





**PLACE IN RETURN BOX** to remove this checkout from your record.  
**TO AVOID FINES** return on or before date due.  
**MAY BE RECALLED** with earlier due date if requested.

| DATE DUE | DATE DUE | DATE DUE |
|----------|----------|----------|
|          |          |          |
|          |          |          |
|          |          |          |
|          |          |          |
|          |          |          |
|          |          |          |
|          |          |          |
|          |          |          |
|          |          |          |
|          |          |          |



**X-RAY CRYSTALLOGRAPHIC STUDIES OF MUTANTS OF CELLULAR  
RETINOIC ACID BINDING PROTEIN (II) TOWARD DESIGNING A MIMIC  
OF RHODOPSIN**

**By**

**Soheila Vaezeslami**

**A DISSERTATION**

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

**DOCTOR OF PHILOSOPHY**

**Department of Chemistry**

**2006**



## **ABSTRACT**

### **X-RAY CRYSTALLOGRAPHIC STUDIES OF MUTANTS OF CELLULAR RETINOIC ACID BINDING PROTEIN (II) TOWARD DESIGNING A MIMIC OF RHODOPSIN**

By

Soheila Vaezeslami

Although it has long been known that 11-*cis*-retinal is the unique visual chromophore in the human eye it is still not clear how the interactions of retinal with the binding pockets of each of the four different opsins lead to the perception of different wavelengths, known as wavelength regulation. Since rhodopsins are transmembrane proteins, spectroscopic and crystallographic experiments on these proteins are very challenging. Therefore, to study the theories about wavelength regulation, a more tractable mimic of rhodopsin would be helpful. Cellular retinoic acid binding protein (CRABP) II, a small cytosolic protein that binds to retinoic acid as its natural substrate, is an attractive candidate. Systematic mutations supported by X-ray crystallographic data allowed conversion of CRABP II from a retinoic acid binding to a retinal binding protein. Crystal structures of several apo, retinoic acid- and retinal-bound mutants of CRABP II at high resolutions elucidate structural changes introduced by different mutations on the structure of the mutants and show the importance of different mutations toward making the mimic. For example, the structure of the mimic not only shows the Schiff base (SB) formation between retinal and a designed Lys residue inside the pocket of the mimic but also shows the importance of a Glu as a counter anion for the protonated Schiff base (PSB), as it was

observed in bovine rod-rhodopsin. Based on these structural data we propose a mechanism for PSB formation inside the pocket of the mimic. Further rational mutations within the pocket of the surrogate protein may reproduce the interactions of different visual rhodopsins and 11-*cis*-retinal.

CRABP II is a small, cytosolic protein that solubilizes and transfers retinoic acid (RA) to the nucleus while also enhancing its transcriptional activity. We have determined the first high-resolution structure of apo-wild type (WT) CRABP II at 1.35 Å. Using three different data sets collected on apo-WT CRABP II we have shown that apo- and holo-CRABP II share very similar structures. Binding of RA appears to increase the overall rigidity of the structure, although the induced structural changes are not as pronounced as previously thought. The enhanced structural rigidity may be an important determinant for the enhanced nuclear localization of the RA-bound protein. Comparison of our apo-WT with the apo-R111M structure shows that mutation of Arg111, a conserved residue of CRABP II and a key residue in RA binding, causes major structural changes in the molecule. We further investigated the structural importance of Arg111 by determining the structures of four other CRABP II mutants. Our structures also demonstrate structural changes induced by crystal packing and show that a crystal can harbor demonstrative structural differences in the asymmetric unit.



*This Dissertation is Lovingly Dedicated to my Mother, Fatemeh Rasian,  
whose Support, Encouragement, and Constant Love  
Have Sustained me throughout my Life.*

## **ACKNOWLEDGEMENTS**

I would like to acknowledge many people for making my doctoral work the most memorable and fruitful experience of my life. I would especially like to thank my advisor, Professor James Geiger, for his generous time and commitment. Throughout my doctoral work Jim encouraged me to develop independent thinking and research skills. He continually stimulated my analytical thinking and greatly assisted me with scientific writing.

I owe a very special thanks to Professor Babak Borhan. First of all, because of his help in following my application for being in this PhD program and second because of all the encouragement, help and support he offered me during the last few years. I really feel fortunate for having the opportunity to collaborate with Babak and his group on my PhD project.

I would like to specially thank Sara, Stacy and Blanka for being my good friends in the lab. Sara and Stacy were the ones who patiently helped me to get started in the lab and learn crystallographic techniques and software. I cannot thank them enough for these and will miss them so very much when I leave here. Also I would like to thank Erika for all her work on this project before I join the lab and Xiaofei for taking care of the crystallographic part of the project after me. Also I would like to thank other former and present members of the lab for being great lab mates and supporting me in any way they could, particularly Jim's wife (Kathy), Marta, Xiangshu, Andy, Suzie and Lei.

I owe a very special thanks to Chrysoula, not only for being one of my best friends and a great collaborator during this project, but also for taking her time to

revise my thesis. I believe it is enough to say that Chrysoula is the best friend and collaborator that someone can wish for.

I would like to thank my good friend, Anne Fischer, for her time and help when I was preparing for my second year seminar. Also I like to give a special thanks to my best friend Prerna Sonthalia. My experience of graduate school would have not been the same without Prerna. She also introduced me to the ACS-Women in Chemistry (ACS-WiC) group in the department, which I really enjoyed participating in.

I owe a very special thanks to Kaveh Jorabchi, my best friend, who has been a great motivation for me to pursue my studies in the graduate school and who has always been there for me whenever I needed help. Kaveh is the most genius person I have ever met in my life; and I am so grateful for our long-term friendship.

I would also like to thank my loving parents and two sisters, Saeideh and Sima, for their love, encouragement and unconditional support. I would also like to express my gratitude to my dear uncle (Mohammad), here in the U.S., for all the love, guidance and support he offered me while I was far from my family. My mom and uncle were the ones who believed in me even more than what I did in myself, and helped me to achieve my dream of getting my PhD and becoming the first “doctor” of the family!

I am in gratitude for the financial support provided by the American Crystallographic Association (ACA), ACS local section, graduate school, and MSU Council of Graduate Students (COGS) to attend the annual scientific meetings to present my research.

## TABLE OF CONTENTS

|                                  |      |
|----------------------------------|------|
| List of Tables                   | x    |
| List of Figures                  | xi   |
| Key to Symbols and Abbreviations | xvii |

### CHAPTER I

#### Introduction

|       |                                                      |    |
|-------|------------------------------------------------------|----|
| 1.1   | Introduction to Vision and Scope of the Project      | 1  |
| 1.1.1 | Background                                           | 1  |
| 1.1.2 | Rod and Cone Photoreceptor Cells                     | 3  |
| 1.1.3 | Rhodopsin System                                     | 7  |
| 1.1.4 | Color Vision with the Same Chromophore               | 16 |
| 1.1.5 | Need for a Protein Mimic of Rhodopsin                | 35 |
| 1.2   | CRABP II: Physiological Importance and the Structure | 37 |
| 1.3   | Literature Cited                                     | 43 |

### CHAPTER II

#### Reengineering CRABP II into a Rhodopsin Surrogate

|       |                                                                                                                 |     |
|-------|-----------------------------------------------------------------------------------------------------------------|-----|
| 2.1   | The Structure of CRABP II Bound to RA as a Starting Point                                                       | 63  |
| 2.2   | Spectroscopic Assays Used for Characterizing the Mutants                                                        | 70  |
| 2.2.1 | Circular Dichroism (CD) Spectroscopy                                                                            | 70  |
| 2.2.2 | Measurement of Extinction Coefficient ( $\epsilon$ )                                                            | 70  |
| 2.2.3 | Fluorescence Titration                                                                                          | 72  |
| 2.2.4 | MALDI-TOF Assays                                                                                                | 73  |
| 2.2.5 | UV-Vis Spectroscopy                                                                                             | 75  |
| 2.3   | X-Ray Crystallography                                                                                           | 75  |
| 2.4   | Enhancing the Fluorescence Titration Measurements                                                               | 83  |
| 2.5   | Design of a Lys Residue Capable of SB Formation (R132K)                                                         | 85  |
| 2.6   | Hydrophobic Tuning of the Pocket and Design of a Favorable Nucleophilic Attack                                  | 87  |
| 2.7   | Facilitation of the Nucleophilic Attack and Design of a Counter Anion for the PSB                               | 101 |
| 2.8   | Restoring Tyr134: Enhancing Formation of a Counter Anion, Orienting Retinal for a Favorable Nucleophilic Attack | 114 |
| 2.9   | Backbone Rotation: Favoring the Nucleophilic Attack                                                             | 124 |
| 2.10  | Proposal of a PSB Formation Mechanism in R132K:R111L:L121E•Retinal                                              | 127 |



|      |                                         |     |
|------|-----------------------------------------|-----|
| 2.11 | The Importance of the PSB Counter Anion | 129 |
| 2.12 | The Importance of the Arg111 Mutation   | 130 |
| 2.13 | Literature Cited                        | 132 |

## **CHAPTER III**

### **Structural Studies of Human Apo- and Holo-Cellular Retinoic Acid Binding Protein Type II: Opening Doors for New Insights into its Structure/Function Relationship**

|      |                                                                           |     |
|------|---------------------------------------------------------------------------|-----|
| 3.1  | Overall Structure of Apo-WT CRABP II                                      | 138 |
| 3.2  | Comparison between Mol A and Mol B of Apo-WT CRABP II                     | 143 |
| 3.3  | Changes Induced by the R111M Mutation                                     | 149 |
| 3.4  | Other Mutants Confirm the Structural Importance of Arg111                 | 151 |
| 3.5  | Overall Structure of CRABP II•RA and Comparison with Apo-WT CRABP II      | 165 |
| 3.6  | Comparison between the Crystal and NMR Structures                         | 173 |
| 3.7  | Comparison between the Crystal Structures of Apo-CRABP I and Apo-CRABP II | 174 |
| 3.8  | The Nuclear Localization Signal in CRABP II•RA                            | 175 |
| 3.9  | Conclusion                                                                | 180 |
| 3.10 | Literature Cited                                                          | 181 |

## **CHAPTER IV**

### **Mutation, Over-expression, Purification, Crystallization, X-ray Diffraction Analysis, Structure Solution and Refinement of CRABP II and Corresponding Complexes**

|        |                                                                |     |
|--------|----------------------------------------------------------------|-----|
| 4.1    | Protein Mutation, Over-expression and Purification of CRABP II | 183 |
| 4.2    | Crystallization and X-Ray Data Collection                      | 192 |
| 4.2.1  | Apo-WT                                                         | 194 |
| 4.2.2  | Apo-F15W                                                       | 195 |
| 4.2.3  | Apo-R132K:Y134F                                                | 196 |
| 4.2.4  | Apo-R132K:Y134F:R111L:T54V:L121E                               | 196 |
| 4.2.5  | Apo-R132K:R111L:L121E                                          | 197 |
| 4.2.6  | WT-CRABP II Bound to RA                                        | 200 |
| 4.2.7  | R132K:Y134F Bound to RA                                        | 201 |
| 4.2.8  | R132K:Y134F Bound to Retinal                                   | 202 |
| 4.2.9  | R132K:Y134F:R111L:T54V:L121E Bound to RA                       | 206 |
| 4.2.10 | R132K:R111L:L121E Bound to Retinal                             | 207 |
| 4.3    | Structure Solution and Refinement                              | 209 |
| 4.3.1  | Apo-WT                                                         | 210 |
| 4.3.2  | Apo-F15W                                                       | 217 |
| 4.3.3  | Apo-R132K:Y134F                                                | 217 |
| 4.3.4  | Apo-R132K:Y134F:R111L:T54V:L121E                               | 220 |

|        |                                          |     |
|--------|------------------------------------------|-----|
| 4.3.5  | Apo-R132K:R111L:L121E                    | 220 |
| 4.3.6  | WT-CRABP II Bound to RA                  | 224 |
| 4.3.7  | R132K:Y134F Bound to RA                  | 224 |
| 4.3.8  | R132K:Y134F Bound to Retinal             | 227 |
| 4.3.9  | R132K:Y134F:R111L:T54V:L121E Bound to RA | 229 |
| 4.3.10 | R132K:R111L:L121E Bound to Retinal       | 229 |
| 4.4    | Literature Cited                         | 234 |

## **Appendices**

|            |                                                    |     |
|------------|----------------------------------------------------|-----|
| Appendix 1 | Summary of Protein Properties for CRABP II Mutants | 237 |
| Appendix 2 | Summary of Binding Properties for CRABP II Mutants | 240 |

## **List of Tables**

### **Chapter I Introduction**

|           |                                             |    |
|-----------|---------------------------------------------|----|
| Table 1.1 | Sequence homology between human rhodopsins. | 27 |
|-----------|---------------------------------------------|----|

### **Chapter II Reengineering CRABP II into a Rhodopsin Mimic**

|           |                                                                                                                                                  |     |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 2.1 | The spectroscopic and binding properties of F15W, L19W and apo-WT CRABP II.                                                                      | 84  |
| Table 2-2 | The spectroscopic and binding properties of WT and engineered CRABP II mutants.                                                                  | 88  |
| Table 2-3 | The spectroscopic and binding properties of CRABP II and its mutants toward engineering the R132K:Y134F:R111L:T54V:L121E.                        | 106 |
| Table 2-4 | The spectroscopic and binding properties of Y134F and Tyr134 series. The Y134F series are shown in black and the Tyr134 series are shown in red. | 117 |

### **Chapter IV Mutation, Over-expression, Purification, Crystallization, X-ray Diffraction Analysis, Structure Solution and Refinement of CRABP II and Corresponding Complexes**

|           |                                                                |     |
|-----------|----------------------------------------------------------------|-----|
| Table 4-1 | PCR conditions for mutation of F15W and L19W CRABP II mutants. | 185 |
| Table 4-2 | FPLC protocol. B: 2M NaCl solution.                            | 191 |
| Table 4-3 | X-ray data collection statistics for the apo-structures.       | 199 |
| Table 4-4 | X-ray data collection statistics for the holo-structures.      | 208 |
| Table 4-5 | Refinement statistics for the apo-structures of CRABP II.      | 223 |
| Table 4-6 | Refinement statistics for the holo-structures of CRABP II.     | 233 |

## List of Figures

### Chapter I Introduction

|             |                                                                                                                                                                                                                                                                                                                                                                                                                      |    |
|-------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1-1  | Electron micrograph of bovine rods and cones. Rods outer segment (OS) are seen in the upper part and two cone OS are seen in the lower part of the picture. The mitochondria rich inner segments of the cones are denoted by IS. From R. N. Frank, H. D. Cavanagh, and K.R. Kenyon. Light-stimulated phosphorylation of bovine visual pigments by adenosine triphosphate. <i>J. Biol. Chem.</i> 248, 596-609 (1973). | 4  |
| Figure 1-2  | Schematic representation of a rod and cone photoreceptor cell.                                                                                                                                                                                                                                                                                                                                                       | 6  |
| Figure 1-3  | Rhodopsins span the membrane disks of the photoreceptor cells.                                                                                                                                                                                                                                                                                                                                                       | 8  |
| Figure 1-4  | Crystal structure of bovine rod rhodopsin solved at 2.8 Å (PDB ID: 1F88). Seven transmembrane helices are attached to each other by six loops of varying lengths.                                                                                                                                                                                                                                                    | 10 |
| Figure 1-5  | Cycle of retinal conformational change after light absorption in rhodopsin. The denoted wavelengths (nm) are the maxima of the absorbance spectra.                                                                                                                                                                                                                                                                   | 13 |
| Figure 1-6  | Phototransduction cascade. Rhodopsin absorbs light and becomes activated, which catalyzes exchange of GDP for GTP in Gt molecules. The $\alpha$ subunit of Gt separates and activates the phosphodiesterases, which convert cGMP to GMP. This conversion polarizes the membrane by closing the ion-gated channels, which leads to the neural signal. <sup>18,63</sup> More details are discussed in the text.        | 15 |
| Figure 1-7  | 11- <i>cis</i> -retinal absorbs at ~ 380 nm in methanol. When bound as a SB to <i>n</i> -butylamine, it absorbs at ~ 365 nm. Protonation of the Schiff base leads to a large bathochromic shift to ~ 440 nm.                                                                                                                                                                                                         | 17 |
| Figure 1-8  | Twisting of the single bonds will reduce the degree of p-orbital overlap, and thus will lead to different maximal wavelengths.                                                                                                                                                                                                                                                                                       | 20 |
| Figure 1-9  | Distance of the counter anion to the PSB and positioning of charges or dipoles along the backbone of the polyene may cause spectral tuning of the chromophore.                                                                                                                                                                                                                                                       | 21 |
| Figure 1-10 | Models of retinal binding site in the blue, green and red cone pigments. Negatively charged residues are shown in red and neutral residues in grey.                                                                                                                                                                                                                                                                  | 34 |
| Figure 1-11 | Holo-CRABP II (PDB ID: 1CBS).                                                                                                                                                                                                                                                                                                                                                                                        | 39 |

### Chapter II Reengineering CRABP II into a Rhodopsin Mimic

|            |                                                                                                                                                |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------|
| Figure 2-1 | Retinal in the binding pocket of bovine rod rhodopsin (PDB ID: 1F88), bound to Lys296 <i>via</i> SB. Residues within 6 Å of Lys296 and retinal |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------|

|             |                                                                                                                                                                                                                                                                                                                   |     |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
|             | (violet) are shown. Residues within 4 Å of PSB are labeled. The distances are in angstroms.                                                                                                                                                                                                                       | 65  |
| Figure 2-2  | (A) RA in the pocket of WT CRABP II; (B) Hydrogen bond interactions of the carboxylate group of RA with the residues inside the pocket. The distances are in angstroms.                                                                                                                                           | 67  |
| Figure 2-3  | The dissociation constant of the chromophore bound to the CRABP II mutants is determined by measuring the fluorescence quenching of Trp residues inside the pocket. (A) There are three Trp residues inside the pocket of CRABP II (Trp7, Trp87, Trp109). (B) A typical fluorescence quenching titration curve.   | 73  |
| Figure 2-4  | Retinal bound to a designed Lys shows a peak at the mass of CRABP II plus 266.                                                                                                                                                                                                                                    | 74  |
| Figure 2-5  | Retinal bound to the Lys residue has a peak at the mass of CRABP II plus 268 when SB is reduced.                                                                                                                                                                                                                  | 74  |
| Figure 2-6  | Red shift of retinal absorption upon PSB formation.                                                                                                                                                                                                                                                               | 75  |
| Figure 2-7  | (A) Overlaid structures of R132K:Y134F•RA (brown) and WT CRABP II•RA (yellow). (B) Water-mediated interaction between Arg111 and Ala36, located on the loop connecting $\alpha 2$ to $\beta B$ , in R132K:Y134F•RA. The distances are in angstroms.                                                               | 93  |
| Figure 2-8  | (A) RA and the neighboring residues in the pocket of R132K:Y134F•RA overlaid on WT CRABP II•RA. The distances are in angstroms. (B) RA in the double mutant tilts by $5.23^\circ$ and the carbonyl carbon moves by 0.92 Å. Only residues of the double mutant are labeled in each picture.                        | 96  |
| Figure 2-9  | (A) Overlaid structures of R132K:Y134F•Retinal (blue) and WT CRABP II•RA (yellow); (B) Retinal inside the pocket of R132K:Y134F•Retinal within its Fo-Fc omit electron density map contoured at $2.0 \sigma$ .                                                                                                    | 98  |
| Figure 2-10 | Retinal and neighboring residues to its carbonyl group inside the binding pocket of R132K:Y134F•Retinal (blue) overlaid on R132K:Y134F•RA (brown). Only residues of R132K:Y134F•Retinal are labeled. The distances are in angstroms and show the hydrogen bond distances in R132K:Y134F•Retinal.                  | 100 |
| Figure 2-11 | (A) Overlaid structures of R132K:Y134F•Retinal (blue) and R132K:Y134F•RA (brown). (B) Water-mediated interaction between Arg111 and Ala36, on the loop connecting $\alpha 2$ to $\beta B$ , in R132K:Y134F•Retinal. The distances are in angstroms.                                                               | 102 |
| Figure 2-12 | Binding pocket of R132K:Y134F•Retinal. Leu121 is located midway between Lys132 and Arg111, making it a good choice for placing the counter anion.                                                                                                                                                                 | 105 |
| Figure 2-13 | (A) The structure of R132K:Y134F:R111L:T54V:L121E•RA (magenta) superimposed on the structure of WT CRABP II•RA (yellow). (B) RA and the neighboring residues to its carboxylate group in the pocket of the penta mutant (magenta) and WT CRABP II•RA (yellow). Only the residues of the penta mutant are labeled. | 109 |

|             |                                                                                                                                                                                                                                                                                                                                                                                                                     |     |
|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 2-14 | The binding pocket of R132K:Y134F:R111L:T54V:L121E•RA. RA and Glu121 make tight dicarboxylic acid interactions with each other, indicating that both groups are protonated. Distances are hydrogen bond distances in angstroms.                                                                                                                                                                                     | 110 |
| Figure 2-15 | Binding pocket of the penta mutant bound to RA: A water-mediated interaction connects Glu121 to the residues on the $\alpha 2$ helix region and stabilizes them. The distances are in angstroms.                                                                                                                                                                                                                    | 112 |
| Figure 2-16 | R132K:Y134F:R111L:T54V:L121E•RA (magenta) superimposed on R132K:Y134F•Retinal (blue). Residues of both structures are labeled.                                                                                                                                                                                                                                                                                      | 113 |
| Figure 2-17 | Superimposed structures of WT CRABP II•RA (yellow), R132K:Y134F•Retinal (blue and green) and R132K:Y134F:R111L:T54V:L121E•RA (magenta). The overlaid structures show that if Tyr134 is restored in the pocket of the penta mutant retinal, it will most probably assume a conformation similar to the green conformation of retinal in the double mutant and hydrogen binds to Tyr134 as RA does in WT CRABP II•RA. | 115 |
| Figure 2-18 | UV-vis spectra of retinal bound to WT CRABP II (blue trace, $\lambda_{\max} = 377$ nm), R132K:Y134F:R111L:L121E (magenta trace, $\lambda_{\max} = 378$ , 438 nm) and R132K:R111L:L121E (red trace, $\lambda_{\max} = 449$ nm).                                                                                                                                                                                      | 118 |
| Figure 2-19 | (A) R132K:R111L:L121E•Retinal (green) superimposed on WT CRABP II•RA (yellow). (B) The Fo-Fc omit electron density map of retinal contoured at $2.2 \sigma$ .                                                                                                                                                                                                                                                       | 122 |
| Figure 2-20 | Binding pocket of R132K:R111L:L121E•Retinal. The distances are in angstroms.                                                                                                                                                                                                                                                                                                                                        | 123 |
| Figure 2-21 | R132K:R111L:L121E•Retinal (green) superimposed on R132K:Y134F•Retinal (blue). The labels of the double mutant residues, are shown in parentheses. The water molecules are shown in the same color as each molecule. The distances are in angstroms.                                                                                                                                                                 | 125 |
| Figure 2-22 | Rotation of the free retinal in R132K:Y134F•Retinal (orange) by $45^\circ$ around the single bond that connects the ionone ring to the backbone and along the long axis of retinal positions retinal in a favorable Bürgi-Dunitz trajectory with respect to Lys132. This rotation leads to complete superimposition of free retinal on the bound retinal in the structure of R132K:R111L:L121E•Retinal.             | 126 |
| Figure 2-23 | Proposed mechanism of PSB formation between Lys132 and retinal in the pocket of R132K:R111L:L121E•Retinal.                                                                                                                                                                                                                                                                                                          | 128 |

### CHAPTER III

#### Structural Studies of Human Apo- and Holo-Cellular Retinoic Acid Binding Protein Type II: Opening Doors for New Insights into its Structure/Function Relationship

|            |                                                                                                                                 |
|------------|---------------------------------------------------------------------------------------------------------------------------------|
| Figure 3-1 | The structure of apo-WT CRABP II. The structure is in the P1 space group with two independent molecules in the asymmetric unit. |
|------------|---------------------------------------------------------------------------------------------------------------------------------|



|             |                                                                                                                                                                                                                                      |     |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
|             | Molecule A (Mol A) and molecule B (Mol B) are shown in green and hot pink, respectively.                                                                                                                                             | 139 |
| Figure 3-2  | The superimposed structures of (A) Mol A's of Xtal1 (green) and Xtal2 (purple); and (B) Mol B's of Xtal1 (hot pink), and Xtal2 (orange). Mol A's are identical. Mol B's differ only at the loop connecting $\alpha 2$ to $\beta B$ . | 142 |
| Figure 3-3  | Superimposed structures of Mol A (green) and Mol B (hot pink) of Xtal1.                                                                                                                                                              | 144 |
| Figure 3-4  | (A) Solvent-mediated interaction between Arg111, and Val33 and Ala36, located on $\alpha 2$ and the loop connecting $\alpha 2$ to $\beta B$ in (A) Xtal1 and (B) Xtal2. The distances are in angstroms.                              | 147 |
| Figure 3-5  | Superimposed structures of (A) Mol A's of R111M (red) and apo-WT (green); and (B) Mol B's of R111M (blue) and apo-WT (hot pink).                                                                                                     | 150 |
| Figure 3-6  | 6 Superimposed structures of (A) Mol A's of apo-F15W (brown) and Xtal1 (green); and (B) Mol B's of apo-F15W (cyan) and Xtal1 (hot pink).                                                                                             | 153 |
| Figure 3-7  | The superimposed structures of (A) Mol A's of apo-R132K:Y134F (magenta) and Xtal1 (green); and (B) Mol B's of apo-R132K:Y134F (yellow) and Xtal1 (hot pink).                                                                         | 155 |
| Figure 3-8  | The water-mediated interaction between Arg111 and Ala36 and Val33 in apo-R132K:Y134F. the distances are in angstroms.                                                                                                                | 156 |
| Figure 3-9  | The superimposed structures of (A) Mol A's of apo-R132K:Y134F:R111L:T54V:L121E (pink) and Xtal1 (green); and (B) Mol B's of apo-R132K:Y134F:R111L:T54V:L121E (blue) and Xtal1 (hot pink).                                            | 159 |
| Figure 3-10 | The water-mediated interaction between Glu121 and Val33 and Ala36 in apo-R132K:Y134F:R111L:T54V:L121E The distances are in angstroms.                                                                                                | 160 |
| Figure 3-11 | The superimposed structures of (A) Mol A's of apo-R132K:R111L:L121E (blue) and Xtal1 (green); and (B) Mol B's of apo-R132K:R111L:L121E (orange) and Xtal1 (hot pink).                                                                | 162 |
| Figure 3-12 | The water-mediated interaction between Glu121 and residues on the $\alpha 2$ helix and the loop connecting this helix to $\beta B$ in apo-R132K:R111L:L121E. The distances are in angstroms.                                         | 163 |
| Figure 3-13 | Superimposed structures of Mol A of apo-R111M (red) and WT CRABP II•RA (yellow).                                                                                                                                                     | 166 |
| Figure 3-14 | The superimposed structures of WT CRABP II•RA (yellow) and (A) Mol A of apo-WT (green); and (B) Mol B of apo-WT (hot pink). RA, in the pocket of WT CRABP II•RA, is shown.                                                           | 168 |
| Figure 3-15 | Superimposed structures of Mol A of apo-WT (green), Mol A of R111M (red) and WT CRABP II•RA (yellow). Residues Arg29, Arg59, and Arg132 of each structure and RA in CRABP II•RA are shown.                                           | 171 |
| Figure 3-16 | The water-mediated interaction between Arg111 and Val33 and Ala36 in WT CRABP II•RA. The distances are in angstroms.                                                                                                                 | 172 |

|             |                                                                                                                                                                                                                                                                       |     |
|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 3-17 | Computed electrostatic surface potentials of apo-R111M and holo-CRABP <sub>II</sub> . Basic, acidic, and neutral charges are denoted by blue, red, and white, respectively. A positively charged patch (arrow) is manifested in holo-CRABP-II.                        | 175 |
| Figure 3-18 | (A) Superimposition of NLS of SV40-antigen on the putative NLS residues of CRABP <sub>II</sub> •RA. (B) The same superimposition performed by Sessler et al.                                                                                                          | 177 |
| Figure 3-19 | (A) Lys20, Arg29, and Lys30 in Mol A of Xtal1 (green), CRABP <sub>II</sub> .RA (yellow) and R111M (red); B) Lys20, Arg29, and Lys30 in our CRABP <sub>II</sub> •RA (2FR3, in yellow) and previously published structure of CRABP <sub>II</sub> •RA (1CBS, in purple). | 179 |

## Chapter IV

### **Mutation, Over-expression, Purification, Crystallization, X-ray Diffraction Analysis, Structure Solution and Refinement of CRABP<sub>II</sub> and Corresponding Complexes**

|             |                                                                                                                                                                                                                                                                               |     |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 4-1  | Overview of the QuikChange site-directed mutagenesis method.                                                                                                                                                                                                                  | 184 |
| Figure 4-2  | Schematic drawing of the apparatus used for running NaCl gradient in PhastQ purification of CRABP <sub>II</sub> . Objects not drawn to scale.                                                                                                                                 | 190 |
| Figure 4-3  | FPLC purification SDS-PAGE.                                                                                                                                                                                                                                                   | 192 |
| Figure 4-4  | Hanging drop method for crystallization of proteins.                                                                                                                                                                                                                          | 192 |
| Figure 4-5  | Crystals of apo-WT CRABP <sub>II</sub> . Photographed through a polarizer.                                                                                                                                                                                                    | 194 |
| Figure 4-6  | A crystal of F15W mutant.                                                                                                                                                                                                                                                     | 195 |
| Figure 4-7  | A Crystal of Apo-R132K:Y134F CRABP <sub>II</sub> .                                                                                                                                                                                                                            | 196 |
| Figure 4-8  | Crystals of R132K:Y134F:R111L:T54V:L121E. Streaking of the drop with the seeds of crushed crystals, from the same drop, produced larger and well diffracting crystals. Among the crystals the smaller ones proved to diffract better than the larger ones.                    | 197 |
| Figure 4-9  | Crystals of R132K:R111L:L121E.                                                                                                                                                                                                                                                | 198 |
| Figure 4-10 | Crystals of CRABP <sub>II</sub> •RA. Since RA is light sensitive a red filter was used to take the picture.                                                                                                                                                                   | 201 |
| Figure 4-11 | A crystal of R132K: Y134F bound to RA.                                                                                                                                                                                                                                        | 201 |
| Figure 4-12 | (A) Overlaid spectrum of the solution containing dissolved co-crystals of R132K:Y134 bound to retinal (blue) and the solution of the mutant mixed with 4 equivalents of retinal (red); (B) Overlaid spectra of drops that produced (blue) and did not produce (red) crystals. | 204 |
| Figure 4-13 | Crystals of R132K:Y134F bound to retinal.                                                                                                                                                                                                                                     | 206 |
| Figure 4-14 | Crystals of R132K:Y134F:R111L:T54V:L121E bound to RA.                                                                                                                                                                                                                         | 207 |
| Figure 4-15 | A crystal of R132K:R111L:L121E bound to retinal.                                                                                                                                                                                                                              | 207 |
| Figure 4-16 | Omit electron density map of residues 35 to 41 of Mol B in Xtal1 (hot pink), contoured at 2.2 $\sigma$ , with Xtal2 (orange) superimposed.                                                                                                                                    | 213 |
| Figure 4-17 | Omit electron density map of residues 35 to 41 of Mol B in Xtal2 (orange), contoured at 2.4 $\sigma$ , with Xtal1 (hot pink) superimposed.                                                                                                                                    | 214 |
| Figure 4-18 | Ramachandran plot of the apo-WT CRABP <sub>II</sub> (Xtal1).                                                                                                                                                                                                                  | 215 |
| Figure 4-19 | Ramachandran plot of the apo-WT CRABP <sub>II</sub> (Xtal2).                                                                                                                                                                                                                  | 216 |

|             |                                                                    |     |
|-------------|--------------------------------------------------------------------|-----|
| Figure 4-20 | Ramachandran plot of the apo-F15W CRABPII.                         | 218 |
| Figure 4-21 | Ramachandran plot of the apo-R132K:Y134F CRABPII.                  | 219 |
| Figure 4-22 | Ramachandran plot of the apo-R132K:Y134F:R111L:T54V:L121E CRABPII. | 221 |
| Figure 4-23 | Ramachandran plot of the apo-R132K:R111L:L121E CRABPII.            | 222 |
| Figure 4-24 | Ramachandran plot of the WT CRABPII•RA.                            | 225 |
| Figure 4-25 | Ramachandran plot of R132K:Y134F bound to RA.                      | 226 |
| Figure 4-26 | Ramachandran plot of R132K:Y134F bound to retinal.                 | 228 |
| Figure 4-27 | Ramachandran plot of the R132K:Y134F:R111L:T54V:L121E bound to RA. | 231 |
| Figure 4-28 | Ramachandran plot of the R132K:R111L:L121E bound to retinal.       | 232 |

## KEY TO SYMBOLS OR ABBREVIATION

### Amino Acids

|        |               |
|--------|---------------|
| Ala, A | Alanine       |
| Arg, R | Arginine      |
| Asn, N | Asparagine    |
| Asp, D | Aspartic Acid |
| Cys, C | Cystein       |
| Gln, Q | Glutamine     |
| Glu, E | Glutamic acid |
| Gly, G | Glycine       |
| His, H | Histidine     |
| Leu, L | Leucine       |
| Ile, I | Isoleucine    |
| 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        |

### Symbols and Abbreviation

|                    |                                               |
|--------------------|-----------------------------------------------|
| Å                  | Angstrom                                      |
| Abs.               | Absorbance                                    |
| ADP                | Adenosine diphosphate                         |
| Amp                | Ampicillin                                    |
| Amp <sup>100</sup> | Ampiciline with a concentration of 100 µg/mL  |
| Arr                | Arrestin                                      |
| asu                | Asymmetric unit                               |
| ATP                | Adenosine triphosphate                        |
| bp                 | Base pair                                     |
| CC                 | Correlation Coefficient                       |
| CCP4               | Collaborative computational project, number 4 |
| CD                 | circular dichroism                            |
| cfu                | colony forming units                          |
| cGMP               | Guanisine 3', 5'-cyclic monophosphate         |
| Clm                | Chloramphenicol                               |

|                   |                                                  |
|-------------------|--------------------------------------------------|
| Clm <sup>25</sup> | Chloramphenicol with a concentration of 25 µg/mL |
| CRABPI            | Cellular retinoic acid binding protein type I    |
| CRABPII           | Cellular retinoic acid binding protein type II   |
| CRABPII•RA        | CRABPII in complex with RA                       |
| CRALBP            | Cellular retinal binding protein                 |
| C-terminal        | Carboxy terminal                                 |
| Da                | Dalton                                           |
| dATP              | Deoxyadenosine Triphosphate                      |
| dCTP              | Deoxycytidine Triphosphate                       |
| dGTP              | Deoxyguanosine Triphosphate                      |
| DMSO              | Dimethylsulphoxide                               |
| DNA               | DeoxyriboNucleic Acid                            |
| dNTP              | Deoxynucleotide triphosphate                     |
| dsDNA             | Double stranded DNA                              |
| DTT               | Dithiothreitol                                   |
| dTTP              | Deoxythymidine Triphosphate                      |
| ε                 | Extinction Coefficient                           |
| <i>E. coli</i>    | <i>Escherichi coli</i>                           |
| FABP              | Fatty acid binding protein                       |
| FPLC              | Fast protein liquid chromatography               |
| FTIR              | Fourier transform infra red                      |
| Gβγ               | βγ subunits of Gt                                |
| GDP               | Guanisine diphosphate                            |
| GMP               | Guanisine monophosphate                          |
| GPCR              | G-protein coupled receptor                       |
| Gt                | G protein transducin                             |
| GTα               | α subunit of the Gt                              |
| GTP               | Guanisine triphosphate                           |
| h                 | Hour                                             |
| iLBP              | Intracellular lipid binding protein              |
| IPTG              | Isopropyl-1-thio-β-D-galactopyranoside           |
| IS                | Inner segment                                    |
| L                 | Liter                                            |
| LB                | Luria broth                                      |
| λ <sub>max</sub>  | Maximal waveklength                              |
| M                 | Molar                                            |
| MES               | 2-Morpholinoethanesulfonic acid, monohydrate     |
| MR                | Molecular replacement                            |

|            |                                                           |
|------------|-----------------------------------------------------------|
| MW         | Molecular weight                                          |
| NLS        | Nuclear localization signal                               |
| NMR        | Nuclear magnetic resonance                                |
| N-terminal | Amino terminal                                            |
| OS         | Outer segment                                             |
| PCR        | Polymerase chain reaction                                 |
| PDE        | Phosphodiesterase                                         |
| PEG        | Polyethylene glycol                                       |
| PSB        | Protonated Schiff base                                    |
| RA         | Retinoic acid                                             |
| RAR        | Retinoic acid receptor                                    |
| R          | Rhodopsin                                                 |
| R*         | Activated rhodopsin                                       |
| RARE       | RAR-response element                                      |
| R-factor   | Reliability factor                                        |
| RK         | Rhodopsin kinase                                          |
| RMSD       | Root mean square standard deviation                       |
| RPE        | Retinal pigment epithelium                                |
| R*-p       | Phosphorylated R*                                         |
| RNA        | Ribonucleic acid                                          |
| RXR        | Retinoid X receptor                                       |
| sa         | Simulated annealing                                       |
| SB         | Schiff base                                               |
| SDS        | Sodium dodecyl sulfate                                    |
| SDS-PAGE   | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| TM         | Transmembrane                                             |
| 7TM        | Seven transmembrane                                       |
| UV         | Ultraviolet light                                         |
| vis        | Visible light                                             |
| <i>vdw</i> | <i>van der waals</i>                                      |
| WT         | Wild type                                                 |

Images in this thesis/dissertation are presented in color.

## **Chapter I**

### **Introduction**

#### **1.1 Introduction to Vision and Scope of the Project**

##### **1.1.1 Background**

Vision is a complex process that begins with the conversion of electromagnetic energy (photons or quanta) into neural signals the brain can analyze. Studying the mechanism of vision and color perception in the human eye dates back to the time of antiquity. Ancient Greeks and Arabian philosophers were the first who have recorded studies of human vision. Since their ideas were unscientific and merely based on their personal opinion, it is not appropriate to credit any of them for proposing a better idea. It was not until the 17<sup>th</sup> century that scientific opinions were proposed.<sup>1</sup>

Plato, in the 4<sup>th</sup> century BC, believed that the soul is the source of vision *via* light rays emitted from the eyes. In the same century, Euclid in his book titled "Optics" mentioned that rays emerge from the eye in a cone (called visual cone) with its vertex at the eye and base at the object and developed a geometrical analysis of the vision problem. Aristotle, one of the most prominent philosophers of ancient Greece in the 4<sup>th</sup> century BC, was among the first who refuted the vision theory of light rays emitted from the eyes. Aristotle's theory of vision stated that the object altered the medium between the object itself and the eye and these changes in the medium were transferred into the eye. Many other philosophers and scientists contributed to the understating of vision problem but until the sixteenth century it was thought that the lens was the receptor of light. It was not until 1604 that Kepler, the chief German founder of modern astronomy, explained the

optics of the eye and proposed the first theory of the retinal image. He could show that the lens was simply an optical component in the eye, which participated in imaging the outside world on the retina.<sup>5</sup>

In 1704 Isaac Newton introduced the scientific thinking of vision and showed that color is not the property of the object but is a property of the eye itself. This was the foundation of modern theories of color vision, however he did not speculate on the basis of color perception in the eye.<sup>1,11</sup> In 1802 Young suggested the three-fold characteristic of color vision and speculated that there are three types of color receptors (red, blue and green) in the eye that make color vision possible, and all other colors are seen by a combination of these primary colors. However his finding was prior to the discovery of cone cells, as photoreceptor cells, by Schultze in 1866.<sup>11</sup> It was Heinrich Müller, who first described a reddish colorant in fresh retinas in 1851 (visual purple, which later was called rhodopsin). The color quickly faded upon peeling of retina from the eyeball. The bleaching effect was attributed to a postmortem effect. It was not until 1876 that Franz Boll discovered that it was the light that bleaches the visual purple. A year later, Willy Kühne for the first time could extract rhodopsin (as he called it) from the outer segment of the bovine rod cells by using a detergent solution. In the 1930's Wald and his coworkers were the first to discover that rhodopsin is a protein and consists of two parts: a protein called opsin and an organic light-absorbing molecule, retinal.<sup>12,13</sup> In the 1950's the same group used chemical methods to separate visual pigments from the retina and measured the wavelength absorption of different pigments in the retina using spectrophotometry experiments. Those experiments verified the Young theory on color perception.<sup>1,5,14,15</sup> In 1958 Hubbard and Kropf discovered that upon absorption of a



photon of light 11-*cis*-retinal photoisomerizes to all-*trans*.<sup>16</sup> Later in 1968 Wald demonstrated that this photoisomerization is the primary event in the phototransduction cascade. George Wald (1906-1997) pioneered our understanding of the molecules responsible for the first steps in the vision process, which led him to win the Nobel Prize in Medicine and Physiology in 1968.<sup>6,17</sup>

However, given the key role of the photoreceptor cells in vision, it was not until the 1980's that Lubert Stryer and Denis Baylor, both at Stanford, discovered how these cells operate.<sup>18-20</sup> Improved methods for making electrical recording from individual photoreceptor cells provided detailed information about the mechanism by which light energy is transduced into neural signals.<sup>5</sup> The day Jeremy Nathans heard of these new discoveries, he became interested in how we see in color. He has spent more than 20 years incorporating molecular genetic approaches to the study of the physiology and development of the retina and to understanding the mechanisms of human retinal diseases. As a result, he has explored the basis of the molecular genetics of the cone pigments and some important aspects of the evolution of human color vision.<sup>5</sup>

### **1.1.2 Rod and Cone Photoreceptor Cells**

Due to extensive research on the vision process our understanding of vision has come a long way. Now we know that although all the parts in the eye are important in perceiving an image on the retina, which is the most inner layer of the eye, it plays a very crucial role in vision process. There are two types of photoreceptor cells in the retina that are called rod and cone cells. These cells are located at the most outer layer of the retina. As shown in Figure 1-1, an electron micrograph of the photoreceptor cells in retina, the

terms rod and cone are derived from the shape of these photoreceptor cells.<sup>2,3</sup> These cells are specialized neural cells that detect the light entering the eye and convert it into an electrical signal that is sent to other neurons of the retina (ganglion cells), which in turn send the signal to the visual section of the brain for further processing.



Figure 1-1 Electron micrograph of bovine rods and cones. Rods outer segment (OS) are seen in the upper part and two cone OS are seen in the lower part of the picture. The mitochondria rich inner segments of the cones are denoted by IS. From R. N. Frank, H. D. Cavanagh, and K.R. Kenyon. Light-stimulated phosphorylation of bovine visual pigments by adenosine triphosphate. *J. Biol. Chem.* 248, 596-609 (1973).<sup>2,3</sup>

Rod cells are responsible for dim light (scotopic) vision but are so sensitive to light that they become overloaded and incapable of signaling in ordinary daylight.<sup>2</sup> Although these cells make it possible to form black-and-white images in dim light, they do not mediate color vision and have a low visual acuity. On the other hand daylight

vision is mediated by the cone cells, which operate successfully at high levels of light (photopic vision). Cones are responsible for high visual acuity (high resolution) and make color vision possible. Although all the vertebrates have rod cells, some do not have any cone cells and therefore cannot see colors.<sup>2,11</sup>

The rod cells are more numerous than the cone cells. The human eye has approximately 120 millions rod and 6 to 7 millions cone cells. Although we have more rods than cones, most of the time we use cone cells because they allow for fine discrimination of the colors and visual acuity. Cones are concentrated in a point of the retina called fovea or “yellow spot”, which is located right across from the pupil of the eye and is 300-700  $\mu\text{m}$  in diameter.<sup>11</sup> Therefore in order for us to see the details, for example while reading a book, our eye constantly moves to focus the image on the small surface of the fovea. This movement takes place so fast that we do not realize we are focused on one point at a time. A schematic drawing of a rod and a cone cell is shown in Figure 1-2.

Each cell is divided into two parts: the outer segment (OS) and the inner segment (IS). The OS contains most of the molecular apparatus for light absorption and generation of an electrical (neural) signal. It is comprised of a membrane that is folded into a densely packed stack of tiny disks, which contain the light receptor protein, rhodopsin.<sup>21</sup> These highly packed disks ensure a dense packing of light-absorbing rhodopsins, and therefore a high probability of a single photon absorption.<sup>20</sup> The structure and function of the rhodopsin system will be discussed in the next section.

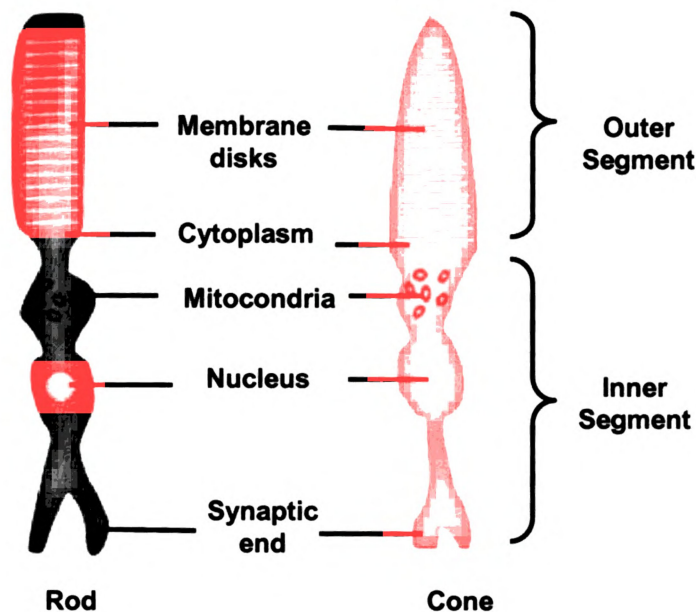


Figure 1-2 Schematic representation of a rod and cone photoreceptor cell.

Rod cells contain only one kind of rhodopsin molecule, rod rhodopsin, which absorbs at  $\sim 500$  nm. Cones, on the other hand, consist of three different types of cells, each "tuned" to maximally absorb at a distinct wavelength. These cells absorb maximally at the short, medium and long wavelength region of visible light and are therefore called blue, green, and red cones, respectively. These different absorbing properties are due to the presence of different photoreceptor cells, red, green, blue opsins that maximally absorb at  $\sim 560$ ,  $530$  and  $410$  nm, respectively. It is the combination of the absorbances of these three different cone cells that allows for color differentiation. Among mammals only primates possess trichromatic color vision ability.<sup>22-24</sup>

The inner segment of the photoreceptor cells contains the nucleus and other cell organelles, such as mitochondria and golgi bodies, that are necessary for the correct cell function. The mitochondria are thought to supply the energy necessary for processes associated with transduction and for the synthesis of visual pigments. In addition to these

components the inner segment is comprised of a synaptic terminus, which sends the electrical signal generated in the outer segment of the photoreceptor cells, to the neurons in the inner retina. These signals are transmitted through an elaborate array of synapses, or neural junctions in the retina and toward the visual section of the brain.<sup>19,25-27</sup>

### 1.1.3 Rhodopsin System

In the 1930's George Wald discovered that rhodopsin consists of two parts: an apo-protein called opsin, bound to a chromophore, 11-*cis*-retinal.<sup>12,13</sup> In 1958 Hubbard and Kropf discovered that 11-*cis*-retinal, an aldehyde derivative of vitamin A, absorbed a photon of light and isomerized to all-*trans*.<sup>16</sup> In 1967 Wald was the first to demonstrate that this photoisomerization is the primary event in visual excitation and triggers the vision transduction process.<sup>2,6,14,15,17,18,28</sup>

Rhodopsins are transmembrane proteins that span the membrane disks of the OS of the photoreceptor cells (Figure 1-3). Over 90% of the proteins in the disk membrane of rod cells are rhodopsins, which are packed with an average packing density of ~ 25,000 rhodopsin/ $\mu\text{m}^2$ .<sup>29</sup> Recent high resolution infrared-laser atomic-force microscopy (AFM) studies on mouse disk membranes revealed that rhodopsins form unevenly and dense packed rows of dimers in their native form. Still it is not clear what impact this high packing density of rhodopsins has on the structure and function of rhodopsins. For example, although this high density may improve the light sensitivity, it may also cause a slow visual response, which is limited by the diffusion controlled protein-coupling in the visual cascade.<sup>30-36</sup>

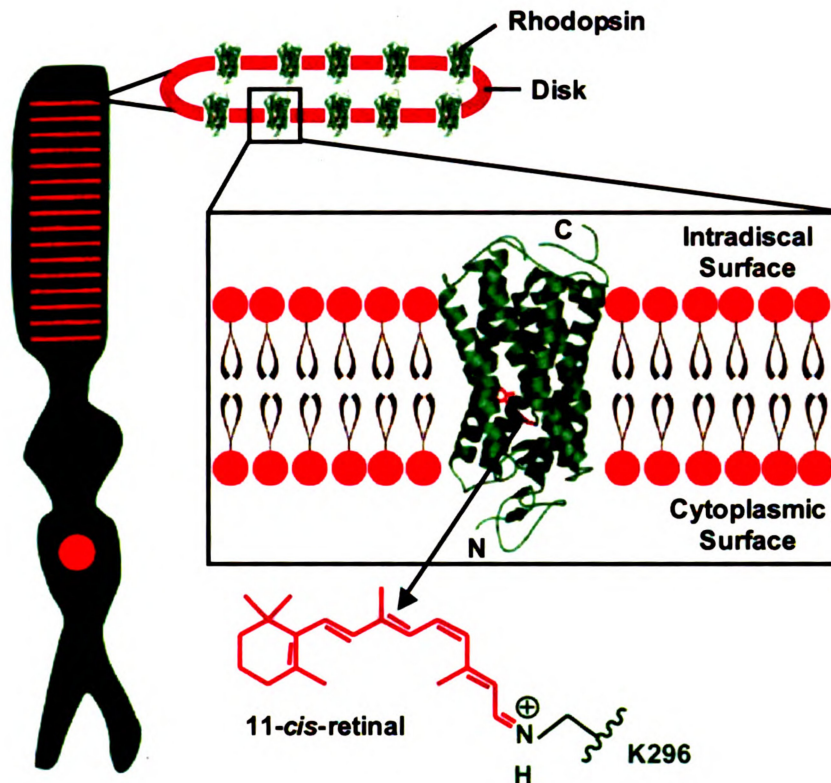


Figure 1-3 Rhodopsins span the membrane disks of the photoreceptor cells.

Since rods are more abundant than cones, most of our knowledge about rhodopsins arises from studies done on the structures and biological functions of rod rhodopsins in human and animals such as bovine, mouse and chicken. Rhodopsins belong to the heterotrimeric guanine nucleotide-binding protein (G protein)-coupled receptor (GPCR) superfamily of proteins. Members of this family are activated by different external stimuli, such as  $\text{Ca}^{2+}$ , amines, hormones, neurotransmitters, odors and light, and in turn activate G proteins. Activation of a G protein initiates an enzymatic cascade that ultimately produces a neural signal.<sup>37,38</sup> Rhodopsins are a distinct group within GPCRs: while all GPCRs bind their agonists through noncovalent interactions, rhodopsins have a covalently bound ligand, 11-*cis*-retinal.<sup>39</sup>

All GPCRs are composed of seven transmembrane (7TM) helices that are connected by six hydrophobic loops of varying lengths. Among GPCRs, bovine rhodopsin was the first member to be purified, its gene was the first to be cloned and sequenced, and its three-dimensional structure was the first determined.<sup>23</sup> The bovine eye has only rod rhodopsins and therefore has no color vision. Rhodopsins can be readily isolated from retina of a number of species, however bovine eyes are readily available from meat packaging plants and are a major source of rhodopsin.<sup>4</sup> Efforts in studying and determining the structures of opsins have been hampered by the difficulties of working with membrane proteins. The only crystal structure of a visual rhodopsin (and a GPCR) is bovine rod rhodopsin, which was determined by Palczewski et al. in 2000 (PDB ID: 1F88)<sup>40</sup> Bovine rhodopsin is a middle size protein among the GPCR family (348 residues and ~ 40 KDa molecular weight), and therefore this structure may be used as a good representative of this superfamily of proteins.<sup>15,23,41-45</sup> The structure is in a P41 space group and has two molecules in the asymmetric unit. A ribbon diagram of chain A, prepared from the deposited coordinates in the Protein Data Bank (PDB ID: 1F88), is shown in Figure 1-4.

The amino-terminus (N) of the protein is located at the extracellular surface and the carboxy-terminus (C) is at the cytoplasm (intracellular). The helices are linked together by 3 intracellular and 3 extracellular loops. The 7TM helical bundle forms a barrel around the chromophore, 11-*cis*-retinal. In both rod and cone cells the chromophore binds covalently to the protein *via* a protonated Schiff base (PSB) with the  $\epsilon$ -amino group of a lysine residue on the seventh helix.<sup>46</sup> In bovine rod rhodopsin Lys296, which is located more toward the extracellular surface of the membrane bilayer,





Figure 1-4 Crystal structure of bovine rod rhodopsin solved at 2.8 Å (PDB ID: 1F88). Seven transmembrane helices are attached to each other by six loops of varying lengths.

forms the PSB linkage with the chromophore.<sup>47,48</sup> The chromophore is positioned parallel to the surface of the membrane and therefore the direction of the light entering the eye is perpendicular to the long axis of the chromophore.<sup>49</sup> This structure makes it possible to have the maximal light absorption and therefore the maximum light sensitivity.<sup>45,50</sup> The retinal binding site is completely buried within the protein and is not accessible to the cytoplasmic or extracellular surfaces. The chromophore interacts with neighboring helices, and is tightly held in place. The structure shows no gap between the helices to let retinal exit the binding site. However, like other ligands, retinal must move in and out of the binding pocket to regenerate the photopigment, therefore it is believed that alterations in the published structure must occur in the functional protein. X-ray data is collected over a long period of time (hours) and is an average of the diffraction of many protein molecules in the crystal. Therefore the crystal structure shows a temporal and spatial average of all the protein molecules in the crystal. The published structure shows an inactive form of rhodopsin, however thermal fluctuations and/or isomerization of retinal might cause structural changes in the protein to allow the passage of the ligand



in and out of the pocket. Although this structure is a ground state it provides valuable insights into the dynamic activation and signaling process of the receptor.<sup>51</sup>

The first step in activation of rhodopsin is absorption of a photon of light by 11-*cis*-retinal, covalently bound to Lys296 *via* forming a PSB, and its photoisomerization to all-*trans*-retinal.<sup>6,17</sup> The quantum yield of the photoisomerization reaction depends on the wavelength of the incident light and decreases going from short wavelengths toward long wavelengths of visible light.<sup>52</sup> The 500 nm light has a high quantum yield of ~ 0.67, which decreases by ~ 5% in 570 nm.<sup>52,53</sup> Therefore two thirds of the 55 Kcal/mol energy of the 500 nm photon (~ 37 Kcal/mol) is stored in the all-*trans*-retinal-opsin complex, which is released later to give rise to the signaling form of the protein, metha-rhodopsin II.<sup>54</sup> Photoisomerized retinal triggers a series of conformational changes in the opsin, which creates an enzymatic site on the cytoplasmic surface of the protein.<sup>55,56</sup> The enzymatically active rhodopsin, although short-lived catalyzes the conversion of several hundred second messengers from an inert to an active state and therefore allows detection of a single photon of visible light. It is this high amplification of a single photon that enables us to see at even very dark environments.<sup>57</sup>

### Conformational Changes of Rhodopsin upon Light Absorption

Within a few picoseconds of absorption of a photon by rhodopsin, the chromophore, 11-*cis*-retinal goes through a series of conformational changes described in Figure 1-5.<sup>6-10</sup> The isomerization can be monitored by various low temperature spectroscopic techniques including UV-vis.<sup>54,58-61</sup> The first spectroscopically characterized species that forms within femtoseconds is called photorhodopsin, which is a

highly distorted *all-trans* PSB. Within picoseconds this species is converted to the first isolable, less distorted intermediate, batho-rhodopsin. In approximately 1 millisecond, bathorhodopsin goes through several additional intermediates to form meta-rhodopsin II (meta II). In meta II, the Schiff base (SB) is deprotonated and the Glu113, which is the counter anion of the PSB is protonated. The conformational changes of the chromophore resulting in the meta II conformation cause a re-packing of the transmembrane helices, that is relayed to changes on the cytoplasmic surface of the cell.<sup>43,55</sup> This allows for the G protein transducin (Gt) to interact with rhodopsin resulting in activation of the phototransduction cascade. At the end of the cycle, as shown in Figure 1-5, the SB imine bond of meta II is protonated (meta III) and hydrolyzed to generate *all-trans*-retinal, which separates from the opsin and travels back to the retinal pigment epithelial cells, where it is converted back to 11-*cis*-retinal *via* several enzymatic steps. At this time the regenerated opsin binds to another 11-*cis*-retinal and the cycle repeats.<sup>43,62</sup> The changes in retinal/opsin interactions accompanying visual transduction and the rhodopsin intermediates have also been elegantly studied spectroscopically using photoaffinity labeling with a rhodopsin analog (diazoketo-rhodopsin).<sup>56</sup>

It is not clear how 11-*cis*-retinal gains access to the binding site and the crystal structure of bovine rhodopsin does not provide any evidence to this mechanism. The binding site is completely buried and inaccessible from outside the protein. Based on accessible surface calculations it appears that movement of several of the helices would permit access to the site from the hydrophobic core of the bilayer.<sup>4</sup>

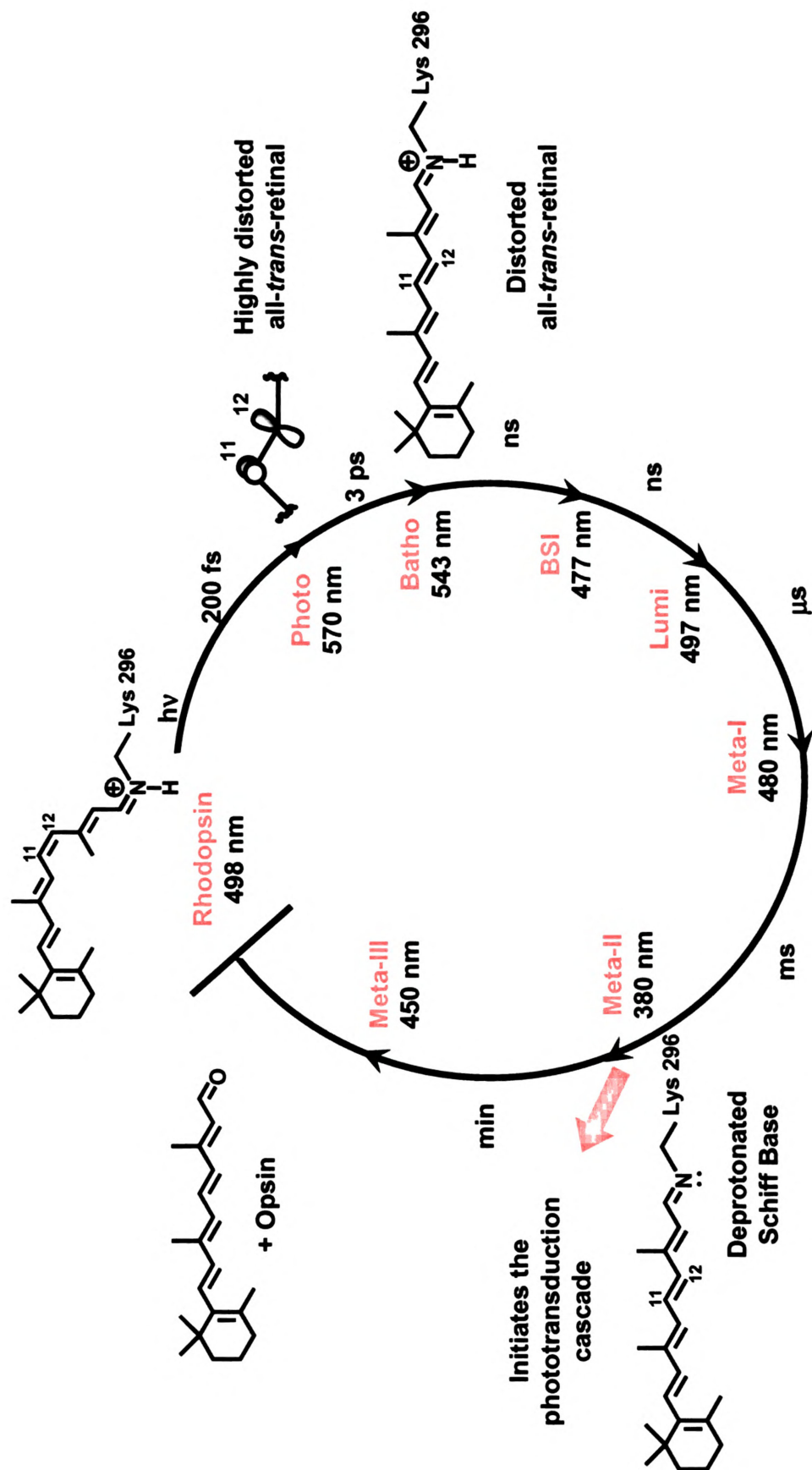


Figure 1-5 Cycle of retinal conformational change after light absorption in rhodopsin. The denoted wavelengths (nm) are the maxima of the absorbance spectra.<sup>6-10</sup>

### The Phototransduction Cascade

As mentioned above upon absorption of a photon of light rhodopsin (R) becomes activated (R\*). Formation of the active form of rhodopsin (meta II) causes structural changes on the 7TM helices of the opsin, which result in a biochemical cascade that involves a G protein, transducin, phosphodiesterase (PDE), and guanine 3', 5'-cyclic monophosphate (cGMP).<sup>62</sup> The biochemical cascade of phototransduction is illustrated in Figure 1-6. R\* binds to GDP-bound transducin (Gt), resulting in an exchange of GDP for GTP. The  $\alpha$  subunit of the Gt (Gt $\alpha$ -GTP) then dissociates from the  $\beta\gamma$  subunits (G $\beta\gamma$ ). Gt $\alpha$ -GTP binds and switches on a cGMP PDE by binding to and removing an inhibitory subunit ( $\gamma$  subunit). The activated PDE converts cGMP to GMP. Reduction in the level of cGMP makes cGMP incapable of keeping the cGMP-gated ion channels open, resulting in hyperpolarization of the membrane and therefore generation of electrical signals that form the basis of vision.<sup>15,18,43</sup> R\* separates from the inactive part of Gt and recycles to activate more transducins.

The binding of Gt-GDP to R\*, exchanging GDP for GTP, and separation of Gt $\alpha$ -GTP from the G $\beta\gamma$  subunits occurs multiple times prior to hydrolysis of the imine bond between all-*trans*-retinal and a lysine of the opsin. Since hydrolysis happens very slowly on the time scale of phototransduction, it gives an R\* molecule enough time to activate on the order of  $10^2$  PDEs, which each can convert on the order of  $10^3$  cGMPs to GMPs. Therefore one R\* is capable of converting  $10^5$  cGMPs to GMPs.<sup>62</sup>

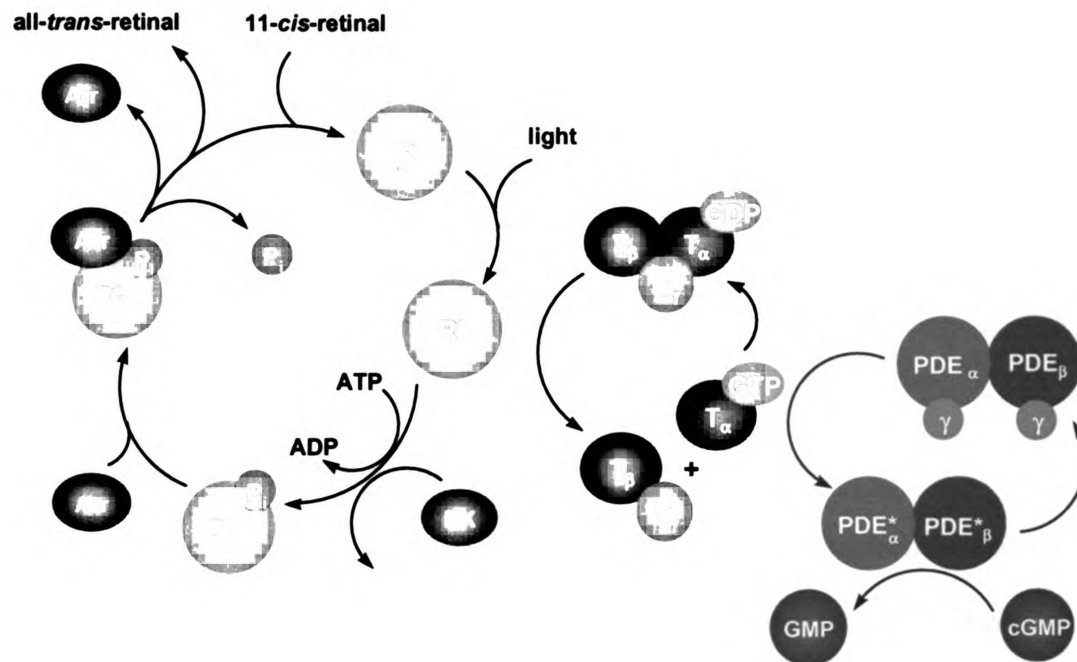


Figure 1-6 Phototransduction cascade. Rhodopsin absorbs light and becomes activated, which catalyzes exchange of GDP for GTP in Gt molecules. The  $\alpha$  subunit of Gt separates and activates the phosphodiesterases, which convert cGMP to GMP. This conversion polarizes the membrane by closing the ion-gated channels, which leads to the neural signal.<sup>18,63</sup> More details are discussed in the text.

During hydrolysis rhodopsin kinase (RK) binds to R\* and phosphorylates several residues in the C-terminal region of the protein. Arrestin (Arr) binds the phosphorylated R\* (R\*-p) to completely shut off the process. At the end of the cycle, *all-trans-retinal* dissociates from the opsin and travels back to the retinal pigment epithelium (RPE) for regeneration to *11-cis-retinal*. The opsin accepts a new *11-cis-retinal*, which was regenerated in RPE and transported to the outer segment of the photoreceptor cells. The regenerated rhodopsin reinitiates the phototransduction cascade.

Rhodopsin is within the disk membranes of the outer segment of the photoreceptor cells, however it is free to diffuse in the membrane. Transducin is at the surface of the disk membrane but when it is activated by R\*, Gtα-GTP can diffuse

toward other neighboring disks and activate many PDEs. PDEs are either membrane bound or free in the cytoplasm. cGMP is within the cytoplasm.<sup>62</sup> When cGMP is hydrolyzed to GMP, the concentration of cGMP in the cytoplasm decreases and therefore more cGMPs are drawn from the plasma membrane toward these sinks, causing the ion-gated channels to close and the membrane to become polarized.<sup>18,43,62-64</sup>

#### 1.1.4 Color Vision with the Same Chromophore

As mentioned earlier all different opsins bind to the same unique chromophore of vision, 11-*cis*-retinal, however they absorb light at different wavelengths. Therefore a long-standing question in vision research has been to understand the mechanisms by which different opsins regulate the wavelength absorption of the same chromophore, and later relate those understandings to the visual ability and evolutionary history of each species.<sup>65</sup>

Free all-*trans* and 11-*cis*-retinal both absorb at ~ 380 nm in organic solvents such as methanol (Figure 1-7). When the SB is formed with *n*-butylamine (to model rhodopsin), it absorbs at ~ 365 nm. However if the SB becomes protonated (PSB) the chromophore will absorb at a significantly higher wavelength, ~ 440 nm (Figure 1-7).<sup>66</sup> With changing the solvent and its concentration this value may shift to 500 nm.<sup>67</sup> Although formation of the PSB can explain the absorption of rod and blue pigments, the ~ 60 nm red shift upon protonation of the SB is not sufficient to explain the red shifted absorption of the green and red rhodopsins.<sup>68</sup> The interaction of the residues within the binding pocket of the opsins with the chromophore must be responsible for the spectral tuning of the chromophore that make it absorb between 400 and 600 nm in different

rhodopsins. These interactions modulate the ground-excited ( $S_0$ - $S_1$ ) electronic transition energy of the retinal PSB. The difference between the  $\lambda_{\max}$  of the pigment from that of the PSB model compound is called “opsin shift”,<sup>69</sup> a term also used to refer to the differences in the  $\lambda_{\max}$  values among pigments.<sup>70</sup> It is the unique interaction of each opsin with the chromophore and coexistence of all three colored rhodopsins in the eye that enable us see the entire visible spectrum.

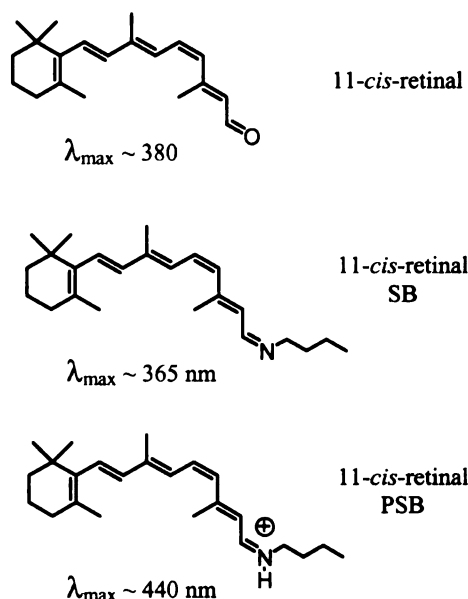


Figure 1-7 11-*cis*-retinal absorbs at  $\sim 380$  nm in methanol. When bound as a SB to *n*-butylamine, it absorbs at  $\sim 365$  nm. Protonation of the Schiff base leads to a large bathochromic shift to  $\sim 440$  nm.

Although opsin shifts have been known for a long time,<sup>71</sup> their physical origin is still a long-standing question, which is addressed in solution experiments,<sup>72</sup> mutagenesis studies,<sup>73</sup> and also high-level quantum-chemistry calculations.<sup>74-76</sup> For a long time it was believed that the interactions of the residues inside the pocket of the opsin with retinal red shift the absorption maximum of the chromophore compared to that in solution.<sup>6,17,77</sup> Although in 1974, Suzuki et al. based on theoretical models had predicted that an isolated PSB retinal absorbs at  $\sim 600$  nm,<sup>78</sup> it was not until very recently that its actual value was measured. In 2005, Andersen et al.<sup>79</sup> measured the absorption spectrum of all-*trans*-

retinal with a protonated *n*-butylamine Schiff base, as a model of PSB in rhodopsin, in vacuo and found a maximum at ~ 610 nm. The new gas-phase absorption data provides a closer model to the actual absorption of the bare chromophore, when it is free from intermolecular interactions, such as hydrogen bonding, dispersion and coulomb interactions with charged and polar groups in solution. So according to this new finding it is believed that the binding pockets of the opsins actually blue shift, and not red shift, the absorption of the bare chromophore.<sup>79</sup>

The either red or blue shifting of the maximum absorption of the chromophore inside the pocket of the opsins can be verified by studying the interactions of the residues inside the binding pocket of each of the opsins with the chromophore. However, rhodopsins are membrane-bound proteins and therefore are very difficult to manipulate and study. Particularly crystallization of membrane proteins has been a bottleneck in studying the structures of these proteins. The first and only crystal structure of a rhodopsin molecule, and in fact a GPCR protein, is of bovine rod rhodopsin. This structure was determined at 2.8 Å by Palczewski et al. in 2000 (PDB ID: 1F88).<sup>40</sup> A refined model added some residues missing from the original structure (PDB ID: 1HZX).<sup>80</sup> The resolution was extended to 2.6 Å in 2002 (PDB ID: 1L9H);<sup>81</sup> and finally in 2004 Okada et al. crystallized this protein in a new condition and refined the structure to 2.2 Å (PDB ID: 1U19).<sup>82</sup> The new structure is in general agreement with the previous data and provides more details on the chromophore binding site including the configuration about the C6-C7 single bond of 11-*cis*-retinal. The map around the C6-C7 bond, which defines the orientation of the β-ionone ring, becomes clearer at 2.2 Å, supporting a 6-*cis* form with substantial negative twist. Most importantly, the C11-C12



double bond is found to be significantly pre-twisted in the ground state, which explains the isomerization around this double bond upon absorption of a photon.<sup>82</sup> As of today, there is no crystal structure of a cone rhodopsin available. Therefore the only structural information on cone rhodopsins are obtained through indirect methods such as photoaffinity NMR labeling, cryoelectron microscopy, site directed mutagenesis and advanced theoretical calculations.<sup>83-89</sup>

### Proposed Hypotheses for the Mechanism of Wavelength Regulation

Although there is no structure of a cone rhodopsin available, over the past 50 years several hypotheses have been proposed to explain the wavelength regulation by the same chromophore. All of these hypotheses are based on the basic fact that photoexcitation of the chromophore increases the delocalization of the  $\pi$ -electrons.<sup>90,91</sup> Since retinal is bound as a SB to the opsin, in addition to normal  $\pi \rightarrow \pi^*$  electronic transitions of conjugated double bonds of the polyene, the lone pair of the nitrogen, which occupies an n-orbital can also be excited and make a second type of transition,  $n \rightarrow \pi^*$ , possible. Protonation of the SB will change the energy gap of the  $\pi \rightarrow \pi^*$  transitions and consequently the absorption spectrum of the chromophore.<sup>92</sup> Also any interactions between the residues inside the pocket of the opsin and 11-*cis*-retinal that favor delocalization, selectively stabilize the excited state ( $S_1$ ) and cause a red shift, i.e., a smaller energy gap between the ground and excited states of the chromophore. On the other hand, any interaction that does not favor delocalization leads to a larger energy difference between the ground and excited state causing a blue shift.

Based on empirical and theoretical studies several major hypotheses have been proposed to explain the mechanism by which opsin shift may occur: (1) Steric factors may lead to twisting around the intervening single bonds of the chromophore altering the p-orbital overlap, which in turn changes the level of delocalization of the  $\pi$ -electrons, or in other words changes the degree of cationic conjugation along the polyene. A change in the level of conjugation will cause a shift in the absorption of the retinal. Binding pockets of different rhodopsins may impose different steric constraints on the chromophore and therefore cause different degrees of twisting about the single bonds. More twisting and less orbital overlap lead to a blue shift and vice versa (Figure 1-8).<sup>93,94</sup>

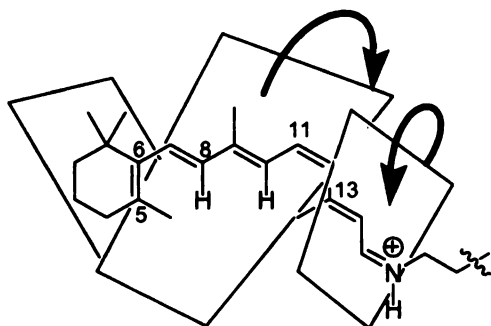


Figure 1-8 Twisting of the single bonds will reduce the degree of p-orbital overlap, and thus will lead to different maximal wavelengths.

Although retinal would rather adopt a planar conformation to maximize the orbital overlap, it has been shown that 11-*cis*-retinal is not a planar molecule. The chromophore has a 6-*s-cis*, 12-*s-trans* conformation, where steric interactions between the C5-Me and C8-H as well as the C10-H and C13-Me force twisting of the C6-C7 and C12-C13 single bonds, respectively.<sup>95,96</sup> (2) Increasing the distance between the PSB and its counter anion may favor delocalization of the  $\pi$ -electrons, which leads to a red-shift. Conversely, decreasing the distance may localize the positive charge of the PSB on the nitrogen and therefore cause a blue shift (Figure 1-9).<sup>97</sup> (3) Point charge model:

Placement of full or partial charges along the polyene can change the degree of delocalization or cationic conjugation along the polyene. For example positioning of one or more anions along the retinal chain, presumably Glu or Asp, leads to increased delocalization (Figure 1-9).<sup>90,98,99</sup> (4) An alternation of the polarity or polarizability of

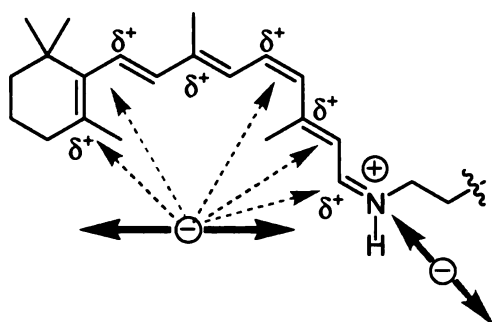


Figure 1-9 Distance of the counter anion to the PSB and positioning of charges or dipoles along the backbone of the polyene may cause spectral tuning of the chromophore.

the binding pocket environment around the chromophore or in other words positioning of the polarizable groups along the chromophore can change the delocalization of the electrons and shift the wavelength. The polar groups can favor or disfavor delocalization depending upon their orientation.<sup>100</sup> On the other hand polarizable groups along the polyene can stabilize the excited state by compensatory electronic movement.<sup>71,101</sup> The last three hypotheses focus on the effects of stereoelectronic factors on shortening or elongating of polyene's cationic conjugation.

In contrast to the variable maximum absorption of the chromophore along the wavelength axis, the shape and width of the rhodopsin absorption curve is determined by chromophore vibration and therefore is not modified by the opsin.<sup>65</sup>

Fully understanding the mechanism of spectral tuning in rhodopsins and verifying the proposed hypotheses are still pending upon determination of the crystal structure of each color rhodopsin, which has been hampered by the challenges of working with membrane proteins and the lack of material.

## Use of Biophysical Assays and Model Studies to Understand the Rhodopsin System

Although the crystal structures of the rhodopsin pigments are not available, the binding pockets of different rhodopsins and retinal chromophore in those pockets have been extensively studied using site-directed mutagenesis and biophysical methods. Magic angle spinning (MAS) NMR techniques have increasingly provided more powerful tools for structure-function analysis of membrane proteins over the past 15 years. These techniques are well-suited to provide information on helix-retinal and helix-helix interactions that may be involved in receptor activation.<sup>102,103</sup> Unfortunately NMR spectroscopy failed to give complete structures of rhodopsins due to the molecular-weight limit of this technique. However, cryo electron microscopy has been used to determine the moderate to low-resolution structures of rhodopsins.<sup>85,87,88,104</sup> Resonance Raman spectroscopy of recombinant human cone pigments has been successfully used to verify the proposed hypotheses of wavelength regulation by positioning of the point charges or altering the polarity around the chromophore.<sup>105</sup> Also Raman spectroscopy has been used to study the *in situ* conformational changes of retinal and its interactions within the opsin binding pocket.<sup>106</sup> Fourier-transform infrared diffraction spectroscopy at low temperature (70 K) has been used to study the vibrational modes and therefore conformational changes of the chromophore and opsin during the phototransduction cycle.<sup>107</sup>

The combination of a wide range of existing experimental data with theoretical calculations has taken our knowledge of vision a long way. In particular, recombinant site-directed mutagenesis of rhodopsins along with the modeling of the binding pockets of different rhodopsins have elicited many of the key structural elements of rhodopsins

and has led to a better understanding of the opsin shift as well as the key residues involved in the wavelength regulation.<sup>92</sup>

The evidence that retinal is bound as a SB via a Lys residue was provided by Bownds in 1957,<sup>47</sup> and Akhtar et al. in 1968.<sup>108</sup> Although in 1958 Hubbard had suggested that the SB must be protonated, it was not until 1974 that Callender and his group confirmed PSB formation through Raman spectroscopy.<sup>106,109</sup> In 1958 Kropf and Hubbard proposed that the wavelength regulation is a result of the interactions between the charged residues inside the binding pocket of the opsins with the chromophore.<sup>90</sup> Following the same principles, in 1972 Blatz et al. proposed the presence of a counter anion to stabilize the positive charge on the PSB.<sup>97</sup> In support of the counter anion idea in 1979 Honig et al.<sup>110</sup> proposed the effect of suitably placed negatively charged residues as a counter anion to stabilize the positive charge on the PSB. Also in agreement with Kropf's and Hubbard's theory, they suggested that positioning of the point charges along the polyene may change the absorption spectrum of the chromophore. They confirmed these ideas by experimental evidence such as measuring the absorption spectra of simple synthetic models.<sup>110,111</sup> Later in the 1980's Mollevanger et al. validated these theories by solid-state NMR studies of bovine rhodopsin.<sup>112</sup> Although it was known that a counter anion is necessary, it was not until 1989 that Sakmar and Oprian and their groups independently showed that Glu113 on the third TM helix, a highly conserved residue in vertebrate opsins, is the counter anion.<sup>73,113,114</sup> Sakmar's group used site-directed mutagenesis to unveil the importance of Glu113 as the counter anion of the PSB and its crucial role in determining the  $pK_a$  and  $\lambda_{max}$  of rhodopsin. Different individual mutants of rhodopsins were prepared in which Glu113 was replaced by different hydrophobic and

hydrophilic residues. It was shown that mutation of Glu113 to neutral residues, or positioning of the counter anion more distant to the PSB dramatically decreased the  $pK_a$  of the PSB in the mutants. The  $pK_a$  of the PSB in rhodopsin is determined to be above 16,<sup>115</sup> while the  $pK_a$  of the mutants were  $\sim 6$ .<sup>114</sup> It is believed that the high  $pK_a$  of the PSB is due to the specific geometrical arrangement of the SB and its counter anion, which allows an efficient bridging of a water molecule to stabilize the ion pair.<sup>116,117</sup>

On the other hand the point mutations modulated the  $\lambda_{max}$  of rhodopsin the same way as predicted by the proposed hypotheses of wavelength regulation. The mutation of Glu113 to a neutral residue such as Ala or Gln or positioning the counter anion more distant with respect to its PSB (E113D mutant) caused a blue shift. These observations indicated that Glu113 definitely influences the protonation state of the SB and its  $pK_a$ . The negative charge of the carboxylate compensates for the positive charge of the PSB. Sakmar et al. introduced a large amount of different solute anions such as chloride to the E113Q and E113A mutants. The anions could fulfill the role of the Glu113 in stabilizing the PSB. Also different degrees of the  $\lambda_{max}$  shift in the presence of different solute anions indicated that the maximum absorptions of different rhodopsins are highly dependant on the interaction between the PSB and its counter anion.<sup>114</sup> In 1993 Oprian and his group successfully showed that human green and red pigments bind chloride ( $Cl^-$ ) ions and undergo a significant amount of red shift in their absorption maxima. They used site-directed mutagenesis to identify the  $Cl^-$ -binding site and demonstrated that this site is conserved in all long wavelength absorbing pigments that had been sequenced and absent from all rod and short wavelength absorbing rhodopsins.<sup>118</sup>

The importance of the distance and angle between the PSB and its counter anion was highlighted by Sheves and his group in 1993. They used model compounds to mimic the PSB of rhodopsin and demonstrated that moving the counter anion further from the PSB caused a significant red shift of the absorption maxima of the rhodopsins. Changing the angle did not influence the maximum absorption, however it significantly perturbed the  $pK_a$  of the PSB. An optimum angle was obtained when one or more water molecules could coordinate to the PSB.<sup>116,119</sup> Model studies have proposed several theories regarding the presence of well-ordered water molecules in the binding pocket of the rhodopsins in the vicinity of the PSB. These models suggest that the water molecules increase the  $pK_a$  of the PSB.<sup>116,120,121</sup> FTIR spectroscopic studies of the intramembrane water molecules upon formation of rhodopsin photointermediates have revealed the structural role of a water molecule during the proton transfer.<sup>122</sup>

Although the orientation of the counter anion with respect to the PSB was shown to shift the  $\lambda_{max}$  of rhodopsin, it cannot by itself explain the 200 nm shift that we see moving from blue to red pigments. In 1973 Blatz et al. proposed the possibility that twisting the chromophore around its single bonds changes the orbital overlap and therefore the  $\lambda_{max}$  of the chromophore (Figure 1-8).<sup>93</sup> As mentioned earlier, 11-*cis*-retinal in rhodopsin is not a planar molecule and has a 6-*S-cis* and 12-*S-trans* conformation. However steric factors imposed by the residues inside the pocket of each rhodopsin make retinal twist around its single bonds in different ways and therefore make PSB bound retinal absorb at a different wavelength in each pigment. According to this theory red rhodopsin should have a more planar retinal molecule and therefore more orbital overlap than green or blue rhodopsins. Model studies performed in solution did not give much

insight into this theory since they showed a multitude of torsional angles possible for the C6-C7 single bond.<sup>123</sup> It should be mentioned that theoretical calculations suggest that twisting around the single bonds by itself cannot be the major reason for the 200 nm red shift between blue and red pigments.<sup>94</sup>

### Molecular Genetics Studies of Rhodopsin

In 1984, Nathans and his coworkers isolated and sequenced the gene encoding human rod rhodopsin. They showed that human and bovine rhodopsins are 93.4% homologous and interestingly the key residues such as Lys296 and those known to form loops on the cytoplasmic side of the cell are perfectly conserved between the two species.<sup>124</sup>

Other groups had shown that the cytoplasmic side of the protein includes: (1) Several residues close to the C-terminus that are subject to phosphorylation by rhodopsin kinase,<sup>125</sup> and (2) A catalytic site that promotes GTP-GDP exchange by transducin.<sup>37</sup>

In 1986, Nathans and his group cloned and characterized the genes that encode human cone pigments. They demonstrated that red and green pigments are highly homologous, 96% identical, but all other pair-wise comparisons showed only a 40%-44% sequence homology. The sequence differences indicate candidate residues responsible for spectral tuning.<sup>24,50</sup> The percentages of homology between different human opsins have been tabulated in Table 1-1. The values below the 100 percent diagonal represent the percentage of amino acids that are identical, while those above this diagonal represent the percentage of amino acids that are identical or homologous.<sup>4,24</sup>



Table 1-1 Sequence homology between human rhodopsins

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

In agreement with the original identification of these three color vision genes by Nathans and his group and based on genetic evidence,<sup>24</sup> in 1991 Oprian et al. designed and chemically synthesized genes for each of the three cone pigments. The genes were expressed in COS cells (COS cells are mammalian cells that are employed frequently as a general purpose mammalian expression system to produce small quantities of recombinant proteins for structural and functional studies) and spectra were measured. These results were the first to directly confirm the presence of the blue, green and red genes in the human genome.<sup>126</sup>

In 1986 Nathans et al. showed that human red and green pigments differ by only 15 amino acids of their 364 amino acids, and yet their  $\lambda_{\text{max}}$  differ by  $\sim 30$  nm.<sup>24</sup> In 1994 Oprian and his coworkers demonstrated that in going from the red to green pigments only 7 amino acid residues are determinant. These 7 residues are as follows: S116Y:S180A:I230T:A233S:Y277F:T285A:Y309F, where the residues on the left are those of the red and the ones on the right are those of the green pigment, respectively.<sup>127</sup> However in 1991 Jacobs and his group showed that substitution of only three non-polar residues at positions 180, 277 and 285 of green pigment with hydroxyl-bearing (polar)

residues can account for the majority of the absorption differences between the two pigments, A180S (~ 4 nm), F277Y (~ 10 nm) and A285T (~ 16nm).<sup>77,128,129</sup> These residues are all located around the ionone ring of 11-*cis*-retinal. In 1992 Sakmar and his group tested this observation by mutating the three equivalent residues in rod rhodopsin ( $\lambda_{\text{max}} = 500$  nm) to the same hydroxyl-bearing residues. The A164S:F261Y:A269T did not bind 11-*cis*-retinal and therefore did not form a pigment to measure its  $\lambda_{\text{max}}$ , however the F261Y:A269T displayed a ~ 20 nm red shift. The A164S mutant just showed a slight red shift (~ 2 nm) and absorbed at ~ 502 nm. So it appears that only two of these positions (Phe277 and Ala285) are responsible for the differences between red and green pigments in human eye. These results clearly show that newly introduced hydroxyl-bearing amino acid residues can interact directly with the chromophore in the red pigment, and the red-shifted  $\lambda_{\text{max}}$  values indicate an enhanced protein-chromophore interaction. The results also show that the spectral tuning of the mutations is not strictly additive.<sup>130</sup>

To study the residues that are responsible for blue shift in the absorption of the chromophore in blue pigments, other opsins have been mutated to promote absorption at the blue wavelengths. In 1998 Sakmar's group demonstrated that mutation of only 9 amino acid residues (M86L:G90S:A117G:E122L A124T:W265Y:A292S:A295S:A299C) in human rhodopsin ( $\lambda_{\text{max}} = 500$  nm) blue shifted the absorption to ~ 438 nm, which accounted for approximately 80% of the opsin shift between the blue pigment ( $\lambda_{\text{max}} = 410$  nm) and rhodopsin. These positions were identified based on sequence alignment of bovine rhodopsin and human, mouse, rat, and bovine blue cone sequences. Although 12 sites were identified that might induce blue shift, mutation of only 9 residues was

required to blue shift the absorption to 438 nm.<sup>70</sup> A similar study by Farrens and his coworkers showed that mutation of T118A:E122D:A292S in rod rhodopsin produces a rhodopsin that absorbs at 453 nm, a 47 nm blue shift as compared to rhodopsin.<sup>131</sup>

### Color Vision Deficiencies by Pigment Gene Defects

Color vision deficiencies are generally congenital and nonprogressive and their existence and hereditary were described as early as the 18<sup>th</sup> century. John Dalton, who had recognized himself and two of his brothers as “red/green blind”, was one of the first people who reported this condition. Red/green dyschromatopsia are often referred to as Daltonism in his honor. Among other early reports of this condition, in 1876 Horner proposed the inheritance of this condition to be sex linked and in 1937 Bell and Haldane proposed the linkage between hemophilia and color blindness.<sup>132,133</sup>

Cloning of the visual pigment genes opened doors to molecular genetic analysis of inherited red/green color vision deficiencies. Inherited color vision deficiencies are known for all three types of color vision. While defects in blue color vision are very rare and affect 1 in 100,000 individuals, red and green color blindness are more common and affect 8% of the male Caucasian population.<sup>127</sup> The red and green genes are located on the X-chromosome with the red gene upstream of one or more green pigment genes.<sup>134</sup> Genomic southern blot analyses of red and green pigments from individuals with inherited red/green color vision deficiencies (dichromats and anomalous trichromats) showed that red and green opsins go through rearrangements, containing segments from both the red and green pigments.<sup>135</sup> Dichromats lack either the red (protanopes) or green pigment sensitivity (deutanopes) and use only two primary colors to see all other

colors. Anomalous trichromats require three primary colors in color-matching tests, but require either more red (protanomalous) or more green (deuteranomalous) than normal individuals.<sup>127</sup> Cys203 is a conserved residue in green pigments. C203R missense mutation of green pigment disrupts the folding and half-life of the green opsin molecule and therefore its ability to absorb light at the appropriate wavelength and activate transducin.<sup>136</sup>

Blue cone monochromacy (BCM) is a rare X-linked recessive condition that results from the lack of both red and green pigments. Different classes of mutations can cause BCM. A large proportion of BCM patients carry only a 5' red-3' green hybrid pigment gene that harbors a C203R missense mutation. Other BCM individuals have been reported who carry up to three pigment genes, each of which contains this lesion. Some other BCM individuals show various sized deletions (0.65-55 kb) of the proximal part of the red/green cone pigment gene cluster.<sup>137</sup> BCM patients suffer from very poor or the absence of color discrimination, photophobia and severely reduced visual acuity. Although these subjects only have blue-cone pigments, it has been shown that some color discrimination can be made in the region between 400 to 500 nm under dim light when both rod and blue-cone are mediating vision. BCM frequency has been estimated to be 1:100,000.<sup>133</sup>

The term "tritanopia" indicates the weak or absent discrimination of short-wavelength blue-yellow stimuli. Analysis of the blue cone pigment in a small group of families with tritanopia has revealed heterozygous missense mutations in this family, G79R, S214P and P264S. Based on few surveys available the frequency of this deficiency has been variously estimated to be 1:500-1:65,000.<sup>133,138</sup>

Finally the most severe color blindness is rod monochromacy (also known as achromatopsia), where no color discrimination happens. These individuals are almost visually handicapped and are considered blind because they have very poor visual acuity and severe photophobia. The frequency of rod-blindness is estimated to be 1:20,000-1:50,000.<sup>133</sup>

Interestingly there is a mutation in the red cone pigment that does not lead to any color vision deficiency. In a recent survey, it was found that 62% of males have a serine in position 190, while 38% have an alanine residue.<sup>139</sup>

#### Modeling of Cone Pigments Based on the Structure of Bovine Rhodopsin

The X-ray crystal structure of bovine rhodopsin has been used to model and perform rational experimental approaches for studying the cone pigments. Homology models of the cone pigments are used to understand the role of different amino acid residues in regulating the wavelength absorption of the chromophore. Figure 1-10 shows the region around retinal in the modeled cone pigments.<sup>4</sup> For clarity residues that are covering the view of the retinal are not shown. As in bovine rhodopsin, the environments around the retinal in all three cone pigments are hydrophobic. Since red and green pigments are more than 96% identical, their retinal binding sites are almost identical. The central residue forming the cavity is Trp281 (analogous to Trp265 in bovine rhodopsin). The counter ion to the PSB is Glu129 (analogous to Glu113 in bovine rhodopsin). Lys312, forms the SB with retinal and there is a second glutamate, Glu102, next to this lysine. The second glutamate is believed to red shift the absorption spectrum

of the chromophore by increasing the negative electron density of Glu129 and therefore decreasing the positive charge delocalization along the chromophore.<sup>4</sup>

In blue pigments Tyr262 is the residue that forms the cavity. Also since Tyr262 is in the proximity of the ionone ring in this pigment, it is believed to be the major factor in the blue shift of this pigment. This factor was tested by observation that the single mutant Y262W leads to a 10 nm red shift.<sup>140</sup> On the other hand, the second glutamic acid residue present in red and green pigments is absent here and causes an additional blue shift. Some other residues have been modeled close to the counter anion in all the three pigments. For example Ser183 in blue pigment is believed to form a hydrogen bond with Glu110, the counter anion. In green and red pigments Ser110 and Ser202 are believed to do the same thing and hydrogen bond to the counter anion. However whether a hydrogen bond forms or not depends on the free energy of this reaction, which is hard to calculate for large systems like rhodopsin.<sup>4</sup>

It has been shown that red and green pigments have a chloride binding site in their cavity which can act as a point charge in modulating the absorption spectrum of the chromophore.<sup>118</sup> Although residues of His197 and Lys200 are believed to be involved in forming this binding site, there are some controversies between the postulated theories and models of red and green pigments on the residues involved in this site. However it is clear that His197 and Lys200 are fully conserved residues of all long-wavelength pigments and are absent in all short-wavelength rhodopsins. Additional experiments need to be done to elucidate their roles in spectral tuning.<sup>4</sup>

Although all this work has resulted in significant improvement of our knowledge about the mechanism of visual transduction and wavelength regulation, in order to

elucidate the precise mechanism of color vision the structures of each of the color pigments are needed. To date, still it is not clear which of the possible theories is most predominant in determining the maximum absorption of the chromophore or how many factors are involved in a specific opsin shift.

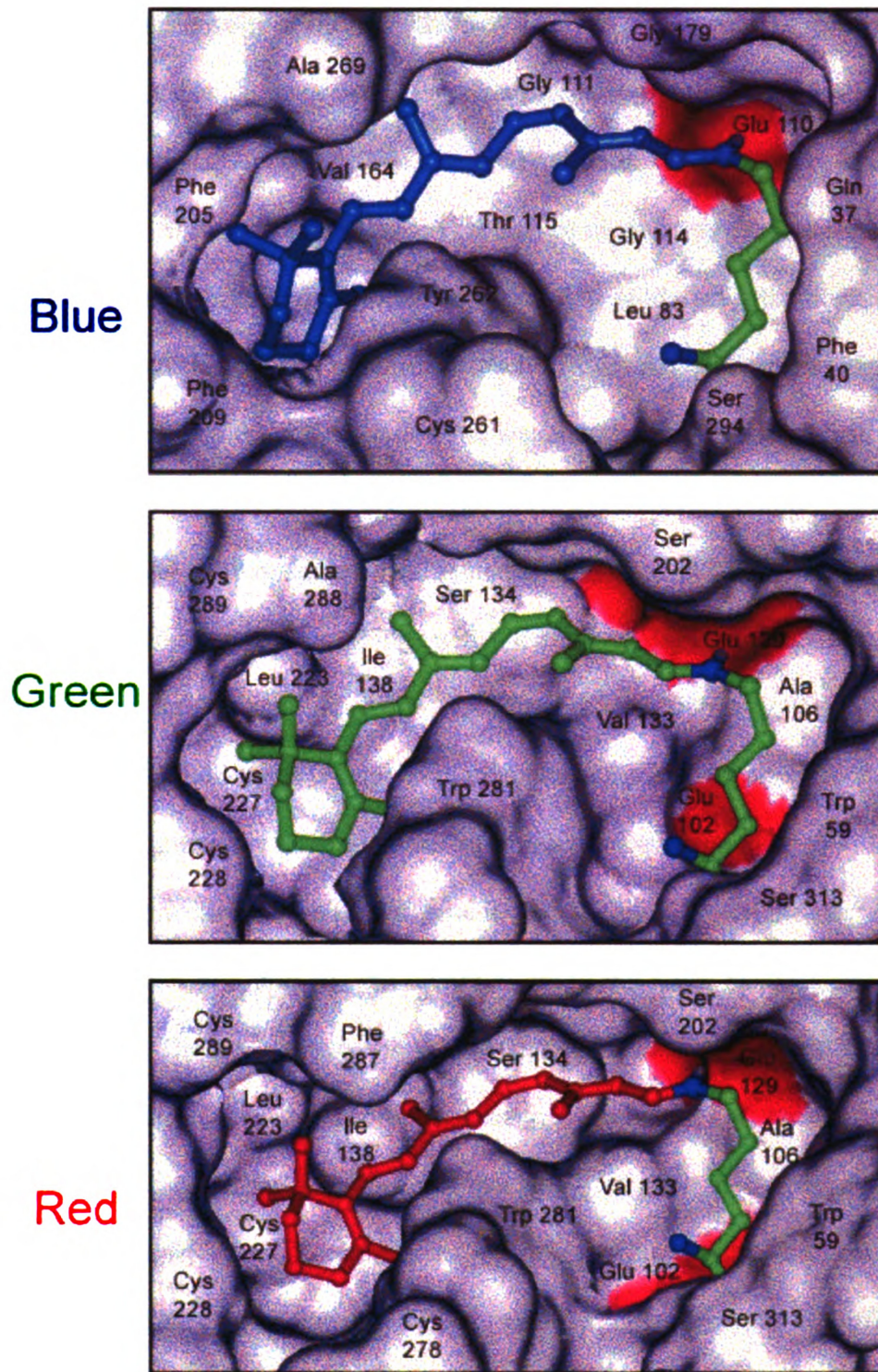


Figure 1-10 Models of retinal binding site in the blue, green and red cone pigments.

Negatively charged residues are shown in red and neutral residues in grey.<sup>4</sup>



### **1.1.5 Need for a Protein Mimic of Rhodopsin**

Over the past few decades various models of visual pigments have been proposed to study the mechanism by which wavelength regulation takes place in each of the different pigments. However the opsins are capable of modulating the wavelength of the chromophore in ways that are difficult or impossible to produce by the solution model systems.<sup>98</sup> Although the crystal structures of each of the cone rhodopsins would be the ultimate answer to our questions regarding the wavelength regulation mechanism in the eye, we believe until the structures are solved, having a “protein mimic” of rhodopsin would be a closer and more realistic model system.

Due to difficulties of working with membrane proteins our collaborators, Professor Babak Borhan and his group (Department of Chemistry, Michigan State University), initiated the idea of reengineering the binding pocket of a protein that is easy to study to make it mimic the binding pocket of visual rhodopsin.

Protein engineering is the use of genetic and chemical techniques to change the structure and function of a protein, and therefore produce a novel product with specific, desired properties. These techniques are very young and recent. Hellinga and coworkers pioneered the field in 1997 and produced several examples of successful protein design.<sup>141-146</sup> Recently there has been an increasing interest in the development of engineered proteins for direct pharmaceutical applications, which can be a milestone in the drug design industry by minimizing the side effects while optimizing the efficiency of the drug.<sup>147,148</sup>

### Characteristic of a Potential Protein Mimic

For a protein to be a good candidate to be engineered into a mimic of rhodopsin, it must have several characteristics. (1) It must be a soluble protein and therefore easy to manipulate and study. (2) Since several site-directed mutations need to be done toward making the potential mimic, the chosen protein has to have a robust structure to withstand the multiple mutations. (3) In order to perform rational mutagenesis studies we need the crystal structures of the intermediate mutants. Thus the protein that we choose must be relatively easy to crystallize and its structure easy to determine. Also since we need to see the details of the interactions between the residues inside the binding pocket with themselves and the chromophore, the candidate protein must be able to produce well-diffracting crystals. (4) Since the majority of the assays that we planned to use to characterize the mutants were spectroscopic assays we needed a protein that could be easily studied spectroscopically. (5) The candidate protein should have an embedded binding pocket to ensure that the changes observed in the structure and the spectroscopic data are only due to the mutations and not the solvent effect. (6) Since we want to design a retinal binding protein that can form a PSB with a lysine residue inside the pocket, as rhodopsin does, it was advantageous if we chose a protein that already binds to one of the retinoid structures as its natural substrate.

In addition, we desired a protein that was easy to express in an *E. coli* expression system. Several proteins were of initial interest, such as retinoid X receptors (RXRs) and cellular retinoid binding proteins (CRBPs). However CRBPs and in particular cellular retinoic acid binding protein type II (CRABPII) are more extensively studied. CRABPII is a small (137 residues and ~ 16 KDa MW) and soluble protein that has high expression

levels. Furthermore the crystal structure of a mutant (R111M) and the retinoic acid (RA)-bound form of this protein had already been determined at the time that we started this project.<sup>149,150</sup> The plasmid of the protein in a pET-17b vector (an *E. coli* expression system) was generously gifted to us by Professor Honggao Yan (Department of Biochemistry, Michigan State University).

To our knowledge this protein mimic is the first example of a protein model of rhodopsin to study the wavelength regulation mechanism. So far CRABP II has been reengineered into a very good retinal binding protein, which binds to retinal by forming a PSB *via* a lysine residue suitably positioned inside the pocket. More mutagenesis studies are being done to substitute the residues interacting with the chromophore with equivalent residues from the rhodopsin pocket that are postulated to regulate the wavelength of each of the cone pigments.

## **1.2 CRABP II: Physiological Importance and the Structure**

Retinoic acid (RA), a biologically active metabolite of vitamin A (retinol), acts as a morphogen during embryonic morphogenesis,<sup>151</sup> and is found to be indispensable for regulating vertebrate cell growth, differentiation and homeostasis.<sup>152-155</sup> It has also been successfully used in the treatment of acute promyelocytic leukemia (APL),<sup>152,156,157</sup> a variety of skin disorders, human cancers and epithelial tumorigenesis.<sup>151,152,158-160</sup>

Developmental gene expression can be disrupted either by an excess or deficiency of RA. Much of the biological role of RA is due to its interaction with the RA receptor (RAR), a member of the steroid/thyroid hormone receptor family of transcription factors. RAR functions as a heterodimer with the retinoid 'X' receptor (RXR). The RAR/RXR

heterodimer recognizes RAR-response elements (RARE) in gene promoters in a ligand-dependant manner and regulates the transcription of these genes.<sup>151,161-166</sup>

Due to its hydrophobic nature RA must be solubilized and transferred to the nucleus, where it binds to RARs. Two homologous cytosolic proteins, cellular RA binding protein type I and II (CRABPI and CRABPII),<sup>167-169</sup> solubilize RA, protect it from isomerization and regulate its effective concentration in the cell by either binding to excess RA and/or metabolizing RA in the cell through interaction with metabolizing enzymes.<sup>170-176</sup> The fact that CRABPs are expressed in tissues that are sensitive to high concentrations of RA supports this hypothesis.<sup>177</sup>

CRABPs are found in virtually all vertebrates.<sup>170</sup> In embryos both CRABPI and II are widely expressed, although they do not co-express in the same cells.<sup>178</sup> In adult rats CRABPI is widely expressed; however, expression of CRABPII is restricted to specific tissues such as skin<sup>170</sup>, testis,<sup>179,180</sup> uterus, ovary,<sup>181,182</sup> choroid plexus,<sup>40</sup> and hematopoietic cells.<sup>157,165</sup>

CRABPs are small (MW ~ 16 kDa), soluble proteins that belong to the family of intracellular lipid-binding proteins (iLBPs). This family is characterized by a well-defined  $\beta$ -barrel formed by two orthogonal five-stranded  $\beta$ -sheets that provide a large, deep and embedded binding cavity. Two short  $\alpha$ -helices act as a cap to the portal of the cavity.<sup>183</sup>

The two isoforms of CRABP are highly homologous in human (74% identity) and among species. For example, CRABPI and II in human and mouse are 99.3% and 93.5% identical, respectively. Higher sequence identity of the same isoform (CRABPI or CRABPII) among different species as compared to different isoforms in the same species

indicates these two homologues should have distinct functions, which explains the conservation of their genes during the course of evolution.<sup>184,185</sup> Although both CRABPs are believed to solubilize and transfer their ligands to the nucleus, experiments show that CRABPII enhances the transcriptional activity of RA,<sup>186</sup> while CRABPI enhances the activity of enzymes that catalyze RA degradation and therefore depresses RA efficacy in the cell.<sup>175,176</sup>

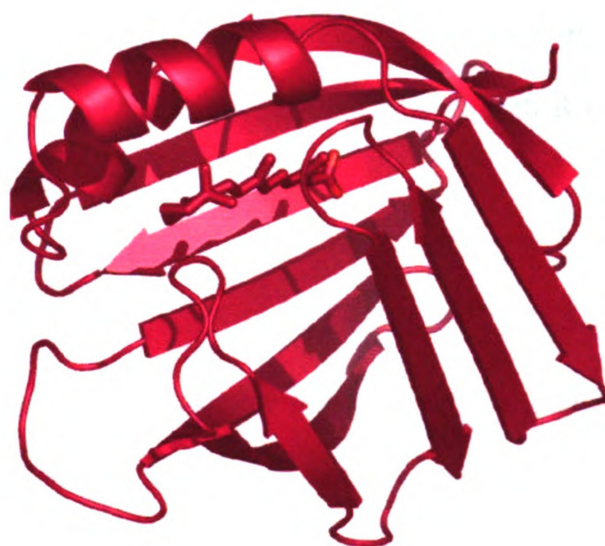


Figure 1-11 Holo-CRABPII (PDB ID: 1CBS).

In 1986 Takase et al. demonstrated that CRABPs are carriers of RA to the nucleus of rat testes by tracing the labeled [ $^3\text{H}$ ]-RA, initially bound to CRABP.<sup>187</sup> In 1998 Gaub et al. clearly showed both CRABPI and II are present in the nucleus.<sup>171</sup> In 1999, Dong et al. successfully showed that the rate of transfer of RA from CRABPII to RAR is highly dependent on the concentration of RAR, and the process follows a first order kinetically controlled mechanism. This finding indicates that CRABPII directly interacts with RAR and forms a protein-protein intermediate complex to directly transfer or “channel” RA to the acceptor.<sup>188</sup> In the same year, Delva et al. showed that CRABPII, and not CRABPI, associates with the RAR $\alpha$ -RXR $\alpha$  complex in a ligand-dependent manner both in *vitro* and

in *vivo* , and enhanced the transcriptional activity of the RAR $\alpha$ -RXR $\alpha$  heterodimer.<sup>157</sup> Recently, Budhu et al. showed that although apo-CRABP II is mostly cytosolic, upon RA binding it was dramatically localized to the nucleus and associated with RAR in a ligand-dependent manner to “channel” RA to RAR. This sensitized cells to RA-induced growth inhibition and enhanced the RA-induced transcriptional activity of RAR $\alpha$ .<sup>185,189</sup> Taken together these observations clearly show the important role CRABP II plays in shuttling RA between the cytoplasm and nucleus of the cell, where RA is needed as a transcriptional regulator. Although CRABP I and II have very high sequence similarity, and their structures are almost identical, CRABP I does not directly interact with RAR and has no effect on the growth inhibition activity of RA.<sup>185,188</sup>

#### Nuclear Localization of Holo-CRABP II

Nuclear localization is believed to occur through the recognition of a nuclear localization signal (NLS). The NLS is a short sequence of basic amino acids that regulates the transport of a protein from the cytoplasm of the cell to the nucleus. Typically deletion of an NLS sequence abrogates nuclear localization and frequently a non-nuclear protein can be localized to the nucleus by fusion to an NLS.<sup>190,191</sup> However, the sequence of CRABP II does not have a recognizable NLS in its primary sequence.<sup>189</sup> Therefore, how these proteins can localize in the nucleus upon RA binding has been a long-standing question.

In an attempt to understand the mechanism of CRABP II nuclear localization upon ligand binding, a comparison was made between the structures of RA-bound CRABP II (PDB ID: 1CBS) and the R111M mutant apo-CRABP II (PDB ID: 1XCA), which was the

only structure of apo-CRABP<sub>II</sub> available at the time.<sup>149,150,189</sup> A significant conformational change was observed in the  $\alpha 2$  helix, which led to a significant change in the electrostatic potential upon RA binding. Most of this change was localized to a change in conformation of three residues, Lys20, Arg29 and Lys30 which upon RA ligation assume a three dimensional structure that could be overlaid with the structure of a classical NLS from the SV40-T antigen. In apparent confirmation of this hypothesis, mutation of all three of these residues results in a CRABP<sub>II</sub> that is incompetent for nuclear localization.<sup>189</sup> However, these conclusions were made based on the structure of a mutant of CRABP<sub>II</sub> (R111M).<sup>149</sup> Arg111 is a conserved residue in CRABP<sub>II</sub> (~80%),<sup>192</sup> and the R111M mutation results in a 45 fold decrease in the RA binding affinity of the protein.<sup>149,193,194</sup> Therefore its mutation is likely to have a significant impact on the tertiary structure of the protein. However since attempts in crystallizing the apo-wild type (WT) had so far been unsuccessful, the apo-R111M mutant was the only apo-CRABP<sub>II</sub> crystal structure available at that time.

### Ligand Entry Problem in CRABP<sub>II</sub>

It has been shown that binding of the ligand to proteins in the iLBP family usually results in major changes in the structure of the protein.<sup>183,195,196</sup> Chen et al. used their structure of the R111M mutant of CRABP<sub>II</sub> to address the ligand entry issue. Comparison of the RA-bound WT CRABP<sub>II</sub> to the apo-R111M mutant CRABP<sub>II</sub> suggested that binding of the ligand is accompanied by major structural changes in the  $\alpha 2$  helix, the  $\beta C$ - $\beta D$  and the  $\beta E$ - $\beta F$  hairpin loops of CRABP<sub>II</sub>. It was thought that these structural changes were necessary to open the binding pocket allowing RA entry.

Further, a three-step mechanism of ligand entry was proposed, which consisted of: opening of the binding pocket, exposure of positive electrostatic potential that directs RA to the binding pocket, and interaction of three residues (Arg111, Arg132, and Tyr134) located deep inside the pocket, with the carboxylic group of RA to stabilize bound RA deep inside the binding cavity.<sup>149</sup>

To test the issues raised by Sessler et al.<sup>189</sup> and Chen et al.<sup>149</sup>, we have crystallized and determined the structure of apo-WT CRABP II at very high resolution (1.35 Å). We have also extended the resolution of the wild type CRABP II bound to RA (CRABP II•RA) complex to 1.48 Å resolution so that a more detailed comparison of the two structures can be made.<sup>150</sup>



### 1.3 Literature Cited

1. Rubin, M. L. & Walls, G. L. (1969). *Fundamentals of Visual Science*, Charles C. Thomas, Springfield, IL.
2. Shichi, H. (1983). *Biochemistry of Vision*, Academic Press, Inc., London, UK.
3. Frank, R. N., Cavanagh, H. D. & Kenyon, K. R. (1973). Light-Stimulated Phosphorylation of Bovine Visual Pigments by Adenosine-Triphosphate. *Journal of Biological Chemistry* **248**, 596-609.
4. Stenkamp, R. E., Filipek, S., Driessen, C. A. G. G., Teller, D. C. & Palczewski, K. (2002). Crystal structure of rhodopsin: a template for cone visual pigments and other G protein-coupled receptors. *Biochimica Et Biophysica Acta-Biomembranes* **1565**, 168-182.
5. Wade, N. (1998). *A Natural History of Vision*, The MIT Press, Cambridge, MA.
6. Wald, G. (1968). Molecular Basis of Visual Excitation. *Science* **162**, 230-239.
7. Bonting, S. J. (1969). The Mechanism of the Visual Process. *Curr. Top. Bioenerg.* **3**, 351-415.
8. Abrahamson, E. W. & Wiesenfeld, J. R. (1972). The Structure, Spectra, and Reactivity of Visual Pigments. In *Photochemistry of Vision* (Dartnall, H. J. A., ed.), Vol. 1. 7 vols. Berlin, Springer-Verlag, NY.
9. Kropf, A. (1972). The Structure and Reactions of Visual Pigments. In *Physiology of Photoreceptor Organs* (Dartnall, H. J. A., ed.), Vol. 2. 7 vols. Berlin, Springer-Verlag, NY.
10. Yoshizawa, T. & Horiuchi, T. (1973). Studies on Intermediates of Visual Pigments by Absorption Spectra at Liquid Helium Temperature And Circular Dichroism at Low Temperatures. In *Biochemistry and Physiology of Visual Pigments* (Langer, H., ed.), pp. 69-81. Berlin, Springer-Verlag, NY.
11. Padgham, C. A. & Saunders, J. E. (1975). *The Perception of Light and Colour*, G. Bell & Sons Ltd, London, UK.

12. Wald, G. (1935). Carotenoids and the visual cycle. *Journal of General Physiology* **19**, 351-371.
13. Wald, G. (1936). Pigments of the retina I. The bull frog. *Journal of General Physiology* **19**, 781-795.
14. Rando, R. R. (1990). The Chemistry of Vitamin-a and Vision. *Angewandte Chemie-International Edition in English* **29**, 461-480.
15. Filipek, S., Stenkamp, R. E., Teller, D. C. & Palczewski, K. (2003). G protein-coupled receptor rhodopsin: A prospectus. *Annual Review of Physiology* **65**, 851-879.
16. Hubbard, R. & Kropf, A. (1958). The Action of Light on Rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* **44**, 130-139.
17. Wald, G. (1968). Molecular Basis of Visual Excitation. *Nature* **219**, 800-807.
18. Stryer, L. (1988). Molecular-Basis of Visual Excitation. *Cold Spring Harbor Symposia on Quantitative Biology* **53**, 283-294.
19. Stryer, L. (1987). The Molecules of Visual Excitation. *Scientific American* **257**, 42-50.
20. Baylor, D. A., Lamb, T. D. & Yau, K. W. (1979). Responses of Retinal Rods to Single Photons. *Journal of Physiology-London* **288**, 613-634.
21. Molday, R. S. (1998). Photoreceptor membrane proteins, phototransduction, and retinal degenerative diseases - The Friedenwald lecture. *Investigative Ophthalmology & Visual Science* **39**, 2493-2513.
22. Kraft, T. W., Makino, C. L., Mathies, R. A., Lugtenburg, J., Schnapf, J. L. & Baylor, D. A. (1990). Cone Excitations and Color-Vision. *Cold Spring Harbor Symposia on Quantitative Biology* **55**, 635-641.
23. Berg, J. M., Tymoczko, J. L. & Stryer, L. (2002). *Biochemistry*. 5th edit, W.H. Freeman and Company, New York, NY.

24. Nathans, J., Thomas, D. & Hogness, D. S. (1986). Molecular-Genetics of Human Color-Vision - the Genes Encoding Blue, Green, and Red Pigments. *Science* **232**, 193-202.
25. Ripps, H. & Weale, R. A. (1969). *The Visual Photoreceptors*. The Eye (Davson, H., Ed.), 2. 6 vols, Academic Press, Inc., New York, NY.
26. Ali, M. A. & Klyne, M. A. (1985). *Vision in Vertebrates*, Plenum Press, New York, NY.
27. Schnapf, J. L. & Baylor, D. A. (1987). How Photoreceptor Cells Respond to Light. *Scientific American* **256**, 40-47.
28. Nathans, J. (1992). Rhodopsin - Structure, Function, and Genetics. *Biochemistry* **31**, 4923-4931.
29. Pugh, E. N. & Lamb, T. D. (1993). Amplification and Kinetics of the Activation Steps in Phototransduction. *Biochimica Et Biophysica Acta* **1141**, 111-149.
30. Fotiadis, D., Liang, Y., Filipek, S., Saperstein, D. A., Engel, A. & Palczewski, K. (2003). Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **421**, 127-128.
31. Fotiadis, D., Liang, Y., Filipek, S., Saperstein, D. A., Engel, A. & Palczewski, K. (2003). Biophysics - Is rhodopsin dimeric in native rods? Reply. *Nature* **426**, 31-31.
32. Liang, Y., Fotiadis, D., Filipek, S., Saperstein, D. A., Palczewski, K. & Engel, A. (2003). Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. *Journal of Biological Chemistry* **278**, 21655-21662.
33. Niu, S. L. & Mitchell, D. C. (2005). Effect of packing density on rhodopsin stability and function in polyunsaturated membranes. *Biophysical Journal* **89**, 1833-1840.
34. Bruckert, F., Chabre, M. & Vuong, T. M. (1992). Kinetic-Analysis of the Activation of Transducin by Photoexcited Rhodopsin - Influence of the Lateral Diffusion of Transducin and Competition of Guanosine Diphosphate and

- Guanosine Triphosphate for the Nucleotide Site. *Biophysical Journal* **63**, 616-629.
35. Kahlert, M. & Hofmann, K. P. (1991). Reaction-Rate and Collisional Efficiency of the Rhodopsin-Transducin System in Intact Retinal Rods. *Biophysical Journal* **59**, 375-386.
  36. Calvert, P. D., Govardovskii, V. I., Krasnoperova, N., Anderson, R. E., Lem, J. & Makino, C. L. (2001). Membrane protein diffusion sets the speed of rod phototransduction. *Nature* **411**, 90-94.
  37. Fung, B. K. K. & Stryer, L. (1980). Photolyzed Rhodopsin Catalyzes the Exchange of Gtp for Bound Gdp in Retinal Rod Outer Segments. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **77**, 2500-2504.
  38. Iiri, T., Farfel, Z. & Bourne, H. R. (1998). G-protein diseases furnish a model for the turn-on switch. *Nature* **394**, 35-38.
  39. Zhukovsky, E. A., Robinson, P. R. & Oprian, D. D. (1991). Transducin Activation by Rhodopsin without a Covalent Bond to the 11-Cis-Retinal Chromophore. *Science* **251**, 558-560.
  40. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M. & Miyano, M. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739-745.
  41. Okada, T., Le Trong, I., Fox, B. A., Behnke, C. A., Stenkamp, R. E. & Palczewski, K. (2000). X-ray diffraction analysis of three-dimensional crystals of bovine rhodopsin obtained from mixed micelles. *Journal of Structural Biology* **130**, 73-80.
  42. Ryba, N. J. P., Hall, M. D. & Findlay, J. B. (1992). Rhodopsin. In *Molecular Biology of G-Protein-Coupled Receptors* (Brann, M. R., ed.), pp. 1-30. Birkhäuser, Boston, MA.
  43. Stenkamp, R. E., Teller, D. C. & Palczewski, K. (2005). Rhodopsin: A structural primer for G-protein coupled receptors. *Archiv Der Pharmazie* **338**, 209-216.

44. Filipek, S., Teller, D. C., Palczewski, K. & Stenkamp, R. (2003). The crystallographic model of rhodopsin and its use in studies of other G protein-coupled receptors. *Annual Review of Biophysics and Biomolecular Structure* **32**, 375-397.
45. Sakmar, T. P. (2002). Structure of rhodopsin and the superfamily of seven-helical receptors: the same and not the same. *Current Opinion in Cell Biology* **14**, 189-195.
46. Sharpe, L., Stockman, A., Jägle, H. & Nathans, J. (1999). *Opsin Genes, Cone Photopigments, Color Vision, and Color Blindness*. Color Vision: From Genes to Perception (Gegenfurtner, K. R. & Sharpe, L. T., Eds.), Cambridge University Press, Cambridge, UK.
47. Bownds, D. (1967). Site of Attachment of Retinal in Rhodopsin. *Nature* **216**, 1178-1181.
48. Wang, J. K., McDowell, J. H. & Hargrave, P. A. (1980). Site of Attachment of 11-Cis-Retinal in Bovine Rhodopsin. *Biochemistry* **19**, 5111-5117.
49. Chabre, M., Breton, J., Michelvillaz, M. & Saibil, H. (1982). Linear Dichroism Studies in the Visible, Uv, and Ir on Oriented Rod Suspensions. *Methods in Enzymology* **81**, 605-616.
50. Nathans, J. (1990). Determinants of Visual Pigment Absorbance - Role of Charged Amino-Acids in the Putative Transmembrane Segments. *Biochemistry* **29**, 937-942.
51. Stenkamp, R. E., Teller, D. C. & Palczewski, K. (2002). Crystal structure of rhodopsin: A G-protein-coupled receptor. *ChemBiochem* **3**, 963-967.
52. Kim, J. E., Tauber, M. J. & Mathies, R. A. (2001). Wavelength dependent cis-trans isomerization in vision. *Biochemistry* **40**, 13774-13778.
53. Dartnall, H. J. (1968). Photosensitivities of Visual Pigments in Presence of Hydroxylamine. *Vision Research* **8**, 339-&.
54. Cooper, A. (1979). Energy Uptake in the 1st Step of Visual Excitation. *Nature* **282**, 531-533.

55. Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L. & Khorana, H. G. (1996). Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **274**, 768-770.
56. Borhan, B., Souto, M. L., Imai, H., Shichida, Y. & Nakanishi, K. (2000). Movement of retinal along the visual transduction path. *Science* **288**, 2209-2212.
57. Okada, T., Ernst, O. P., Palczewski, K. & Hofmann, K. P. (2001). Activation of rhodopsin: new insights from structural and biochemical studies. *Trends in Biochemical Sciences* **26**, 318-324.
58. Konig, B., Welte, W. & Hofmann, K. P. (1989). Photoactivation of Rhodopsin and Interaction with Transducin in Detergent Micelles - Effect of Doping with Steroid Molecules. *Febs Letters* **257**, 163-166.
59. Degrip, W. J., Gray, D., Gillespie, J., Bovee, P. H. M., Vandenberg, E. M. M., Lugtenburg, J. & Rothschild, K. J. (1988). Photoexcitation of Rhodopsin - Conformation Changes in the Chromophore, Protein and Associated Lipids as Determined by Ftir Difference Spectroscopy. *Photochemistry and Photobiology* **48**, 497-504.
60. Birge, R. R. (1981). Photophysics of Light Transduction in Rhodopsin and Bacteriorhodopsin. *Annual Review of Biophysics and Bioengineering* **10**, 315-354.
61. Mathies, R. (1982). Resonance Raman-Spectroscopy of Rhodopsin and Bacteriorhodopsin Isotopic Analogs. *Methods in Enzymology* **88**, 633-643.
62. Leibovic, K. N. (1990). Vertebrate Photoreceptors. In *Science of Vision* (Leibovic, K. N., ed.), pp. 16-52. Springer-Verlag, New York, NY.
63. Teller, D. C., Stenkamp, R. E. & Palczewski, K. (2003). Evolutionary analysis of rhodopsin and cone pigments: connecting the three-dimensional structure with spectral tuning and signal transfer. *Febs Letters* **555**, 151-159.
64. Fung, B. K. K., Hurley, J. B. & Stryer, L. (1981). Flow of Information in the Light-Triggered Cyclic-Nucleotide Cascade of Vision. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **78**, 152-156.

65. Nathans, J. (1999). The evolution and physiology of human color vision: Insights from molecular genetic studies of visual pigments. *Neuron* **24**, 299-312.
66. Blatz, P. E., Baumgart, N., Balasubramanian, V., Balasubramanian, P. & Stedman, E. (1971). Wavelength Regulation in Visual Pigment Chromophore - Large Induced Bathochromic Shifts in Retinol and Related Polyenes. *Photochemistry and Photobiology* **14**, 531-549.
67. Erickson, J. O. & Blatz, P. E. (1968). N-Retinylidene-1-Amino-2-Propanol - a Schiff Base Analog for Rhodopsin. *Vision Research* **8**, 1367-1375.
68. Kochendoerfer, G. G., Lin, S. W., Sakmar, T. P. & Mathies, R. A. (1999). How color visual pigments are tuned. *Trends in Biochemical Sciences* **24**, 300-305.
69. Nakanishi, K., Balogh-Nair, V., Arnaboldi, M., Tsujimoto, K. & Honig, B. (1980). An External Point-Charge Model for Bacteriorhodopsin to Account for Its Purple Color. *Journal of the American Chemical Society* **102**, 7945-7947.
70. Lin, S. W., Kochendoerfer, G. G., Carroll, H. S., Wang, D., Mathies, R. A. & Sakmar, T. P. (1998). Mechanisms of spectral tuning in blue cone visual pigments - Visible and raman spectroscopy of blue-shifted rhodopsin mutants. *Journal of Biological Chemistry* **273**, 24583-24591.
71. Irving, C. S., Byers, G. W. & Leermaker, P. A. (1970). Spectroscopic Model for Visual Pigments - Influence of Microenvironmental Polarizability. *Biochemistry* **9**, 858-864.
72. Albeck, A., Livnah, N., Gottlieb, H. & Sheves, M. (1992). C-13 Nmr-Studies of Model Compounds for Bacteriorhodopsin - Factors Affecting the Retinal Chromophore Chemical-Shifts and Absorption Maximum. *Journal of the American Chemical Society* **114**, 2400-2411.
73. Zhukovsky, E. A. & Oprea, D. D. (1989). Effect of Carboxylic-Acid Side-Chains on the Absorption Maximum of Visual Pigments. *Science* **246**, 928-930.
74. Gonzalez-Luque, R., Garavelli, M., Bernardi, F., Merchán, M., Robb, M. A. & Olivucci, M. (2000). Computational evidence in favor of a two-state, two-mode model of the retinal chromophore photoisomerization. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 9379-9384.

75. Vreven, T. & Morokuma, K. (2003). Investigation of the S-0 -> S-1 excitation in bacteriorhodopsin with the ONIOM(MO : MM) hybrid method. *Theoretical Chemistry Accounts* **109**, 125-132.
76. Rajamani, R. & Gao, J. L. (2002). Combined QM/MM study of the opsin shift in bacteriorhodopsin. *Journal of Computational Chemistry* **23**, 96-105.
77. Neitz, M., Neitz, J. & Jacobs, G. H. (1991). Spectral Tuning of Pigments Underlying Red-Green Color-Vision. *Science* **252**, 971-974.
78. Suzuki, H., Komatsu, T. & Kitajima, H. (1974). Theory of Optical Property of Visual Pigment. *Journal of the Physical Society of Japan* **37**, 177-185.
79. Andersen, L. H., Nielsen, I. B., Kristensen, M. B., El Ghazaly, M. O. A., Haacke, S., Nielsen, M. B. & Petersen, M. A. (2005). Absorption of Schiff-base retinal chromophores in vacuo. *Journal of the American Chemical Society* **127**, 12347-12350.
80. Teller, D. C., Okada, T., Behnke, C. A., Palczewski, K. & Stenkamp, R. E. (2001). Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry* **40**, 7761-7772.
81. Okada, T., Fujiyoshi, Y., Silow, M., Navarro, J., Landau, E. M. & Shichida, Y. (2002). Functional role of internal water molecules in rhodopsin revealed by x-ray crystallography. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 5982-5987.
82. Okada, T., Sugihara, M., Bondar, A. N., Elstner, M., Entel, P. & Buss, V. (2004). The retinal conformation and its environment in rhodopsin in light of a new 2.2 angstrom crystal structure. *Journal of Molecular Biology* **342**, 571-583.
83. Sakmar, T. P. (1998). Rhodopsin: A prototypical G protein-coupled receptor. *Progress in Nucleic Acid Research and Molecular Biology, Vol 59* **59**, 1-34.
84. Schertler, G. F. X. & Hargrave, P. A. (1995). Projection Structure of Frog Rhodopsin in 2 Crystal Forms. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 11578-11582.



85. Unger, V. M., Hargrave, P. A., Baldwin, J. M. & Schertler, G. F. X. (1997). Arrangement of rhodopsin transmembrane alpha-helices. *Nature* **389**, 203-206.
86. Baldwin, J. M., Schertler, G. F. X. & Unger, V. M. (1997). An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors. *Journal of Molecular Biology* **272**, 144-164.
87. Unger, V. M. & Schertler, G. F. X. (1995). Low-Resolution Structure of Bovine Rhodopsin Determined by Electron Cryomicroscopy. *Biophysical Journal* **68**, 1776-1786.
88. Schertler, G. F. X., Villa, C. & Henderson, R. (1993). Projection Structure of Rhodopsin. *Nature* **362**, 770-772.
89. Herzyk, P. & Hubbard, R. E. (1998). Combined biophysical and biochemical information confirms arrangement of transmembrane helices visible from the three-dimensional map of frog rhodopsin. *Journal of Molecular Biology* **281**, 741-754.
90. Kropf, A. & Hubbard, R. (1958). The Mechanism of Bleaching Rhodopsin. *Annals of the New York Academy of Sciences* **74**, 266-280.
91. Mathies, R. & Stryer, L. (1976). Retinal Has a Highly Dipolar Vertically Excited Singlet-State - Implications for Vision. *Proceedings of the National Academy of Sciences of the United States of America* **73**, 2169-2173.
92. Pepe, I. M. (1999). Rhodopsin and phototransduction. *Journal of Photochemistry and Photobiology B-Biology* **48**, 1-10.
93. Blatz, P. E. & Liebman, P. A. (1973). Wavelength Regulation in Visual Pigments. *Experimental Eye Research* **17**, 573-580.
94. Kakitani, H., Kakitani, T., Rodman, H. & Honig, B. (1985). On the Mechanism of Wavelength Regulation in Visual Pigments. *Photochemistry and Photobiology* **41**, 471-479.
95. Tan, Q., Lou, J. H., Borhan, B., Karnaukhova, E., Berova, N. & Nakanishi, K. (1997). Absolute sense of twist of the C12-C13 bond of the retinal chromophore

in bovine rhodopsin based on exciton-coupled CD spectra of 11,12-dihydroretinal analogues. *Angewandte Chemie-International Edition in English* **36**, 2089-2093.

96. Verdegem, P. J. E., Bovee-Geurts, P. H. M., de Grip, W. J., Lugtenburg, J. & de Groot, H. J. M. (1999). Retinylidene ligand structure in bovine rhodopsin, metarhodopsin-I, and 10-methylrhodopsin from internuclear distance measurements using C-13-labeling and 1-D rotational resonance MAS NMR. *Biochemistry* **38**, 11316-11324.
97. Blatz, P. E., Mohler, J. H. & Navangul, H. V. (1972). Anion-Induced Wavelength Regulation of Absorption Maxima of Schiff-Bases of Retinal. *Biochemistry* **11**, 848-&.
98. Honig, B., Greenberg, A. D., Dinur, U. & Ebrey, T. G. (1976). Visual-Pigment Spectra - Implications of Protonation of Retinal Schiff-Base. *Biochemistry* **15**, 4593-4599.
99. Honig, B., Ebrey, T., Callender, R. H., Dinur, U. & Ottolenghi, M. (1979). Photoisomerization, Energy-Storage, and Charge Separation - Model for Light Energy Transduction in Visual Pigments and Bacteriorhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 2503-2507.
100. Hays, T. R., Lin, S. H. & Eyring, H. (1980). Wavelength Regulation in Rhodopsin - Effects of Dipoles and Amino-Acid Side-Chains. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **77**, 6314-6318.
101. Irving, C. S., Byers, G. W. & Leermake, Pa. (1969). Effect of Solvent Polarizability on Absorption Spectrum of All-Trans-Retinylpyrrolidininium Perchlorate. *Journal of the American Chemical Society* **91**, 2141-2143.
102. Creemers, A. F. L., Klaassen, C. H. W., Bovee-Geurts, P. H. M., Kelle, R., Kragl, U., Raap, J., de Grip, W. J., Lugtenburg, J. & de Groot, H. J. M. (1999). Solid state N-15 NMR evidence for a complex Schiff base counterion in the visual G-protein-coupled receptor rhodopsin. *Biochemistry* **38**, 7195-7199.
103. Eilers, M., Reeves, P. J., Ying, W. W., Khorana, H. G. & Smith, S. O. (1999). Magic angle spinning NMR of the protonated retinylidene Schiff base nitrogen in rhodopsin: Expression of N-15-lysine- and C-13-glycine-labeled opsin in a stable cell line. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 487-492.

104. Krebs, A., Edwards, P. C., Villa, C., Li, J. D. & Schertler, G. F. X. (2003). The three-dimensional structure of bovine rhodopsin determined by electron cryomicroscopy. *Journal of Biological Chemistry* **278**, 50217-50225.
105. Kochendoerfer, G. G., Wang, Z. Y., Oprian, D. D. & Mathies, R. A. (1997). Resonance Raman examination of the wavelength regulation mechanism in human visual pigments. *Biochemistry* **36**, 6577-6587.
106. Oseroff, A. R. & Callender, R. H. (1974). Resonance Raman-Spectroscopy of Rhodopsin in Retinal Disk Membranes. *Biochemistry* **13**, 4243-4248.
107. Bagley, K. A., Baloghnaïr, V., Croteau, A. A., Dollinger, G., Ebrey, T. G., Eisenstein, L., Hong, M. K., Nakanishi, K. & Vittitow, J. (1985). Fourier-Transform Infrared Difference Spectroscopy of Rhodopsin and Its Photoproducts at Low-Temperature. *Biochemistry* **24**, 6055-6071.
108. Akhtar, M., Blossé, P. T. & Dewhurst, P. B. (1968). Studies on Vision - Nature of Retinal-Op sin Linkage. *Biochemical Journal* **110**, 693-&.
109. Hubbard, R. E. (1958). *Proc. Nat. Phys. Lab. Symp., London, UK*.
110. Honig, B., Dinur, U., Nakanishi, K., Baloghnaïr, V., Gawinowicz, M. A., Arnaboldi, M. & Motto, M. G. (1979). External Point-Charge Model for Wavelength Regulation in Visual Pigments. *Journal of the American Chemical Society* **101**, 7084-7086.
111. Sheves, M., Nakanishi, K. & Honig, B. (1979). Through-Space Electrostatic Effects in Electronic-Spectra - Experimental-Evidence for the External Point-Charge Model of Visual Pigments. *Journal of the American Chemical Society* **101**, 7086-7088.
112. Mollevanger, L. C. P. J., Kentgens, A. P. M., Pardoën, J. A., Courtin, J. M. L., Veeman, W. S., Lugtenburg, J. & Degrip, W. J. (1987). High-Resolution Solid-State C-13-Nmr Study of Carbons C-5 and C-12 of the Chromophore of Bovine Rhodopsin - Evidence for a 6-S-Cis Conformation with Negative-Charge Perturbation near C-12. *European Journal of Biochemistry* **163**, 9-14.
113. Sakmar, T. P., Franke, R. R. & Khorana, H. G. (1989). Glutamic Acid-113 Serves as the Retinylidene Schiff-Base Counterion in Bovine Rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 8309-8313.

114. Sakmar, T. P., Franke, R. R. & Khorana, H. G. (1991). The Role of the Retinylidene Schiff-Base Counterion in Rhodopsin in Determining Wavelength Absorbency and Schiff-Base Pka. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 3079-3083.
115. Steinberg, G., Ottolenghi, M. & Sheves, M. (1993). Pk(a) of the Protonated Schiff-Base of Bovine Rhodopsin - a Study with Artificial Pigments. *Biophysical Journal* **64**, 1499-1502.
116. Gat, Y. & Sheves, M. (1993). A Mechanism for Controlling the Pk(a) of the Retinal Protonated Schiff-Base in Retinal Proteins - a Study with Model Compounds. *Journal of the American Chemical Society* **115**, 3772-3773.
117. Vogel, R., Fan, G. B., Siebert, F. & Sheves, M. (2001). Anions stabilize a metarhodopsin II-like photoproduct with a protonated Schiff base. *Biochemistry* **40**, 13342-13352.
118. Wang, Z. Y., Asenjo, A. B. & Oprian, D. D. (1993). Identification of the Cl<sup>-</sup> Binding Site in the Human Red and Green Color-Vision Pigments. *Biochemistry* **32**, 2125-2130.
119. Livnah, N. & Sheves, M. (1993). Model Compounds Can Mimic Spectroscopic Properties of Bovine Rhodopsin. *Journal of the American Chemical Society* **115**, 351-353.
120. Sampogna, R. V., Yang, A. S. & Honig, B. H. (1994). Calculation of Pkas of Titratable Groups in Bacteriorhodopsin. *Biophysical Journal* **66**, A45-A45.
121. Deng, H., Huang, L. W., Callender, R. & Ebrey, T. (1994). Evidence for a Bound Water Molecule Next to the Retinal Schiff-Base in Bacteriorhodopsin and Rhodopsin - a Resonance Raman-Study of the Schiff-Base Hydrogen/Deuterium Exchange. *Biophysical Journal* **66**, 1129-1136.
122. Kandori, H., Yamazaki, Y., Sasaki, J., Needleman, R., Lanyi, J. K. & Maeda, A. (1995). Water-Mediated Proton-Transfer in Proteins - an Ftir Study of Bacteriorhodopsin. *Journal of the American Chemical Society* **117**, 2118-2119.
123. Honig, B., Hudson, B., Sykes, B. D. & Karplus, M. (1971). Ring Orientation in Beta-Ionone and Retinals (Theoretical/Nmr/Semi-Empirical

Approach/Overhauser Effect/S-Cis Conformation). *Proceedings of the National Academy of Sciences of the United States of America* **68**, 1289-&.

124. Nathans, J. & Hogness, D. S. (1984). Isolation and Nucleotide-Sequence of the Gene Encoding Human Rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **81**, 4851-4855.
125. Kuhn, H., Cook, J. H. & Dreyer, W. J. (1973). Phosphorylation of Rhodopsin in Bovine Photoreceptor Membranes - Dark Reaction after Illumination. *Biochemistry* **12**, 2495-2502.
126. Oprian, D. D., Asenjo, A. B., Lee, N. & Pelletier, S. L. (1991). Design, Chemical Synthesis, and Expression of Genes for the 3 Human Color-Vision Pigments. *Biochemistry* **30**, 11367-11372.
127. Asenjo, A. B., Rim, J. & Oprian, D. D. (1994). Molecular Determinants of Human Red/Green Color Discrimination. *Neuron* **12**, 1131-1138.
128. Yokoyama, R. & Yokoyama, S. (1990). Convergent Evolution of the Red-Like and Green-Like Visual Pigment Genes in Fish, *Astyanax-Fasciatus*, and Human. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 9315-9318.
129. Deeb, S. S. (2004). Molecular genetics of color-vision deficiencies. *Visual Neuroscience* **21**, 191-196.
130. Chan, T., Lee, M. & Sakmar, T. P. (1992). Introduction of Hydroxyl-Bearing Amino-Acids Causes Bathochromic Spectral Shifts in Rhodopsin - Amino-Acid Substitutions Responsible for Red-Green Color Pigment Spectral Tuning. *Journal of Biological Chemistry* **267**, 9478-9480.
131. Janz, J. M. & Farrens, D. L. (2001). Engineering a functional blue-wavelength-shifted rhodopsin mutant. *Biochemistry* **40**, 7219-7227.
132. Francois, J. (1979). Importance of Heredity in Ophthalmology. *Journal Francais D Ophtalmologie* **2**, 569-578.

133. Wissinger, B. & Sharpe, L. T. (1998). New aspects of an old theme: The genetic basis of human color vision. *American Journal of Human Genetics* **63**, 1257-1262.
134. Vollrath, D., Nathans, J. & Davis, R. W. (1988). Tandem Array of Human Visual Pigment Genes at Xq28. *Science* **240**, 1669-1672.
135. Nathans, J., Piantanida, T. P., Eddy, R. L., Shows, T. B. & Hogness, D. S. (1986). Molecular-Genetics of Inherited Variation in Human Color-Vision. *Science* **232**, 203-210.
136. Kazmi, M. A., Sakmar, T. P. & Ostrer, H. (1997). Mutation of a conserved cysteine in the X-linked cone opsins causes color vision deficiencies by disrupting protein folding and stability. *Investigative Ophthalmology & Visual Science* **38**, 1074-1081.
137. Nathans, J., Maumenee, I. H., Zrenner, E., Sadowski, B., Sharpe, L. T., Lewis, R. A., Hansen, E., Rosenberg, T., Schwartz, M., Heckenlively, J. R., Traboulsi, E., Klingaman, R., Bechhansen, N. T., Laroche, G. R., Pagon, R. A., Murphey, W. H. & Weleber, R. G. (1993). Genetic-Heterogeneity among Blue-Cone Monochromates. *American Journal of Human Genetics* **53**, 987-1000.
138. Weitz, C. J., Miyake, Y., Shinzato, K., Montag, E., Zrenner, E., Went, L. N. & Nathans, J. (1992). Human Tritanopia Associated with 2 Amino-Acid Substitutions in the Blue-Sensitive Opsin. *American Journal of Human Genetics* **50**, 498-507.
139. Winderickx, J., Lindsey, D. T., Sanocki, E., Teller, D. Y., Motulsky, A. G. & Deeb, S. S. (1992). Polymorphism in Red Photopigment Underlies Variation in Color Matching. *Nature* **356**, 431-433.
140. Fasick, J. I., Lee, N. & Oprian, D. D. (1999). Spectral tuning in the human blue cone pigment. *Biochemistry* **38**, 11593-11596.
141. Coldren, C. D., Hellinga, H. W. & Caradonna, J. P. (1997). The rational design and construction of a cuboidal iron-sulfur protein. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 6635-6640.
142. Pinto, A. L., Hellinga, H. W. & Caradonna, J. P. (1997). Construction of a catalytically active iron superoxide dismutase by rational protein design.

*Proceedings of the National Academy of Sciences of the United States of America* **94**, 5562-5567.

143. Wisz, M. S., Garrett, C. Z. & Hellinga, H. W. (1998). Construction of a family of Cys(2)His(2) zinc binding sites in the hydrophobic core of thioredoxin by structure-based design. *Biochemistry* **37**, 8269-8277.
144. Dwyer, M. A., Looger, L. L. & Hellinga, H. W. (2003). Computational design of a Zn<sup>2+</sup> receptor that controls bacterial gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 11255-11260.
145. Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. (2003). Computational design of receptor and sensor proteins with novel functions. *Nature* **423**, 185-190.
146. Drexler, K. E. (1981). Molecular Engineering - an Approach to the Development of General Capabilities for Molecular Manipulation. *Proceedings of the National Academy of Sciences of the United States of America-Physical Sciences* **78**, 5275-5278.
147. Chirino, A. J., Ary, M. L. & Marshall, S. A. (2004). Minimizing the immunogenicity of protein therapeutics. *Drug Discovery Today* **9**, 82-90.
148. Marshall, S. A., Lazar, G. A., Chirino, A. J. & Desjarlais, J. R. (2003). Rational design and engineering of therapeutic proteins. *Drug Discovery Today* **8**, 212-221.
149. Chen, X., Tordova, M., Gilliland, G. L., Wang, L. C., Li, Y., Yan, H. G. & Ji, X. H. (1998). Crystal structure of apo-cellular retinoic acid-binding protein type II (R111M) suggests a mechanism of ligand entry. *Journal of Molecular Biology* **278**, 641-653.
150. Kleywegt, G. J., Bergfors, T., Senn, H., Lemotte, P., Gsell, B., Shudo, K. & Jones, T. A. (1994). Crystal-Structures of Cellular Retinoic Acid-Binding Protein-I and Protein-Ii in Complex with All-Trans-Retinoic Acid and a Synthetic Retinoid. *Structure* **2**, 1241-1258.
151. Norris, A. W., Cheng, L., Giguere, V., Rosenberger, M. & Li, E. (1994). Measurement of Subnanomolar Retinoic Acid-Binding Affinities for Cellular

Retinoic Acid-Binding Proteins by Fluorometric Titration. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **1209**, 10-18.

152. De Luca, L. M. (1991). Retinoids and Their Receptors in Differentiation, Embryogenesis, and Neoplasia. *Faseb Journal* **5**, 2924-2933.
153. Gudas, L. J., Sporn, M. B. & A.B., R. (1994). Cellular Biology and Biochemistry of the Retinoids. In *The Retinoids: Biology, Chemistry, and Medicine*. 2 edit. (Sporn, M. B. R., A.B.; Goodman D.S., ed.), pp. 443-520. Raven Press, New York.
154. Lotan, R. (1980). Effects of vitamin A and its analogs (retinoids) on normal and neoplastic cells. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **605**, 33-91.
155. Chambon, P. (1996). A decade of molecular biology of retinoic acid receptors. *Faseb Journal* **10**, 940-954.
156. Chomienne, C., Fenaux, P. & Degos, L. (1996). Retinoid differentiation therapy in promyelocytic leukemia. *Faseb Journal* **10**, 1025-1030.
157. Delva, L., Bastie, J. N., Rochette-Egly, C., Kraiba, R., Balitrand, N., Despouy, G., Chambon, P. & Chomienne, C. (1999). Physical and functional interactions between cellular retinoic acid binding protein II and the retinoic acid-dependent nuclear complex. *Molecular and Cellular Biology* **19**, 7158-7167.
158. Tallman, M. S. (1996). Differentiating therapy in acute myeloid leukemia. *Leukemia* **10**, 1262-1268.
159. Fisher, G. J. & Voorhees, J. J. (1996). Molecular mechanisms of retinoid actions in skin. *Faseb Journal* **10**, 1002-1013.
160. Hong, W. K. & Itri, L. M. (1994). Retinoids and human cancer. In *The Retinoids: Biology, Chemistry, and Medicine*. 2 edit. (M B Sporn, A. B. R., D S Goodman, ed.), pp. 597-630. Raven Press, New York.
161. Morriskay, G. (1993). Retinoic Acid and Craniofacial Development - Molecules and Morphogenesis. *Bioessays* **15**, 9-15.



162. Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. & Evans, R. M. (1995). The Nuclear Receptor Superfamily - the 2nd Decade. *Cell* **83**, 835-839.
163. Leroy, P., Krust, A., Kastner, P., Mendelsohn, C., Zelent, A. & Chambon, P. (1992). Retinoic Acid Receptors. In *Retinoids in Normal Development and Teratogenesis* (Morris-Kay, G., ed.), pp. 7-25. Oxford University Press, Oxford.
164. Leid, M., Kastner, P., Lyons, R., Nakshatri, H., Saunders, M., Zacharewski, T., Chen, J. Y., Staub, A., Garnier, J. M., Mader, S. & Chambon, P. (1992). Purification, Cloning, and R $\alpha$ r Identity of the Hela-Cell Factor with Which R $\alpha$ r or Tr Heterodimerizes to Bind Target Sequences Efficiently. *Cell* **68**, 377-395.
165. Kreutz, M., Fritsche, J., Andreesen, R. & Krause, S. W. (1998). Regulation of cellular retinoic acid binding protein (CRABP II) during human monocyte differentiation in vitro. *Biochemical and Biophysical Research Communications* **248**, 830-834.
166. Klierer, S. A., Umesono, K., Mangelsdorf, D. J. & Evans, R. M. (1992). Retinoid X-Receptor Interacts with Nuclear Receptors in Retinoic Acid, Thyroid-Hormone and Vitamin-D3 Signaling. *Nature* **355**, 446-449.
167. Sani, B. P. & Hill, D. L. (1974). Retinoic Acid - Binding-Protein in Chick-Embryo Metatarsal Skin. *Biochemical and Biophysical Research Communications* **61**, 1276-1282.
168. Ong, D. E. & Chytil, F. (1975). Retinoic Acid-Binding Protein in Rat Tissue - Partial-Purification and Comparison to Rat Tissue Retinol-Binding Protein. *Journal of Biological Chemistry* **250**, 6113-6117.
169. Bailey, J. S. & Siu, C. H. (1988). Purification and Partial Characterization of a Novel Binding-Protein for Retinoic Acid from Neonatal Rat. *Journal of Biological Chemistry* **263**, 9326-9332.
170. Ong, D. E., Newcomer, M. E. & Chytil, F. (1994). Cellular Retinoid-Binding Proteins. In *The Retinoids : Biology, Chemistry, and Medicine* 2 edit. (Michael B. Sporn, A. B. R., DeWitt S. Goodman., ed.), pp. 283-318. Raven Press, New York.
171. Gaub, M. P., Lutz, Y., Ghyselinck, N. B., Scheuer, I., Pfister, V., Chambon, P. & Rochette-Egly, C. (1998). Nuclear detection of cellular retinoic acid binding

- proteins I and II with new antibodies. *Journal of Histochemistry & Cytochemistry* **46**, 1103-1111.
172. Napoli, J. L. (1994). Retinoic Acid Homeostasis. Prespective Roles of B-Carotene, retinol, CRBP and CRABP. In *Vitamin A in Health and Disease* (Blomhoff, R., ed.), pp. 135-188. Marcel Dekker, New York.
  173. Napoli, J. L. (1996). Retinoic acid biosynthesis and metabolism. *Faseb Journal* **10**, 993-1001.
  174. Napoli, J. L. (1999). Interactions of retinoid binding proteins and enzymes in retinoid metabolism. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids* **1440**, 139-162.
  175. Boylan, J. F. & Gudas, L. J. (1991). Overexpression of the Cellular Retinoic Acid Binding Protein-I (Crabp-I) Results in a Reduction in Differentiation-Specific Gene-Expression in F9 Teratocarcinoma Cells. *Journal of Cell Biology* **112**, 965-979.
  176. Boylan, J. F. & Gudas, L. J. (1992). The Level of Crabp-I Expression Influences the Amounts and Types of All-Trans-Retinoic Acid Metabolites in F9 Teratocarcinoma Stem-Cells. *Journal of Biological Chemistry* **267**, 21486-21491.
  177. Dolle, P., Ruberte, E., Leroy, P., Morriskay, G. & Chambon, P. (1990). Retinoic Acid Receptors and Cellular Retinoid Binding-Proteins .1. A Systematic Study of Their Differential Pattern of Transcription During Mouse Organogenesis. *Development* **110**, 1133-1151.
  178. Maden, M. (1994). Roles of Retinoids in Embryonic Development. In *Vitamin A in Health and Disease* (Blomhoff, R., ed.), pp. 289-322. Marcel Dekker, New York.
  179. Zheng, W. L., Bucco, R. A., Schmitt, M. C., Wardlaw, S. A. & Ong, D. E. (1996). Localization of cellular retinoic acid-binding protein (CRABP) II and CRABP in developing rat testis. *Endocrinology* **137**, 5028-5035.
  180. Ong, D. E. & Chytil, F. (1978). Cellular Retinol-Binding Protein from Rat-Liver - Purification and Characterization. *Journal of Biological Chemistry* **253**, 828-832.

181. Wardlaw, S. A., Bucco, R. A., Zheng, W. L. & Ong, D. E. (1997). Variable expression of cellular retinol- and cellular retinoic acid-binding proteins in the rat uterus and ovary during the estrous cycle. *Biology of Reproduction* **56**, 125-132.
182. Zheng, W. L. & Ong, D. E. (1998). Spatial and temporal patterns of expression of cellular retinol-binding protein and cellular retinoic acid-binding proteins in rat uterus during early pregnancy. *Biology of Reproduction* **58**, 963-970.
183. Banaszak, L., Winter, N., Xu, Z. H., Bernlohr, D. A., Cowan, S. & Jones, T. A. (1994). Lipid-Binding Proteins - a Family of Fatty-Acid and Retinoid Transport Proteins. In *Advances in Protein Chemistry, Vol 45*, Vol. 45, pp. 89-151.
184. Budhu, A., Gillilan, R. & Noy, N. (2001). Localization of the RAR interaction domain of cellular retinoic acid binding protein-II. *Journal of Molecular Biology* **305**, 939-949.
185. Budhu, A. S. & Noy, N. (2002). Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest. *Molecular and Cellular Biology* **22**, 2632-2641.
186. Jing, Y. K., Waxman, S. & MirayLopez, R. (1997). The cellular retinoic acid binding protein II is a positive regulator of retinoic acid signaling in breast cancer cells. *Cancer Research* **57**, 1668-1672.
187. Takase, S., Ong, D. E. & Chytil, F. (1986). Transfer of Retinoic Acid from Its Complex with Cellular Retinoic Acid-Binding Protein to the Nucleus. *Archives of Biochemistry and Biophysics* **247**, 328-334.
188. Dong, D., Ruuska, S. E., Levinthal, D. J. & Noy, N. (1999). Distinct roles for cellular retinoic acid-binding proteins I and II in regulating signaling by retinoic acid. *Journal of Biological Chemistry* **274**, 23695-23698.
189. Sessler, R. J. & Noy, N. (2005). A ligand-activated nuclear localization signal in cellular retinoic acid binding protein-II. *Molecular Cell* **18**, 343-353.
190. Stryer, L. (1988). *Biochemistry*. Third edit, W.H. Freeman and Company, New York, NY.

191. Cokol, M., Nair, R. & Rost, B. (2000). Finding nuclear localization signals. *Embo Reports* **1**, 411-415.
192. Gunasekaran, K., Hagler, A. T. & 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.
193. Wang, L. C. & Yan, H. G. (1998). NMR study suggests a major role for Arg111 in maintaining the structure and dynamical properties of type II human cellular retinoic acid binding protein. *Biochemistry* **37**, 13021-13032.
194. Wang, L. C., Li, Y. & Yan, H. G. (1997). Structure-function relationships of cellular retinoic acid-binding proteins - Quantitative analysis of the ligand binding properties of the wild-type proteins and site-directed mutants. *Journal of Biological Chemistry* **272**, 1541-1547.
195. Winter, N. S., Bratt, J. M. & Banaszak, L. J. (1993). Crystal-Structures of Holo and Apo-Cellular Retinol-Binding Protein-Ii. *Journal of Molecular Biology* **230**, 1247-1259.
196. Xu, Z. H., Bernlohr, D. A. & Banaszak, L. J. (1992). Crystal-Structure of Recombinant Murine Adipocyte Lipid-Binding Protein. *Biochemistry* **31**, 3484-3492.

## **Chapter II**

### **Reengineering CRABPII into a Rhodopsin Surrogate**

#### **2.1 The Structure of CRABPII Bound to RA as a Starting Point**

As mentioned in Chapter I, our goal is to study the mechanism by which wavelength regulation takes place in each of the different rhodopsin pigments in eye. Rhodopsin itself is a membrane protein and therefore is hard to study. Opsins are capable of modulating the wavelength of the chromophore in ways that are difficult to impossible to produce by solution model systems.<sup>1</sup> The only structure of a vertebrate rhodopsin is of bovine rod rhodopsin,<sup>2-4</sup> which by itself cannot address the questions regarding the mechanism of wavelength regulation in the eye. We believe that having a “protein mimic” of rhodopsin would be a closer and more realistic model system.

Since many rational mutations are envisioned in achieving a mimic of the rhodopsins, the utmost important factor was to choose a protein that is highly tolerant of mutations and at the same time is easy to over-express and crystallize. A vast amount of work has been done on studying CRABPs using spectroscopic techniques while both the NMR and the X-ray crystal structures have been determined. These reports suggest that CRABPs have a robust structure and are easy to manipulate and study.<sup>5-12</sup> Therefore we chose CRABPII to be engineered as a rhodopsin mimic.

CRABPII, a small (137 residues and ~ 16 kDa MW) cytosolic protein with a large (~ 600 Å<sup>3</sup>) and embedded binding pocket, is comprised of two orthogonal five-stranded β-sheets and a helix-turn-helix motif capping the portal of the binding pocket (Figure 1-11). These features make CRABPII a very good candidate for our purposes. The

embedded binding pocket ensures that absorption and binding properties will be sensitive to mutation and the large binding pocket allows the chromophore to assume a favorable conformation in the pocket and not have steric hindrance by the residues within the pocket.

Figure 2-1 shows 11-*cis*-retinal bound to Lys296 *via* PSB inside the pocket of bovine rod rhodopsin. Residues within 6 Å of retinal and Lys296 are shown and residues within 4 Å of the PSB are labeled. Rhodopsin has a very hydrophobic pocket and the  $pK_a$  of its PSB is estimated to be  $\sim 16$ .<sup>13</sup> It has been shown that Glu113, which is a highly conserved residue among all vertebrate opsins,<sup>14</sup> plays the important role of counter anion for the PSB, stabilizing its positive charge. In bovine rhodopsin this residue is 3.22 Å away from the nitrogen of the PSB, however it is postulated that the counter anion may reside at varying locations in other opsins.<sup>15-17</sup> Mutation of Glu113 to Gln, led to a dramatic blue shift in the maximum absorption of retinal, from 490 nm to 380 nm, indicating deprotonation of the SB.<sup>18</sup>

Therefore in order to reengineer the pocket of CRABP II into a mimic of rhodopsin, first we need to design a Lys residue, inside the pocket, that can perform a nucleophilic attack on the carbonyl carbon of retinal and form a SB with it. We also need to position a counter anion for the PSB, a mimic of Glu113 in rhodopsin. In addition, the hydrophilic residues in close proximity to the chromophore may need to be mutated to more hydrophobic residues, mimicking the hydrophobic pocket of rhodopsin, in order to lower the  $pK_a$  of Lys and therefore making it a good nucleophile.

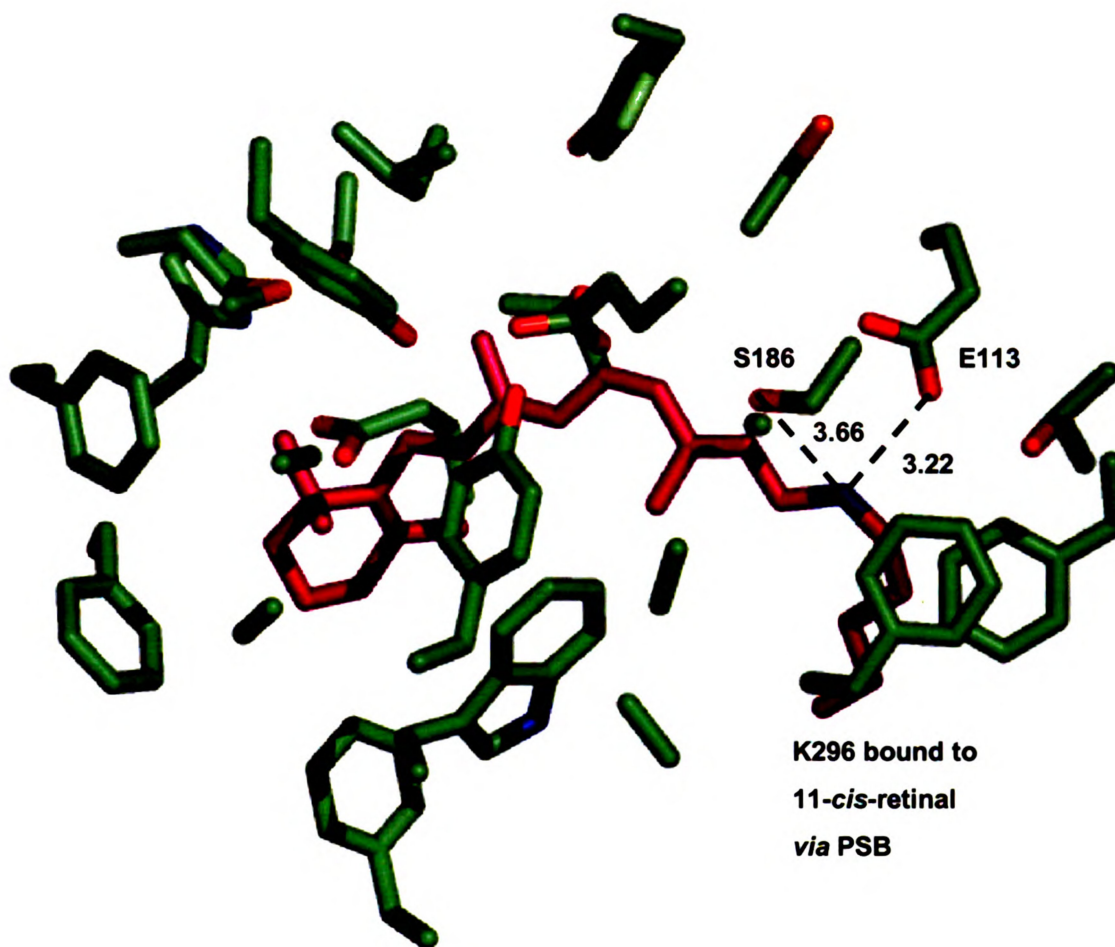


Figure 2-1 Retinal in the binding pocket of bovine rod rhodopsin (PDB ID: 1F88), bound to Lys296 *via* SB. Residues within 6 Å of Lys296 and retinal (violet) are shown. Residues within 4 Å of PSB are labeled. The distances are in angstroms.

#### Crystal Structure of WT CRABP II Bound to RA (PDB ID: 2FR3)

We began with the structure of CRABP II bound to RA (CRABP II•RA) as our starting model. This structure had been solved for the first time by Kleywegt et al. at 1.80 Å (PDB ID: 1CBS, Figure 1-11),<sup>5</sup> however we crystallized it in a different crystallization condition; and extended the resolution to 1.48 Å (PDB ID: 2FR3). The

structure was refined with very good crystallographic R-factors of 12.30% and 17.09% for  $R_{\text{work}}$  and  $R_{\text{free}}$ , respectively. Our structure is identical to the published one, but the higher resolution of our data makes a better and more detailed definition of the model possible. The structure is in the  $P2_12_12_1$  space group and has one molecule in the asymmetric unit. The data collection and refinement statistics are reported in Tables 4-4 and 4-6 in Chapter IV, respectively. All-*trans*-RA binds inside the pocket with its hydrophilic end buried deep inside the cavity, and its ionone ring partially solvent exposed at the portal of the pocket (Figure 2-2-A). RA is mostly surrounded by hydrophobic residues except for the carboxylate group region, which makes direct and tight hydrogen bonds with Arg132 and Tyr134 from one side and water-mediated hydrogen bonds with Arg111 and Thr54 from the other side of the pocket (Figure 2-2-B). These tight hydrogen bonds make RA a very good binder to the protein ( $K_d \sim 2$  nM). Consistent with the structural data, a novel competitive binding assay was developed, by Yan and his co-workers for measuring the relative dissociation constants of the site-directed mutants of CRABPs. Arg111 and Arg132 of CRABPII were replaced with Met by site-directed mutagenesis. The relative dissociation constants of R111M and R132M ( $K_d$  (R111M)/  $K_d$  (CRABP-II) and  $K_d$  (R132M)/  $K_d$  (CRABP-II)) were determined to be 40-45 and 6-8, respectively.<sup>12</sup> As further discussed in this Chapter, our mutants also show the importance of these residues in RA binding.



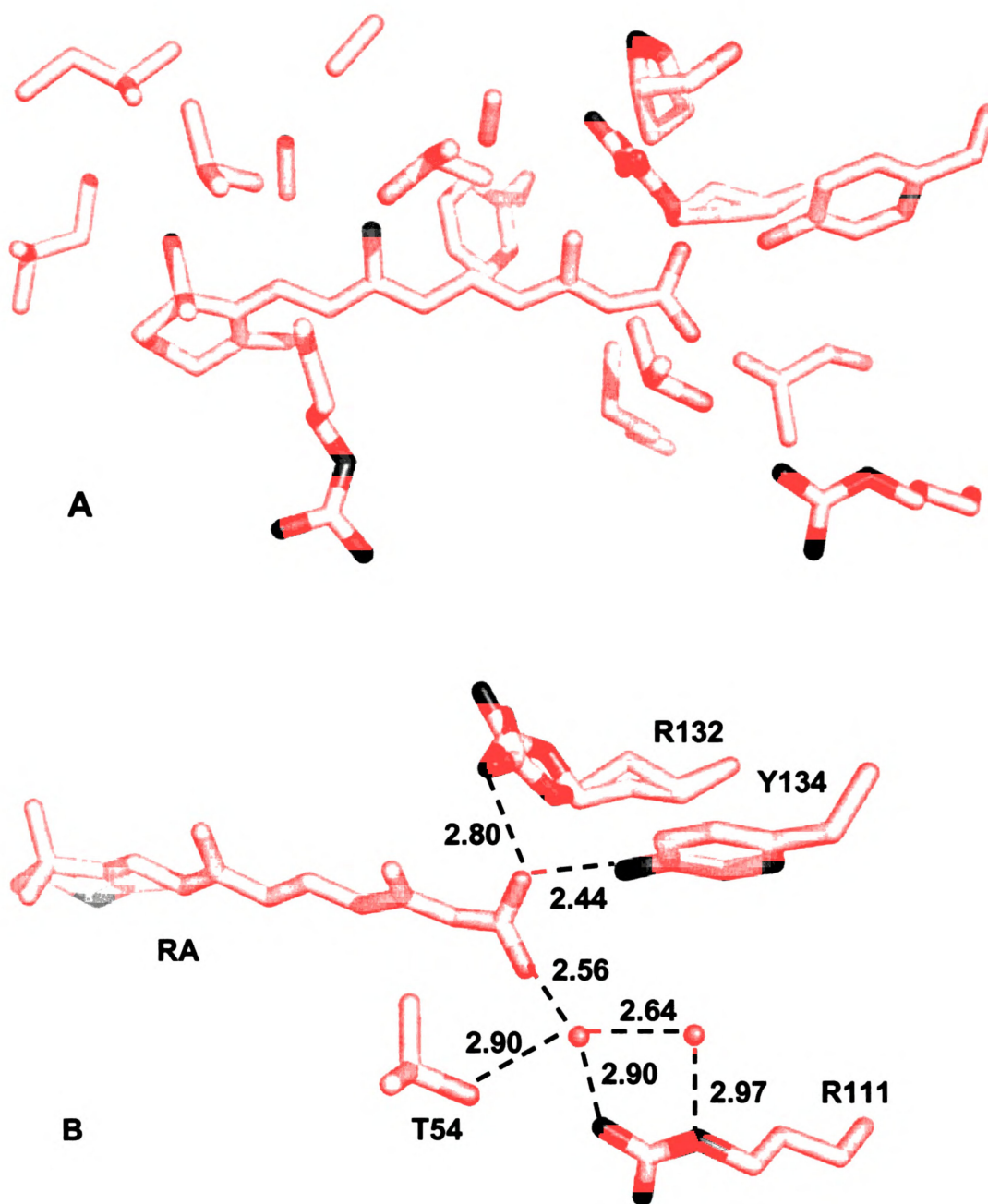


Figure 2-2 (A) RA in the pocket of WT CRABP1; (B) Hydrogen bond interactions of the carboxylate group of RA with the residues inside the pocket. The distances are in angstroms.

We believed that since all-*trans*-retinal and all-*trans*-RA have similar structures, retinal would occupy a very similar position as RA does in the pocket of CRABP<sub>II</sub>. We used all-*trans*-retinal because it is much more stable and less light sensitive compared to its *cis* isomer (the natural substrate of visual rhodopsin). Although the wavelength regulation may be slightly different by all-*trans*- vs. 11-*cis*-retinal, the interactions of the chromophore and the electrostatic environment should be very similar in both isomers, and follow the same principles. Thus, first we will test the proposed hypotheses using the *trans* isomer and after achieving the potential mimic we will substitute it with the *cis* isomer to have a closer mimic to rhodopsin.

In order to perform rational mutagenesis studies toward making the mimic, we needed the crystal structures of the intermediate mutants. The structures not only could depict the picture of the interactions of the residues with each other and the chromophore, inside the pocket, but also could show the effect of the mutations on the overall structures of the mutants. Experimental details of the crystallographic experiments are in Chapter IV.

Since this project involves expression and purification of a significant number of protein mutants, we decided to add an affinity purification tag, His<sub>6</sub>-tag, to the protein to minimize the amount of work and time spent on the purification process.<sup>19</sup> This tag consists of six His residues that will bind to a Ni<sup>2+</sup> resin column (Ni-NTA, Novagen), and allow a fast and clean purification. This was done by cloning CRABP<sub>II</sub> into a pET-Blue2 plasmid *via* NcoI and XhoI restriction sites. Details of the cloning procedure can be found in Dr. Crist's thesis.<sup>20</sup>

However, early on during crystallization trials it was out that the His<sub>6</sub>-tag was hampered crystallization and leading to only poor diffracting crystals. Since CRABP II is known to produce well-diffracting crystals in the absence of a tag,<sup>5,8</sup> we removed the His<sub>6</sub>-tag, which resulted in very well-diffracting crystals. This was not to our surprise because there were other cases reported in the literature in which the His<sub>6</sub>-tag had hindered crystallization.<sup>21</sup> It should be mentioned that there have also been cases where the His<sub>6</sub>-tag has had no negative effect on crystallization,<sup>22</sup> but the general belief is that if both tagged and non-tagged proteins have a comparable expression level, it is better to work with a non-tagged protein because it can potentially modify the structure of the protein, as reported in the case of  $\beta$ -lactamase of a thermophilic *Bacillus licheniformis* strain.<sup>23</sup> So our collaborators re-made all the mutant proteins without a His<sub>6</sub>-tag.

Since His<sub>6</sub>-tagged proteins are easier and faster to purify, Dr. Crist continued most of the mutations and assays with His<sub>6</sub>-tagged proteins on a pET-Blue2 vector. For the best and most successful mutants Dr. Vasileiou made the similar mutant on a pET-17b vector. Eventually we gave up on pET-Blue2 mutations and all the mutations were performed on pET-17b vector. Details about the mutation protocols can be found in Drs. Crist and Vasileiou dissertations.<sup>20,24</sup> Having in hand two series of the same mutants, tagged and non-tagged, verified that the extra residues resulting from the tag did not change the binding and spectroscopic properties of the mutants. Therefore for the rest of this discussion we make no distinction between His<sub>6</sub>-tagged and non-His<sub>6</sub>-tagged mutants.

## **2.2 Spectroscopic Assays Used for Characterization of the Mutants**

Once the mutants were prepared they were characterized by spectroscopic assays to verify SB or PSB formation between retinal and the designed Lys residue. Also these assays were used to measure the binding constants of RA or retinal to each mutant. Here, I briefly discuss each of these assays to explain how they were used to characterize the mutants. More details about the spectroscopic assays can be found in the doctoral theses of Drs. Crist and Vasileiou, who performed the assays.<sup>20,25</sup>

### **2.2.1 Circular Dichroism (CD) Spectroscopy**

Circular Dichroism (CD) spectroscopy was used to make sure that the protein had reserved its structural integrity.<sup>20</sup> The CD result will show the presence of  $\alpha$  helices and  $\beta$  sheets in the structure, but will not tell us exactly what percent of the protein molecules in the sample has  $\alpha$  helical and  $\beta$  sheet properties.<sup>24</sup> Our further crystallization analysis proved that CRABPII is a very robust protein upon mutation, except for mutation of some fully conserved residues that proved to affect the structure of parts of the protein.

### **2.2.2 Measurement of Extinction Coefficient ( $\epsilon$ )**

In the beginning we used the Bradford assay to measure the protein concentration, which is based on the absorbance shift in Coomassie Brilliant Blue G-250 (CBBG) when bound to Arg and aromatic residues.<sup>26</sup> However making a standard curve is very important in this assay, which itself is nonlinear over a wide range of concentrations. Also since response to different proteins can vary widely, choice of standard was very important. Soon we realized that measuring concentration directly using Beer's law

(Abs.= $\epsilon$ bc) gives us more accurate and producible results. To measure the concentrations using the Beer's law, we need the value of the extinction coefficient ( $\epsilon$ ) for each mutant. These values are calculated following an accurate method (to  $\pm 5\%$  in most cases) developed by Gill and Vonhippel in 1989.<sup>27</sup> The method is based on using the calculated  $\epsilon$  of denatured proteins in 6 M Guanidine HCl ( $\epsilon_{\text{den}}$ ) to calculate  $\epsilon$  for the native proteins ( $\epsilon_{\text{nat}}$ ).  $\epsilon_{\text{den}}$  is dependant primarily on the number of Tyr, Trp and Cys residues present and can be calculated according to equation:

$$\epsilon_{\text{den}} = a\epsilon_{\text{Tyr}} + b\epsilon_{\text{Trp}} + c\epsilon_{\text{Cys}} \quad (2.1)$$

where a, b, and c represent the number of respective amino acids per molecule of protein, and their  $\epsilon$  values have been determined experimentally ( $\epsilon_{\text{Trp}} = 5690 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\epsilon_{\text{Tyr}} = 1280 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\epsilon_{\text{Cys}} = 120 \text{ M}^{-1} \text{ cm}^{-1}$ , wild type CRABPII protein contains 3 Trp, 2 Tyr, and 3 Cys residues).

Two protein solutions are prepared, at identical concentrations, one in the native buffer (4 mM  $\text{NaH}_2\text{PO}_4$ , 16 mM  $\text{Na}_2\text{HPO}_4$ , 150 mM NaCl, pH = 7.3) and the other in a denaturing buffer (6 M guanidine HCl, 4 mM  $\text{NaH}_2\text{PO}_4$ , 16 mM  $\text{Na}_2\text{HPO}_4$ , 150 mM NaCl), and the Abs<sub>280</sub> for each sample is measured.

The  $\epsilon_{\text{nat}}$ , the extinction coefficient for the native protein, can be determined according to Beer's law:

$$\text{Abs}_{\text{den}} + \epsilon_{\text{den}} = c_{\text{den}} \quad (2.2)$$

$$\text{Abs}_{\text{nat}} + \epsilon_{\text{nat}} = c_{\text{nat}} \quad (2.3)$$

where, Abs is the UV absorbance at 280 nm, under both native and denatured conditions. Since the concentration of both samples is the same, we can equate the two equations and solve for the extinction coefficient by:

$$\epsilon_{nat} = \frac{(Abs_{nat})(\epsilon_{den})}{(Abs_{den})} \quad (2.4)$$

The measurement must be repeated until a consistent number is obtained, since the calculated number is what is used for the determination of the concentration of each CRABP II mutant for all the following spectroscopic assays.

This method has proved very useful since it is simply based on the knowledge of the amino acid sequence analysis. These values, measured by our collaborators, are listed in Appendix 1.

### 2.2.3 Fluorescence Titration

Fluorescence quenching is used to measure the dissociation constant values for binding retinal or RA to each mutant. We followed a method developed by Li and coworkers,<sup>28</sup> which was also used by other groups.<sup>12</sup> The fluorescence emission of Trp residues in the protein is monitored as a function of the added chromophore. When chromophore is present inside the pocket, it quenches the fluorescence emission of the Trp residues. The chromophore (RA or retinal) is added to the protein solution until no change is observed in the fluorescence emissions of the Trp residues, or in other words the pocket is fully saturated. There are three Trp residues in the pocket of CRABP II: Trp7, Trp87 and Trp109, which are 10.7, 17.8 and 9.4 Å, away from the nearest point on

RA, respectively (Figure 2-3-A). A typical graph obtained from these measurements is shown in Figure 2-3-B. Dissociation constant ( $K_d$ ) values of RA or retinal for each mutant are calculated using non-linear least square regression analysis of this data. These values, measured by our collaborators, are listed in Appendix 2.

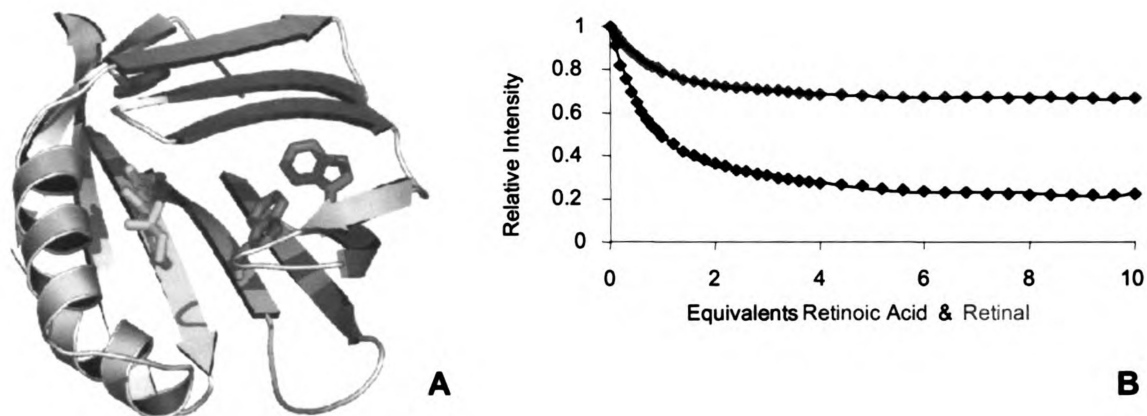


Figure 2-3 The dissociation constant of the chromophore bound to the CRABP II mutants is determined by measuring the fluorescence quenching of Trp residues inside the pocket. (A) There are three Trp residues inside the pocket of CRABP II (Trp7, Trp87, Trp109). (B) A typical fluorescence quenching titration curve.

#### 2.2.4 MALDI-TOF Assays

MALDI-TOF assays were performed on each mutant to verify whether a SB has formed between the designed Lys residue and retinal. In order to perform these assays two types of experiments were performed, which will be briefly discussed here.

### Incubation of Protein with Retinal Followed by MALDI-TOF

The protein is incubated with retinal. After giving the mixture sufficient time (1-2 h) to form a SB and equilibrate, the sample is washed with ethanol and concentrated. The putative covalent bond formed between retinal and the designed Lys residue (Figure 2-4), can be detected as an  $[M+266]^+$  peak in the MALDI-TOF spectrum.<sup>29</sup>

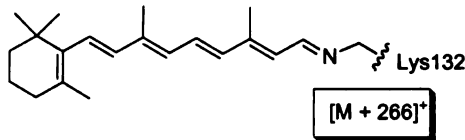


Figure 2-4 Retinal bound to a designed Lys shows a peak at the mass of CRABPII plus 266.

### Reductive Amination of the SB Followed by MALDI-TOF

First the putative SB is reduced using sodium cyanoborohydride ( $\text{NaCNBH}_3$ ) to covalently trap retinal inside the binding pocket and then the sample is analyzed using MALDI-TOF mass spectrometry. SB formation is confirmed if a peak is observed at  $[M+268]^+$  (Figure 2-5).

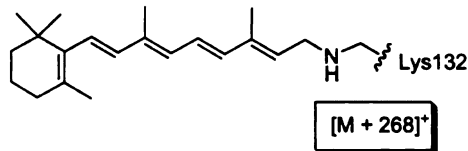


Figure 2-5 Retinal bound to the Lys residue has a peak at the mass of CRABPII plus 268 when SB is reduced.

There is a necessity for both types of experiments, since while some mutants have very stable SB, in some others the SB can be hydrolyzed and therefore reductive amination will help to confirm SB formation. On the other hand in some other mutants the SB cannot tolerate the reductive amination condition but can easily be



detected by simply incubating the mutant with retinal. The results of MALDI assays, measured by our collaborators, are listed in Appendix 2.

### 2.2.5 UV-vis Spectroscopy

UV-vis spectroscopy was used as our main tool to distinguish whether a PSB has formed between the designed Lys residue and retinal. Free retinal absorbs at  $\sim 380$  nm, either in ethanol or TRIS buffer (storage buffer). When the chromophore migrates to the binding pocket of the protein we do not expect to see a significant change in the absorption. However a  $\sim 10$  nm blue shift is observed, probably due to the more hydrophobic environment inside the pocket. We are sure there is

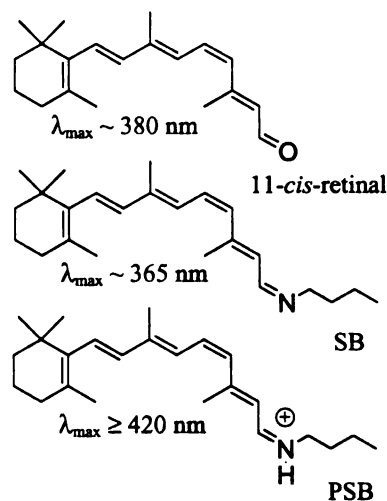


Figure 2-6 Red shift of retinal absorption upon PSB formation

no SB formation taking place because MALDI indicates no SB formation. Upon formation of a SB we do not expect a very large change in the  $\lambda_{\text{max}}$  because only the carbonyl oxygen of retinal is replaced with a nitrogen, which does not make a very large difference in the absorption of the conjugated system,  $\sim 365$  nm (Figure 2-6). However, protonation of the SB increases the levels of cationic conjugation along the chromophore, and therefore results in a dramatic red shift in the absorption of the chromophore (above  $\sim 420$  nm).

### 2.3 X-ray Crystallography

Having the X-ray crystal structures of the CRABP II mutants was a key factor in performing rational mutagenesis studies toward making a mimic of rhodopsin.

When a crystal is exposed to X-rays, constructive interferences between rays scattered from successive planes in the crystal will only take place if the path differences between the rays are equal to an integral number of wavelengths. This is known as Bragg's law:

$$2d \sin \theta = n\lambda \quad (2.5)$$

The Bragg equation gives the condition for diffraction so that if a crystal is rotated in a beam of X-rays, the scattering pattern is a series of intensity maxima. In a crystal, electrons in atoms are the scatters, and each atom has a different effectiveness as a scatter. Consequently, when an experiment is carried out, a set of diffraction maxima of different intensities are observed. The crystal is rotated to obtain the scattering intensity at various angles. The scattering intensity depends on the scattering effectiveness of the individual atoms and the phase of the wave from each scattering source, known as individual structure factor ( $f_j$ ). The structure factor,  $\mathbb{F}(hkl)$ , for each plane ( $hkl$ ) can be defined as the sum of the structure factors for individual atoms,  $f_j$ , times a phase factor,  $\alpha(hkl)$ , for each atom. In other words, the structure factor can be represented by its amplitude and phase:

$$\mathbb{F}(hkl) = \sum f_j e^{2\pi i(hx_j + ky_j + lz_j)} = F(hkl) e^{i\alpha(hkl)} \quad (2.6)$$

where,  $F(hkl)$  is the amplitude and  $\alpha(hkl)$  is the phase.

When the diffracted X-ray is recorded, all information on the phase is lost and only a measurement of the intensity of the diffracted beam is recorded. The intensity in each spot of the diffraction pattern is given by:

$$I(hkl) = [\mathbb{F}(hkl)]^2 \quad (2.7)$$

The electron density,  $\rho(r)$ , is a function of the coordinates of the scattering centers (the atoms) and has a maximum around the position of each atom. What is desired is to convert the measured structure factors into atomic coordinates. This is done by taking the Fourier transform of equation 2.6. In this case, the Fourier transform takes the structure factors, which are functions of the electron density, and inverts the functional dependence so that the electron density is expressed as a function of the structure factors:

$$\rho(r) = \int_{\text{diffraction space}} \mathbb{F}(s) e^{-2\pi i r \cdot s} dv_s \quad (2.8)$$

where,  $dv_s$  is a small unit of volume in diffraction space. The integration can be replaced by a summation since  $\mathbb{F}(s)$  is not continuous and is non-zero only at the reciprocal lattice points. Therefore:

$$\begin{aligned} \rho(xyz) &= \frac{1}{V} \sum_h \sum_k \sum_l \mathbb{F}(hkl) e^{-2\pi i(hx+ky+lz)} \\ &= \frac{1}{V} \sum_h \sum_k \sum_l F(hkl) e^{i\alpha(hkl)} e^{-2\pi i(hx+ky+lz)} \end{aligned} \quad (2.9)$$

where,  $\rho(xyz)$  is the electron density at any point  $x,y,z$  and  $F(hkl)$  is the amplitude, which is proportional to the square root of the measured intensity of each reflection, labeled  $hkl$ .

The problem is that only the intensity, which is the square of the amplitude, can be directly derived from the measured intensity of the diffracted beam. Since only intensities but not phases are measured in the recorded diffraction pattern it is impossible to determine the electron density, and therefore a structure, directly from a recorded diffraction pattern. This is a major problem in crystallography and is referred to as the “*phase problem*”.<sup>30,31</sup>

There are few methods by which the phase problem can be solved. (1) The Patterson summation: This is a Fourier summation based on the experimentally observed  $[F(hkl)]^2$ . It is basically a vector map of the structure and is applied for relatively small molecules. (2) Direct methods: In this method mathematical relationship between the reflections can be used to provide phase information. (3) Heavy atom isomorphous replacement: In this method a heavy atom is introduced to a structure to provide phase information. (4) Anomalous scattering: In this method phase information is obtained from the information contained in the scattering by an atom, which its natural absorption frequency is close to the frequency of the incident radiation. (5) Molecular replacement (MR) method: A known structure will be used to find the phase of the unknown structure. However the use of this method is limited because to have a good chance of success in finding a solution, the search and target molecules must have a reasonable sequence identity (>25%). Likewise, having data with high completeness can be crucial. Most of

the time, but not always, molecular replacement seems to be relatively easier than the other methods and is the first choice in solving the phase problem.

Since two crystal structures of CRABP II have already been solved, CRABP II•RA (PDB ID: 1CBS)<sup>5</sup> and the apo-R111M mutant of CRABP II (PDB ID: 1XCA)<sup>8</sup>, we could easily use molecular replacement to determine the initial phases of each of the structures of CRABP II complexes and its mutants.

Generally there are two steps in molecular replacement known as the rotation and translation functions that will be briefly discussed here.

### Rotation Function

The rotation function should allow the orientation of the search molecule, which produces a maximal overlap with the target structure to be determined in the absence of any phases for the unknown structure. To do this, it compares the Patterson self-vectors of the known and unknown structures at different orientations of the search model. It should be noted that Patterson functions can be calculated from the amplitudes only and using Patterson space means that the translation vector is irrelevant, since all intramolecular vectors are shifted to the origin. The rotation function is usually calculated as a function of Eulerian angles,  $\alpha$ ,  $\beta$  and  $\gamma$ . The molecule is placed in an orthogonal coordinate system with the axis of highest symmetry along Z (about  $\alpha$ ) to reduce the amount of computation.

### Translation Function

Having determined the angles,  $\alpha$ ,  $\beta$  and  $\gamma$ , from the rotation search the rotation matrix can be used and applied to the coordinates of the search molecule. The shift vector, which is required to position the search molecule correctly relative to the symmetry elements of the target molecule, can be determined by one of a number of translation searches. Patterson methods can be used to measure the overlap of the target Patterson cross-vectors with those calculated for the oriented search molecule as it ranges through the target cell. The simpler way to solve the translation problem is the reliability factor (R-factor) search. It involves the calculation of an R-factor as the search molecule and its symmetry mates are moved through the unit cell of the target crystal. The correct position should give the lowest R-factor defined as:

$$R = \frac{\sum ||F_{\text{obs}}| - |F_{\text{cal}}||}{\sum |F_{\text{obs}}|} \quad (2.10)$$

where,  $F_{\text{obs}}$  and  $F_{\text{cal}}$  are the observed and calculated structure factors, respectively. Basically when a crystal structure is known, values of  $F(hkl)$  can be calculated and a way to test the correctness of the structure is how well the calculated values of  $F(hkl)$  agree with the observed ones. Any random collection of atoms in the cell has been shown to result in an R-factor of 83% and 59% for centric and acentric space groups, respectively. Therefore any model that gives R-factors approaching these values is just a little better than a collection of atoms randomly placed in a cell. An R-factor around 45% tells us that the solution is not hopeless but major changes are needed to fit the model. An R-factor of  $\sim 35\%$  is likely to be a correct solution. An R-factor of 25% and below

indicates that the model is most probably correct except for small (1.0 Å) atomic shifts and changes in temperature factors (B-factors).

Other parameters such as the Correlation Coefficient (CC) can be used to measure the agreement between the  $F_{\text{obs}}$  and  $F_{\text{cal}}$  as the search model is moved around. CC can be expressed as:

$$CC = \frac{\sum \left( |F_{\text{obs}}|^2 - \overline{|F_{\text{obs}}|^2} \right) \left( |F_{\text{cal}}|^2 - \overline{|F_{\text{cal}}|^2} \right)}{\left[ \sum \left( |F_{\text{obs}}|^2 - \overline{|F_{\text{obs}}|^2} \right)^2 \sum \left( |F_{\text{cal}}|^2 - \overline{|F_{\text{cal}}|^2} \right)^2 \right]^{\frac{1}{2}}} \quad (2.11)$$

The CC runs from  $-1$  (perfect inverse correlation), through  $0$  (no correlation), to  $+1$  (perfect correlation). It has the advantages of being almost independent of scaling between  $F_{\text{obs}}$  and  $F_{\text{cal}}$ , and is much more sensitive than the R-factor in the region where the R-factor approaches its random limits. Conversely, the coefficient approaches  $1.0$  closely as the R-factor goes below  $20\%$ , and thus becomes of limited value.<sup>32</sup>

If successful, a preliminary model of the target structure will be obtained by correctly orienting and positioning the search molecule in the target cell. Subsequently, this model can be refined and optimized by rigid body refinement.

Since our structures were just point mutations of the previously determined structures we could simply use rigid body refinement, to determine the phases in each structure.

### Structure Refinement

Once one has overcome the phase problem and a solution is found for the structure,  $2F_{\text{obs}} - F_{\text{cal}}$  and  $F_{\text{obs}} - F_{\text{cal}}$  Fourier maps are calculated, and the atomic positions are taken as the locations of the electron density function. The advent of high-speed computers has led to widespread use of the method of least squares, which automatically adjusts the parameters so as to minimize some functions such as  $\sum (F_{\text{obs}} - F_{\text{cal}})^2$ . Each residue will be fitted manually in the structure until the bias introduced by the starting model is reduced considerably. A correct structure should have a satisfactory R-factor with no major unexplained discrepancies.<sup>33</sup> A satisfactory value for the R-factor depends on the resolution of the data. The higher the resolution the lower the R-factor must be. Before refining the model, a fraction of the reflections, usually 5-10%, are randomly chosen and put aside for cross-validation. This set of randomly selected reflections is called the test set. All the fitting of the residues and refinement will be done on the working set (95% or 90% of the data). Then two separate R-factors are calculated. The R-factor calculated for the working set is called “R<sub>work</sub>” or “R<sub>real</sub>” and that of the test set is called “R<sub>free</sub>”. As the R<sub>real</sub> decreases upon model fitting, the R<sub>free</sub> should concurrently decrease. Generally R<sub>free</sub> is higher than R<sub>real</sub>, however they should not differ by more than 6-7%.<sup>31,32,34,35</sup>

Values of R<sub>work</sub> and R<sub>free</sub> are measures of the agreement between the values of the observed structure factors, given by these equations:

$$R_{\text{work}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{cal}}||}{\sum |F_{\text{obs}}|} \quad (2.12)$$

$$R_{\text{free}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{cal}}||}{\sum |F_{\text{obs}}|} \quad (2.13)$$



For the resolution range of our structures (between 1.2-1.8 Å) an R-factor ( $R_{\text{work}}$  or  $R_{\text{real}}$ ) below 20% is acceptable but below 18% is more desirable.<sup>32,33</sup> All the R-factors for the mutant structures that we worked on are between 12-18%, which indicates that the models are well-defined in all structures.

Details about the expression, purification, crystallization and X-ray data collection of each of the mutants are discussed in Chapter IV.

## **2.4 Enhancing the Fluorescence Titration Measurement**

In order to increase the fluorescence emission of the Trp residues inside the pocket and around the chromophore, which will enhance the fluorescence quenching experiments, we decided to incorporate an additional Trp residue in the binding pocket closer to retinal. F15W and L19W mutants were designed, in which the introduced Trp residues are 4.5 and 3.7 Å away from the nearest point on retinal, respectively. The experimental details on mutating these two residues are discussed in Chapter IV. Phe15 and Leu19 are both located on the  $\alpha 1$  helix (residues 15-21). The two single mutants, F15W and L19W led to very poor binding of RA as compared to apo-wild type (apo-WT) CRABP II (Table 2-1). They maximally absorb at similar wavelengths to WT CRABP II. The absorption is similar, because mutation of Phe or Leu to Trp does not change the hydrophobicity of the pocket, significantly.

Table 2-1 The spectroscopic and binding properties of F15W, L19W and apo-WT CRABP II

| Protein       | RA $K_d$<br>(nM) | Retinal $K_d$<br>(nM) | Retinal $\lambda_{max}$<br>(nm) | SB<br>(MALDI) | SB<br>Red. Am. |
|---------------|------------------|-----------------------|---------------------------------|---------------|----------------|
| <b>Apo-WT</b> | $2.0 \pm 1.2$    | $6600 \pm 360$        | 377                             | No            | No             |
| <b>F15W</b>   | $40 \pm 4$       | $5957 \pm 355$        | 372                             | No            | No             |
| <b>L19W</b>   | $170 \pm 22$     | $900 \pm 38$          | 378                             | No            | No             |

We undertook an effort to crystallize both of these mutants to determine whether or not these mutations change the overall structure of the mutants. F15W produced beautiful, orthogonal very well-diffracting crystals, which diffracted to 1.51 Å (see Chapter IV). However attempts in crystallizing L19W only resulted in small and non-diffracting crystals.

#### Crystal Structure of F15W (PDB ID: 2FRS)

This mutant was crystallized at room temperature (25 °C) and in a crystallization condition containing 0.2 M bis-TRIS propane pH 6.5 and 30% PEG 4000. The structure of F15W was determined using the Molecular Replacement (MR) method and the structure of R111M mutant (PDB ID: 1XCA)<sup>8</sup> as the search model. It should be mentioned that rigid body refinement could be used to determine the phases of the structure. However MR produced better R-factors and thus a better solution. The structure was refined at 1.51 Å with crystallographic R-factors of  $R_{real} = 17.40$  and  $R_{free} = 24.00$ . The crystallization, structure solution and refinement details are discussed in Chapter IV.

The crystal structure of the F15W mutant (Figure 3-6 Chapter III) demonstrated that mutation of Phe15 to a Trp completely disrupted the residues in the  $\alpha 2$  helix region of Mol A (residues 25-37). Since Phe15 is a fully conserved residue (100%) in CRABP II, its mutation probably affects the integrity of the structure. More details on the structure and comparing this structure with apo-WT CRABP II are discussed in Chapter III.

Our efforts in producing diffracting quality crystals of L19W mutant were unsuccessful. However since Leu19 is also one of the fully conserved residues in CRABP II, we believe its mutation will also disrupt the structure. Thus further mutants containing the F15W and/or L19W mutations were not pursued since they change the three dimensional structure of the protein. Later we realized that there is no need for enhancement of the fluorescence emission in the vicinity of the chromophore and the results from the fluorescence titration assays are satisfactory.

## **2.5 Design of a Lys Residue Capable of SB Formation**

In order to engineer CRABP II into a mimic of rhodopsin it is necessary to design a Lys residue that is capable of forming a SB with retinal. There are two conditions that need to be met: First the amino group of the putative Lys should be positioned within the *van der Waals* (*vdw*) distance of the carbonyl group of retinal. Second, this amino group should be in a proper direction with respect to the carbonyl plane for the attack to happen. Bürgi and Dunitz through elegant crystallographic experiments, studying amine nucleophiles attacking electrophilic carbonyls, have demonstrated that nucleophiles attack carbonyl groups at a tetrahedral angle ( $105 \pm 5^\circ$ ) rather than perpendicular.<sup>36</sup> This

probably stems from a compromise between the perpendicular approach to the carbonyl, maximizing overlap of the nucleophile HOMO with  $\pi^*$ , and the electronic repulsion between the nucleophile and  $\pi$ -electrons. Therefore one of our guiding principles throughout the reengineering of CRABP II was to apply the same principle, namely, orienting the bound retinal to the active site Lys with a favorable Bürgi-Dunitz trajectory.

Based on our crystal structures and the work done by other groups we knew that Arg111, Arg132 and Tyr134 are the three key residues in binding of RA to CRABP II.<sup>5,8</sup> *In silico* mutagenesis and minimizations of the structure of CRABP II•RA (performed by our collaborators) led us to believe that Arg132 would be an ideal choice to be mutated to a Lys residue which is capable of nucleophilic attack. The amino group of the modeled Lys residue resided  $\sim 3$  Å away from the carbonyl carbon of retinal, modeled in similar position as RA. On the other hand the minimized model indicated that Lys132 can approach the carbonyl carbon of retinal from the side and an almost Bürgi-Dunitz-like trajectory (an angle of  $\sim 135^\circ$ ). Tyr134 was also considered as a possible site to design the Lys residue, however a series of mutants containing the Y134K mutation did not result in a favorable PSB formation (see Dr. Crist's dissertation).<sup>20</sup>

The binding constant of RA is much lower for the R132K mutant than WT ( $K_d$  of apo-WT CRABP II =  $2 \pm 1.2$  nM,  $K_d$  of R132K mutant =  $65 \pm 14$  nM), however, retinal becomes a better binder and its dissociation constant decreases dramatically from  $6600 \pm 360$  nM in apo-WT to  $280 \pm 17$  nM in the R132K mutant. The increased affinity of the R132K mutant toward retinal is either due to SB formation with the designed Lys residue at position 132 or to other considerations.<sup>37,38</sup>

The maximal absorption of the mutant bound to retinal is ~ 379 nm, which is very similar to that of apo-WT (~ 377 nm). Since the R132K mutant did not red shift the absorption of retinal, it must therefore not form a PSB. However, both incubation and reductive amination MALDI experiments verified the SB formation in this mutant. Similar MALDI experiments on apo-WT CRABP II did not show any SB formation between retinal and the protein. Thus the SB in the R132K mutant indeed must be between retinal and the designed Lys residue. For detailed experimental procedures and spectra see Dr. Vasileiou's dissertation.<sup>25</sup> We did not study this mutant by X-ray crystallography.

## **2.6 Hydrophobic Tuning of the Pocket and Design of a Favorable Nucleophilic Attack**

The binding pocket of this mutant was modeled using minimization techniques (see Dr. Vasileiou's dissertation).<sup>25</sup> In this model the only polar residue within 4 Å of the amino group of the engineered Lys residue was Tyr134 (3.39 Å). The hydrophobic index of Tyr and Lys are -1.30 and -3.9, respectively.<sup>37</sup> The more negative this number the more hydrophilic is the residue. Due to the close proximity of Tyr134 to the designed Lys at position 132, the hydroxyl of Tyr134 may make the environment of Lys132 more hydrophilic. This increased hydrophilicity might make Lys132 protonated and therefore a weak nucleophile.

On the other hand, as shown in Figure 2-2-B, the hydroxyl group of Tyr134 makes a direct and tight hydrogen bond interaction with the carboxylate group of RA (2.44 Å). If the same interaction exists between the carbonyl group of retinal and

Tyr134, it might prevent free rotation of retinal to form a favorable Bürgi-Dunitz trajectory with respect to the amino group of Lys132. Thus mutation of Tyr134 to Phe was assumed to help SB formation in two ways: First, by increasing the nucleophilicity of Lys132; and second, by giving retinal freedom to rotate and assume a favorable orientation for the nucleophilic attack. Measuring the binding properties of this mutant verified a significant increase in binding affinity toward retinal ( $K_d = 120 \pm 5$  nM) and a decreased affinity toward RA ( $K_d = 100 \pm 7$  nM), which were both expected because we have mutated Tyr134, a key residue in binding of RA, to a hydrophobic residue. The spectroscopic and binding properties of WT, R132K and R132K:Y134F are shown in Table 2-2.

Table 2-2 The spectroscopic and binding properties of WT and engineered CRABP II mutants

| Protein            | RA $K_d$ (nM) | Retinal $K_d$ (nM) | Retinal $\lambda_{max}$ (nm) | SB formation. |
|--------------------|---------------|--------------------|------------------------------|---------------|
| <b>WT</b>          | $2.0 \pm 1.2$ | $6600 \pm 360$     | 377                          | No            |
| <b>R132K</b>       | $65 \pm 14$   | $280 \pm 17$       | 379                          | Yes           |
| <b>R132K:Y134F</b> | $100 \pm 7.1$ | $120 \pm 4.9$      | 404                          | No            |

However, MALDI assays showed an unexpected result for this mutant. Analysis of the protein sample after incubation with retinal resulted in a very small peak, which was due to the covalent adduct that formed between retinal and Lys132. This data indicated that a very small portion of the bound RA formed a SB. On the other hand, reductive amination of this putative SB followed by MALDI did not support any SB formation. This experiment was repeated many times while varying the reaction

conditions, such as incubation time and amount and reaction time of reducing agent (NaCNBH<sub>3</sub>) used, but the same result was obtained. It was interesting to explain why while R132K mutant formed a SB, R132K:Y134 did not. We were concerned that SB had formed, however it was not stable enough to be detected by MALDI.

UV-vis data also did not indicate PSB formation. The R132K:Y134F mutant bound to retinal (R132K:Y134F•Retinal) maximally absorbed at ~ 404 nm, which is a ~ 25 nm red shift compared to the R132K mutant. This is not a red shift consistent with PSB formation (based on solution studies, formation of a stable PSB leads to ~ 60 nm red shift). So the red shift observed in this mutant must be due to making the pocket more hydrophobic. In fact another mutant of CRABP II, R132L:R111L, is in agreement with this hypothesis. In this mutant both Arg132 and Arg111 are mutated to a hydrophobic residue (Leu). Although there is no engineered Lys residue which can form a PSB, the mutant red shifts the maximum absorption of retinal to ~ 400 nm. In addition, MALDI experiments indicate no evidence of SB formation in this mutant. Therefore these observations suggest that the red shift observed in the R132K:Y134F mutant must be a result of positioning retinal in a more hydrophobic pocket. It is not yet clear how making the pocket more hydrophobic leads to a bathochromic shift.

There have been other reports of red shifted absorption of retinal in the binding pocket of proteins, without SB formation. For example, in the case of cellular retinaldehyde binding protein (CRALBP), which plays a fundamental role in vitamin A metabolism in the retina and retinal pigment epithelium (RPE), a bathochromic shift was observed from 380 to 425 nm. A series of reductive aminations in the presence of retinal failed to produce a covalently bound protein-ligand species by MALDI mass

spectrometry. Thus the observed red shift was attributed to other interactions, such as unspecified electrostatic interactions between retinaldehyde and the hydroxyl-bearing (Thr, Tyr and Ser) and ionizable amino acids inside the pocket of CRALBP.<sup>39-42</sup> A similar red shift was observed in the spectrum of retinol inside the pocket of cellular retinol binding protein (CRBP). CRBP has a very similar structure to CRABP II and also belongs to the iLBP family. Free all-*trans*-retinol absorbs at ~ 325 nm, however when it binds inside the pocket of CRBP the  $\lambda_{\text{max}}$  shifts to ~ 350 nm (a ~ 25 nm red shift). Although the role of electrostatic interactions should not be underestimated, it is believed that retinol assumes a more planar conformation when it is inside the protein pocket than in solution and therefore the polyene is further conjugated.<sup>43,44</sup> Since retinol is an alcohol, unlike retinal, it cannot form a SB with the residues inside the pocket.

The above examples cannot provide a solid explanation for the red shift observed when the chromophore is inside the hydrophobic pocket of the protein but can suggest some possible explanations.

We undertook an effort to crystallize R132K:Y134F bound to all-*trans*-retinal. Contrary to the spectroscopic assays, which can only provide indirect evidence of SB formation, the crystal structure of the mutant bound to retinal can directly depict the binding pocket of the protein and therefore verify whether a SB has formed. Also the crystal structure can help us perform more rational mutations toward designing a mimic of rhodopsin.

Both RA and retinal are light sensitive molecules (retinal is much more light sensitive than RA), therefore all the manipulations and crystallization of RA- and retinal-



bound CRABP<sub>II</sub> were performed in the dark and under red light (see Chapter IV for more details).

Our initial attempts in crystallizing the retinal-bound CRABP<sub>II</sub> were unsuccessful. However we successfully crystallized both the apo- and RA-bound R132K:Y134F mutant to gain an insight into the binding pocket of this mutant and also to study the effects of these two mutations on the overall structure of the protein. The apo-structure of this mutant has a very similar overall structure to apo-WT CRABP<sub>II</sub> (RMSDs between Mol A's and Mol B's are 0.168 and 0.343 Å, respectively), which indicates that these two mutations do not change the overall fold of the structure. More details about the apo-structures are discussed in Chapter III.

#### Crystal Structure of R132K:Y134F Bound to RA (PDB ID: 2G78)

The structure of R132K:Y134F bound to RA (R132K:Y134F•RA) was determined and refined at 1.70 Å with good crystallographic R-factors of  $R_{\text{work}} = 14.50\%$  and  $R_{\text{free}} = 20.28\%$ . The structure is in the  $P2_12_12_1$  space group, has one molecule in the asymmetric unit and has a unit cell almost identical to that of CRABP<sub>II</sub>•RA. The data collection and refinement statistics are reported in Tables 4-4 and 4-6 in Chapter IV, respectively. This structure, superimposed on the structure of WT CRABP<sub>II</sub>•RA, is shown in Figure 2-7-A, which shows that the structure of R132K:Y134F•RA is almost identical to that of CRABP<sub>II</sub>•RA (RMSD = 0.182 Å). This similarity between the mutant and WT indicates that the mutations do not change the integrity of the structure.

As shown in Figure 2-7-B, interestingly also in this structure, we observed a very similar water-mediated network that connects Arg111 to Ala36, located on the loop

connecting  $\alpha 2$  to  $\beta B$ . This observation suggests the importance of Arg111 and water-mediated interactions in the structural integrity of CRABP II. The path of the water-mediated network is as follow: guanidine of Arg111  $\rightarrow$  water 11  $\rightarrow$  carboxylate group of RA  $\rightarrow$  amino group of Lys132  $\rightarrow$  carbonyl of Ala36. Unlike WT CRABP II•RA and apo-WT structure, Arg111 does not have water-mediated interaction with Val33, on the  $\alpha 2$  helix. This suggests that possibly interaction with Ala36 is more determinant for the structure.

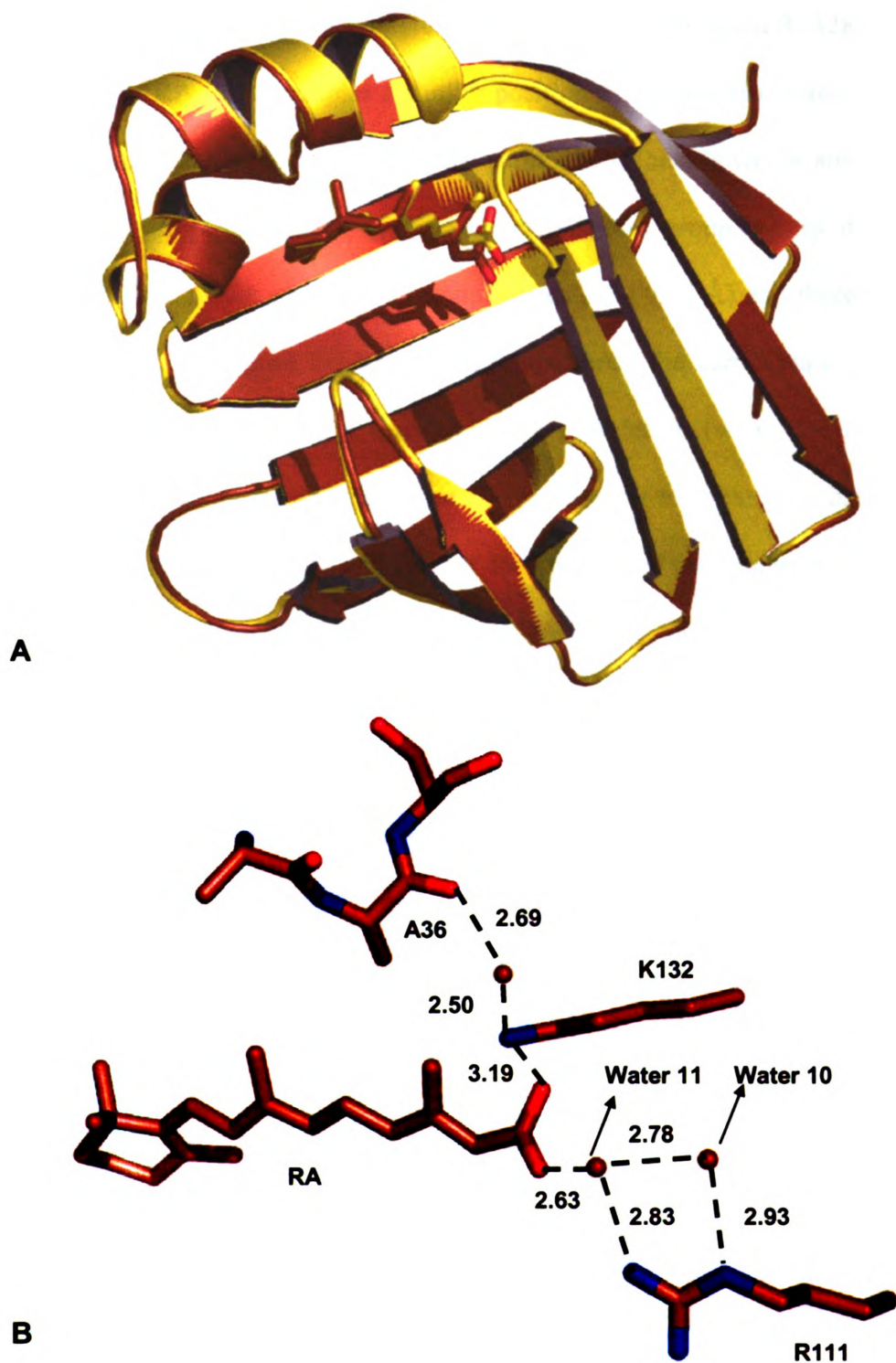


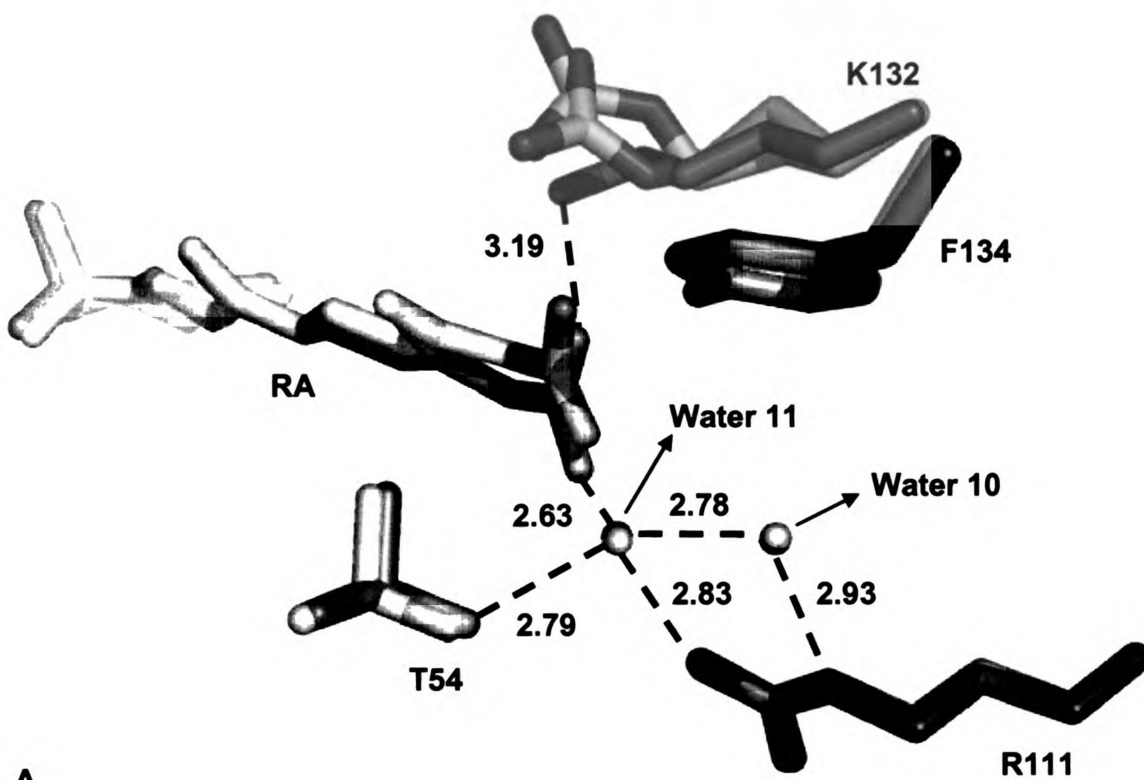
Figure 2-7 (A) Overlaid structures of R132K:Y134F•RA (brown) and WT CRABP II•RA (yellow). (B) Water-mediated interaction between Arg111 and Ala36, located on the loop connecting  $\alpha 2$  to  $\beta B$ , in R132K:Y134F•RA. The distances are in angstroms.

Figure 2-8 shows an overlay picture of WT CRABP II•RA and R132K:Y134F•RA in the vicinity of RA. RA occupies a similar position in the two structures. The ionone ring and most of the backbone occupy very similar positions, however since Tyr134 is mutated to a hydrophobic residue, Phe, the carboxylate group moves down toward Arg111 (the carbon of the carboxylate group moves by  $\sim 0.92$  Å) and therefore tilts the backbone by  $\sim 5.23^\circ$  (Figure 2-8-B). In both structures, the carboxylate group of RA interacts with Arg111 and Thr54 through a water molecule (water 11 in the double mutant). This water mediated interaction is very similar to what we observed in the structure of WT CRABP II•RA, and in fact the two water molecules that mediate the hydrogen bond interaction between Arg111 and RA occupy identical positions (Figure 2-8-A). In the double mutant the carboxylate oxygen of RA is further from N<sup>ε</sup> of Lys132 (3.19 Å, Figure 2-8-A) than it is to the guanidino group of Arg132 in WT (2.67 Å, Figure 2-2-B). Possibly due to the more hydrophobic pocket of the double mutant compared to WT, the amino group of Lys132 is not protonated in the double mutant and, therefore, has a weaker interaction. The angle formed between the N<sup>ε</sup> of Lys132 and the carbonyl is  $52.99^\circ$  for the carboxylate oxygen pointing towards Lys132 and  $162.86^\circ$  for the carboxylate oxygen pointing away.

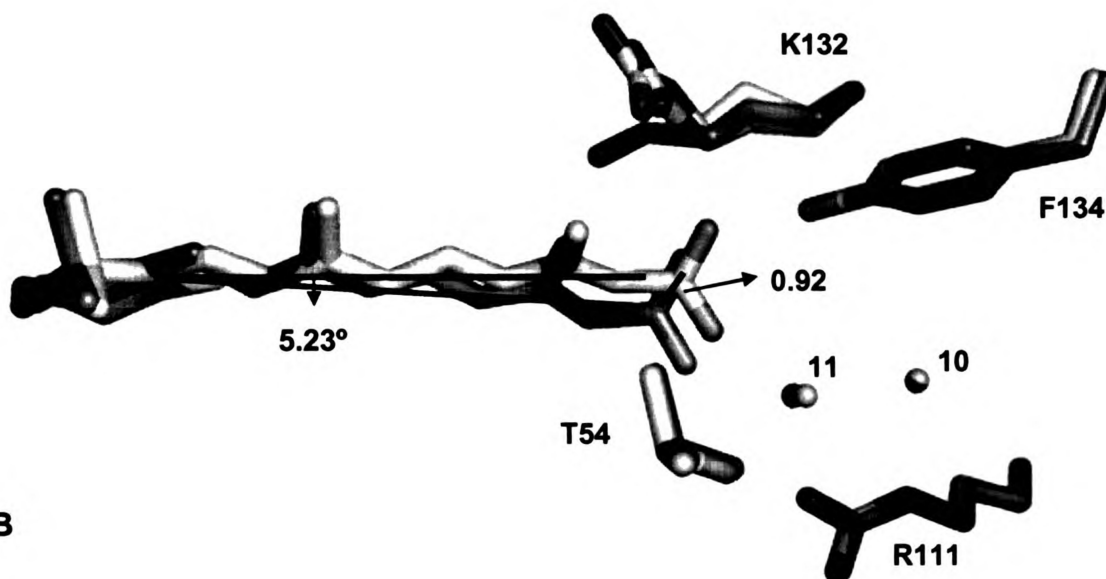
One can expect that since retinal has a carbonyl group instead of a carboxylate it will only interact with Arg111 and Thr54 through a water molecule and not with Lys132. This interaction will position the carbonyl group of retinal a large distance away from N<sup>ε</sup> of Lys132 and also in an unfavorable Bürgi-Dunitz trajectory, which will prevent SB formation. Also, if retinal is bound to Arg111 and Thr54, by forming tight hydrogen

bonds as RA does, it will not be free to rotate and form a favorable position for the nucleophilic attack by Lys132.

As further discussed in Chapter IV, we learned that retinal in the protein solution, not only is light sensitive but also, is temperature sensitive. As a result, we successfully crystallized R132K:Y134F•Retinal in the dark (under red light) and at 4 °C.



A



B

Figure 2-8 (A) RA and the neighboring residues in the pocket of R132K:Y134F•RA overlaid on WT CRABP II•RA. The distances are in angstroms. (B) RA in the double mutant tilts by 5.23° and the carbonyl carbon moves by 0.92 Å. Only residues of the double mutant are labeled in each picture.

### Crystal Structure of R132K:Y134F Bound to Retinal (PDB ID: 2G79)

The structure of R132K:Y13F bound to retinal (R132K:Y134F•Retinal) was determined and refined at 1.69 Å with crystallographic R-factors of  $R_{\text{work}} = 15.52\%$  and  $R_{\text{free}} = 21.37\%$ . The unit cell is virtually identical to that of the previously bound structures. The data collection and refinement statistics are reported in Tables 4-4 and 4-6 in Chapter IV, respectively. The overall structure is almost identical to the structures of WT CRABP II•RA (RMSD = 0.449 Å) and R132K:Y134F•RA (RMSD = 0.373 Å). The overlaid structures of the R132K:Y134F•Retinal and WT CRABP II•RA are shown in Figure 2-9-A.

This structure shed light on whether retinal is bound as a SB or free retinal inside the pocket. As discussed earlier, reductive amination of the complex followed by MALDI suggests that when retinal binds inside the pocket, it does not form a SB. However this was an unexpected result because the mutant has a designed Lys residue which is expected to form a SB with retinal, as the R132K mutant does. Figure 2-9-B shows retinal within its unbiased Fo-Fc omit electron density map contoured at 2  $\sigma$ . The high resolution map clearly shows that retinal does not form a SB, and in fact is in agreement with the MALDI experiments.

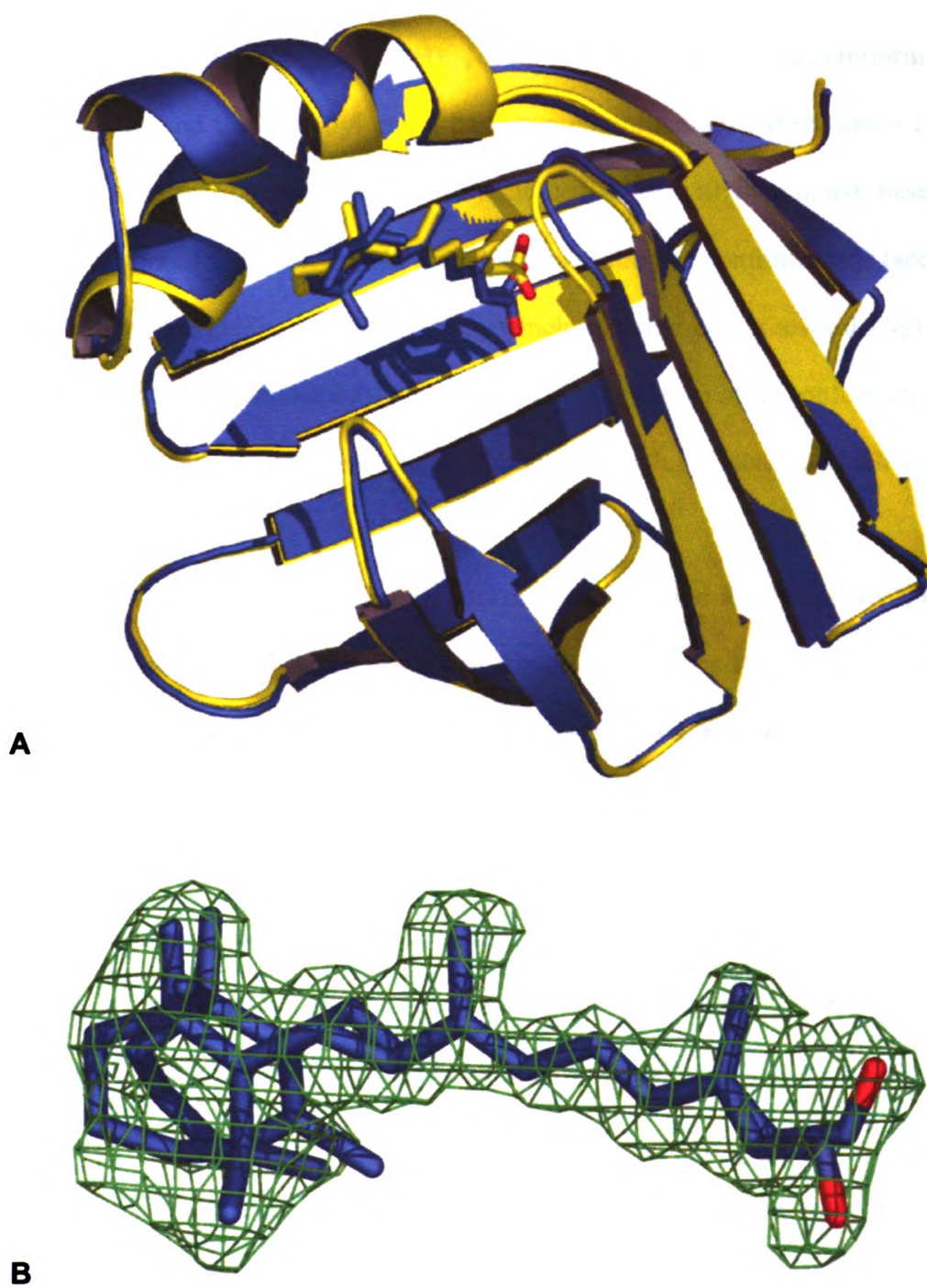


Figure 2-9 (A) Overlaid structures of R132K:Y134F-Retinal (blue) and WT CRABP II-RA (yellow); (B) Retinal inside the pocket of R132K:Y134F-Retinal within its Fo-Fc omit electron density map contoured at 2.0  $\sigma$



Interestingly the Fo-Fc omit map of retinal shows two unevenly occupied conformations for retinal (Figure 2-9-B). Most of the backbone has one conformation but the ionone ring and the carbonyl group have two conformations, which are  $\sim 180^\circ$  from each other. The two conformations were built as 70% and 30% occupied, based on the overall R-factors of the structure and the B-factors of the two conformations (see Chapter IV for more details). The conformation in which carbonyl points toward Lys132 (30% occupied) makes a  $70^\circ$  trajectory with N $^\epsilon$  of Lys132, while the other conformation makes a  $146^\circ$  angle. Lys132, itself, has two conformations in this structure, which seem to be evenly occupied (50%).

Figure 2-10 shows the binding pocket of retinal- and RA-bound R132K:Y134F mutant. As shown, retinal occupies a very similar position as RA in this mutant. This observation suggests that our assumption that retinal and RA occupy a very similar position in the same mutant, is a valid one. As we had predicted from the RA-bound structure of this mutant, the carbonyl group of retinal makes a bond to Arg111 and Thr54 through a water molecule (water 10 in R132K:Y134F•Retinal), as RA does in the pocket of R132K:Y134F (and WT). The two water molecules that are bound to Arg111 and mediate the interactions between Arg111 and the chromophore occupy identical positions in the two structures (Figure 2-10).

Note that the double mutant does not form a SB with the Lys residue in this mutant presumably because the carbonyl cannot adopt a favorable orientation for nucleophilic attack. The N $^\epsilon$  of Lys 132 is 3.55 Å distant from the oxygen of the carbonyl group of retinal (and 3.76 Å distant from the carbonyl carbon) and therefore can no longer form a tight hydrogen bond with retinal. As mentioned in the structure of

R132K:Y134F•RA, since the pocket is more hydrophobic in the double mutant compared to WT, the amino group of Lys132 is probably not protonated and cannot interact with the carbonyl group of retinal. Therefore, the majority of the time, the carbonyl of retinal is pointing towards Arg111 and forms a tight hydrogen bond with it through a water molecule (water 10). Therefore Lys132 cannot perform a nucleophilic attack on the carbonyl carbon of retinal in this mutant because, for a majority of the time, retinal is positioned in an unfavorable position for nucleophilic attack.

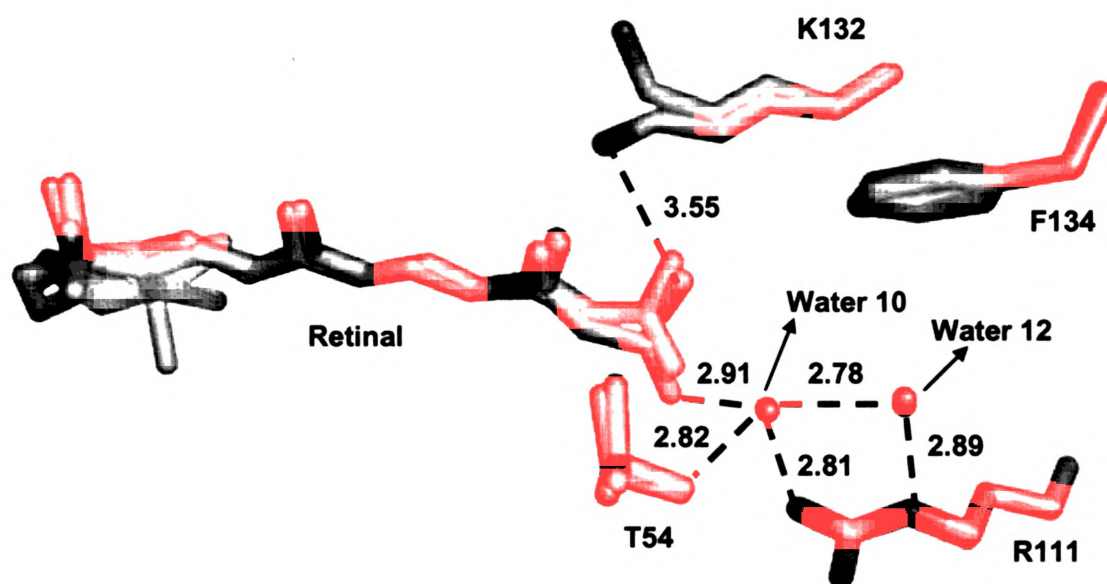


Figure 2-10 Retinal and neighboring residues to its carbonyl group inside the binding pocket of R132K:Y134F•Retinal (blue) overlaid on R132K:Y134F•RA (brown). Only residues of R132K:Y134•Retinal are labeled. The distances are in angstroms and show the hydrogen bond distances in R132K:Y134F•Retinal.

The overlaid structures of R132K:Y134F•Retinal and R132K:Y134F•RA are shown in Figure 2-11-A. The  $\alpha 2$  helix and the loop connecting this helix to  $\beta B$  have identical conformation in the two structures, as observed in apo-WT, apo-R132K:Y134F and WT CRABP II•RA. A similar water-mediated interaction, as observed in the previous structures, connects Arg111 to Ala36 on the loop region. The path of the network is as follow: Arg111  $\rightarrow$  water 10  $\rightarrow$  carbonyl of retinal  $\rightarrow$  N $^{\epsilon}$  of Lys132  $\rightarrow$  carbonyl group of Ala36 (Figure 2-11-B).

It is clear from the latter two crystal structures that the plane of the carbonyl has to be rotated ( $\sim 45^{\circ}$  counterclockwise, considering the chromophore with the carbonyl oxygen pointing toward Lys132) along the long axis of the chromophore in order to achieve a favorable trajectory for nucleophilic attack.

## **2.7 Facilitation of the Nucleophilic Attack and Design of a Counter Anion for the PSB**

The crystal structure of the double mutant bound to retinal indicates that Arg111 hydrogen bonds to retinal and positions it in an unfavorable position for the nucleophilic attack by Lys132 (the N $^{\epsilon}$  of Lys132 is 3.76 Å away from the carbonyl carbon of retinal in the 30% occupied conformation). Therefore we decided to remove the unfavorable interaction of retinal with Arg111 and Thr54 by removing water 10 (in R132K:Y134F•Retinal structure, Figure 2-10). Water 10 was eliminated by mutating Arg111 to Leu (the R132K:Y134F:R111L mutant) and Thr54 to Val resulting in the tetra mutant, R132K:Y134F:R111L:T54V.

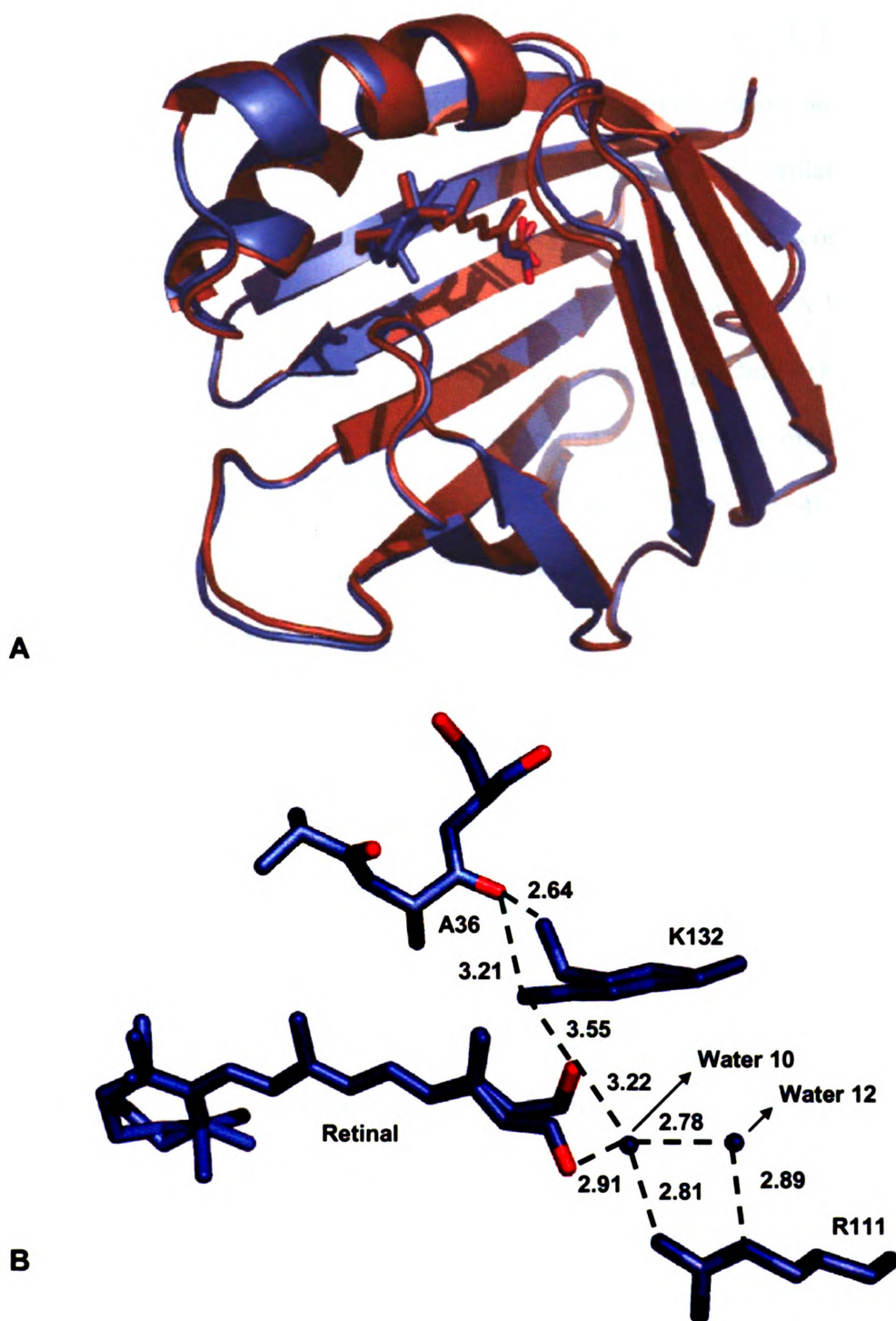


Figure 2-11 (A) Overlaid structures of R132K:Y134F•Retinal (blue) and R132K:Y134F•RA (brown). (B) Water-mediated interaction between Arg111 and Ala36, on the loop connecting  $\alpha 2$  to  $\beta B$ , in R132K:Y134F•Retinal. The distances are in angstroms.

The R132K:Y134F:R111L (or triple) and R132K:Y134F:R111L:T54V (or tetra) mutants were studied using spectroscopic assays. The triple mutant bound to retinal maximally absorbs at  $\sim 400$  nm. The maximum absorption is very similar to that of the double mutant. RA binding is down 10 fold ( $K_d = 1000 \pm 28$  nM) compared to the double mutant, which is expected because Arg111 a key residue in RA binding and is mutated to a hydrophobic residue. However the retinal binding affinity ( $K_d = 160 \pm 6.7$  nM) does not change significantly as compared to that of the double mutant. Since the triple mutant has a very similar  $\lambda_{max}$  to the double mutant (R132K:Y134F) we conclude that it must also be unable to form a PSB. However, both incubation and reductive amination MALDI experiments verified SB formation in this mutant.

The tetra mutant, R132K:Y134F:R111L:T54V absorbed maximally at  $\sim 400$  nm, similar to double and triple mutants. The binding of RA did not change considerably compared to the triple mutant ( $K_d = 900 \pm 64$  nM), which indicates that Arg111 plays a more important role in RA binding compared to Thr54. However, retinal is a two fold better ligand ( $K_d = 84 \pm 12$  nM) compared to the triple mutant, which must be due to the more hydrophobic pocket resulting from the T54V mutation. Again a PSB formation was not indicated by a red shift in the UV-vis data, however both incubation and reductive amination MALDI experiments verified SB formation in this mutant.

As mentioned above, our goal is to form a stable PSB between retinal and Lys132 to mimic PSB formation in the binding pocket of rhodopsin. As shown in Figure 2-1, the active site Lys residue in rhodopsin, Lys296, is surrounded by hydrophobic residues and is not in close contact with any polar residue with the exception of Glu113 and Ser186. These two residues are 3.22 Å and 3.66 Å away from the nitrogen of PSB, respectively.

Glu113 plays the important role of being the counter anion for the PSB, stabilizing its positive charge. Although the previously discussed mutants verified SB formation, none of them indicated any PSB formation between retinal and Lys132. Therefore we decided to introduce a negatively charged residue close to the SB to stabilize a positive charge on the nitrogen of the SB.

As previously mentioned, in rhodopsin the position of the counter anion can change the maximum absorption of the chromophore. Several positions were considered for the placement of the PSB counter anion. Minimized models were studied in which the counter anion (Glu) was placed in positions 54, 121, and 111, the residues around the SB. The models suggested that position 121, which is 2.59 Å away from the SB nitrogen, is an ideal position for placing the counter anion. Since we believe retinal will rotate to assume a favorable Bürgi-Dunitz trajectory with respect to N<sup>ε</sup> of Lys132, positioning of the counter anions at the other two positions will place them at a far distance from the PSB (see Dr. Vaileious's dissertation)<sup>25</sup>.

Besides being a counter anion, Glu121 may play an important role in positioning retinal in a favorable orientation for the nucleophilic attack. As shown in the crystal structure of R132K:Y134F•Retinal in Figure 2-12, Leu121 is positioned midway between Lys132 and Arg111. Therefore Glu121 interaction with the oxygen of the carbonyl can position retinal in a closer position to Lys132 and lead to a more favorable Bürgi-Dunitz trajectory, necessary for nucleophilic attack. In fact the importance of the L121E mutation in positioning retinal for a favorable nucleophilic attack was confirmed by the R132K:Y134F:L121E mutant. Although the R132K:Y134F mutant did not form a SB with retinal, mutation of Leu121 to Glu restored the ability of the protein to form a SB.

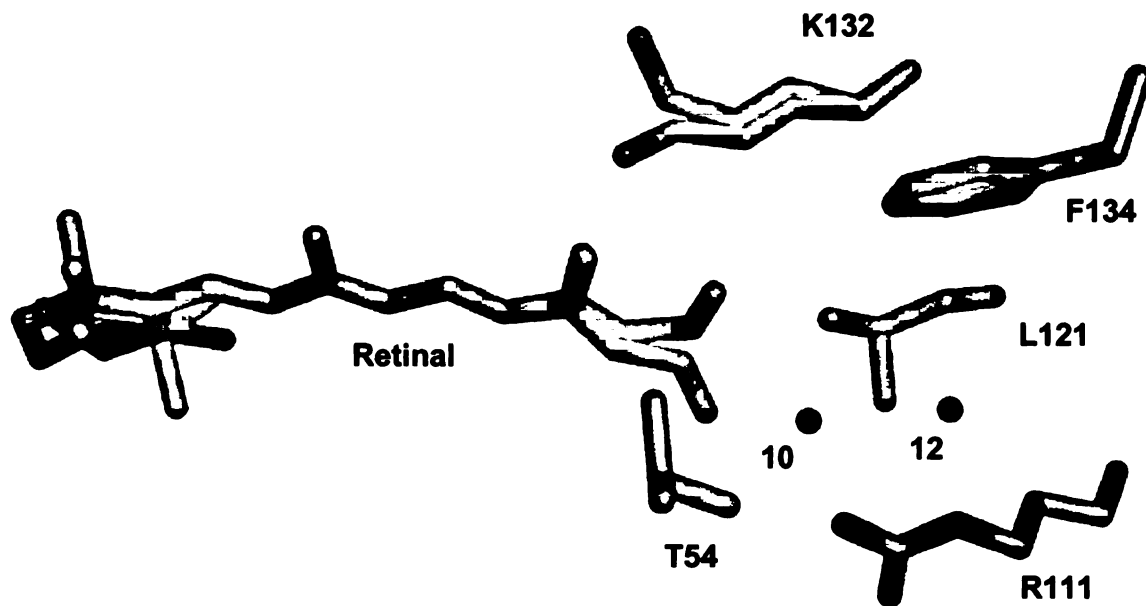


Figure 2-12 Binding pocket of R132K:Y134F•Retinal. Leu121 is located midway between Lys132 and Arg111, making it a good choice for placing the counter anion.

The resulting penta mutant, R132K:Y134F:R111L:T54V:L121E, when bound to retinal, maximally absorbs at 382 and 446 nm. The significant amount of red shift is a strong indication that a PSB has formed between retinal and Lys132. However the mixture of the two peaks can be attributed to a mixture of protonated and non-protonated SB. This observation suggests that Glu121 possibly is not a strong counter anion for the PSB. Both incubation and reductive amination MALDI experiments verified the SB formation, which is in agreement with what we expected based on UV-vis data that suggest PSB formation. Retinal binds to this mutant with a very high binding affinity ( $K_d = 2.7 \pm 7$  nM), which is almost 30 times better than the tetra mutant (R132K:Y134F:R111L:T54V). This must be due to facilitation of SB formation by positioning Glu in position 121 and stabilization of the putative PSB. The penta mutant becomes a slightly better binder toward RA ( $K_d = 250 \pm 19$  nM) upon introduction of the L121E mutation. This

observation could be explained from the crystal structure of the penta mutant bound to RA and will be discussed later in this chapter. The summarized spectroscopic and binding properties of the mutants that we discussed thus far are shown in Table 2-3.

Table 2-3 The spectroscopic and binding properties of CRABPII and its mutants toward engineering the R132K:Y134F:R111L:T54V:L121E

| <b>Protein</b>                      | <b>RA <math>K_d</math><br/>(nM)</b> | <b>Retinal <math>K_d</math><br/>(nM)</b> | <b>Retinal<br/><math>\lambda_{max}</math> (nm)</b> | <b>SB<br/>formation.</b> |
|-------------------------------------|-------------------------------------|------------------------------------------|----------------------------------------------------|--------------------------|
| <b>WT</b>                           | $2.0 \pm 1.2$                       | $6600 \pm 360$                           | 377                                                | No                       |
| <b>R132K</b>                        | $65 \pm 14$                         | $280 \pm 17$                             | 379                                                | Yes                      |
| <b>R132K:Y134</b>                   | $100 \pm 7.1$                       | $120 \pm 4.9$                            | 404                                                | No                       |
| <b>R132K:Y134F:R111L</b>            | $1000 \pm 28$                       | $160 \pm 6.7$                            | 400                                                | Yes                      |
| <b>R132K:Y134F:R111L:T54V</b>       | $900 \pm 64$                        | $84 \pm 12$                              | 400                                                | Yes                      |
| <b>R132K:Y134F:R111L:T54V:L121E</b> | $250 \pm 19$                        | $2.7 \pm 7.0$                            | 446                                                | Yes                      |

We undertook an effort to crystallize the penta mutant bound to retinal. The structure could first of all depict the SB formation between retinal and Lys132 and also give us an insight into the interactions inside the pocket of the mutant. Unfortunately all attempts in co-crystallizing the penta mutant bound to retinal were unsuccessful. We performed the crystallization experiments in the dark, both at room temperature and at 4 °C. 1, 2, 5 and 10 equivalents of retinal were tried. Both ~ 10 and ~ 20 mg/mL protein concentrations were tried.

Since co-crystallization experiments did not produce any crystals, we tried soaking apo crystals of the penta mutant in retinal solutions. We tried different soaking



times and different equivalents of retinal (as explained in chapter IV). The soaking experiments were performed in the dark and at 4 °C. Data were collected for the soaked crystals and structures were determined for numerous crystals with various soaking times. None of the soaked crystals showed any evidence of retinal inside the pocket. However, interestingly, soaking significantly improved the resolution of the crystals. Although the highest resolution crystals of apo-penta mutant diffracted to 1.60 Å, soaking improved the resolution to 1.20 Å. Therefore one of the best diffracting soaked crystals was used to determine the structure of the apo-penta mutant. This structure is discussed in Chapter III. Since our efforts in crystallizing the penta mutant bound to retinal were unsuccessful, we decided to crystallize the mutant bound to RA to gain an insight into the designed binding pocket. We could successfully crystallize and determine the structure of the RA-bound mutant.

#### Crystal Structure of R132K:Y134F:R111L:T54V:L121E bound to RA (PDB ID: 2G7A)

The structure of R132K:Y134F:R111L:T54V:L121E bound to RA (R132K:Y134F:R111L:T54V:L121E•RA) was determined and refined at 1.57 Å and crystallographic R-factors of  $R_{\text{work}} = 12.17\%$  and  $R_{\text{free}} = 16.81\%$ . The structure is in the  $P2_12_12_1$  space group and has one molecule in the asymmetric unit. The data collection and refinement statistics are reported in Tables 4-4 and 4-6 in Chapter IV, respectively. The overall structure is almost identical to that of WT CRABP II•RA (RMSD = 0.492 Å) and R132K:Y134F•RA (RMSD = 0.318 Å). The overlaid structures of the penta mutant bound to RA and WT CRABP II•RA are shown in Figure 2-13-A.

As shown in Figure 2-13-B, RA occupies a similar position in the penta mutant as in the WT protein. The ionone ring and most of the backbone of the chromophore occupy identical positions, however the carboxylate group moves toward Glu121 tilting the backbone by 5.90° and moving the carboxylate carbon by 1.27 Å (Figure 2-13-B). This motion is due to different interactions than what was observed in the pocket of WT. In the latter, RA hydrogen bonds to Arg111, Lys132 and Tyr134. However in the penta mutant these three residues are mutated to hydrophobic residues.

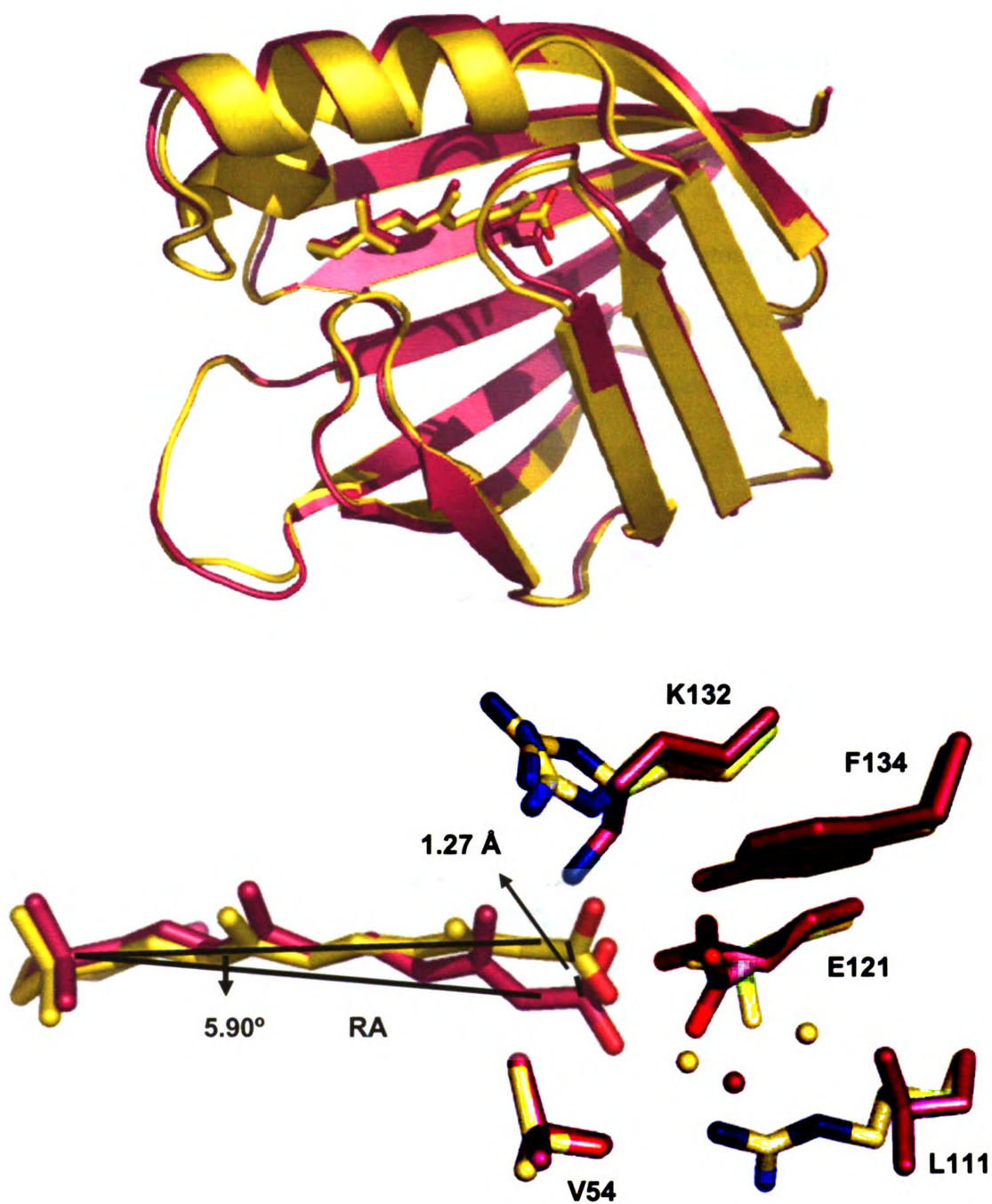


Figure 2-13 (A) The structure of R132K:Y134F:R111L:T54V:L121E-RA (magenta) superimposed on the structure of WT CRABPII-RA (yellow). (B) RA and the neighboring residues to its carboxylate group in the pocket of the penta mutant (magenta) and WT CRABPII-RA (yellow). Only the residues of the penta mutant are labeled.

Interestingly the crystal structure of the RA-bound penta mutant showed that RA binds to this protein by forming a tight dicarboxylic acid interactions with Glu121 (Figure 2-14). This observation could explain the increased affinity of RA for the penta mutant ( $K_d = 250 \pm 19$ ), as compared to the tetra mutant, R132K:Y134F:R111L:T54, ( $K_d = 900 \pm 64$ ). As shown in Figure 2-14, the carboxylate group of Glu121 has two conformations that each are 50% occupied. The oxygens of the two carboxylate groups are  $\sim 2.6$  Å away from each other, at their shortest distances. This tight dicarboxylic

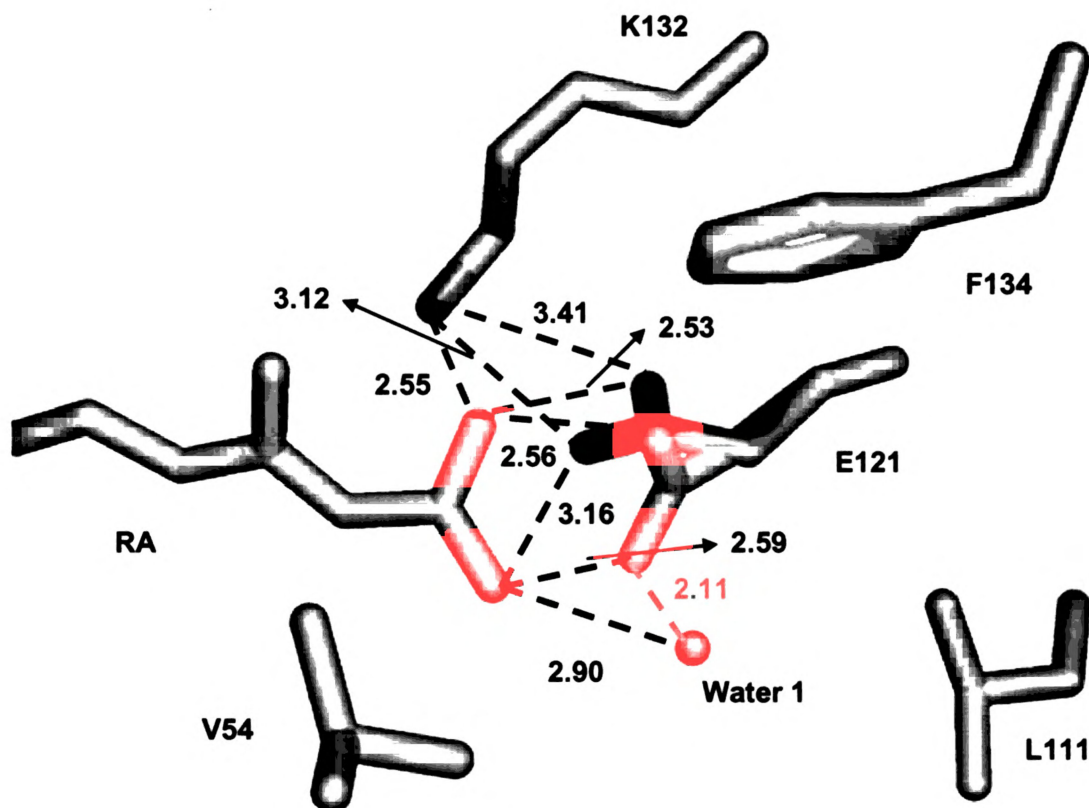


Figure 2-14 The binding pocket of R132K:Y134F:R111L:T54V:L121E•RA. RA and Glu121 make tight dicarboxylic acid interactions with each other, indicating that both groups are protonated. Distances are hydrogen bond distances in angstroms.

interaction indicates that both RA and Glu121 must be protonated. This is not to our surprise because the pocket of the penta mutant is highly hydrophobic. Considering the fact that a hydrophobic environment cannot tolerate charges, it is expected that both carboxylate groups become protonated in this pocket. Previous reports on similar cases have shown that within a hydrophobic cavity the  $pK_a$  of a Glu residue can be increased by 5 units.<sup>45</sup>

As shown in Figures 2-13-B and 2-14, the hydrogen bond interaction observed between Arg111 and RA is broken upon R111L mutation. A 50% occupied water molecule forms a hydrogen bond with the carboxylate group of RA. Probably the other half of the time this water molecule is disordered and that is why it cannot be seen in the crystal structure as a fully occupied water. This water cannot be fully occupied because it is positioned too close to one of the conformations of the carboxylate group of Glu121. The distance of 2.11 Å, shown in red, does not exist and water 1 will be disordered when the carboxylate moves toward it.

In this structure, unlike other structures, Lys132 has moved toward the RA and Glu121 and makes direct hydrogen bonds with both of them. Therefore Glu121 seems to help facilitate the nucleophilic attack by positioning Lys132 in a closer proximity of retinal and also possibly by protonating the oxygen of the carbonyl group of retinal. Besides, since Glu121 is positioned midway between Lys132 and Arg111 (Figure 2-12), its interaction with the oxygen of the carbonyl group of retinal through a hydrogen bond (as in the RA-bound structure) could rotate the carbonyl plane in a more favorable arrangement to achieve the necessary Bürgi-Dunitz trajectory.

As will be discussed in Chapter III, comparing the apo-mutants of CRABP II in which Arg111 is mutated into a hydrophobic residue shows that this mutation destabilizes the structure, particularly at the  $\alpha 2$  helix and the loop connecting this helix to  $\beta B$ . However, although Arg111 is mutated to Leu, the structure of the penta mutant bound to RA shows an almost identical structure to WT bound to RA. Investigating the interactions inside the pocket shows a water-mediated interaction between the designed Glu121 and Ala36 (the first residue on the loop connecting  $\alpha 2$  to  $\beta B$ ), and also Val33 (on the  $\alpha 2$ ), as observed in other structures between Arg111 and the two latter residues. The water-mediated interactions are as follow: carboxylate of Glu121  $\rightarrow$  carboxylate of RA  $\rightarrow$  N $^{\epsilon}$  of Lys132  $\rightarrow$  water 189  $\rightarrow$  water 62  $\rightarrow$  carbonyl of Ala36; and carboxylate of Glu121  $\rightarrow$  carboxylate of RA  $\rightarrow$  N $^{\epsilon}$  of Lys132  $\rightarrow$  water 189  $\rightarrow$  water 62  $\rightarrow$  water 65  $\rightarrow$  carbonyl of Val33 (Figure 2-15).

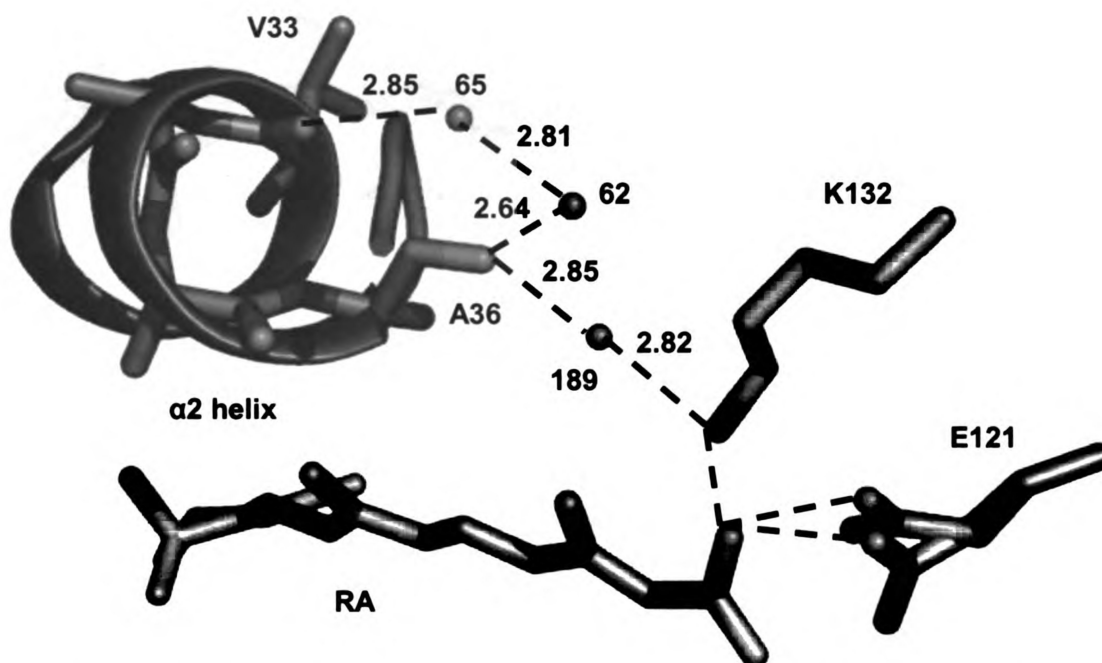


Figure 2-15 Binding pocket of the penta mutant bound to RA: A water-mediated interaction connects Glu121 to the residues on the  $\alpha 2$  helix region and stabilizes them. The distances are in angstroms.

Figure 2-16 shows the superimposed binding pockets of the penta mutant bound to RA and R132K:Y134F bound to retinal. Interestingly the overlay shows that retinal in the double mutant occupies almost an identical position as RA in the penta mutant. Residues of both mutants are labeled. Lys 132 has two conformations in the double mutant, however in the penta mutant since Leu121 is mutated to a Glu, Lys132 moves toward Glu121 to form a hydrogen bond with it and therefore assumes only one conformation (see Figure 2-14 for the hydrogen bond distances). The overlaid structures show that L121E and R111L mutations do not change the position of the chromophore.

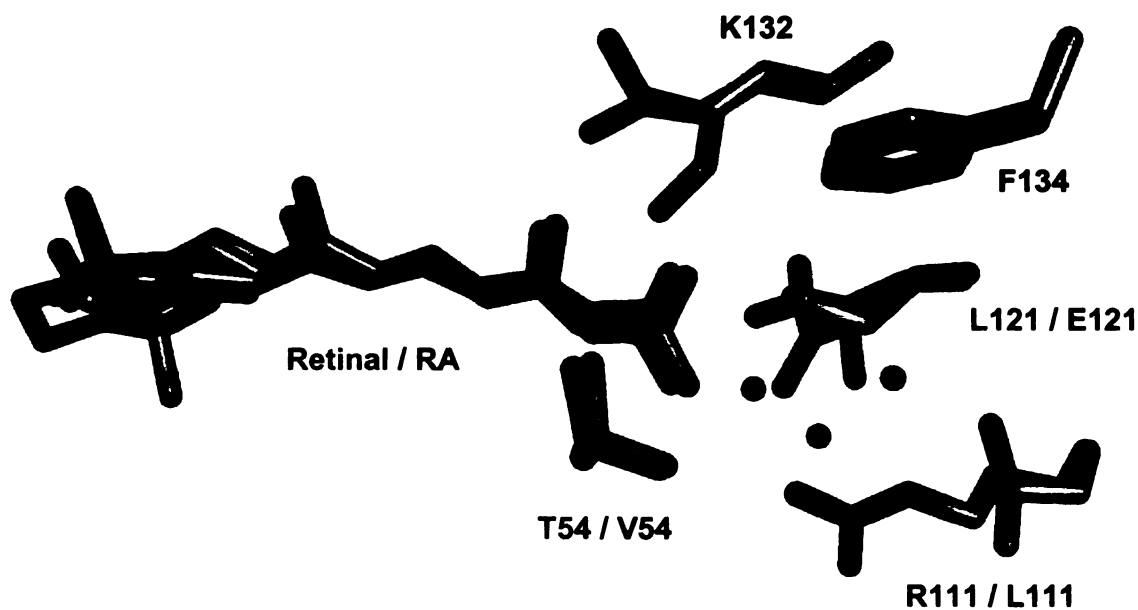


Figure 2-16 R132K:Y134F:R111L:T54V:L121E•RA (magenta) superimposed on R132K:Y134F•Retinal (blue). Residues of both structures are labeled.

## **2.8 Restoring Tyr134: Enhancing Formation of a Counter Anion, Orienting Retinal for a Favorable Nucleophilic Attack**

Although Glu121 was proven essential for PSB formation, the mixture of the peaks in the UV-vis absorption of the penta mutant bound to retinal, at 382 and 466 nm, indicates that Glu121 is not a strong counter anion for the putative PSB. In addition the crystal structure of the mutant bound to RA suggests that Glu121 is at a protonated state in this pocket. However, in order to act as a good counter anion Glu121 needs to be deprotonated. Therefore both the structural and UV-vis data indicate that possibly, in the penta mutant, the pocket has become “too hydrophobic” in the vicinity of Glu121.

As apparent from the crystal structure of WT CRABP II•RA, Tyr134 is in close proximity of Leu121 (Figure 2-2-B), and therefore if it is restored in the pocket of the penta mutant would also be in close proximity to Glu121. Based on this analogy and in order to increase the hydrophilicity of the pocket to help deprotonation of Glu121, we decided to restore Tyr at position 134.

The idea of restoring Tyr134 at position 134 was pursued for an additional reason. As mentioned earlier, although the R132K mutant can form a SB with retinal, R132K:Y134F does not form a SB because upon Y134F mutation retinal is positioned from Lys132 and interacts with Arg111 (Figure 2-10). As the crystal structure of R132K:Y134F•Retinal (Figure 2-10) shows, the motion of retinal toward Arg111 will position retinal in an unfavorable position for the nucleophilic attack by Lys132. SB formation was restored through additional mutations resulting in the penta mutant, capable of forming a PSB. As shown in the superimposed crystal structures of WT and R132K:Y134F bound to RA (Figure 2-8), Lys132 and Tyr134 are in close proximity of



each other. Therefore Tyr134 could possibly position retinal in a favorable position for the nucleophilic attack, by positioning the electrophile and nucleophile next to each other. In addition, Tyr134 may help protonation of the carbonyl for the nucleophilic attack by Lys132.

Figure 2-17 shows the overlaid structures of WT CRABPII•RA, R132K:Y134F•Retinal and penta mutant bound to RA (from WT CRABPII•RA only Tyr134 and from R132K:Y134F•Retinal only retinal are shown. The rest of the residues are from the penta mutant bound to RA). As discussed earlier, retinal in the double

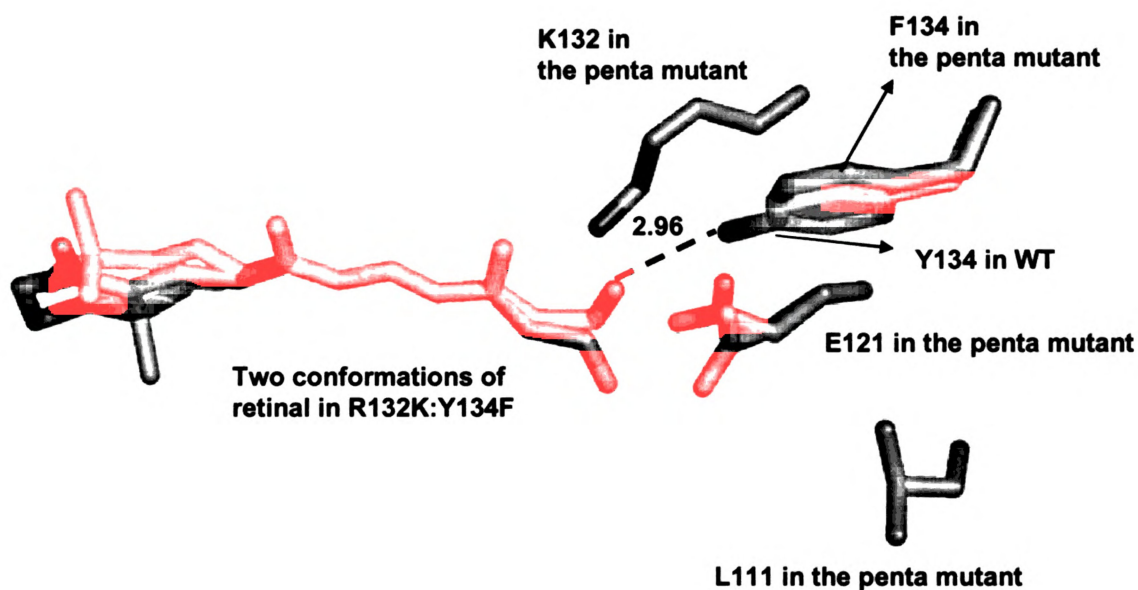


Figure 2-17 Superimposed structures of WT CRABPII•RA (yellow), R132K:Y134F•Retinal (blue and green) and R132K:Y134F:R111L:T54V:L121E•RA (magenta). The overlaid structures show that if Tyr134 is restored in the pocket of the penta mutant retinal, it will most probably assume a conformation similar to the green conformation of retinal in the double mutant and hydrogen binds to Tyr134 as RA does in WT CRABPII•RA.

mutant occupies two unevenly occupied conformations. For clarity the conformations facing toward and away of Lys132 are shown in orange and blue, respectively. As shown in Figure 2-17, if retinal is positioned in the pocket of the penta mutant while Tyr is restored at position 134, most probably it assumes a position similar to the orange conformation and forms a direct and tight hydrogen bond with Tyr134, as RA does in WT CRABP II (Figure 2-2-B). Based on these superimposed structures, the carbonyl of retinal will be positioned 2.96 Å away from the phenolic group of Tyr134.

Therefore using the knowledge acquired from our previous mutations, a new series of mutants, that restores Tyr at position 134, was produced. We refer to these proteins as the “Tyr134 series” to distinguish them from the old series (Y134F series). Tables 2-2 and 2-3 summarize the spectroscopic and binding properties of Y134F and Tyr134 series, respectively.

As shown in Table 2-4, the triple (R132K:R111L:L121E) and the tetra (R132K:R111L:L121E:T54V) mutants are very similar in their binding and spectroscopic properties. Since the tetra mutant does not show any advantage over the triple mutant it appears that Thr54 does not play a major role in the binding and absorption properties of the Tyr134 series.

Table 2-4 The spectroscopic and binding properties of Y134F and Tyr134 series. The Y134F series are shown in black and the Tyr134 series are shown in red.

| Protein                      | RA $K_d$<br>(nM) | Retinal $K_d$<br>(nM) | Retinal<br>$\lambda_{max}$ (nm) | SB<br>formation |
|------------------------------|------------------|-----------------------|---------------------------------|-----------------|
| WT                           | $2.0 \pm 1.2$    | $6600 \pm 360$        | 377                             | No              |
| R132K                        | $65 \pm 14$      | $280 \pm 17$          | 379                             | Yes             |
| R132K:Y134F                  | $100 \pm 7.1$    | $120 \pm 4.9$         | 404                             | No              |
| R132K:Y134F:R111L            | $1000 \pm 28$    | $160 \pm 6.7$         | 400                             | Yes             |
| R132K:R111L                  | $736 \pm 114$    | $567 \pm 36$          | 401                             | Yes             |
| R132K:Y134F:R111L:T54V       | $900 \pm 64$     | $84 \pm 12$           | 400                             | Yes             |
| R132K:Y134F:R111L:T54V:L121E | $250 \pm 19$     | $2.7 \pm 7.0$         | 446                             | Yes             |
| R132K:R111L:L121E:T54V       | $400 \pm 36$     | $1.99 \pm 4.1$        | 449                             | Yes             |
| R132K:R111L:L121E            | $426 \pm 47$     | $1.36 \pm 4.9$        | 449                             | Yes             |

The Tyr134 series has a significant advantage over the Y134F series. Although the penta mutant (R132K:Y134F:R111L:T54V:L121E) bound to retinal from the Y134F series shows a mixture of two peaks in its UV-vis absorption spectrum, which indicates that the SB is at an equilibrium between protonated and non-protonated states, the R132K:R111L:L121E and R132K:R111L:L121E:T54V from the new series (Tyr134 series) show a single red shifted peak. Similarly, R132K:Y134F:R111L:L121E•Retinal maximally absorbs at 378 and 438 nm (data not shown in the table), while R132K:R111L:L121E•Retinal shows one red shifted peak at 449 nm (Figure 2-18). This is a major improvement toward making a PSB. These results indicate that restoring

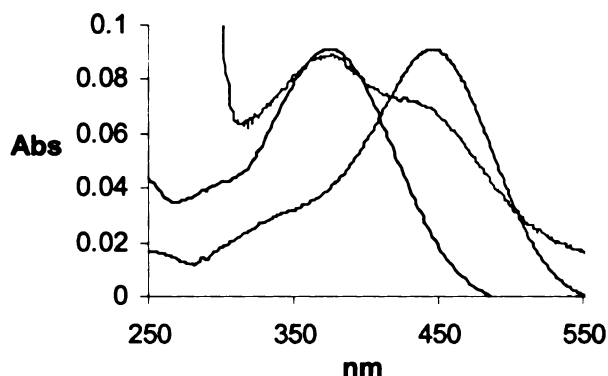


Figure 2-18 UV-vis spectra of retinal bound to WT CRABP II (blue trace,  $\lambda_{\text{max}} = 377$  nm), R132K:Y134F:R111L:L121E (magenta trace,  $\lambda_{\text{max}} = 378, 438$  nm) and R132K:R111L:L121E (red trace,  $\lambda_{\text{max}} = 449$  nm).

Tyr134 is critical in enhancing the polarity of the pocket in the vicinity of the PSB, which can result in deprotonation of Glu121 and therefore formation of a strong counter anion for the PSB. See Dr. Vasileious's dissertation for more details on the spectroscopic data.<sup>25</sup>

Since the PSB in the R132K:R111L:L121E mutant is stable at physiological pH, the apparent  $\text{pK}_a$  of the PSB was determined using an acid/base pH titration. The  $\text{pK}_a$  value was determined to be 8.7 which was 2 units higher than that of R132K:Y134F:R111L:T54V:L121E•Retinal ( $\text{pK}_a = 6.5$ ), verifying the positive effect of Tyr134 on the stabilization of the PSB.

Although the estimated  $\text{pK}_a$  of 8.7 for the PSB in the triple mutant is significantly lower than that of rhodopsin ( $\text{pK}_a \sim 16$ )<sup>13,46,47</sup>, the fact that the PSB formed between Lys132 and retinal is stable at physiological pH allows us to use the R132K:R111L:L121E as a potential mimic of rhodopsin, in order to test the proposed hypotheses on the wavelength regulation mechanism.

The  $\text{pK}_a$  of the triple mutant compares better with what is observed for invertebrate rhodopsins (PSB  $\text{pK}_a$  of octopus rhodopsin = 10.4; PSB  $\text{pK}_a$  of gecko cone opsin = 9.9).<sup>48,49</sup> It has been reported that this drop in the  $\text{pK}_a$  value is due to the lack of a

carboxylate counter anion in the vicinity of the SB. A Tyr residue has been located close to the SB in octopus rhodopsin, with spectroscopic results indicating that the Tyr is not the anionic form.<sup>50</sup>

In conclusion, comparison of R132K:Y134F:R111L:T54V:L121E•Retinal ( $pK_a = 6.5$ ) with R132K:R111L:L121E•Retinal ( $pK_a = 8.7$ ), further supports that restoring Tyr134 helps the deprotonation of Glu121, which therefore becomes a more efficient counter anion. The  $pK_a$  was measured using spectroscopic acid/base titration of the mutant and can only give a rough estimate of the  $pK_a$ . The ideal method to measure the  $pK_a$  of the PSB is by using NMR spectroscopy. However this technique requires the use of specifically labeled reagents and/or protein.<sup>51</sup> The measured  $pK_a$  values of Lys residues buried inside the protein pockets lie within a wide range of 5.5-8.5, depending on the neighboring residues.<sup>45,52-54</sup>

#### The Crystal Structure of R132K:R111L:L121E Bound to Retinal (PDB ID: 2G7B)

We decided to crystallize R132K:R111L:L121E bound to retinal (R132K:R111L:L121E•Retinal) to verify the SB formation between retinal and Lys132 and also to depict the interactions of the residues inside the binding pocket. Based on what we learned from crystallizing R132K:Y134F•Retinal, since retinal is both light and temperature sensitive all the manipulations and crystallizations of retinal bound to the protein were performed in the dark and at 4 °C.

The crystals diffracted to extraordinarily high resolution for macromolecular structures (1.10 Å). The structure was determined and refined at 1.18 Å and with R-factors of  $R_{work} = 12.85\%$  and  $R_{free} = 15.54\%$ . The structure is in the  $P2_12_12_1$  space group

with one molecule in the asu, similar to other holo-structures that we discussed in this document. The data collection and refinement statistics are reported in Tables 4-4 and 4-6 in Chapter IV, respectively. The overall structure is similar to the structures of WT CRABP II•RA (RMSD = 0.705 Å). The superimposed structures of R132K:R111L:L121E•Retinal and WT CRABP II•RA are shown in Figure 2-19-A.

The two structures are very similar with most of the differences concentrated in the  $\alpha 2$  helix region. The RMSD of residues 26-42 ( $\alpha 2$  and the loops connecting this helix to  $\alpha 1$  and  $\beta B$ ) between the two structures is 0.989 Å. Excluding these residues from the RMSD calculations between the two structures results in an RMSD of 0.472 Å. The helix in the triple mutant is well defined and the secondary structure calculations (using DSSP program)<sup>55</sup> shows that like other holo-structures residues 26-36 comprise the second helix. However the helix shows some motions compared to that in WT structure. The maximum xyz deviation between the structures lies at Ala36 (1.764 Å). This residue is the one that has water-mediated interaction with Arg111 (or Glu121) in the previous structures. In this structure Ala36 is not making any water-mediated interaction with Glu121. Possibly this is the reason for the motions observed in the  $\alpha 2$  helix in this structure. None of the residues on the  $\alpha 2$  helix region interact with Glu121 or its neighboring residues.

As further discussed in Chapter III, we have also determined the structure of the apo-R132K:R111L:L121E. Interestingly similar to R132K:R111L:L121E•Retinal, the  $\alpha 2$  helix has a different conformation from all other structures in which Arg111 is not mutated and even different from the penta mutant (R132K:Y134F:R111L:T54V:L121E). Interestingly, in the apo-R132K:R111L:L121E mutant, unlike other apo-structures, and

similar to R132K:R111L:L121E•Retinal, Ala36 does not have any water-mediated interaction with Glu121. The C-terminus of the second helix in this structure is completely unwound and Ala36 moves significantly compared to that in the WT structure. The secondary structure calculation using DSSP shows that the helix is shorter in this structure (residues 26-33) and that the last helical turn is lost. Different arrangements of water molecules in this pocket should be the reason for the different interactions. In apo-R132K:R111L:L121E, a different water-mediated interaction connects Glu121 to Ala32 and Ala34 on the second helical turn of  $\alpha 2$ , stabilizing the helix at this new conformation. More details on the apo-structure and the paths of the water networks are discussed in Chapter III.

The unbiased omit Fo-Fc electron density map of retinal contoured at  $2.2 \sigma$  is shown in Figure 2-19-B. Thanks to the high resolution of the data (1.18 Å) we observed a very well defined omit electron density map, which clearly shows a *cis* SB (imine) bond between retinal and Lys132. Although MALDI and UV-vis data had suggested the SB formation, the crystal structure directly depicted the SB formation.



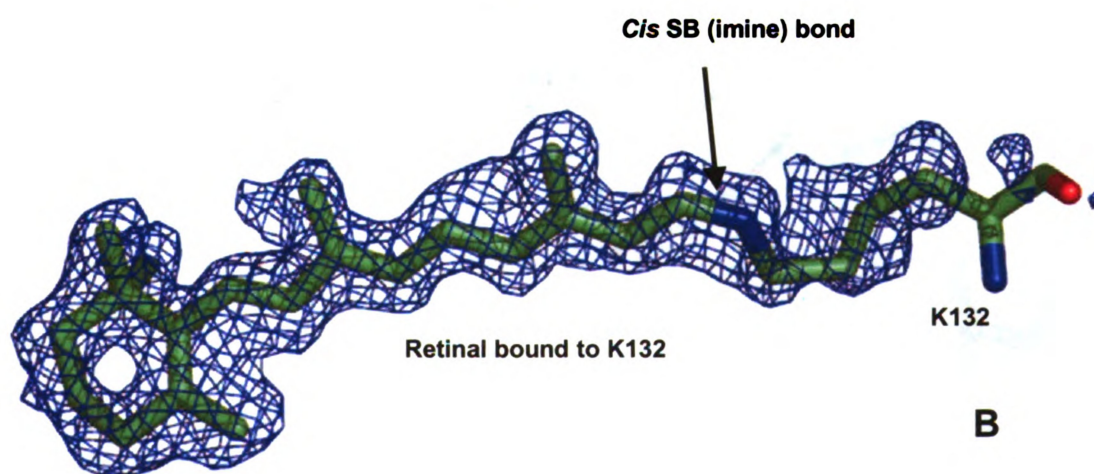


Figure 2-19 (A) R132K:R111L:L121E•Retinal (green) superimposed on WT CRABP II•RA (yellow).

(B) The Fo-Fc omit electron density map of retinal contoured at 2.2  $\sigma$ .



Figure 2-20 shows the binding pocket of R132K:R111L:L121E•Retinal. Glu121, is positioned 2.55 Å away from the nitrogen of the PSB. This verifies that indeed Glu121 plays the important role of being the counter anion for the PSB. Tyr134 is 2.77 Å away from Glu121 and therefore makes a direct and tight hydrogen bond interaction with it. Therefore Tyr134 must be protonated in this pocket. Because of the direct interaction of Tyr134 with Glu121, it can reduce the  $pK_a$  of Glu121 or, in other words, deprotonate it. This is consistent with the UV-vis data for the R132K:R111L:L121E•Retinal, which reveals one single red shifted peak at 449 nm vs. the mixture of the peaks for R132K:Y134F:R111L:T54V:L121E•Retinal, in which Glu121 is probably only partially protonated. Tyr134 itself is 3.75 Å away from the SB nitrogen and, therefore, does not make any hydrogen bond interaction with the SB nitrogen.

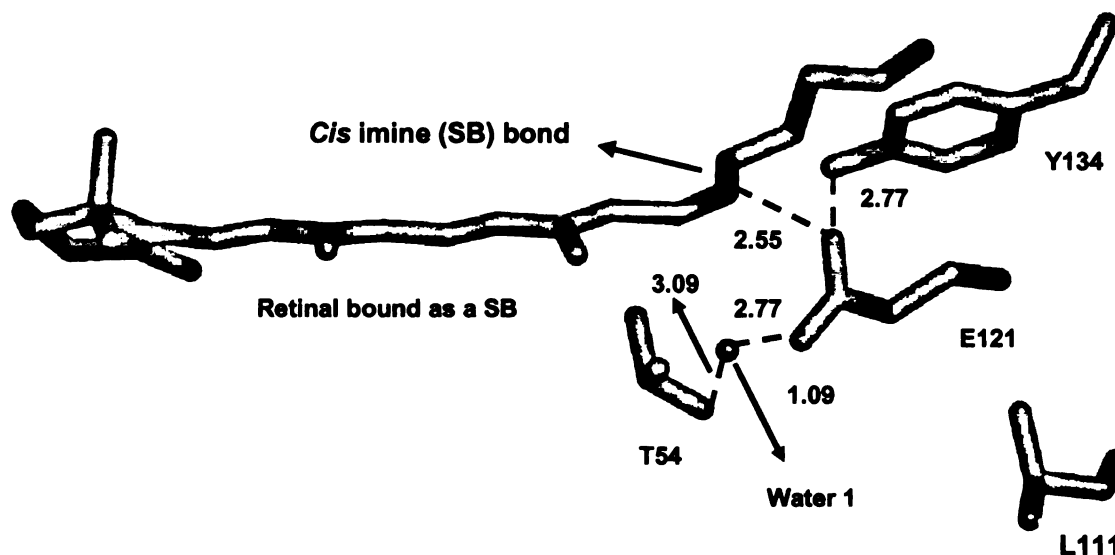


Figure 2-20 Binding pocket of R132K:R111L:L121E•Retinal. The distances are in angstroms.

In addition the structure shows that the ordered water molecules that mediate the interaction between RA (Figures 2-2, 2-7) or retinal (Figure 2-10) move upon R111L mutation. One of the water molecules (water 12 in R132K:Y134F•Retinal) moves 1.79 Å toward Thr54 and is positioned between Thr54 and Glu121 (Figure 2-21) while the other one (water 10 in R132K:Y134F•Retinal) is not observed. Therefore this structure suggests that removal of the ordered water molecule between retinal and Arg111, by the R111L mutation, aids in SB formation. This gives retinal freedom to bind Tyr134, rather than Arg111, which positions retinal in a more favorable position for the nucleophilic attack by Lys132 (Figure 2-17).

The SB that forms is a *cis* imine bond, while that of visual rhodopsin is a *trans* imine. Looking at the pocket shows that Met123 is in the *vdw* distance of the PSB and has multiple conformations in this structure. The conformations are not well defined, which suggests that Met123, which is well defined in other structures, has destabilized by retinal. Possibly mutation of Met123 to a smaller residue will prevent the steric hindrance between retinal and Met123.

## **2.9 Backbone Rotation: Favoring the Nucleophilic Attack**

The overlaid structures of R132K:Y134•Retinal and R132K:R111L:L121E•Retinal are shown in Figure 2-21. The two conformations of retinal in R132K:Y134F•Retinal are shown in two different colors; orange for the conformation in which the carbonyl oxygen faces toward Lys132 and blue for the conformation that the carbonyl oxygen faces away. Since in the triple mutant Arg111 is mutated to a hydrophobic residue while Tyr134 is maintained, most probably when

retinal enters the pocket it occupies a conformation similar to the orange conformation, hydrogen binding with Tyr134 (Figure 2-21). The overlaid structures suggest a distance of 2.90 Å for the hydrogen bond that can form between Tyr134 and the carbonyl group of retinal. This conformation of retinal is also in agreement with our observation for RA in the pocket of WT CRABP II•RA (Figure 2-2-B). Lys132 also has two conformations in R132K:Y134F•Retinal, which are shown in blue and orange. The orange conformer is the one which has favorable orientation for the nucleophilic attack on the carbonyl carbon of retinal.

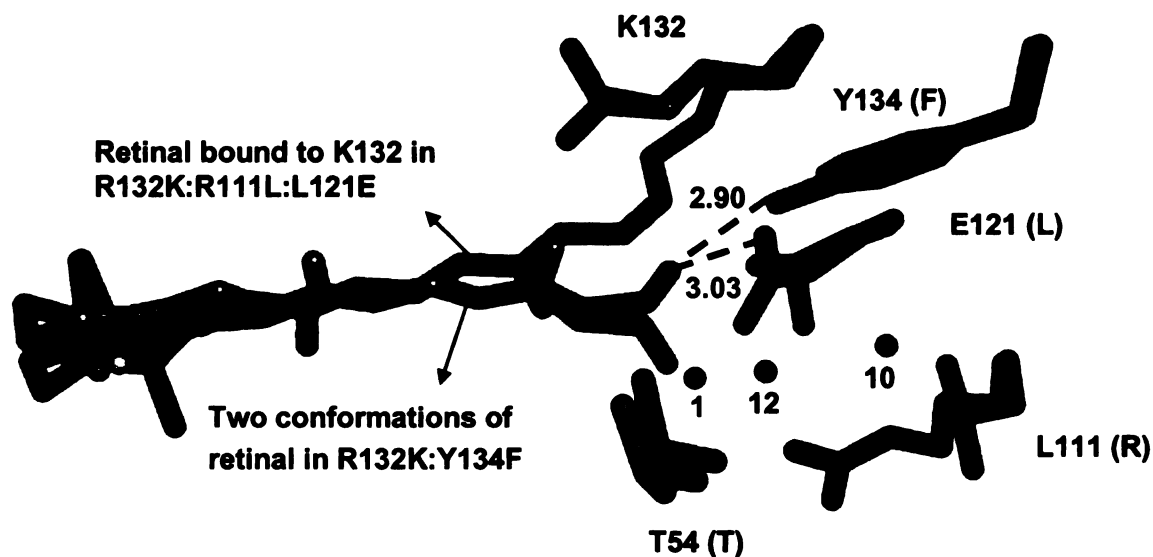


Figure 2-21 R132K:R111L:L121E•Retinal (green) superimposed on R132K:Y134F•Retinal (blue). The labels of the double mutant residues, are shown in parentheses. The water molecules are shown in the same color as each molecule. The distances are in angstroms.

Comparing the position of retinal in the two molecules shows that the ionone ring occupies a very similar position in the two structures, while the backbone has rotated by ~ 45°, counterclockwise, along its long axis and around the single bond that connects the

ionone ring to the backbone (Figure 2-22). In fact it is this backbone rotation which positions retinal in a favorable Bürgi-Dunitz trajectory with respect to Lys132 and makes the nucleophilic attack possible. The carbonyl carbon is 2.30 Å away from the N<sup>ε</sup> of Lys132 and N<sup>ε</sup> forms a ~ 107° trajectory with the plane of the carbonyl. In fact the 45° backbone rotation completely superimposes the free retinal in R132K:Y134F•Retinal on the bound retinal in the triple mutant (R132K:R111L:L121E•Retinal), as shown in Figure 2-22.

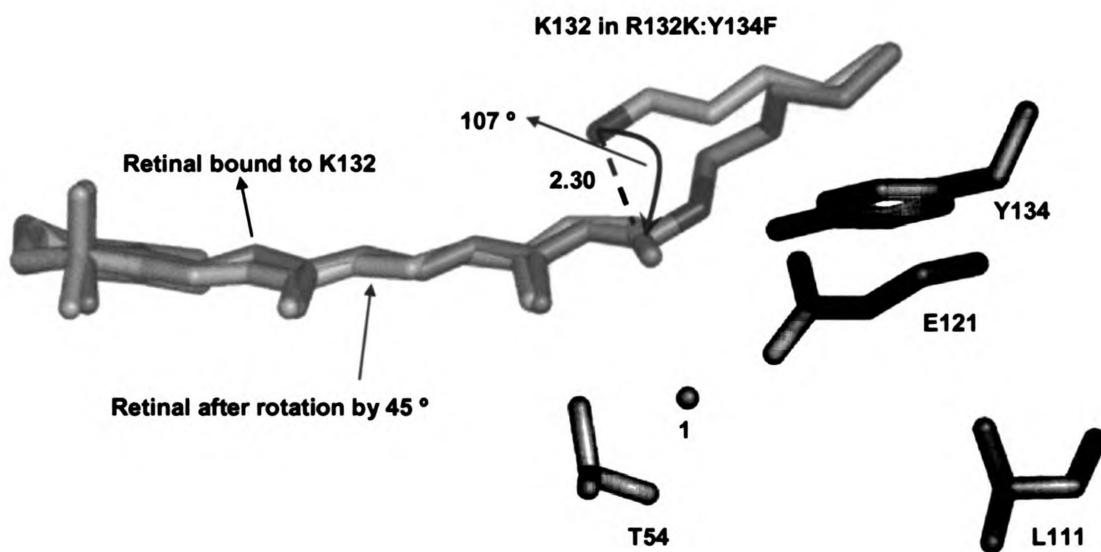


Figure 2-22 Rotation of the free retinal in R132K:Y134F•Retinal (orange) by 45 ° around the single bond that connects the ionone ring to the backbone and along the long axis of retinal positions retinal in a favorable Bürgi-Dunitz trajectory with respect to Lys132. This rotation leads to complete superimposition of free retinal on the bound retinal in the structure of R132K:R111L:L121E•Retinal.

## **2.10 Proposal of a PSB Formation Mechanism in R132K:R111L:L121E Bound to Retinal**

Based on the previous crystal structures we can propose a mechanism for PSB formation between retinal and the R132K:R111L:L121E mutant (Figure 2-23). When retinal enters the pocket, it will hydrogen bond to Tyr134, as RA does in WT CRABP<sub>II</sub>•RA. As shown in the overlaid structures of R132K:Y134F•Retinal with R132K:R111L:L121E•Retinal (Figure 2-21), the carboxylate group of Glu121 is within hydrogen bond distance of the oxygen of the carbonyl group of retinal. In fact the overlaid structures of WT CRABP<sub>II</sub>•RA also shows that the oxygen of the carboxylate of RA is within hydrogen bond distance of Glu121 (picture not shown). As shown in Figure 2-21 the carbonyl oxygen also hydrogen bonds to Tyr134 (2.90 Å). To activate the carbonyl for a nucleophilic attack either Tyr134 or Glu121 will protonate the carbonyl oxygen. Since Glu121 has a lower  $pK_a$  than Tyr134 at the physiological pH ( $pK_a$  of Glu = 4.4,  $pK_a$  of Tyr = 10)<sup>56</sup>, most probably in the hydrophobic pocket, like that of the R132K:R111L:L121E•Retinal, Glu121 is still more acidic than Tyr134 and protonates the carbonyl. Besides based on the crystal structure of the R132K:R111L:L121E•Retinal we know that Glu121 must be deprotonated, because it is within 2.55 Å of the PSB (Figure 2-20). Either at the same time or subsequent to this protonation, Lys132 would perform a nucleophilic attack on the carbonyl carbon of retinal. The carbonyl oxygen would get a hydrogen atom from the amino group of Lys132 and water elimination occurs and finally a PSB forms. As shown by the crystal structure of R132K:R111L:L121E•Retinal, Glu121 is 2.55 Å away and plays the important role of being the counter anion for the PSB.

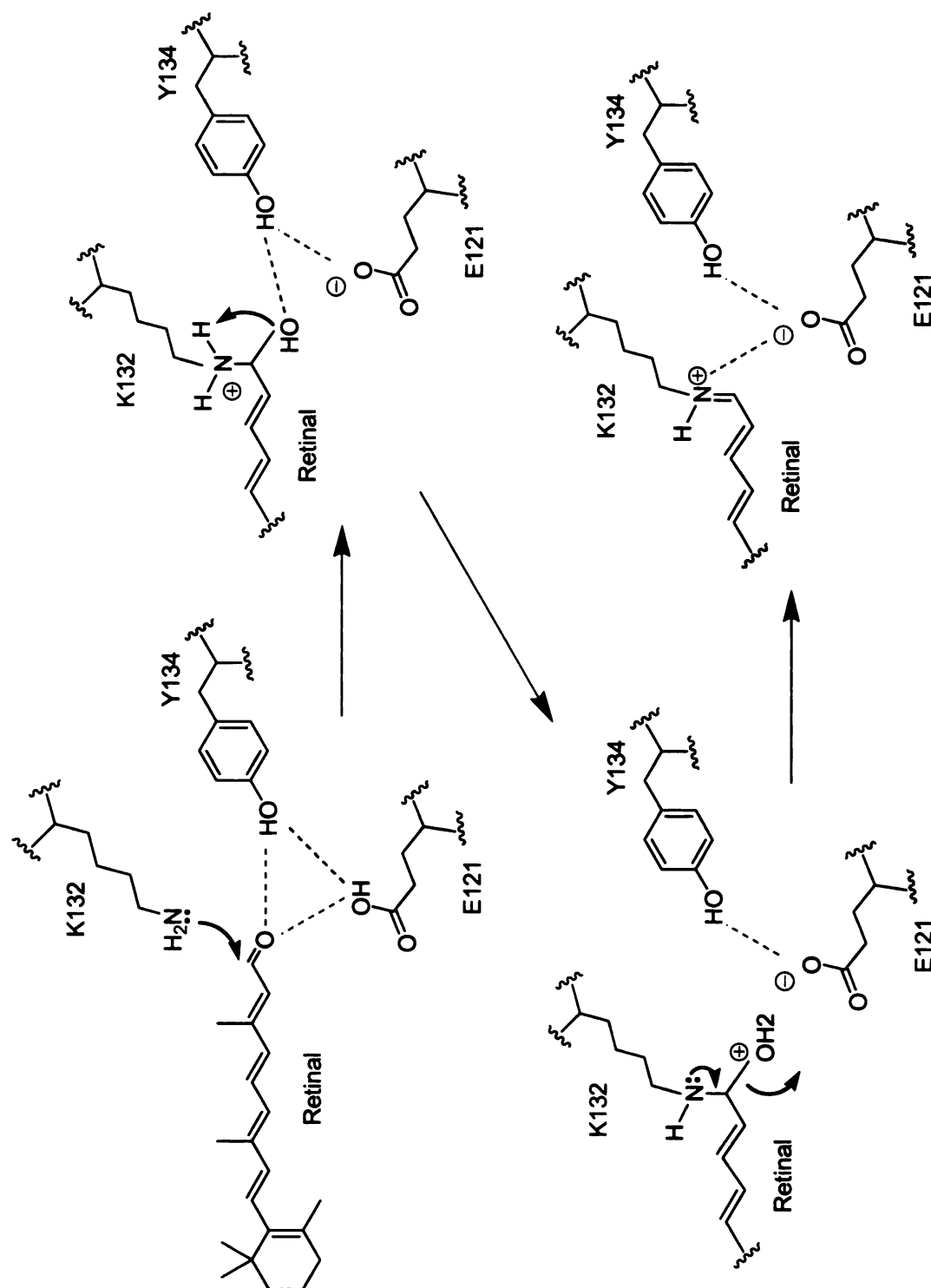


Figure 2-23 Proposed mechanism of PSB formation between Lys132 and retinal in the pocket of R132K:R111L:L121E•Retinal.

## **2.11 The Importance of the PSB Counter Anion**

The crystal structure of R132K:R111L:L121E•Retinal clearly shows the important role of Glu121 in serving as a counter anion for the PSB. To verify that Glu121 is indeed the counter anion several mutants were designed, in which Leu121 was mutated to other residues. For example the R132K:R111L:L121Q•Retinal could clearly form a SB when tested by MALDI, however it only showed a single and non-red shifted peak at 376 nm. This experiment was also performed for the R132K:R111L:L121Q:T54V•Retinal mutant to make sure Thr54 is not hindering the PSB formation. The tetra mutant bound to retinal did not show any red shift either, and maximally absorbed at 375 nm. However it could clearly form a SB with retinal based on the MALDI experiments. These experiments could pinpoint the importance of Glu21 in stabilizing the positive charge on the PSB.

Similar results were observed in rhodopsin since mutation of Glu113 in the pocket of bovine rhodopsin, which is the counter anion in this molecule, results in a ~ 110 nm blue shift (from 490 nm to 380 nm).<sup>17</sup> More details on studying the relationship between the distance and position (angle) of the counter anion, with respect to PSB, and the spectroscopic behavior of the CRABPII mutants bound to retinal are discussed in the Dr. Vasileiou's dissertation.<sup>25</sup> Also more examples of destabilization of the structure upon mutation of the conserved residues, in order to design a counter anion at positions 134 and 41, are discussed.

In summary, so far Glu121 is the best place to design the counter anion. Other positions either destabilize the mutant or result in an unstable PSB species.

## **2.12 The Importance of the Arg111 Mutation**

As we learned from the crystal structures of R132K:Y134F•Retinal and R132K:R111L:L121E•Retinal, removal of Arg111 is critical for SB formation. When Arg111 is mutated a hydrophobic residue the water-mediated interaction between Arg111 and retinal is removed, which gives retinal freedom to assume a favorable conformation for the nucleophilic attack by Lys132. However in R132K:Y134F mutant, Tyr134 which hydrogen bonds to the carbonyl group of retinal, is mutated. We believe Tyr134 not only positions retinal for a favorable nucleophilic attack but also activates the carbonyl carbon for the attack. Therefore a series of mutants were designed in which Tyr134 was restored but Arg111 was mutated to various hydrophobic and hydrophilic residues. The results show that mutation of Arg111 into a hydrophobic residue clearly assists the SB formation. R111V and R111M mutations formed a stable PSB in the presence of Glu121. However, regardless of the residue at position 111, when Glu121 is absent there is no indication of PSB formation. For more details see Dr. Vasileious's dissertation.

### **Closing Remarks**

Overall, using modern techniques of rational protein design and following the basic principles of organic chemistry we have successfully converted CRABP<sub>II</sub>, a small and soluble protein which is easy to manipulate and study, into a retinal binding protein. The designed binding pocket of CRABP<sub>II</sub>, in R132K:R111L:L121E•Retinal, has the characteristics to mimic the binding cavity of visual rhodopsin. Namely, it forms a PSB with Lys132 and Glu121 stabilizes the positive charge on the PSB as Glu113 does in the



pocket of rhodopsin. The mimic will be used to test the different hypotheses that are proposed to be involved in the wavelength regulation mechanism.

The high resolution crystal structures of the intermediate mutants proved to be invaluable in planning rational mutagenesis studies toward making this mimic. The picture of the binding pocket provided by crystal structures were critical to the successful design of the mimic; for example R132K:Y134F•RA or R132K:Y134F•Retinal suggested that Arg111 hindered the SB formation.

One important factor that should not be underestimated, while performing the mutations, is that mutations of the conserve residues might result in destabilizing the structure. For example the structure of the F15W mutant proved that mutation of the Phe15, which is a fully conserved residue in the iLBPs family, disrupted the entire  $\alpha 2$  helix. Although the integrity of each mutant was tested by CD experiments, these assays cannot indicate all the structural changes. For example although CD experiments of F15W mutant did not indicate the loss of the  $\alpha 2$  helix, the crystal structure of this mutant clearly shows that the F15W mutation destabilizes the structure and entirely disrupts the second helix. More details on this structure are discussed in Chapter III.

The crystal structures will be even more important in the studies that we have underway to test the proposed hypotheses of wavelength regulation. For example the crystal structures can depict the interactions of the designed charged or bulky residues in the vicinity of the chromophore and explain the resulting spectroscopic behavior of the mutants.

## 2.13 Literature Cited

1. Honig, B., Greenberg, A. D., Dinur, U. & Ebrey, T. G. (1976). Visual-Pigment Spectra - Implications of Protonation of Retinal Schiff-Base. *Biochemistry* **15**, 4593-4599.
2. Stenkamp, R. E., Filipek, S., Driessen, C. A. G. G., Teller, D. C. & Palczewski, K. (2002). Crystal structure of rhodopsin: a template for cone visual pigments and other G protein-coupled receptors. *Biochimica Et Biophysica Acta-Biomembranes* **1565**, 168-182.
3. Stenkamp, R. E., Teller, D. C. & Palczewski, K. (2002). Crystal structure of rhodopsin: A G-protein-coupled receptor. *Chembiochem* **3**, 963-967.
4. Stenkamp, R. E., Teller, D. C. & Palczewski, K. (2005). Rhodopsin: A structural primer for G-protein coupled receptors. *Archiv Der Pharmazie* **338**, 209-216.
5. Kleywegt, G. J., Bergfors, T., Senn, H., Lemotte, P., Gsell, B., Shudo, K. & Jones, T. A. (1994). Crystal-Structures of Cellular Retinoic Acid-Binding Protein-I and Protein-II in Complex with All-Trans-Retinoic Acid and a Synthetic Retinoid. *Structure* **2**, 1241-1258.
6. Chaudhuri, B. N., Kleywegt, G. J., Broutin-L'Hermite, I., Bergfors, T., Senn, H., Le Motte, P., Partouche, O. & Jones, T. A. (1999). Structures of cellular retinoic acid binding proteins I and II in complex with synthetic retinoids. *Acta Crystallographica Section D-Biological Crystallography* **55**, 1850-1857.
7. Wang, L. C., Li, Y., Abildgaard, F., Markley, J. L. & Yan, H. G. (1998). NMR solution structure of type II human cellular retinoic acid binding protein: Implications for ligand binding. *Biochemistry* **37**, 12727-12736.
8. Chen, X., Tordova, M., Gilliland, G. L., Wang, L. C., Li, Y., Yan, H. G. & Ji, X. H. (1998). Crystal structure of apo-cellular retinoic acid-binding protein type II (R111M) suggests a mechanism of ligand entry. *Journal of Molecular Biology* **278**, 641-653.
9. Rizo, J., Liu, Z. P. & Gierasch, L. M. (1994). H-1 and N-15 Resonance Assignments and Secondary Structure of Cellular Retinoic Acid-Binding Protein with and without Bound Ligand. *Journal of Biomolecular Nmr* **4**, 741-760.

10. Rizo, J., Liu, Z. P. & Gierasch, L. M. (1994). Structure and Folding of Cellular Retinoic Acid-Binding Protein, a Predominantly Beta-Sheet Protein. *Biophysical Journal* **66**, A397-A397.
11. Banaszak, L., Winter, N., Xu, Z. H., Bernlohr, D. A., Cowan, S. & Jones, T. A. (1994). Lipid-Binding Proteins - a Family of Fatty-Acid and Retinoid Transport Proteins. In *Advances in Protein Chemistry, Vol 45*, Vol. 45, pp. 89-151.
12. Wang, L. C., Li, Y. & Yan, H. G. (1997). Structure-function relationships of cellular retinoic acid-binding proteins - Quantitative analysis of the ligand binding properties of the wild-type proteins and site-directed mutants. *Journal of Biological Chemistry* **272**, 1541-1547.
13. Steinberg, G., Ottolenghi, M. & Sheves, M. (1993). Pk(a) of the Protonated Schiff-Base of Bovine Rhodopsin - a Study with Artificial Pigments. *Biophysical Journal* **64**, 1499-1502.
14. Nathans, J., Thomas, D. & Hogness, D. S. (1986). Molecular-Genetics of Human Color-Vision - the Genes Encoding Blue, Green, and Red Pigments. *Science* **232**, 193-202.
15. Zhukovsky, E. A. & Oprian, D. D. (1989). Effect of Carboxylic-Acid Side-Chains on the Absorption Maximum of Visual Pigments. *Science* **246**, 928-930.
16. Nathans, J. (1990). Determinants of Visual Pigment Absorbance - Role of Charged Amino-Acids in the Putative Transmembrane Segments. *Biochemistry* **29**, 937-942.
17. Sakmar, T. P., Franke, R. R. & Khorana, H. G. (1989). Glutamic Acid-113 Serves as the Retinylidene Schiff-Base Counterion in Bovine Rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 8309-8313.
18. Sakmar, T. P., Franke, R. R. & Khorana, H. G. (1991). The Role of the Retinylidene Schiff-Base Counterion in Rhodopsin in Determining Wavelength Absorbency and Schiff-Base Pka. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 3079-3083.
19. Bornhorst, J. A. & Falke, J. J. (2000). Purification of proteins using polyhistidine affinity tags. *Applications of Chimeric Genes and Hybrid Proteins, Pt A* **326**, 245-254.

20. Crist, R. M. (2004). Exploring the underlying chemical basis for color vision: designing a protein mimic of rhodopsin. Doctor of Philosophy, Michigan State University.
21. Kim, K. M., Yi, E. C., Baker, D. & Zhang, K. Y. J. (2001). Post-translational modification of the N-terminal His tag interferes with the crystallization of the wild-type and mutant SH3 domains from chicken src tyrosine kinase. *Acta Crystallographica Section D-Biological Crystallography* **57**, 759-762.
22. Mueller, U., Buessow, K., Diehl, A., Bartl, F. J., Niesen, F. H., Nyarsik, L. & Heinemann, U. (2003). Rapid purification and crystal structure analysis of a small protein carrying two terminal affinity tags. *J. Struct. Funct. Genomics* **4**, 217-225.
23. Ledent, P., Duez, C., Vanhove, M., Lejeune, A., Fonze, E., Charlier, P., RhaziFilali, F., Thamm, I., Guillaume, G., Samyn, B., Devreese, B., VanBeeumen, J., LamotteBrasseur, J. & Frere, J. M. (1997). Unexpected influence of a C-terminal-fused His-tag on the processing of an enzyme and on the kinetic and folding parameters. *Febs Letters* **413**, 194-196.
24. Venyaminov, S. Y. & Vassilenko, K. S. (1994). Determination of Protein Tertiary Structure Class from Circular-Dichroism Spectra. *Analytical Biochemistry* **222**, 176-184.
25. Vasileiou, C. (2005). Protein design: reengineering cellular retinoic acid binding protein II into a retinal binding protein. Doctor of Philosophy, Michigan State University.
26. Bradford, M. M. (1976). Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry* **72**, 248-254.
27. Gill, S. C. & Vonhippel, P. H. (1989). Calculation of Protein Extinction Coefficients from Amino-Acid Sequence Data. *Analytical Biochemistry* **182**, 319-326.
28. Norris, A. W., Cheng, L., Giguere, V., Rosenberger, M. & Li, E. (1994). Measurement of Subnanomolar Retinoic Acid-Binding Affinities for Cellular Retinoic Acid-Binding Proteins by Fluorometric Titration. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **1209**, 10-18.

29. Hermanson, G. T. (1996). *Bioconjugate Techniques*, Academic Press, San Diego, CA.
30. Hammes, G. G. (2005). *Spectroscopy for the Biological Sciences*, John Wiley & Sons Inc., Hoboken, NJ.
31. McPherson, A. (2003). *Introduction to Macromolecular Crystallography*, John Wiley & Sons, Inc., Hoboken, NJ.
32. Stout, G. H. & Jensen, L. H. (1989). *X-ray Structure Determination*. 2nd edit, John Wiley & Sons, New York, N.Y.
33. Sands, D. E. (1975). *Introduction to Crystallography*, Dover Publications, Inc., Mineola, N.Y.,.
34. Ladd, M. & Palmer, R. (2003). *Structure Determination by X-ray Crystallography*. Fourth edit, Kluwer Academic/Plenum Publishers, New York, NY.
35. McRee, D. E. (1999). *Practical Protein Crystallography*. Second edit, Academic Press, San Diego, CA.
36. Burgi, H. B., Dunitz, J. D., Lehn, J. M. & Wipff, G. (1974). Stereochemistry of Reaction Paths at Carbonyl Centers. *Tetrahedron* **30**, 1563-1572.
37. Kyte, J. & Doolittle, R. F. (1982). A Simple Method for Displaying the Hydropathic Character of a Protein. *Journal of Molecular Biology* **157**, 105-132.
38. Hopp, T. P. & Woods, K. R. (1981). Prediction of Protein Antigenic Determinants from Amino-Acid-Sequences. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **78**, 3824-3828.
39. Crabb, J. W., Nie, Z. Q., Chen, Y., Hulmes, J. D., West, K. A., Kapron, J. T., Ruuska, S. E., Noy, N. & Saari, J. C. (1998). Cellular retinaldehyde-binding protein ligand interactions - GLN-210 and LYS-221 are in the retinoid binding pocket. *Journal of Biological Chemistry* **273**, 20712-20720.

40. Crabb, J. W., Carlson, A., Chen, Y., Goldflam, S., Intres, R., West, K. A., Hulmes, J. D., Kapron, J. T., Luck, L. A., Horwitz, J. & Bok, D. (1998). Structural and functional characterization of recombinant human cellular retinaldehyde-binding protein. *Protein Science* **7**, 746-757.
41. Crabb, J. W., Chen, Y., Kapron, J. T., West, K. A., Bredberg, D. L. & Saan, J. C. (1996). Analysis of the retinoid-binding site of CRALBP. *Investigative Ophthalmology & Visual Science* **37**, 3691-3691.
42. Crabb, J. W., Intres, R., Goldflam, S., Chen, Y., Johnson, C. & West, K. (1992). Recombinant Production of Visual Cycle Cellular Retinaldehyde-Binding Protein. *Faseb Journal* **6**, A528-A528.
43. Ong, D. E. & Chytil, F. (1978). Cellular Retinol-Binding Protein from Rat-Liver - Purification and Characterization. *Journal of Biological Chemistry* **253**, 828-832.
44. Folli, C., Calderone, V., Ottonello, S., Bolchi, A., Zanotti, G., Stoppini, M. & 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.
45. Fitch, C. A., Karp, D. A., Lee, K. K., Stites, W. E., Lattman, E. E. & Garcia-Moreno, B. (2002). Experimental pK(a) values of buried residues: Analysis with continuum methods and role of water penetration. *Biophysical Journal* **82**, 3289-3304.
46. Wald, G. (1938). On rhodopsin in solution. *Journal of General Physiology* **21**, 795-832.
47. Lythgoe, R. J. (1937). The absorption spectra of visual purple and of indicator yellow. *Journal of Physiology-London* **89**, 331-358.
48. Liang, J., Steinberg, G., Livnah, N., Sheves, M., Ebrey, T. G. & Tsuda, M. (1994). The Pk(a) of the Protonated Schiff-Bases of Gecko Cone and Octopus Visual Pigments. *Biophysical Journal* **67**, 848-854.
49. Koutalos, Y., Ebrey, T. G., Gilson, H. R. & Honig, B. (1990). Octopus Photoreceptor-Membranes - Surface-Charge Density and Pk of the Schiff-Base of the Pigments. *Biophysical Journal* **58**, 493-501.

50. Nakagawa, M., Iwasa, T., Kikkawa, S., Tsuda, M. & Ebrey, T. G. (1999). How vertebrate and invertebrate visual pigments differ in their mechanism of photoactivation. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6189-6192.
  
51. Song, J. K., Laskowski, M., Qasim, M. A. & Markley, J. L. (2003). NMR determination of pK(a) values for Asp, Glu, His, and Lys mutants at each variable contiguous enzyme-inhibitor contact position of the turkey ovomucoid third domain. *Biochemistry* **42**, 2847-2856.
  
52. Barbas, C. F., Heine, A., Zhong, G. F., Hoffmann, T., Gramatikova, S., Bjornestedt, R., List, B., Anderson, J., Stura, E. A., Wilson, I. A. & Lerner, R. A. (1997). Immune versus natural selection: Antibody aldolases with enzymic rates but broader scope. *Science* **278**, 2085-2092.
  
53. Schmidt, D. E. & Westheim, F. H. (1971). pK of Lysine Amino Group at Active Site of Acetoacetate Decarboxylase. *Biochemistry* **10**, 1249-&.
  
54. Daopin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W. & Matthews, B. W. (1991). Structural and Thermodynamic Consequences of Burying a Charged Residue within the Hydrophobic Core of T4 Lysozyme. *Biochemistry* **30**, 11521-11529.
  
55. Kabsch, W. & Sander, C. (1983). Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **22**, 2577-2637.
  
56. Stryer, L. (1988). *Biochemistry*. Third edit, W.H. Freeman and Company, New York, NY.

## **Chapter III**

### **Structural Studies of Human Apo- and Holo-Cellular Retinoic Acid Binding Protein Type II: Opening Doors for New Insights into its Structure/Function Relationship**

The physiological importance and function of CRABP II has been discussed in Chapter I. In order to get a better understanding of the structure/function relationship of CRABP II we crystallized and determined the first crystal structure of apo-WT CRABP II. In addition we have determined and studied the high resolution crystal structures of several apo-CRABP II mutants, which led to a better understanding of the importance of conserved residues on the structural integrity of CRABP II. Our structure of apo WT-CRABP II shed light on some recently proposed functional properties of this protein, which were based on a mutant of this protein (R111M).<sup>1-3</sup>

The crystallization conditions, data collection and refinement statistics of each structure is discussed in Chapter IV. Data collection and refinement statistics of all the apo-structures are reported in Tables 4-3 and 4-5, respectively, in Chapter IV. Analysis of Ramachandran plots are also shown and discussed in Chapter IV

#### **3.1 Overall Structure of Apo-WT CRABP II**

Here we report the first high resolution crystal structure of human apo-WT CRABP II. The structure was determined at 1.35 Å with crystallographic R-factors of  $R_{\text{work}} = 14.35\%$  and  $R_{\text{free}} = 20.05\%$ . The structure is monomeric and in the P1 space group with 2 independent molecules in the asymmetric unit, identical to that of the



R111M structure solved by Chen et al.<sup>2</sup> The two molecules are labeled Mol A and Mol B consistent with Chen et al.<sup>2</sup> (Figure 3-1).

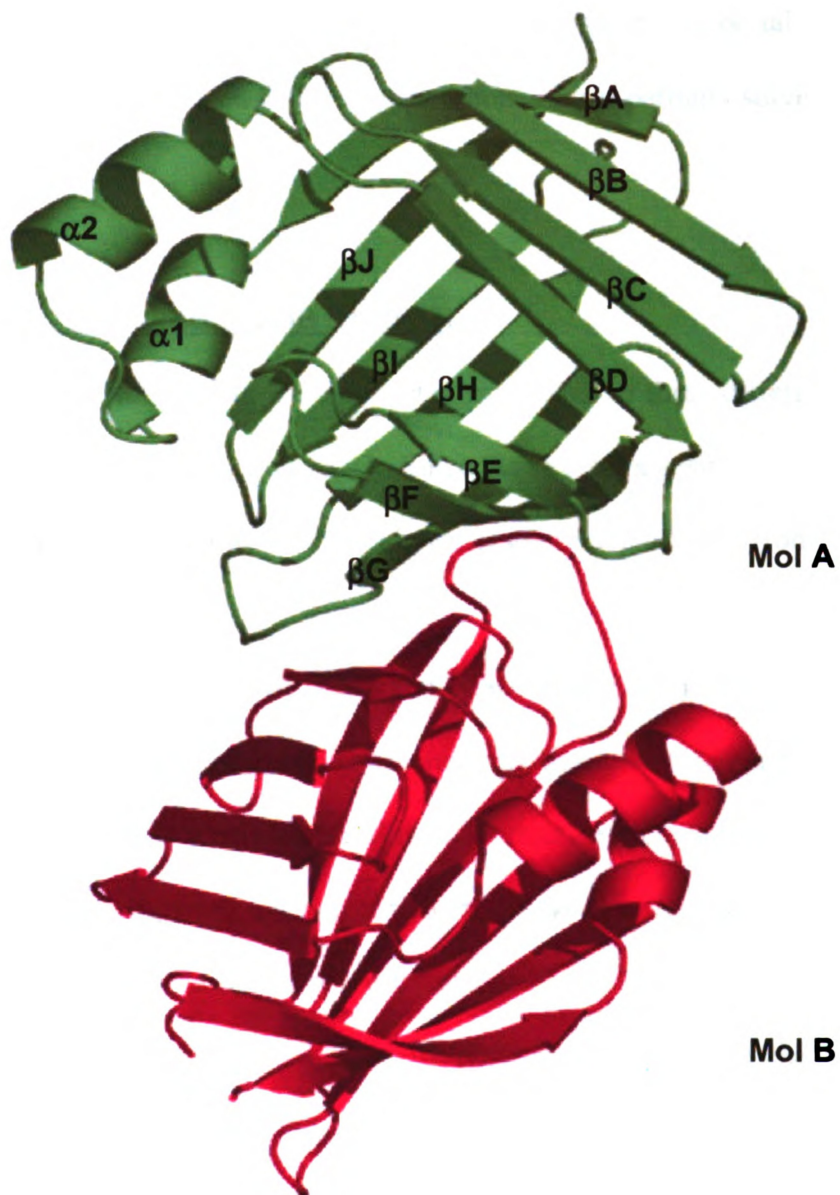


Figure 3-1 The structure of apo-WT CRABP II. The structure is in the P1 space group with two independent molecules in the asymmetric unit. Molecule A (Mol A) and molecule B (Mol B) are shown in green and hot pink, respectively.

Apo-WT CRABP II, like other iLBPs, has two  $\alpha$  helices and a  $\beta$  barrel that comprises its binding pocket. The  $\beta$  barrel is formed by two five-stranded  $\beta$  sheets that are almost orthogonal to each other and form a deep, large and embedded binding pocket. The  $\alpha$ 1-loop- $\alpha$ 2 motif and  $\beta$ C- $\beta$ D and  $\beta$ E- $\beta$ F hairpin loops form the portal of the pocket. RA binds deep inside the pocket, while only its ionone ring is partially solvent exposed at the portal of the pocket.<sup>4</sup>

#### Xtal1 (PDB ID: 2FS6) vs. Xtal2 (PDB ID: 2FS7)

Complete diffraction data were collected from each of three different crystals of apo-WT CRABP II, Xtal1, Xtal2 and Xtal3. Since the structures resulting from Xtal2 and Xtal3 were virtually identical, only the higher resolution data set (from Xtal2) was completely refined.

The secondary structures of Xtal1 and Xtal2 were determined using the DSSP program.<sup>5</sup> Both structures show a similar secondary structure. To compare the structures of Xtal1 and Xtal2, Mol A's and Mol B's of each structure were superimposed on each other (Figure 3-2). Comparing the two structures shows that the Mol A's are virtually identical (the overall RMSD between the C $^{\alpha}$  atoms of the Mol A's is only 0.198 Å). However Mol B's in Xtal1 and Xtal2 show distinct differences in the loop region connecting  $\alpha$ 2 to  $\beta$ B (residues 36-40), while  $\alpha$ 2 has an identical conformation in both molecules (Figures 3-2-A and 3-2-B). The unbiased Fo-Fc electron density maps of the loop region in Mol B's of Xtal1 and Xtal2 are shown in Figures 4-17 and 4-18, respectively, in Chapter IV. The overall RMSD between the C $^{\alpha}$  atoms of the Mol B's is 0.689 Å. However, most of this difference comes from the loop region (residues 36-40),

as the RMSD between the C<sup>α</sup> atoms of these residues is 3.35 Å. Excluding these residues from the RMSD calculation gives an overall RMSD of 0.26 Å. Therefore, although these crystals were grown in an identical condition and have identical unit cells and crystal-packing, there is a substantial difference in conformation between two crystallographically identical molecules.

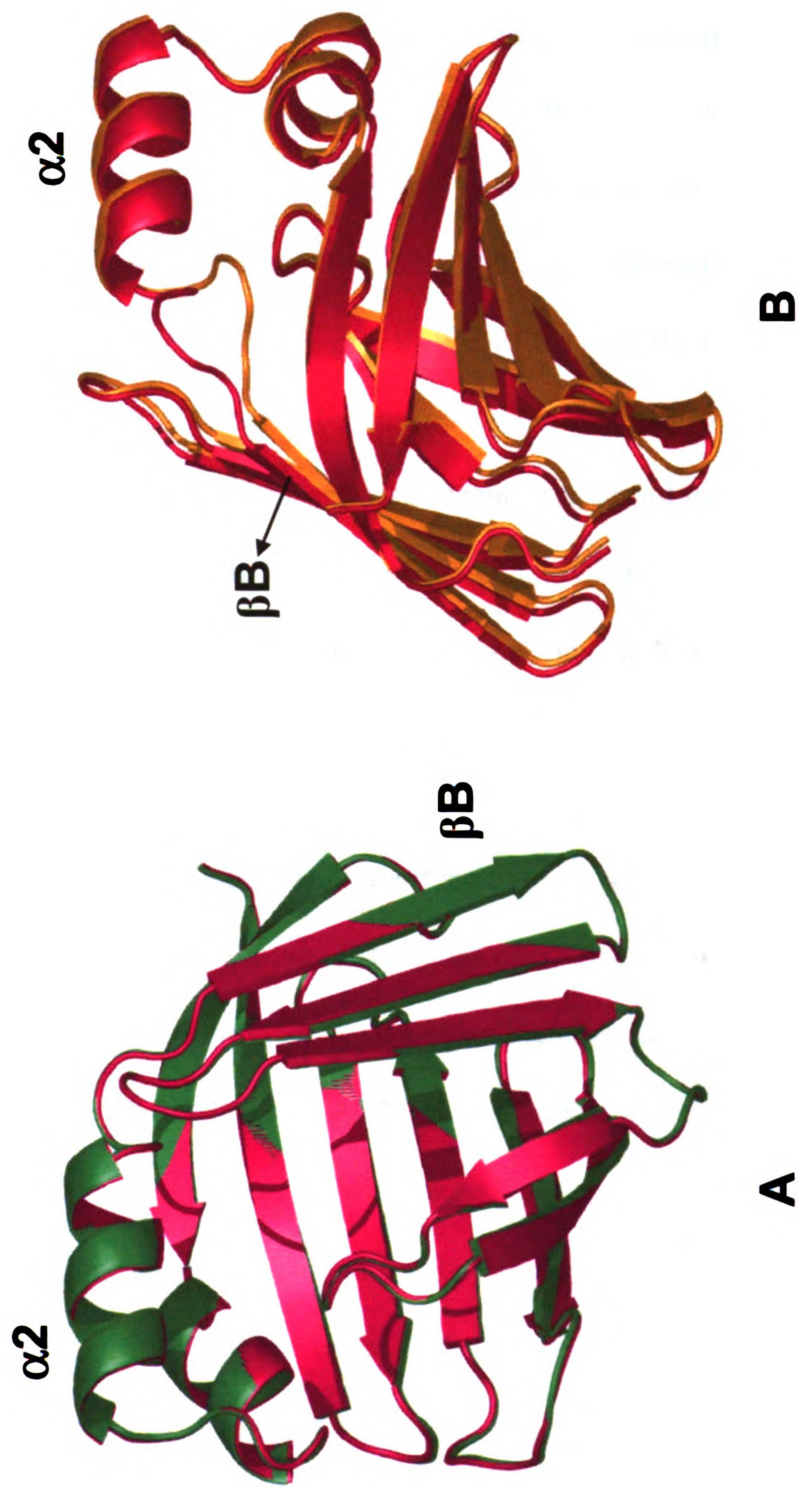


Figure 3-2 The superimposed structures of (A) Mol A's of Xtal1 (green) and Xtal2 (purple); and (B) Mol B's of Xtal1 (hot pink), and Xtal2 (orange). Mol A's are identical. Mol B's differ only at the loop connecting  $\alpha 2$  to  $\beta B$ .

### 3.2 Comparison between Mol A and Mol B of Apo-WT CRABPII

We refer to Xtal1 when comparing the two molecules in the asymmetric unit, because it has higher resolution and better data statistics. The two molecules have very similar structures except for major structural changes in  $\alpha 2$ , the loop connecting  $\alpha 2$  to  $\beta B$ , and the  $\beta C$ - $\beta D$  hairpin loop. There are also minor differences in the  $\beta F$ - $\beta G$  and  $\beta I$ - $\beta J$  hairpin loops (Figure 3-3). The overall RMSD between the  $C^\alpha$  atoms of Mol A and Mol B is 0.922 Å. The differences become noticeable at Lys20, located at the C-terminus of  $\alpha 1$  (xyz deviation of 0.910 Å), and increase moving toward the C-terminus of  $\alpha 2$  and the loop connecting  $\alpha 2$  to  $\beta B$ . Ser37, in the middle of the loop, has a maximum xyz deviation (3.384 Å) between equivalent  $C^\alpha$  positions in Mol A and Mol B. The differences decrease at the N-terminus of  $\beta B$ . If residues 20 to 40 (the loop- $\alpha 2$ -loop region) are excluded from the RMSD calculation between the two molecules, the overall RMSD falls to 0.73 Å. The RMSD for residues 20 to 40 is equal to 1.60 Å.

As shown in Figure 3-3, the C-terminus of  $\alpha 2$  moves toward the binding site of RA, and away from  $\alpha 1$  (maximum xyz deviation between the  $C^\alpha$  atoms of the  $\alpha 2$  in Mol A and Mol B is equal to 2.953 Å at Ala35, the last residue of  $\alpha 2$ ). Since residues on the  $\beta C$ - $\beta D$  hairpin loop directly interact with the residues on the third helical turn of  $\alpha 2$ , the  $\beta C$ - $\beta D$  hairpin loop moves concurrently with the movement of the  $\alpha 2$  helix and away from the binding cavity. This movement of the hairpin loop causes the C-terminus of  $\beta C$  and the N-terminus of  $\beta D$  to move away from the binding cavity as well. The second helical turn of  $\alpha 1$  shows some minor movement compared to Mol A. The maximum xyz

deviation between the  $C^\alpha$  atoms of  $\alpha 1$  in Mol A and Mol B is equal to 1.247 Å at Leu21, the last residue of  $\alpha 1$ .

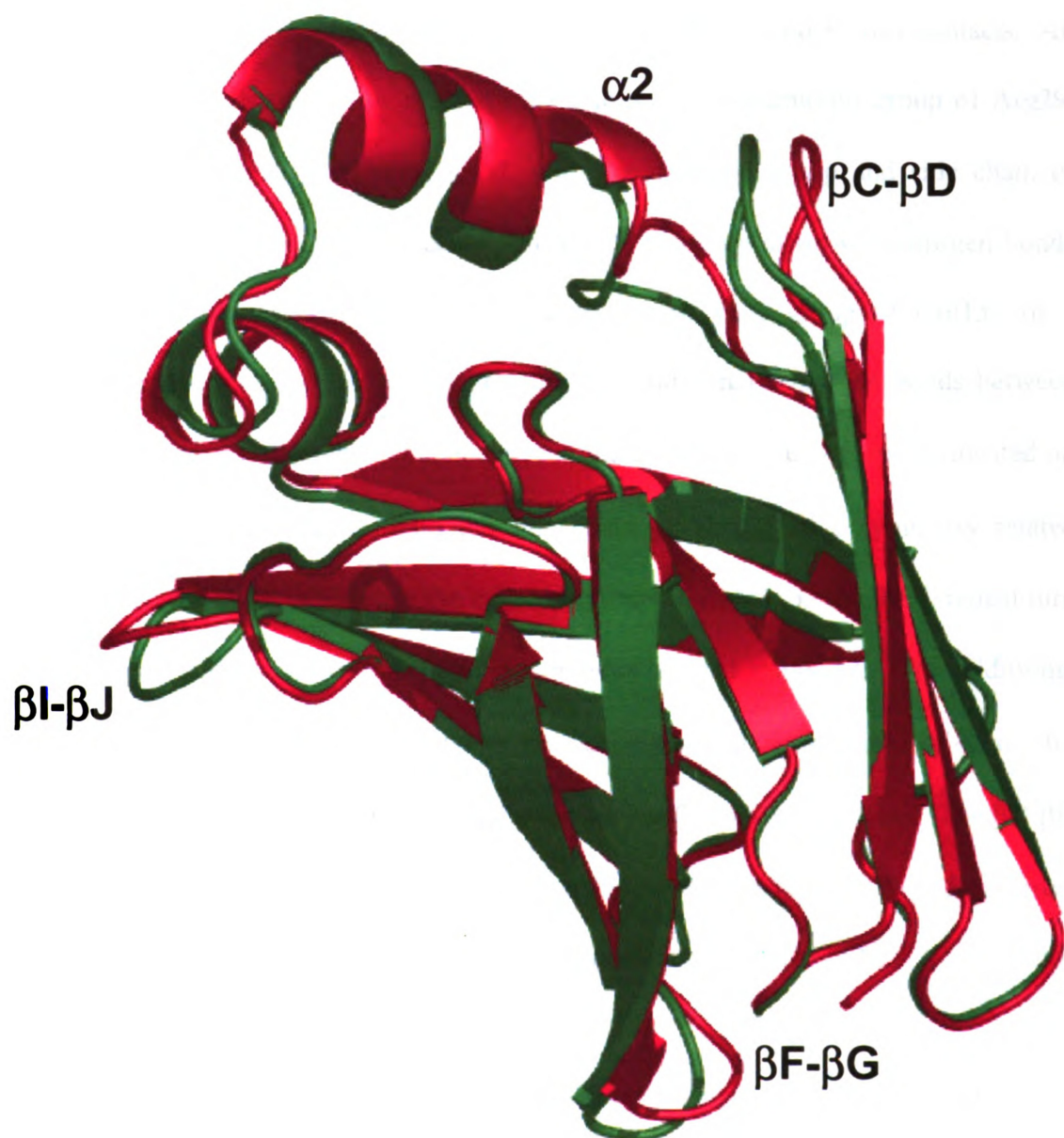


Figure 3-3 Superimposed structures of Mol A (green) and Mol B (hot pink) of Xtal1.

The dissimilarities between the two molecules indicate regions of flexibility and must be due to different crystal-packing environments for the two molecules. As assessed by the program CONTACT in the CCP4 suite,<sup>6</sup> the  $\alpha 2$  helix in Mol A is involved in 3 hydrogen bonds and 16 *van der Waals* (*vdw*) contacts with neighboring molecules, while that of Mol B is involved in 4 hydrogen bonds and 59 *vdw* contacts. All three hydrogen bonds of Mol A involve interactions of the guanidino group of Arg29, located on the second helical turn of  $\alpha 2$ , with the carbonyl group and side chain of Glu137 of a symmetry related molecule. In Mol B there are also two hydrogen bonds between the guanidino group of Arg29 and the carboxylate group of Glu137 of a symmetry related molecule, however there are two additional hydrogen bonds between  $\alpha 2$  and neighboring molecules. One is between the carbonyl group of Lys30, located on the second helical turn of  $\alpha 2$ , with the side chain of Thr131 of a symmetry related molecule; and the other is between the carbonyl group of Ala34, on the third helical turn of  $\alpha 2$ , and the side chain of Asn14 of a symmetry related molecule. The additional hydrogen bonds and *vdw* contacts between the second and third helical turns of  $\alpha 2$  in Mol B with neighboring molecules drag the helix toward the  $\beta C$ - $\beta D$  hairpin loop. The  $\beta C$ - $\beta D$  hairpin loop moves concurrently with the movement of the helix. Therefore direct crystal-packing interactions can explain differences in the structures of Mol A and Mol B in the asymmetric unit.

Crystal-packing also has an impact on the interactions of the residues within the binding pocket of the protein. In Mol A but not in Mol B, the guanidino group of Arg111 interacts with  $\alpha 2$  and the loop connecting  $\alpha 2$  to  $\beta B$  through two short solvent-mediated interactions. The paths of these two interactions are shown in Figure 3-4-A and are: (1)



guanidino group of Arg111 → water 204 → carboxylate of acetate 4 → guanidino group of Arg132 → carbonyl of Ala36; and (2) guanidino group of Arg111 → water 204 → carboxylate of acetate 4 → guanidino group of Arg132 → water 33 → carbonyl of Val33. Val33 is located on the third helical turn of  $\alpha 2$  and Ala36 is on the loop connecting  $\alpha 2$  to  $\beta B$ . Acetate is one of the crystallization reagents and penetrates into the binding cavity either through the portal of the pocket or the gap between the  $\beta D$  and  $\beta E$  strands. This gap is due to the lack of hydrogen bonding between these two  $\beta$  strands.

One might think the solvent-mediated interaction depends on the acetate from the crystallization reagent. However the same solvent-mediated interaction exists in Mol A of Xtal2, but a chloride ion occupies the position of acetate in Xtal1, confirming that the contact does not rely on acetate (Figure 3-4-B). Xtal3 has an acetate molecule exactly as observed in Xtal1.



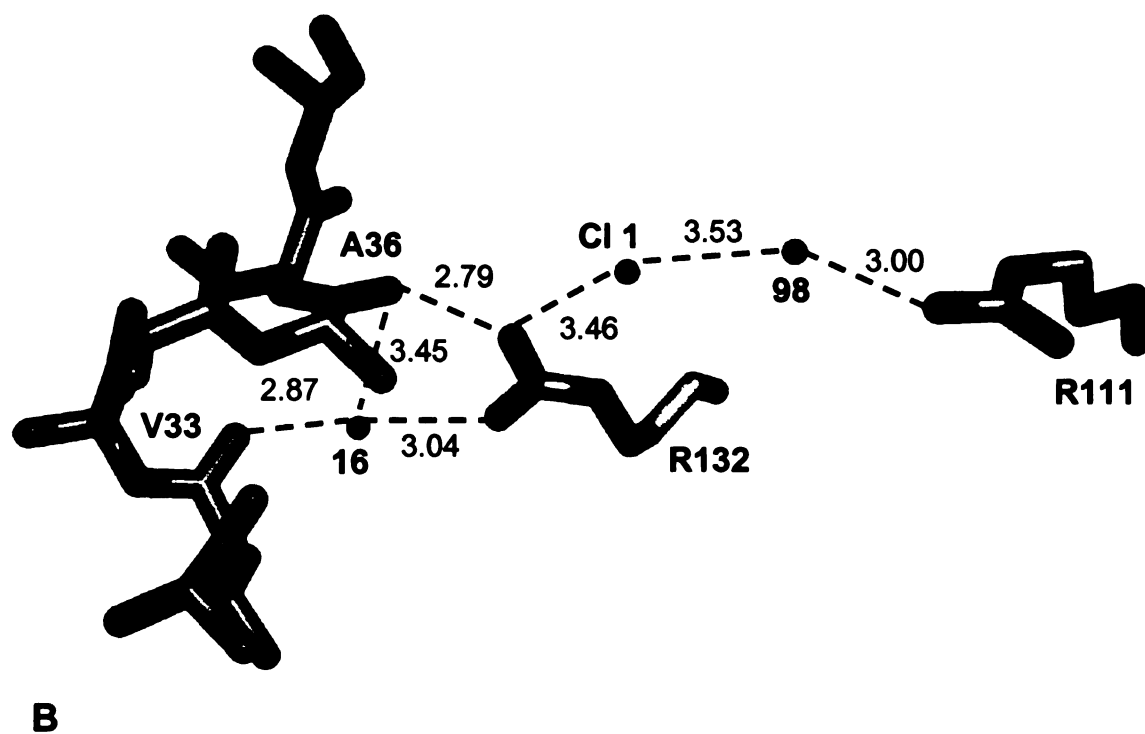
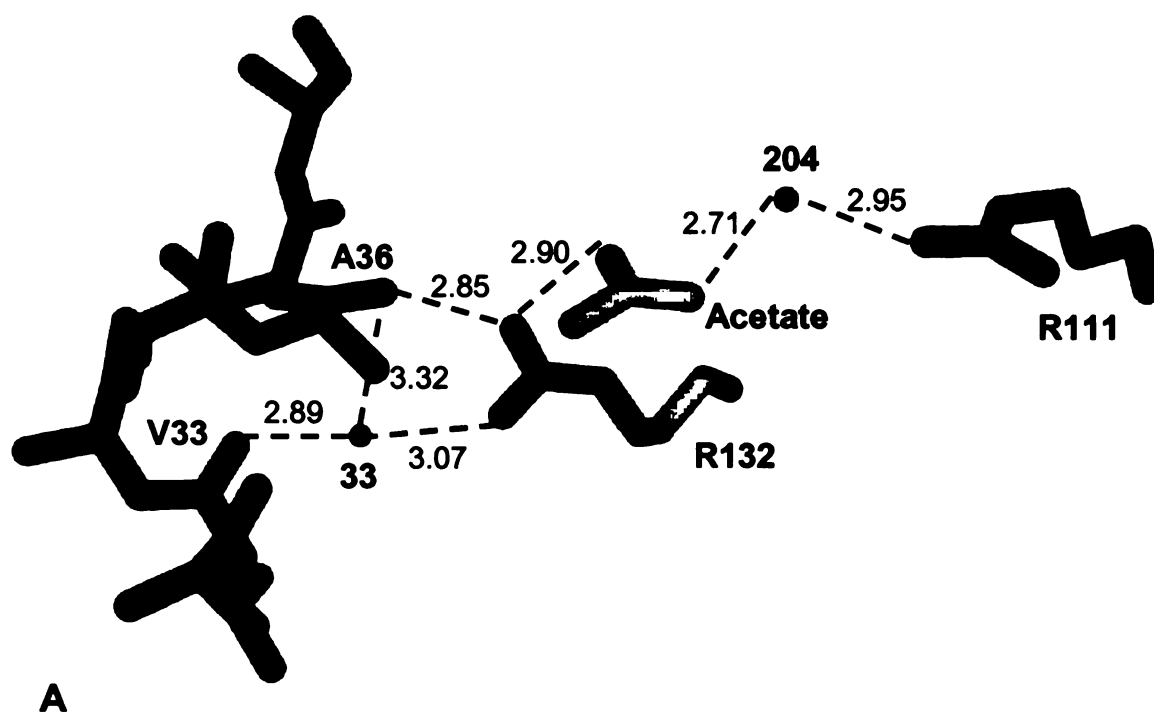


Figure 3-4 (A) Solvent-mediated interaction between Arg111, and Val33 and Ala36, located on  $\alpha 2$  and the loop connecting  $\alpha 2$  to  $\beta B$  in (A) Xtal1 and (B) Xtal2. The distances are in angstroms.

Arg111 has exactly the same conformation in Mol A and Mol B. However, Arg132 has distinct conformations in each molecule. In Mol A it faces toward the  $\alpha 2$  helix and hydrogen bonds with Val33 and Ala36, while in Mol B it looks away from the loop and binds to the carbonyl group of Val76, on the  $\beta C$ - $\beta D$  hairpin loop, and the oxygen of Ser12, the last residue of  $\beta A$  through other water networks.

In Mol B of Xtal2 and Xtal3, and not of Xtal1, the amino group of Arg132 binds to the main chain nitrogen of Lys38, located on the loop connecting  $\alpha 2$  to  $\beta B$ , through a solvent molecule (water or chloride). In Xtal1 the loop connecting  $\alpha 2$  to  $\beta B$  has a different conformation compared to that in Xtal2 and Xtal3 and therefore is positioned further away from Arg132 (Figure 3-2-B). Thus in Mol B of Xtal1 there is no interaction between Arg132 and the loop. This observation indicates that since in Mol A, Arg132 faces the loop and makes a direct and tight hydrogen bond with Ala36, it fixes the position of the  $\alpha 2$  helix and the loop in this molecule. However in Mol B it cannot make such a tight hydrogen bond with the loop and therefore the loop becomes flexible. It seems that interaction of Arg111 with the carbonyl group of Ala36 plays an important role in preserving the structural integrity of the protein in Mol A, while in Mol B crystal-packing interactions between the  $\alpha 2$  and loop regions with their neighboring molecules overcome this interaction, and therefore they assume a different conformation than Mol A. The conformations of these residues are identical in Xtal1, Xtal2 and Xtal3 (data not shown for Xtal3).

But why does Arg132 have two different conformations in the two molecules? Different intermolecular interactions of the C-terminus of  $\alpha 2$  in Mol B as compared to that of Mol A eliminate the interactions of Arg132 with Ala33 and Val36 inside the

pocket and therefore Arg132 assumes a different conformation to hydrogen bond to Val76 and Ser12 through different water-mediated interactions.

### **3.3 Changes Induced by the R111M Mutation**

Arg111 not only is one of the conserved residues of CRABP II,<sup>7</sup> but also one of the key residues in binding of RA to this protein,<sup>2</sup> and therefore mutation of this residue could potentially cause major changes in its structure.

The first crystal structure of an apo-CRABP II has been that of the R111M mutant. The crystals were grown in a very similar condition to our apo-WT crystals and the unit cell is virtually identical to that of apo-WT.

Comparing Mol A of apo-WT with that of apo-R111M shows that although the overall folds of the molecules are very similar, the  $\alpha 2$  helix and  $\beta C$ - $\beta D$  hairpin loop are strikingly different in R111M (Figure 3-5-A). However, the same comparison for Mol B's shows that the  $\alpha 2$  helices have identical conformations (Figure 3-5-B). The major difference between Mol B of apo-WT and R111M lies in the  $\beta C$ - $\beta D$ ,  $\beta G$ - $\beta H$ , and  $\beta I$ - $\beta J$  hairpin loops. The loop connecting  $\alpha 2$  to  $\beta B$ , residues 36 to 39, in R111M has a very similar conformation to that of Mol B of Xtal1 (Figure 3-5-B).

It was reported by Chen et al.<sup>2</sup> that  $\alpha 2$  of Mol A in R111M is involved in no intermolecular hydrogen bond interactions and only 8 intermolecular *vdw* contacts, while that of Mol B is involved in 7 hydrogen bonds and 43 *vdw* contacts. Therefore it was suggested that structural differences between Mol A and Mol B of R111M were due to the fact that the structure of Mol B is an artifact of crystal-packing while that of Mol A is the actual conformation of apo-CRABP II. However, contrary to these observations, the

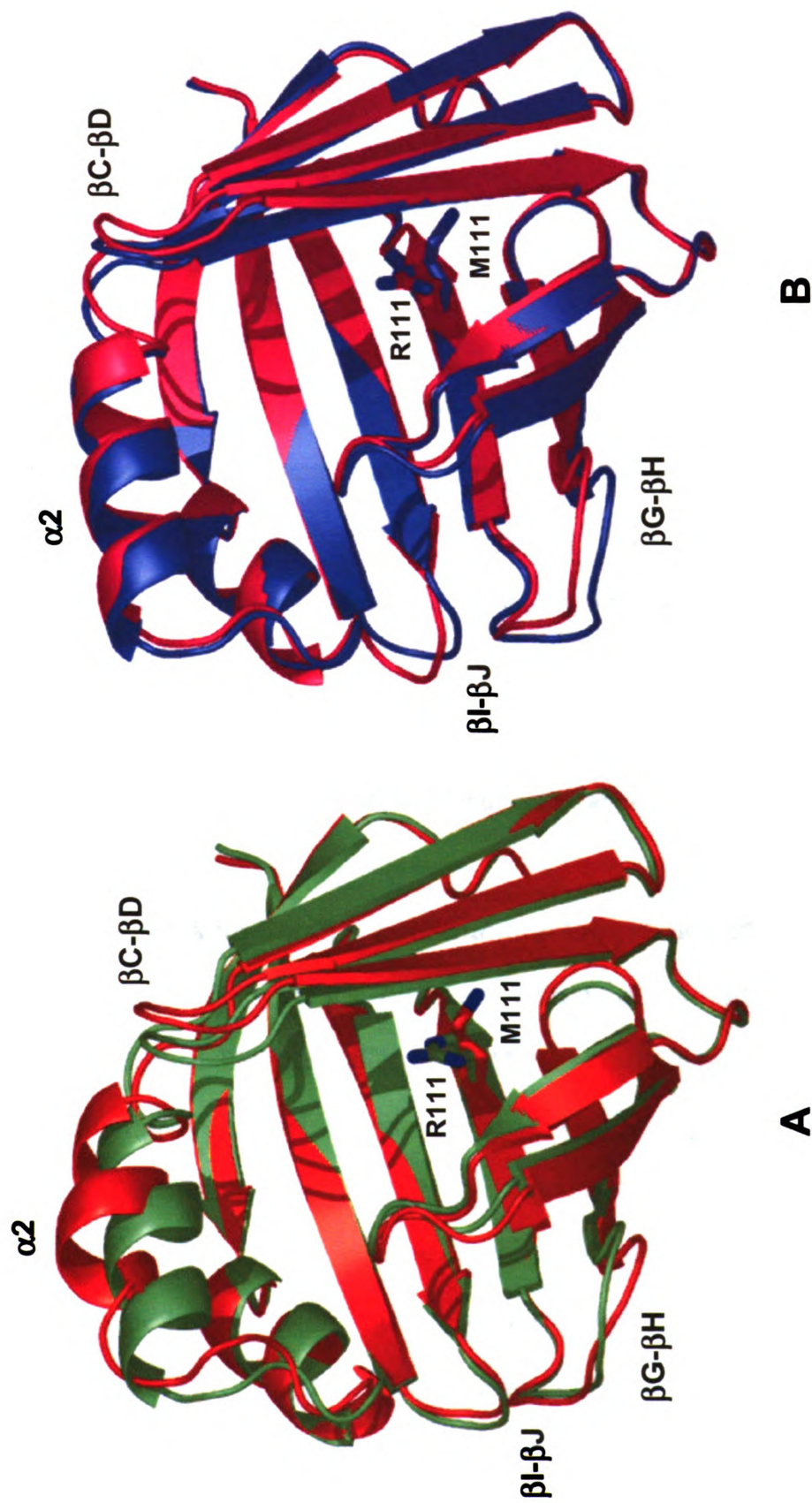


Figure 3-5 Superimposed structures of (A) Mol A's of R111M (red) and apo-WT (green); and (B) Mol B's of R111M (blue) and apo-WT (hot pink).

$\alpha 2$  of Mol A and Mol B of apo-WT CRABP<sub>II</sub> are both involved in crystal-packing hydrogen bonds and *vdw* contacts.

As we discussed here, in Xtal1 Arg111 interacts with carbonyl groups of Val33 and Ala36 through two short solvent-mediated interactions (Figure 3-4). These residues are located on the  $\alpha 2$  helix and the loop connecting this helix to  $\beta B$ , respectively. The water network interaction keeps these regions tight at their place. Therefore it is expected that mutation of Arg111 could allow the helix and loop to be mobile, and therefore cause major structural changes. Investigating the interactions inside the pocket of the R111M mutant shows that there is no interaction between Met111 and residues on the helix or the loop, which indicates that movement of the  $\alpha 2$  helix of Mol A of R111M as compared to that of apo-WT must be due to mutation of this key residue. Also Arg132 has a distinct conformation in this structure from that in Mol A of apo-WT, however still interacts with the carbonyl group of Ala36, but unlike apo-WT, *via* a water molecule.

### **3.4 Other Mutants Confirm the Structural Importance of Arg111**

As mentioned in Chapter II, several mutants of CRABP<sub>II</sub> were designed to optimize the retinal PSB formation in CRABP<sub>II</sub>. To understand the impact of the mutations on the structural integrity of apo-CRABP<sub>II</sub> we determined the crystal structures of apo-F15W, apo-R132K:Y134F, apo-R132K:Y134F:R111L:T54V:L121E, and apo-R132K:R111L:L121E mutants of CRABP<sub>II</sub>. In fact, these structures further characterize the importance of Arg111 and other conserved residues on the structural integrity of apo-CRABP<sub>II</sub>. Crystallization experiments and conditions are discussed in Chapter IV. Also data collection and refinement statistics of these mutants are reported

in Tables 4-3 and 4-5, respectively, in Chapter IV. All the apo-structures of CRABP II that we discuss in the following, are in the same P1 unit cell and have two independent molecules in the asymmetric unit.

#### Crystal Structure of Apo-F15W (PDB ID: 2FRS)

To better understand the importance of conserved residues to the structural integrity of the iLBP fold, we determined the structure of another CRABP II mutant. Among all 52 members of iLBPs Phe15, the first residue of the  $\alpha 1$  helix, is a fully conserved residue (100%). Therefore we decided to conservatively mutate this fully conserved residue to Trp and study the impact of this mutation on the structure. The F15W mutant was crystallized and the structure was determined at 1.51 Å. The structure was refined at crystallographic R-factors of  $R_{\text{work}} = 17.15\%$  and  $R_{\text{free}} = 24.25\%$ .

Interestingly we observed that the F15W mutation completely disrupted the  $\alpha 2$  helix in Mol A (Figure 3-6-A). Though  $\alpha 2$  in Mol B is well defined, this helix and the loop connecting  $\alpha 2$  to  $\beta B$  have a different conformation from that in Mol B of apo-WT and R111M (Figure 3-6-B). In fact Mol B of this mutant has a very similar conformation to that of Mol A of apo-WT. Therefore this structure further confirms the high mobility of this loop in CRABP II while also providing further evidence that mutations of conserved residues in the iLBP fold result in significant structural variance. This mutant is still capable of RA binding, though the affinity is significantly reduced ( $K_d = 40 \pm 4$  nM) compared to WT-CRABP II ( $K_d = 2.0 \pm 1.2$  nM).

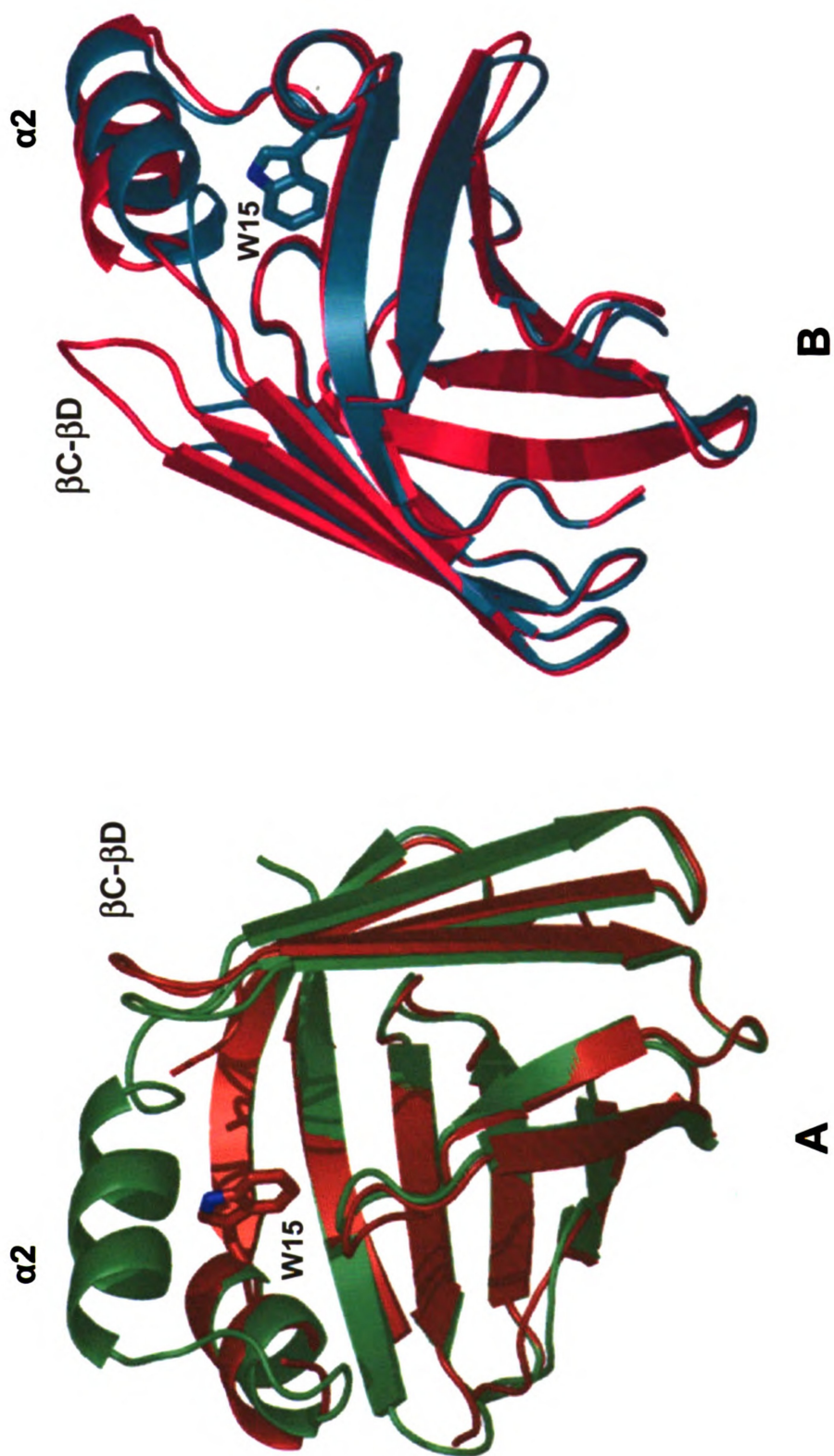


Figure 3-6 Superimposed structures of (A) Mol A's of apo-F15W (brown) and Xtal1 (green); and (B) Mol B's of apo-F15W (cyan) and Xtal1 (hot pink).

### Crystal Structure of Apo-R132K:Y134F (PDB ID: 2FRU)

Arg132 is a fully conserved residue in CRABP<sub>II</sub>, and is strongly conserved as either an Arg or Asp in the entire iLPB's family. Tyr134 is also a highly conserved aromatic residue in this family. Lys and Phe are conserved residues for Arg and Tyr, respectively. To understand whether these conserved mutations affect the structural integrity of CRABP<sub>II</sub> we determined the structure of the apo-R132K:Y134F mutant. The structure was refined at 1.70 Å with R-factors of  $R_{\text{work}} = 15.35\%$  and  $R_{\text{free}} = 22.18\%$ , in the same P1 unit cell and identical crystal-packing as apo-WT CRABP<sub>II</sub>. The result was a structure that was virtually identical to the WT protein. The RMSDs between Mol A's and Mol B's of the two structures are 0.168 and 0.343 Å, respectively, which indicate that these two mutations do not change the overall folding of the structure. The superimposed structures of Mol A's and Mol B's of apo-R132K:Y134F and apo-WT CRABP<sub>II</sub> are shown in Figure 3-7.



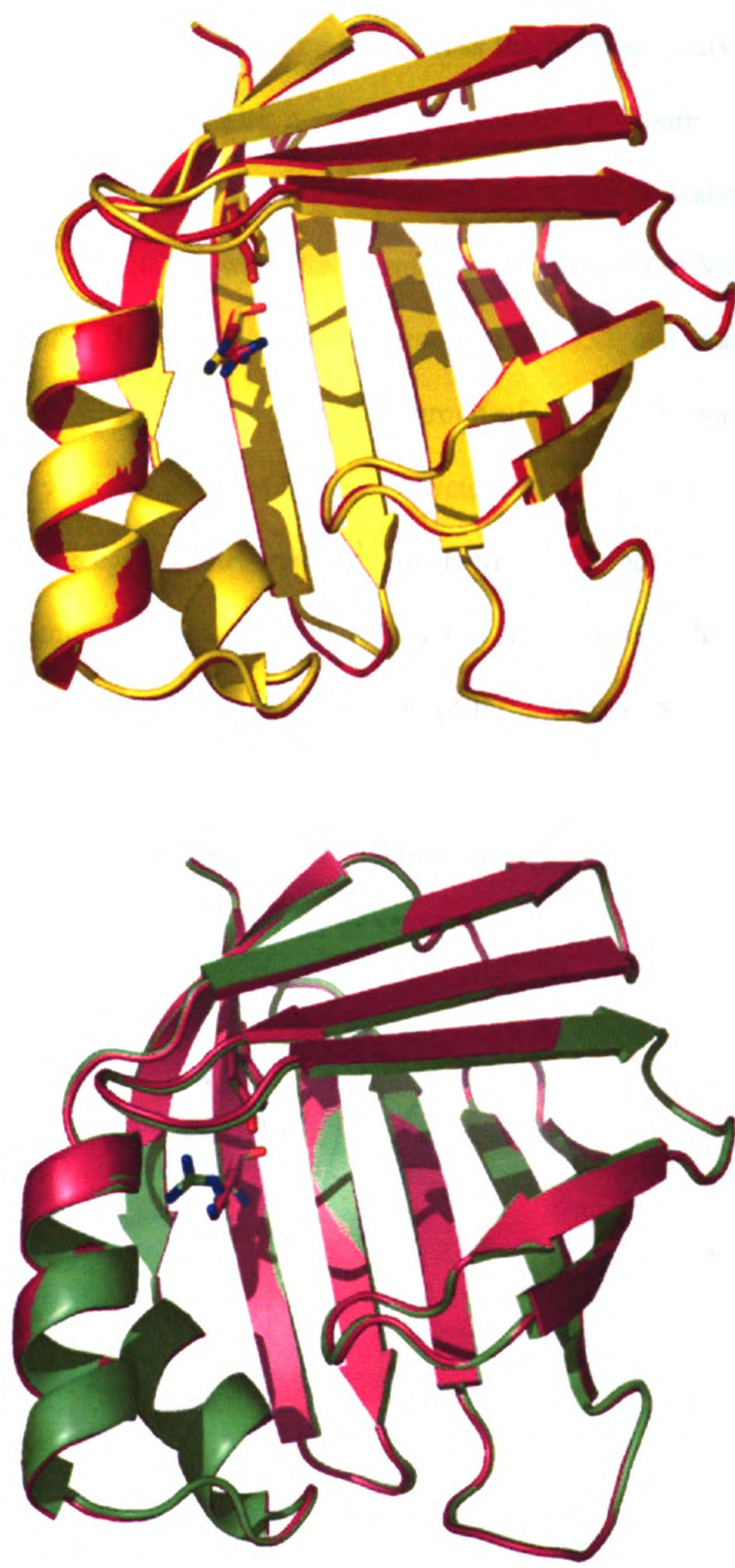


Figure 3-7 The superimposed structures of (A) Mol A's of apo-R132K:Y134F (magenta) and Xtal1 (green); and (B) Mol B's of apo-R132K:Y134F (yellow) and Xtal1 (hot pink).

In fact, a very similar solvent-mediated interaction bridged Lys132 with the carbonyl groups of Val 33 and Ala36 across the binding cavity, as was observed in the structure of apo-WT. The networks are shown in Figure 3-8 and include: (1) the guanidino group of Arg111  $\rightarrow$  water 121  $\rightarrow$  water 225  $\rightarrow$  water 290  $\rightarrow$  amino group of Lys132  $\rightarrow$  water 202  $\rightarrow$  water 83  $\rightarrow$  water 44  $\rightarrow$  carbonyl of Val33; also there is a direct hydrogen bond between the amino group of Lys132 and water 83 but it is a weak interaction (3.40 Å); (2) the guanidino group of Arg111  $\rightarrow$  water 121  $\rightarrow$  water 225  $\rightarrow$  water 290  $\rightarrow$  amino group of Lys132  $\rightarrow$  water 202  $\rightarrow$  carbonyl of Ala36.

This elegant water-mediated interaction was not observed in Mol B of this mutant, similar to the Mol B of Apo-WT. Lys132 has two evenly occupied conformations in Mol B. Although one of these two conformers interacts with Arg111 through two water molecules, it does not interact with the loop region. This structure shows that the crystal-packing interactions in the loop region of Mol B can dictate the interactions inside the pocket, as in Mol B of apo-WT. Mol B of this mutant is almost

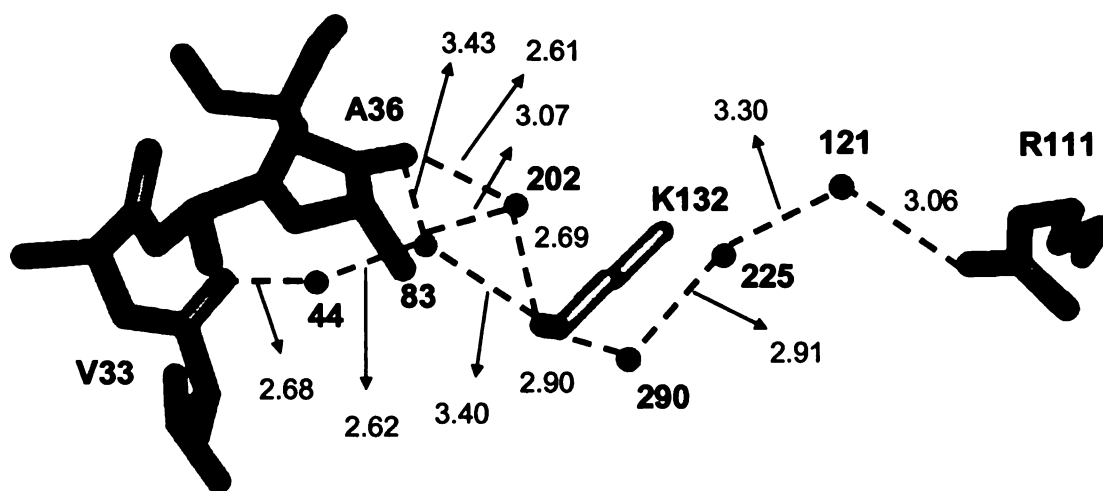


Figure 3-8 The water-mediated interaction between Arg111 and Ala36 and Val33 in apo-R132K:Y134F. The distances are in angstroms.

identical to Mol B of apo-WT (RMSD = 0.343) and has similar intermolecular interactions with its neighboring symmetry molecules, at this region (see the discussion on comparison of Mol A and Mol B of apo-WT).

More importantly, this structure shows that a very similar structure to apo-WT, with very similar solvent-mediated interactions in the binding cavity, is obtained when Arg111 is intact, and other residues in the binding cavity are conservatively mutated.

#### The Crystal Structure of apo-R132K:Y134F:R111L:T54V:L121E (PDB ID: 2FS0)

The structure of the apo-R132K:Y134F:R111L:T54V:L121E (the penta) mutant was determined at very high resolution of 1.20 Å with crystallographic R-factors of  $R_{\text{work}} = 13.80\%$  and  $R_{\text{free}} = 17.62\%$ . The superimposed structures of Mol A's and Mol B's of the penta mutant and apo-WT CRABPII are shown in Figure 3-9. The RMSDs between Mol A's and Mol B's of the two structures are 0.256 and 0.530 Å, respectively. Comparing the two structures indicate that these five mutations do not change the overall fold of the structure. However early on during the refinement we realized that although the  $\alpha 2$  helix has an identical conformation to that in apo-WT, it has an additional conformation.

To generate an unbiased map of the loop and  $\alpha 2$  helix regions, residues 26-40, which are located on the  $\alpha 2$  and the loop connecting this helix to  $\beta B$ , were deleted from map calculations. The Fo-Fc omit electron density map was calculated for this region, which clearly showed one conformation for the helix. This conformation was identical to that in apo-WT. However after the helix and loop were modeled in the omit map, some extra positive electron density map was generated. The map was not definitive enough to

build a second conformation. The B-factors of the atoms at this region (residues 26 to 40) were  $\sim 20 \text{ \AA}^2$ , which are close to the average B-factor of this structure ( $18.13 \text{ \AA}^2$ ). This observation indicates that although Arg111 is mutated to a hydrophobic residue, the helix and loop are not as severely destabilized as was observed in the R111M structure. In addition calculation of the secondary structure using DSSP program confirms that as in apo-WT, residues 26-35 comprise the second helix and it has three full helical turns.<sup>5</sup>

Leu121 is not a conserved residue in CRABP II and is therefore not expected to be critical to the structural integrity of the protein. We believe that either R111L or T54V mutation (or both at the same time) is responsible for partially destabilizing the helix and the loop regions. Thr54 is a conserved residue (80%), and has a similar solvent-mediated interaction with the loop region as observed for Arg111.

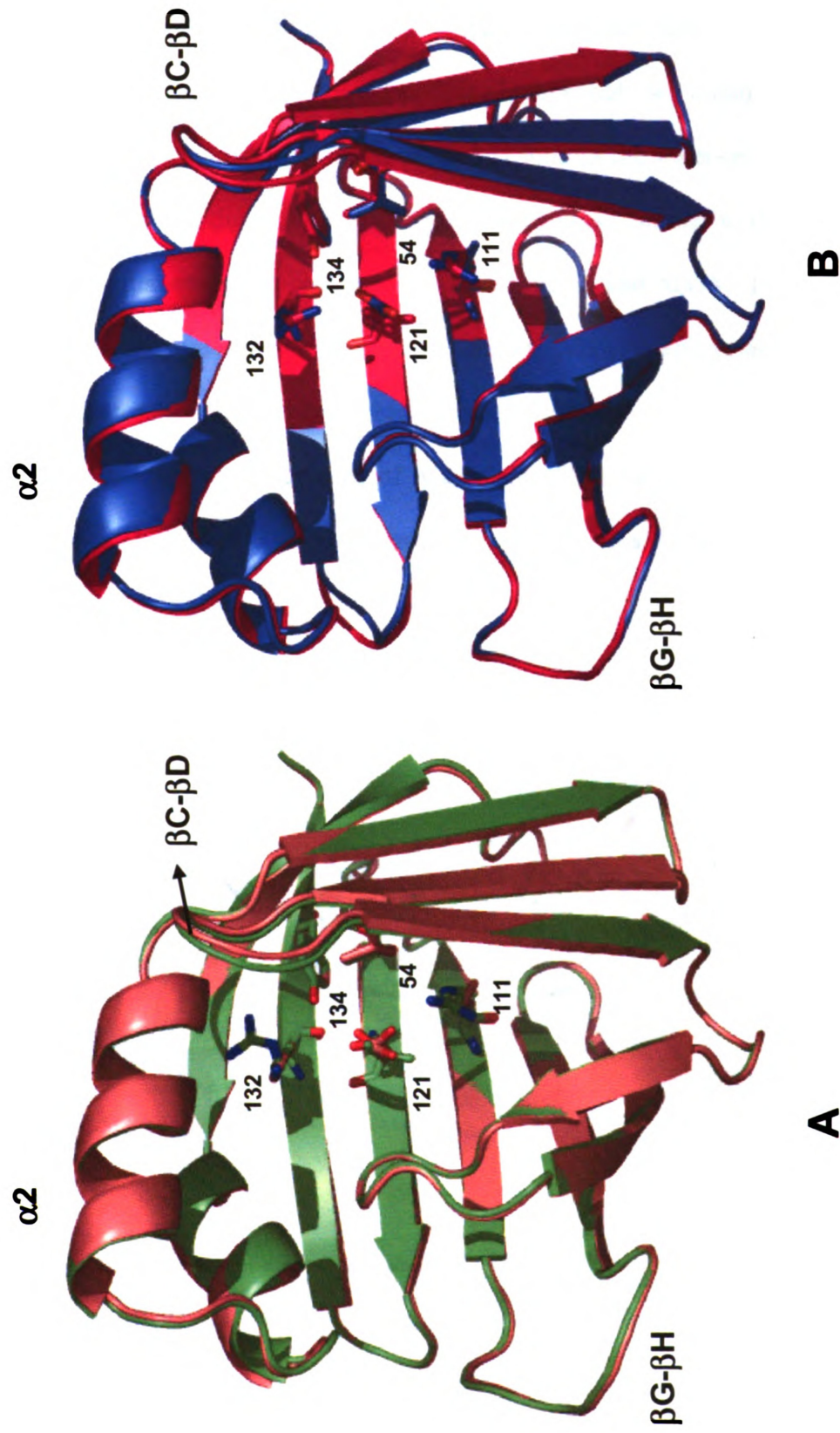


Figure 3-9 The superimposed structures of (A) Mol A's of apo-R132K:Y134F:R111L:T54V:L121E (pink) and Xtal1 (green); and (B) Mol B's of apo-R132K:Y134F:R111L:T54V:L121E (blue) and Xtal1 (hot pink).

To understand why the R111M structure has a destabilized  $\alpha 2$  but the penta mutant has a well-defined  $\alpha 2$ , we investigated the interaction of the residues inside the binding pocket of the latter structure. Very interestingly we found that although Arg111 is mutated, the engineered Glu at position 121 mimics the water-mediated interaction that Arg111 reveals in the previously discussed structures. Similar to these structures, water mediated interactions are between Glu121 and Val33 and Ala36. The paths of these two water networks are shown in Figure 3-10 and include: (1) carboxylate of Glu121  $\rightarrow$  water 393  $\rightarrow$  water 392 (in two closely located conformations)  $\rightarrow$  water 391  $\rightarrow$  amino group of Lys132  $\rightarrow$  water 83  $\rightarrow$  carbonyl of Ala36; and (2) carboxylate of Glu121  $\rightarrow$  water 393  $\rightarrow$  water 392  $\rightarrow$  water 391 amino group of Lys132  $\rightarrow$  water 83  $\rightarrow$  water 89  $\rightarrow$  water 230  $\rightarrow$  carbonyl of Val33.

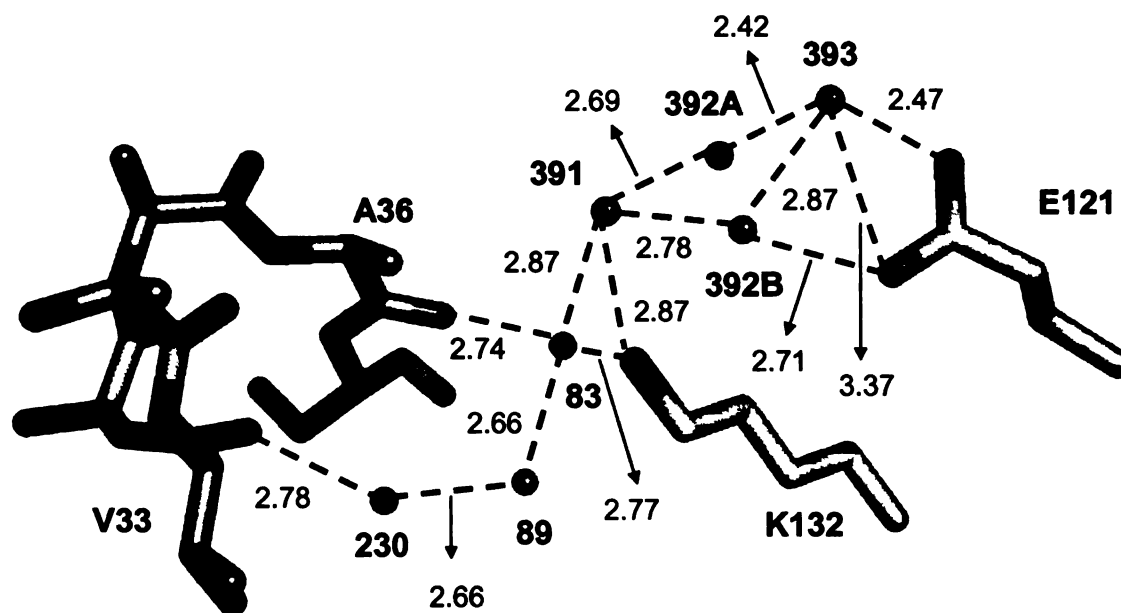


Figure 3-10 The water-mediated interaction between Glu121 and Val33 and Ala36 in apo-R132K:Y134F:R111L:T54V:L121E The distances are in angstroms.

Similar to other structures that we discussed thus far, the water-mediated interaction was not observed in Mol B of this mutant. Glu121 interacts with Lys132 through a water molecule, however Lys132 does not have any interaction with the second helix and loop region.

#### The Crystal Structure of Apo-R132K:R111L:L121E (PDB ID: 2FRR)

The structure of the apo-R132K:R111L:L121E (the triple) mutant was determined at of 1.50 Å and refined with R-factors of  $R_{\text{work}} = 18.22\%$  and  $R_{\text{free}} = 23.03\%$ . The superimposed structures of Mol A's and Mol B's of this mutant and apo-WT CRABPII are shown in Figure 3-11. The RMSDs between Mol A's and Mol B's of the two structures are 1.160 and 0.465 Å, respectively.

The structure of the apo-R132K:R111L:L121E mutant is consistent with the hypothesis that mutation of Arg111 leads to a different conformation of the  $\alpha 2$  helix. In this mutant the last helical turn of the  $\alpha 2$  helix of Mol A is completely unwound at its C-terminus, and the  $\beta C$ - $\beta D$  hairpin loop moves concurrently (Figure 3-11-A). The differences between the Mol A of this mutant and apo-WT become noticeable at the N-terminus of  $\alpha 1$  and reach a maximum divergence at Ala36, the first residue on loop connecting  $\alpha 2$  to  $\beta B$ . The secondary structure calculation of this mutant using DSSP shows that the  $\alpha 2$  helix has lost its last helical turn in this structure and unlike apo-WT that residues 26 to 35 comprise the  $\alpha 2$  helix, here only residues 26 to 33 form the helix.



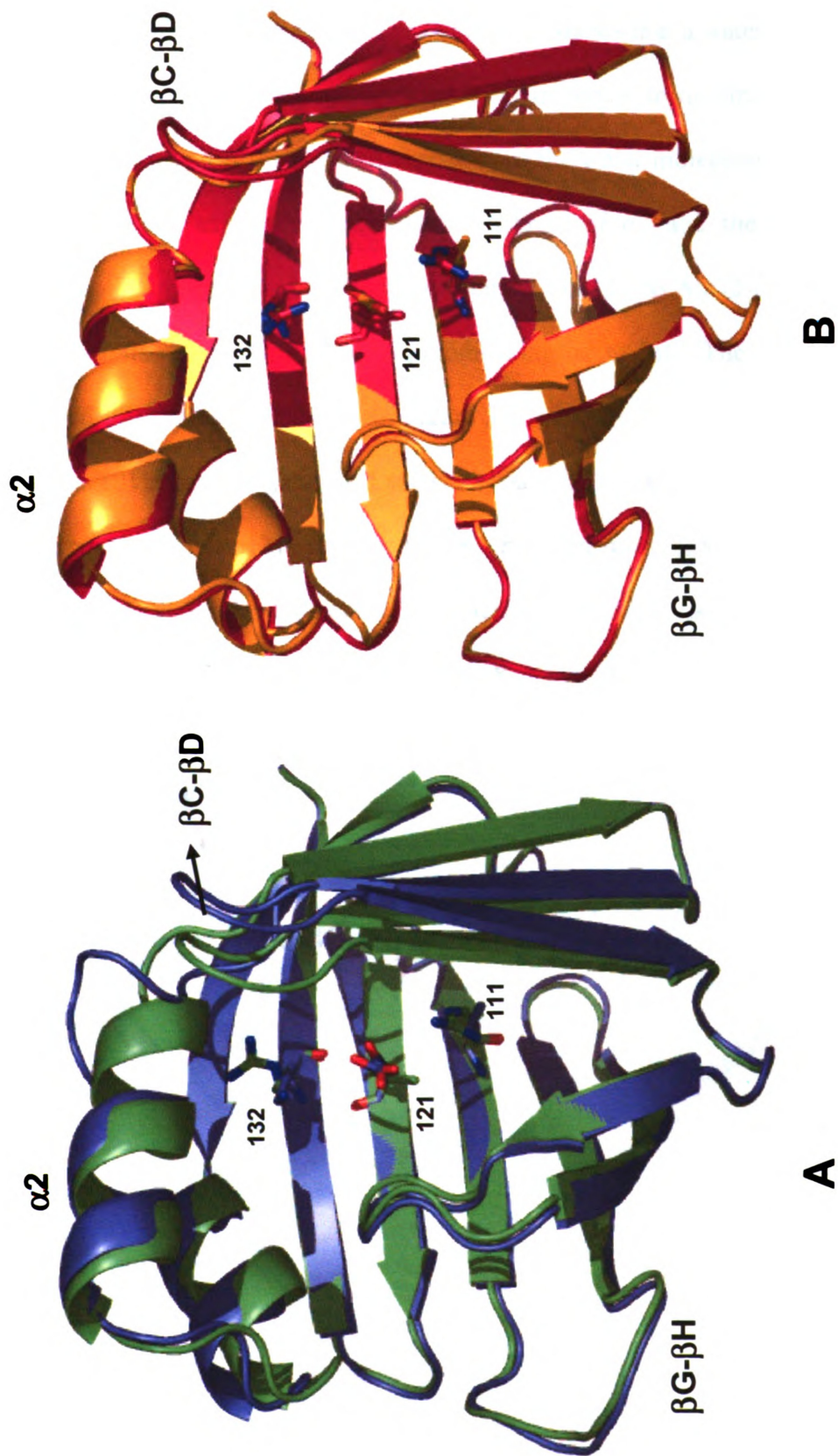


Figure 3-11 The superimposed structures of (A) Mol A's of apo-R132K:R111L:L121E (blue) and Xtal1 (green); and (B) Mol B's of apo-R132K:R111L:L121E (orange) and Xtal1 (hot pink).



Investigating the interaction between the  $\alpha 2$  helix and the loop region with residues at the deep end of the binding pocket (like Glu121) shows that a water mediated interaction exists between these regions, however it is different from other structures that we discussed thus far. A different arrangement of the water molecules inside the pocket of this mutant can explain different conformations of  $\alpha 2$  and the loop region in this structure. Water-mediated interactions connect Glu121 to Ala32 and Ala34, on the  $\alpha 2$  helix, and also to ser37, on the loop connecting  $\alpha 2$  to  $\beta B$ . The paths of these water-mediated interactions are shown in Figure 3-12, and include: (1) carboxylate of Glu121  $\rightarrow$  water 101  $\rightarrow$  amino group of Lys132  $\rightarrow$  water 13  $\rightarrow$  water 102  $\rightarrow$  carbonyl of ser37; (2) carboxylate of Glu121  $\rightarrow$  water 101  $\rightarrow$  amino group of Lys132  $\rightarrow$  water 13  $\rightarrow$  water 102  $\rightarrow$  carbonyl of Ala32; and (3) carboxylate of Glu121  $\rightarrow$  water 101  $\rightarrow$  amino group of Lys132  $\rightarrow$  water 13  $\rightarrow$  water 102  $\rightarrow$  water 18  $\rightarrow$  carbonyl of Ala34.

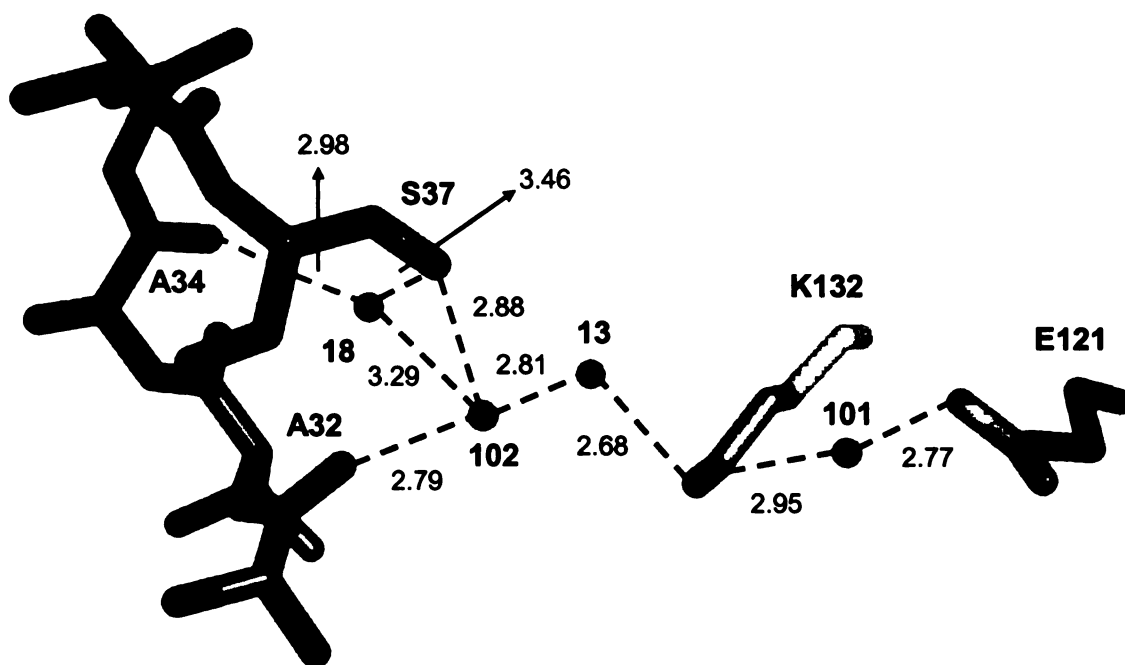


Figure 3-12 The water-mediated interaction between Glu121 and residues on the  $\alpha 2$  helix and the loop connecting this helix to  $\beta B$  in apo-R132K:R111L:L121E. The distances are in angstroms.

Since R132K is a conservative mutation and Leu121 is not a conserved residue, this structure further verifies the importance of Arg111 in defining the structural integrity of CRABP II, particularly of  $\alpha 2$ . Although  $\alpha 2$  and the loop in the apo-R132K:R111L:L121E mutant have different conformations compared to WT, they are well defined. The B-factors of the atoms in these regions (residues 26-40) are  $\sim 22 \text{ \AA}^2$ , which is only slightly higher than the average B-factor of the structure ( $19.251 \text{ \AA}^2$ ).

This observation is contrary to what was observed in the structure of R111M mutant, determined by Chen et al.<sup>2</sup> In R111M, although Arg132 interacts with Ala36 it has no interaction with the other residues. The helix in this structure is highly destabilized and has a very high B-factor of  $\sim 75 \text{ \AA}^2$ . As reported by the authors the electron density map of the helix region in Mol A was not definitive enough to build the helix. So the residues 24 to 37 of Mol A were deleted from the model and all the rest of the residues were built to improve the map in this region. At the end of the refinement, the electron density for these residues was still of insufficient quality to fit the missing segment. At this stage, the least-squares refinement procedure was employed with the program GPRLSA.<sup>8-10</sup> All the reflection data was included and geometry was tightly restrained. Several rounds of such refinement significantly improved the electron density of the missing segment of Mol A. Residues 24-32 were unambiguously built however more rounds of the least-squares refinement had to be run to generate a definitive electron density map for residues 33-37.

In apo-R132K:R111L:L121E (and our other structures discussed here) the electron density map of these residues was very well defined from the beginning of the refinement, which indicates that the region is stable despite the R111L mutation. It

appears that Glu121 is stabilizing the helix in the triple mutant. Therefore although the R111L mutation led to a different conformation for the helix and loop, they are not destabilized. In other words, in this structure Glu121 mimics the role of Arg111 in apo-WT, however through different water-mediated interaction, which leads to a different conformation in the loop and helix region.

### **3.5 Overall Structure of CRABPII•RA and Comparison with Apo-WT CRABPII**

As explained in Chapter II, we have redetermined and refined the structure of CRABPII•RA (PDB ID: 2FR3) to an improved resolution of 1.48 Å with very good crystallographic R-factors of 12.30% and 17.09% for  $R_{\text{work}}$  and  $R_{\text{free}}$ , respectively. The data collection and refinement statistics are reported in Tables 4-4 and 4-6, respectively.

The CRABPII•RA complex was previously crystallized in a different crystallization condition and the structure was refined at 1.80 Å by Kleywegt et al. in 1994 (PDB ID: 1CBS).<sup>3</sup> Our structure is identical to the previously published structure of CRABPII•RA, with the same space group ( $P2_12_12_1$ ) and similar cell constants. RA binds deep inside the binding pocket and the carboxylic group of RA interacts with Arg111, Arg132, and Tyr134. Only the  $\beta$  ionone ring is partially solvent exposed at one of its edges.<sup>4</sup>

As mentioned earlier, Chen et al.<sup>2</sup> believed that Mol A of apo-R111M shows the actual structure of apo-CRABP II and therefore chose Mol A over Mol B for comparing apo- and holo-CRABP II. Superimposition of Mol A of apo-R111M on CRABP II•RA showed that there were significant structural differences between the two molecules in the  $\alpha 2$  helix,  $\beta C$ - $\beta D$  and  $\beta E$ - $\beta F$  hairpin loops, which were thought to be necessary for opening the binding pocket and making RA-entry possible (Figure 3-13, also Figure 3 in Chen et al.<sup>2</sup>).

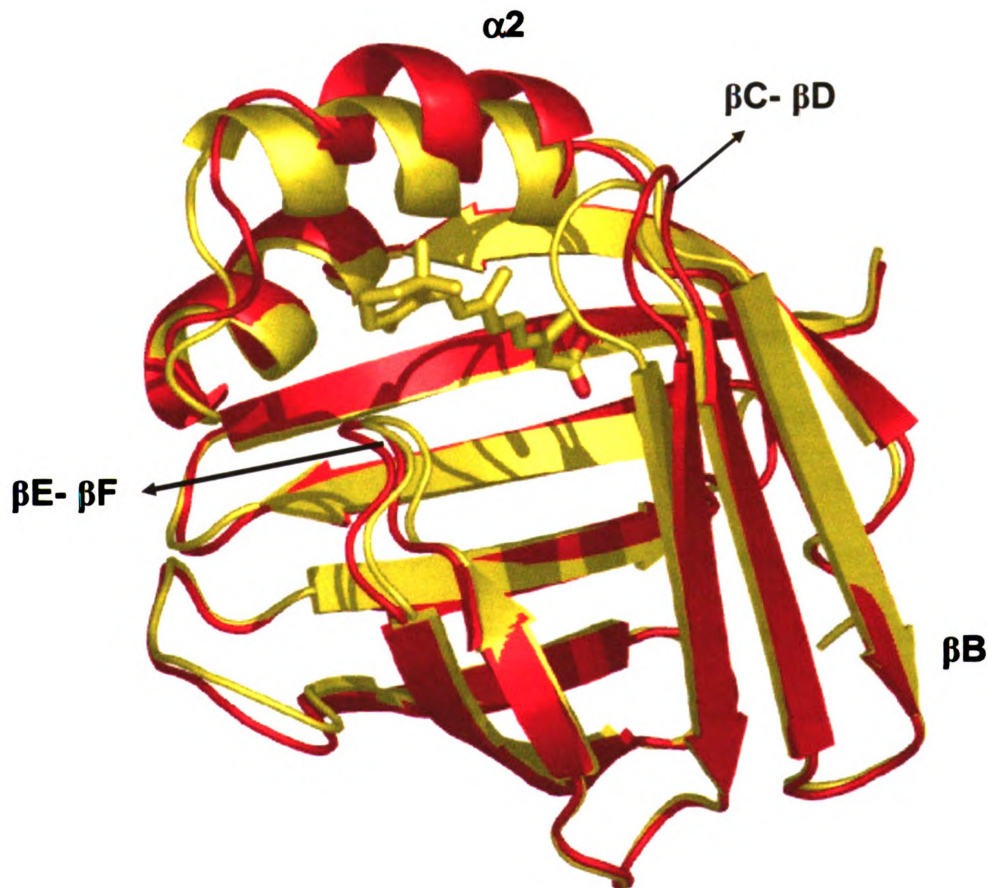


Figure 3-13 Superimposed structures of Mol A of apo-R111M (red) and WT CRABP II•RA (yellow).<sup>2</sup>

Here having the “real” structure of apo-CRABP<sub>II</sub> in hand, we compared the structures of both molecules of apo-CRABP<sub>II</sub> to that of CRABP<sub>II</sub>•RA (Figure 3-14). Comparing Mol A and CRABP<sub>II</sub>•RA (Figure 3-14-A) shows that in fact these two structures are almost identical (RMSD = 0.500) except in the  $\beta$ C- $\beta$ D hairpin loop region. In the apo structure, Arg59 on this hairpin loop has a water-mediated interaction with the carbonyl group of Ala32, which keeps Arg59 close to the portal of the pocket and in the path of RA entry. However, in the holo structure RA pushes the mediating waters away and Arg59 moves away from the portal to give RA space to enter the pocket (Figure 3-15). Also Val58 on the  $\beta$ C- $\beta$ D hairpin loop moves closer to the binding pocket to form *vdw* contacts with C19 of RA. It should be mentioned that the same interactions were observed in the published structure of CRABP<sub>II</sub>•RA (PDB ID: 1CBS).

Although the  $\beta$ C- $\beta$ D hairpin loop shows some minor differences, they are not as significant as what had been observed when the structures of apo-R111M and CRABP<sub>II</sub>•RA were compared (Figure 3-13). Chen et al.<sup>2</sup> had suggested that this hairpin loop opens in apo-CRABP<sub>II</sub> to allow RA entry and then closes when RA enters the pocket, while apo-WT-CRABP<sub>II</sub> does not show such a movement for the hairpin loop (Figure 3-14-A). The  $\alpha$ 2 helix has an identical conformation in both structures, which indicates that this helix does not change its conformation upon ligation, contrary to the R111M structure. So the major differences observed when comparing the structures of apo-R111M and CRABP<sub>II</sub>•RA were actually the result of the R111M mutation, not RA binding.

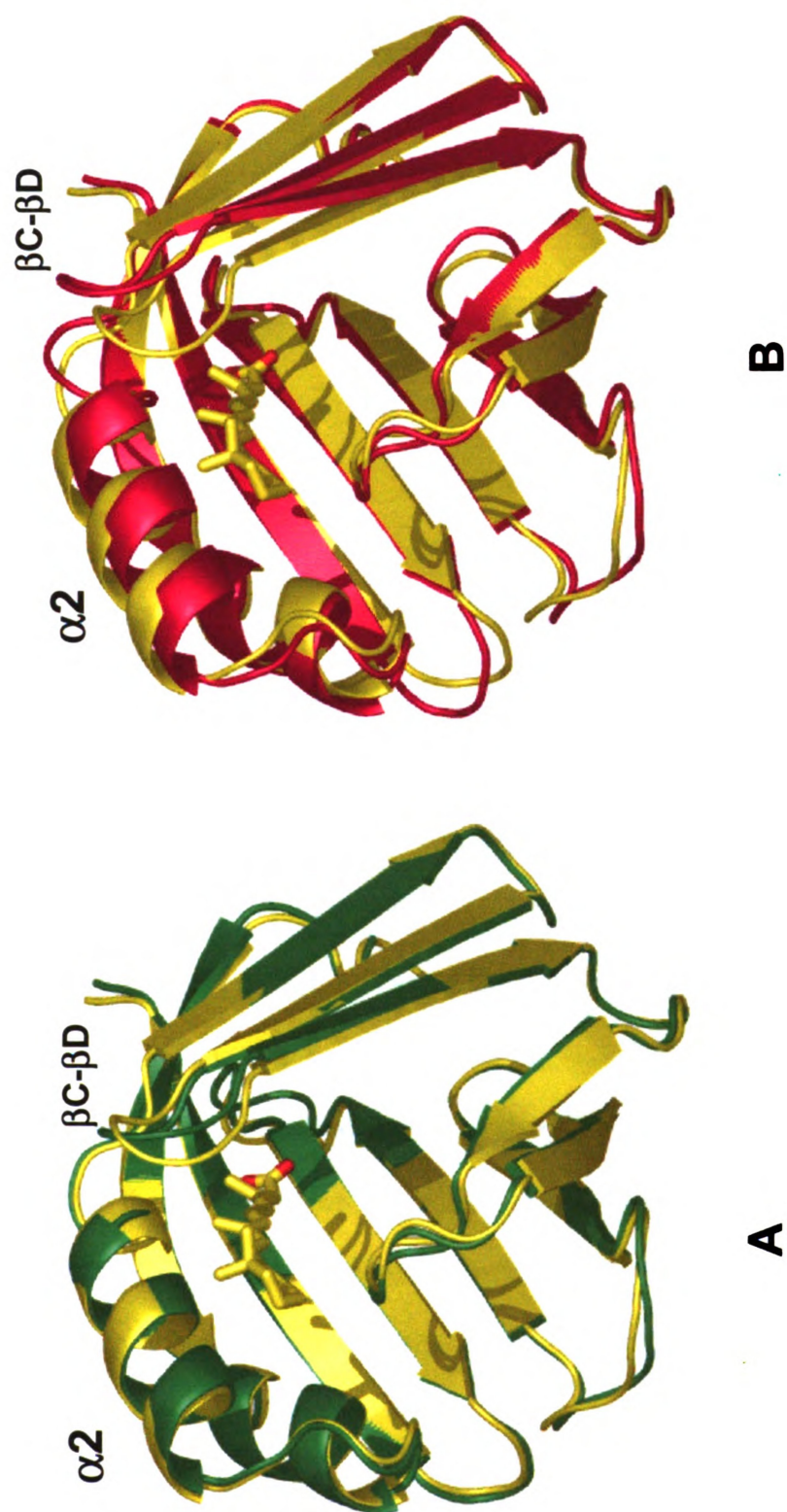


Figure 3-14 The superimposed structures of WT CRABPII•RA (yellow) and (A) Mol A of apo-WT (green); and (B) Mol B of apo-WT (hot pink). RA, in the pocket of WT CRABPII•RA, is shown.

Based on the conformational differences between apo-R111M and CRABPII•RA, Chen et al.<sup>2</sup> suggested a three-step mechanism of ligand entry into CRABPII (Figure 3-13). This mechanism involved first portal opening involving unwinding of the  $\alpha 2$  helix and motion of the  $\beta C$ - $\beta D$  and  $\beta E$ - $\beta F$  hairpin loops, where unwinding of the  $\alpha 2$  helix was thought to be essential for RA entry. Three positively charged residues, Arg29, Arg59 and Arg132, were then thought to "direct" the RA carboxylate group into the binding cavity. These residues would then flip out of the binding cavity in the last binding step. Our structure of apo-WT CRABPII, however, is not consistent with this hypothesis because we do not see the same conformational changes between the apo and bound CRABPII (Figure 3-14). First, the  $\alpha 2$  helix in apo-WT CRABPII and the CRABPII•RA complex is identical in most of our structures, in contrast to the large conformational change seen in the R111M structure (compare Figures 3-13 and 3-14-A). Furthermore, as shown in Figure 3-15, both Arg29 and Arg132 have identical conformations in the bound and apo forms of the protein, indicating that conformational change of these residues is not required to direct RA binding. Arg59, the last residue of the  $\beta C$ - $\beta D$  hairpin loop (Figure 3-15), located at the portal of the pocket, has the same conformation in Xtal1 and Xtal2 but is different from CRABPII•RA and R111M. In both apo-WT and CRABPII•RA, the side chain of Arg59 is located toward the exterior of the pocket, however in CRABPII•RA it is forced away from the portal of the pocket to provide enough space for the gem-dimethyl group of the ionone ring. In contrast, in the R111M structure, the conformation of Arg59 is strikingly different and unlike other structures it bends toward the interior of the pocket (Figure 3-15) and has a water network interaction with Arg132.<sup>2</sup> Therefore we believe that the three-step mechanism of RA entry proposed

by Chen et al.,<sup>2</sup> based on comparing the structure of apo-R111M and holo-CRABP II is not consistent with the structures of apo-WT CRABP II we have determined.

We conclude that, in fact, the most significant conformational differences between the structures of apo-R111M and the WT CRABP II•RA are due to the mutation of the very conserved Arg111 residue.

Comparing the structure of Mol B and CRABP II•RA, Figure 3-14-B, shows that the C-terminus of the  $\alpha 2$  helix and the loop connecting  $\alpha 2$  to  $\beta B$  are significantly different in the two structures. As we explained above, the crystal-packing environment of Mol B results in the altered conformation of this region and therefore its altered conformation is an artifact of crystal-packing. The conformation of the  $\alpha 2$  helix in Mol B does not leave enough space for RA to enter and bind inside the pocket: the second helical turn of the helix is very close to where RA binds inside the pocket and its residues would collide with RA. Also Arg132 has a different conformation in Mol B than in Mol A and is also in the way of RA entry. Therefore we believe that Mol A shows the “real” conformation of apo-WT CRABP II (or at least the closest that we can get to “real” from a crystal structure) and is compatible with RA entry.

Here using three independent sets of diffraction data, collected on three different crystals of apo-WT we demonstrated that apo-WT has an almost identical structure to CRABP II•RA with only minor differences in the  $\beta C$ - $\beta D$  hairpin loop. This hairpin loop is at the portal of the binding cavity and tends to close up the opening of the pocket upon RA ligation (Figure 3-14-A).



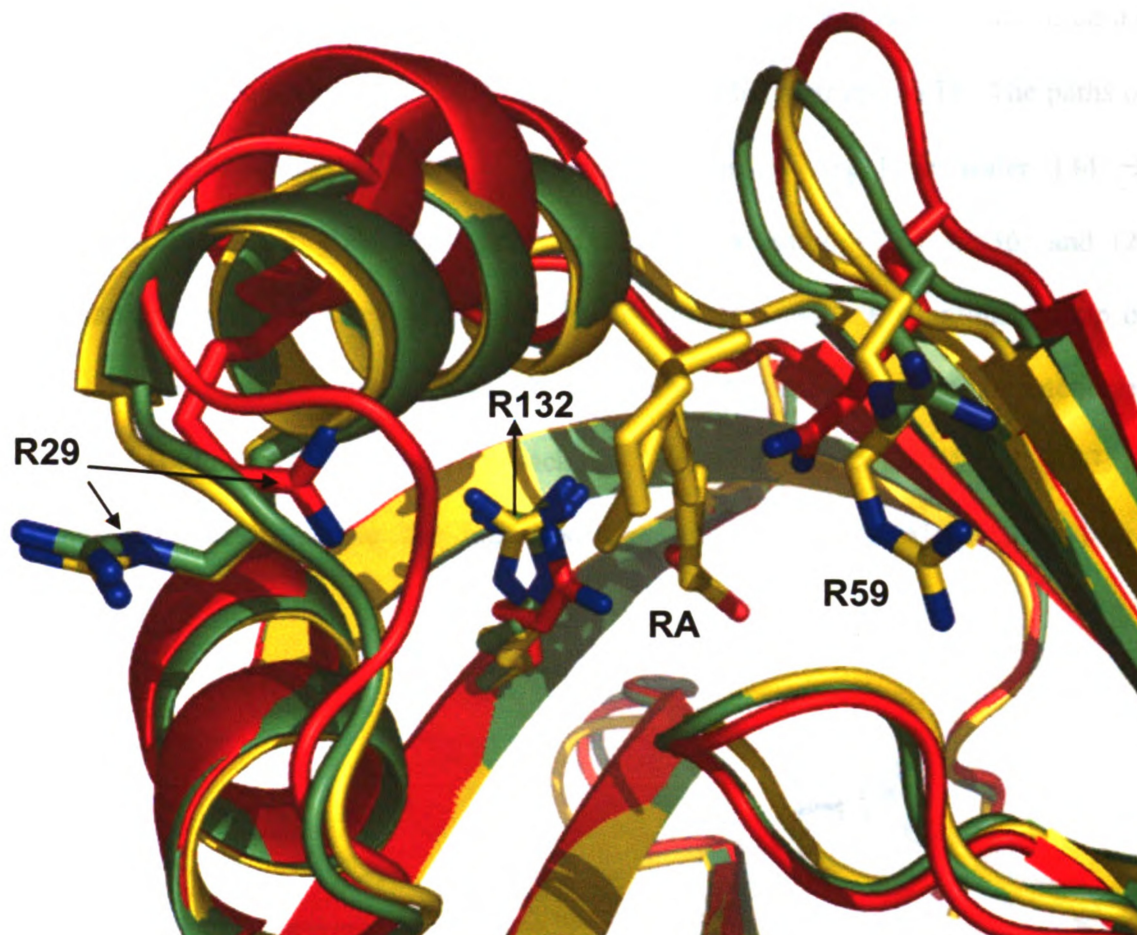


Figure 3-15 Superimposed structures of Mol A of apo-WT (green), Mol A of R111M (red) and WT CRABP II•RA (yellow). Residues Arg29, Arg59, and Arg132 of each structure and RA in CRABP II•RA are shown.

It is interesting to note that very similar structures of these molecules, especially in the  $\alpha 2$  helix and loop region correlate with the very similar solvent-mediated interactions in the binding cavity of both CRABP II•RA and apo-WT Mol A. This occurs because an acetate (Xtal1) or a chloride (Xtal2) ion replaces the carboxylate group of RA in Mol A of the apo-WT structure. This carboxylate or chloride serves to bridge Arg111 to the opposite side of the binding cavity, stabilizing this region. Both acetate and RA bind inside the cavity and form tight hydrogen bonds with Arg132 and Tyr134 (Figures

3-4 and 3-16). The network starts at Arg111 and ends at Val33 and Ala36, located on the  $\alpha 2$  helix and the loop region (similar to the network in Mol A of apo-WT). The paths of the networks (Figure 3-16) are: (1) guanidino group of Arg111  $\rightarrow$  water 134  $\rightarrow$  carboxylate of RA  $\rightarrow$  guanidino group of Arg132  $\rightarrow$  carbonyl of Ala36; and (2) guanidino group of Arg111  $\rightarrow$  water 134  $\rightarrow$  carboxylate of RA  $\rightarrow$  guanidino group of Arg132  $\rightarrow$  water 28  $\rightarrow$  carbonyl of Val33. In contrast, this water-mediated interaction is lost in Mol B of apo-WT due to crystal-packing. The differences in Mol A and Mol B of the apo structures illustrate the increased flexibility of apo-CRABP II over holo-CRABP II.

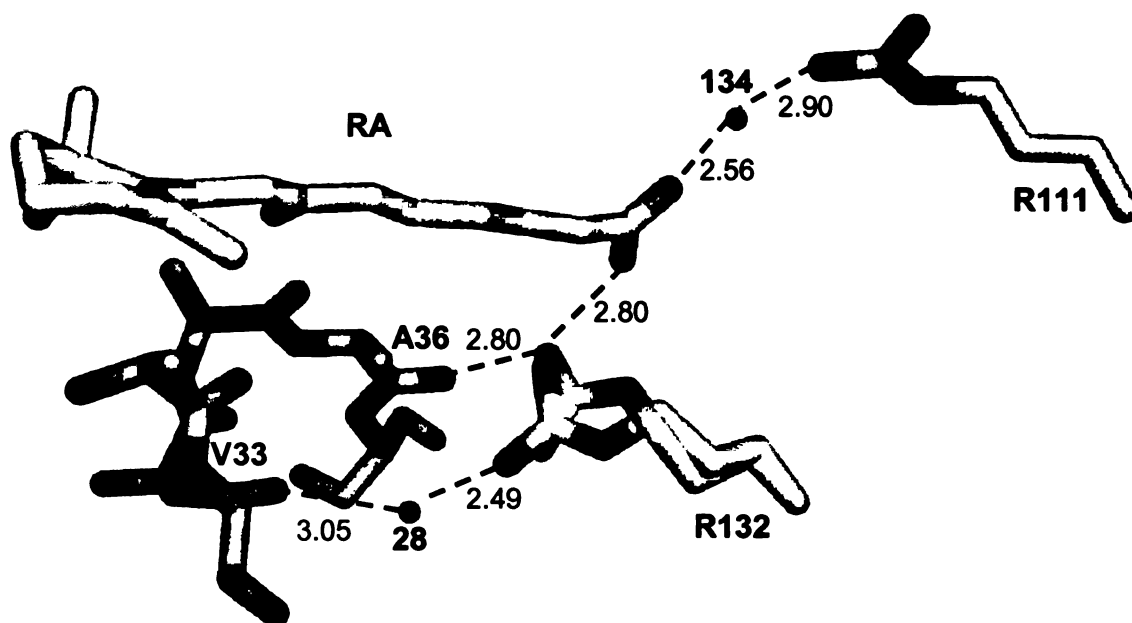


Figure 3-16 The water-mediated interaction between Arg111 and Val33 and Ala36 in WT CRABP II-RA. The distances are in angstroms.

### 3.6 Comparison between the Crystal and NMR Structures

The structures of apo-WT and the R111M mutant of CRABP<sub>II</sub> have both been determined using NMR by Wang et al.<sup>4,11</sup> The authors have compared these two structures with the previously published crystal structure of CRABP<sub>II</sub>•RA, and shown that the structure of R111M is more similar to holo-CRABP<sub>II</sub> than apo-CRABP<sub>II</sub> in both structural and dynamic properties.<sup>4</sup> We compared these two solution structures to our crystal structure of apo-WT CRABP<sub>II</sub>. This comparison has shown that the NMR structure of the R111M mutant (PDB ID: 1BM5) is similar (with a high RMSD of 3.56 Å) to that of apo-WT, which itself is identical to holo-CRABP<sub>II</sub>. However the NMR solution of apo-WT CRABP<sub>II</sub> (PDB ID: 1BLR) is dramatically different from our crystal structure of this protein. The  $\alpha$ 1-loop- $\alpha$ 2 region and the  $\beta$ C- $\beta$ D,  $\beta$ E- $\beta$ F and  $\beta$ G- $\beta$ H loops have significantly different conformations in this structure. Particularly  $\alpha$ 2 is completely unwound and has opened up toward the exterior of the pocket. We have calculated the secondary structure of this molecule using DSSP program,<sup>5</sup> which shows that in fact the molecule does not have a second helix. These changes were thought to be necessary for opening the pocket and ligand entry. However we believe that since the ionone ring of RA is exposed at the portal, the ligand seems to find its way toward the interior of the pocket without requiring significant structural differences between the apo- and holo-CRABP<sub>II</sub>. In other words we believe that loosing the integrity of the structure, as observed in the NMR structure of apo-WT, should not be necessary for RA entry. Our three different crystal structures of apo-WT show that the apo- and holo-CRABP<sub>II</sub> are very similar and that the portal of the pocket is large enough to let RA enter the pocket without causing dramatic structural changes. In light of both the NMR structure of apo-

R111M and of our numerous structures of apo-WT CRABP II, the NMR structure of the apo-WT protein is puzzling. In fact the structures of apo-CRABPI (PDB ID: 1CBI),<sup>12</sup> apo-M-FABP (PDB ID: 1FTP)<sup>13</sup>, apo-IFABP (PDB IDs: 1IFB and 1IFC)<sup>14,15</sup>, apo-CRBP II (PDB ID: 1OPA)<sup>16</sup> and apo-ALBP (PDB IDs: 1ALB and 1LIB)<sup>17,18</sup>, the only other members of the iLBPs protein family whose structures are known, are also quite similar to our apo-CRABP II crystal structure and none display the extreme structural changes seen in the apo-WT CRABP II NMR-determined structure.

### **3.7 Comparison between the Crystal Structures of Apo-CRABPI and Apo-CRABP II**

Our apo-CRABP II is monomeric while in apo-CRABPI, unlike other iLBPs, the two molecules in the asymmetric unit form a dimer. Five new hydrogen bonds are formed between the  $\beta$ D strands of Mol A and Mol B which make an intermolecular  $\beta$ -sheet between the two molecules. The movement of the  $\beta$ C- $\beta$ D hairpin loop of Mol A toward the exterior of the pocket was believed to be essential for the formation of the intermolecular  $\beta$ -sheet resulting in dimerization. Further it was suggested that formation of this 20-stranded double  $\beta$ -barrel is essential for opening the portal of the pocket and may be the ligand entry mechanism in CRABPI.<sup>12</sup>

We compared the two molecules of our apo-WT CRABP II with those of apo-CRABPI. The two structures are very similar except for the  $\beta$ C- $\beta$ D hairpin loop of Mol A. Unlike apo-CRABPI, this loop does not move toward Mol B to form an intermolecular hydrogen bond with its  $\beta$ D and therefore does not form a dimer. In fact the loop has a very similar conformation in apo- and holo-CRABP II, as explained above

(see comparison of the structures of apo- and holo-CRABP<sub>II</sub>). Therefore our crystal structure of apo-WT CRABP<sub>II</sub> is very similar to the structure of apo-CRABP<sub>I</sub>, except that it does not dimerize.

### 3.8 The Nuclear Localization Signal in CRABP<sub>II</sub>•RA

Although CRABP<sub>II</sub> is a cytosolic protein, upon binding to RA it translocates into the nucleus to transfer RA to the nuclear receptor RAR.<sup>19</sup> However, CRABP<sub>II</sub> does not have a recognizable nuclear localization signal (NLS) in its primary structure, leaving open the question of how CRABP<sub>II</sub> is recognized by adaptor proteins ( $\alpha$  importins) and why holo-CRABP<sub>II</sub> is far more efficiently localized in the nucleus than apo-CRABP<sub>II</sub>. Sessler et al.<sup>1</sup> observed structural differences between Mol A of apo-R111M and holo-CRABP<sub>II</sub> and calculated electrostatic surface potentials of these two proteins, (Figure 3-16, or Figure 3-B, Sessler et al.<sup>1</sup>).

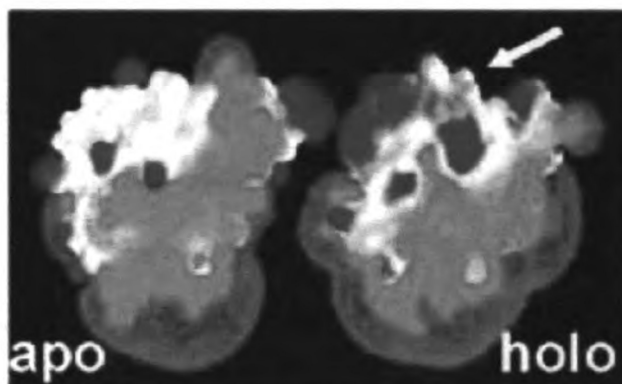


Figure 3-17 Computed electrostatic surface potentials of apo-R111M and holo-CRABP<sub>II</sub>. Basic, acidic, and neutral charges are denoted by blue, red, and white, respectively. A positively charged patch (arrow) is manifested in holo-CRABP-II.<sup>1</sup>

This comparison showed that CRABPII•RA has a positively charged patch on the two  $\alpha$  helices that is not present in apo-R111M. Three basic residues, Lys20 (on the second helical turn of the  $\alpha$ 1), Arg29 and Lys30 (both on the second helical turn of  $\alpha$ 2) make up this positive patch (Figure 3-17). These residues assume different conformations in holo-CRABPII vs. apo-R111M that lead to a positive electrostatic surface potential on holo-CRABPII and a neutral surface potential on apo-R111M. Comparison of the two structures by Sessler et al.<sup>1</sup> showed that upon RA binding these three residues assume an NLS-like structure, similar to what is observed in a classical NLS: SV40-T antigen. They showed that three SV40 peptide residues (Lys128, Lys129 and Lys131) could be superimposed on the three basic residues in CRABPII•RA (Figure 3-18-B and also Figure 3-G of Sessler et al.<sup>1</sup>). It should be mentioned that it seems the residues are rotated to be superimposed on each other and it is impossible to superimpose these three residues as shown by Sessler et al. We tried to reproduce the same overlaid picture (Figure 3-18), however as shown in the picture the best superimposition that we could produce positions residues very far from each other. The best superimposition was within a wide range of 2-10 Å, and is therefore is questionable.

However comparing the conformations of these residues in apo-WT with CRABPII•RA (our data) shows that Lys20, Arg29 and Lys30 have very similar conformations in apo- and holo-CRABPII (Figure 3-19-A). As shown in Figure 3-19-B, Arg29 and Lys20 have very similar conformations in both CRABPII•RA structures (our data vs. 1CBS). Though the side chain of Lys30 has a different orientation between our structure and 1CBS, it is solvent-exposed, and has a higher than average B factor in both structures, indicating significant side chain flexibility. Considering the fact that these



residues are long and located at the surface of the molecules, conformational differences and high B-factors in the different structures are not unexpected. Interestingly Lys20, Arg29 and Lys30 are well defined and have exactly the same conformation in Xtal1 and Xtal2.

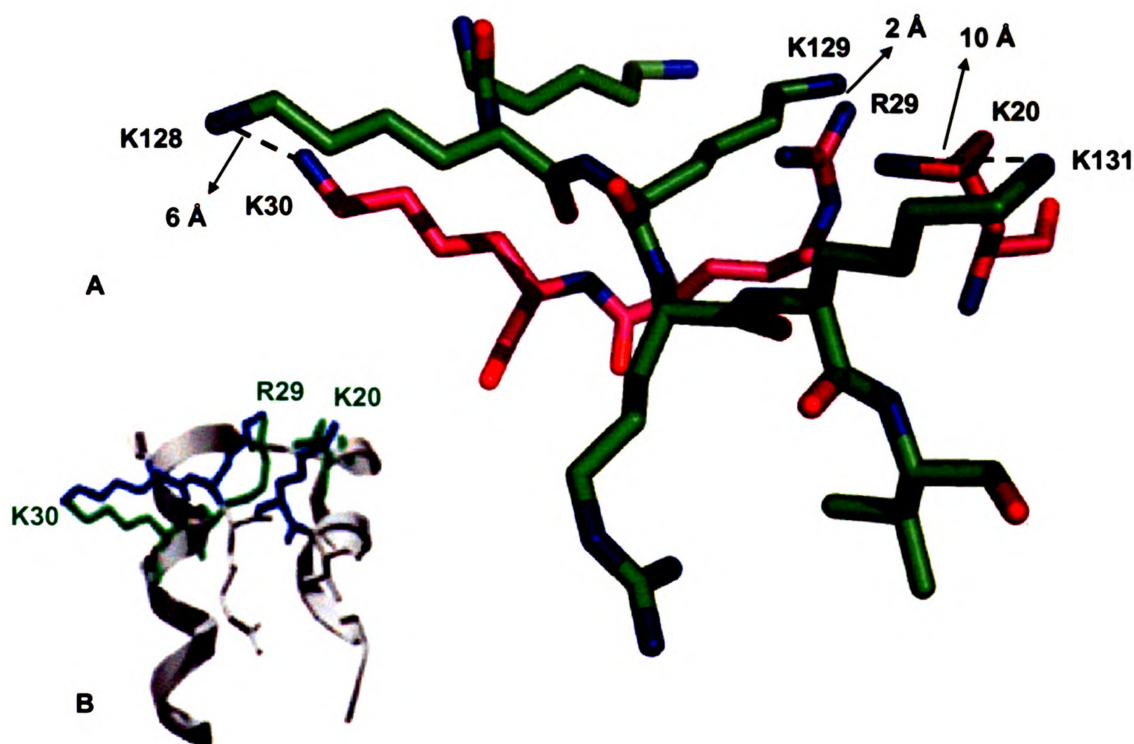


Figure 3-18 (A) Superimposition of NLS of SV40-antigen on the putative NLS residues of CRABP II·RA. (B) The same superimposition performed by Sessler et al.<sup>1</sup>

In short, the differences observed in the conformation of these three basic residues by Sessler et al.<sup>1</sup> were due to structural differences that are induced by the R111M mutation. As shown in Figure 3-19-A, movement of the  $\alpha 2$  helix in Mol A of the R111M structure causes dramatic conformational changes of Arg29 and Lys30. Consistent with these observations, apo-CRABP II also localizes somewhat to the nucleus, though clearly not as efficiently as holo-CRABP II. Partial nuclear localization of the apo-CRABP II is

evident in the experiments performed by Sessler et al.<sup>1</sup> It is clear that RA binding rigidifies the molecule, particularly in this region, perhaps decreasing the entropic barrier to importin binding and improving the efficiency of the translocation.



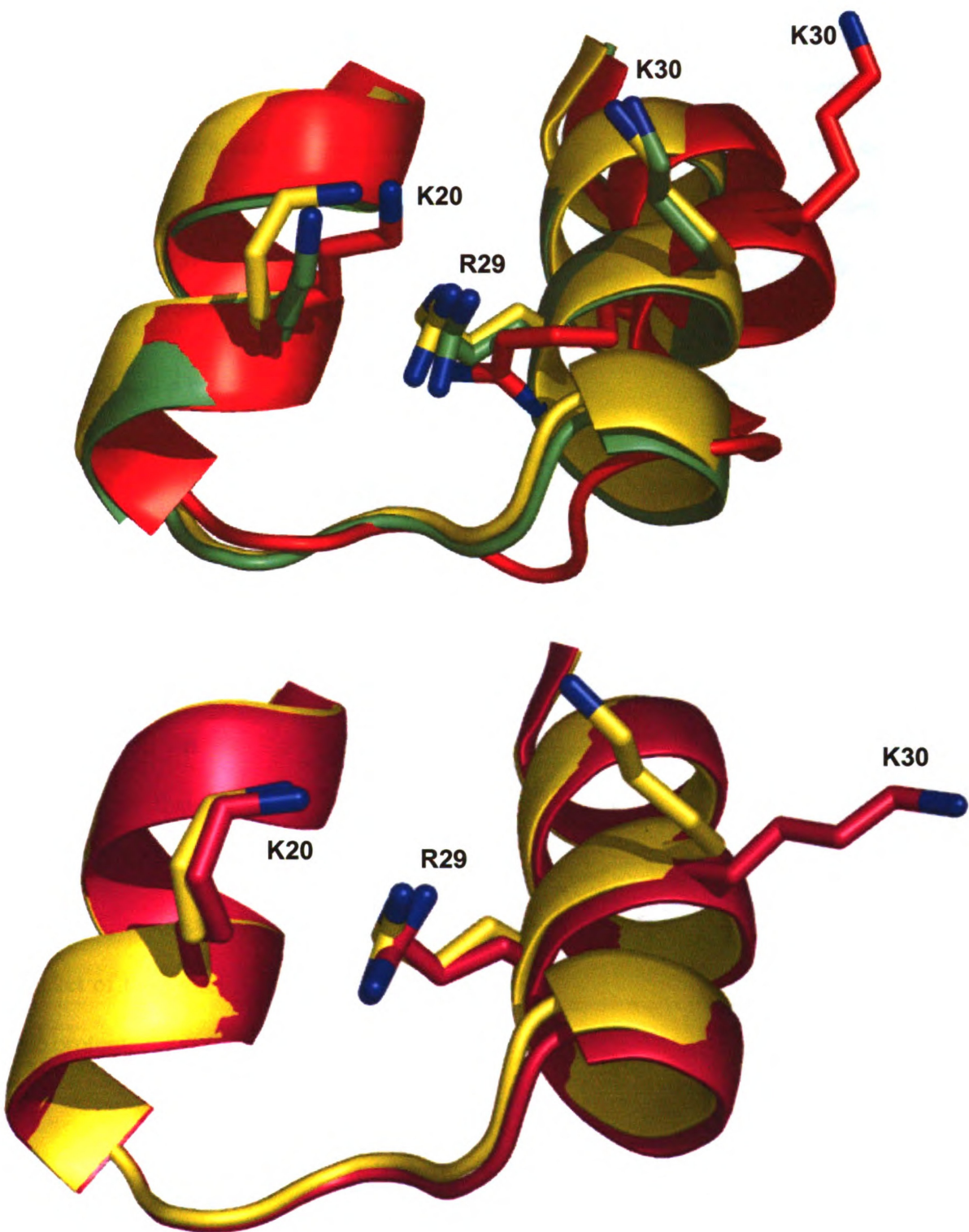


Figure 3-19 (A) Lys20, Arg29, and Lys30 in Mol A of Xtal1 (green), CRABPII.RA (yellow) and R111M (red); B) Lys20, Arg29, and Lys30 in our CRABPII.RA (2FR3, in yellow) and previously published structure of CRABPII.RA (1CBS, in purple)

### **3.9 Conclusions**

We have determined the first high-resolution structure of apo-CRABP<sub>II</sub>. Using three different data sets collected on apo-WT CRABP<sub>II</sub> we have shown that apo- and holo-CRABP<sub>II</sub> have very similar structures and that the apo-structure is capable of increased flexibility in several regions. Comparison of our apo-WT with the reported apo-R111M structure shows that mutation of Arg111, one of the conserved residues of CRABP<sub>II</sub> and a key residue in binding RA to the protein, causes major structural changes in the molecule that had previously been attributed to ligand binding. Our structures demonstrate that ligand binding leads to much less pronounced structural changes than previously thought, but does lead to overall increased rigidity of the structure. This decrease in flexibility may be an important determinant in the increased nuclear localization efficiency of the RA-bound protein. In addition our structures have demonstrated structural changes induced by crystal-packing in the molecule and have also shown that an apparently identical crystal can harbor demonstrative structural differences in the asymmetric unit.

The structures show the importance of water-mediated interactions inside the pocket of the proteins and their role as connectors to preserve the structural integrity of the protein.

### 3.10 Literature Cited

1. Sessler, R. J. & Noy, N. (2005). A ligand-activated nuclear localization signal in cellular retinoic acid binding protein-II. *Molecular Cell* **18**, 343-353.
2. Chen, X., Tordova, M., Gilliland, G. L., Wang, L. C., Li, Y., Yan, H. G. & Ji, X. H. (1998). Crystal structure of apo-cellular retinoic acid-binding protein type II (R111M) suggests a mechanism of ligand entry. *Journal of Molecular Biology* **278**, 641-653.
3. Kleywegt, G. J., Bergfors, T., Senn, H., Lemotte, P., Gsell, B., Shudo, K. & Jones, T. A. (1994). Crystal-Structures of Cellular Retinoic Acid-Binding Protein-I and Protein-II in Complex with All-Trans-Retinoic Acid and a Synthetic Retinoid. *Structure* **2**, 1241-1258.
4. Wang, L. C. & Yan, H. G. (1998). NMR study suggests a major role for Arg111 in maintaining the structure and dynamical properties of type II human cellular retinoic acid binding protein. *Biochemistry* **37**, 13021-13032.
5. Kabsch, W. & Sander, C. (1983). Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **22**, 2577-2637.
6. Bailey, S. (1994). The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallographica Section D-Biological Crystallography* **50**, 760-763.
7. Gunasekaran, K., Hagler, A. T. & 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.
8. Furey, W., Wang, B. C. & Sax, M. (1982). Crystallographic Computing on an Array Processor. *Journal of Applied Crystallography* **15**, 160-166.
9. Hendrickson, W. A. (1985). Stereochemically Restrained Refinement of Macromolecular Structures. *Methods in Enzymology* **115**, 252-270.
10. Konnert, J. H. (1976). Restrained-Parameter Structure-Factor Least-Squares Refinement Procedure for Large Asymmetric Units. *Acta Crystallographica Section A* **32**, 614-617.

11. Wang, L. C., Li, Y., Abildgaard, F., Markley, J. L. & Yan, H. G. (1998). NMR solution structure of type II human cellular retinoic acid binding protein: Implications for ligand binding. *Biochemistry* **37**, 12727-12736.
12. Thompson, J. R., Bratt, J. M. & Banaszak, L. J. (1995). Crystal-Structure of Cellular Retinoic Acid-Binding Protein-I Shows Increased Access to the Binding Cavity Due to Formation of an Intermolecular Beta-Sheet. *Journal of Molecular Biology* **252**, 433-446.
13. Tan, Q., Lou, J. H., Borhan, B., Karnaukhova, E., Berova, N. & Nakanishi, K. (1997). Absolute sense of twist of the C12-C13 bond of the retinal chromophore in bovine rhodopsin based on exciton-coupled CD spectra of 11,12-dihydroretinal analogues. *Angewandte Chemie-International Edition in English* **36**, 2089-2093.
14. Sacchettini, J. C., Gordon, J. I. & Banaszak, L. J. (1989). Refined Apoprotein Structure of Rat Intestinal Fatty-Acid Binding-Protein Produced in Escherichia-Coli - (Protein-Structure X-Ray Crystallography Fatty Acid-Protein Interactions). *Proceedings of the National Academy of Sciences of the United States of America* **86**, 7736-7740.
15. Yee, V. C., Pratt, K. P., Cote, H. C. F., LeTrong, I., Chung, D. W., Davie, E. W., Stenkamp, R. E. & Teller, D. C. (1997). Crystal structure of a 30 kDa C-terminal fragment from the gamma chain of human fibrinogen. *Structure* **5**, 125-138.
16. Winter, N. S., Bratt, J. M. & Banaszak, L. J. (1993). Crystal-Structures of Holo and Apo-Cellular Retinol-Binding Protein-Ii. *Journal of Molecular Biology* **230**, 1247-1259.
17. Xu, Z. H., Bernlohr, D. A. & Banaszak, L. J. (1992). Crystal-Structure of Recombinant Murine Adipocyte Lipid-Binding Protein. *Biochemistry* **31**, 3484-3492.
18. Xu, Z. H., Bernlohr, D. A. & Banaszak, L. J. (1993). The Adipocyte Lipid-Binding Protein at 1.6-a Resolution - Crystal-Structures of the Apoprotein and with Bound Saturated and Unsaturated Fatty-Acids. *Journal of Biological Chemistry* **268**, 7874-7884.
19. Budhu, A. S. & Noy, N. (2002). Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest. *Molecular and Cellular Biology* **22**, 2632-2641.

## Chapter IV

# Mutation, Over-expression, Purification, Crystallization, X-ray Diffraction Analysis, Structure Solution and Refinement of CRABPII and Corresponding Complexes

### 4.1 Protein Mutation, Over-expression and Purification of CRABPII

#### Protein Mutation

We made the F15W and L19W mutants of CRABPII and the rest of the of the mutantions were prepared by our collaborators in the laboratory of Professor Babak Borhan (Chemistry Department, Michigan State University).<sup>1,2</sup> The mutations were constructed using the Stratagene QuikChange<sup>®</sup> site-directed mutagenesis kit and protocol, which is a simplified method to perform point mutations or delete/insert amino acids using a thermal cycling technique in combination with *Dpn* I digestion. This is a rapid, one day procedure, which results in a high rate of positive mutants. This method is performed using *PfuTurbo*<sup>®</sup> DNA polymerase and a temperature cycler. *PfuTurbo* replicates both plasmid strands without displacing the mutant oligonucleotide primers. The basic procedure, shown in Figure 4-1, is as follows: The method utilizes a DNA vector with an insert of interest and two synthetic oligonucleotide primers, carrying the desired mutation. The mixture along with dNTP (a generic term referring to the four deoxyribonucleotides: dATP, dCTP, dGTP and dTTP) and *PfuTurbo* polymerase is put through temperature cycling PCR reaction. As a result both primers, each complementary to opposite strands of the vector, are extended by *PfuTurbo*.

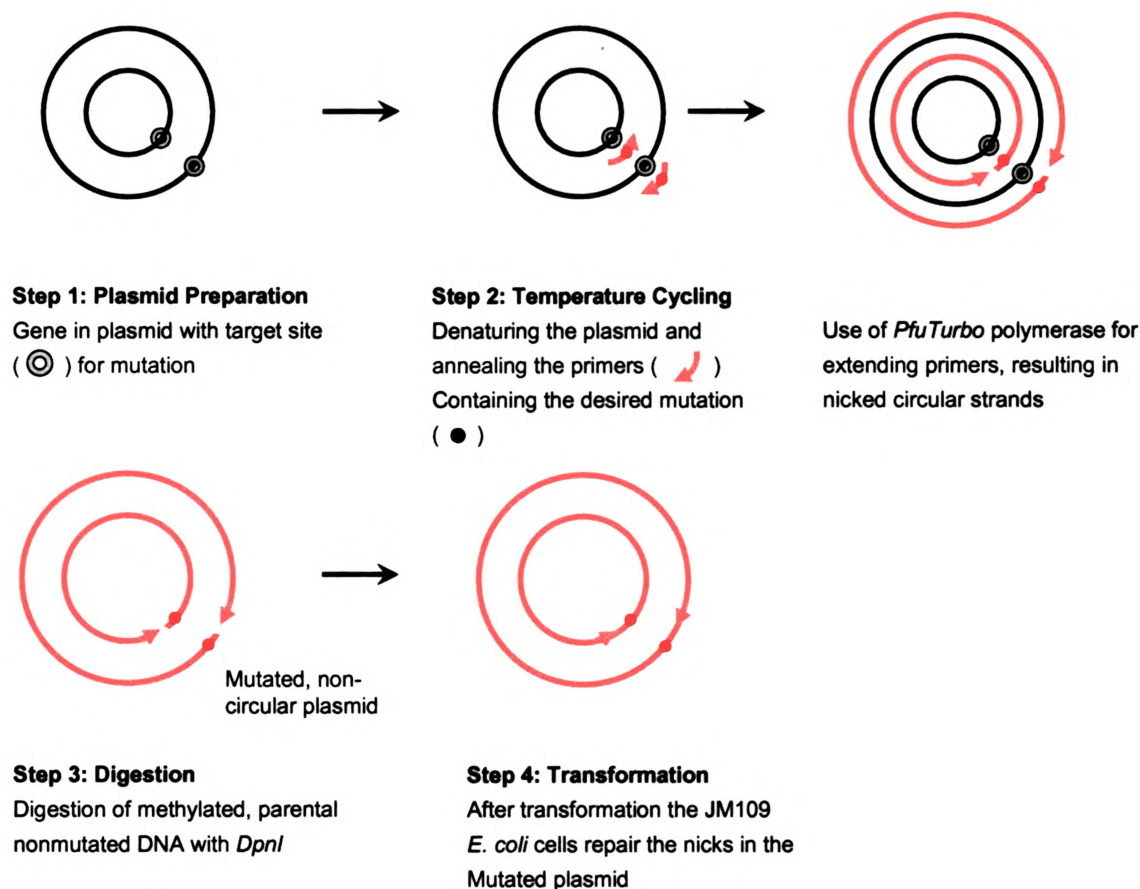


Figure 4-1 Overview of the QuikChange site-directed mutagenesis method

This step produces a mutated plasmid with staggered nicks. The product is treated with *Dpn I* endonuclease. *Dpn I* is used for digestion of parental, unmutated DNA and targets GATC sequence (blunt cut between A and T). However it is only specific for methylated DNA and where methylated adenine residues exist. DNA isolated from almost all *E. coli* strains is dam methylated and therefore susceptible to *Dpn I* digestion. The *Dpn I* restriction enzyme effectively destroys the parental DNA and leaves the linear, mutated plasmid which does not carry any methylated adenine. The nicked vector containing the desired mutation is then transformed into the host strain of choice, XL1-Blue or JM109 *E. coli* (we used JM109 cells, Stratagene) to make a circular vector and repair the nicks. The small amount of parental DNA required for this method, the high fidelity of

*PfuTurbo* and the low number of thermal cycles all contribute to the high mutation efficiency and decreased potential for generating random mutations during the reaction.

### PCR Reaction

Table 4-1 summarizes the PCR reagents and conditions for mutation of F15W and L19W mutants of CRABP II.

Table 4-1 PCR conditions for mutation of F15W and L19W CRABP II mutants

| Reactant                | Amount   | PCR Sequences |                      |
|-------------------------|----------|---------------|----------------------|
| Template DNA (10 ng/μL) | 5 μL     | # Cycles      | Temperature and Time |
| Primer 1                | 15 pmol  | 1 X           | 94 °C for 5 min      |
| Primer 2                | 15 pmol  | 16 X          | 94 °C for 2 min      |
| dNTP                    | 1 μL     |               | 55 °C for 2 min      |
| <i>PfuTurbo</i>         | 1 μL     |               | 72 °C for 2 min      |
| 10 x pfu buffer         | 10 μL    | 1 X           | 72 °C for 10 min     |
| Water                   | 50 – rxn | 1 X           | 23 °C for 10 min     |

The primers used were as follows:

|       |         |                                |
|-------|---------|--------------------------------|
| F15W: | bbb0049 | 5' CGATCGGAAACTGGGAGGAATTGCTC  |
|       | bbb0050 | 5' GAGCAATTCCTCCCAGTTTTCCGATCG |
| L19W: | bbb0047 | 5' CGAGGAATTGTGGAAAGTGCTGGGG   |
|       | bbb0048 | 5' CCCCAGCACTTCCACAATTCCTCG    |

### Melting Temperature Calculation for Primers

Primers should ideally have melting points  $\geq 78$  °C, and end in at least one, if not more, GC base pairs.

$$\text{Primer melting point} = 81.5\text{ °C} + (0.41) \times (\%GC) - 675 / N - \% \text{ of mismatch}$$

where, N is the length of the primer.

After the PCR is completed, 10  $\mu$ L of the sample was removed for gel analysis, and 1  $\mu$ L *Dpn* I was added to the remaining solution to digest the parental DNA. The reaction was incubated for 1-2 hours at 37 °C, and then transformed into JM109 *E. coli* competent cells (Stratagene). For best results, super competent cells were used (Stratagene,  $10^6 - 10^8$  cfu/ $\mu$ g) as self prepared competent cells were usually not efficient enough ( $10^3$  cfu/ $\mu$ g). These steps are standard practice, and were performed after every mutagenesis PCR.

The mutants were screened by DNA sequencing. In order to ensure that there were no unintended mutations, the entire sequences of the mutated genes were determined.

#### Transformation of Mutated Plasmid into Competent Cells

We used a modified protocol developed by Sambrook et al.<sup>3</sup> Sterile conditions must be maintained throughout this protocol. An aliquote (0.1 mL) of competent cells was thawed on wet ice for approximately 5 minutes. 1-10  $\mu$ L of the plasmid DNA (10-100 ng) was added to the competent cells, and mixed **gently** by tapping the tube (**competent cells should never be vortexed**). The samples were returned to the ice bath immediately. It is very important to remember that the amount of plasmid DNA must never exceed 0.1 the volume of the component cells; i.e. use less than 10  $\mu$ L of plasmid



for 100  $\mu$ L of component cells. Exceeding this volume will significantly reduce the transformation efficiency. The samples were allowed to stand on ice for 30 minutes. Then they were heat shocked in a water bath at 42  $^{\circ}$ C for 120 seconds. The samples were returned on ice for 1-2 minutes. 0.5 mL of LB medium was added to each sample and mixed gently by inversion. Then samples were incubated at 37  $^{\circ}$ C for 1 hour. Each tube was microfuged for 30 seconds to pellet the cells. Using a pipettman approximately 0.5 mL of supernatant solution was removed, while being careful not to disturb the cell pellet. Using the pipettman the pellet was resuspended in the remaining liquid. The suspension was transferred to an LB/agar plate containing the appropriate antibiotic, Amp (100  $\mu$ g/mL), and the solution was spread over the surface of the plate. Finally the plates were incubated for 16-20 hours at 37  $^{\circ}$ C. A good practice in transforming a ligation mixture is to not transfer the entire transformation onto a single plate. Instead it is better to split the ligation into two parts and plate 90  $\mu$ L of sample onto one plate and 10  $\mu$ L of it onto a second plate. This way if there are a lot of cells that grow up, we can still have a good chance of getting single colonies on the plate containing the lower volume of transformation solution.

### Plasmid Purification

Following successful transformation into JM109 competent cells, a colony was grown up in 500 mL LB (containing Amp<sup>100</sup>) and purified using the Qiagen<sup>®</sup> maxi plasmid purification kit and following the “maxi prep” protocol. Upon completion of the Qiagen<sup>®</sup> purification, the recovered DNA was resuspended in approximately 400  $\mu$ L sterile water, and analyzed by UV-vis for determination of its concentration and purity.

Concentration of DNA sample ( $\mu\text{g}/\mu\text{L}$ ) =

$$\text{Abs.}^{260} \times (50 \mu\text{g} / 1000 \mu\text{L}) \times (\text{volume of DNA used} / \text{total volume UV sample})$$

$$\text{Purity of DNA sample} = \text{Abs.}^{260} / \text{Abs.}^{280}$$

Purity of 1.8 is considered a pure sample, above 1.8 having RNA contamination and below 1.8 having protein contamination. In practice we will consider a sample pure enough, when this value is  $\sim 1.7$ .

#### Sample Preparation for DNA Sequencing

In a sterilized 0.5 mL eppendorf tube, 1  $\mu\text{g}$  DNA and 30 pmol primer (5'-TGTTAGCAGCCGGATCTGCTCG-3' for pET-17b plasmid, approximately 75 base pairs upstream from the desired sequence starting position) were combined and the sample was made up to a total volume of 12  $\mu\text{L}$  with sterile water (not buffer). The DNA was purified by a Qiagen column and resuspended in sterile water. Failure to do this will result in poor sequence data.

#### Protein Over-expression

Human recombinant CRABP II was expressed and purified following the procedure described by Wang et al.<sup>4</sup> Expression vector pET-17b containing the CRABP II gene was transformed into *E. coli* strain BL21(DE3)pLysS cells (Stratagene). The transformed cells were grown at 37 °C in an LB agar plate containing both Amp<sup>100</sup> (CRABP II plasmid is Amp resistant) and Clm<sup>25</sup> (the BL21(DE3)pLysS cells are Clm resistant). A single colony was inoculated in 100 ml LB containing the same amount of

the two antibiotics and grown overnight with shaking at 250 rpm at 37 °C. This culture was then transferred to 1 liter of LB with the same antibiotics. After OD<sup>600</sup> reached between 0.6 to 0.8, samples were induced by 0.4 mM IPTG (isopropyl-1-thio-β-D-galactopyranoside) and growth was continued for another 4 to 5 hours at 30 °C. Then cells were harvested, (6000 rpm, 20 min, 4 °C) by centrifugation and frozen in aliquotes, containing 2 L of growth, at -80 °C for later use.

### Protein Purification

The frozen cells obtained from a 2 L growth were thawed on ice and resuspended in 100 ml of 10mM TRIS-HCl pH 8.0. The suspended cells were lysed by five 45-second bursts of sonication (probe sonicator, 60% power) on ice and centrifuged 15 minutes at 4 °C. The supernatant containing CRABPII protein was subjected to FastQ (Q-sepharose Fast Flow Resin) column chromatography (80 ml of resin with a capacity ~ 20 mg of protein, Amersham Biosciences).

FastQ, which is a quaternary ammonium functionalized resin, is a strong anion exchanger. The column was pre-equilibrated with 200 mL of 10 mM TRIS-HCl pH 8.0 buffer. If the column was previously used, it had to be cleaned with ~ 300 mL of 2 M NaCl and then equilibrated with the same TRIS-HCl buffer. The supernatant was loaded and flow through was collected. Then the column was washed using the same buffer as used for equilibrating the column and the flow through was collected again. These samples were later assayed for protein content using the Bradford assay,<sup>5</sup> and SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Little protein should be present in the flow through, otherwise the column is overloaded. If the column is overloaded the flow

through must be concentrated and run through the column again. Finally proteins were eluted using an NaCl gradient of 0 to 200 mM and eluant was collected in 5 mL fractions in plastic test tubes. The apparatus, used to run the gradient, is shown in Figure 4-2.

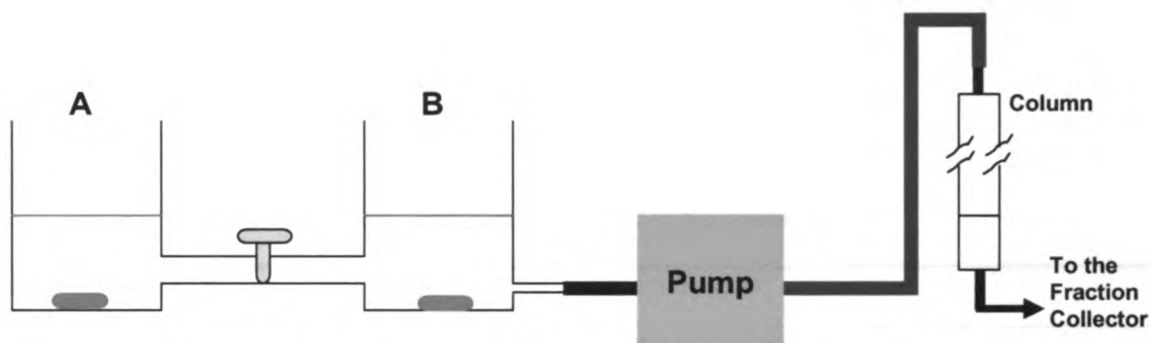


Figure 4-2 Schematic drawing of the apparatus used for running NaCl gradient in PhastQ purification of CRABP II. Objects not drawn to scale.

While the valve between the two compartments of the apparatus was closed, the more concentrated buffer (10 mM TRIS-HCl pH 8.0 and 1 M NaCl) was poured into compartment A and the more dilute buffer (10 mM TRIS-HCl pH 8.0 and 50 mM NaCl) in compartment B. The apparatus was placed on a stirrer plate and a magnetic stir bar was placed in each compartment. The apparatus was attached to a pump, which itself was connected to the top of the column. The pump was used to facilitate the flow of the buffers to the column, which significantly reduced the time required to run this column. CRABP II was eluted mostly at 100-150 mM NaCl. The fractions containing 16 kDa CRABP II were located by measuring the absorbance of the fractions using Bradford assay at 600 nm. These fractions were analyzed on a 20% SDS-PAGE to test for the purity of the sample. The most pure fractions were pooled together and desalted using a Vivaspin concentrator (Vivascience) with a 5,000 MW cutoff (10,000 MW cutoff concentrators showed some leak through the membrane).

Final chromatographic separation of the desalted protein solution was carried out on a FPLC system (BioLogics Duo Flow, BioRad) using a Source15Q (Amersham Biosciences) ion exchange column. Source 15Q is a quaternary ammonium and strong anion exchanger resin with a binding capacity ~ 20 mg per 80 mL of the resin. The FPLC protocol that we used is shown in Table 4-2.

Table 4-2 FPLC protocol. B: 2M NaCl solution

| Step | Volume (mL) | Description                                         |
|------|-------------|-----------------------------------------------------|
| 1    | 0           | Collect 3.0 mL fractions                            |
| 2    | 0           | Isocratic Flow pH = 8.1, 0%B, 20 mL, 5 mL/min       |
| 3    | 20          | Quad tec autozero                                   |
| 4    | 20          | Load/Inject 50 mL, 5 mL / min                       |
| 5    | 70          | Isocratic Flow pH = 8.1, 0% B, 20 mL, 5 mL/min      |
| 6    | 90          | Linear Gradient pH = 8.1, 0-4% B, 20 mL, 5 mL/min   |
| 7    | 110         | Isocratic Flow pH = 8.1, 4% B, 20 mL, 5 mL/min      |
| 8    | 130         | Linear Gradient pH = 8.1, 4-8% B, 20 mL, 5 mL/min   |
| 9    | 150         | Linear Gradient pH = 8.1, 8-20% B, 20 mL, 5 mL/min  |
| 10   | 170         | Linear Gradient pH = 8.1, 20-75% B, 20 mL, 5 mL/min |
| 11   | 190         | Isocratic Flow pH = 8.1, 100% B, 20 mL, 5 mL/min    |
| 12   | 210         | Isocratic Flow pH = 8.1, 0% B, 20 mL, 5 mL/min      |
| 13   | 230         | End of Protocol                                     |

The protocol has been optimized for CRABP<sup>II</sup> purification. CRABP<sup>II</sup> proteins elute at ~ 4% of 2 M NaCl.

Fractions were located by measuring the Abs.<sup>280</sup> of the fractions and analyzed on a 20% SDS-PAGE (Figure 4-3). The most pure fractions (more than ~ 95% pure) were

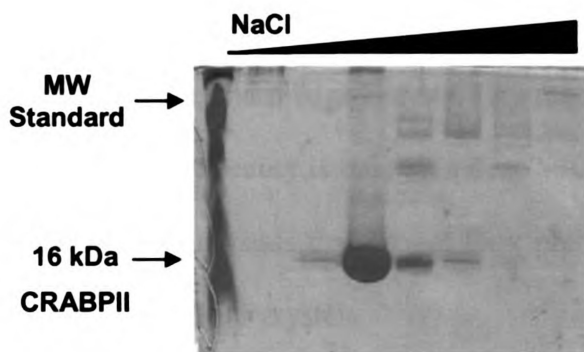


Figure 4-3 FPLC purification SDS-PAGE.

pooled together, concentrated and buffer exchanged (100 mM TRIS-HCl pH 8.0) in Vivaspin concentrators to a concentration of ~ 16-37 mg/ml. The concentration was determined by measuring the Abs.<sup>280</sup> of the protein in the buffer solution. The final concentration of the samples depended on the crystallization experiments. Some proteins crystallized in lower and some in higher concentrations, which were found by trial and error. The pure, concentrated protein solution was divided into 100  $\mu$ l aliquots and stored at -80 °C. All the steps of PhastQ and FPLC purification were performed at 4 °C.

## 4.2 Crystallization and X-Ray Data Collection

All the crystals were grown using the hanging drop (vapor diffusion) method (Figure 4-4). In this method, a drop (2-5  $\mu$ L) containing a mixture of protein and reservoir solution is equilibrated against the reservoir (500-1000  $\mu$ L). As a standard procedure, at least 215 different crystallization

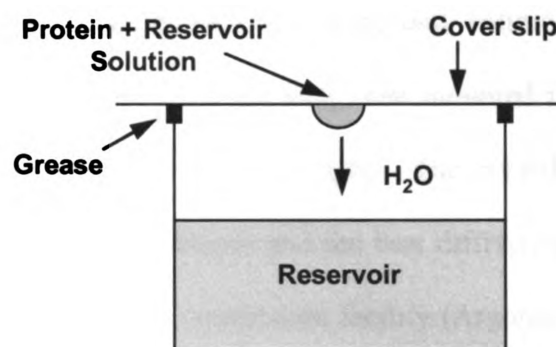


Figure 4-4 Hanging drop method for crystallization of proteins

conditions were tried for each protein to find the conditions which could produce crystals.

The crystallization condition was then optimized to produce better diffracting quality crystals. Single, sharp-edged crystals are most desirable but we should remember that in crystallography “beauty is only skin deep”. Good looking crystals might be salt crystals and sometimes crystals that do not look pretty diffract to high resolution as it was the case for some of our crystals.

One way that was used to improve the quality of crystals was changing the concentrations of the crystallization reagents of the condition that produced the initial crystals. Another way was to change the protein concentration by varying the volume of the protein drop with respect to the reservoir drop on the cover slip. If no crystal was obtained in these trials, the protein was concentrated to a higher or lower concentration. Also sometimes increasing the volume of the drop (mixture of protein and reservoir) produced crystals by decreasing the rate of equilibration and increasing the amount of protein in the drop. All apo-WT and mutants of CRABP II were crystallized at room temperature. Often crystals appeared in less than three days and reached their maximum size in a week.

Crystals were cryo-protected by quick soaking in a cryoprotectant solution consisting of the reservoir solution and 30% glycerol. Crystals were then mounted in nylon loops (Hampton Research) and flash-frozen using liquid nitrogen. The crystals were screened at home using a Rigaku RU-200 X-ray generator and the best diffracting crystals were sent to the Advanced Photon Source (APS) synchrotron facility (Argonne, IL) for better data collection. All the data for different crystals were collected under a stream of nitrogen gas at  $\sim -160^{\circ}\text{C}$ .

CRABP<sub>II</sub> crystals proved to diffract strongly at very high resolution. Therefore in order to avoid overloaded spots at high resolutions, two data sets were collected at high and low resolutions and were then merged together. Data sets were integrated using DENZO and scaled and merged using SCALEPACK from the HKL package.<sup>6</sup> The diffraction data statistics for all the apo- and holo-structures are shown in Tables 4-3 and 4-4, respectively.

#### 4.2.1 Apo-WT

Apo-WT CRABP<sub>II</sub> crystals were successfully grown at room temperature (25 °C). 2  $\mu$ L of the final purified protein (~ 16 mg/mL) in 100 mM TRIS-HCl pH 8.0 was mixed with an equal volume of a reservoir solution. The best crystals appeared using reservoir solutions containing either one of these solutions: 1) 0.2 M sodium acetate, 0.1 M TRIS-HCl pH 8.0 (or 8.5), 30% (w/v) PEG 4000 (or 8000), which is identical to reagent # 22 of crystal screen 1 (Hampton Research Inc.); or 2) 0.1 M bis-TRIS propane pH 6.5, 30% (w/v) PEG 4000. Crystals were relatively small (0.03 mm  $\times$  0.05 mm  $\times$  0.05mm) but sharp-edged (Figure 4-5).

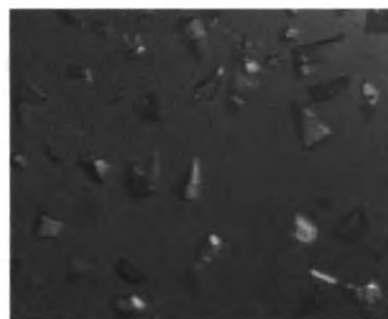


Figure 4-5 Crystals of apo-WT CRABP<sub>II</sub>. Photographed through a polarizer.

It is worth mentioning that our initial attempts in crystallizing the apo-WT CRABP<sub>II</sub> containing His<sub>6</sub>-tag, only resulted in poorly diffracting crystals. Removing the His<sub>6</sub>-tag led to well diffracting crystals. Therefore we designed all the mutants in the



pET-17b vector containing the CRABP<sub>II</sub> plasmid. This vector does not carry any purification tag.

Complete data were collected from each of three different crystals of apo-WT CRABP<sub>II</sub>, Xtal1, Xtal2 and Xtal3. Data were collected at the BioCars beamline (section 14-ID-B) of the advanced photon source (APS) synchrotron facility (Argonne, IL) using a MAR-CCD (165-mm) detector, at a wavelength of 1.00 Å, and crystal to detector distances of 60.00, 101.45, and 100.00 mm, respectively. 200 images were collected at 1° oscillation. Crystals obtained from 0.1 M TRIS-HCl pH 8.0, 0.2 M sodium acetate and 30% PEG 8000 were the best diffracting crystals (Xtal1, which diffracted to 1.35 Å), although the other crystals also diffracted to high resolutions (1.5-1.6 Å). Data collection statistics for Xtal1 and Xtal2 are reported in Table 4-3.

#### 4.2.2 Apo-F15W

1.5 µL of protein (~ 20 mg/mL) was mixed with 1.5 µL of the reservoir containing 0.2 M bis-TRIS propane pH 6.5 and 30% (w/v) PEG 4000 (room temperature). These conditions were identical to one of the apo-WT CRABP<sub>II</sub> crystallization conditions. The mutant produced relatively large (0.1 mm × 0.2 mm × 0.2mm), beautiful, orthogonal and sharp-edged crystals (Figure 4-6). Data were

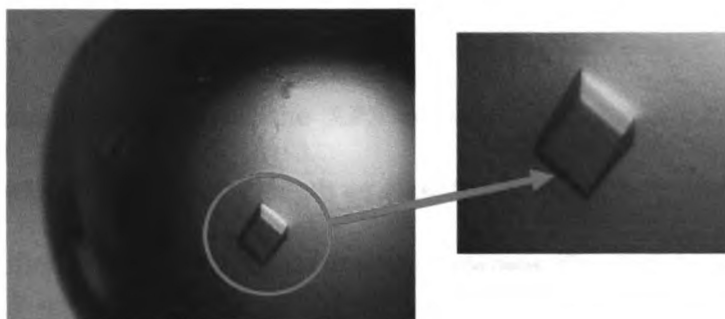


Figure 4-6 A crystal of F15W mutant.

collected at the SBC (19-ID) beamline at the APS synchrotron facility. 239 frames of images were collected at 1 ° oscillation at a wavelength of 0.99 Å. The crystal to detector distance was 120 mm. Data collection statistics are reported in Table 4-3.

#### 4.2.3 Apo-R132K:Y134F

1.5 µL of the mutant (~ 30 mg/mL) was mixed with 1.5 µL of the reservoir. The protein was crystallized in the same condition as the best diffracting crystals of apo-WT CRABPII (0.2 M sodium acetate, 0.1 M TRIS-HCl pH 8.5, 30% (w/v) PEG 4000) and at room temperature (Figure 4-7). Data was collected at the COM-CAT beamline (32-ID) at the APS synchrotron facility. 360 frames were collected at 1° oscillation at a wavelength of 1.00000 Å. Data collection statistics are reported in Table 4-3.



Figure 4-7 A Crystal of Apo-R132K:Y134F CRABPII

#### 4.2.4 Apo-R132K:Y134F:R111L:T54V:L121E

The mutant (~ 20 mg/mL) was crystallized in 0.1 M sodium acetate pH 5.8, 20% PEG 6000 and 5% ethanol and at room temperature. The ethanol proved to be necessary as an additive to form well-diffracting crystals. Crystals obtained without ethanol were not single and formed clusters. Diffraction data was collected at the COM-CAT beamline (32-ID) at the APS synchrotron facility. 300 frames were collected at 1° oscillation at a wavelength of 1.00 Å and crystal to detector distance of 80.0 mm. Data collections statistics are reported in Table 4-3.

As is shown in Figure 4-8, seeding helped the crystallization of this mutant. Small, non-single crystals that had formed at the edge of the crystal were crushed with a nylon loop and the loop was streaked over the cover slip. Larger and single crystals formed in less than three days. Although the large crystals were prettier than the smaller ones, the smaller crystals were the ones that diffracted to highest resolution. In an effort to produce retinal-bound crystals of this mutant, the crystals were soaked for 10 hours in 5-10% retinal in ethanol solution at 4 °C. The soaked crystals diffracted to even a higher resolution of 1.2 Å, although our effort in producing retinal-bound crystals was unsuccessful

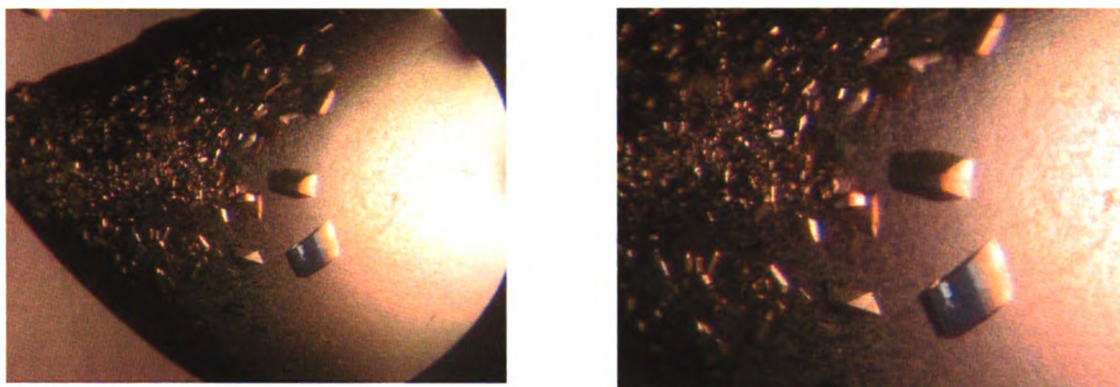


Figure 4-8 Crystals of R132K:Y134F:R111L:T54V:L121E. Streaking of the drop with the seeds of crushed crystals, from the same drop, produced larger and well diffracting crystals. Among the crystals the smaller ones proved to diffract better than the larger ones.

#### **4.2.5 Apo-R132K:R111L:L121E**

The triple mutant (~ 27 mg/mL) was crystallized in 0.1 M bis-TRIS propane pH 6.5 and 30% PEG 4000 at room temperature (identical to the apo-WT condition). Data was collected at the COM-CAT beamline (32-ID) at the APS synchrotron facility. 300

frames were collected at 1° oscillation at a wavelength of 1.00 Å. Data collection statistics are reported in Table 4-3.



Figure 4-9 Crystals of R132K:R111L:L121E

Table 4-3 X-ray data collection statistics for the apo-structures

|                                                               | Apo-CRABP II<br>Xtal 1   |    | Apo-CRABP II<br>Xtal 2   |    | Apo-F15W                  |    | Apo-<br>R132K:Y134F       |    | R132K:Y134F:R111L<br>T54V:L121E |    | Apo-<br>R132K:R111L:L121E |    |
|---------------------------------------------------------------|--------------------------|----|--------------------------|----|---------------------------|----|---------------------------|----|---------------------------------|----|---------------------------|----|
| Space group                                                   | P1                       | P1 | P1                       | P1 | P1                        | P1 | P1                        | P1 | P1                              | P1 | P1                        | P1 |
| <b>Unit cell dimensions</b>                                   |                          |    |                          |    |                           |    |                           |    |                                 |    |                           |    |
| <b>a (Å)</b>                                                  | 34.46                    |    | 34.48                    |    | 34.58                     |    | 34.590                    |    | 34.591                          |    | 34.553                    |    |
| <b>b (Å)</b>                                                  | 37.70                    |    | 37.22                    |    | 37.03                     |    | 37.230                    |    | 36.897                          |    | 37.220                    |    |
| <b>c (Å)</b>                                                  | 56.56                    |    | 57.60                    |    | 58.69                     |    | 58.770                    |    | 57.811                          |    | 61.058                    |    |
| <b>α (°)</b>                                                  | 72.63                    |    | 73.09                    |    | 102.00                    |    | 73.90                     |    | 73.28                           |    | 73.40                     |    |
| <b>β (°)</b>                                                  | 76.35                    |    | 75.74                    |    | 106.12                    |    | 74.82                     |    | 75.91                           |    | 73.30                     |    |
| <b>γ (°)</b>                                                  | 87.29                    |    | 87.22                    |    | 93.00                     |    | 87.90                     |    | 87.76                           |    | 89.73                     |    |
| <b>Wavelength (Å)</b>                                         | 1.00                     |    | 1.00                     |    | 0.99                      |    | 1.00                      |    | 1.00                            |    | 1.00                      |    |
| <b>Resolution range (Å)</b>                                   | 50.0-1.35<br>(1.40-1.35) |    | 50.0-1.55<br>(1.61-1.55) |    | 50.00-1.51<br>(1.56-1.51) |    | 50.00-1.70<br>(1.76-1.70) |    | 40.00-1.20<br>(1.24-1.20)       |    | 55.90-1.50<br>(1.55-1.50) |    |
| <b>Average I/σ</b>                                            | 14.56 (1.77)             |    | 23.70 (3.36)             |    | 14.38 (1.69)              |    | 20.44 (1.55)              |    | 18.17 (2.55)                    |    | 8.3 (2.40)                |    |
| <b>Total reflections</b>                                      | 512,955                  |    | 196,753                  |    | 356,236                   |    | 254,023                   |    | 539,004                         |    | 112,496                   |    |
| <b>Unique reflections</b>                                     | 56,961                   |    | 38,968                   |    | 42,559                    |    | 29,672                    |    | 82,503                          |    | 42,752                    |    |
| <b>Completeness (%)</b>                                       | 88.4 (56.5)              |    | 92.3 (68.1)              |    | 92.8 (75.0)               |    | 89.7 (48.9)               |    | 93.8 (89.9)                     |    | 96.1 (93.0)               |    |
| <b>R<sub>sym</sub> (%)</b>                                    | 4.2 (26.1)               |    | 4.2 (15.7)               |    | 3.2 (26.1)                |    | 5.2 (24.0)                |    | 6.6 (27.1)                      |    | 4.03 (29.76)              |    |
| Values in the parentheses refer to the last resolution shell. |                          |    |                          |    |                           |    |                           |    |                                 |    |                           |    |

$$R_{\text{sym}} = \frac{\sum ||I| - \langle I \rangle||}{\sum |I|}, \text{ where } I \text{ is the observed intensity and } \langle I \rangle \text{ is the average intensity obtained from multiple observations of symmetry-related reflections.}$$

#### 4.2.6 WT CRABP II Bound to RA

All-*trans*-RA was purchased from Sigma. The complex of CRABP II with all-*trans*-RA (CRABP II•RA) was prepared fresh and before each crystallization experiment, following almost the same procedure described by Bergfors et al.<sup>7</sup> but with some modifications as explained here. Since RA is light sensitive, all the manipulation of holo-CRABP II was performed in dark and under red light. We found that crystals of apo-WT cannot tolerate more than ~ 12-16% (v/v) ethanol. Therefore a saturated RA solution (2.7 mg/mL; ~  $1.0 \times 10^{-2}$  M) in ethanol was prepared. To keep the percentage of ethanol below 10%, while adding ~ 10 equivalents of the ligand, the protein solution was diluted almost 10 times with 100 mM TRIS-HCl pH 8.0 buffer to a concentration of ~ 1.5 mg/mL (or lower). 10 equivalents of RA, which was dissolved in ethanol and stored at – 20 °C, was added and the solution was re-concentrated to ~ 15 mg/mL, using Vivaspinn concentrators. For some crystallization trials the protein was concentrated to ~ 37 mg/mL. The dilute and concentrated solutions both could produce single and well diffracting crystals, but in different crystallization conditions. CRABP II•RA complex was stored in light-proof tubes (or alternatively tubes were covered with aluminum foil) and at 4 °C for crystallization. Crystallization was performed following the apo-CRABP II crystallization procedure, however since RA is light sensitive all the manipulations of CRABP II•RA were performed in dark and under red light. Boxes were wrapped by aluminum foil to be protected from light. The best crystals were obtained using a reservoir solution containing 30% (w/v) PEG 4000, 0.1 M sodium citrate/citric acid pH 5.4, and 0.2 ammonium acetate. This condition is different from that of the previously published structure (0.1 M sodium citrate pH 4.8, 20% (w/v) PEG 5000 and



10% dimethylsulphoxide, DMSO). Tetragonal crystals with sharp edges were obtained in less than a week and reached their maximum size in 10 days (Figure 4-10). Stabilization and freezing of the crystals were performed in dark and under red light following the same procedure as explained for apo-WT CRABP<sub>II</sub>. After the crystals were frozen in liquid nitrogen it was safe to expose them to light. Diffraction data were collected at the COM-CAT beamline (32-ID) of the APS synchrotron facility. 60 frames were collected at 2° oscillation at a wavelength of 1.00 Å and crystal to detector distance of 98.23 mm. Data collection statistics are reported in Table 4-4.

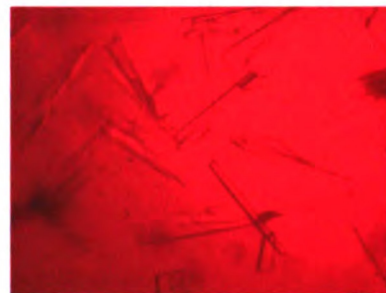


Figure 4-10 Crystals of CRABP<sub>II</sub>•RA. Since RA is light sensitive a red filter was used to take the picture.

#### 4.2.7 R132K:Y134F Bound to RA

The complex of the R132K:Y134F mutant with RA was formed as described for WT CRABP<sub>II</sub>•RA. R132K:Y134F (~ 25 mg/mL) bound to RA was crystallized in 0.1 M sodium citrate pH 4.8, 20% PEG 5000 and 10% DMSO, identical to one of the crystallization conditions of WT CRABP<sub>II</sub>•RA (Figure 4-11). The complex was crystallized at room temperature and under white light. This mutant showed that although RA is light sensitive, it was stable enough to be exposed to light for short periods of time, while the crystallization box was set up. However



Figure 4-11 A crystal of R132K:Y134F bound to RA

the tubes containing the complex were covered with aluminum foil and kept on ice during the crystallization experiments. The crystallization boxes were wrapped with aluminum foil and kept in dark cabinets. Data was collected at the COM-CAT beamline (32-ID) at the APS synchrotron facility. 120 frames were collected at 1° oscillation at a wavelength of 1.00 Å. Data collection statistics are reported in Table 4-4.

#### **4.2.8 R132K:Y134F Bound to Retinal**

Retinal is much more light-sensitive than RA and should not be exposed to white light at any time. Although 11-*cis*-retinal is the natural substrate of rhodopsin we used a more stable isomer of this molecule, all-*trans*-retinal, for our initial experiments toward making a mimic of rhodopsin. Working with 11-*cis*-retinal requires very strict dark conditions, however the all-*trans* isomer is stable as long as it is not directly exposed to light.

Due to instability of retinal in the protein solution, our initial attempts in crystallizing this complex were unsuccessful. We tried co-crystallization experiments in which different equivalents of retinal in the protein solution, both at room temperature and 4 °C, were added. Although we could produce well-diffracting crystals of this mutant there was not any retinal molecule present inside the pocket.

In order to make sure that retinal is bound inside the pocket of the protein before it is used for crystallization, the  $\lambda_{\text{max}}$  of the protein mixed with retinal was measured. The spectra verified the presence of retinal by showing ~ 27 nm red shift and a peak at ~ 404 nm (Figure 4-12-A). Since 4 equivalents of retinal was added the  $\lambda_{\text{max}}$  of bound retinal at 404 nm was covered with the peak of excess and free retinal at 380 nm. However the



peak at 404 nm is observed when a more accurate analysis of the spectrum was performed by calculating the second derivative of the data. Also we tested whether retinal was bound in the crystals before we exposed them to X-rays. 14-16 co-crystals were washed with mother liquor (2% higher concentration of precipitating agent was used to stabilize the crystals) and dissolved in 600  $\mu$ L of 10 mM TRIS-HCl buffer. The maximum absorption of this solution was measured. The spectrum did not show any peak corresponding to retinal (380 nm) and had a strong base line, which must be due to the crystallization reagents and glycerol from cryoprotectant. This spectrum overlaid on the spectrum of the double mutant mixed with 4 equivalents of retinal is shown in Figure 4-12-A.

In another experiment the drop containing the co-crystals of the double mutant with retinal was tested for the presence of retinal. No peak was observed at 380 nm or higher indicating that there is no retinal present in the drop. We also tested the drops that did not produce any crystals for comparison. These drops did not show any evidence of retinal in the drop either. The overlaid spectra are shown in Figure 4-12-B. As apparent from the spectra, although both drops originally had the same amount of protein, the one that produced crystals has an expected lower concentration of the protein. This is because most of the protein was crystallized. Both drops do not show any peak for retinal at 380 nm or higher. These experiments verified that retinal was unstable and somehow degraded during the crystallization period (4-5 days).

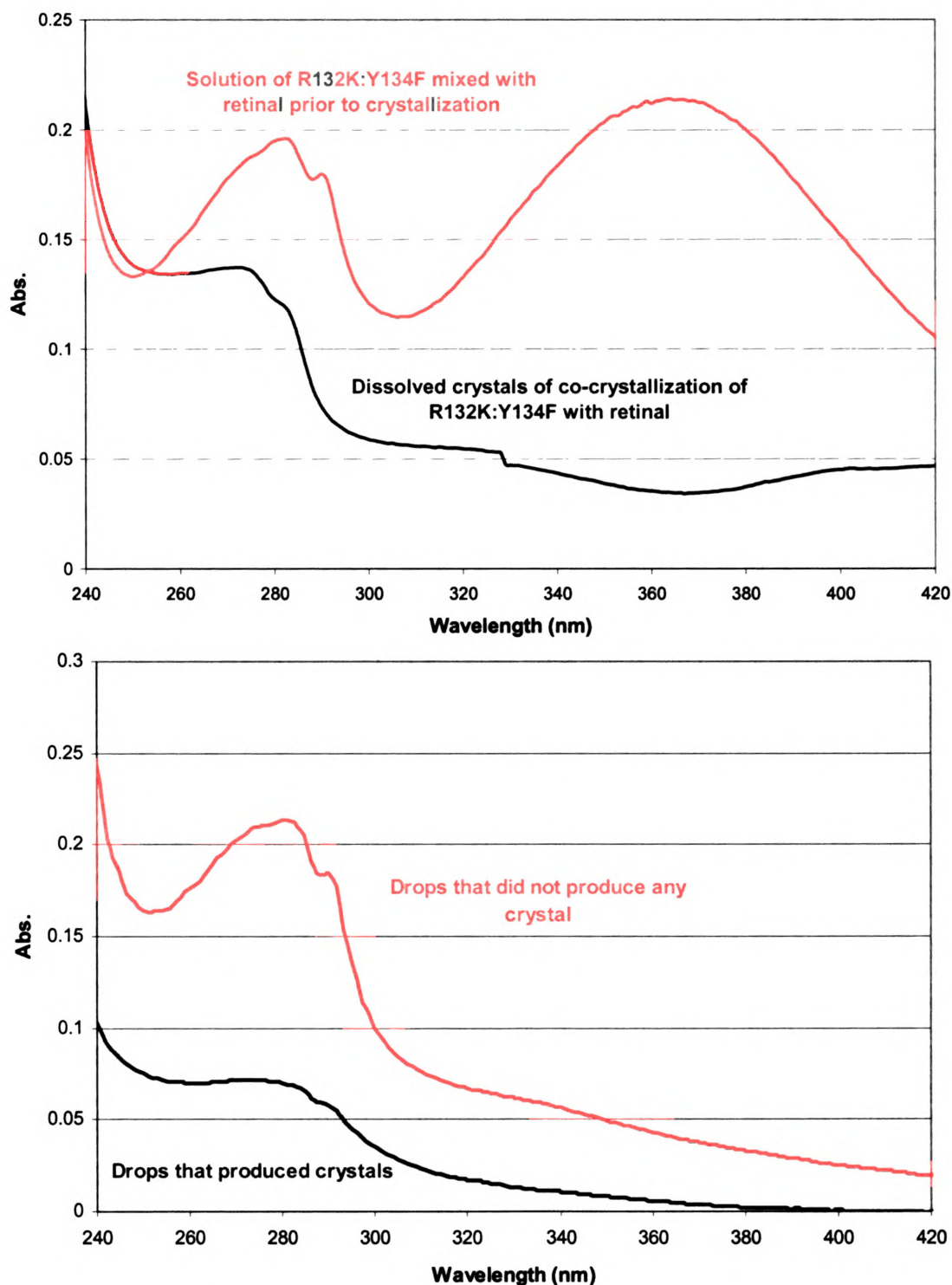


Figure 4-12 (A) Overlaid spectrum of the solution containing dissolved co-crystals of R132K:Y134 bound to retinal (blue) and the solution of the mutant mixed with 4 equivalents of retinal (red); (B) Overlaid spectra of drops that produced (blue) and did not produce (red) crystals.

Therefore we decided to try soaking the apo-crystals in a solution of retinal. The crystals were transferred to new drops containing reservoir solution plus retinal. To stabilize the crystals the amount of precipitating agent was increased by 2% in the new drops. Retinal is dissolved and stored in pure ethanol. Soaking of the crystals in different amounts of retinal showed that they could not tolerate more than 12-16% (v/v) of ethanol. Therefore a saturated solution of retinal ( $\sim 5.5 \times 10^{-2}$  M) was used to maximize the amount of retinal that could be added, while keeping the amount of ethanol up to 12% in the drop (up to 4 equivalents of retinal). Soaking experiments were performed both at room temperature and at 4 °C. A range of retinal concentrations and soak times (up to 24 hours, with 4 hours intervals) were tried however there was no electron density map for retinal evident from any of the soaking experiments.

#### Remedy for Crystallization of CRABPII Bound to Retinal

We next tested the stability of retinal in the protein solution both at room temperature and at 4 °C (experiments performed in dark). Interestingly we found that although retinal was stable for more than a week at 4 °C it was only stable for 12 h at room temperature. This finding explained why retinal co-crystallization experiments with retinal had so far failed to produce a complex. Therefore co-crystallization experiments were performed in the dark and at 4 °C. 4 equivalents of retinal were added to the protein solution ( $\sim 28$  mg/mL). This way,  $\sim 12\%$  (v/v) of ethanol was added. The crystallization boxes were wrapped with aluminum foil and kept in dark cabinets. The crystals appeared within 3 days, against a reservoir containing 30% (w/v) PEG 8000 and 0.2 M ammonium sulfate, and reached a reasonable size within 6 days (Figure 4-13). The

crystals were cryo-protected solution and flash frozen using liquid nitrogen in the dark and at 4 °C. After crystals were frozen it was safe to expose them to light. Data was collected at the COM-CAT beamline (32-ID) at the APS synchrotron facility. 80 frames were collected at 2° oscillation at a wavelength of 1.00 Å. The crystal to detector distance was set to 98.17 mm. Data collection statistics are reported in Table 4-4.

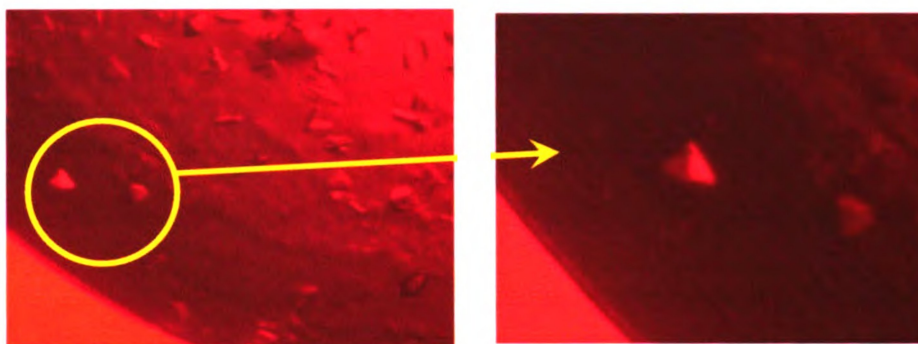


Figure 4-13 Crystals of R132K:Y134F bound to retinal.

#### 4.2.9 R132K:Y134F:R111L:T54V:L121E Bound to RA

The complex of the R132K:Y134F:R111L:T54V:L121E mutant with RA was prepared as described for CRABP II•RA. The penta mutant (~ 20 mg/mL) bound to RA was crystallized in reagent # 26 of Crystal Screen II (Hampton Research Inc.), 0.1 M MES pH 6.5, 30% (w/v) PEG 5000, 0.2 M ammonium sulfate (Figure 4-14). The complex was crystallized at room temperature. The crystallization boxes were wrapped with aluminum foil and kept in dark cabinets. Data was collected at the COM-CAT beamline (32-ID) at the APS synchrotron facility. 200 frames were collected at 1° oscillation at a wavelength of 1.00 Å. The crystal to detector distance was set to 108.00 mm. Data collection statistics are reported in Table 4-4.



Figure 4-14 Crystals of  
R132K:Y134F:R111L:T54V:L121E bound to RA.

#### 4.2.10 R132K:R111L:L121E Bound to Retinal

The complex of R132K:R111L:L121E with retinal was prepared by adding 4 equivalents of a saturated solution of retinal ( $\sim 5.5 \times 10^{-2}$  M) to a  $\sim 27$  mg/mL solution of the mutant. Since retinal is light and temperature sensitive, all the crystallization experiments were performed in the dark and at 4 °C. The crystals appeared within 3 days against a reservoir containing 18% (w/v) PEG 6000 and 0.1 M sodium acetate pH 5.8, and reached a maximum size in 6 days (Figure 4-15). The crystals were cryo-protected and flash frozen in the dark and at 4 °C before they were exposed to light for X-ray data collection. Data was collected at the COM-CAT beamline (32-ID) at the APS synchrotron facility. 180 frames were collected at 2° oscillation at a wavelength of 1.00 Å. The crystal to detector distance was set to 80.00 mm. Data collection statistics are reported in Table 4-4.



Figure 4-15 A crystal of  
R132K:R111L:L121E  
bound to retinal

Table 4-4 X-ray data collection statistics for the holo-structures

|                                                               | CRABPII-RA                                    | R132K:Y134F.RA                                | R132K:Y134F.Retinal                           | R132K:Y134F:R111L:<br>T54V:L121E.RA           | R132K:R111L:L121E.Retinal                     |
|---------------------------------------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| <b>Space group</b>                                            | P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> | P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> | P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> | P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> | P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> |
| <b>Unit cell dimensions</b>                                   |                                               |                                               |                                               |                                               |                                               |
| <b>a (Å)</b>                                                  | 44.15                                         | 44.702                                        | 45.493                                        | 45.787                                        | 45.459                                        |
| <b>b (Å)</b>                                                  | 46.59                                         | 46.251                                        | 46.325                                        | 44.860                                        | 45.514                                        |
| <b>c (Å)</b>                                                  | 77.55                                         | 77.553                                        | 73.597                                        | 77.865                                        | 77.747                                        |
| <b>α (°)</b>                                                  | 90.00                                         | 90.00                                         | 90.00                                         | 90.00                                         | 90.00                                         |
| <b>β (°)</b>                                                  | 90.00                                         | 90.00                                         | 90.00                                         | 90.00                                         | 90.00                                         |
| <b>γ (°)</b>                                                  | 90.00                                         | 90.00                                         | 90.00                                         | 90.00                                         | 90.00                                         |
| <b>Wavelength (Å)</b>                                         | 1.00                                          | 1.00                                          | 1.00                                          | 1.00                                          | 1.00                                          |
| <b>Resolution range (Å)</b>                                   | 39.81-1.48<br>(1.53-1.48)                     | 40.00-1.70<br>(1.76-1.70)                     | 50.00-1.69<br>(1.75-1.69)                     | 50.00-1.60<br>(1.66-1.60)                     | 40.00-1.18<br>(1.22-1.18)                     |
| <b>Average I/σ</b>                                            | 28.19 (5.08)                                  | 17.66 (1.91)                                  | 17.03 (1.74)                                  | 27.28 (9.04)                                  | 25.60 (1.68)                                  |
| <b>Total reflections</b>                                      | 50,640                                        | 205,027                                       | 178,818                                       | 394,601                                       | 768,796                                       |
| <b>Unique reflections</b>                                     | 27,332                                        | 18,359                                        | 18,002                                        | 21,863                                        | 54,241                                        |
| <b>Completeness (%)</b>                                       | 95.5 (96.4)                                   | 95.6 (60.5)                                   | 99.6 (99.9)                                   | 99.8 (100.0)                                  | 96.7 (74.6)                                   |
| <b>R<sub>sym</sub> (%)</b>                                    | 4.4 (13.9)                                    | 4.6 (36.1)                                    | 7.6 (74.1)                                    | 5.0 (15.7)                                    | 7.2 (33.2)                                    |
| Values in the parentheses refer to the last resolution shell. |                                               |                                               |                                               |                                               |                                               |

$$R_{\text{sym}} = \frac{\sum ||I| - \langle I \rangle||}{\sum |I|}$$
 , where  $I$  is the observed intensity and  $\langle I \rangle$  is the average intensity obtained from multiple observations of symmetry-related reflections.

### 4.3 Structure Solution and Refinement

All the structure determination and refinement calculations were performed using the CCP4 suite (Collaborative Computational Project, Number 4, 1994).<sup>9</sup> The structure of apo-WT CRABP II was determined using rigid body refinement with the structure of the apo-R111M mutant of CRABP II (PDB ID: 1XCA)<sup>10</sup> used as the starting model. The rest of the apo-structures were determined using the refined, high resolution structure of apo-WT CRABP II (Xtal1). The RA- and retinal-bound structures were determined using the previously published structure of CRABP II•RA (PDB ID: 1CBS)<sup>8</sup>, with RA removed, as the starting model.

In cases where the crystallographic axes of the model and our data were not identical, molecular replacement (MR) was used to determine the phases of the structures.

The models were visualized and manually rebuilt using the program TURBO-FRODO in each structure.<sup>11</sup> Version 5.2.0005 of REFMAC, the maximum-likelihood refinement program, was used to refine the structure against 90% of the data,<sup>12</sup> while 5-10% of the data was chosen randomly for cross-validation,  $R_{\text{free}}$ .<sup>13</sup> The refinement statistics and analysis of Ramachandran plots of all the apo- and holo-structures are shown in Tables 4-5 and 4-6, respectively, and are also discussed in this chapter. Values of  $R_{\text{work}}$  and  $R_{\text{free}}$  reported in the tables are measures of the agreement between the values of the observed structure factors and the calculated ones:

$$R_{\text{work}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{cal}}||}{\sum |F_{\text{obs}}|}$$

$$R_{\text{free}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{cal}}||}{\sum |F_{\text{obs}}|}$$

where, reflections belong to a test set of 5-10% randomly selected data.

All the apo-structures consistently had a P1 space group with two molecules in the asymmetric unit (asu) and similar cell constants, while the holo-structures had a P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> space group with one molecule in their asu and similar cell constants.

#### 4.3.1 Apo-WT

As mentioned above three complete data sets were collected on three different crystals of apo-WT CRABP<sub>II</sub>. However, after determining the structure from each of the three data sets, we realized that Xtal1 is different from Xtal2 and Xtal3. Since the structures resulting from Xtal2 and Xtal3 were virtually identical, between these two data, only the higher resolution data set (from Xtal2) was completely refined. The structures of both Xtal1 and Xtal2 were determined using rigid body refinement in the CCP4 suite.<sup>9</sup> The coordinates of the apo-R111M (both of the molecules) after removing the water molecules were used as the original model. The difference between Xtal1 and Xtal2 was in the loop connecting the α2 to the βB in Mol B (residues 35-41). To generate an unbiased Fo-Fc map for this region, residues 35-41 were deleted in Mol B of R111M. The Fo-Fc maps were calculated based on this unbiased model, which clearly showed that the loop connecting the α2 to the βB in Mol B is significantly different in the two structures of Xtal1 and Xtal2. In order to generate an unambiguous Fo-Fc electron density map of the deleted region, residues 35 to 41 of Mol B were built in each structure



only after most of the water molecules and residues were built into the electron density and refined.

To confirm that we see two different conformations of the loop region in the Mol B's of Xtal1 and Xtal2, the occupancies of residues in the loop region and neighboring residues before and after the loop (residues 35 to 41) were set to zero, 15 cycles of REFMAC refinement were run and an Fo-Fc electron density map was calculated using the CCP4 suite.<sup>9</sup> This is a REFMAC approach to calculate an unbiased omit map. The resulting Fo-Fc omit map for Xtal1 (Figure 4-16) and Xtal2 (Figure 4-17) clearly confirmed that region 36 to 40 in Mol B has two distinct conformations in these two crystals. The two structures are superimposed for comparison. Further, double conformation of the loop was verified by calculating the simulated annealing-omit (sa-omit) map of this region using the CNS program.<sup>14</sup> The sa-omit map provided us with additional evidence of having two conformations for this loop in the two structures. Although this map was not as definitive as the REFMAC map, it still showed the loop region being different in the two structures. Interestingly the omit maps show some additional electron density map for the other conformation of the loop in each structure, but the maps were not definitive enough to build an alternative conformation of the loop in each structure.

Ramachandran analyses of the structures were calculated by the program Procheck,<sup>15</sup> and shows that in Xtal1 and Xtal2 94.3% and 93.5% of the residues are in the most favored region, respectively. In both structures the only residue in the disallowed region is Asp126 A, and the only residue in the generously allowed region is Asp126 B. These residues are located at the sharp turn of the  $\beta$ I- $\beta$ J hairpin loop and have higher than

average B-factors and poor 2Fo-Fc electron density maps. These two residues are reported to be in the same Ramachandran plot regions in the R111M structure and other reinoid-binding proteins.<sup>10</sup> This can be due to the flexibility of the  $\beta$ I- $\beta$ J hairpin loop that makes this residue have a different conformation from what is expected based on commonly observed ones. Ramachandran plots of Xtal1 and Xtal2 are represented in Figures 4-18 and 4-19, respectively. Based on an analysis of 118 structures of resolution of at least 2.0 Å and R-factor no greater than 20%, a good quality model would be expected to have over 90% of its residues in the most favored regions.

In the following Ramachandran plots the red areas show the most favored, bright-yellow areas show the additional allowed, dull-yellow areas show the generously allowed and white areas show the disallowed regions of the plot. The other areas are the generously allowed regions. Refinement statistics and the pdb codes are listed in Table 4-5.

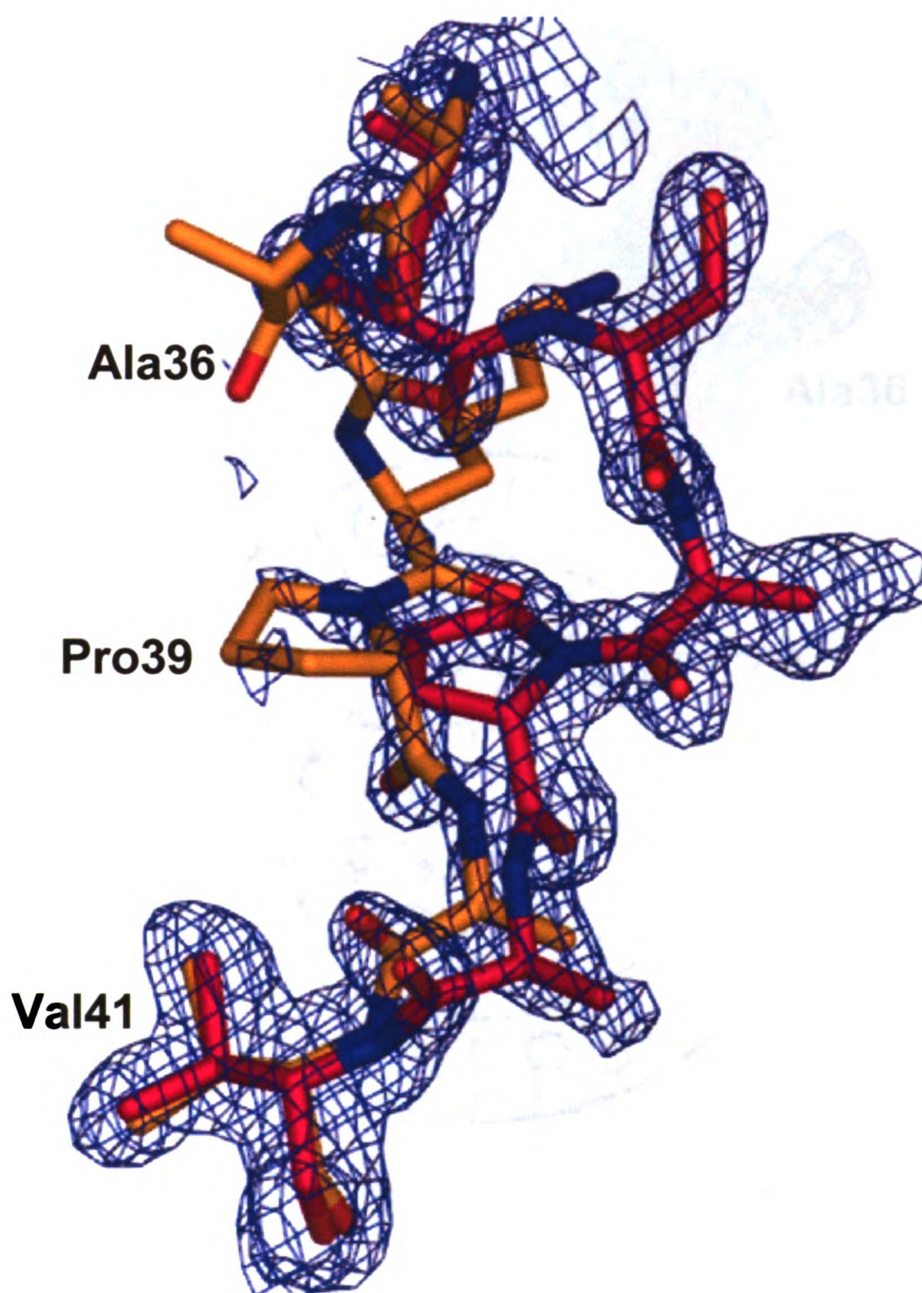


Figure 4-16 Omit electron density map of residues 35 to 41 of Mol B in Xtal1 (hot pink), contoured at 2.2  $\sigma$ , with Xtal2 (orange) superimposed.

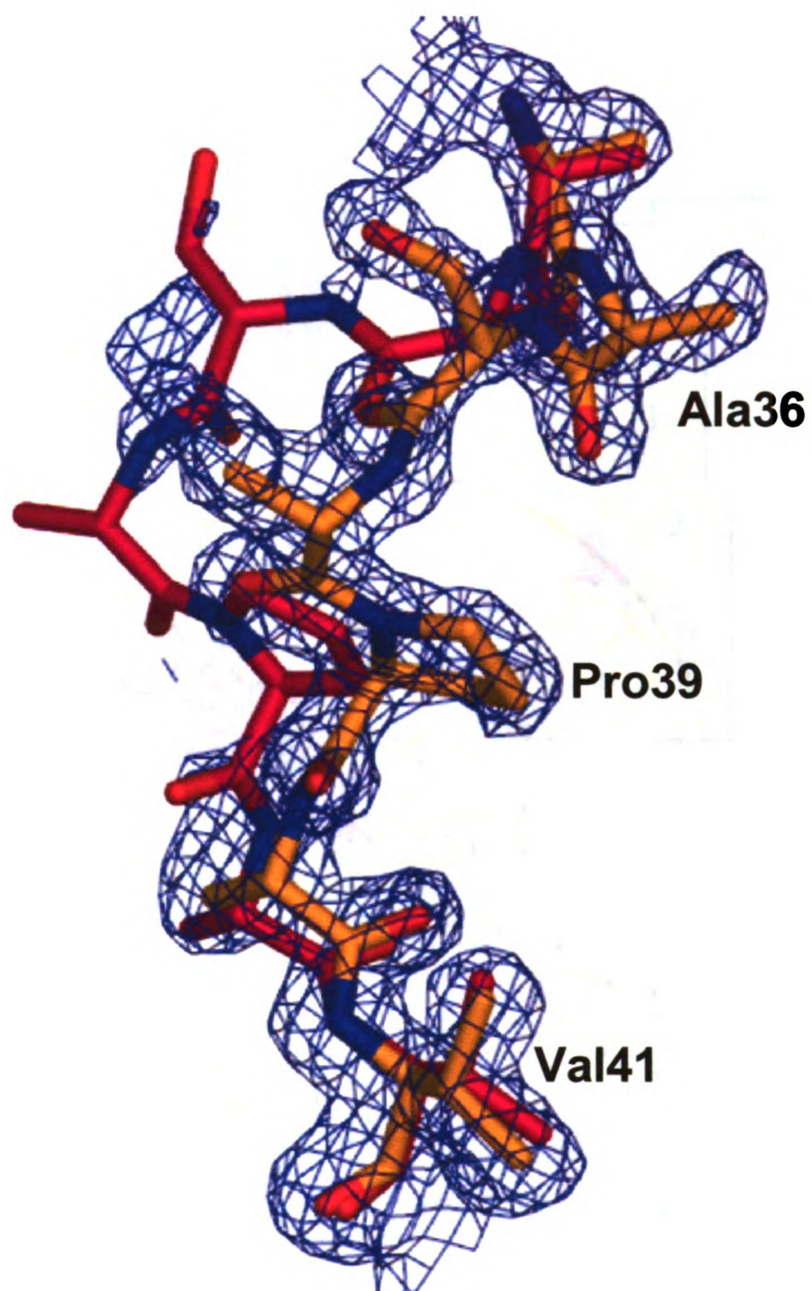


Figure 4-17 Omit electron density map of residues 35 to 41 of Mol B in Xtal2 (orange), contoured at 2.4  $\sigma$ , with Xtal1 (hot pink) superimposed.

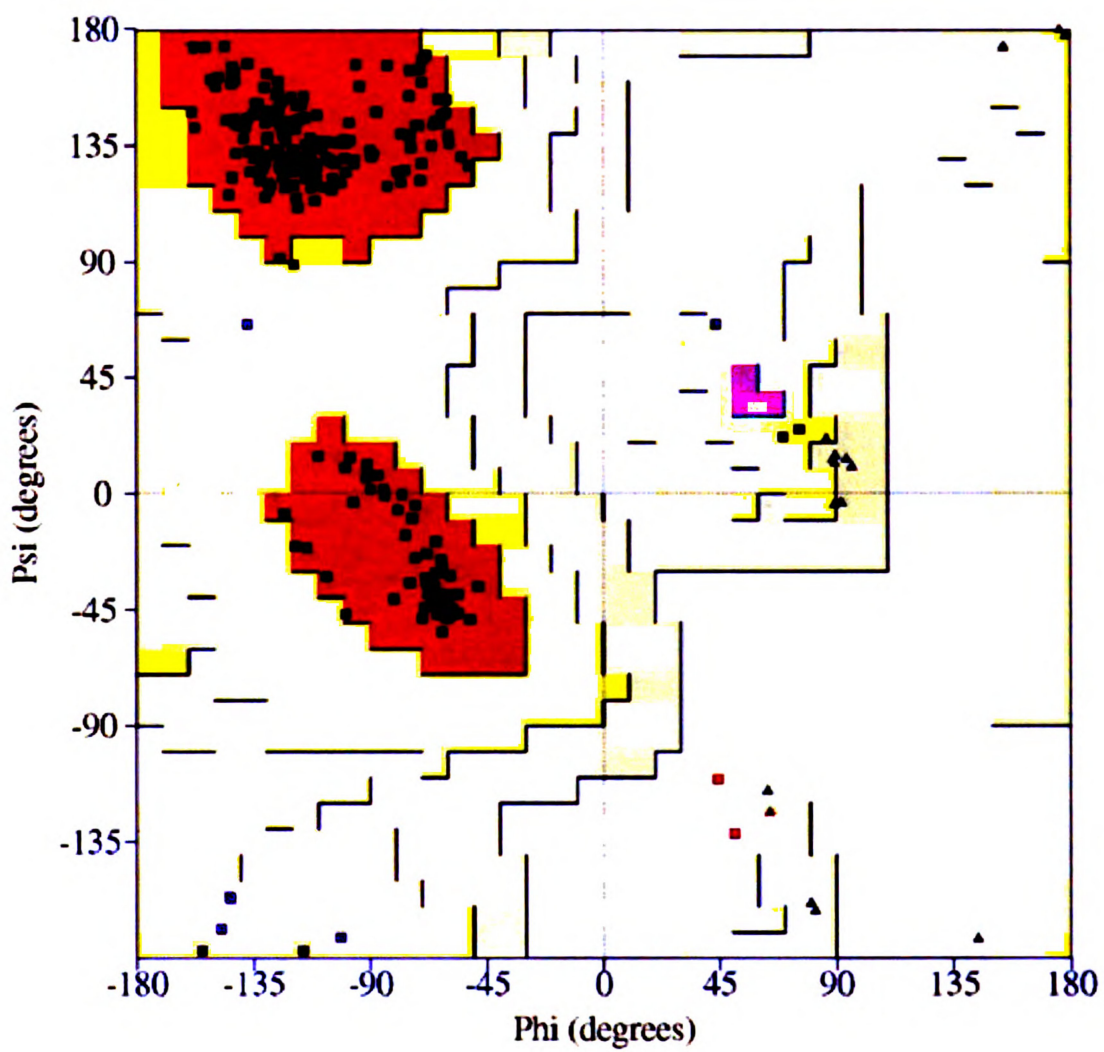


Figure 4-18 Ramachandran plot of the apo-WT CRABP II (Xtal1)

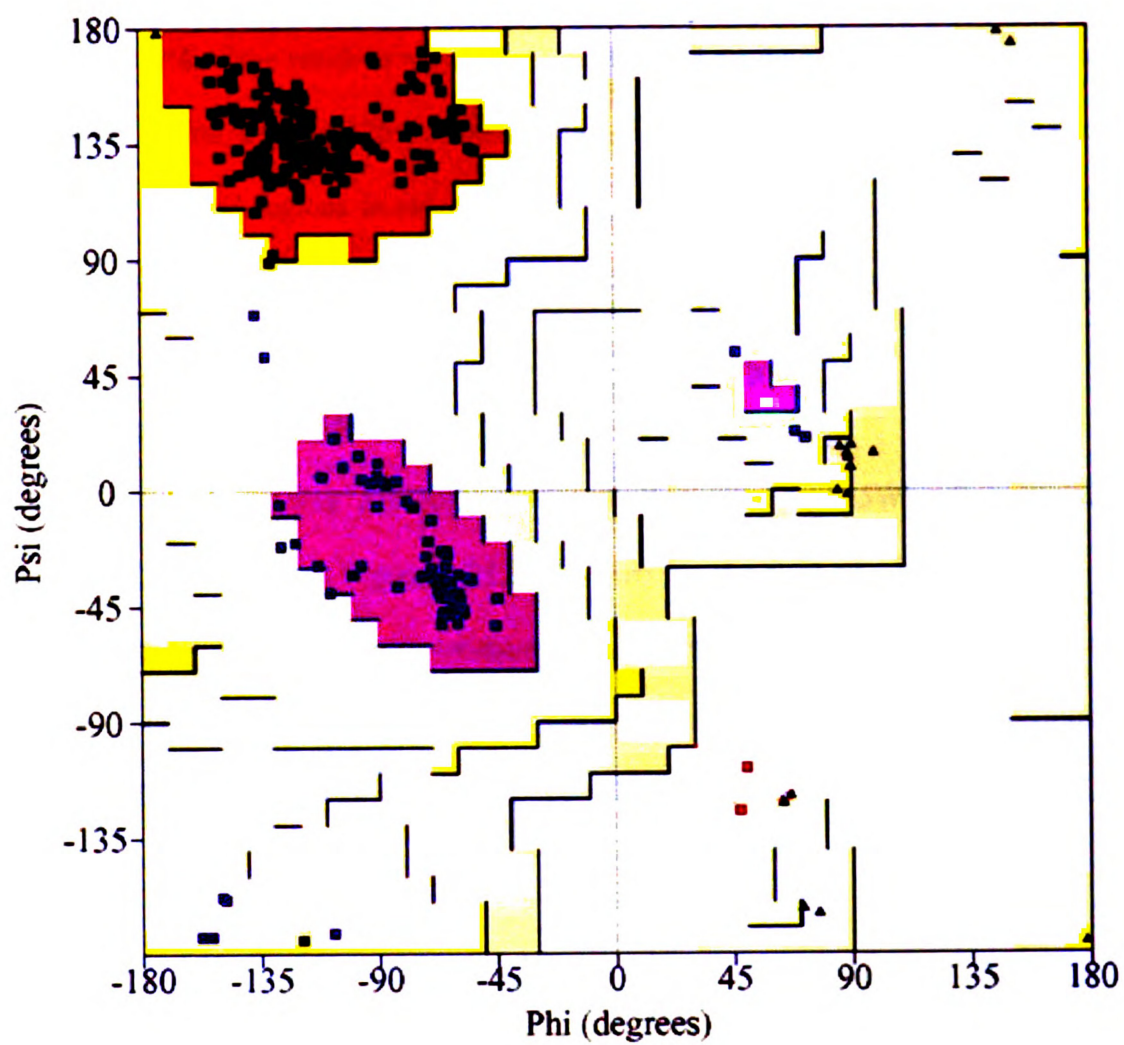


Figure 4-19 Ramachandran plot of the apo-WT CRABP II (Xtal2)



### **4.3.2 Apo-F15W**

The structures of apo-F15W were determined using the coordinates of Xtal1, as the starting model, and rigid body refinement in the CCP4 suite.<sup>9</sup> Ramachandran analysis of the structure, Figure 4-20, calculated by program Procheck,<sup>15</sup> shows that in this structure 93.8% of the residues are in the most favored region. The only residues in the disallowed region are Asp126 A and Asp126 B, which are usually in disallowed or generously allowed regions in retinoid-binding proteins (see Ramachandran analysis of Xtal1 and Xtal2). Refinement statistics and the pdb code are listed in Table 4-5.

### **4.3.3 Apo-R132K:Y134F**

The structure was determined as explained for apo-F15W. Ramachandran analysis of the structure, Figure 4-21, calculated by program procheck,<sup>15</sup> shows that similar to apo-WT (Xtal2), 93.5% of the residues are in the most favored region and only Asp126 A is in disallowed region, while Asp126 B resides in a generously allowed region. Refinement statistics and the pdb code of this structure are listed in Table 4-5.

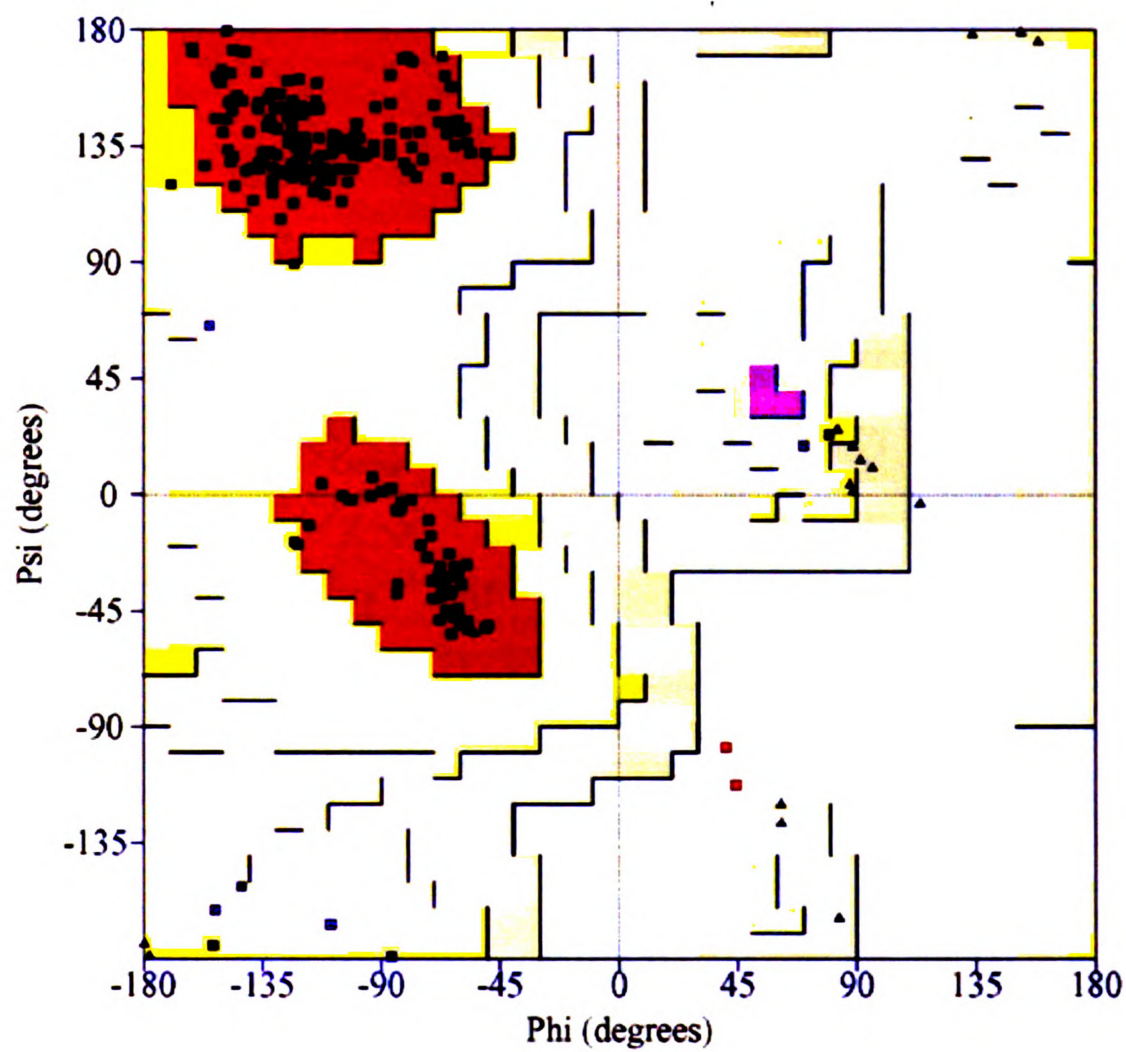


Figure 4-20 Ramachandran plot of the apo-F15W CRABPII



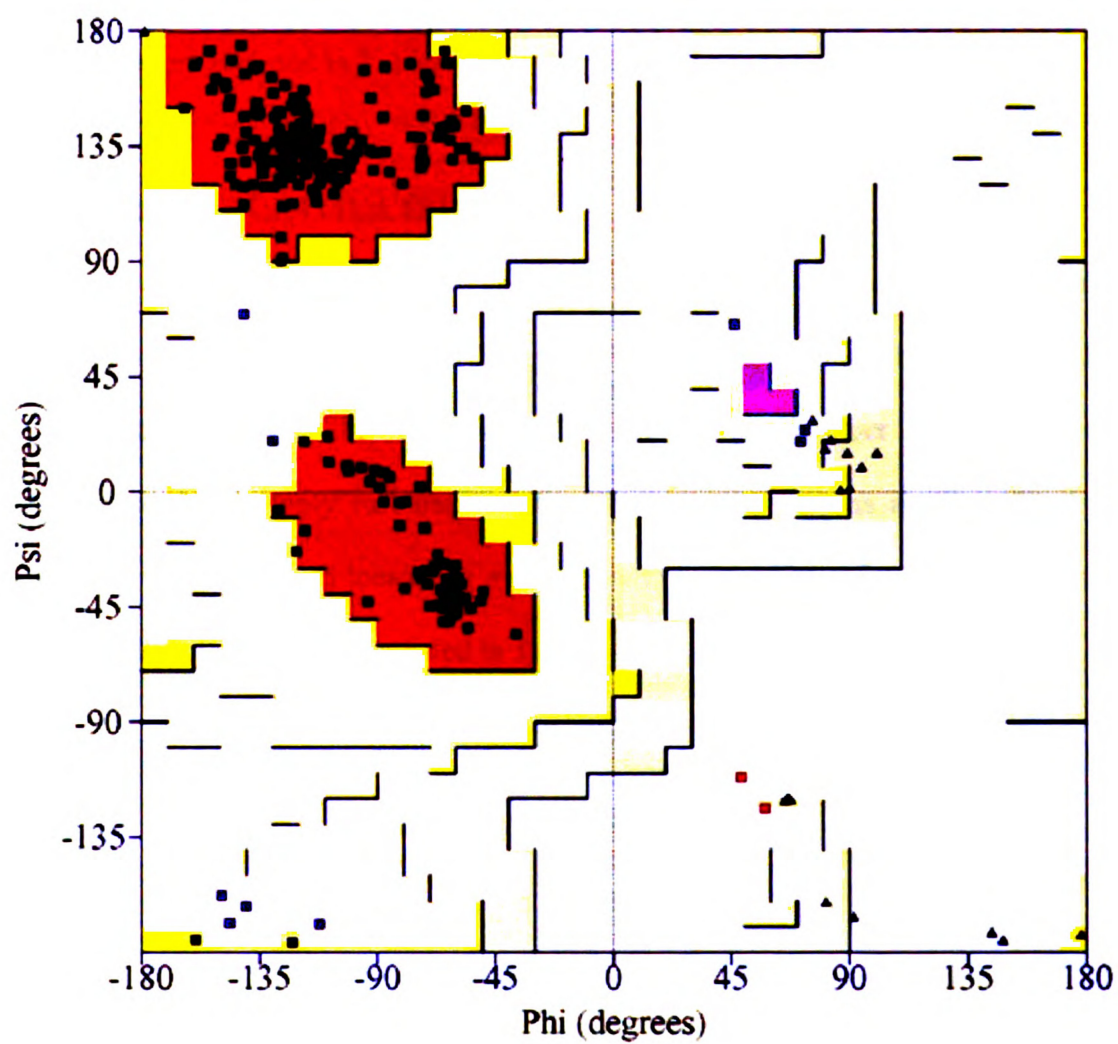


Figure 4-21 Ramachandran plot of the apo-R132K:Y134F CRABPII

#### **4.3.4 Apo-R132K:Y134F:R111L:L121E:T54V**

The structure was determined as explained for apo-F15W. The Ramachandran plot shows that 92.7% of the residues reside in the most favored region of the plot and only Asp126 A is located in a disallowed region (Figure 4-22). Refinement statistics and the pdb code are reported in Table 4-5.

#### **4.3.5 Apo-R132K:R111L:L121E**

The structure was determined as described for the apo-F15W structure. Analysis of the Ramachandran plot, Figure 4-23, shows that similar to apo-F15W, 94.0% of the residues are in the most favored regions and similar to the other structures that we discussed so far the only residues that have the most deviations are Asp126 A and Asp126 B, which are both located in a disallowed region. The refinement statistics and the pdb code of this structure is listed in Table 4-5.

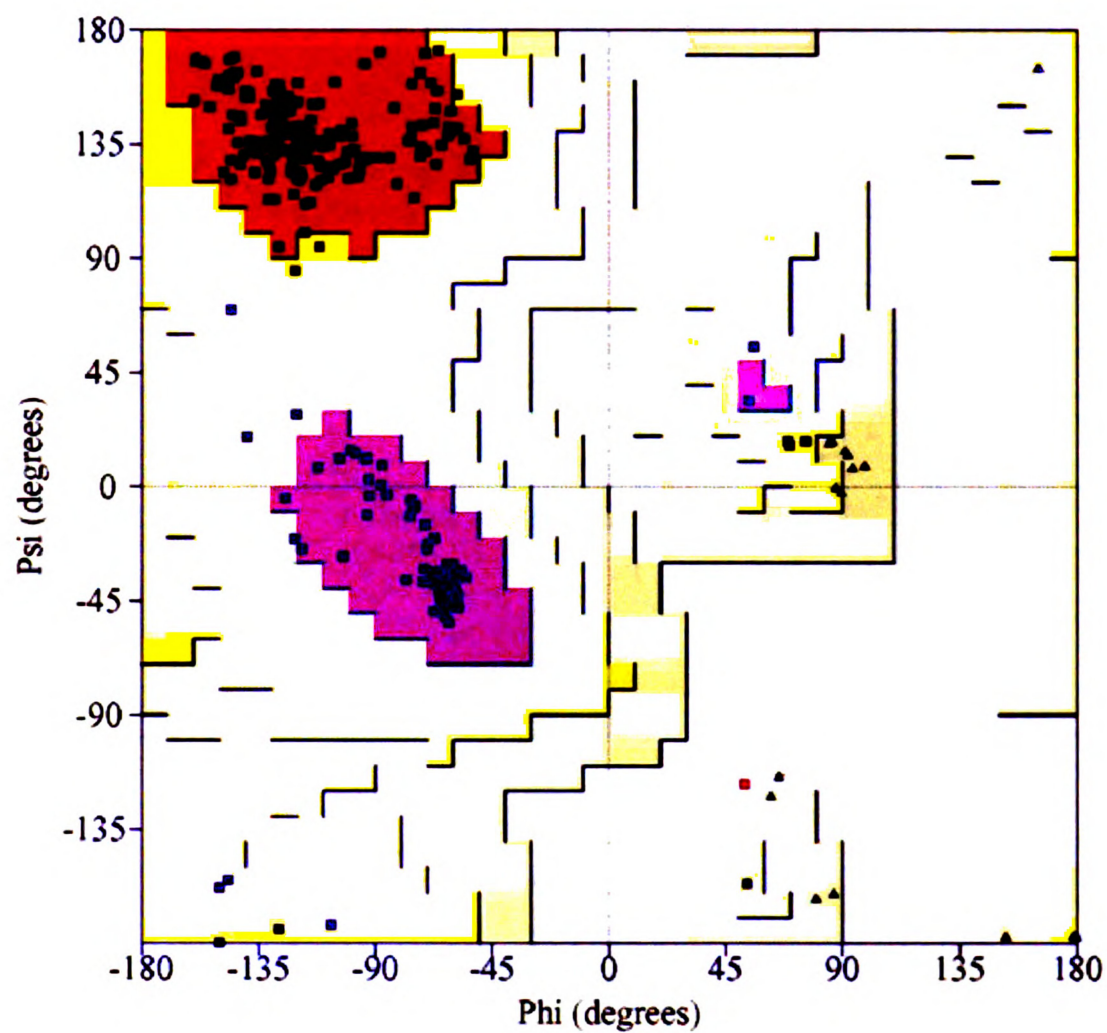


Figure 4-22 Ramachandran plot of the apo-R132K:Y134F:R111L:T54V:L121E CRABPII

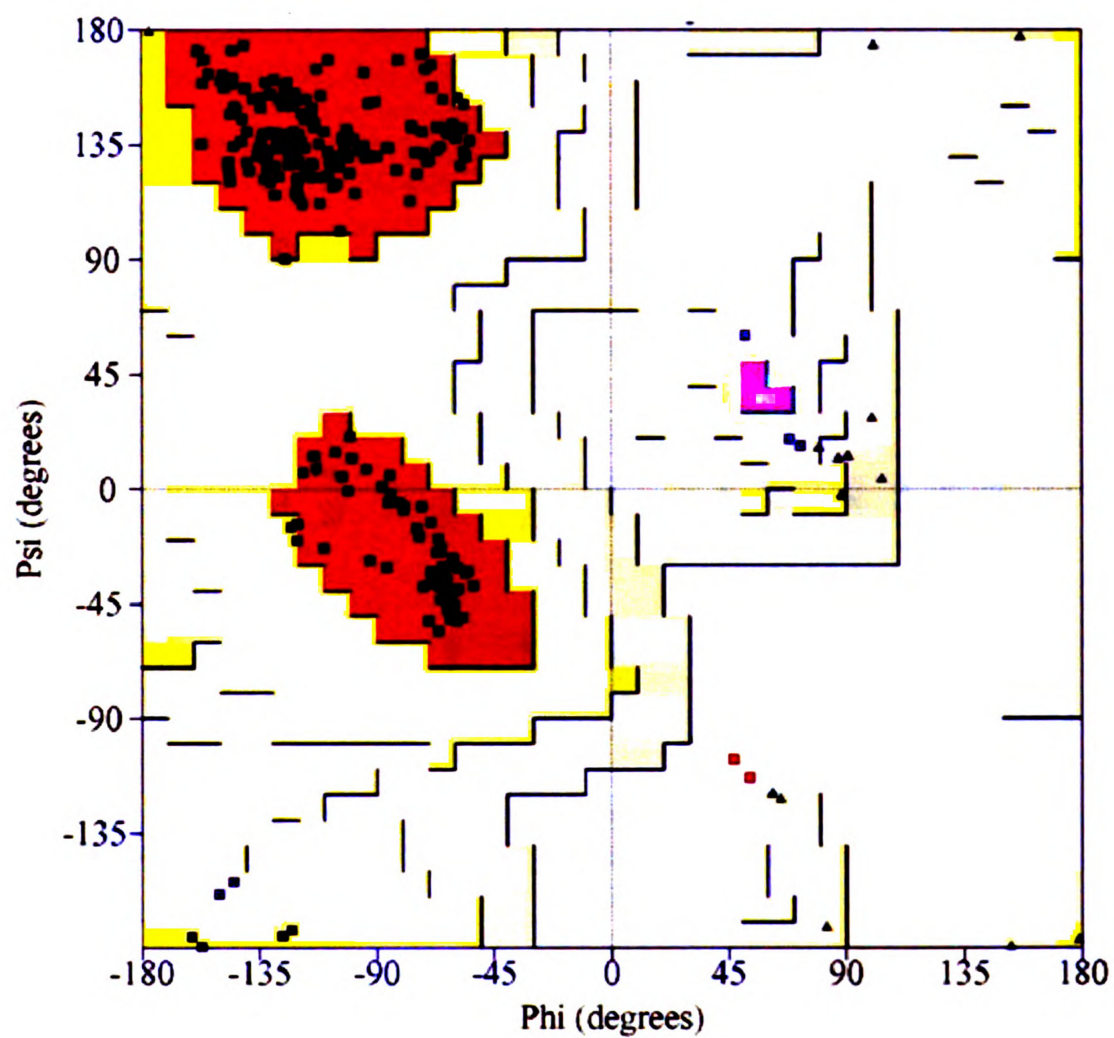


Figure 4-23 Ramachandran plot of the apo-R132K:R111L:L121E CRABPII

Table 4-5 Refinement statistics for the apo-structures of CRABPII

|                                                     | Apo-<br>CRABPII<br>Xtal1 | Apo-<br>CRABPII<br>Xtal2 | Apo-F15W | Apo-<br>R132K:Y134F | Apo-<br>R132K:Y134F:R111L:<br>T54V:L121E | Apo-<br>R132K:R111L:L121E |
|-----------------------------------------------------|--------------------------|--------------------------|----------|---------------------|------------------------------------------|---------------------------|
| <b>Average B-Factor (<math>\text{\AA}^2</math>)</b> | 16.78                    | 18.82                    | 17.96    | 23.20               | 18.13                                    | 18.82                     |
| <b>Resolution (<math>\text{\AA}</math>)</b>         | 1.35                     | 1.55                     | 1.51     | 1.70                | 1.20                                     | 1.50                      |
| <b><math>R_{\text{work}}</math> (%)</b>             | 14.35                    | 13.94                    | 17.15    | 15.35               | 13.80                                    | 18.22                     |
| <b><math>R_{\text{free}}</math> (%)</b>             | 20.05                    | 19.67                    | 24.25    | 22.18               | 17.62                                    | 23.03                     |
| <b>Number of water molecules</b>                    | 386                      | 299                      | 455      | 362                 | 399                                      | 364                       |
| <b>Total reflections used</b>                       | 50,364                   | 35,954                   | 39,495   | 25,713              | 75,296                                   | 41,632                    |
| <b>Total reflections in work set</b>                | 45,282                   | 32,359                   | 35,549   | 23,415              | 67,736                                   | 37,465                    |
| <b>Total reflections in test set</b>                | 5,082                    | 3,595                    | 3,946    | 2,298               | 7,560                                    | 4,167                     |
| <b>rmsd from ideality</b>                           |                          |                          |          |                     |                                          |                           |
| <b>Bond lengths (<math>\text{\AA}</math>)</b>       | 0.020                    | 0.019                    | 0.016    | 0.020               | 0.018                                    | 0.009                     |
| <b>Bond Angles (<math>^\circ</math>)</b>            | 1.888                    | 1.750                    | 1.681    | 1.779               | 1.687                                    | 1.270                     |
| <b><i>Ramachandran plot</i></b>                     |                          |                          |          |                     |                                          |                           |
| <b>Most favored (%)</b>                             | 94.3                     | 93.5                     | 93.8     | 93.5                | 92.7                                     | 94.0                      |
| <b>Allowed (%)</b>                                  | 4.9                      | 5.7                      | 5.3      | 5.6                 | 6.9                                      | 5.2                       |
| <b>Generously allowed (%)</b>                       | 0.4                      | 0.4                      | 0.0      | 0.4                 | 0.0                                      | 0.0                       |
| <b>Disallowed (%)</b>                               | 0.4                      | 0.4                      | 0.9      | 0.4                 | 0.4                                      | 0.8                       |
| <b>PDB code</b>                                     | 2FS6                     | 2FS7                     | 2FRS     | 2FRU                | 2FS0                                     | 2FRR                      |

#### **4.3.6 WT CRABPII Bound to RA**

The structure was determined using rigid body refinement and the previously published structure of CRABPII•RA (PDB ID: 1CBS), as the phasing model, after removing the water and RA molecules. Ramachandran analysis of the data, Figure 4-24, shows that 95.2% of the residues are in the most favored region of the plot and only Asp126 is in the generously allowed region (there is only one molecule in the a.s.u.). Unlike apo-structures there is no residue in the disallowed region. Refinement statistics and the pdb code of the structure are reported in Table 4-6.

#### **4.3.7 R132K:Y134F Bound to RA**

The structure was determined as explained for WT-CRABPII•RA. The Ramachandran plot analysis shows that 92.7% of the residues are located in the most favored region of the plot and Asp126 is the only residue in the disallowed region (Figure 4-25). The refinement statistics and pdb code of the structure are reported in Table 4-6.

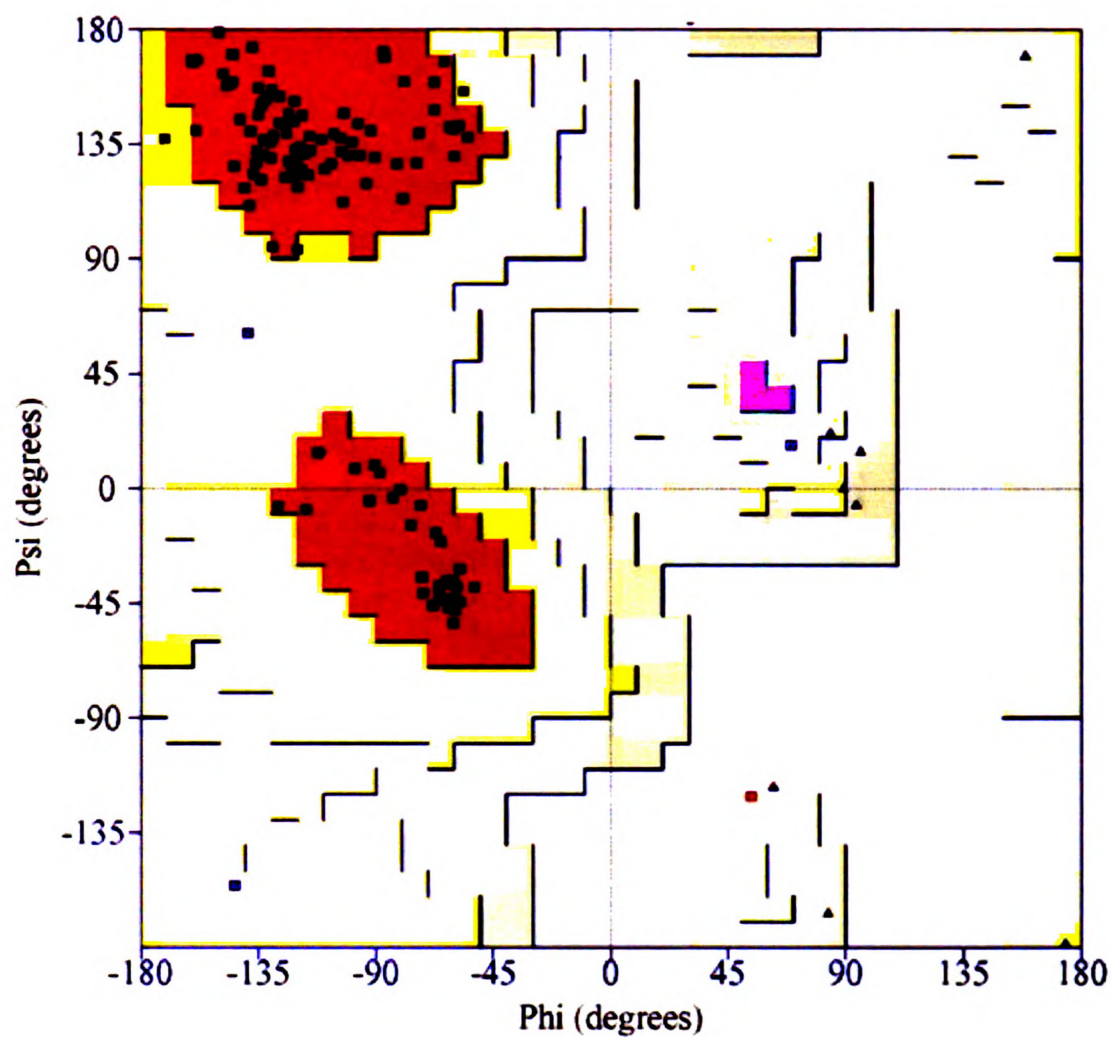


Figure 4-24 Ramachandran plot of the WT CRABP II•RA

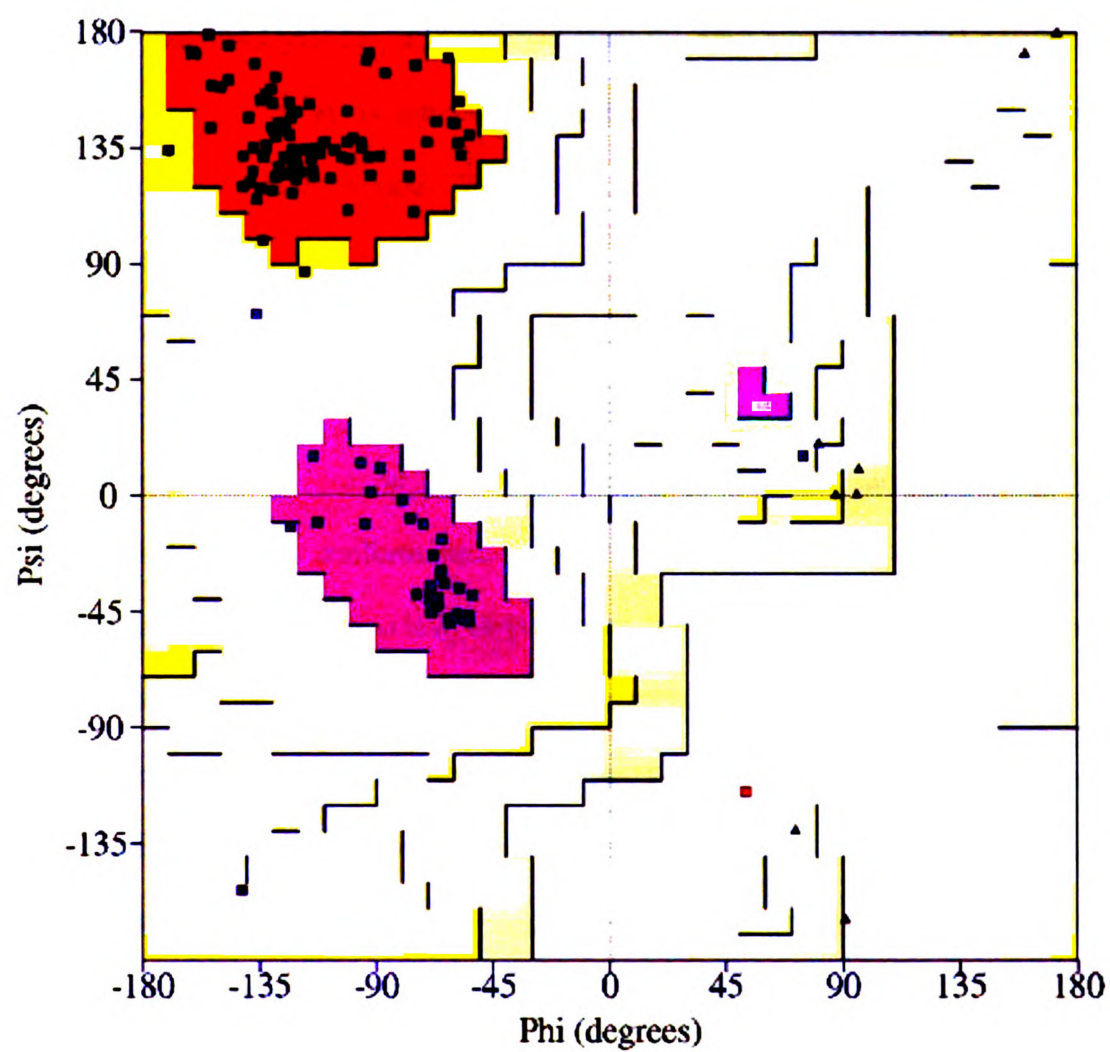


Figure 4-25 Ramachandran plot of R132K:Y134F bound to RA



#### **4.3.8 R132K:Y134F Bound to Retinal**

The structure was determined as explained for WT CRABP II•RA. Inspection of the unbiased, high resolution (1.60 Å) omit Fo-Fc electron density map inside the binding pocket clearly shows that retinal is bound as a free retinal inside the pocket (no SB forms between retinal and Lys132). It is apparent from the omit map that retinal occupies two unequal conformations at its ionone ring and carbonyl group, while it has only one conformation along its backbone. The two conformations are definitely not equal since the omit electron density map is asymmetric at the carbonyl group region. The map has more weight for the conformation of carbonyl group that faces the Arg111 than the one that faces Lys132. Toward the end of the refinement the best combination for the two occupancies was found to be ~ 70% and 30%. It is impossible to assign exact occupancies for the two conformations, based on the electron density map, however it is apparent that one conformation is more occupied. Retinal inside the pocket and within its omit electron density map is shown in Figure 2-9-B in chapter II.

Ramachandran analysis of the data shows that 91.9% of the residues reside in the most favored region of the plot and is similar to other structures, Asp126 A is the only residue that is in a disallowed region (Figure 4-26). The refinement statistics and pdb code of the structure are reported in Table 4-6.

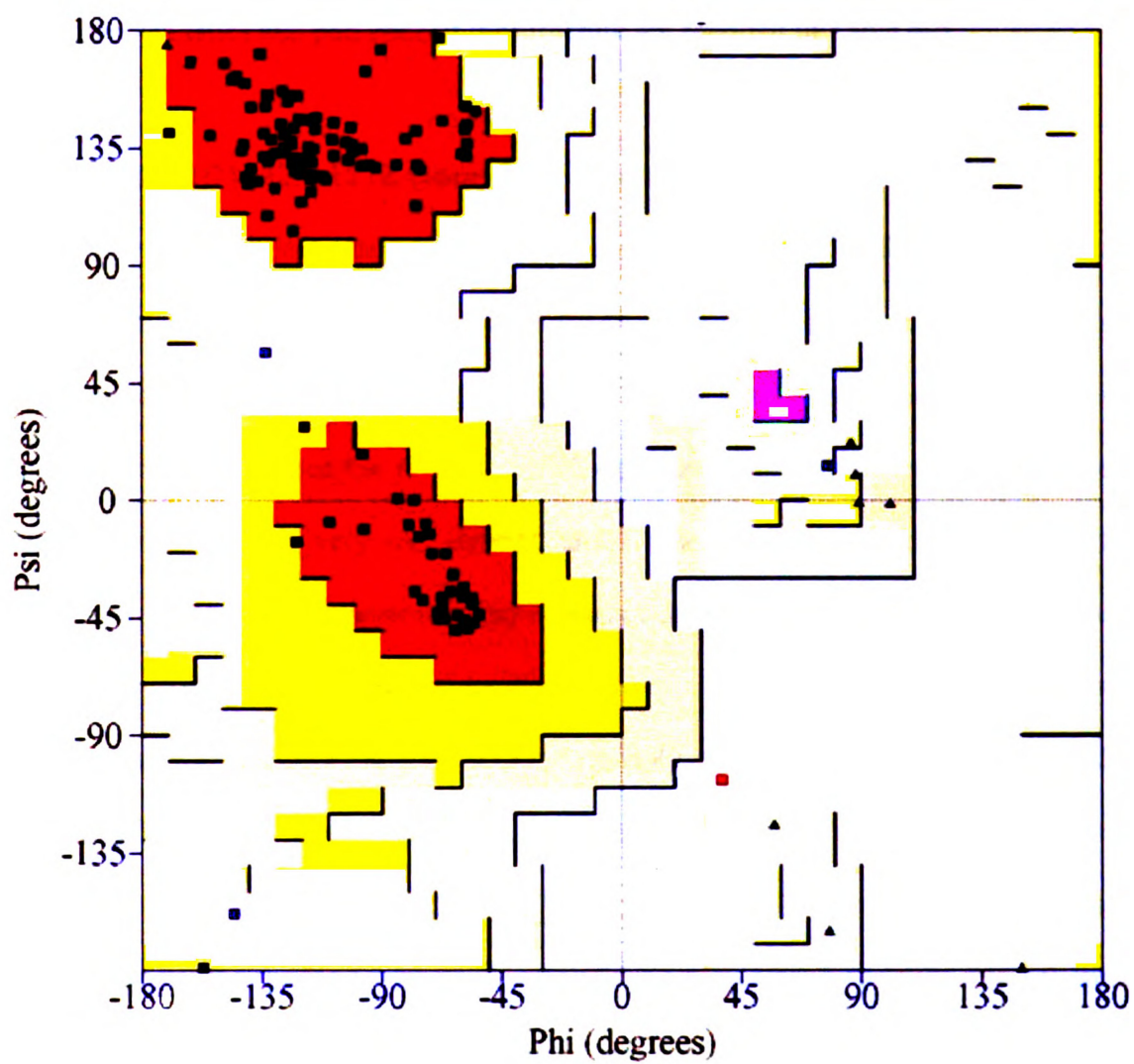


Figure 4-26 Ramachandran plot of the R132K:Y134F bound to retinal

#### **4.3.9 R132K:Y134F:R111L:T54V:L121E Bound to RA**

The structure was determined as described for WT CRABP<sub>II</sub>•RA. Analysis of its Ramachandran plot shows that 94.4% of residues are located in the most favored region of the plot and Asp 126 is the only residue in the disallowed region (Figure 4-27). The refinement statistics and pdb code of the structure are reported in Table 4-6.

#### **4.3.10 R132K:R111L:L121E Bound to Retinal**

The structure was determined as described for WT CRABP<sub>II</sub>•R. Inspection of the electron density map inside the pocket clearly shows that retinal is bound to Lys132 *via* SB formation. The very high resolution data, at 1.20 Å, leaves us no doubt that an imine bond has formed between the N<sup>ε</sup> of Lys 132 and carbonyl carbon of retinal. The very high resolution map is very well-defined around the SB bond and clearly shows the position of the atoms. However the map is broken around the last two methylene groups of Lys132. After building retinal into its omit map some extra positive electron density map formed, which shows the trace of free Lys in the apo-structure of this mutant. Also some negative density was built around the SB bond. This indicates that retinal does not occupy the modeled position 100% of the time and some of the crystallized molecules in the crystal have free lysine residues. However it is apparent from the omit map of retinal that the bound molecules are indeed predominant and therefore retinal was modeled as ~80% occupied, based on the best values obtained for the B-factors of retinal and Lys132, and the R-factors of the structure. Retinal bound to Lys132 inside the pocket and within its omit electron density map is shown in Figure 2-19-B in chapter II. Ramachandran analysis of the data shows that 91.7% of the residues are in the most

favored region. There is no residue in the disallowed region of the plot and only Asp 126 is in the generously allowed region.

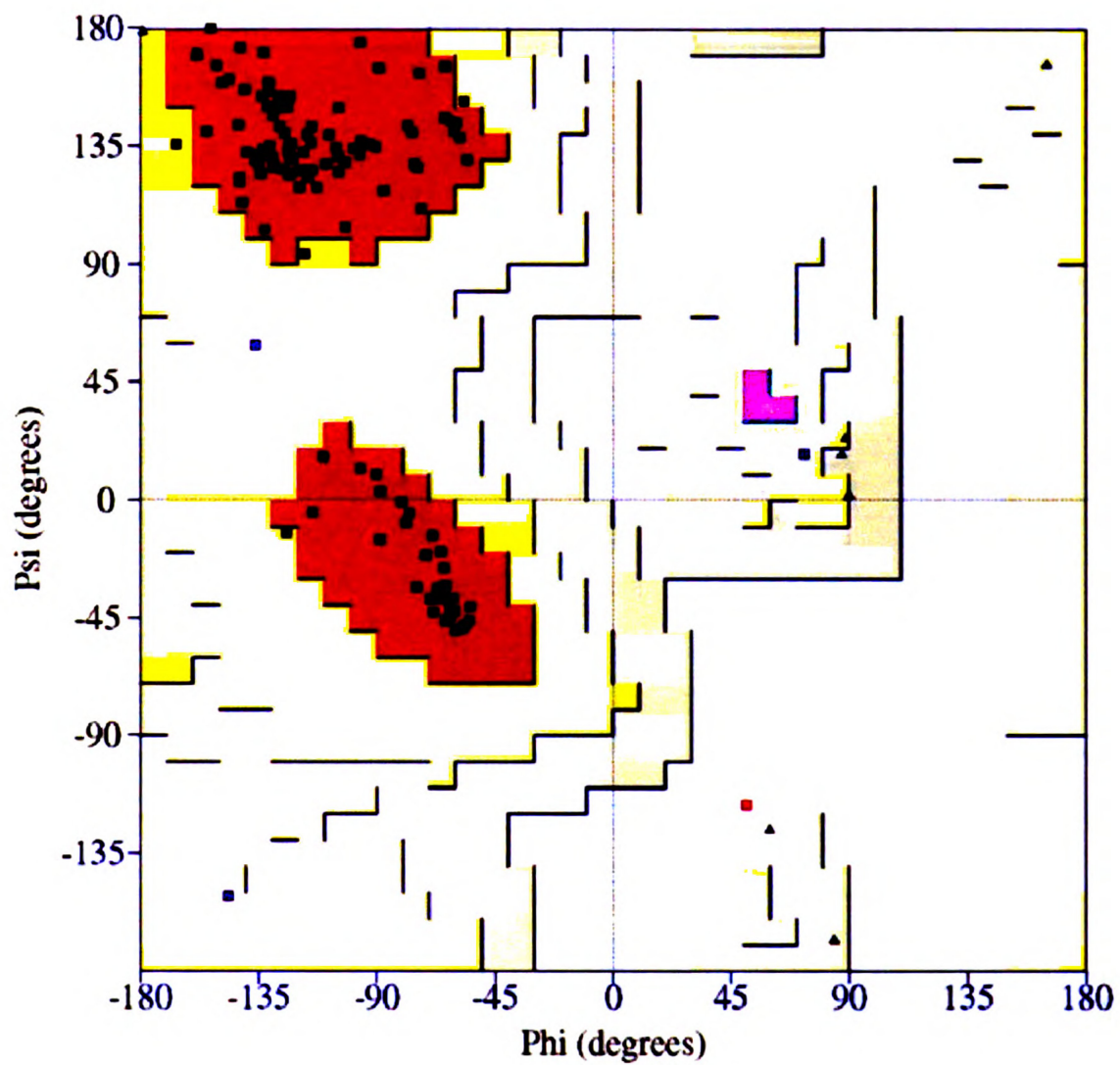


Figure 4-27 Ramachandran plot of the R132K:Y134F:R111L:T54V:L121E bound to RA

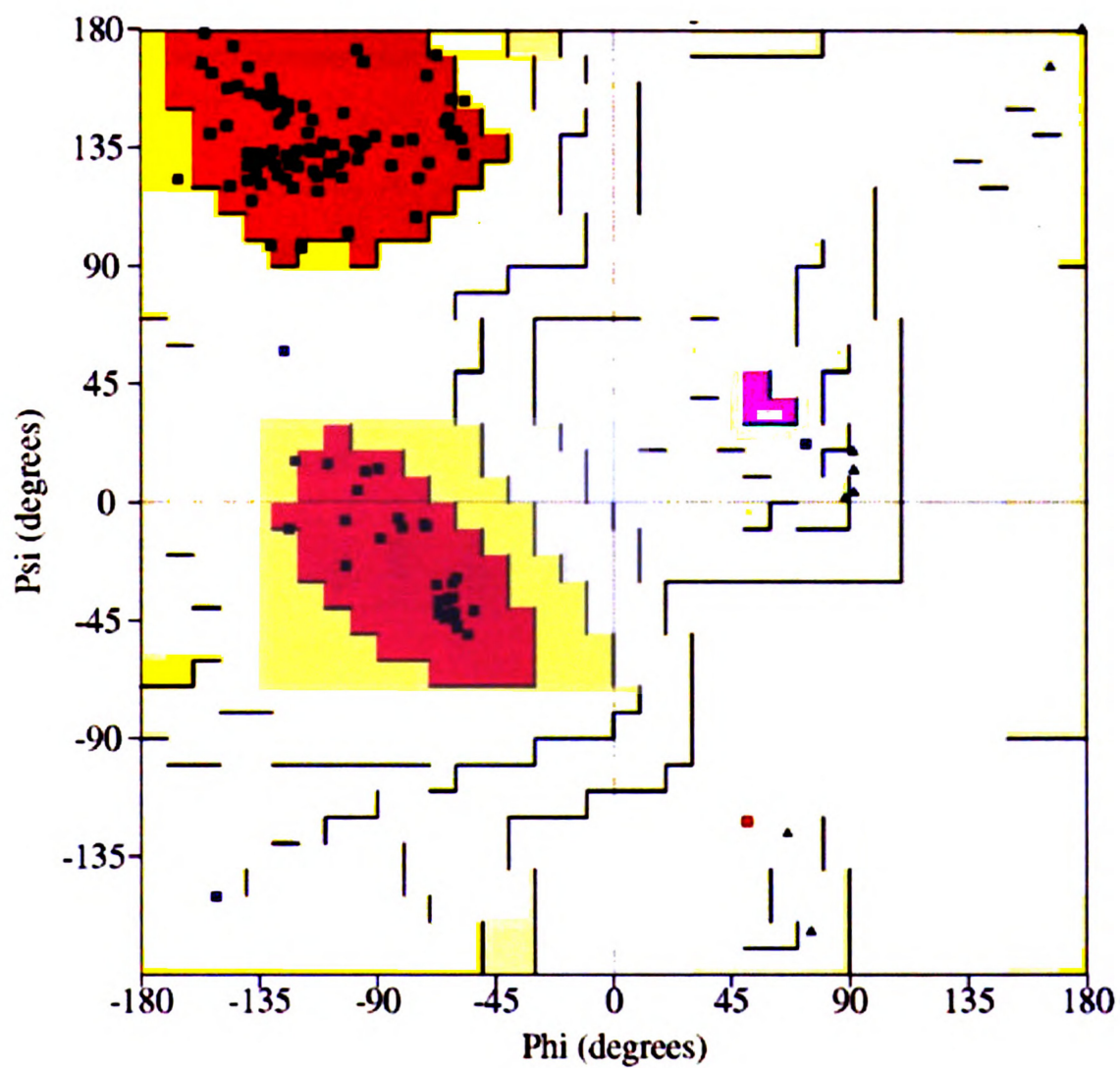


Figure 4-28 Ramachandran plot of the R132K:R111L:L121E bound to retinal

Table 4-6 Refinement statistics for the holo-structures of CRABPII

|                                     | CRABPII-RA | R132K:Y134F-RA | R132K:Y134F-Retinal | R132K:Y134F:R111L:<br>T54V:L121E-RA | R132K:R111L:L121E-Retinal |
|-------------------------------------|------------|----------------|---------------------|-------------------------------------|---------------------------|
| Average B-Factor ( $\text{\AA}^2$ ) | 15.99      | 21.27          | 21.178              | 14.134                              | 12.51                     |
| Resolution ( $\text{\AA}$ )         | 1.48       | 1.70           | 1.69                | 1.60                                | 1.18                      |
| $R_{\text{work}}$ (%)               | 12.30      | 14.50          | 15.52               | 12.06                               | 13.07                     |
| $R_{\text{free}}$ (%)               | 17.09      | 20.28          | 21.37               | 16.79                               | 15.94                     |
| Number of water molecules           | 209        | 187            | 153                 | 221                                 | 253                       |
| Total reflections used              | 26,145     | 17,030         | 17,868              | 21,611                              | 52,405                    |
| Total reflections in work set       | 23,523     | 15,424         | 16,053              | 19,463                              | 49,485                    |
| Total reflections in test set       | 2,622      | 1,606          | 1,815               | 2,148                               | 2,920                     |
| rmsd from ideality                  |            |                |                     |                                     |                           |
| Bond lengths ( $\text{\AA}$ )       | 0.019      | 0.017          | 0.019               | 0.020                               | 0.019                     |
| Bond Angles ( $^\circ$ )            | 1.776      | 1.675          | 1.819               | 1.850                               | 1.805                     |
| <i>Ramachandran plot</i>            |            |                |                     |                                     |                           |
| Most favored (%)                    | 95.2       | 92.7           | 91.9                | 94.4                                | 91.7                      |
| Allowed (%)                         | 4.0        | 6.5            | 7.3                 | 4.8                                 | 7.4                       |
| Generously allowed (%)              | 0.8        | 0.0            | 0.0                 | 0.0                                 | 0.8                       |
| Disallowed (%)                      | 0.0        | 0.8            | 0.8                 | 0.8                                 | 0.0                       |
| PDB code                            | 2FR3       | 2G78           | 2G79                | 2G7A                                | 2G7B                      |

#### 4.4 Literature Cited

1. Vasileiou, C. (2005). Protein design: reengineering cellular retinoic acid binding protein II into a retinal binding protein. Doctor of Philosophy, Michigan State University.
2. Crist, R. M. (2004). Exploring the underlying chemical basis for color vision: designing a protein mimic of rhodopsin. Doctor of Philosophy, Michigan State University.
3. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). 2nd edit. Molecular Cloning: A Laboratory Manual, 1, Cold Spring Harbor Press, New York, NY.
4. Wang, L. C., Li, Y. & Yan, H. G. (1997). Structure-function relationships of cellular retinoic acid-binding proteins - Quantitative analysis of the ligand binding properties of the wild-type proteins and site-directed mutants. *Journal of Biological Chemistry* **272**, 1541-1547.
5. Bradford, M. M. (1976). Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry* **72**, 248-254.
6. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Method Enzymol* (Carter Jr. , C. W. & Sweet, R. M., eds.), Vol. 276, pp. 307-326. Academic Press, New York.
7. Bergfors, T., Kleywegt, G. J. & Jones, T. A. (1994). Crystallization and Preliminary-X-Ray Analysis of Recombinant Bovine Cellular Retinoic Acid-Binding Protein. *Acta Crystallographica Section D-Biological Crystallography* **50**, 370-374.
8. Kleywegt, G. J., Bergfors, T., Senn, H., Lemotte, P., Gsell, B., Shudo, K. & Jones, T. A. (1994). Crystal-Structures of Cellular Retinoic Acid-Binding Protein-I and Protein-II in Complex with All-Trans-Retinoic Acid and a Synthetic Retinoid. *Structure* **2**, 1241-1258.
9. Bailey, S. (1994). The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallographica Section D-Biological Crystallography* **50**, 760-763.



10. Chen, X., Tordova, M., Gilliland, G. L., Wang, L. C., Li, Y., Yan, H. G. & Ji, X. H. (1998). Crystal structure of apo-cellular retinoic acid-binding protein type II (R111M) suggests a mechanism of ligand entry. *Journal of Molecular Biology* **278**, 641-653.
11. Roussel, A. & Cambillau, C. (1989). *Turbo-Frodo. Silicon Graphics Geometry Partners Directory.*, Silicon Graphics, Mountain View.
12. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D-Biological Crystallography* **53**, 240-255.
13. Brunger, A. T. (1993). Assessment of Phase Accuracy by Cross Validation - the Free R-Value - Methods and Applications. *Acta Crystallographica Section D-Biological Crystallography* **49**, 24-36.
14. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica Section D-Biological Crystallography* **54**, 905-921.
15. Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *Journal of Applied Crystallography* **26**, 283-291.

## **Appendices**

## Appendix 1 Summary of Protein Properties for CRABPII Mutants

|    | Protein                       | Calculated Mass | MALDI mass | $\epsilon$ (cm <sup>-1</sup> M <sup>-1</sup> ) | Yield | CD  |
|----|-------------------------------|-----------------|------------|------------------------------------------------|-------|-----|
| 1  | *WT                           | 16829.2         | 16821.9    | 20,215                                         | --    | 221 |
| 2  | *R132K                        | 16801.2         | 16803.6    | 19,202                                         | --    | 225 |
| 3  | *R132K:Y134F                  | 16785.2         | 16788.2    | 18,317                                         | --    | 221 |
| 4  | *R132K:Y134F:L121E            | 16801.2         | 167943.8   | 18,422                                         | --    | 224 |
| 5  | *R132K:Y134F:R111L            | 16742.2         | 16740.6    | 16,654                                         | --    | 220 |
| 6  | *R132K:Y134F:R111E            | 16758.1         | 16759.4    | 16,698                                         | --    | 218 |
| 7  | R132K:Y134F:R111L:L121E       | 15490.7         | 15490.5    | 21,500                                         | 27    | 218 |
| 8  | *R132K:Y134F:R111L:T54E       | 16670.2         | 16770.8    | 18,430                                         | --    | 224 |
| 9  | *R132K:Y134F:R111L:L121D      | 16744.1         | 16753.5    | 19,138                                         | --    | 218 |
| 10 | *R132K:Y134F:R111L:L121E:T54V | 16756.2         | 16765.8    | 19,362                                         | --    | 222 |
| 11 | *R132K:Y134F:R111L:L121D:T54V | 16742.1         | 16740.0    | 22,056                                         | --    | 226 |
| 12 | *R132K:Y134F:R111L:T54V       | 16740.2         | 16740.3    | 17,798                                         | --    | 218 |
| 13 | R132K:Y134F:T54V              | 15515.8         | 15519.8    | 20,934                                         | 40    | 218 |
| 14 | *R132K:Y134F:R111E:T54V       | 16756.2         | 16759.6    | 17,214                                         | --    | 220 |
| 15 | *R132K:Y134F:R111L:L121Q      | 16757.1         | 16752.3    | 17,397                                         | --    | 229 |
| 16 | *R132K:Y134F:R111L:L121Q:T54V | 16755.2         | 16754.6    | 22,821                                         | --    | 224 |
| 17 | *R132K:Y134F:R111L:L121N      | 16743.1         | 16742.3    | 20,266                                         | --    | 222 |

(\*) Indicates His-tagged proteins

## Appendix 1 Summary of Protein Properties for CRABPII Mutants

|    | Protein                       | Calculated<br>Mass | MALDI mass | $\epsilon$ (cm <sup>-1</sup> M <sup>-1</sup> ) | Yield | CD  |
|----|-------------------------------|--------------------|------------|------------------------------------------------|-------|-----|
| 18 | *R132K:Y134F:R111L:L121N:T54V | 16741.1            | 16740.6    | 21,766                                         | ---   | 225 |
| 19 | *R132L:Y134F:R111L:L121E      | 16743.1            | 16747.0    | 19,415                                         | ---   | 220 |
| 20 | *Y134F:R111L:L121E            | 16786.1            | 16781.7    | 19,432                                         | ---   | 218 |
| 21 | R132K:R111L                   | 15490.8            | 15491.9    | 20,846                                         | 22    | 218 |
| 22 | R132K:L121E                   | 15549.8            | 15551.2    | 17,981                                         | 21    | 215 |
| 23 | R132K:R111L:L121E             | 15506.7            | 15505.6    | 23,132                                         | 15    | 218 |
| 24 | R132K:R111L:L121D             | 15492.7            | 15492.9    | 20,636                                         | 11    | 218 |
| 25 | R132K:R111L:L121E:T54V        | 15504.8            | 15505.8    | 19,706                                         | 20    | 218 |
| 26 | R132K:R111L:L121Q             | 15505.8            | 15507.9    | 16,953                                         | 8     | 218 |
| 27 | R132K:R111L:L121Q:T54V        | 15503.8            | 15501.9    | 19,627                                         | 22    | 218 |
| 28 | R132L:R111L:L121E             | ---                | ---        | 20,136                                         | 15    | --- |
| 29 | R132K:R111M                   | 15508.8            | 15507.5    | 23,108                                         | 2     | 218 |
| 30 | R132K:R111M:L121E             | 15524.8            | 15526.7    | 18,678                                         | 24    | 218 |
| 31 | R132K:R111V:L121E             | 15492.7            | 15501.1    | 19,027                                         | 19    | 218 |
| 32 | R132K:R111K                   | 15505.8            | ---        | 20,389                                         | 15    | 218 |
| 33 | R132K:R111K:L121E             | 15521.7            | 15523.1    | 15,068                                         | 12    | 218 |
| 34 | R132K:R111H                   | 15514.7            | 15517.5    | 20,768                                         | 7     | 218 |

(\*) Indicates His-tagged proteins

## Appendix 1 Summary of Protein Properties for CRABPII Mutants

|    | Protein                      | Calculated Mass | MALDI mass | $\epsilon$ (cm <sup>-1</sup> M <sup>-1</sup> ) | Yield | CD      |
|----|------------------------------|-----------------|------------|------------------------------------------------|-------|---------|
| 35 | R132K:R111H:L121E            | 15530.7         | 15532.2    | 19,550                                         | 17    | 218     |
| 36 | R132K:R111E                  | 15506.7         | ---        | 19,748                                         | 6     | 215     |
| 37 | R132K:R111E:L121E            | 15522.7         | ---        | 17,611                                         | 6     | 219     |
| 38 | R132K:R111L:C130D            | 15502.7         | 15504.1    | 23,874                                         | 10    | 217     |
| 39 | R132K:R111L:L121E:C130D      | 15518.7         | 15518.6    | 17,898                                         | 15    | 218     |
| 40 | R132K:R111L:L121E:V41E       | 15536.7         | ---        | 21,871                                         | 5     | 208/217 |
| 41 | R132K:R111L:Y134D            | 15442.7         | ---        | 19,720                                         | 3     | ---     |
| 42 | R132K:R111L:L121E:Y134E      | 15472.7         | ---        | 19,742                                         | 5     | ---     |
| 43 | R132K:R111L:L121E:Y134D      | 15458.6         | ---        | 19,482                                         | 6     | ---     |
| 44 | R132K:R111L:L121E:Y134E:T54V | 15488.8         | ---        | 18,564                                         | 5     | ---     |
| 45 | R132K:W109L                  | 15460.7         | ---        | 16,789                                         | 7     | ---     |
| 46 | R132K:W109L:L121E            | 15476.7         | 15478.1    | 15,289                                         | 4     | ---     |
| 47 | R132K:E73A                   | 15475.8         | 15479.0    | 18,369                                         | 4     | 218     |
| 48 | R132K:E73A:L121E             | 15491.7         | 15494.2    | 20,298                                         | 5     | 218     |
| 49 | R132K:I52D                   | 15535.7         | ---        | 19,800                                         | 5     | ---     |
| 50 | R132K:L121E:I52D             | 15551.7         | ---        | 20,100                                         | 8     | ---     |

(\*) Indicates His-tagged proteins

## Appendix 2 Summary of Binding Assay Results for CRABP II Mutants

|    | Protein                       | RA K <sub>d</sub> (nM) | Rt K <sub>d</sub> (nM) | UV λ <sub>max</sub> (nm)  | UV 2 <sup>nd</sup> der               | Incubation | Reductive Amination |
|----|-------------------------------|------------------------|------------------------|---------------------------|--------------------------------------|------------|---------------------|
| 1  | *WT                           | 2.0±1.2                | 6600±360               | 377                       | 390                                  | No         | No                  |
| 2  | *R132K                        | 65±14                  | 280±17                 | 392, 2 min<br>376, 15 min | 417(1.2), 377(1)<br>415(1), 376(2.2) | Yes        | Yes                 |
| 3  | *R132K:Y134F                  | 100±7.1                | 120±4.9                | 396                       | 405                                  | Yes        | No                  |
| 4  | *R132K:Y134F:L121E            | 1400±63                | 160±9.8                | 381                       | 381                                  | No         | Yes                 |
| 5  | *R132K:Y134F:R111L            | 1000±28                | 160±6.7                | 398                       | 399                                  | Yes        | Yes                 |
| 6  | *R132K:Y134F:R111E            | 530±31                 | 260±12                 | 385                       | 393                                  | No         | No                  |
| 7  | *R132K:Y134F:R111L:L121E      | 770±61                 | 200±8.3                | 378, 438                  | 372(5), 449(1)                       | Yes        | Yes                 |
| 8  | *R132K:Y134F:R111L:T54E       | 2180±43                | 80±3.4                 | 381                       | 379                                  | Yes        | No                  |
| 9  | *R132K:Y134F:R111L:L121D      | 200±25                 | 2.6±7.9                | 375, 440                  | 371(9), 450(1)                       | Yes        | Yes                 |
| 10 | *R132K:Y134F:R111L:L121E:T54V | 250±19                 | 2.7±7.0                | 382, 430                  | 370(1), 447(1.4)                     | Yes        | Yes                 |
| 11 | *R132K:Y134F:R111L:L121D:T54V | 760±43                 | 30±5.5                 | 379                       | 373(26), 457(1)                      | Yes        | Yes                 |
| 12 | *R132K:Y134F:R111L:T54V       | 900±64                 | 84±12                  | 400                       | 412                                  | Yes        | Yes                 |
| 13 | R132K:Y134F:T54V              | 565±79                 | 81±14                  | 417                       | 422                                  | ---        | Yes                 |
| 14 | *R132K:Y134F:R111E:T54V       | 860±36                 | 210±8.2                | 395                       | 399                                  | Yes        | No                  |
| 15 | *R132K:Y134F:R111L:L121Q      | 240±22                 | 530±44                 | 386                       | 390                                  | Yes        | Yes                 |
| 16 | *R132K:Y134F:R111L:L121Q:T54V | 1739±94                | 600±66                 | 377                       | 386                                  | Yes        | Yes                 |
| 17 | *R132K:Y134F:R111L:L121N      | 3050±240               | 1270±49                | 385                       | 387                                  | Yes        | Yes                 |

(\*) Indicates His-tagged proteins; UV results in the presence of 0.1 equiv all-*trans*-retinal; Numbers in parenthesis represent the relative ratio

## Appendix 2 Summary of Binding Assay Results for CRABPII Mutants

|    | Protein                       | RA K <sub>d</sub> (nM) | Rt K <sub>d</sub> (nM) | UV λ <sub>max</sub> (nm) | UV 2 <sup>nd</sup> der | Incubation | Reductive Amination |
|----|-------------------------------|------------------------|------------------------|--------------------------|------------------------|------------|---------------------|
| 18 | *R132K:Y134F:R111L:L121N:T54V | 420±34                 | 180±15                 | 404                      | 400                    | Yes        | Yes                 |
| 19 | *R132L:Y134F:R111L:L121E      | 160±20                 | 180±14                 | 395                      | 396                    | No         | No                  |
| 20 | *Y134F:R111L:L121E            | 220±22                 | 570±32                 | 384                      | 385                    | No         | No                  |
| 21 | R132K:R111L                   | 736±114                | 567±36                 | 401, 2 min 377, 50 min   | 408 376                | —          | Yes                 |
| 22 | R132K:L121E                   | 1260±73                | 104±11                 | 457                      | 350(1), 458(6.6)       | —          | Yes                 |
| 23 | R132K:R111L:L121E             | 426±47                 | 1.36±4.9               | 449                      | 449                    | Yes        | Yes                 |
| 24 | R132K:R111L:L121D             | 2639±189               | 133±11                 | 441, 15 min              | 444(18), 341(1)        | Yes        | Yes                 |
| 25 | R132K:R111L:L121E:T54V        | 400±36                 | 1.99±4.1               | 449                      | 451                    | Yes        | Yes                 |
| 26 | R132K:R111L:L121Q             | 306±33                 | 1526±83                | 376                      | 370                    | Yes        | Yes                 |
| 27 | R132K:R111L:L121Q:T54V        | 1458±119               | 637±67                 | 375                      | 369                    | Yes        | Yes                 |
| 28 | R132L:R111L:L121E             | —                      | —                      | 377                      | 370                    | —          | —                   |
| 29 | R132K:R111M                   | 699±71                 | 1142±47                | 378                      | 375                    | Yes        | Yes                 |
| 30 | R132K:R111M:L121E             | 742±75                 | 164±32                 | 448                      | 450                    | Yes        | Yes                 |
| 31 | R132K:R111V:L121E             | 2313±254               | 30.9±63                | 446                      | 448                    | No         | Yes                 |
| 32 | R132K:R111K                   | 5016±633               | 112±25                 | 388, 2 min 372, 15 min   | 392, 2 min 372, 15 min | —          | Yes                 |
| 33 | R132K:R111K:L121E             | 486±31                 | 1619±65                | 365, 455                 | 357, 460               | Yes        | Yes                 |
| 34 | R132K:R111H                   | 1362±95                | 293±34                 | 385, 2 min 376, 20 min   | 386, 2 min 376, 20 min | —          | No                  |

(\*) Indicates His-tagged proteins; UV results in the presence of 0.1 equiv all-*trans*-retinal; Numbers in parenthesis represent the relative ratio

## Appendix 2 Summary of Binding Assay Results for CRABP II Mutants

|    | Protein                      | RA $K_d$ (nM) | Rt $K_d$ (nM) | UV $\lambda_{max}$ (nm) | UV 2 <sup>nd</sup> der | Incubation | Reductive<br>Amination |
|----|------------------------------|---------------|---------------|-------------------------|------------------------|------------|------------------------|
| 35 | R132K:R111H:L121E            | 264±12        | 4.1±4.7       | 446                     | 455                    | --         | Yes                    |
| 36 | R132K:R111E                  | 1088±95       | 965±17.5      | 386                     | 380                    | --         | --                     |
| 37 | R132K:R111E:L121E            | 608±53        | 1435±90       | 382                     | 378(1), 410(1)         | --         | --                     |
| 38 | R132K:R111L:C130D            | 6105±699      | 103±10        | 377                     | 380                    | --         | --                     |
| 39 | R132K:R111L:L121E:C130D      | 258±36        | 52±10         | 440                     | 446                    | No         | No                     |
| 40 | R132K:R111L:L121E:V41E       | 490±43        | 87±20         | 448, 15 min             | 451, 15 min            | --         | --                     |
| 41 | R132K:R111L:Y134D            | 7419±404      | 5937±397      | 379                     | 387                    | --         | --                     |
| 42 | R132K:R111L:L121E:Y134E      | 3297±265      | 4755±253      | 381                     | 385                    | --         | --                     |
| 43 | R132K:R111L:L121E:Y134D      | 2718±188      | 5403±339      | 378                     | 381                    | --         | --                     |
| 44 | R132K:R111L:L121E:Y134E:T54V | 3665±248      | 4802±273      | 380                     | 385                    | --         | --                     |
| 45 | R132K:W109L                  | 3196±252      | 986±107       | 381                     | 381                    | --         | --                     |
| 46 | R132K:W109L:L121E            | 93±9          | 191±18        | 370, 455                | 367(1), 464(2)         | --         | --                     |
| 47 | R132K:E73A                   | 564±56        | 528±42        | 378                     | 379                    | --         | --                     |
| 48 | R132K:E73A:L121E             | 353±35        | 526±73        | 374, 452                | 371(1), 461(1.3)       | --         | --                     |
| 49 | R132K:I52D                   | 2742±158      | 662±62        | 386                     | 380                    | --         | --                     |
| 50 | R132K:L121E:I52D             | --            | --            | 382                     | 383                    | --         | --                     |

(\*) Indicates His-tagged proteins; UV results in the presence of 0.1 equiv all-*trans*-retinal; Numbers in parenthesis represent the relative ratio



MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 02845 1155