Vasileiou, C., Borhan, B., and Geiger, J. (2014) Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal, *Acta Crystallogr. Sect. D. Biol. Crystallogr. 70*, 3226-3232.

83. Shameer, K., Shingate, P. N., Manjunath, S. C. P., Karthika, M., Pugalenthi, G., and Sowdhamini, R. (2011) 3DSwap: curated knowledgebase of proteins involved in 3D domain swapping, *Database-the Journal of Biological Databases and Curation*.

84. Sillitoe, I., Cuff, A. L., Dessailly, B. H., Dawson, N. L., Furnham, N., Lee, D., Lees, J. G., Lewis, T. E., Studer, R. A., Rentzsch, R., Yeats, C., Thornton, J. M., and Orengo, C. A. (2013) New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures, *Nucleic Acids Research 41*, D490-D498.

85. Englander, S. W., and Mayne, L. (2014) The nature of protein folding pathways, *Proc Natl Acad Sci U S A 111*, 15873-15880.

86. Englander, S. W., Mayne, L., and Krishna, M. M. (2007) Protein folding and misfolding: mechanism and principles, *Quarterly reviews of biophysics 40*, 287-326.

87. Krishna, M. M., and Englander, S. W. (2007) A unified mechanism for protein folding: predetermined pathways with optional errors, *Protein science : a publication of the Protein Society 16*, 449-464.

88. Mayne, L., and Englander, S. W. (2000) Two-state vs. multistate protein unfolding studied by optical melting and hydrogen exchange, *Protein science : a publication of the Protein Society 9*, 1873-1877.

89. Roder, H., Elove, G. A., and Englander, S. W. (1988) Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR, *Nature 335*, 700-704.

90. Martins, N. F., and Santoro, M. M. (1999) Partially folded intermediates during trypsinogen denaturation, *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas 32*, 673-682.

91. Medina, E., Cordova, C., Villalobos, P., Reyes, J., Komives, E. A., Ramirez-Sarmiento, C. A., and Babul, J. (2016) Three-Dimensional Domain Swapping Changes the Folding Mechanism of the Forkhead Domain of FoxP1, *Biophysical journal 110*, 2349-2360.

92. Zwanzig, R. (1997) Two-state models of protein folding kinetics, *Proceedings of the National Academy of Sciences of the United States of America 94*, 148-150.

93. Sanson, B., Wang, T., Sun, J., Wang, L., Kaczocha, M., Ojima, I., Deutsch, D., and Li, H. (2014) Crystallographic study of FABP5 as an intracellular endocannabinoid transporter, *Acta crystallographica. Section D, Biological crystallography 70*, 290-298.

94. Assar, Z., Nossoni, Z., Wang, W., Santos, Elizabeth M., Kramer, K., McCornack, C., Vasileiou, C., Borhan, B., and Geiger, James H. Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates, *Structure*.

95. Silvaroli, J. A., Arne, J. M., Chelstowska, S., Kiser, P. D., Banerjee, S., and Golczak, M. (2016) Ligand Binding Induces Conformational Changes in Human Cellular Retinol-binding Protein 1 (CRBP1) Revealed by Atomic Resolution Crystal Structures, *The Journal of biological chemistry 291*, 8528-8540.

96. Gillilan, R. E., Ayers, S. D., and Noy, N. (2007) Structural basis for activation of fatty acid-binding protein 4, *J Mol Biol 372*, 1246-1260.

97. Sessler, R. J., and Noy, N. (2005) A ligand-activated nuclear localization signal in cellular retinoic acid binding protein-II, *Molecular cell 18*, 343-353.

98. Armstrong, E. H., Goswami, D., Griffin, P. R., Noy, N., and Ortlund, E. A. (2014) Structural Basis for Ligand Regulation of the Fatty Acid-binding Protein 5, Peroxisome Proliferator-activated Receptor β/δ (FABP5-PPARβ/δ) Signaling Pathway, *The Journal of biological chemistry 289*, 14941-14954.

99. Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, *Nature Protocols 1*, 2527-2535.

100. Ramos, C. R. R., Oyama, S., Jr., Sforca, M. L., Pertinhez, T. A., Ho, P. L., and Spisni, A. (2008) 2POA: Schistosoma mansoni Sm14 Fatty Acid-Binding Protein: improvement of protein stability by substitution of the single Cys62 residue, *Worldwide Protein Data Bank*.

101. Gill, S. C., and Vonhippel, P. H. (1989) CALCULATION OF PROTEIN EXTINCTION COEFFICIENTS FROM AMINO-ACID SEQUENCE DATA, *Analytical Biochemistry 182*, 319-326.

102. 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., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular

structure solution, *Acta Crystallographica Section D-Biological Crystallography 66*, 213-221.

103. Bailey, S. (1994) THE CCP4 SUITE - PROGRAMS FOR PROTEIN CRYSTALLOGRAPHY, *Acta Crystallographica Section D-Biological Crystallography 50*, 760-763.

104. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L.-W., Jain, S., Kapral, G. J., Kunstleve, R. W. G., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2011) The Phenix software for automated determination of macromolecular structures, *Methods 55*, 94-106.

105. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallographica Section D-Biological Crystallography 53*, 240-255.

106. Ong, D. E., and Page, D. L. (1987) Cellular retinol-binding protein (type two) is abundant in human small intestine, *Journal of Lipid Research 28*, 739-745.

107. Haunerland, N. H., and Spener, F. (2004) Fatty acid-binding proteins – insights from genetic manipulations, *Progress in Lipid Research 43*, 328-349.

108. Gillilan, R. E., Ayers, S. D., and Noy, N. (2007) Structural Basis for Activation of Fatty Acid-binding Protein 4, *Journal of Molecular Biology 372*, 1246-1260.

109. Kaczocha, M., Glaser, S. T., and Deutsch, D. G. (2009) Identification of intracellular carriers for the endocannabinoid anandamide, *Proceedings of the National Academy of Sciences of the United States of America 106*, 6375-6380.

110. Sulsky, R., Magnin, D. R., Huanga, Y., Simpkins, L., Taunk, P., Patel, M., Zhu, Y., Stouch, T. R., Bassolino-Klimas, D., Parker, R., Harrity, T., Stoffel, R., Taylor, D. S., Lavoie, T. B., Kish, K., Jacobson, B. L., Sheriff, S., Adam, L. P., Ewing, W. R., and Robl, J. A. (2007) Potent and selective biphenyl azole inhibitors of adipocyte fatty acid binding pirotein (aFABP), *Bioorganic & Medicinal Chemistry Letters 17*, 3511-3515.

111. Yapici, I., Lee, K. S. S., Berbasova, T., Nosrati, M., Jia, X. F., Vasileiou, C., Wang, W. J., Santos, E. M., Geiger, J. H., and Borhan, B. (2015) "Turn-On" Protein Fluorescence: In Situ Formation of Cyanine Dyes, *Journal of the American Chemical Society 137*, 1073-1080.

112. Berbasova, T., Nosrati, M., Vasileiou, C., Wang, W., Lee, K. S., Yapici, I., Geiger, J. H., and Borhan, B. (2013) Rational design of a colorimetric pH sensor from a soluble retinoic acid chaperone, *Journal of the American Chemical Society 135*, 16111-16119.

113. Abraham, M. H., Grellier, P. L., Abboud, J. L. M., Doherty, R. M., and Taft, R. W. (1988) Solvent Effects in Organic-Chemistry - Recent Developments, *Canadian Journal of Chemistry-Revue Canadienne De Chimie 66*, 2673-2686.

114. Ando, Y., Homma, Y., Hiruta, Y., Citterio, D., and Suzuki, K. (2009) Structural characteristics and optical properties of a series of solvatochromic fluorescent dyes displaying long-wavelength emission, *Dyes and Pigments 83*, 198-206.

115. Carlotti, B., Flamini, R., Kikas, I., Mazzucato, U., and Spalletti, A. (2012) Intramolecular charge transfer, solvatochromism and hyperpolarizability of compounds bearing ethenylene or ethynylene bridges, *Chemical Physics 407*, 9-19.

116. Kamlet, M. J., and Taft, R. W. (1976) Solvatochromic Comparison Method 1. Beta-Scale of Solvent Hydrogen-Bond Acceptor (Hba) Basicities, *Journal of the American Chemical Society 98*, 377-383.

117. Yutani, K., Ogasahara, K., Tsujita, T., and Sugino, Y. (1987) Dependence of Conformational Stability on Hydrophobicity of the Amino-Acid Residue in a Series of Variant Proteins Substituted at a Unique Position of Tryptophan Synthase Alpha-Subunit, *Proceedings of the National Academy of Sciences of the United States of America 84*, 4441-4444.

118. Noguchi, J., Chaen, K., Nhuan Thi, V., Akasaka, T., Shimada, H., Nakashima, T., Nishi, A., Satoh, H., Omori, T., Kakuta, Y., and Kimura, M. (2011) Crystal structure of the branching enzyme I (BEI) from Oryza sativa L with implications for catalysis and substrate binding, *Glycobiology 21*, 1108-1116.

119. Feng, L., Fawaz, R., Hovde, S., Gilbert, L., Chiou, J., and Geiger, J. H. (2015) Crystal Structures of Escherichia coli Branching Enzyme in Complex with Linear Oligosaccharides, *Biochemistry 54*, 6207-6218.

120. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities., *Biochemical Journal 280*, 309-316.

121. Jespersen, H. M., MacGregor, E. A., Henrissat, B., Sierks, M. R., and Sevensson, B. (1993) Starch- and Glycogen-Debranching and Branching Enzymes: Prediction of Structural Features of the Catalytic (α/β)8-Barrel Domain and Evolutionary Relationship to Other Amylolytic Enzymes, *Journal of Protein Chemistry 12*, 791-805.

122. Baba, T., Kimura, K., Mizuno, K., Etoh, H., Ishida, Y., Shida, O., and Arai, Y. (1991) SEQUENCE CONSERVATION OF THE CATALYTIC REGIONS OF AMYLOLYTIC ENZYMES IN MAIZE BRANCHING ENZYME-I, *Biochem. Biophys. Res. Commun. 181*, 87-94.

123. Boyer, C. D., and Preiss, J. (1981) Evidence for Independent Genetic Control of the Multiple Forms of Maize Endosperm Branching Enzymes and Starch Synthases, *Plant Physicology 67*, 1141-1145.

124. Gao, M., Fisher, D. K., Kim, K. N., Shannon, J. C., and Guiltinan, M. J. (1996) Evolutionary conservation and expression patterns of maize starch branching enzyme I and llb genes suggests isoform specialization, *Plant Molecular Biology 30*, 1223-1232.

125. Guan, H., Li, P., Imparl-Radosevich, J., Preiss, J., and Keeling, P. (1997) Comparing the Properties of *Escherichia coli* Branching Enzyme and Maize Branching Enzyme, *Archives of biochemistry and biophysics 342*, 92-98.

126. Guan, H. P., Baba, T., and Preiss, J. (1994) Expression of Branching Enzyme I of Maize Endosperm in *Escherichia coli*, *Plant Physicology 104*, 1449-1453.

127. Guan, H. P., Kuriki, T., Sivak, M., and Preiss, J. (1995) Maize Branching Enzyme Catalyzes synthesis of Glycogen-like Polysaccharide in GlgB-Deficient Escherichia Coli, *Proceedings of the National Academy of Sciences in the United States of America 92*, 964-967.

128. Guan, H. P., and Preiss, J. (1993) Differentiation of the Properties of the Branching lsozymes from Maize (Zea mays), *Plant Physiology 102*, 1269-1273.

129. Kuriki, T., Guan, H., Sivak, M., and Preiss, J. (1996) Analysis of the Active Center of Branching Enzyme II from Maize Endosperm, *J Protein Chem 15*, 305-313.

130. Kuriki, T., Stewart, D. C., and Preiss, J. (1997) Construction of Chimeric Enzymes out of Maize Endosperm Branching Enzymes I and II:: ACTIVITY AND PROPERTIES, *Journal of Biological Chemistry 272*, 28999-29004.

131. Takeda, Y., Guan, H.-P., and Preiss, J. (1993) Branching of amylose by the branching isoenzymes of maize endosperm, *Carbohhydr. Res. 240*, 253-263.

132. Kawasaki, T., Mizumo, K., Baba, T., and Shimada, H. (1993) Molecular analysis of the gene encoding a rice starch branching enzyme, *Molecular and General Genetics 237*, 10-16.

133. Mizuno, K., Kawasaki, T., Shimada, H., Satoh, H., Kobayashi, E., Okumura, S., Arai, Y., and Baba, T. (1993) Alteration of the Structural Properties of Starch Components by the Lack of an Isoform of Starch Branching Enzyme in Rice Seeds, *The Journal of biological chemistry 268*, 19084-19091.

134. Mizuno, K., Kimura, K., Arai, Y., Kawasaki, T., Shimada, H., and Baba, T. (1992) Starch Branching Enzymes from Immature Rice Seeds, *Journal of Biochemistry 112*, 643-651.

135. Mizuno, K., Kobayashi, E., Tachibana, M., Tsutomu, K., Fujimura, T., Funane, K., Kobayashi, M., and Baba, T. (2001) Characterization of an isoform of rice starch branching enzyme, RBE4, in developing seeds, *Plant Cell Physiol. 42*, 349-357.

136. Nakamura, Y., Takeichi, T., Kawaguchi, K., and Yamanouchi, H. (1992) Purification of two forms of starch branching enzyme (Q-enzyme) from developing rice endosperm, *Plant Physiology 84*, 329 –335.

137. Nakamura, Y., Utsumi, Y., Sawada, T., Aihara, S., Utsumi, C., Yoshida, M., and Kitamura, S. (2010) Characterization of the reactions of starch branching enzymes from rice endosperm, *Plant and Cell Physiology 51*, 776-794.

138. Morell, M. K., Blennow, A., Kosar-Hashemi, B., and Samuel, M. (1997) Differential Expression and Properties of Starch Branching Enzyme lsoforms in Developing Wheat Endosperm, *Plant Physiol. 113*, 201-208.

139. Rahman, S., Li, Z., Abrahams, S., Abbott, D., Appels, R., and Morell, M. K. (1999) Characterisation of a gene encoding wheat endosperm starch branching enzyme-I, *Theor. Appl. Genet. 98*, 156-163.

140. Repellin, A., Båga, M., and Chibbar, R. N. (2001) Characterization of a cDNA encoding a type I starch branching enzyme produced in developing wheat (Triticum aestivum L.)kernels, *Journal of Plant Physiology 158*, 91-100.

141. Sadequr Rahman*, A. R., Zhongyi Li, Yasuhiko Mukai, Maki Yamamoto, Behjat Kosar-Hashemi, Sharon Abrahams, and Matthew K. Morell. (2001) Comparison of Starch-Branching Enzyme Genes Reveals Evolutionary Relationships Among Isoforms. Characterization of a Gene for Starch-Branching Enzyme IIa from the Wheat D Genome Donor Aegilops tauschii1, *Plant Physiology 125*.

142. Sun, C., Sathish, P., Ahlandsberg, S., Deiber, A., and Jansson, C. (1997) Identification of four starch-branching enzymes in barley endosperm- partial purification of forms I, Ila and IIb, *New Phytol. 137*, 215-222.

143. ChuanXin, S., Sathish, P., Deiber, A., and Jansson, C. (1995) Multiple starch branching enzymes in barley endosperm, *Photosynthesis: from light to biosphere Volume V. Proceedings of the Xth International Photosynthesis Congress, Montpellier, France*, 317-320.

144. Sun, C., Sathish, P., Ahlandsberg, S., and Jansson, C. (1998) The Two Genes Encoding Starch-Branching Enzymes IIa and IIb Are Differentially Expressed in Barley, *Plant Physiol. 118*, 37–49.

145. Rydberg, U., Andersson, L., Andersson, R., A˚man2, P., and Larsson, H. (2001) Comparison of starch branching enzyme I and II from potato, *Eur. J. Biochem. 268*, 6140-6145.

146. Visko-Nielsen, A., and Blennow, A. (1998) Isolation of starch branching enzyme I from potato using gamma-cyclodextrin affinity chromatography, *J. Chromatogr. 800*, 382-385.

147. Salehuzzaman, S. N. I. M., Jacobsen, E., and Visser, R. G. F. (1992) Cloning, partial sequencing and expression of a cDNA coding for branching enzyme in cassava, *Plant Molecular Biology 20*, 809-819.

148. Amith, A. M. (1988) Major differences in isoforms of starch-branching enzyme between developing embryos of round- and wrinkled-seeded peas (Pisum sativum L.), *Planta 175*, 270-279.

149. Tomlinson, K. L., Lloydt, J. R., and Smith, A. M. (1997) Importance of isoforms of starch-branching enzyme in determining the structure of starch in pea leaves, *The Plant Journal 11*, 31-43.

150. Zhu, Z. P., and Chen, S. Q. (1992, 20-24 July) Starch accumulation and branching enzyme gene expression during seed development of soybean and pea, *Proceedings of the Fourth International Workshop on Seeds: basic and applied aspects of seed biology, Angers, France 1*, 127-133.

151. Nakamura, Y. (2002) Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants, *Plant Cell Physiol. 43*, 718-725.

152. Chaen, K., Noguchi, J., Omori, T., Kakuta, Y., and Kimura, M. (2012) Crystal structure of the rice branching enzyme I (BEI) in complex with maltopentaose, *Biochemical and Biophysical Research Communications 424*, 508-511.

153. Vu, N. T., Shimada, H., Kakuta, Y., Nakashima, T., Ida, H., Omori, T., Nishi, A., Satoh, H., and Kimura, M. (2008) Biochemical and Crystallographic Characterization of the Starch Branching Enzyme I (BEI) from *Oryz sativa* L, *Biosci. Biotechnol. Biochem. 72*, 2858-2866.

154. Boyer, C. D., and Preiss, J. (1978) MULTIPLE FORMS OF (1,4)-a-D-GLUCAN, (1,4)-a-D-GLUCAN-6- GLYCOSYL TRANSFERASE FROM DEVELOPING Zea mays L. KERNELS, *Carbohydrate Research 61*, 321-334.

155. Hizukuri, S., Takeda, Y., Yasuda, M., and Suzuki, A. (1981) MULTI-BRANCHED NATURE OF AMYLOSE AND THE ACTION OF DE- BRANCHING ENZYMES, *Carbohydrate Research 94*, 205-213.

156. Binderup, K., and Preiss, J. (1998) Glutamate-459 Is Important for Escherichia coli Branching Enzyme Activity, *Biochemistry 37*, 9033-9037.

154. Boyer, C. D., and Preiss, J. (1978) MULTIPLE FORMS OF (1,4)-a-D-GLUCAN, (1,4)-a-D-GLUCAN-6- GLYCOSYL TRANSFERASE FROM DEVELOPING Zea mays L. KERNELS, *Carbohydrate Research 61*, 321-334.

155. Hizukuri, S., Takeda, Y., Yasuda, M., and Suzuki, A. (1981) MULTI-BRANCHED NATURE OF AMYLOSE AND THE ACTION OF DE- BRANCHING ENZYMES, *Carbohydrate Research 94*, 205-213.

156. Binderup, K., and Preiss, J. (1998) Glutamate-459 Is Important for Escherichia coli Branching Enzyme Activity, *Biochemistry 37*, 9033-9037.

156. Binderup, K., and Preiss, J. (1998) Glutamate-459 Is Important for Escherichia coli Branching Enzyme Activity, *Biochemistry 37*, 9033-9037.