DOMAIN SWAPPING IN THE ILBP FAMILY

By

Nona Ehyaei

A DISSERTATION

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

Chemistry- Master of Science

2018

ABSTRACT

DOMAIN SWAPPING IN THE ILBP FAMILY

By

Nona Ehyaei

Intracellular lipid binding proteins (iLBPs) are small soluble proteins in cytosol responsible for transport of different insoluble hydrophobic molecules. Members of this family have relatively the same structures including ten stranded beta barrel and two alpha helices located at the mouth of the internal binding cavity like a cap for binding pocket of these proteins. Studies that have been done on folding pathway of iLBP reports the early beta barrel formation in the folding pathway for this family of proteins.

Human cellular retinol binding protein II (CRBPII) is a member of the iLBP family and is responsible for transport of retinal and our group have been using this protein in an effort to create rhodopsin mimics to create Schiff base with retinal. During studies on hCRBPII for protein design, my former lab mates discovered a domain swapped dimer for this protein. Domain swapping is a process by which two or more monomers exchange an identical part of their structures to form dimer or higher order oligomers. The swapped region in this protein is three beta strands with two alpha helixes which is about half of the protein.

Surprisingly, mutational studies on hCRBPII have shown that with single mutation, we can change the ratio of monomer and domain swapped dimer which are folding products in folding pathway of this protein. We tried to find the effect of different residues on domain using our strategy called “Phase-problem”. Existence of domain swapping for hCRBPII, can have physiological relevant and may affect the folding pass way for this family of proteins.

In the meanwhile, another member of the iLBP family, Fatty acid binding protein 5 (FABP5), has reported to have a very similar domain swapped dimer. This protein is found mostly in epidermal cells, but also in other tissues, such as brain, liver, kidney, lung, and adipose tissue. Existence of domain swapping in Holo FABP5 is very interesting for us, since this protein is in the same family as hCRBPII and we are interested to investigate on the structure of Apo FABP5. In the course of these studies, we resolved two domain swapped structures of FABP5 bind to palmitic acid and studied the melting points of monomer and dimer FABP5 relative to each other.

Existence of domain swapping for hCRBPII and FABP5 may lead to allosteric regulation, also huge effect on the folding passway for this family of proteins. Previously, all studies in our lab for domain swapping was through bacterial expression, however, in order to check the physiological relevance of this phenomena we made studies on mammalian expression of hCRBPII. We successfully expressed this protein in HeLa cells and got closer to find the size of this protein in mammalian expression by using size exclusion chromatography and western blotting.

Dedicated to my family and James Geiger for his support.

ACKNOWLEDGEMENTS

First, I want to thank my advisor, Prof. James H. Geiger. He teaches me how to start new projects and how to work independently. He also helps me to find my way through the barriers to achieve my goals and supports me through all of my important decisions in these three years. He is one the nicest people I have ever seen in my life and I am really glad to have him as my advisor.

I also want to thank my committee members for their support. I am glad to have Prof. Heedeok Hong as my second reader. He helps me a lot during my graduate study here. He was the first one to teach me protein folding and piqued my interest in biophysical science. I am very thankful and honored to have the support of Prof. David Weliky. I attended two of his courses which were the best courses that I have had during my educational background. Prof. Weliky, thank you so much for teaching me how to think and participate in scientific discussions. I would like to thank Prof. Jian Hu for supporting me during my seminar and oral presentation. I also want to appreciate all of my committee and my advisor for reading this dissertation.

I am also very thankful for knowing Prof. Babak Borhan. He always helps me to improve my research and I have received many great advices from him. I also want to thank Dr. Chrysoula Vasileiou for her helpful ideas in our meetings. I enjoy our collaboration with Borhan's group especially Raheleh Esmatpour for letting me to broaden my knowledge.

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

During my graduate studies, I have learned many new techniques by my lab mates. I have learned many new techniques by Dr. Zahra Assar-nossoni. She always helps me to improve in my research and I am glad to know her and have her as one my best friends. Thank you very much Zahra. I also thankful for working with Hadi Nayebi in the same lab. He helps me a lot for having better understanding of biochemistry and he always is available for me to help. Thank you Hadi for your support and kindness. I also want to thank Alireza Ghanbarpour for being helpful and I have enjoyed having scientific discussion with him. I would like to thank Rafida Nossoni for helping me in crystallography and being so understanding during hard times. Thank you Matison Pawlowski, Joelle Eaves, Lukus Sun, Yiwei Zhao and Bryan Qiu for helping me during the experiments.

I am so thankful for finding great friends at MSU and have their support during these three years.

Finally, I want to thank my kind parents Minoo and Reza and my smart brothers Danial and Dana. I am so grateful to have their support and encouragements during all of the stages of my life. I have learned many things from them. I have learned to be strong, rational and respectful to myself and others. This is best family that I can ever asked for. Having them by my side helps me to achieve anything that I hope for in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO SYMBOLS AND ABBREVIATIONS

Å	Angstrom
PDB ID	Protein Data Bank Identifier
PBS	Phosphate-buffered saline
iLBP	intracellular Lipid Binding Proteins
hCRBPII	human Cellular Retinol Binding Protein II
FABP5	Fatty acid binding protein 5
σ	Sigma
WT	Wild type
IPTG	Isopropyl &-D-1-thiogalactopyranoside Polymerase
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Kd	Kilo Dalton
E. Coli's	Escherichia coli
min	Minute
S	Second
mM	Milimolar
μM	Micromolar
nM	Nanomolar
mol	mole
mmol	Milimole
mg	Miligram
ml	Mililitre

DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
rpm	Deoxynucleotide triphosphates
°C	Degrees of centigrade
K	Degrees of kelvin
pH	Logarithmic scale of hydrogen ion activity
NaOH	Sodium hydroxide
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

Chapter I: Study of domain swapping on Human cellular retinol binding protein II (hCRBPII).

I-1 introduction

Domain swapping is a process by which two or more monomers exchange an identical part of their structures to form dimer or higher order oligomers[1] **(figure I-1.)**. The first evidence of domain swapping was seen for RNase A and it was first discovered by Eisenberg and his group for diphteria toxin.[2 3 4] So far, domain swapping has been reported for 40 different proteins. This process can be dynamic, where the barrier between monomer and domain swapped dimer or oligomer is low and it can be static, where the barrier is very high.[5 6] The swapped region can be very small or it can be as large as large as half of a protein domain. In domain swapping, the only difference between monomer and dimer is called the hinge region.[2 7] The exchanged region in most of the cases is either the N terminus or C terminus of the protein but it can also be in the middle of the protein sequence. Domain swapping can exist in some proteins naturally, but sometimes it can be the consequence of unusual conditions like pH, temprature and prescence of denaturant. Domain swapping is responsible for evolution of a number of large proteins.[6 8] Also, it may lead to aggregation that is implicated for many diseases, such as Parkinson's disease, Alzheimer disease, diabetes and many others.[4 9 10]

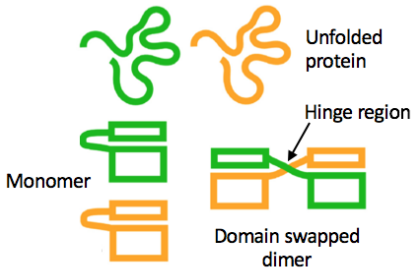

Figure I- 1. Shows the concept of domain swapping

Intracellular lipid binding proteins (iLBPs) are small cytosolic proteins responsible for transport of various insoluble hydrophobic molecules, such as retinoids and fatty acids, inside cells.[11] [12] ILBPs are a subfamily of the lipocallin family. All of the family members have similar structures including a ten stranded beta barrel and two alpha helices located at the mouth of the internal binding cavity **(figure I-2)**. Structures of almost all of the members of the iLBP family have been reported which include fatty acid binding proteins and retinoic acid binding proteins.[12] [13] Most of the studies on the folding pathway of iLBPs have been carried out in the Gierasch lab. Their results which are mostly on human Cellular Retinoic acid binding I (hCRBPI), reports the early beta barrel formation in the folding pathway of this protein. The formation of the beta barrel as a meta stable intermediate during the folding process of iLBPs prevent this family from having aggregation and amyloid formation.[14]

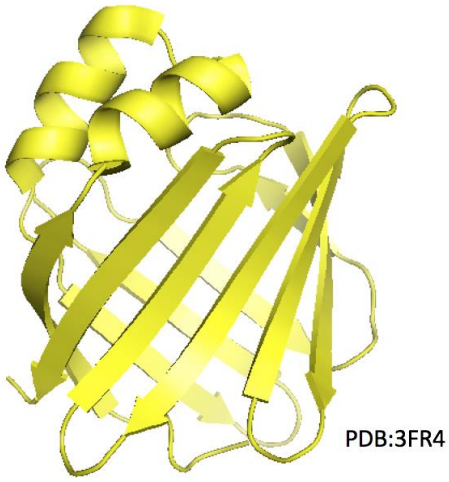

Figure I- 2. 10 antiparallel beta strands capped with two alpha helixes

Human cellular retinol binding protein II (CRBPII) is a member of the iLBP family and is responsible for transport of retinal through the cytosol of the cell.[12] Having a large binding cavity and being resilient to mutation make this protein a very interesting target for protein design applications.[5] [15] Our group has been using this protein in an effort to create rhodopsin

mimics.[16,17,18] Over 400 different mutations for hCRBPII have been studied in our group and in collaboration with Prof. Borhan's group to study protein chromophore interactions. In the process of using hCRBPII for protein design, my former lab colligues discovered a domain swapped dimer for this protein[5] **(figure I-3)**. The swapped region in this protein is three beta strands with two alpha helixes which is about half of the protein. The existence of domain swapping for hCRBPII can be a biologically relevant fold for the iLBP family, since it provides more binding sites for the protein.

Figure I- 3. Structure of domain swapped dimer for hCRBPII

Surprisingly, mutational studies on hCRBPII have shown that with single mutation, we can change the ratio of monomer and domain swapped dimer which are folding products in the folding pathway of this protein. The existence of a domain swapped dimer for hCRBPII, has led us to suggest a new folding pathway for this protein.[19 5 20] In our hypothesis based on our structural work, we predict at least one stable intermediate in the folding pathway of this protein. In our proposed mechanism, hCRBPII folds via an "open monomer" and both monomer and domain swapped dimer are derived from this intermediate. [5 21 22] **(figure I-4)**. Our lab suggested that the N- terminal and C-terminal halves of hCRBPII are capable of at least partially folding independently, and make the "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”. [5]

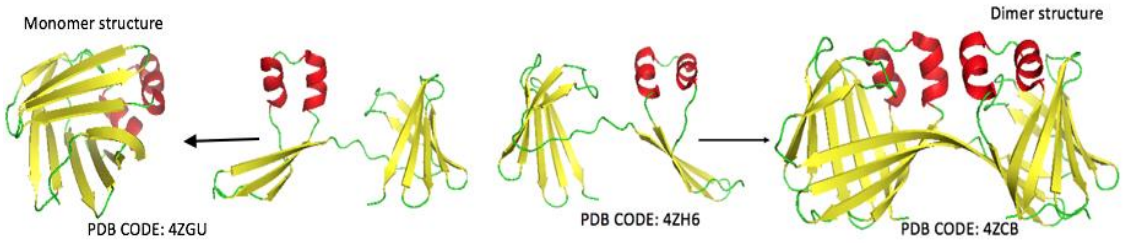

Figure I- 4. Proposed folding path way for hCRBPII. Open monomer is shown in the middle of the figure.

I-2 Phase relationship and study the occurrence of domain swapping by mutational analysis of hCRBPII.

Structure determination on hCRBPII dimer varients shows that in all cases in domain swapping of hCRBPII the loop between beta strand 3 and 4 in the monomer become straight, resulting in a single beta strand stretching the length of the dimer. However, it results in a “phase problem”. In the monomer structure, all of the odd residues in beta strand 3 and even residues of beta strand 4 are inside the binding pocket. However, in the dimer structure, since the beta strand 3 and beta strand 4 makes a single beta strand, the conformation of even residues of beta strand 4 should be toward the solvent this time **(figure 1-5)**.

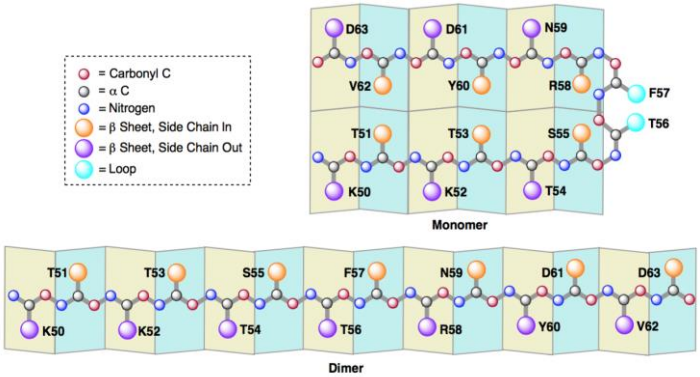

Figure I- 5. The "phase problem" in domain swap dimerization. Top figure is for monomer and bottom one is for dimer.

This means that for the formation of this single beta strand in the domain swapped dimer, the C terminal strand should re-phase and put the even side chains in phase with the odd numbered side chains of strand.

Different studies have been done on domain swapping and suggest some factors that can give rise to domain swapping. For instance, a study on hCRBPII indicates that precense of some hydrophobic residues inside the binding pocket can increase the domain swapping for this protein.[3] The way that mutant (for example, Y60L and Y60W which make dimers) solve the phase problem might be key to the domain swapping mechanism.

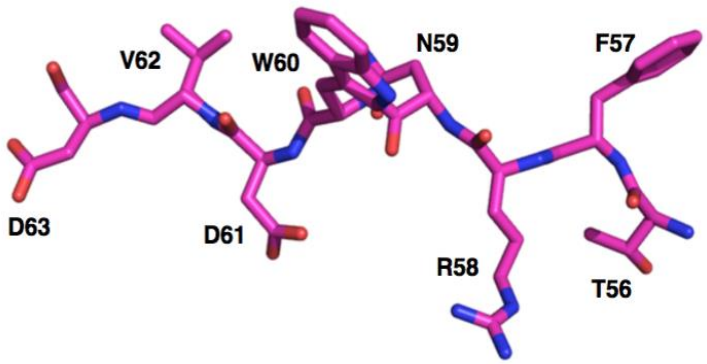

Figure I- 6. Residues 56-63 in the B subunit of Y60W-hCRBPII

Another important structural detail that we thought it might be important for domain swapping is the conformation of residue Y60 and D61. In all of the domain swapped dimers of hCRBPII mutants, these two residues are pointed toward the solvent which re-phase the strand **(figure I-6)**. However, in wild type hCRBPII which gave mostly monomer, residue Y60 is inside the binding pocket. This can be another significant factor in the mechanism for dimerization in hCRBPII.[5,12]

In order to predict the occurrence of domain swapping by amino acid sequence, Dr. Zahra Assar-nossoni and Alireza Ghanbarpour have carried out many mutations on hCRBPII (for

example: Y60L, Q108K:K40L:T51F, Q108K:T51D, E72A and Q108K:K40L:T51W) and their relative orientations of amino acids have been studied by them **(figure I-7)**.

		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

Figure I- 7. Relative orientation of angles in different mutant of hCRBPII. I: Inside; O: Outside; S: Neither Inside or Outside, but Solvent exposed; L: loop

These data showed us the large hinge motion at Thr56, subsequent proper orientation of the N and C terminus for dimer formation and reforming the connecting strand are required for domain swapping in hCRBPII.

Other important residues for domain swapping are in position 58 and 59, since the relative orientation of these two residues in the domain swapped dimer are different from the monomer. In order to test the effect of residue 59 on domain swapping, we made hCRBPII N59L mutant. In the domain swapped dimer this residue is toward the inside of the barrel and we thought by mutating this position to a more hydrophobic residue, we can increase the dimerization. However, expression of this protein at room temperature was low and it was mostly monomer. The SDS PAGE shows a small band for our protein in 15kD **(figure I-8)**.

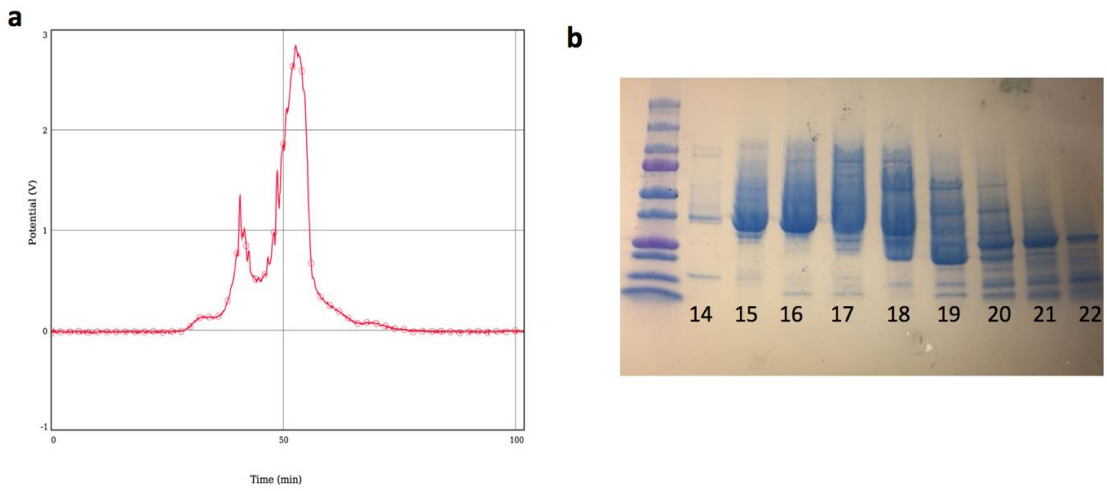

Figure I- 8. a) SEC chromatogram for hCRBPII N59L. b) SDS PAGE from fraction 14-22 of SEC.

Another important residue was Arg58. We made R58L and R58Q to see the result after mutating this residue to both hydrophobic and hydrophilic residues. However, both results show just monomer protein for this mutant. **(figure I-9)**.

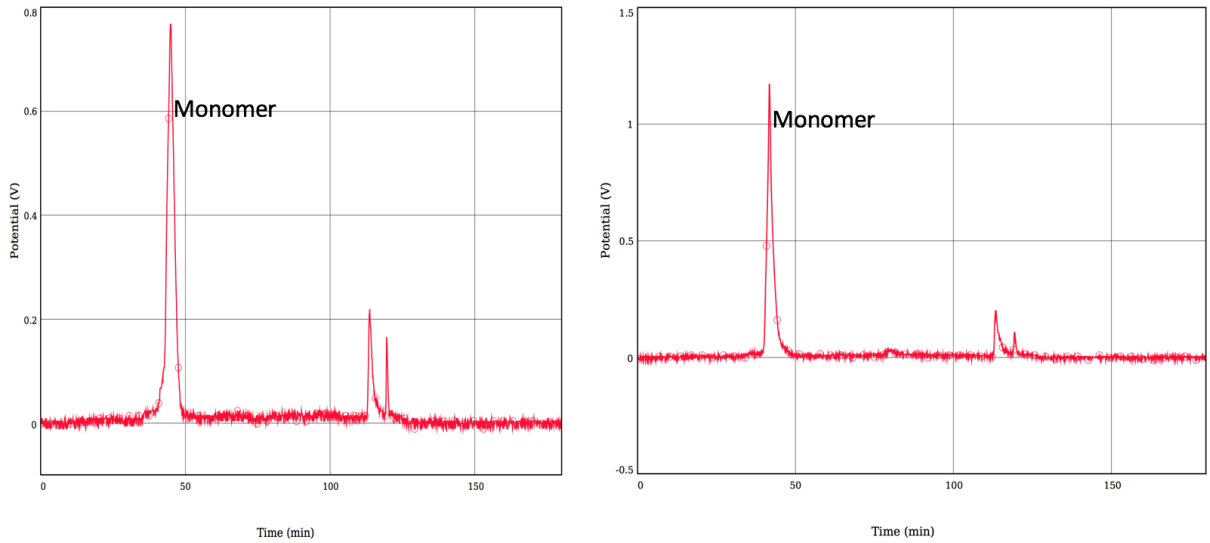

Figure I- 9. SEC chromatogram of a) R58L hCRBPII b) R58Q hCRBPII

I-3 Conduct an in vitro refolding experiment to investigate the effect of ligand binding on the folding pathway of human cellular retinol binding protein II.

Previously, the effect of protein concentration has been studied in our group. In vitro refolding experiments for some mutations of hCRBPII indicates that in higher concentrations of protein, the ratio of dimer to monomer will be increased.[5] However, the effect of ligand binding by using in vitro refolding has never been studied in this protein. For this experiment, the mutation Q108K:T51D was chosen, since the structure of both the Apo and Holo dimer of this mutation have been reported by our group and the ligand, which is retinal, can make Schiff base with Q108K. Chromatogram of size exclusion chromatography showed more dimer form for Holo form of the protein. Also, in high concentration of the protein, the amount of the monomer was still very high. **(figure I- 10)**. The concentration was 6 mg/mL for both before the start of experiment.

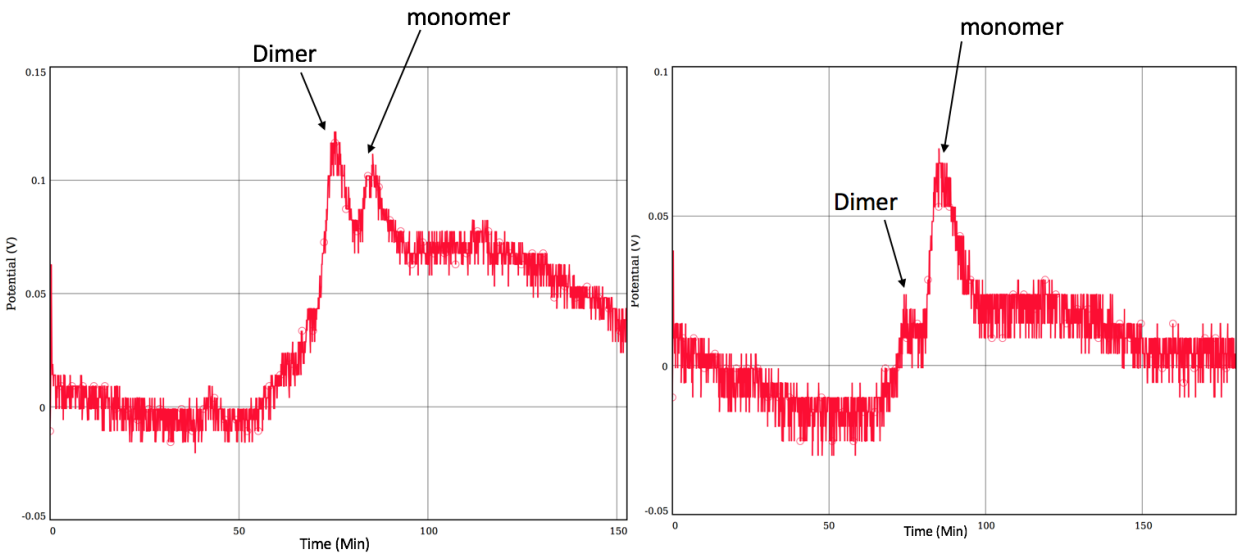

Figure I- 10. Chromatogram of size exclusion chromatography after conducting the refolding experiment in high concentration. The left figure is the chromatogram for refolding with presence of ligand and the right one is for Apo form of the protein.

These results were reproducible, but since the concentration of protein changed significantly after refolding with ligand, it needs more investigation.

1-4 Experimental

I-4-1 Site directed mutagenesis

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

Table I- 1. PCR protocol for mutagenesis

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

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

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

Primers

T59L hCRBPII

Forward: 5**'**- CACATTCCGCTTATATGATGTGGATTTC-3**'**

Reverse: 5**'**- CTAGTGGTTTTTGTCTTG -3**'**

R58L hCRBPII

Forward: 5'-TAGCACATTCTTAAACTATGATGTGGATTTC-3'

Reverse: 5**'**- GTGGTTTTTGTCTTGAAG -3**'**

R58Q hCRBPII

Forward: 5'-TAGCACATTCCAAAACTATGATGTG-3'

Reverse: 5'- GTGGTTTTTGTCTTGAAG-3'

Q108K hCRBPII

Forward: 5'-CCGCGGCTGGAAGAAGTGGATTGAGGGGG-3'

Reverse: 5'-CCCCCTCAATCCACTTCTTCCAGCCGCGG-3'

K40L hCRBPII

Forward: 5'-CTCACTCAGACGCTGGTTATTGATCAAGATGG -3'

Reverse: 5’-CCATCTTGATCAATAACCAGCGTCTGAGTGAG-3’

T51D hCRBPII

Forward: 5’-GGTGATAACTTCAAGGATAAAACCACTAGCAC-3’

Reverse: 5’-GTGCTAGTGGTTTTATCCTTGAAGTTATCACC-3’

I-4-2 Material and Method: Protein Expression and Purification

The hCRBPII gene was purchased from ATCC and cloned into pET17b vector by Dr. Wenjing Wang. The NdeI and XhoI restriction enzyme was used as cutting site in the N- and C terminus of the vector. For transformation step, 1 μL of the plasmid was transformed in 50 μL of *E. coli* DH5α competent cells from Novagen® company. The cells were incubated for 30 min in ice and heat shocked at 42°C for 45 seconds, then 400 μL of Luria- Bertani broth was added and the mixture were incubated at 37°C for two to three hours. After that, the cells were spread on an LB agar plate treated with ampicillin and incubated at 37°C for 16-20 hours. A single colony from the plate was added to 10 mL of LB media contains 100 μg/mL ampicillin. The cell culture was grown for 12-16 hours at 37°C, then the media were centrifuged at 14000 rpm for 1 min. DNA extraction and isolation from the cell pellet was done according to the manufacturer’s instructions from Promega Wizard and SV Miniprep (A1330) DNA purification kit. For protein expression, the construct was transformed into BL2 *E. coli* competent cells from Invitrogen company and spread on the plate treated with ampicillin. A single colony was picked from the plate and transferred into 1L of LB media with ampicillin (100 mg/L) and incubated at 37°C until OD_{600} reached 0.7-1. 0. 1 mM isopropyl β-D- 1 thiogalactopyranoside (IPTG) from Gold Biotechnology was induced to the cells and the cells incubate overnight at 25°C in the shaker. The transfected cells were harvested by centrifugation at 5000 rpm for 20 min. The cells were resuspended in

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

For purification, the solution of the protein passed through Q Sepharose Fast Flow resin (GE Health Sciences) and the bound protein was eluted with 10 mM Tris, 150 mM NaCl, pH 8.0. The purity of the elution was checked with SDS PAGE and the pure fractions desalted by dialysis against 10 mM Tris pH=8.0 buffer. The desalted mixture was then loaded on a 15Q anion exchange column (GE Health Sciences), using the program described in **Table I-2**.

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

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

I-4-3 In vitro refolding

For initiating this experiment this protein was expressed in *E-coli* at 16ºC for two days and purified by using fast Q column, ion exchanged chromatography from GH Health Sciences and size SEC column which was Superdex S75 16/600 HiLoad column from GE healthcare company. Chromatogram of ion exchanged and size exclusion chromatography shows mostly dimer form for this protein. After the purification step, most of the fractions that contains dimer of the protein was collected and concentrated to concentration 8 mg/ML. Then, four equivalents of the ligand were added to half of the protein and incubated for a couple of hours at 4°C to make sure that the ligand was bound to the protein. After that, 1mL of the protein solution from Apo and Holo form were inserted to 6 separate dialysis cassettes (3 of them used for Holo form of the protein and the other ones used for Apo form). Also, 2mL of 8M urea was added to each cassette; Therefore, of the concentration of the urea was 5.4 in each caset with is enough to unfold all of the protein. When the caset became ready, they were put inside the folding buffer and every hour the buffer was replaced with the new buffer. We exchanged the folding buffer for seven times to make sure that most of the urea was removed from the solution of the protein. In the last step, we added the solution of the protein to the size exclusion chromatography.

BIBLOGRAPHY

BIBLOGRAPHY

1. Mascarenhas, N. M. & Gosavi, S. Understanding protein domain-swapping using structure-based models of protein folding. *Prog. Biophys. Mol. Biol.* **128,** 113–120 (2017).

2. Bennett, M. J., Schlunegger, M. P. & Eisenberg, D. 3D domain swapping: A mechanism for oligomer assembly. *Protein Sci.* **4,** 2455–2468 (1995).

3. Liu, Y. & Eisenberg, D. 3D domain swapping: As domains continue to swap. *Protein Sci.* **11,** 1285–1299 (2002).

4. Bennett, M. J., Choe, S. & Eisenberg, D. Refined structure of dimeric diphtheria toxin at 2.0 Å resolution. *Protein Sci.* **3,** 1444–1463

5. Assar, Z. *et al.* Domain-Swapped Dimers of Intracellular Lipid-Binding Proteins: Evidence for Ordered Folding Intermediates. *Struct. Lond. Engl. 1993* **24,** 1590–1598 (2016).

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

7. Mascarenhas, N. M. & Gosavi, S. Understanding protein domain-swapping using structure-based models of protein folding. *Prog. Biophys. Mol. Biol.* **128,** 113–120 (2017).

8. O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D. & Zhang, K. Y. Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus. *Struct. Lond. Engl. 1993* **9,** 1017–1027 (2001).

9. Fändrich, M., Fletcher, M. A. & Dobson, C. M. Amyloid fibrils from muscle myoglobin. *Nature* **410,** 165–166 (2001).

10. Cleary, J. P. *et al.* Natural oligomers of the amyloid-beta protein specifically disrupt cognitive function. *Nat. Neurosci.* **8,** 79–84 (2005).

11. Haunerland, N. H. & Spener, F. Fatty acid-binding proteins--insights from genetic manipulations. *Prog. Lipid Res.* **43,** 328–349 (2004).

12. Nossoni, Z. *et al.* Structures of holo wild-type human cellular retinol-binding protein II (hCRBPII) bound to retinol and retinal. *Acta Crystallogr. D Biol. Crystallogr.* **70,** 3226–3232 (2014).

13. Marcelino, A. M. C., Smock, R. G. & Gierasch, L. M. Evolutionary coupling of structural and functional sequence information in the intracellular lipid-binding protein family. *Proteins Struct. Funct. Genet.* **63,** 373–384 (2006).

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

15. Yapici, I. *et al.* "Turn-On" Protein Fluorescence: In Situ Formation of Cyanine Dyes. *J. Am. Chem. Soc.* **137,** 1073–1080 (2015).

16. Vaezeslami, S., Mathes, E., Vasileiou, C., Borhan, B. & Geiger, J. H. The Structure of Apo-wild-type Cellular Retinoic Acid Binding Protein II at 1.4 Å and its Relationship to Ligand Binding and Nuclear Translocation. *J. Mol. Biol.* **363,** 687–701 (2006).

17. Banaszak, L. *et al.* Lipid-Binding Proteins: A Family of Fatty Acid and Retinoid Transport Proteins. 89–151 (1994). doi:10.1016/S0065-3233(08)60639-7

18. Carroll, S. F., Barbieri, J. T. & Collier, R. J. Dimeric form of diphtheria toxin: purification and characterization. *Biochemistry (Mosc.)* **25,** 2425–2430 (1986).

19. Schreckenbach, T., Walckhoff, B. & Oesterhelt, D. Studies on the Retinal-Protein Interaction in Bacteriorhodopsin. *Eur. J. Biochem.* **76,** 499–511

20. Zwanzig, R. Two-state models of protein folding kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 148–150 (1997).

21. Rousseau, F., Schymkowitz, J. & Itzhaki, L. S. Implications of 3D Domain Swapping for Protein Folding, Misfolding and Function. in *Protein Dimerization and Oligomerization in Biology* 137–152 (Springer, New York, NY, 2012).

22. Masters, J. R. HeLa cells 50 years on: the good, the bad and the ugly. *Nat. Rev. Cancer* **2,** 315 (2002).

Chapter II: Domain swapping of Fatty acid binding protein 5 (FABP5)

II-1 Introduction

The finding of domain swapping for hCRBPII has led us to ask if other members of the iLBP family can have undergo domain swapping. Meanwhile, another member of the iLBP family, Fatty acid binding protein 5 (FABP5), was reported to produce a very similar domain swapped dimerv (figure II-1).[1] Interestingly, the sequence identity between hCRBPII and FABP5 is low. Existence of a domain swapped dimer in different members of the iLBP family is likely to have physiological importance; Therefore, investigating the mechanism of domain swap dimerization in FABP5 is also critical for us.[2,3,4]

Fatty acid binding proteins are a subfamily of iLBPs that are responsible for transport of fatty acids through the cell.[5,6,7] FABPs are found in different parts of the body and the nomenclature of them is based on the tissues in which they have been discovered. Among the members of the FABP family, 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.[5,8,9] Recently, domain swapping has been reported for FABP5.[1]

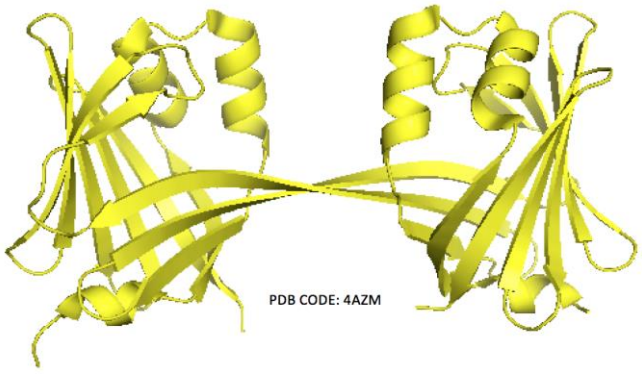

Figure II- 1. Structure of domain swapped dimer for FABP5 bound to 2-arachidonoylglycerol.

The existence of domain swapping in FABP5 is very interesting for us, since this protein is in the same family as hCRBPII.

Before the first domain swapped structure for human and mouse FABP5 which is reported by Sanson et al, all the other structures for this protein had been reported as monomer. Since the presence or absence of ligand does not affect the amount of dimer and monomer after expression, Sanson and his group suggested that mouse and human FABP5 protein structures may not be stable and demonstrate conformational dynamics. However, from our data, we know that in hCRBPII the dimer and monomer do not convert to each other under normal conditions and the dimer appears to be a kinetically trapped product that does not readily convert to monomer at room temprature.[10]

Sanson and his group got the structure FABP5 binds to endocannabinoids anandamide (AEA) and 2-arachidonoylglycerol (2-AG) **(figure II-2)**.

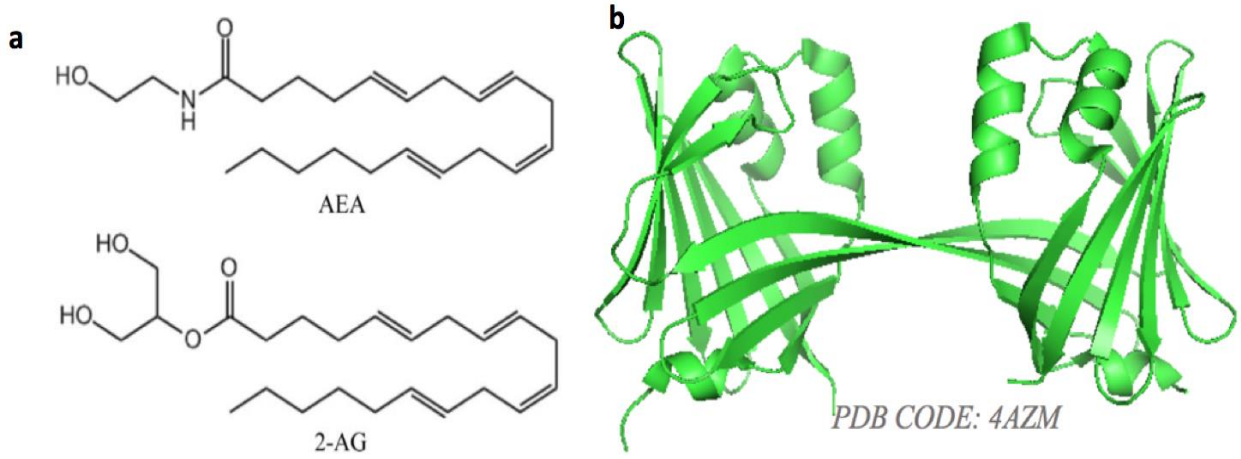

Figure II- 2. a) structure of AEA and 2-AG b) structure of human FABP5 bound to AEA.

Overlay of these two structures shows huge conformational change **(figure II-3)**.

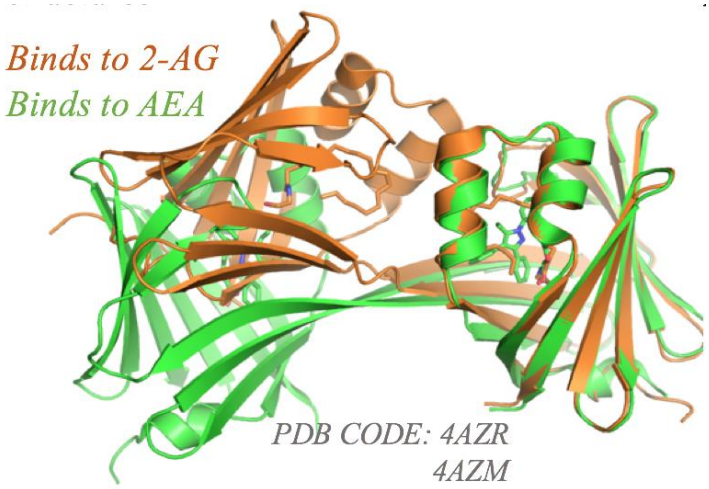

Figure II- 3. Overlay of FABP5 bound to AEA (in green) and binds to 2-AG (in orange)

As we found out domain swapping for hCRBPII, we are interested in investigating the structure of Apo FABP5. By having a structure of Apo FABP5 we can compare the conformation of the residues in this structure with the Holo forms. In this case, if any conformational change is detected, it may have physiological relevant and it can lead to the allosteric regulation for this family of proteins.[2,3,4]

II-2 Monomer structure of FABP5

In order to investigate domain swapping in hFABP5, Dr. Zahra Assar Nossoni and I studeid domain swapping in the Apo protein. Since the size of the hCRBPII is similar to that of FABP5 (15 KD), we used chromatogram of the hCRBPII to see in which fractions dimer and monomer comes out of the column. Surprisingly, the chromatogram of the size exclusion chromatography for FABP5 showed different results in each time that we expressed the protein. Most of the times

after expression, we got a very large monomer peak and very small peaks for the dimer and higher oligomers. However, one time we were able to get a significant amount of dimer after purification (The reason for getting different results for expression is not clear yet.) After that, we used SDS PAGE to make sure that we had our purified protein on those fractions **(figure II-4)**. We also ran a native gel from fractions 18 19 20 21 and we tried to compare the result with the monomer and dimer of hCRBPII; however, we were not able to say in which fractions we have just dimer. Therefore, we mixed fractions 18, 19, 20 that we thought we had dimer and put the crystallization box for them. We were able to obtain crystals that diffracted and determine the structure in $P2_12_12_1$ space group **(Table II-1)**. However, it was not the domain swapped dimer as we expected, instead it was two monomers per asymmetric unit. The structure of the Apo monomer Fabp5 that we got has a different conformation compared to monomer structures that have been reported before and it has been discussed in part II-4.[11]

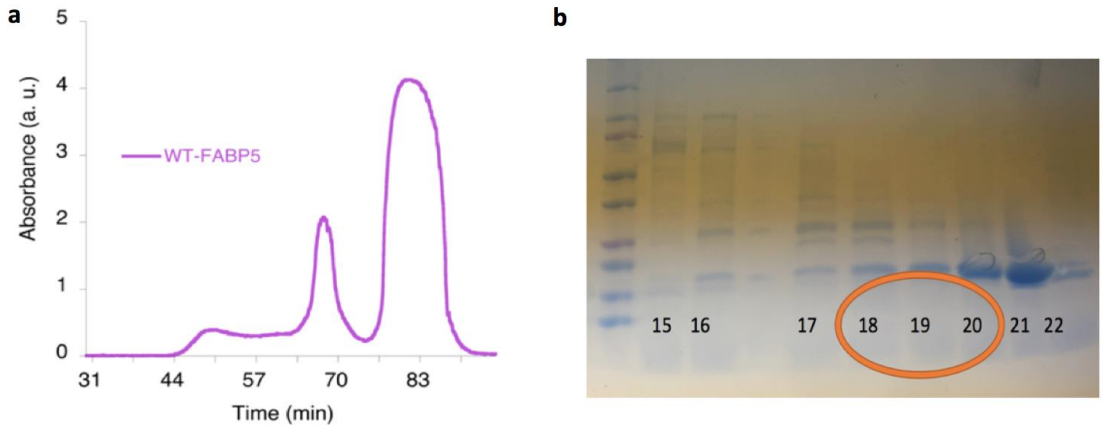

Figure II- 4. a) chromatogram of size exclusion chromatography. b) SDS PAGE of fractions 15 to 22. Fractions 18, 19, 20 were mixed for putting the crystallization box.

II-3 Domain Swapped dimer structures for FABP5 bound to palmitic acid

we expressed hFABP5 multiple times and one time we obtained dimer peak again. This time we had enough protein in each fraction; therefore, we concentrated fraction 18 separately and put separate crystallization box for that. Surprisingly, this time we observed an electron density for hinge loop and the structure was domain swapped dimer in $P6_322$ space group **(Table II-2)**. However, we saw an extra electron density in the binding pocket of the protein when we were trying to solve the structure. The electron density seems to be for a fatty acid and since we did not add any ligand, we predicted that it should be one of the fatty acids in E-Coli BL21 cells. Therefore, we overlaid various fatty acids in this electron density, such as endocannabinoids anandamide (AEA) that can bind to hFABP5 to get information about the conformation and size of this fatty acid **(figure 11-5)**.

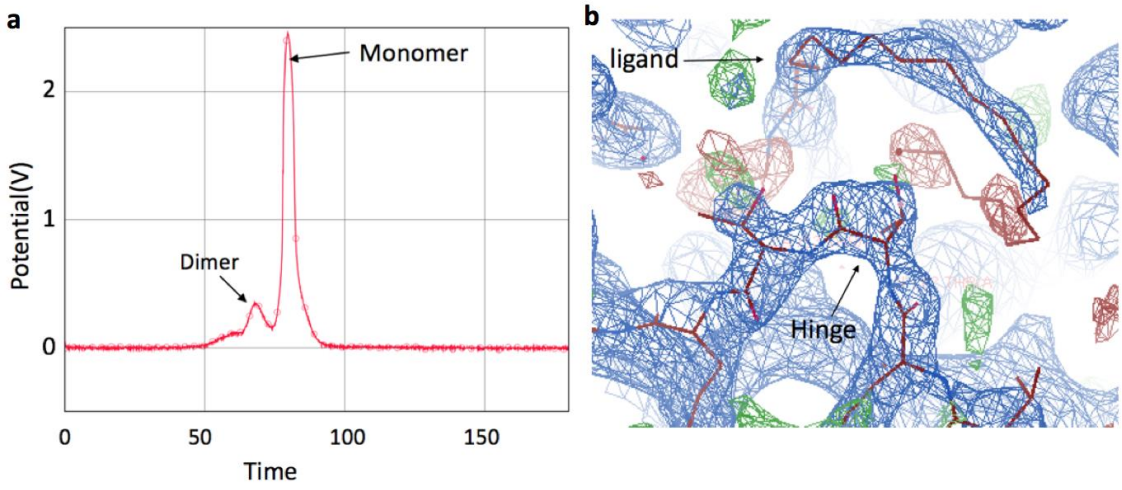

Figure II- 5. a) SEC for FABP5 b) crystal structure of FABP5 domain swapped dimer with AEA as a ligand in extra electron density.

After overlying different ligands, Palmitic acid was matched to our electron density. Palmitic acid has 16 carbons in its formula and it is one of the most common fatty acids produced in bacteria **(figure II-6)**.[12]

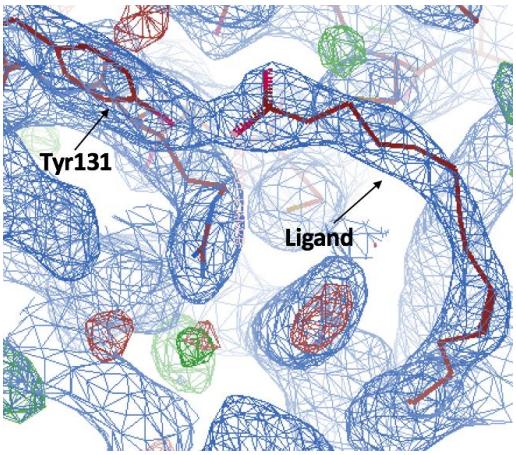

Figure II- 6. structure of FABP5 with palmitic acid in the extra electron density.

We expressed FABP5 protein multiple times and obtained again. This time we treated the dimer fraction from SEC column with lipidex column. Lipidex resin contains many lipids and based on the hydrophobic interactions any lipids should binds to the resin. Delipidation by lipidex resin is for removing all the non-specific lipids from the protein. Therefore, we used it to remove the palmitic acid from the binding pocket of our protein. However, our new domain swapped structure for FABP5 showed the same extra electron density in the binding pocket as before and it is probably palmitic acid again **(figure II-7)**. The aforementioned crystal is in the$C222_1$ space group **(Table II-3)**.

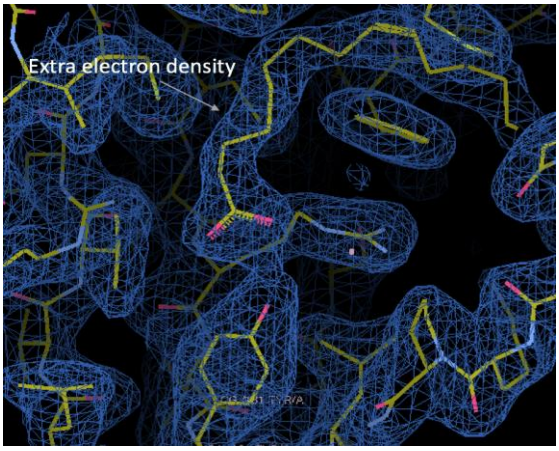

Figure II- 7. Second crystal structure of FABP5 with palmitic acid in the binding pocket.

Surprisingly, the alignment between these two domain swapped structures shows different conformation in hinge region **(figure II-8)**.

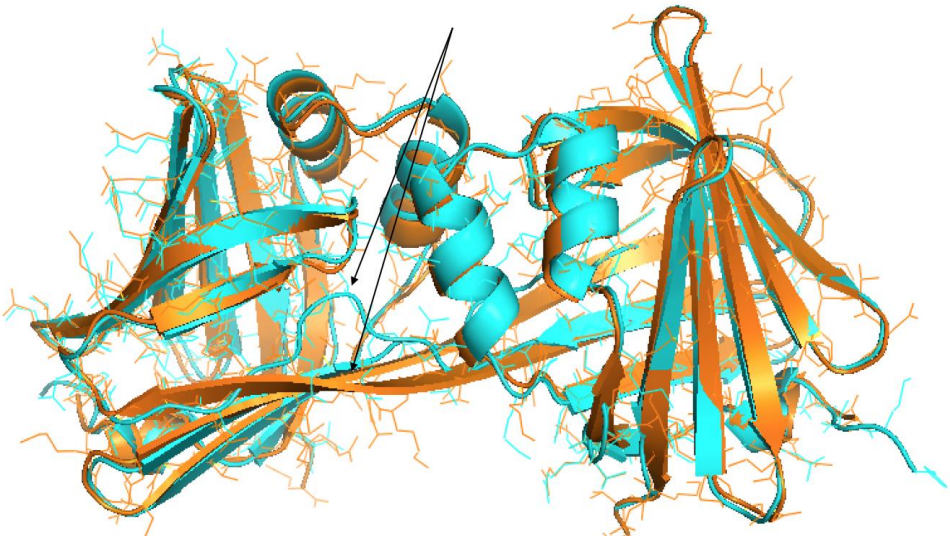

Figure II- 8. Overlay of two FABP5 domain swapped structures bound to palmitic acid (first structure is in cyan and the most recent one is in orange). The arrows point to the hinge loops.

II-4 Conformational change between FABP5 domain swapped structures.

As mentioned before, there is a huge conformational change between FABP5 domain swapped structures bound with AEA and nound with 2-AG. We overlaid also our FABP5 domain

swapped structure with palmitic acid with these two structures and we saw conformational change between these structures. The relative orientation of two domains are more similar in AEA and palmitic acid case, since both of them have more similar structures compare to 2-AG **(figure II-9)**. These results suggest the domain swapped dimer structure to be sensitive to ligand binding, undergoing conformational change. The change in relative orientation of two domains and hinge loop may lead to allosteric regulation for this protein.

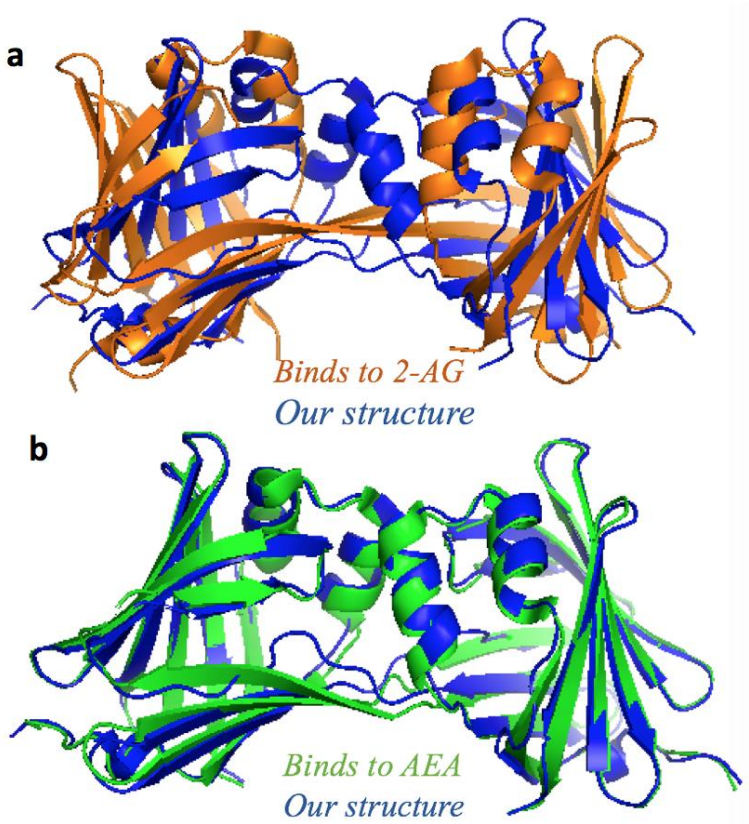

Figure II- 9. Overlay between a) structures of FABP5 binds to 2-AG (orange) and FABP5 binds to palmitic acid (blue) b) structures of FABP5 binds to AEA (green) and binds to palmitic acid (blue).

II-5 Determination of melting point of FABP5 by using Thermal Shift Assay (TSA)

TSA is a proper method for determination of melting temperature (T_m) of proteins. In this method, a fluorophore binds nonspecifically to hydrophobic surfaces and water strongly quenches

the assorted fluorescence. By unfolding the protein, the hydrophobic surface binds to the dye and water molecules are excluded from the fluorophore; Therefore, the fluorescence will increase. In this case, we can obtain thermal melting curve which shows the change of florescence by increasing the temperature for the protein, and T_m (midpoint of the stability curve) can be obtained from this curve.[13,14] In order to find the melting temperature of both FABP5 monomer and dimer we expressed and purified the protein. After that we ran the SEC column for protein and fractions for both monomer (mostly fractions 21 22) and dimer (mostly fraction 18, 19) were collected and were sent to Dr. Zahra Assar-Nossoni for TSA analysis in Cayman chemical company. The melting point for fraction 21 and 22 were 63.3°C and 61.2°C respectively **(figure II-10)**. Surprisingly, for dimer, two melting points were reported for each fraction. For fraction 18 they were 42.5°C and 63.5°C and for fraction 19 they were 43°C and 65°C **(figure II-11)**.

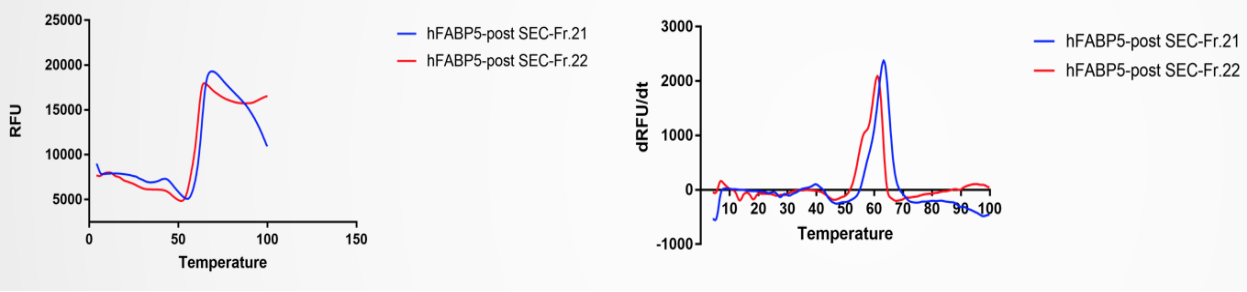

Figure II- 10. Thermal stability assay for monomer FABP5 (fraction 21 in blue and 22 in red).

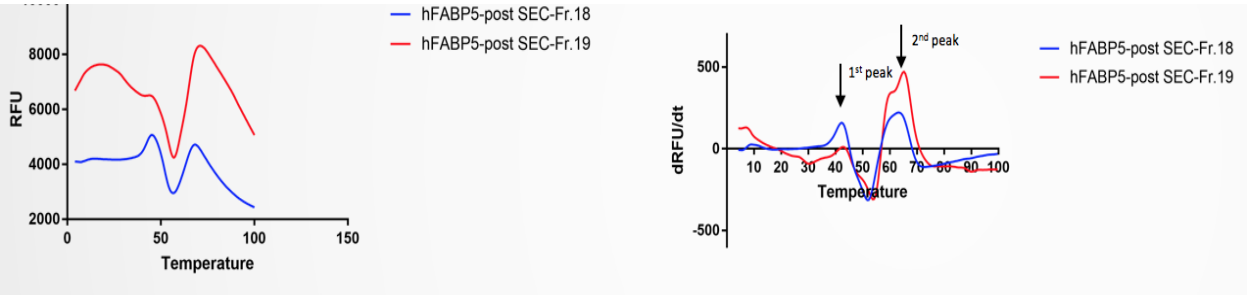

Figure II- 11. Thermal stability assay for dimer FABP5 (fraction 18 in blue and 19 in red).

Based on TSA profile of dimer fractions, we speculated that the 1st peak suggests dissociation of monomer molecules within dimers (half of dimer) and the 2nd peak might be related to the stability of each monomer (similar Tm for them, between 63-65 °C). Another explanation can be with the effect of ligand binding. If the dimer fractions are a mixture of Apo and Holo, then bound form (Holo) may have higher melting temperature. In order to examine our suggestion, we should express more dimer of this protein first. Then we can heat the dimer to 42°C and run SEC column to see if we have monomer or not. Also, we can saturate the dimer with ligand and measure the meting point.

II-6 Study the occurrence of domain swapping by mutational analysis of human FABP5.

Investigating domain swapping in amino acid sequence of hCRBPII indicates some important residues which can increase the ratio of dimerization significantly. One of these residues is Tyr60, since mutation Y60L and Y60W were reported to increase the ratio of dimer to monomer. The reason for more dimerization for these mutations seems to be cause by removal of the hydrogen bond between Glu72 and Y60L. Here, in our FABP5 project, we are looking for mutations that can increase the ratio of dimerization of this protein. In order to do this, we compared the structure of the monomer and dimer of hFABP5. In the monomer, there is a hydrogen via a bond through water network between T64 and T57; however, in the dimer there is a hydrogen bond between the water network between Q65 of one chain and T57 of the other chain **(figure II-12)**.[11]

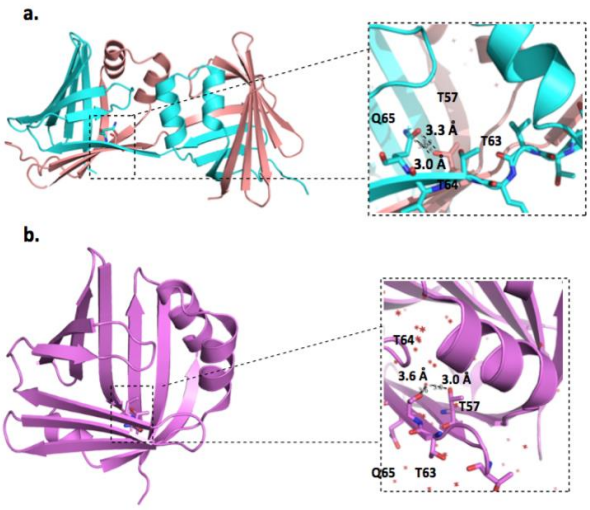

Figure II- 12. a) Domain swapped dimer of FABP5 (PDB 4AZR), there is a hydrogen bond between Q65 and T57. b. monomer structure (purple, PDB ID 4LKP) Gln65 is pointing out of toward the solvent and T64 makes a hydrogen bond to Thr57 via water network.

We thought that by removing the hydrogen bond between Q65 with T57 in dimer, the formation of the monomer would be more probable because mutated glutamine to alanine which is a small hydrophobic residue. The chromatogram of size exclusion chromatography shows that in mutant Q65A we will obtained about the same amount of dimer compare to the wild type FABP5 which was not as we predicted **(figure II-13)**. We could obtain small crystals of this mutant but they did not diffract. Also, we made Q65W and Q65M mutations as well which in both mutations the monomer form of FABP5 was only expressed. We also applied the phase relationship for FABP5 to increase dimerization and we have done different mutations like T64L and T64E; however, our hypothesis seems not to be right and we obtained monomer form by expressing these mutations **(figure II-14)**.

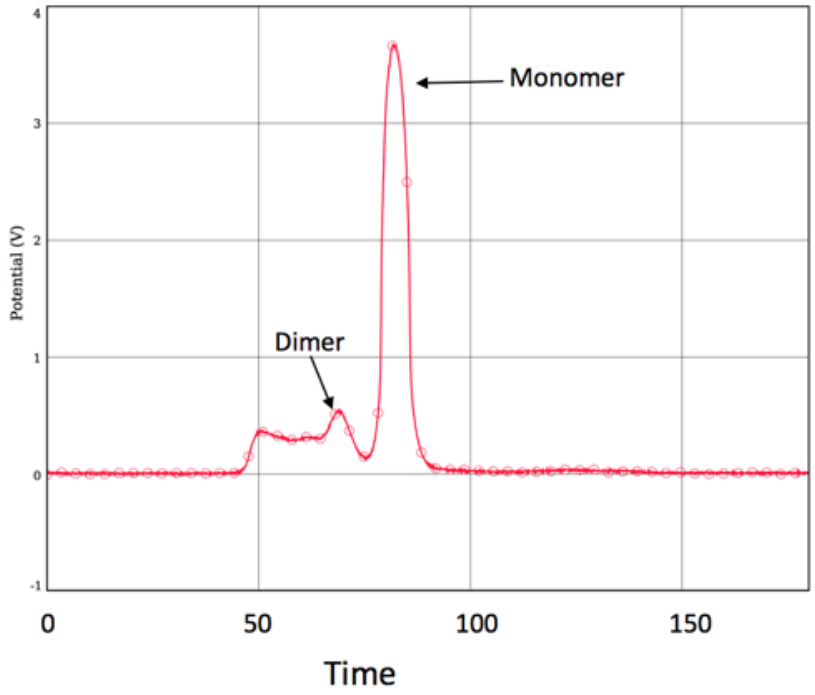

Figure II- 13. Shows SEC after expression of Q65A FABP5.

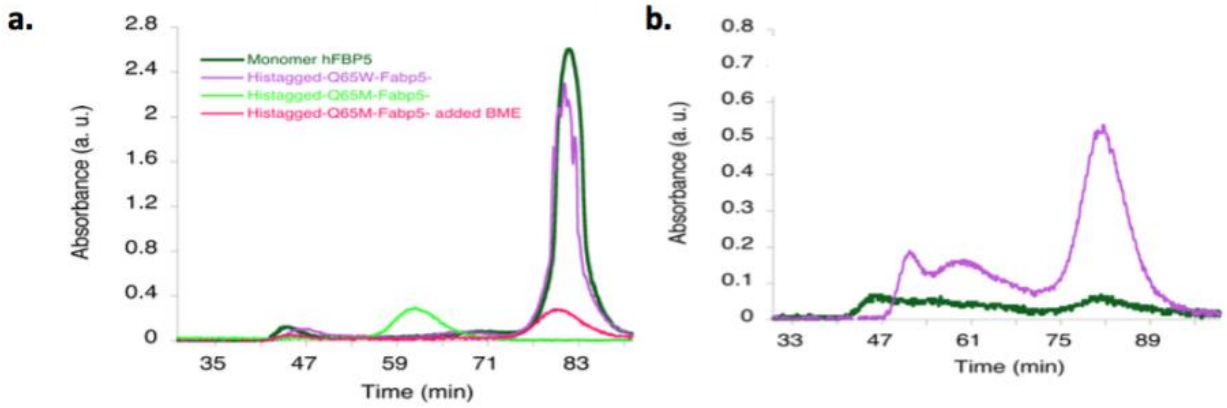

Figure II- 14. a) Size exclusion chromatogram of Gln65 mutants of FABP5. b. SEC of T64 mutants of FABP5 (Thr63E pink, T63L green).

II-7 Investigate the reasons behind the dimer expression

In order to get the structure of Apo FABP5 domain swapped dimer, we investigated the reasons behind the different expression levels of dimer and monomer in different trials. Since the

expression process was the same every time, we were suspicious that different colonies may make different amounts of domain swapped dimer each time. We thought colonies which make more dimer may have mutations in the sequence. In this case, we thought if we got dimer from that expression, we could do DNA purification and check for sequencing to see if our protein were mutated. Therefore, we collected a small portion of the E. Coli cells before adding the IPTG at OD_{600} and we did DNA purification. The sequence was exactly the same as before and there was no mutation in the sequence.

Another suggestion to get more dimer was changing the media. In this case, we used the premade LB broth containing casein digest peptone, sodium chloride, yeast and tris HCl. Also, we made LB broth by mixing the trypton, sodium chloride and yeast extract. However, in both cases the results were the same.

Changing the amount of IPTG also did not enhance dimerization. The IPTG were added to the final concentration of 1mM, 0.5mM and 0.25mM. The difference in these concentrations seems not to have an effect on increasing the dimer FABP5.

We also tried changing the temperature to increase the dimer ratio. We expressed FABP5 at 20°C and 25°C. However, the ratio of the dimer was low in both cases **(figure II-15)**. Increasing the expression temperature to 37°C does not help dimerization.

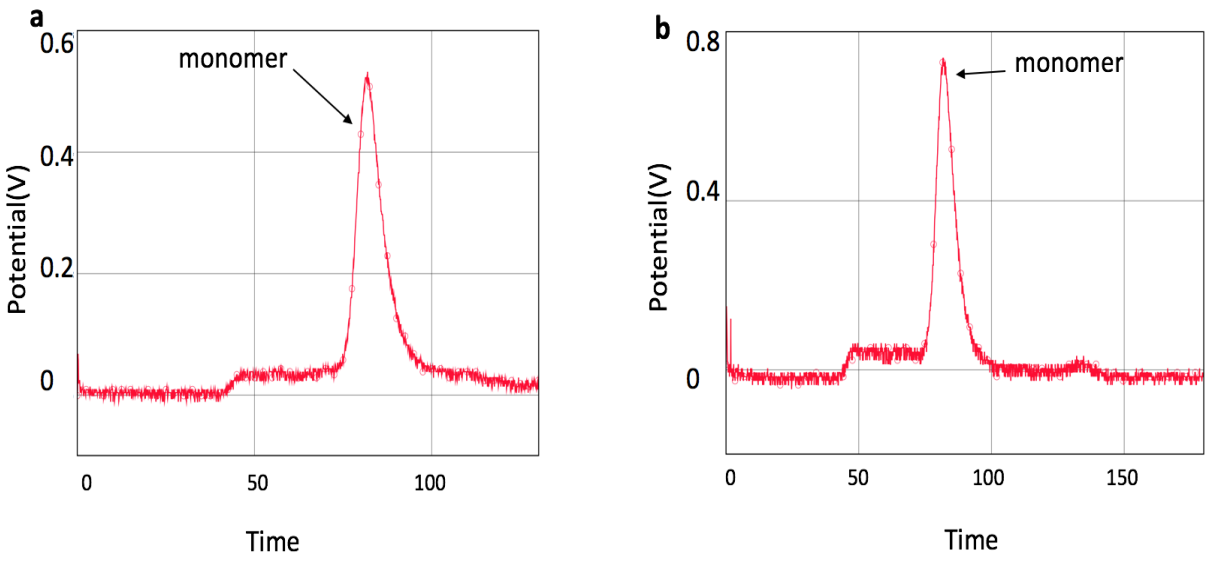

Figure II- 15. Chromatogram of SEC column after a) expression of FABP5 in 25°C. b) expression in 20°C.

We asked Sanson and his group to send us their construct for FABP5. It seems that they got the dimer every time that they expressed the FABP5. We amplified their construct and did the protein blast with our sequence to compare their sequence with ours **(figure II-16)**. There were two differences in these two constructs. There is a mutation on residue 61 (Leu to Val) and there are extra residues before the N terminus of our construct. In our construct for FABP5 residue 61 is Val and this residue is Leu in homo sapiens species.[1]

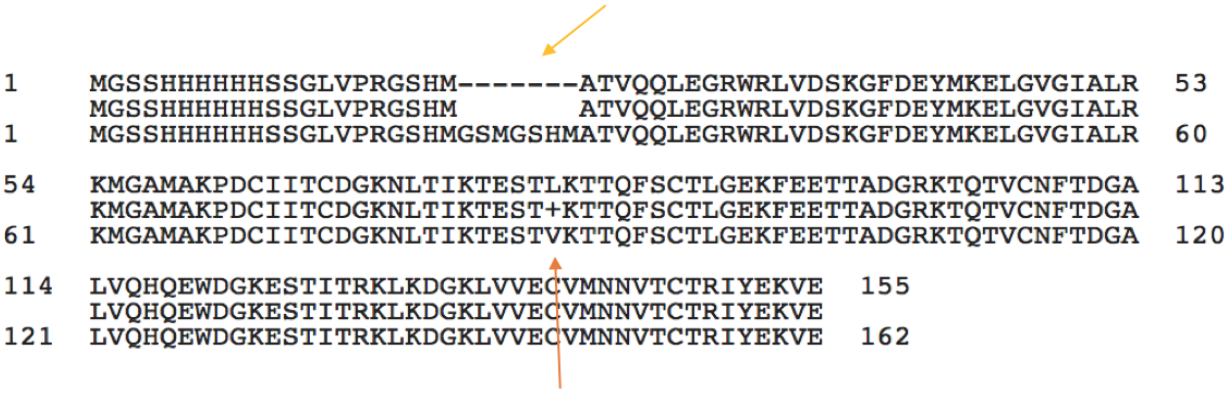

Figure II- 16. Blast between FABP5 from Sanson group (the top sequence) and our construct (the blew sequence) the arrows show differences.

We expressed their construct at 25°C. The level of expression was low, but the dimer ratio changed significantly compare to the expression with our construct **(figure II-17)**.

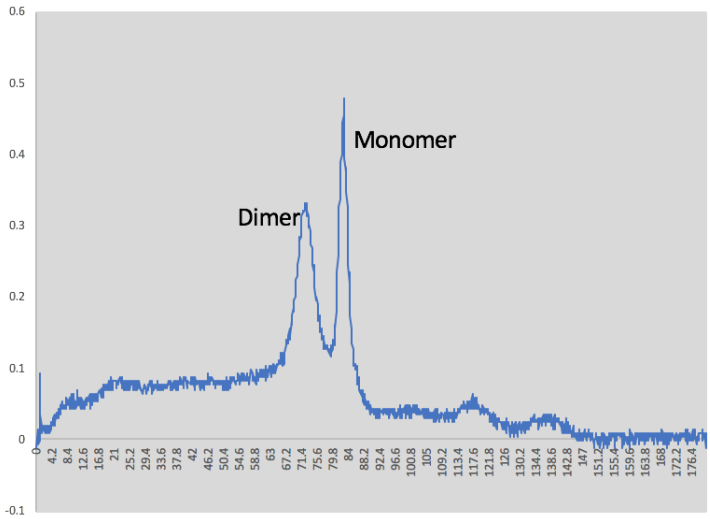

Figure II- 17. Chromatogram of SEC column after expression of FABP5 in 25°.

This result was very interesting for us and we plan to mutate the new construct to L61V human FABP5 to see if this mutation affects dimerization or not.

II-8 Experimental

II-8-1 Site-Directed Mutagenesis

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

Amino acid sequence of FABP5:

HHHHHHSSGLVPRGSHMGSMGSHMATVQQLEGRWRLVDSKGFDEYMKELGVGIALR KMGAMAKPDCIITCDGKNLTIKTESTVKTTQFSCTLGEKFEETTADGRKTQTVCNFTDG ALVQHQEWDGKESTITRKLKDGKLVV ECVMNNVTCTRIYEKVE

Primers

T64A-FABP5

Forward: 5’- CGTGAAGACGGCGCAGTTTTCAT-3’

Reverse: 5’- GTCGATTCAGTTTTAATAGTTAAGGTTCTTA-3’

T60P-FABP5

Forward: 5’- AACTGAATCGCCCGTGAAGACGA -3’

Reverse: 5’- TTAATAGTTAAGTTCTTACCATCGC-3’

T64E-FABP5

Forward: 5’- CGTGAAGACGGAGCAGTTTTCATG-3’

Reverse: 5’- GTCGATTCAGTTTTAATAGTTAAG-3’

T64L-FABP5

Forward: 5’- CGTGGAGACGCTGCAGTTTTCATG- 3’

Reverse: 5’- GTCGATTCAGTTTTAATAGTTAAG- 3’

Q65W-FABP5

Forward: 5’- GAAGACGACGTGGTTTTCATGCAC-3’

Reverse: 5’-ACGGTCGATTCAGTTTTAATAG -3’

Q65M-FABP5

Forward: 5’-GAAGACGATGTTTTCATGCAC-3’

Reverse: 5’-ACGGTCGATTCAGTTTTAATAG-3’

II-8-2 Protein Expression and Purification of FABP5

We expressed proteins by using Escherichia coli BL21 cells and using the T7 expression system. After expressing the cells in LB media at OD_{600} of 0.4-0.7, the IPTG were added to a final concentration of 0.4 mM. Then, the mixture was incubated at both 20°C and 20°C for 20-24 h and cells were collected by centrifugation at 5000rpm at 4°C for 20 min after that. Cells were lysed using the FABP5 lysate buffer (20mM Tris pH 8.5, 200mM NaCl), by sonication on ice.

Then by 20 min centrifugation at 10,000rpm at 4°C, the supernatant was collected and loaded onto a Ni–NTA column from GE healthcare company. After mixing and incubating the protein with the Ni resin, the wash buffers containing lysate buffer and 20 mM imidazole, 30mM imidazole and 50 mM imidazole were added, respectively. The proteins were then eluted with buffer containing 100, 150 and 200 mM imidazole. For removing the lipid from the protein (delipidation), we mixed Lipidex-5000 resin with our protein and the mixture were incubated for 1 hour at 37°C. In case of removing the 6 His tagged from the N terminus, the purified proteins were incubated with thrombin from GE Healthcare Life Sciences at 10 units per milligram of protein at 4°C overnight. Then the cleaved proteins were loaded onto an Ni–NTA columns and the cleaved proteins collected. The samples were concentrated to around 6-8mg/mL and 3mL of the concentrated protein were loaded onto our SEC column which was Superdex S75 16/600 HiLoad column from GE healthcare company and equilibrated with 1CV PBS pH 8.5. The peak fractions were collected and SDS PAGEs were run to check the purity of the protein. Then, purified proteins were concentrated to for crystallization.

II-8-3 Crystallization and refinement

II-8-3-1 Crystallization of Apo FABP5

The concentrated Apo FABP5 with His tag was concentrated to 6 mg/mL in lysate buffer containing 20 mM Tris, 200 mM NaCl, pH = 8.5. We used hanging drop vapor diffusion method using 1 μL of protein solution, 1 μL of crystallization solution in the drop and 1 mL of crystallization solution in the reservoir for crystal grow. This protein crystallized in c. We collected our data at the Advanced Photon Source (APS) (Argonne National Laboratory IL) LS-CAT, (sector 21-ID-D, beamline F and G) using a MAR300 detector and 1.00Å wavelength radiation at

100K. Using the HKL2000 software, the diffraction data were indexed, processed and scaled.[15] The structure was solved by molecular replacement using PHASER in PHENIX and FABP5 (PDB: 4AZR as a search mode.[16] The electron density map was produced by REFMAC5 in the CCP4 package or Phaser-MR in PHENIX.[17,18,19] Model rebuilding and water replacement were done using COOT program.[23] The structures were refined using PHENIX program packages.

II-8-3-2 crystallization of FABP5 domain swapped dimer binds to palmitic acid (first structure)

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

II-8-3-3 Crystallization of FABP5 domain swapped dimer binds to palmitic acid (second structure)

The concentrated FABP5 with His tag was treated with lipidex column for 1 h in 37°C. Then the protein concentrated to 6 mg/mL in lysate buffer containing 20 mM Tris, 200 mM NaCl, pH = 8.5. We used mosquito® Robot from TTP labtech for crystal grow. The Holo protein binds to palmitic acid is crystallized in %20 PEG 2000, 0.2M Trimethylamine N-oxide dihydrate, 0.1 tris pH= 8.5 which was different condition as before. For Data collection and refinement please refer to II-8-3-1.

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

	WT-hFABP5
Space group	P21 21 21
a (Å)	40.89
b (Å)	55.10
c (Å)	126.36
α (°)	90
β (°)	90
δ (°)	90
Molecules per asymmetric unit	2
Total reflection	1524152
Unique reflection	4421
Completeness (%)	99.41 (92) [a]
Average I/σ	30
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

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

Space group a (Å)	$P6_322$
a (Å)	97.999
b (Å)	97.999
c (Å)	65.907
α (°)	90.00
β (°)	90.00
δ (°)	120.00
Number of reflections	6870
after outlier rejection	6870
Completeness	99.88%
R-work	0.2312
R-free	0.3198
Rotamer outliers	0.85 %
C-beta deviations	0
RMS(bonds)	0.0081
RMS(angles)	0.99
MolProbity score	1.87
Resolution	2.50

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

Space group a (Å)	$C222_1$
a (Å)	66.265
b (Å)	114.735
c (Å)	107.729
α (°)	90.00
β (°)	90.00
δ (°)	90.00
Number of reflections	21127
after outlier rejection	21127
Completeness	99.05%
R-work	0.2753
R-free	0.3107
Rotamer outliers	0.0 %
C-beta deviations	0
RMS(bonds)	0.0078
RMS(angles)	0.85
MolProbity score	1.33
Resolution	2.2

BIBLOGRAPHY

BIBLOGRAPHY

1. Sanson, B. *et al.* Crystallographic study of FABP5 as an intracellular endocannabinoid transporter. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **70,** 290–298 (2014).

2. Armstrong, E. H., Goswami, D., Griffin, P. R., Noy, N. & Ortlund, E. A. Structural basis for ligand regulation of the fatty acidbinding protein 5, peroxisome proliferator-activated receptor β/δ (FABP5-PPARβ/ δ) signaling pathway. *J. Biol. Chem.* **289,** 14941–14954 (2014).

3. Sessler, R. J. & Noy, N. A ligand-activated nuclear localization signal in cellular retinoic acid binding protein-II. *Mol. Cell* **18,** 343–353 (2005).

4. Nossoni, Z. *et al.* 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 (2014).

5. Haunerland, N. H. & Spener, F. Fatty acid-binding proteins - Insights from genetic manipulations. *Prog. Lipid Res.* **43,** 328–349 (2004).

6. Ong, D. E. & Page, D. L. Cellular retinol-binding protein (type two) is abundant in human small intestine. *J. Lipid Res.* **28,** 739–745 (1987).

7. Kaczocha, M., Glaser, S. T. & Deutsch, D. G. Identification of intracellular carriers for the endocannabinoid anandamide. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 6375–80 (2009).

8. Liu, Y. & Eisenberg, D. 3D domain swapping: As domains continue to swap. *Protein Sci.* **11,** 1285–1299 (2002).

9. Sulsky, R. *et al.* Potent and selective biphenyl azole inhibitors of adipocyte fatty acid binding protein (aFABP). *Bioorganic Med. Chem. Lett.* **17,** 3511–3515 (2007).

10. Koharudin, L. M. I., Liu, L. & Gronenborn, A. M. Different 3D domain-swapped oligomeric cyanovirin-N structures suggest trapped folding intermediates. *Proc. Natl. Acad. Sci.* **110,** 7702–7 (2013).

11. Assar, Zahra. Elucidation of ILBP family folding pathway and study of reengineering them as florescent protein tags via structural analysis, PhD thesis, Michigan State University (2018)

12. Bloch, K. Unsaturated Fatty Acids in Microorganisms *. **237,** 2064–2069 (1962).

13. Lo, M. C. *et al.* Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem.* **332,** 153–159 (2004).

14. Kohlstaedt, M., Von Der Hocht, I., Hilbers, F., Thielmann, Y. & Michel, H. Development

of a Thermofluor assay for stability determination of membrane proteins using the Na^+/H^+ antiporter NhaA and cytochrome c oxidase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71,** 1112–1122 (2015).

15. Abergel, C. Molecular replacement: Tricks and treats. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69,** 2167–2173 (2013).

16. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 22–25 (2010).

17. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 213–221 (2010).

18. Project, C. C. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763 (1994).

19. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 486–501 (2010).

Chapter III: Mammalian expression of hCRBPII

III-1 Introduction

III-1-1 Mammalian expression

Highly purified and low-cost expression of the majority of recombinant protein for biophysical studies can be achieved by bacterial expression. Although many mammalian proteins have been produced in bacteria, yeast, or insect cells, expression in mammalian cells is still needed to investigate the function and right posttranslational modifications for these proteins. Mammalian expression is used for studying the function of a particular protein in physiologically relevant environment. This expression method allows for the highest level of posttranslational processing and functional activity of the protein. This system is commonly used for the production of antibodies and therapeutic proteins. [1,2,3]

III-1-2 HeLa cells

Many eukaryotic expression systems are available for expression of mammalian proteins including those that contain unique post-translational modifications. The most common methods currently include yeast, baculovirus expression vector systems and mammalian cell systems. One of the most common mammalian cells used for mammalian culture is HeLa cells. HeLa cell are cervical cancer cells which come from a sample taken from a woman called Henrietta Lacks. These cells are the oldest and most commonly used human cell line which is used for scientific research. For instance, they were used to test the first polio vaccine in the 1950s, test the infection of human cells by parvo virus in 1953, and also for testing the heptamethine dye IR-808 and etc. The genome of Hela cells was sequenced in 2013. These cells can proliferate rapidly, even faster than cancer cells. They have an altered version of telomerase, which can prevent aging and cell death. The

adaptation of HeLa cells to grow in tissue culture plates make them difficult to control and through the improper maintenance, they can lead to contamination and interfere with other cell cultures in the same laboratory **(figure III-1)**.[4,5,6,7]

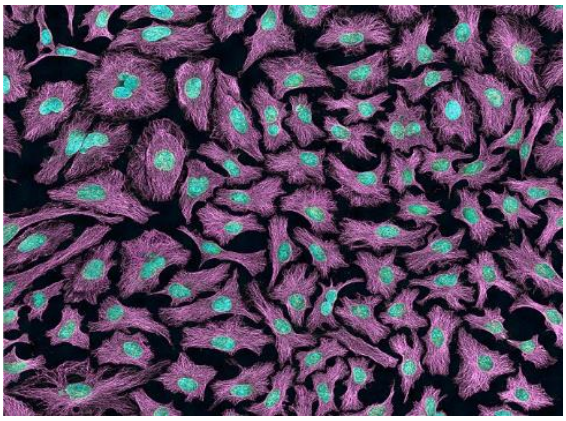

Figure III- 1. Multiphoton fluorescence image of HeLa cells with cytoskeletal microtubules (magenta) and DNA (cyan).

III-1-3 Western blotting as a way for detecting protein expression in mammalian cells.

Western blotting is an analytical technique for detecting specific proteins from a complex mixture of proteins extracted from cells. This method was introduced by Towbin, et al. in 1979 In this technique, an extract from cells which is a mixture of proteins is separated by molecular weight, through addition of sodium dodecyl sulfate and gel electrophoresis (SDS PAGE gel). Then samples from a gel are transferred to a nitrocellulose or polyvinylidene difluoride (PVDF) membrane. The term "blotting" refers to this step which means transferring biological samples from a gel to a membrane. For detection of proteins, the membrane is then incubated with label antibodies specific to the protein of interest. The target protein can be identified by the specificity of the antibody-antigen interaction. [8,9,10,11]

III-1-4 Study Physiological relevance of Domain swapping

The existence of domain swapping for hCRBPII and FABP5 may lead to allosteric regulation, also huge effect on the folding pathway for this family of proteins. Previously, all studies in our lab for domain swapping was through bacterial expression, however, in order to check the physiological relevance of this phenomena we decided to investigate the existence of domain swapping in mammalian cell line. Since we had higher information on domain swapping of hCRBPII, we used this protein as our first target for this study.[12] Most post translational modifications occur in endoplasmic reticulum where there are enzymes that facilitates proper folding for proteins.[13] Therefore, mammalian expression of hCRBPII can help us through studying the folding pathway for this protein.

III-2 Mammalian expression and western blotting of hCRBPII

For mammalian expression and detection of our target protein, we used a HeLa cell line and western blotting, respectively. We cloned hCRBPII into CMV(cytomegalovirus) with a flag tag on the N terminus of our target gene. Flag tag is a peptide tag that can help us to identify our target protein from other proteins from the extract by using a proper antibody. We used Q108K-T51D mutation of hCRBPII, because this variant gave virtually all the domain swapped dimer form of the protein in bacterial expression. After preparation of the aforementioned plasmid, we prepared 6 well culture plate of HeLa cells and did transfection step. For transfection, we had to try different concentrations from the construct to find the best concentration. The first three wells were as controls; the first well was blank without changing the media, the second was blank with changing the media through transfection and the third one was transfected with 350 ng of empty cmv-flag tag vector. The rest of the wells were transfected with different amount of the target

construct; 350 ng, 750 ng, 1050 ng of the target plasmid, respectively. After this step, for identifying the hCRBPII from a mixture of proteins in the extracts, western blotting was done. For western blotting, the extract of each well was collected separately and 20ng of each extract was run through SDS PAGE and transferred to the membrane. The membrane was soaked for 24 h with anti-flag tag antibody and 1 h with goat anti mouse hrp as a secondary antibody for detection of the target protein. Although we thought the antibodies should just bind to our flag tagged protein, the exposed X-ray film shows many other proteins after developing. This may be because of the high amount of extracts (20 ng) which was load to the gel. The bands in 15 KD seems to be our target protein. The 15KD light bands in sample 2 and 3 (the extracts from control samples which were without our target plasmid) probably were because of contamination with sample 4 (20Mg of transfected extract) **(figure III-2)**.

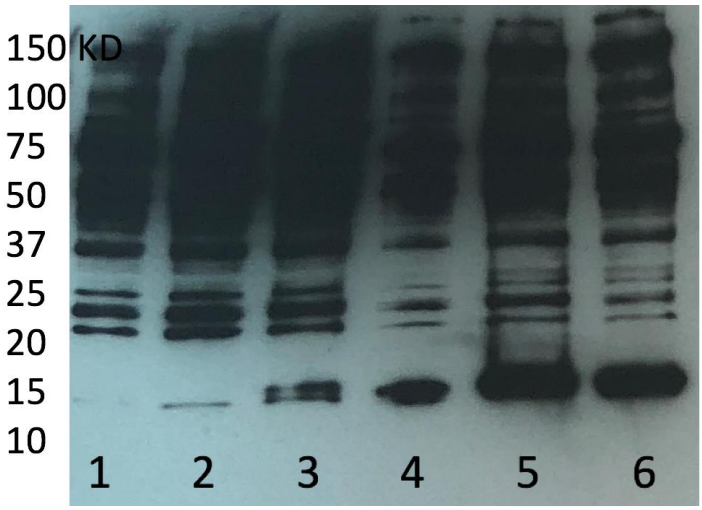

Figure III- 2. Exposed x-ray film from the western blotting. Row 1, 2 and 3 are extracts from controls (cells without changing media, cells with changing the media in transfection, transfection with 350 ng of empty cmv-flag tag vector, respectively). sample 4,5 and 6

The X-ray film from the western blotting shows darker 15KD band in sample 5 which was transfected by 750ng of the plasmid; Therefore, we decided to use the same amount of the plasmid

(750ng) for our next transfections in each well. For improving the results of western blotting and increasing the specificity of the antibody to the flag tagged protein, we decided to lower the amount of extract that we inserted to SDS PAGE. This time 8Mg of the controls (one without the transfection and one transfected by empty pCMV-flag tag), 8Mg,4Mg and 2Mg of the extract transfected by plasmid were used in SDS PAGE. Also, the first and secondary antibody were reused from last time and the membrane soaked for 1 h instead of 24 h with first antibody. The results showed substantially less background. These data showed us that by using around 4-8Mg of the extract and soaking 1h with first anti body, we can improve the western blotting results **(figure III-3)**.

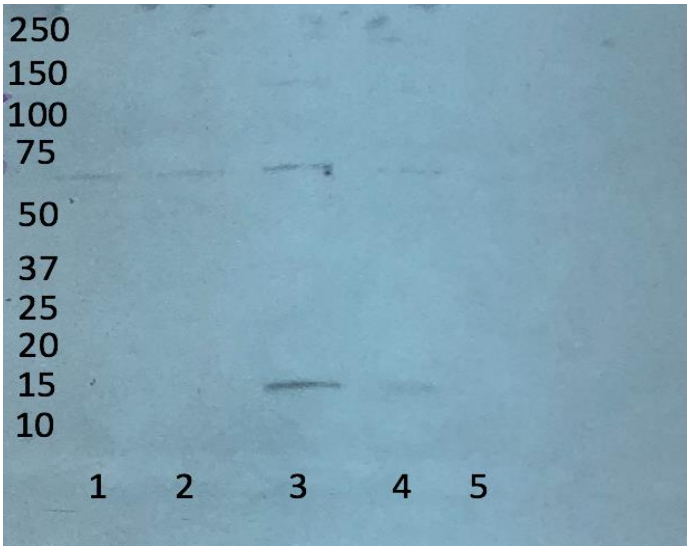

Figure III- 3. Exposed x-ray film from western blotting. Sample 1 and 2 are from 8Mg extract of control and 8Mg extract of empty pCMV-flag tag vector. Sample 3,4 and 5 are from 8Mg,4Mg and 2Mg extract of infection with 750ng plasmid, respectively.

III-3 Investigate the size of hCRBPII in mammalian expression

Through our last data, we found out the expression of hCRBPII in mammalian cell. But in order to check the existence of domain swapping, we needed a way to detect the folded form of protein. We were looking to compare the hCRBPII in mammalian expression to monomer and

dimer form from bacterial expression. The first solution was running a native gel instead of SDS PAGE; however, flag tag is negatively charged and even if the protein folded in domain swapped dimer form, we could not compare it with dimer without flag tag. Another solution for detecting the dimer was running the size exclusion chromatography from the extract and running the western blot of the fractions from the column. By western blotting, we were able to find the fractions in which the protein came out and then check if they were the same as the fractions for the domain swapped dimer of hCRBPII. We tried the second way. We first expressed monomer and dimer hCRBPII in bacteria and ran it through a superdex 75 size exclusion chromatography column which is a different column that used is last chapters. The monomer (Q108k: K40L: T51V hCRBPII) came out at fractions 24,25 and 26. The dimer (Q108K: T51DhCRBPII) came out mostly at fractions 20 and 21 **(figure III-4)**.

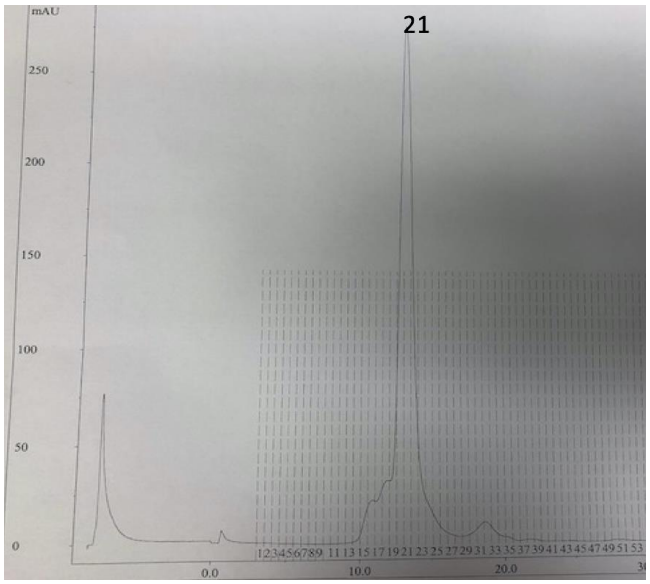

Figure III- 4. Chromatogram of size exclusion chromatography for 1.5mg of domain swapped dimer Q108K:T51D hCRBPII.

Since the amount of protein in each fraction is very low as a result of dilution in size exclusion chromatography column, we had to express more proteins in mammalian cells. So

another 6 well culture plate were transfected by 750ng plasmid for each well. The resulted extracts (0.5 mg) from the 6 wells of culture plate were mixed and run through size exclusion chromatography column **(figure III-5)**.

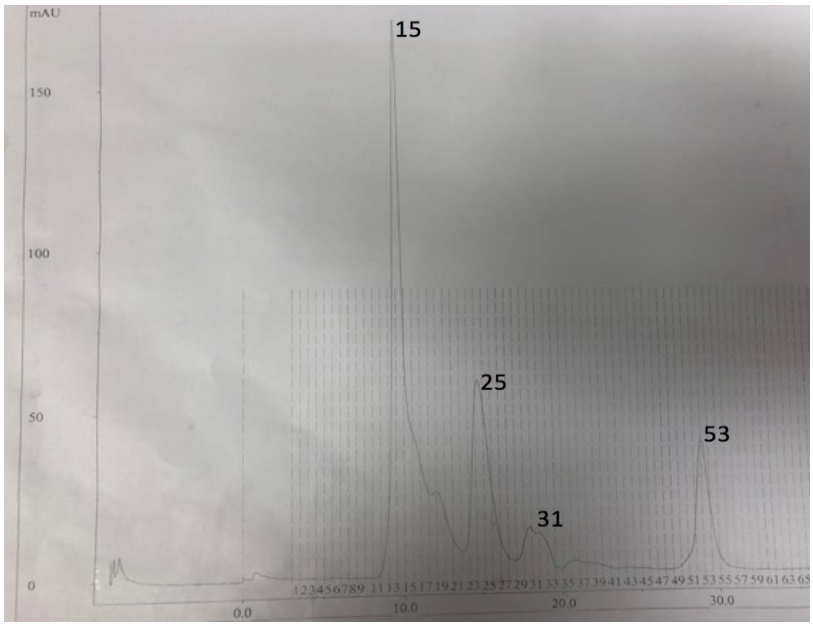

Figure III- 5. Chromatogram of size exclusion chromatography for 0.5mg extract from transfection by 750ng plasmid in 6 well culture plates.

Fractions 13-28 from size exclusion chromatography were concentrated and run through SDS PAGE. The western blot data surprisingly shows the 15KD band only in fraction 16 **(figure III-6)**. This can be a higher order oligomer form for the hCRBPII or it can be our target protein bound to another protein. In order to find more information, first, we have to run ladder through the size exclusion chromatography to find the exact size for fraction16. Also, we will clone the flag tag to the N termini of hCRBPII gene in our bacterial construct (vector pET 17b). In this way, both of proteins from mammalian and bacterial expression have flag tag and we are able to compare the results from native gel.

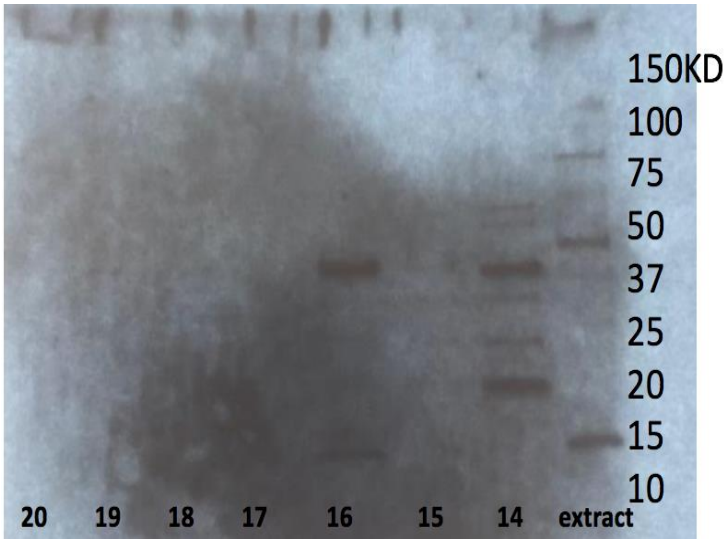

Figure III- 6. Exposed x-ray film from western blotting for extract before running the SEC column and fraction 14-20 after SEC column.

III-4 Experimental

III-4-1 Site directed mutagenesis

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

Amino acid sequence of hCRBPII with flag tag in N terminus

MDYKDDDDKLADRMTRDQNGTWEMESNENFEGYMKALDIDFATRKIAVRLTQTKVID
QDGDNFKDKTTSTFRNYDVDFTVGVEFDEYTKSLDNRHVKALVTWEGDVLVCVQKGE
KENRGWKKWIEGDKLYLELTCGDQVCRQVFKKK

Primers

Q108K-hCRBPII

Forward: 5’-CCGCGGCTGGAAGAAGTGGATTGAGGGGG -3’

Reverse: 5’-CCCCCTCAATCCACTTCTTCCAGCCGCGG-3’

K40L-hCRBPII

Forward: 5’-CTCACTCAGACGCTGGTTATTGATCAAGATGG -3’

Reverse: 5’-CCATCTTGATCAATAACCAGCGTCTGAGTGAG-3’

T51D-hCRBPII

Forward: 5’-GGTGATAACTTCAAGGATAAAACCACTAGCAC-3’

Reverse: 5’-GTGCTAGTGGTTTTATCCTTGAAGTTATCACC-3’

T51V-hCRBPII

Forward: 5’-GGTGATAACTTCAAGGTAAAAACCACTAGCAC-3’

Reverse: 5’-GTGCTAGTGGTTTTTACCTTGAAGTTATCACC-3’

III-4-2 Mammalian expression

In order to do transient transfection, all of these steps needed to be done in culture room. 2.5 mL HeLa cells (concentration 0.8×10^5) in DMEM were added to each well in 6 well culture plate. The DMEM contained 5% FBS and Penicillin/Streptomycin (P/S). After that, the aforementioned plate incubated at 37˚C for 24h (cells should be 90% confluent at this time). The next step is preparing two solutions for transfection. 14mL polypropylene tubes were labeled for A and B solutions. 250ML DMEM (without FBS and P/S) were added to each tube. 2.5ML

Lipofectamine 3000 reagent from Invitrogen company were added to solution B tubes and incubated for 15 min. Meanwhile, 1ML of P3000 from Invitrogen company and appropriate amounts of DNA as explained in part III-2 were added to solution A tubes and incubated for 15min. Then, solution B was added to solution A and the mixture was incubated for another 15min. Meanwhile, the old media were removed from the wells from culture plate and the plates washed with 2mL PBS. We aspirated the PBS and 500ML DMEM without FBS and P/S were added to each well. Each mixture of solution A/B were added drop wise to each well accordingly and the plate was incubated in 37˚C/5% CO2 for 5-7 hrs. The cells were harvested after 48 hrs. The media was removed from the wells. 1-1.5mL of PBS was added to each well and the cells were scrapped from the plate by scrapper. The mixture of cells with PBS were collected in Eppendorf tubes and spin for 7 min in 2000rpm. The PBS solution were removed from the cells. 50ML of the extraction buffer were added to pellet and mixed well. Then, the mixtures were incubated in ice for 20 min and then spin for 13000rpm for 7min in cold room. The supernatants (extract) were transferred to new tubes and were frozen in -80 ˚C.

BIBLOGRAPHY

BIBLOGRAPHY

1. Khan, K. H. Gene expression in mammalian cells and its applications. *Adv. Pharm. Bull.* **3,** 257–263 (2013).

2. Nettleship, J. E., Assenberg, R., Diprose, J. M., Rahman-Huq, N. & Owens, R. J. Recent advances in the production of proteins in insect and mammalian cells for structural biology. *J. Struct. Biol.* **172,** 55–65 (2010).

3. Portolano, N. *et al.* Recombinant Protein Expression for Structural Biology in HEK 293F Suspension Cells: A Novel and Accessible Approach. *J. Vis. Exp.* **1,** 1–8 (2014).

4. Rahbari, Raheleh et al. "A Novel L1 Retrotransposon Marker for HeLa Cell Line Identification." *BioTechniques* 46.4. 277–284. (2009)

5. Capes-Davis, A. *et al.* Check your cultures! A list of cross-contaminated or misidentified cell lines. *Int. J. Cancer* **127,** 1–8 (2010).

6. nigma.nit.gov

7. Masters, J. R. HeLa cells 50 years on: the good, the bad and the ugly. *Nature Reviews Cancer* **2,** 315 (2002).

8. Towbin, H., Staehelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci.* **76,** 4350–4354 (1979).

9. Mahmood, T. & Yang, P.-C. Western Blot: Technique, Theory, and Trouble Shooting. *N Am J Med Sci* **4,** 429–434 (2012).

10. Renart, J., Reiser, J. & Stark, G. R. Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. *Proc. Natl. Acad. Sci.* **76,** 3116–3120 (1979).

11. Alwine, J C, D J Kemp, and G R Stark. "Method for Detection of Specific RNAs in Agarose Gels by Transfer to Diazobenzyloxymethyl-Paper and Hybridization with DNA Probes." *Proceedings of the National Academy of Sciences of the United States of America* 74.12, 5350–5354. (1997)

12. Nossoni, Z. *et al.* 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 (2014).

13. Braakman, Ineke, and Daniel N. Hebert. "Protein Folding in the Endoplasmic Reticulum." *Cold Spring Harbor Perspectives in Biology* 5.5 (2013)