

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

STUDIES ON 15-15'- β -CAROTENE DIOXYGENASE AND REENGINEERING
CELLULAR RETINOIC ACID BINDING PROTEIN II INTO A RETINAL
BINDING PROTEIN AND ITS INTERACTION WITH RETINAL MIMICS

VOLUME I

By

Montserrat Rabago-Smith

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

STUDIES ON 15-15'- β -CAROTENE DIOXYGENASE AND REENGINEERING CELLULAR RETINOIC ACID BINDING PROTEIN II INTO A RETINAL BINDING PROTEIN AND ITS INTERACTION WITH RETINAL MIMICS

By

Montserrat Rabago-Smith

Vitamin A and its derivatives are C_{20} isoprenoids and its metabolic pathway is tied to the activity of β -Carotene Dioxygenase (BCDOX). Although it has been known since 1930 that vitamin A is derived *in vivo* from β -carotene, the enzymatic origin of β -carotene cleavage is not fully understood. The current view is that central cleavage is clearly the predominant pathway but there is also evidence that excentric cleavage of carotenoids occurs in plants and microorganisms. Even though in the last decade a large amount of information has been gathered for BCDOX, many questions remain unanswered in regard to its mechanism of action. Further proof of whether the cleavage of β -carotene occurs via a monooxygenase or dioxygenase mechanism is needed. Because of the importance of β -carotene as the main source of retinoids *in vivo*, and because the oxidation of β -carotene is such a unique example for regiospecific oxidation, we were interested in studying the mode of control for such oxidation, as well as how such oxidation occurs for an inactivated olefin. Therefore the BCDOX protein was expressed in *E. coli*. When a weak promoter was used the protein was expressed in low yield as soluble protein. This protein had an activity of 194 pmol retinal/mg BCDOX•min and a $K_M=3 \mu M$. Due to the low expression yield the protein was cloned into a plasmid that contained a stronger promoter, but the protein was only obtained as inclusion bodies. Our interest in the mechanistic aspects

of BCDOX led us to synthesize the tritiated photoaffinity label [10'-³H]-8'-*apo*-β-carotenoic acid. In an attempt to develop a partition assay for BCDOX the following tritiated compounds were prepared *trans*-[15,15'-³H]-β-carotene, all-*trans*-[15-³H]-retinal and all-*trans*-[15-³H]-retinol were prepared. The mechanism by which we see colors has been long studied, but even today, it is not fully understood. So far it is known that there are four proteins in the eye which allow us to see: rod rhodopsin and three cone rhodopsins, red, blue, and green. Each of the four proteins binds the same compound, 11-*cis*-retinal, as a protonated Schiff base (PSB) with an active site lysine residue. The key to color vision is due to different interactions between 11-*cis*-retinal and each one of the four proteins in the eye. Understanding these protein substrate interactions is the topic of study for many researchers. Various theories have been developed in an attempt to explain this phenomenon. But due to the difficulties in expressing, handling, and studying rhodopsin proteins, these studies have been hampered. To solve this problem a rhodopsin surrogate that is easier to handle, manipulate and crystallize was engineered. By using rational protein design and through site directed mutagenesis CRABP_{II} has been converted into a retinal binding protein. The 11-*cis*-retinal was synthesized and binds to the engineered CRABP_{II} protein as a protonated Schiff base through an engineered lysine residue. Two different all-*trans*-retinal analogs were synthesized, and the mode of binding has been found to be very similar to what is observed in the all-*trans*-retinal system. The synthesis of [6-¹³C]-L-Lysine was performed. This labeled amino acid would be used to calculate the pK_a of the Lysine residue that forms the PSB with retinal.

**Dedicated to my parents
for their love and unconditional support.**

ACKNOWLEDGMENTS

My deepest thanks to all the people that helped me accomplish the completion of my degree, in all aspects research physically, emotionally, or spiritually.

First I would like to thank Babak, for his mentoring and patience. For the time that we spent together, I am very grateful for the times when he taught me how to trust myself. He always pushed me to the end of my limits. He taught me how to approach the problems that I encountered in my research on my own, and allowed me to make my own decisions and mistakes, in order to learn from them. Also I appreciate all his ability to motivate me when times were difficult. Babak has he been an excellent mentor, and a great friend. He has helped to make my time at MSU a wonderful and fulfilling experience.

I must also thank Dr. Chrysoula Vasileiou and Dr. Rachael Crist for always being there, both as friends and excellent co-workers. My progress and level of understanding for this research have benefited from the intellectual foresight of Chrysoula and Rachel. I the years that I have been here, they have been a constant source of wisdom and insight, as well as friendship. Special thanks are due for certain people for their efforts and assistance in various aspects of my research. I thank Dr. Qifei Yang, Dr. Radha Narayan, Marina Tasanova and Dr. Meenakshi Sivakumar for their useful discussions we had covering all aspects of life and science. I also would like to thank all the members of the lab for their help entertainment and friendship. Dan, Jennifer, Jun Shang, Sonmath, Stewart, Tao, Xiaofei, Xiaoyong, Ali, Golala. I thank Dr. Pulgam Veera Reddy for the collaboration in the synthesis of 10^{-3}H-apo

carotenoic acid. I cannot forget the undergraduate students Laura, Sarah and Rocky for their efforts and enthusiasm both in learning all aspects of the project helping out wherever possible. Without the collaborative efforts of all these individuals, the progress of this research would have been severely hampered.

My thanks also goes out to our collaborators on the protein crystallography efforts, Professor James Geiger (Department of Chemistry, Michigan State University) and students Erika Mathes and Soheila Vaezslami. I also wish to thank Professor Honggao Yan (Department of Biochemistry, Michigan State University) who provided both the instrumentation and helpful staff for acquiring all mass spectral data. The genomics facility (Michigan State University) provided all DNA sequencing analysis and the Macromolecular Structure Facility (Department of Biochemistry, Michigan State University) provided all the primers.

A special mention is due for some close friend who, throughout the years provided the emotional support necessary for maintaining an acceptable level of sanity. These people are Chrysoula, Glenn, Mapitso, Rachael and Connie.

I must also thank all of the friends and family who have provided a constant stream of support. My Mom and Dad have always been there for me their love encouragement of my dreams. Specifically, I wish to thank my Mom for her unwavering confidence in me, as well as for her great sense of humor and strength. I thank my Dad for his example and continuous encouragement to pursue my interests. My brother also supported me by helping me to keep sight on what is really important in life. I thank my Grandma and aunt Amelia Smith for all their love and support and for being proud of me.

Finally, I would like to thank my husband Steve for all the moments we share during this time. We have laughed together, we have cried together and trained together. His patience and understanding have been key to overcome all of my difficulties. His love and interior peace had helped me to see the light at the end of the tunnel. Thank you for all your love and support.

TABLE OF CONTENTS

List of Tables.....	xi
List of Figures.....	xiii
List of Schemes.....	xx
Key to Symbols and Abbreviations.....	xxiii

Part I.

Chapter 1

I. Introduction

1.1	Carotenoids as netabolic source of retinoids.....	1
1.2	β -Carotene Dioxygenase.....	5
1.3	15,15' β -Carotene Dioxygenase.....	8
1.4	Mechanism of action.....	13
1.5	References.....	23

Chapter 2

II. Expression of BCDOX

2.1	Expression of using a T5 promoter.....	30
2.2	β -Carotene Dioxygenase activity.....	33
2.3	Expression of BCDOX with rare codons.....	36
2.4	Expression of BCDOX <i>in vivo</i>	43
2.5	Expression of BCDOX using a T7 promoter.....	46
2.6	Cloning of BCDOX into a T7 promoter vector.....	46
2.7	Expression of pET-29b(+)-BCDOX.....	52
2.8	Optimization of pBCDOX expression.....	52
2.9	Attempts to express the recombinant BCDOX in a soluble form.....	53
2.9.1	Decrease the rate of expression, use of different temperatures, time and IPTG.....	53
2.9.2	Different <i>E. coli</i> strain.....	56
2.9.3	Over expression of chaperones.....	58
2.10	Activity <i>in vivo</i>	61

2.11	Expression of the native BCDOX under the control of a T7 promoter.....	63
2.11.1	Mutation of BCDOX to add a stop codon to avoid translation of the 6xHis tag and thrombine cut site.....	63
2.11.2	Cloning BCDOX.....	64
2.11.3	Expression of pBCDOX-3; native BCDOX.....	66
2.12	Isolation of inclusion bodies.....	67
2.13	Materials and Methods.....	71
2.14	References.....	85

Chapter 3

III. Studies on BCDOX

3.1	Introduction.....	92
3.2	Synthesis of all- <i>trans</i> -[10'- ³ H]-8'apo- β -carotenoic acid.....	95
3.3	Enzymatic studies using of all- <i>trans</i> -8'apo- β -carotenoic acid and BCDOX.....	98
3.4	Failed photolabeling attempts using all- <i>trans</i> -8'apo- β -carotenoic acid and BCDOX.....	100
3.5	Attempts to develop a rapid and efficient assay for BCDOX.....	102
3.6	Materials and Methods.....	108
3.7	References.....	125

Part II

Chapter 4

IV Introduction

4.1	Vision.....	129
4.1.1	Process of vision.....	130
4.1.2	Color vision.....	136
4.2	Designing of a rhodopsin surrogate.....	145
4.3	Wild Type CRABP II binding properties.....	149
4.4	References.....	162

Chapter 5

V Attempts to label Lys132 to determine the pK_a of CRABP II mutants

5.1	Introduction.....	178
5.2	Non selective labeling of CRABP II.....	183
5.2.1	Use of ¹⁵ N-NMR to determine the pK _a of Lys.....	183

5.2.2	Use of ^{13}C -NMR.....	184
5.2.3	Synthesis of 6- ^{13}C -lysine of vision.....	185
5.3	Specific labeling of Lysine	195
5.3.1	Chemical ligation.....	195
5.3.2	Native chemical ligation.....	196
5.3.3	Expression of CRABPII fused to a mini intein.....	201
5.3.4	Attempts to solubilize fusion protein.....	202
5.3.5	Refolding of <i>Mxe</i> GyrA intein system.....	203
5.3.6	Use different intein.....	208
5.4	Labeling of Lys with ϵ - ^{15}N -Lys or 6- ^{13}C -Lys.....	213
5.4.1	Conversion of Cys to Thio Lys.....	214
5.4.2	Synthesis of the alkylating agent to convert Cys to Thio-Lys.....	224
5.5	Materials and Methods.....	230
5.6	References.....	252

Chapter 6

VI Interaction of 11-*cis* retinal and other chromophores with the protein rhodopsin mimics

6.1	Introduction.....	263
6.2	Preparation of mutants to study wavelength regulation.....	273
6.3	Studies using 11- <i>cis</i> -retinal.....	276
6.3.1	Synthesis using 11- <i>cis</i> -retinal.....	277
6.3.2	Protein substrate interactions using 11- <i>cis</i> -retinal.....	283
6.4	Retinal analogs.....	288
6.4.1	Synthesis of different chromophores.....	292
6.4.2	Binding studies wit merocyanine 3 and azulene 4.....	300
6.5	Materials and Methods.....	313
6.6	References.....	338

Chapter 7

VII Attempts to develop an assay to identify rhodopsin surrogates

7.1	Introduction.....	347
7.2	<i>In vitro</i> test by using pORANGE, pBCDOX-1 and pCRABPII.....	348
7.3	<i>In vivo</i> test for protonated PBS formation using pORANGE pRET-CRABPII.....	353
7.4	Materials and Methods.....	362
7.5	References.....	368

LIST OF TABLES

Table 1-1.	Different substrates tested for BCDOX activity.....	20
Table 2-1.	Comparison of protein activity and expression after addition of glucose.....	35
Table 2-2.	Protein expression after addition of rare codons using XL1-Blue as host.....	40
Table 2-3.	Expression of recombinant pBCDOX-2 in <i>Rosetta(DE3)pLysS</i> , induced at different concentrations of IPTG and different temperatures.....	55
Table 2-4.	Expression of recombinant pBCDOX-2 in <i>Rosetta(DE3)pLysS</i> , induced by addition of different portions IPTG of OD ₆₀₀ ~ 1.0 for overnight at 26 °C.....	55
Table 2-5.	Expression of recombinant pBCDOX-2 using different host induced by addition of IPTG at different portions.....	57
Table 2-6.	Expression of recombinant BCDOX in the presence of its native substrate β -carotene.....	58
Table 2-7.	Overexpression of chaperones to improve the solubility of the recombinant BCDOX.....	60
Table 2-8.	Attempts to test the recombinant mice β -carotene-15-15' dioxygenase pBCDOX-2 <i>in vivo</i>	62
Table 2-9.	Expression of recombinant pBCDOX-3 in <i>Rosetta(DE3)pLysS</i> , induced at different concentrations of IPTG and temperatures.....	66
Table 5-1.	Attempts to selectively reduce N-CBZ-Glutamic acid α -methyl ester.....	187
Table 5-2.	Attempts to hydrogenate nitrile 19.....	188
Table 5-3.	Reduction of nitrile using Raney-Ni.....	193
Table 5-4.	Use of different <i>E. coli</i> . hosts to test solubility of the fusion intein protein.....	202
Table 6-1.	Spectroscopic characteristics of rhodopsin proteins of different species.....	272

Table 6-2.	Reduction of alkyne 14 to obtain 11- <i>cis</i> -retinal (1).....	280
Table 6-3.	Titration of different CRABPII mutants with 11- <i>cis</i> -retinal and all- <i>trans</i> -retinal.....	288
Table 6-4.	HWE between ylide 30 and aldehyde 27 to obtained chloro ester 31	295
Table 6-5.	Spectroscopic data of merocyanine 3 and the bacteriorhodopsin (bRh).....	302
Table 6-6.	Titration of different CRABPII mutants with merocyanine 3 and all- <i>trans</i> -retinal.....	305
Table 6-7.	Titration of different CRABPII mutants with azulene 4 and all- <i>trans</i> -retinal.....	311
Table 7-1.	Expression of carotenoid genes from <i>E. uredoovora</i> in different <i>E. coli</i> . strain.....	353
Table 7-2.	Results from the attempt to develop an assay CRABPII-mutants that form PSB with all- <i>trans</i> -retinal.....	358

LIST OF FIGURES

Note: Select figures in this manuscript are presented in color.

Figure 1-1.	Various retinoids.....	1
Figure 1-2.	Representation of Schiff base formed in rhodopsin.....	2
Figure 1-3.	Activation of the GPCR rhodopsin, the G-protein transducin, and the enzyme PDE as the visual process is initiated.....	3
Figure 1-4.	Sequence comparison of human BCDOX (H-BCDOX), mouse BCDOX (M-BCDOX), chicken BCDOX (C-BCDOX), and <i>Drosophila</i> BCDOX (D-BCDOX), and human RPE65.....	10
Figure 1-5.	Different substrates used to evaluate activity in chicken BCDOX.....	19
Figure 2-1.	Depiction of pQE30 plasmid with mice BCDOX gene subcloned in.....	30
Figure 2-2.	Gel depiction of the purification of His-tagged BCDOX protein.....	31
Figure 2-3.	HPLC chromatographs of enzymatic reaction of BCDOX and β -carotene.....	34
Figure 2-4.	Gel depiction of the purification of His-tagged BCDOX protein with impurities.....	37
Figure 2-5.	Depiction of the sequence of mice BCDOX gene showing its 29 rare codons.....	38
Figure 2-6.	Comparison of relative activity and expression of recombinant mice BCDOX using different <i>E. coli</i> hosts.....	42
Figure 2-7.	Depiction of pORANGE plasmid containing a gene cluster to express β -carotene.....	43
Figure 2-8.	<i>In vivo</i> activity of recombinant BCDOX.....	45
Figure 2-9.	Comparison of the amino acid sequence of the reported mice BCDOX and the obtained BCDOX.....	48
Figure 2-10.	Depiction of pET-29b(+) plasmid with mice BCDOX gene subcloned in.....	51

Figure 2-11.	Depiction of pET-29b(+) plasmid with mice BCDOX gene subcloned in, with out a His tag.....	65
Figure 2-12.	Gel depiction of the purification of BCDOX protein as inclusion bodies.....	68
Figure 3-1.	Inhibition studies of BCDOX using all- <i>trans</i> -[10'- ³ H]-8' <i>apo</i> - β -carotenoic acid	99
Figure 3-2.	Depiction of a partition assay to test BCDOX activity.....	103
Figure 3-3.	Comparison of extraction of retinal, retinol and β -carotene using a partition assay	105
Figure 4-1.	Schematic representation of the human eye.....	129
Figure 4-2.	Schematic representation of the rod and cone cells in retina.....	130
Figure 4-3.	Schematic representation of rhodopsin in rod cells.....	131
Figure 4-4.	Representation of Schiff base formed in rhodopsin.....	132
Figure 4-5.	Schematic representation of the intermediates observed in the isomerization of 11- <i>cis</i> -retinal to all- <i>trans</i> -retinal in rhodopsin.....	133
Figure 4-6.	Schematic representation of the change observed in rhodopsin after the isomerization of 11- <i>cis</i> -retinal to all- <i>trans</i> -retinal.....	134
Figure 4-7.	Activation of the GPCR rhodopsin, the G-protein transducin, and the enzyme PDE as the visual process is initiated.....	135
Figure 4-8.	Schematic representation of the visual transduction pathway.....	136
Figure 4-9.	Representation of the visible region of the electromagnetic spectrum and the rhodopsin absorbance patterns that result in color vision.....	137
Figure 4-10.	Comparison of absorbances of 11- <i>cis</i> -retinal as free aldehyde, as a Schiff base (SB) and as a protonated Schiff base (PSB).....	138

Figure 4-11.	Distance of the counter anion and placement of other charges or dipoles along the backbone of the polyene may be a possible mode of wavelength regulation.....	139
Figure 4-12.	Absorption maxima and apparent pK_a values of Schiff base (SB) and protonated SB (PSB) of retinal with various amino acids. The relative position of the carboxylic acid dictates the λ_{max} and the pK_a values.....	141
Figure 4-13.	Twisting of the planes about the double bonds of 11- <i>cis</i> -retinal may be a possible mode of wavelength regulation	142
Figure 4-14.	Representation of 11- <i>cis</i> -retinal, this molecule is not planar due to steric interactions.....	143
Figure 4-15.	Representation of the crystal structure of CRABPII bound to all- <i>trans</i> -retinoic acid.....	148
Figure 4-16.	Representation of interactions between wild type CRABPII and all- <i>trans</i> -retinoic acid.....	149
Figure 4-17.	Fluorescence and UV-vis data for CRABPII-WT.....	150
Figure 4-18.	Fluorescence and UV-vis data for CRABPII-R132K.....	152
Figure 4-19.	Fluorescence, UV-vis and MALDI-TOF data for CRABPII-R132K::Y134F.....	154
Figure 4-20.	Representation of the crystal structure of CRABPII-R132K::Y134F bound to all- <i>trans</i> -retinoic acid.....	156
Figure 4-21.	Representation of the conformation of all- <i>trans</i> -retinoic acid bound to CRABPII-R132K::Y134F.....	157
Figure 4-22.	Fluorescence, UV-vis and MALDI-TOF data for CRABPII-R132K::R111L::L121E.....	158
Figure 4-23.	Representation of the crystal structure of CRABPII-R132K::R111L::L121E bound to all- <i>trans</i> -retinal.....	159
Figure 4-24.	Representation of the proposed mechanism for the Schiff base formation between CRABPII- R132K::R111L::L121E and all- <i>trans</i> -retinal.....	160

Figure 5-1.	Comparison of UV-vis spectrum of CRABPII-WT, CRABPII-R132K::R111L::L121E and CRABPII-R132K::Y134F::R111L::L121E.....	178
Figure 5-2.	Acid/base titration of the triple mutant CRABPII-R132K::R111L::L121E.....	180
Figure 5-3.	Schematic representation of the intein-mediated native chemical ligation process.....	197
Figure 5-4.	Schematic representation of the formation of ¹⁵ N-Lys132 CRABPII <i>via</i> intein-mediated native chemical ligation.....	199
Figure 5-5.	Gel depiction of the over expression of the small intein fusion CRABPII protein.....	201
Figure 5-6.	Gel depiction of the small intein fusion CRABPII protein after dialysis.....	205
Figure 5-7.	Gel depiction of the over expression of the large intein fusion CRABPII protein.....	210
Figure 5-8.	Gel depiction of the binding of over expression of the large intein fusion CRABPII protein to chitin beads.....	211
Figure 5-9.	Gel depiction of the cleavage with MESNA and DTT of large intein fusion CRABPII protein.....	212
Figure 5-10.	Gel depiction of the cleavage with MESNA and DTT of large intein fusion CRABPII protein.....	213
Figure 5-11.	Comparison of lysine and thiolysine.....	214
Figure 5-12.	Representation of the three Cys in CRABPII.....	216
Figure 5-13.	UV-vis data for the titration of CRABPII-R132C::R111L::L121E and all- <i>trans</i> -retinal.....	218
Figure 5-14.	UV-vis data for the titration of CRABPII-R132C::R111L::L121E and all- <i>trans</i> -retinal.....	221
Figure 5-15.	Fluorescence, UV-vis and MALDI-TOF data for CRABPII-R132C::R111L::L121E after alkylation.....	223
Figure 6-1.	Representation of 11- <i>cis</i> -retinal and all- <i>trans</i> -retinal.....	263

Figure 6-2.	Binding pocket of the rod rhodopsin from its crystal structure showing the PSB between the 11- <i>cis</i> -retinal and Lys296, the important interactions between amino acids have been labeled.....	266
Figure 6-3.	Binding site of ion pump bacteriorhodopsin.....	270
Figure 6-4.	Representation of the important amino acids in the crystal structures of retinal binding site of bacteriorhodopsin.....	271
Figure 6-5.	Distance of the counter anion and placement of other charges or dipoles along the backbone of the polyene may be a possible mode of wavelength regulation.....	273
Figure 6-6.	Representation of the amino acid residues within 7 of the polyene in the triple mutant CRABPII-R132K::R111L::L121E.....	274
Figure 6-7.	Representation of 11- <i>cis</i> -retinal, this molecule is not planar due to steric interactions.....	277
Figure 6-8.	Separation of alkyne precursors of 11- <i>cis</i> -retinal.....	279
Figure 6-9.	UV-vis spectra of 11- <i>cis</i> -retinal.....	282
Figure 6-10.	Representation of two possible conformation of 11- <i>cis</i> -retinal, in the triple mutant CRABPII-R132K::R111L::L121E.....	284
Figure 6-11.	Comparison of absorbances of 11- <i>cis</i> -retinal as free aldehyde, as a Schiff base (SB) and as a protonated Schiff base (PSB).....	285
Figure 6-12.	Model structure of the binding cavity of the triple mutant CRABPII-R132K::R111L::L121E and the mutated amino acids 11- <i>cis</i> -retinal.....	286
Figure 6-13.	UV-vis absorbances of the ligand 11- <i>cis</i> -retinal with different rhodopsin surrogates.....	287
Figure 6-14.	Acid/base titration of the triple mutant CRABPII-R132K::L121E.....	287
Figure 6-15.	Representation of merocyanines that form a very stable PSB due to the fact that the positive charge is stabilized all around the polyene.....	291
Figure 6-16.	Representation of azulene and its remarkable polarizability.....	291
Figure 6-17.	Representation of the stabilization of positive charges in azulenenes..	292

Figure 6-18.	Different aldehydes commercially available.....	297
Figure 6-19.	Fill in model of binding of merocyanine 3 to the triple mutant CRABPII-R132K::R111L::L121E.....	301
Figure 6-20.	UV-vis data for the titration of merocyanine 3 with the triple mutant CRABPII-R132K::R111L::L121E.....	303
Figure 6-21.	Comparison of UV-vis data for the titration of merocyanine 3 with CRABPII-R132K::R111L::L121E, CRABPII-R132K::R111L::L121E::R59W, CRABPII-WT.....	304
Figure 6-22.	Comparison of UV-vis data for the titration of merocyanine 3 with CRABPII-R132K::R111L::L121E, CRABPII-R132K::R111L::L121Q, CRABPII-R132K::R111L.....	306
Figure 6-23.	UV-vis data for the titration of merocyanine 3 with CRABPII-R132K::R111L::L121E::R59W.....	307
Figure 6-24.	Fill in model of binding of azulene 4 to the triple mutant CRABPII-R132K::R111L::L121E.....	308
Figure 6-25.	Comparison of UV-vis data for the titration of azulene 4 with CRABPII-R132K::R111L::L121E, CRABPII-R132K::R111L CRABPII-WT.....	309
Figure 6-26.	Acid / base titration of the triple mutant CRABPII-R132K::R111L::L121E bound to azulene 4.....	310
Figure 7-1.	Representation of different position in CRABPII to possibly insert the nucleophilic lysine.....	347
Figure 7-2.	Absorption differences between retinal and the PSB of retinal.....	347
Figure 7-3.	Depiction of pQE30 plasmid with mice BCDOX gene subcloned in and a pORANGE plasmid that contains the genes to express β -carotene <i>in vivo</i>	348
Figure 7-4.	Production of retinal <i>in vivo</i> in XL1-Blue <i>E. coli</i>	349
Figure 7-5.	Production of retinal <i>in vivo</i> in different <i>E. coli</i>	350
Figure 7-6.	Representation of an assay to identify mutants that form a PSB between retinal and CRABPII mutants.....	351

Figure 7-7.	Depiction of pET17b plasmid with CRABP II gene subcloned in.....	352
Figure 7-8.	Fluorescence, UV-vis data for CRABP II -R132K::Y134F::R111L::L121E and CRABP II -WT.....	355
Figure 7-9.	Depiction of pQE30 plasmid with CRABP II gene and mice BCDOX subcloned in.....	357
Figure 7-10.	Representation of an <i>E. coli</i> system to assay CRABP II mutants that form PSB with retinal.....	357

LIST OF SCHEMES

Scheme 1-1.	Biosynthesis of retinoids, β -carotene oxidation provides retinoids.....	4
Scheme 1-2.	β -Carotene oxidation by BCDOX. Central cleavage (15-15' oxidation) vs excentric cleavage.....	6
Scheme 1-3.	9- <i>cis</i> -Epoxycarotenoid dioxygenase VP14, excentric oxidative cleavage.....	7
Scheme 1-4.	General mechanism for activation of O ₂ in dioxygenases.....	14
Scheme 1-5.	Carotene dioxygenase mechanism.....	15
Scheme 1-6.	Carotene monooxygenase mechanism.....	17
Scheme 1-7.	Alternate carotene dioxygenase mechanism.....	18
Scheme 2-1.	Kinetics of the recombinant β -carotene 15-15' dioxygenase. Monitoring the conversion of β -carotene to retinal.....	35
Scheme 2-2.	Biosynthesis of β -carotene in <i>E. coli</i> . starts with the condensation between DMAPP and IPP.....	44
Scheme 3-1.	The enzymatic mechanism of BCDOX is intriguing in that utilizes a molecule of oxygen to cleave an electronically undistinguished double bond with a high degree of selectivity.....	92
Scheme 3-2.	Postulated mechanisms. A) Dioxygenase, suggest the involvement of a dioxetane intermediate. B) Monooxygenase, suggest the involvement of an epoxide intermediate.....	93
Scheme 3-3.	NaB ³ H ₄ was chosen to be the source of tritium and thus required a reductive step to incorporate the label.....	95
Scheme 3-4.	Synthesis of all- <i>trans</i> -[10'- ³ H]-8' <i>apo</i> - β -carotenoic acid. Part I.....	96
Scheme 3-5.	Synthesis of all- <i>trans</i> -[10'- ³ H]-8' <i>apo</i> - β -carotenoic acid. Part II.....	97
Scheme 3-6.	The proposed mechanism for cross linking of $\alpha\beta$ -unsaturated acids involves a diradical intermediate.....	101
Scheme 3-7.	Synthesis of ³ H labeled β -carotene and retinal.....	104
Scheme 5-1.	In the formation of a protonated Schiff base various equilibrium are involved.....	177

Scheme 5-2.	Retrosynthetic analysis for the preparation of [6- ¹³ C]-L-Lysine.....	185
Scheme 5-3.	Synthesis of N-CBZ-Glutamic acid α-methyl ester 12.....	186
Scheme 5-4.	Selective reduction of N-CBZ-Glutamic acid α-methyl ester 12.....	189
Scheme 5-5.	Preparation of nitrile 18.....	190
Scheme 5-6.	Proposed mechanism for the hydrogenation of nitriles.....	191
Scheme 5-7.	Reduction of nitrile using Raney-Ni, Pd/C.....	194
Scheme 5-8.	Conversion of Cys to Thiolsine via alkylation of cysteine 25.....	215
Scheme 5-9.	Formation of a PSB between thiolsine and all- <i>trans</i> -retinal.....	220
Scheme 5-10.	The label ethanolamine could be used to prepare iodoethyl-trifluoroacetamide. and bromo ethyl amine.....	225
Scheme 5-11.	Synthesis of alkylating agents using ethanol amine as the starting material	227
Scheme 5-12.	Possible mechanisms involved in the bromination of the label ethanolamine.....	228
Scheme 6-1.	Synthesis of acetyline (9) intermediate in the synthesis of 11- <i>cis</i> -retinal.....	277
Scheme 6-2.	Synthesis of acetyline (14) precursor of the Zn mediated reduction critical step in the synthesis of 11- <i>cis</i> -retinal.....	277
Scheme 6-3.	Retrosynthesis of ylide 22.....	293
Scheme 6-4.	Synthesis of aldehyde 27.....	293
Scheme 6-5.	Synthesis of ylide 30 starting from 3,3, dimethyl acrylic acid (24)....	294
Scheme 6-6.	Synthesis of ylide 33.....	296
Scheme 6-7.	Synthesis of various aldehydes.....	297
Scheme 6-8.	HWE using ylide 33 and various aldehydes.....	298
Scheme 6-9.	Synthesis of merocyanine 3.....	299

Scheme 6-10. Synthesis of azulene 4	300
--	-----

KEY TO SYMBOLS AND ABBREVIATIONS

Å	Angstrom
ϵ	Extinction coefficient
μg	Micro grams
μL	Micro liters
μM	Micro molar
λ_{max}	Maximal wavelength
Abs, A	Absorbance
ALBP	Adipocyte Lipid Binding Protein

Amino Acids:

Ala, A	Alanine
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartic acid
Cys, C	Cysteine
Gln, Q	Glutamine
Glu, E	Glutamic acid
His, H	Histidine
Leu, L	Leucine
Lys, K	Lysine
Phe, F	Phenylalanine
Pro, P	Proline

Thr, T	Threonine
Trp, W	Tryptophan
Tyr, Y	Tyrosine
Val, V	Valine
Amp	Ampicillin
AMP	Adenine Monophosphate
BCDOX	β -Carotene dioxygenase
bRh	Bacteriorhodopsin
BSA	Bovine Serum Albumin
bp	Base pairs
br	Broad
CAP	Catabolite activation
CD	Circular Dichroism
CDNA	Deoxyribonucleic acid
C-BCDOX	Chicken β -Carotene dioxygenase
cGMP	Cyclic-Guanidine Monophosphate
CIP	Calf Intestine Phosphatase
Clm	Chloramphenicol
CRABPI	Human Cellular Retinoic Acid Binding Protein I
CRABPII	Human Cellular Retinoic Acid Binding Protein II
CRBP	Cellular Retinol-Binding Proteins
d	Doublet
Da	Daltons

D-BCDOX	<i>Drosophila</i> β -Carotene dioxygenase
DCM	Dichloromethane
DIBAL	Diisobutyl aluminum hydride
DMAPP	Dimethyl allyl pyrophosphate
DMF	Dimethyl formamide
DMSO	Dimethyl sulfoxide
DTT	dithiothreitol
DNA	Deoxyribonucleic acid
<i>E</i>	<i>trans</i>
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetra acetic acid
Et ₂ O	Ether
EtOAc	Ethyl acetate
GDP	Guanidine Diphosphate
GFP	Green Fluorescent Protein
GMP	Guanidine Monophosphate
GTP	Guanidine Triphosphate
GPCR	G-Protein Coupled Receptor
h	Hour
H-BCDOX	Human β -Carotene dioxygenase
Hex	Hexanes
HMQC	Heteronuclear Multiple Quantum Coherence
HPLC	High pressure liquid chromatography

HSPS	heat shock proteins
HWE	Horner-Wadsworth-Emmons
Hz	Hertz
iLBPs	Intracellular lipid-binding proteins
IMPACT™-CN System	Intein Mediated Purification with an Affinity Chitin-binding Tag, New England Biolabs)
IPTG	Isopropylthiogalactoside
IPP	Isopentyl pyrophosphate
IUPAC	International Union of Pure and Applied Chemists
<i>J</i>	<i>J</i> coupling constant
KDa	Kilo daltons
Kb	Kilo-Base pairs
K_d	Dissociation constant
K_M	Michaelis Menten constant
LB	Luria bertani
LPA	Lipid-protein aggregate
m	Multiplet
MALDI-TOF	Matrix assisted laser desorption ionization – Time of flight
M-BCDOX	Mice β-Carotene dioxygenase
MESNA	2-mercaptoethanesulfonic acid
MeOH	Methanol
mL	Mililiters
MHz	Mega hertz

Min	Minute
MRNA	Messenger ribonucleic acid
MSU	Michigan State University
mCi	Mili curies
N-CBZ	N-benzyloxycarbonloxy)
nm	Nanometer
nM	Nanomolar
NMA	N-methylmercaptoacetamide
NMR	Nuclear Magnetic Resonance
ON	Over night
Op	Opsin
PBS	Phosphate buffered saline
PCC	Pyridinium chlorochromate
PCR	Polymerase chain reaction
PDE	Phosphodiesterase
pmol	pico moles
PSB	Protonated Schiff base
RA	Retinoic acid
RARs	Retinoic acid receptors
RBP	Serum Retinol-Binding Proteins
Rh	Rhodopsin
RPE65	Retinal Pigment Epithelium
Rt	Retinal

RT	Room temperature
RXR _s	Retinoid X receptors
s	Singlet
SB	Schiff base
SDS-PAGE	Sodium dodecyl sulfate – Polyacrylamide Gel Electrophoresis
t	Triplet
THF	Tetrahydrofuran
TLC	Thin layer chromatography
Tris	2-amino-hydroxymethyl-1,3-propanediol
tRNA	Transfer ribonucleic acid
V _{MAX}	Maximum velocity
WT	Wild type
Z	<i>Cis</i>

Mutant proteins are designated by the one letter abbreviation for the wild type amino acid, followed by the amino acid position number, followed by the one letter abbreviation for the new mutant amino acid.

Note: Select Figures in this manuscript are presented in color

Chapter 1

Introduction

1.1 Carotenoids as a metabolic source of retinoids

Carotenoids are C_{40} isoprenoids with a wide variety of structures and biological activities. Carotenoid biosynthesis is mostly found in plants, fungi and several bacterial species.^{3,4} Carotenoids contain a characteristic polyene structure, which is responsible for light absorption as well as for singlet oxygen quenching or inactivation of aggressive radicals. More importantly, carotenoids and their

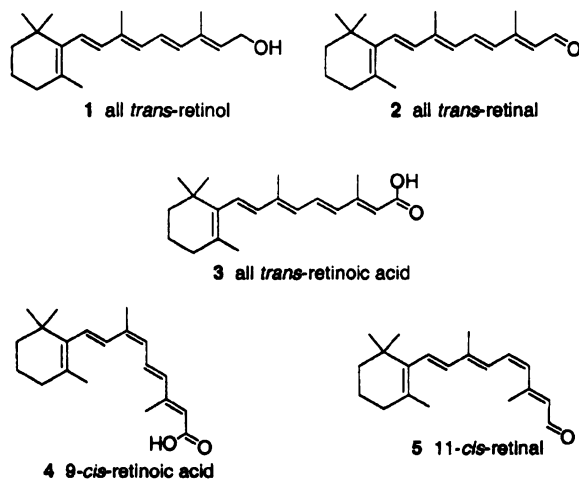
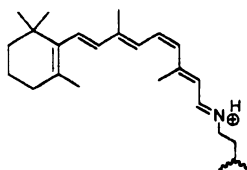


Figure 1-1. Different retinoids. Retinoids are isoprene derivatives.

oxidation products (retinoids) play an essential role in various pathways such as cellular physiology, visual transduction,⁵ gene regulatory control, cancer⁶ and other disease prevention.⁷⁻⁹

Retinoids are a group of molecules derived from four isoprene units joined head to tail (Figure 1.1).¹⁰ Retinol is known as vitamin A. 11-*Cis*-retinal **5** is essential for vision in the entire animal

kingdom. The visual pigments (rhodopsin) of animals are composed of a retinal chromophore (vitamin A aldehyde) bound to a lysine in the protein (opsin) via a Schiff



base as shown in Figure 1-2.^{5,11} Light induces isomerization of the chromophore, and thus leading to a conformational change activating the visual pigments (rhodopsin).

Figure 1-2. A Schiff base is formed between 11-*cis*-retinal and a lysine residue in Rhodopsin.

Activation of rhodopsin triggers a G protein-coupled receptor cascade that leads to changes in the permeability of the photoreceptor cell membranes, as shown in Figure 1-3. It is also known that different physiological concentrations of retinoic acid induce cell differentiation at the embryonic stage. The retinoic acid receptors (RARs) and retinoid X receptors (RXRs) are members of the steroid receptor super family of proteins, which function as ligand dependent transcription factors.¹² All *trans*-retinoic acid **3** and 9-*cis*-retinoic acid **4** can dictate signal transduction pathways *in vivo*, which is key in controlling

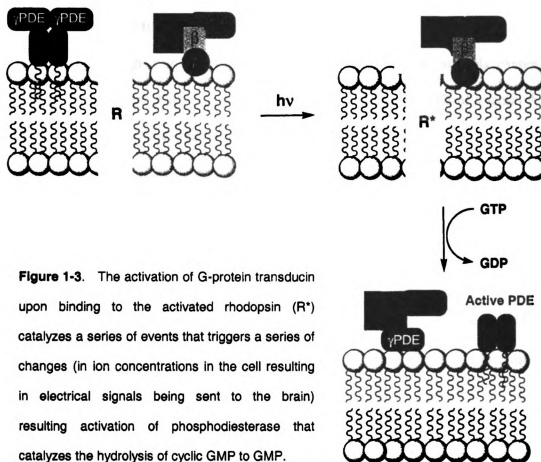


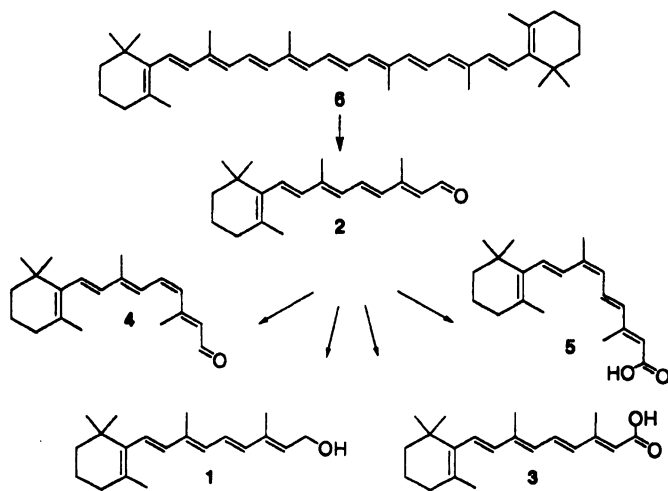
Figure 1-3. The activation of G-protein transducin upon binding to the activated rhodopsin (R*) catalyzes a series of events that triggers a series of changes (in ion concentrations in the cell resulting in electrical signals being sent to the brain) resulting activation of phosphodiesterase that catalyzes the hydrolysis of cyclic GMP to GMP.

transcriptional activity of RARs and RXRs. Therefore, the concentration of retinoic acid in cells must be tightly controlled.

It has been suggested that the concentration of retinoids in the cell is regulated by a number of specific carrier or binding proteins that belong to the super-family of intracellular lipid-binding proteins (iLBPs). These proteins include: Serum Retinol-Binding Proteins (RBP), transporters of retinol from storage in the liver; Cellular Retinol-Binding Proteins (CRBP), specific carriers of the alcoholic form of vitamin A 1; Cellular Retinoic Acid-Binding Protein

(CRABP), carriers of retinoic acid; Cellular Retinal-Binding Protein and Interstitial Retinol Binding Proteins.¹³⁻¹⁶

Even though retinoids are essential for protein function, animals cannot biosynthesize them. However, it has been estimated that up to 80% of vitamin A in mammals is accumulated through the oxidative metabolism of carotenoids. Of all the known carotenoids 50-60 display pro-vitamin A activity, in particular, β -carotene (1).^{17,18} β -Carotene is oxidized to form retinal which can be reoxidized to produce retinoic acid (3, 4),¹⁹⁻²¹ or reduced to form retinol 1, (Scheme 1-1). Therefore, β -carotene is considered to be the inlet for many physiologically important retinoid compounds, which affect a multitude of biological systems.^{8,22-26} β -Carotene seems to be absorbed intact by passive diffusion at the intestinal



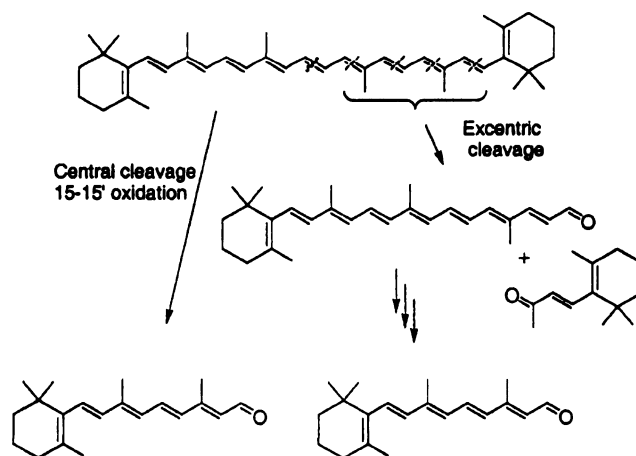
Scheme 1-1. Biosynthesis of retinoids, β -carotene oxidation provides retinoids.

brush border and it shows no toxicity at high doses.²⁷ This suggests that β -carotene could be used as a safe precursor for the *in vivo* production of retinoids in mammals. Except vision, retinoic acid (*trans*-retinoic acid **3** and 9-*cis*-retinoic acid **4**) exerts most of the physiological functions. It has been suggested that oxidation of retinal is the most important metabolic pathway for the production of these retinoic acids.¹⁹⁻²¹

1.2 β -Carotene Dioxygenase

Vitamin A and its derivatives are C₂₀ isoprenoids and its metabolic pathway is tied to the activity of β -Carotene Dioxygenase (BCDOX). Although it has been known since 1930²⁸ that vitamin A is derived *in vivo* from β -carotene, the enzymatic origin of β -carotene cleavage was only shown in 1965 when Goodman and Olson were able to demonstrate that cell free enzymatic preparations isolated from rat intestinal mucus convert β -carotene to retinal.^{29,30} Up to the mid 1960s, the central cleavage of β -carotene to two molecules of vitamin A (aldehyde) was the most popular pathway for the oxidation of β -carotene to afford two molecules of retinal (Scheme 1-2).^{29,30} *In vitro* studies have demonstrated a stoichiometric formation of two moles of retinal per mole of β -carotene cleaved, suggesting a 15-15'C oxidative cleavage.^{31,32}

However, it was not until the nineties that Wang *et al.* provided evidence for *in vivo* and *in vitro* production of *apo*-carotenals from β -carotene. These results supported a mechanism previously proposed by Glover,³³ involving the excentric cleavage of β -carotene to *apo*-carotenoids. Thus, a renewed interest in the



Scheme 1-2. β -Carotene oxidation by BCDOX. Central cleavage (15-15' oxidation) vs excentric cleavage.

stepwise cleavage of β -carotene to one molecule of vitamin A has recently resurged, (Scheme 1-2).^{31,32,34-38} The current view is that central cleavage is clearly the predominant pathway in healthy animals and human subjects,³¹ although stepwise cleavage is enhanced in intestinal tissue by oxidative stress,^{35,39,40} and there is also evidence that excentric cleavage of carotenoids occurs in plants and microorganisms.^{41,42}

In plants, a protein called VP14 catalyzes the oxidative cleavage of 9-*cis*-violaxanthin, a molecule similar to β -carotene as part of the biosynthetic pathway of the plant hormone abscisic acid as revealed in Scheme 1-3.^{41,43} VP14 was the first carotenoid dioxygenase that was molecularly identified; and is an example of an excentric carotenoid cleavage. There are several other examples for excentric cleavage reactions, such as the formation of saffron⁴⁴ in crocus or *apo*-carotenals

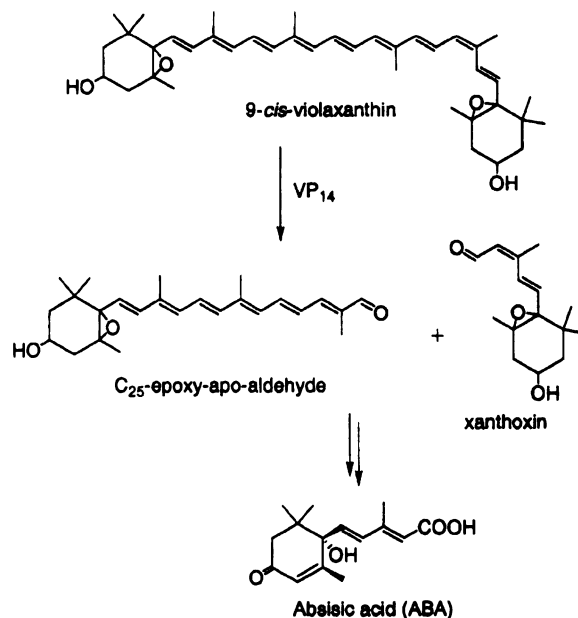
in citrus fruits;
however, whether or
not this process
occurs in mammals
remains to be seen.

Dimitrovskii has
reported the presence
of an *apo*-carotene
dioxygenase that
cleaves *apo*-
carotenoids to β -*apo*-
14'-carotenal. In

2001, von Lintig

reported the discovery of a β -carotene dioxygenase that cleaves at the 9'-10' double bond of β -carotene, resulting in the formation of β -*apo*-10'-carotenal and β -ionone.⁴⁵

The fact that different oxidation products are obtained with different enzyme preparations suggest that the presence of the observed *apo*-carotenoids could be due to the presence of an β -carotene dioxygenase isozyme that cleaves β -carotene at a different double bond.⁴⁶ Therefore, the enzyme responsible of the central cleavage of β -carotene to retinal was called 15-15'- β -carotene dioxygenase (15-15'-BCDOX).



Scheme 1-3. 9-cis-Epoxycarotenoid dioxygenase VP14, excentric oxidative cleavage.

1.3 15,15'- β -Carotene Dioxygenase

The past few years have witnessed a resurgence of activity in the study of BCDOX, mainly as a result of the identification and isolation of the gene responsible for BCDOX in a number of organisms.^{41,45,47-49} Von Lintig and co-workers expressed the BCDOX from *Drosophila melanogaster* into the genetic background of an *E. coli* that accumulates β -carotene,⁵⁰ which resulted in the formation of retinoids that could be monitored spectroscopically.⁴⁷ The cDNA obtained encodes a protein of 620 amino acids with an estimated molecular mass of 70 kDa and an apparent K_M of 5 μ M. Blaner and colleagues used the cDNA reported by Von Linting to obtain a cDNA encoding a mouse BCDOX and found a protein of 566 amino acids with an estimated molecular mass of 64 kDa and an apparent K_M of 0.95 μ M. The human BCDOX was characterized and purified by Andersson and Lindqvist. They purified the protein from baculovirus-infected *Spodoptera frugiperda* 9 insect cells. The human BCDOX had an estimated molecular mass of 62 kDa and an apparent K_M of 7 μ M.⁵¹ This protein expresses highly in the Retinal Pigment Epithelium (RPE65), kidney, intestine, liver, brain, stomach and testis. The gene is ~20 Kbp and is composed of 11 exons and 10 introns that maps to chromosome 16, q21-q23. The open reading frame of human BCDOX contains 547 amino acids. It is interesting to notice that the position of the BCDOX gene is close to the BB52 locus for Bardet Bield syndrome (an autosomal recessive disease associated with pigmentary retinopathy, mental retardation, polydactyly, obesity and hypogenitalims). Andersson reported that expression of human BCDOX in bacteria was problematic, thus they used insect

cells to express BCDOX. All the previous reports suggest BCDOX to be a monomer, however their observations suggest that the active BCDOX was a tetramer. Although this is the first evidence that BCDOX might not be a monomer, the expression of BCDOX as a tetramer could be an artifact of expression in insects, or it could be just simply due to a different form of the protein in different organisms.

Many other BCDOX's have been isolated and cloned, and so far all the reported proteins verify that BCDOX's are cytosolic proteins that have an open reading frame (ORF) of about 1.7 to 3.1 Kbp, 544-620 amino acids and a molecular weight of 62-70 kDa.^{41,45,47-49} So far all the reported proteins are monomers with the exception of human BCDOX, which as mentioned previously was suggested to be a tetramer. The enzymes are dependant on molecular O₂ and iron cation; the pH for optimum activity is slightly alkaline (7.8-8.2). It has been reported that this enzyme is inhibited by sulfhydryl alkylating reagents, and it is either protected or activated by the presence of thiols such as cysteine or glutathione.²⁹ The Michaelis-Menten constant for the native BCDOX enzymes of a variety of animal species relative to β -carotene has been estimated to be in the range of 0.95-10 μ M. It has also been suggested that BCDOX might have an amphiphilic character; some reports suggest that this enzyme is associated to a high-molecular weight lipid-protein aggregate (LPA), which might be necessary for optimal activity.³²

In chickens, mice and humans, it has been established that the mRNA for BCDOX was localized primarily in duodenal viles as well as in tubular structures

Sequence comparison

H-BCDOX	MDI-IFG---	RN---R---	-K-----EQL	-----	-----	13
M-BCDOX	.E.-----	Q.-----K---	-----	-----	-----	13
C-BCDOX	.ET--.N---	..-----K---	-E-----HP	-----	-----	13
D-BCDOX	.AAGV.KSFM	.DFFAVKYDE	QRNDPQA.R.	DGNGRLYPNC	SSDVWLRSC	50
H-RPE65	.S.-QVE---	HPAGGYK---	-.LFETV.E.	S-----	-----	23
A						
H-BCDOX	----EPVRAK	VTGKIPAWLQ	GTDLKNGFDM	HTVGESRYNH	WGLALLHS	59
M-BCDOX	-----Q..	...S.....K...	59
C-BCDOX	-----IK.E	.Q.QL.T...	.V.....	.I.DTK...	59
D-BCDOX	REIVD.IEGH	HS.H..K.IC	.S.....S	WK..DMTFG.	L..CS....R	100
H-RPE65	----S.LT.H	...R..L..T	.S.....L	FE..SEPFY.	L..Q....K	69
H-BCDOX	FTIRDGEVYY	RSKYLRSDTY	NTNIEANRIV	VSEFGTMAYP	DPCKNIFSKA	109
M-BCDOX	.S.....F.Q....	IA.....	109
C-BCDOX	.TFKN.....C.....A..	109
D-BCDOX	.A..N.R.T.	QNRFDVTE.L	RK.RS.Q...	.T....A.V.	...HS..DR-	149
H-RPE65	.DFKE.H.T.	HRRFI.T.A.	VRAMTEK...	IT....C.F.RF	119
H-BCDOX	FSYLSHTIPD	FTDNCLINIM	KCGEDFYATS	ETNYIRKINP	QTLETLEKVD	159
M-BCDOXTD.	159
C-BCDOXET.D.Y...	...F....D.D...	159
D-BCDOX	.AAIFRPDSG	...SM.S.Y	PF.DQY.TFT	..PFMHR...	C..A.EARIC	198
H-RPE65	...F-RGV-E	V...A.V.VY	PV...Y..CT	..F.T....	E....IKQ..	167
H-BCDOX	YRKYAVNLA	TSHPHYDEAG	NVLNMGTSIV	EKGKTKYVIF	KIPATVPEGK	209
M-BCDOXV.	D..R.....DS.	209
C-BCDOX	.S.....S.	.I.....I.	D..R....L.	...SS...KE	209
D-BCDOX	TTDF.G.VNH	...VLPS.	T.Y.L..TMT	RS.PA..T.L	SF.--HG.QM	245
H-RPE65	LCN..S..G.	.A...IEND.	T.Y.I.NCFG	KNFSIA.N.V	...PLQADKE	217
B						
H-BCDOX	KQGKSPWKHT	EVFCSIPSR	LLSPSYH	EQPFRLDILK	259	
M-BCDOX	.K...V..AS...K....	259		
C-BCDOX	.K...CF..L	.V.....K...V.	258		
D-BCDOX	----FEDA	H.VATL.C.W	K.H.G.M.T.	...LSVSLTE	289	
H-RPE65	---D.ISKS	.IVVQF.CSD	RFK...V...	.T.VKINLF.	263	
H-BCDOX	MATAY-IRRM	SWASCLAFHR	EKTYIHID	QRTQPVQTK	FYTDAMVVF	308
M-BCDOX-M.GVMS.D.	.D.....K..P..P....	308
C-BCDOX	L.....GV	N.....S.HK	.D..WF.FV.	RK.KKE.S..L.LY.	307
D-BCDOX	YIK.Q-LGGQ	NLSA..KWFE	DRP.LP.L..	RVSGKL...	YESE.PFYL	337
H-RPE65	FLSSWSLWGA	NYMD.FESNE	TMGVWL..A.	KKRKKYLNN.	YR.SPFNL.	313

Figure 1-4A. Sequence comparison of human BCDOX (H-BCDOX), mouse BCDOX (M-BCDOX), chicken BCDOX (C-BCDOX), and *Drosophila* BCDOX (D-BCDOX), and human RPE65. The most conserved regions are shaded and labeled A-B. Part I. These homology was generated by Zack and colleagues using GeneWorks 2.5.1 (Oxford Scientific).²

H-BCDOX	C HEDGC	IVFDVIAYED	-NSLYQLFY-	-LANL--NQD	FKE---NSRL	350
M-BCDOX	VL.....	-S.....	-.....K.	.E---K...		350
C-BCDOX	H V...IV..R.	-N...DM..	-KK.--DK.	.EV---.NK.		349
D-BCDOX	H V.V.ICS.RN	PEMINCMYLE	AI..MQT.PN	YATLFRGRP.		387
H-RPE65	DN.F	LIV.LCCWKG	FEFV.NYL.-	-----REN	WE.VKK.A.K	359
H-BCDOX	A TSVPTLRRFA	B VPLHVDKNAE	VGTNLIKVAS	TTATALKEED	QGVYQCPEFL	400
M-BCDOXD..	..S..V..S.K.	.H.....V.		400
C-BCDOX	...I..CK..V	...QY..D..	..S..V.L-P	.S...V..K.	.SI.....I.	398
D-BCDOX	RF.LP.GTIP	PASIAKRLV	KSFS.AGLSA	PQVSRTMKHS	VSQ.ADIITYM	437
H-RPE65	APQ.EV..YV	L..NI...-D	T.K..VTLPNILCS.	ETIWLE..V.	408
H-BCDOX	-YEG----L-	-ELPRVNY--	A-HNGKQYRY	VFATGVQWSP	IPTKIIKYDI	440
M-BCDOX	- - - - -	-...T...--	-Y...P...	I..AE.....	V....L....	440
C-BCDOX	-C...--I-	...V...--	D-Y...K.K.	.Y.TE.....	V....A.LNV	438
D-BCDOX	PTN.KQATAG	E.S.KRDAKR	GRYEEENLVN	LVTMEGSQAE	AFQGTNGIQL	487
H-RPE65	-FS.PRQAF.	-F.QI...--	QKYC..P.T.	AYGL.LNHF-	V.DRLC.LNV	452
H-BCDOX	LTKSSSLKWRE	D DDCWPAELF	E VHAP---GA	F KDEEDVILS	AIVSTDP-QK	485
M-BCDOXS.	ES.....	..T....	485
C-BCDOX	Q EV.H.G.	.H.S..I..S.	D..S..E..V.T	CV.VSE.-N.		483
D-BCDOX	RPEMLCD.GC	ETPRIYKMK	MCKNYRYFY	ESS.VDAVNP	GTLIKVDVWN	537
H-RPE65	K..ETWV.Q.	P.SY.S..I..	.SH....-D	..V..	VV..PGAG..	498
H-BCDOX	LPFLILIDAK	SPTELARASV	DVDMHMDLHG	LFITDMWDWT	KKQ--AASE-	532
M-BCDOXA...L....	...P.A..NA	V...-TPA.T	533
C-BCDOX	A.....	T.K..G..T.	N.E..L....	M..PQN.LGA	ETE-----	526
D-BCDOX	KSC.TWCEEN	VYPSEIFVP	SP.PKSEDD.	VILAS.VLGG	LNDRYVGLIV	587
H-RPE65	PAY....N.	DLS.V...E.	EINIPVTF..	-----	..S-----	533
H-BCDOX	---EQDRAS	D-----	-----CHGA	PLT		547
M-BCDOX	QEV.NS.HPT	.PTAPELSHS	ENDFTAG..G	SSL		566

Figure 1-4B. Sequence comparison of human BCDOX (H-BCDOX), mouse BCDOX (M-BCDOX), chicken BCDOX (C-BCDOX), and *Drosophila* BCDOX (D-BCDOX), and human RPE65. The most conserved regions are shaded and labeled A-D. Part II. These homology was generated by Zack and colleagues using GeneWorks 2.5.1 (Oxford Scientific).²

of lungs and kidneys. But its activity has also been found in intestines, brains, stomachs and testes. Comparison of the deduced amino acid sequences revealed that humans, chickens, mice and the *Drosophila* enzymes share overall sequence homology with a distinct pattern of highly identical conserved stretches (Figure 1-4a and 1-4b).

Most of the highly conserved regions are located in the terminal section of the protein.^{47,49,53,54} A sequence alignment performed by Yan and colleagues between BCDOX from different organisms and human RPE65 reveals four areas of increased sequence conservation. It is possible that those regions could belong to the binding site of β -carotene.² It is also possible that these conserved regions are associated to the binding site of the iron cation. In particular, region A contains a very conserved arginine, a couple of histidines and a carboxylate that could form the binding site for the iron cation. In contrast to the vertebrate enzyme, the insect enzyme possesses a long extension close to the C-terminus. The role of this extension is not clear. Also, it is interesting to notice that alignment shows that around 8% of the highly hydrophobic amino acids are conserved in the enzyme, suggesting that these amino acids participate in the formation of the substrate-binding pocket.⁵¹ Besides the homology amongst each other, the various BCDOX proteins share significant sequence homology to RPE65. Whether the homology between BCDOX and RPE65 is due to a conserved carotenoid/retinoid-binding motif or because of a related biochemical function remains to be investigated.

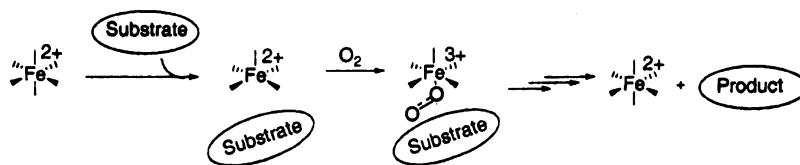
Interestingly, a homology analysis of human, mouse, chicken and fly BCDOX and the RPE proteins, performed by Cunningham and co workers, shows that ten residues are absolutely conserved, 4 histidines and 6 acidic residues. It is also possible that those 4 histidines are involved in the binding of the iron cation. Also it has been suggested that a conserved region (consensus sequence EDDGVVLSXVVS) at residues 469-480 of mouse BCDOX, might be considered a family signature sequence.⁵⁵ A detail examination of the different domains in the protein does not provide any obvious evidence for a transmembrane domain, or any potential site for N-linked glycosilation. This supports the observed results that BCDOX is a cytosolic protein. However, it does not eliminate the possibility that it might be a tetramer or a lipid associated protein.

1.4 Mechanism of action

15-15' β -Carotene Dioxygenase belongs to the enzyme family of oxygenases. Oxygenases are enzymes that catalyze the incorporation of oxygen from molecular oxygen (O_2) into an organic substrate. There are two classes of oxygenases: dioxygenases in which both oxygen atoms from molecular oxygen (O_2) are incorporated into the substrate(s), and monooxygenases in which one oxygen atom from molecular oxygen (O_2) is incorporated into the substrate while the other is reduced to water. As previously mentioned, it has been observed that BCDOX is an iron dependant enzyme. Oxygenases commonly use iron (Fe) as a cofactor to activate the molecular oxygen. The iron in oxygenases could be part of a heme group or a non-heme form binding to some residues in the protein. 15-

15' β -Carotene Dioxygenase is a nonheme iron-containing enzyme. In most of the non-heme Fe^{II} dependant oxygenases the iron has a five or six coordination sphere that binds to the molecular oxygen.

Oxidation by dioxygenases first requires the activation of molecular oxygen. However, to date, the precise mode for the activation of molecular oxygen has eluded researchers. Up to now, most researchers agree that first; Fe^{II} exists in a six coordinate state when it is relatively unreactive toward O_2 . When the substrate binds, one of the ligands dissociates from Fe^{II} , generating a 5 coordinate complex. O_2 binds to this complex generating a highly reactive iron-oxygen intermediate, which could oxidize the substrate; (Scheme 1-4).^{56,57}



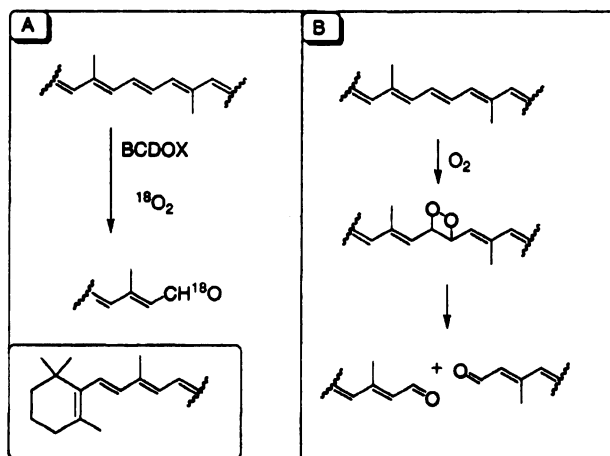
Scheme 1-4. General mechanism for activation of O_2 in dioxygenases.

The ligands present in non-heme enzymes so far identified are tyrosine, histidine and glutamate.^{57,58} It is known that BCDOX is a non-heme Fe^{II} oxidase. However, identification of the amino acids that bind to Fe^{II} has not been accomplished. Therefore, the mechanism of action remains unknown.

Oxidative cleavage of β -carotene is an unprecedented process for a dioxygenase enzyme. Typically, dioxygenases require a hydrophilic handle such as a hydroxyl group to proceed.⁵⁹ Enzyme/substrate interactions that lead to stereospecific catalytic action through hydrogen bonding and dipolar

communication are common. However, the enzymatic mechanism of BCDOX is intriguing in that it uses a molecule of oxygen to cleave an electronically undistinguished double bond with a high degree of selectivity.

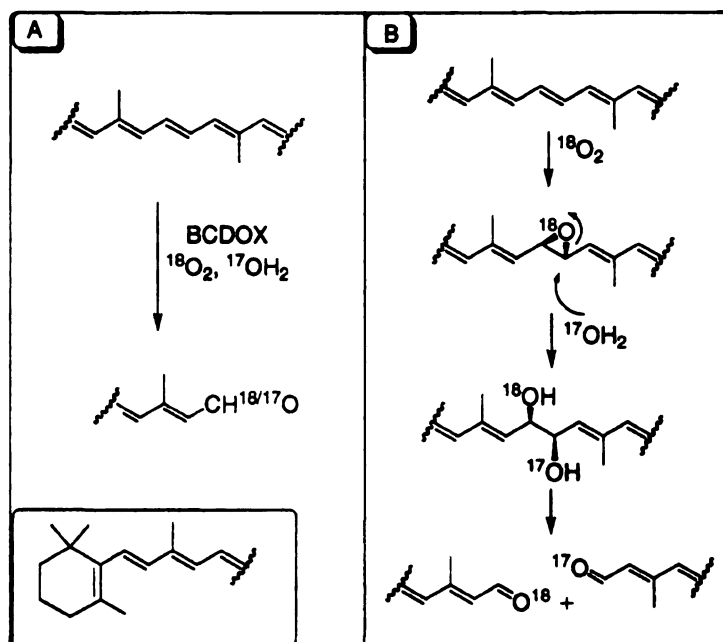
It was not until after 40 years of the discovery of carotene dioxygenase in 1965, that Olson and Hayaishi demonstrated that molecular oxygen is incorporated into vitamin A. When $^{18}\text{O}_2$ gas was used, a retinal adduct that contained an M+2 mass peak was observed, but not when H_2^{18}O was used. Since then the β -carotene oxygenase was believed to be a dioxygenase (Scheme 1-5A). Later, Goodman postulated a mechanism that suggested the intermediacy of a dioxetane intermediate. This mechanism supposes a 2+2 cycloaddition of molecular oxygen to the central double bond of β -carotene, followed by a retro



Scheme 1-5. Carotene dioxygenase mechanism. A) Both atoms from molecular oxygen ($^{18}\text{O}_2$) are added to retinal. B) Proposed mechanism of oxidation, dioxetane intermediate.

2+2 cleavage of the carbon-carbon bond (Scheme 1-5B). Although this could be a viable mechanism, cleavage of inactivated olefins to aldehydes by dioxygenases is unprecedented. Also the postulated dioxetane is a high energy intermediate and its decomposition is highly exothermic (~ 70 kcal/mol),⁶⁰ followed by chemiluminescence. This is the result of producing a triplet state oxygen, which upon relaxation emits light. This is the proposed mechanism for bioluminescence for the firefly luciferase reaction.⁶⁰

For another 40 years the dioxygenase mechanism was accepted as the true mode of action for β -carotene oxygenase. It was not until 2001 that Woggon and co-workers challenged the oxidation mechanism. Woggon provided evidence for the incorporation of one ^{18}O atom of molecular oxygen and the incorporation of one ^{17}O atom from labeled water (Scheme 1-6A). Accordingly, and in contrast to earlier findings, the reaction mechanism of enzymatic central β -carotene cleavage does not agree with a dioxygenase catalyzed mechanism. The proposed mechanism by Goodman would require the incorporation of one complete oxygen molecule into the product aldehydes and no oxygen originated from labeled water should be present (Scheme 1-6B). The discrepancy could be due to the different conditions used in the experiments or the different origins of the BCDOX used. Olson and Hayaishi used crude enzyme preparations from rat liver and rat intestine.²⁹ Woggon and colleagues used baby hamster kidney cells.⁶¹ Woggon and colleagues proposed an oxidation mechanism that first involves the epoxidation of the 15-15' double bond; they suggest that the oxygen from this epoxidation comes from molecular oxygen. As a second step, the epoxide formed

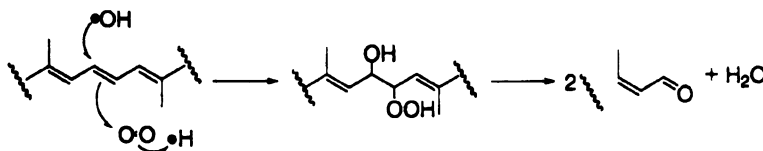


Scheme 1-6. Monooxygenase mechanism. A) One atom from molecular oxygen is added to one retinal ($^{18}\text{O}_2$) and one molecule of water is added to retinal H_2^{17}O . B) Proposed mechanism of oxidation, epoxidation by molecular oxygen follow by addition of a molecule of water.

is attacked by a molecule of water to afford a diol; this diol could fragment to afford two molecules of retinal.

Alternate mechanisms for oxidation of β -carotene can be envisioned in which two molecules of oxygen are used to cleave the olefin. The presence of Fe^{2+} can lead to the generation of hydroperoxy radicals that can attack the central olefinic bond, and via capture of another oxygen molecule, yield a vicinal bis-hydroperoxy intermediate. This intermediate could undergo fragmentation to

produce retinal. Similarly, attack of a hydroxyl radical and capture of an oxygen molecule can result in a hydroxy-hydroperoxy intermediate as shown in Scheme 1-7.



Scheme 1-7. Alternate mechanisms. A hydroperoxy radical could be formed, and attack the central olefinic bond. Capture of another oxygen molecule can yield a vicinal bis-hydroperoxy intermediate. This intermediate could undergo fragmentation to produce retinal. Similarly, attack of a hydroxyl radical and capture of an oxygen molecule can result in a hydroxy-hydroperoxy intermediate.

Woggon and co-workers synthesized a few substrate analogues of β -carotene.¹ When a single bond was replaced at the 15-15'C, no activity was observed, but a slow oxidation of (b) to β -carotene (a) was observed, followed by oxidative cleavage to retinal (15%). A change in the configuration of the 15-15'C=C bond in β -carotene (*trans* to *cis*) or substitution to a single C-C or a C \equiv C bond yields no product at all. Thus it was concluded that any deviation from the 'rod like' β -carotene structure is not tolerated (Figure 1-5 d, e and f). When the positions of the methyl groups were modified (Figure 1-5 a, b and c) a significant change in the activity was observed. The position and presence of methyl groups attached to the polyene chain is significant. Also it was evident that at least one

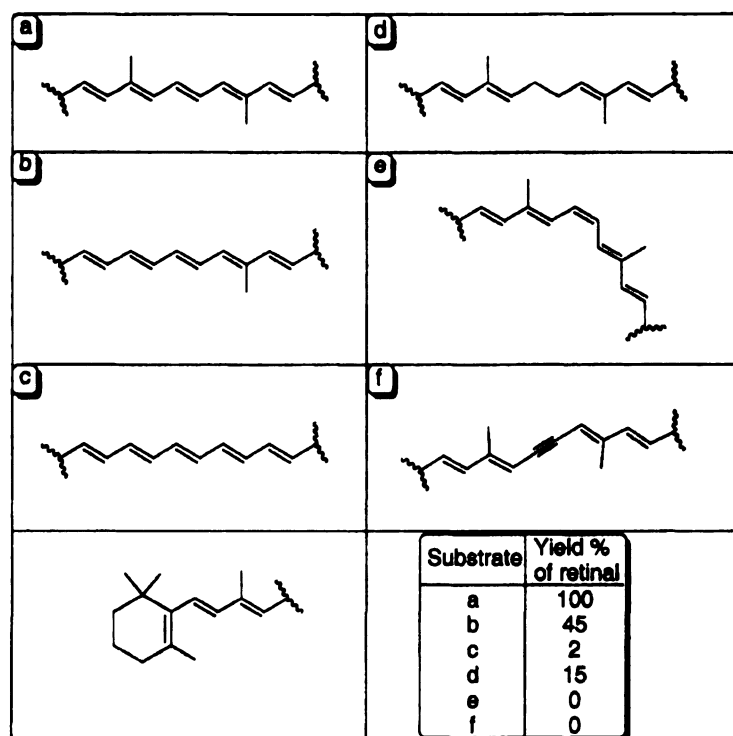


Figure 1-5. Different substrates used to evaluate activity in chicken BCDOX.

a-c were used to evaluate the importance of methyl groups. Variation of the geometry in 15-15'C d-f was used to evaluate the effect in function of BCDOX.¹

'natural', unsubstituted β -ionone ring was required for the oxidative cleavage to take place. Woggon suggested a hydrophobic barrel-like substrate-binding site, in which the protein's amino acid residues interact with the methyl groups to direct the central double bond (15-15') to the correct place for optimal distance to the active site metal-oxo center, thus supporting the unique observed regiospecificity in the cleavage of β -carotene to retinal.

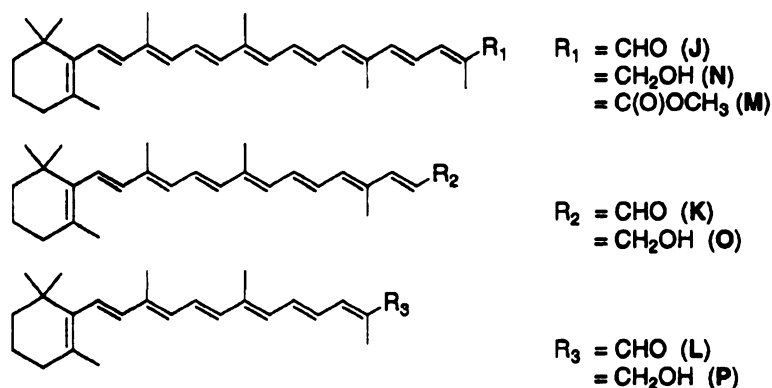
Cama and Sing performed a number of substrate specificity experiments with BCDOX using different apo-carotenals and carotenoids.⁶² They found that

the BCDOX from pig and rabbit showed remarkable similarity with respect to cleavage of carotenoids. With respect to apo-carotenoids, it was found that 10'-apo-carotenoid alcohol and aldehyde forms are more readily cleaved, as compared with other apo-carotenals tested (Table 1-1). However, all of these experiments were performed using purified protein from intestines that was not completely pure.

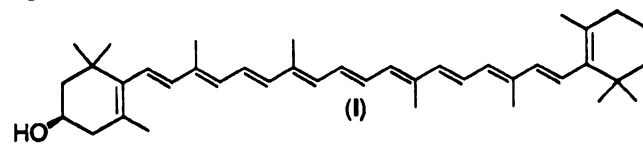
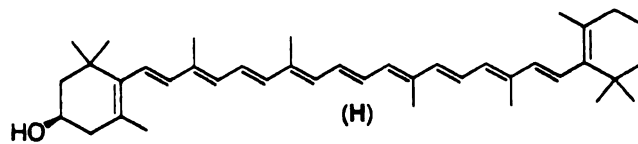
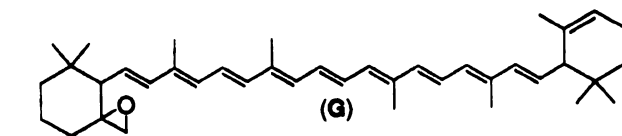
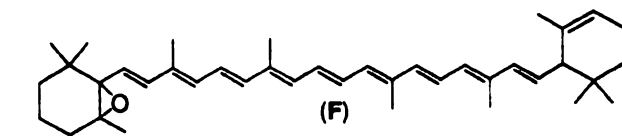
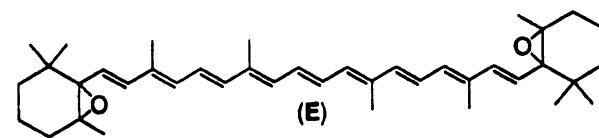
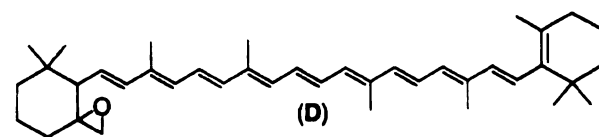
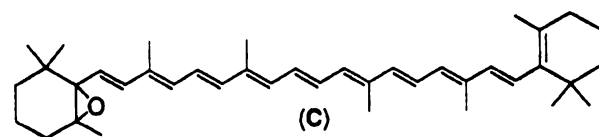
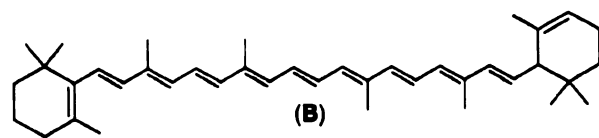
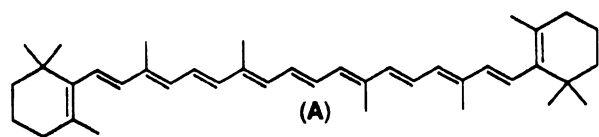
Even though in the last decade a large amount of information has been gathered for BCDOX, many questions remain unanswered in regard to its mechanism of action. More specifically, it is not understood how the enzyme directs the oxidation of an undistinguished double bond. In addition, further proof

Table 1-1. Different substrates tested for BCDOX activity.

Substrate	Guinea pig relative activity	Rabbit relative activity
A β - β -carotene	1.00	1.00
B β - ϵ -carotene	0.72	0.56
C 5,6-epoxy-5,6-dihydro- β - β -carotene	0.32	0.25
D 5,8-epoxy-5,8-dihydro- β - β -carotene	0.19	0.11
E 5,8,5',8'-diepoxy-5,8,5'8'-tetrahydro- β - β -carotene	0.02	0.02
F 5,6-epoxy-5,6-dihydro- β - ϵ -carotene	0.25	0.22
G 5,8-epoxy-5,8-dihydro- β - ϵ -carotene	0.09	0.09
H β - β -carotene-3'-ol	0.04	0.03
I 3',4'-didehydro- β - β -carotene-3-ol	0.03	0.02
J 8'-apo- β -carotenal	0.09	0.03
K 10'-apo- β -carotenal	0.57	0.55
L 12'-apo- β -carotenal	0.12	0.07
M 8'-apo- β -caroten-8'-yl acetate	0.16	0.10
N 8'-apo- β -carotenol	0.07	0.02
O 10'-apo- β -carotenol	0.50	0.43
P 12'-apo- β -carotenol	0.08	0.07



of whether the cleavage of β -carotene occurs via a monooxygenase or dioxygenase mechanism is needed. Because of the importance of β -carotene as the main source of retinoids *in vivo*, and because the oxidation of β -carotene is such a unique example for regiospecific oxidation, we were interested in understanding the mode of control for such oxidation, as well as how such oxidation occurs for an inactivated olefin. Finally we were interested in discovering whether the oxidation mechanism occurred via a dioxygenase or monooxygenase mechanism.



1.5 References

1. G. M. Wirtz; C. Bornemann; A. Giger; R. K. Muller; H. Scheinder; G. Schlotterbeck; G. Schiefer; W. Woggon, "The substrate specificity of beta carotene 15-15'monooxygenase." *Helvetica Chimica Acta* **2001**, *84*, 2301-2315.
2. W. Yan; G.-F. Jang; F. Haeseleer; N. Esumi; J. Chang; K. Palcewski; D. Zack, "Cloning and characterization of a human beta-carotene." *Genomics* **2001**, *72*, 193-202.
3. G. A. Armstrong, "Genetics of eubacterial carotenoid biosynthesis: A colorful tale." *Annual Review of Microbiology* **1997**, *51*, 629-659.
4. G. E. Bartley; P. A. Scolnik, "Plant carotenoids: Pigments for photoprotection, visual attraction, and human health." *Plant Cell* **1995**, *7*, 1027-1038.
5. R. A. Morton, *Photochemistry of Vision*. ed.; Springer-Verlag: New York, 1972; 'Vol.' p 33.
6. P. Palozza, "Regulation of cell cycle progression and apoptosis by beta-carotene in undifferentiated and differentiated HL-60 leukemia cells: Possible involvement of a redox mechanism." *International Journal of Cancer* **2002**, *97*, 593-600.
7. L. J. Gudas, *In The retinoids: Biology chemistry and medicine*. ed.; New York, 1994; 'Vol.' p 443-520.
8. A. C. Ross, "Cellular-metabolism and activation of retinoids-roles of cellular retinoid-binding proteins." *FASEB Journal* **1993**, *7*, (2), 317-327.
9. C. Merlet-Benichou; J. Vilar; M. Lelievre-Pegorier; T. Gilbert, "Role of retinoids in renal development: pathophysiological implication." *Current Opinion in Nephrology and Hypertension* **1999**, *8*, (1), 39-43.

10. T. E. Gundersen; R. Blomhoff, "Qualitative and quantitative liquid chromatographic determination of natural retinoids in biological samples." *Journal of Chromatography A* **2001**, 935, 13-43.
11. L. Stryer, "Vision: From photon to perception." *Proceedings of the National Academy of Sciences of the United States of America* **1996**, 93, (2), 557-559.
12. D. J. Mangelsdor; R. M. Evans, "The RXR heterodimers and orphan receptors." *Cell* **1995**, 83, (6), 841-850.
13. D. E. Ong, "A novel retinol-binding protein from rat." *The Journal of Biological Chemistry* **1984**, 259, (10), 1476-1482.
14. J. E. Smith; D. S. Goodman, "Retinol-binding protein and the regulation of vitamin A transport." *Federal Proceedings* **1979**, 38, 2504-2509.
15. S. Futterman, and Saari, J. C., "Occurrence of 11-cis-retinal-binding protein restricted to the retina." *Investigative Ophthalmology and Visual Science* **1977**, 16, 768-771.
16. Y. L. W. Lai, B., Liu, Y.P., and Chader, G. J., "Interphotoreceptor retinol-binding proteins: possible transport vehicles between compartments of the retina." *Nature* **1982**, (298), 848-859.
17. J. A. Olson; N. I. Krinsky, "Introduction: the colorful, fascinating world of the carotenoids: important physiologic modulators." *FASEB Journal* **1995**, 9, 1547-1550.
18. R. S. Parker, "Absorption, metabolism, and transport of carotenoids." *FASEB Journal* **1996**, 10, 542-551.
19. A. A. Dmitrovskii, *Metabolism of natural retinoids and their functions. In : New trends in Biological Chemistry.* ed.; Japan Science Society: Tokyo, Japan, 1991; 'Vol.' p 297-308.
20. J. A. Olson, "Some aspects of vitamin A metabolism." *Vitamines and Hormones* **1968**, 26, 1-63.

21. G. A. J. Pitt, *Carotenoids*. ed.; Isler, O: Birkhauser Verlag, Basel, 1971; 'Vol.' p 717-742.
22. T. R. Evans; S. B. Kaye, "Retinoids: present role and future potential." *Cancer* **1999**, (80), 1-8.
23. S. A. Kliewer; J. M. Lehmann; T. M. Willson, "Orphan nuclear receptors: Shifting endocrinology into reverse." *Science* **1999**, *284*, 757-760.
24. G. M. W. Morris-Kay, S. J., "Retinoids and mammalian development." *International Reviews in Cytology* **1999**, *188*, 73-131.
25. M. D. M. Collins, G. E, "Teratology of retinoids." *Annual Reviews in Pharmacology and Toxicology* **1999**, *39*, 399-430.
26. A. B. Goodman, "Three independent lines of evidence suggest retinoids as causal to schizophrenia." *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 7240-7244.
27. D. Hollander, Ruble, P.E. Jr, "beta-Carotene intestinal absorption: bile, fatty acid, pH, and flow rate effects on transport." *American Journal of Physiology* **1978**, *235*, E686-E691.
28. T. Moore, "Vitamin A and carotene." *Biochemical Journal* **1930**, *24*, 692-702.
29. J. A. Olson; O. Hayaishi, "The enzymatic cleavage of beta-carotene into vitamin A by soluble enzymes of rat liver and intestine." *Proceedings of the National Academy of Sciences of the United States of America* **1965**, *54*, 1364-1370.
30. D. S. Goodman; H. S. Huang, "Biosynthesis of vitamin A with rat intestinal enzymes." *Science* **1965**, *149*, 879-880.
31. A. Nagao; A. During; C. Hoshino; J. Terao; J. A. Olson, "Stoichiometric conversion of all trans-beta-carotene to retinal by pig intestinal extract." *Archives of Biochemistry and Biophysics* **1996**, *328*, (1), 57-63.

32. J. Devery; B. V. Milborrow, "beta-Carotene-15,15'-dioxygenase (Ec-1.13.11.21) isolation reaction-mechanism and an improved assay procedure." *British Journal of Nutrition* **1994**, 72, (3), 397-414.
33. J. Glover, "The conversion of beta-carotene into vitamin A." *Vitamines and Hormones* **1960**, 18, 371-386.
34. M. R. Lakshman; I. Mychkovsky; M. Attlesey, "Enzymatic conversion of all-trans-beta-carotene to retinal by a cytosolic enzyme from rabbit and rat intestinal mucosa." *Proceedings of the National Academy of Sciences of the United States of America* **1989**, 86, 9124-9128.
35. K. Wang, "The bioconversion of beta-carotene into retinoids." *Subcellar Biochemistry* **1998**, 30, 159-180.
36. T. V. Vliet; R. V. Schaik; H. V. D. Berg; W. H. P. Schreurs, "Effect of vitamin A and beta-carotene intake on dioxygenase activity in rat intestine." *Annals of the New York Academy of Sciences* **1993**, 691, 220-222.
37. A. Yang; R. K. Tume, "Comparison of beta-carotene-splitting activity isolated from intestinal mucosa of pasture-grazed sheep, goats and cattle." *Biochemistry. Molecular Biology International* **1993**, 30, 209-217.
38. G. Scita; G. W. Aponte; G. Wolf, "Uptake and cleavage of beta-carotene by cultures of rat small intestinal cells and human lung fibroblasts." *Journal of Nutritional Biochemistry* **1992**, 3, 118-123.
39. J. Olson, *In Modern Nutrition in Health and Disease*. 9th ed.; Williams and Wilkins: Baltimore, MD, 1999; 'Vol.' p 193-221.
40. J. A. Olson, "The effects of iron and copper status and of dietary carbohydrates on the activity of rat intestinal beta-carotene 15,15' dioxygenase." *British Journal of Nutrition* **2000**, 84, 3-4.
41. S. H. Schwartz; B. C. Tan; D. A. Gage; J. A. D. Zeevaart; D. R. McCarty, "Specific oxidative cleavage of carotenoids by VP14 of maize." *Science* **1997**, 276, (5320), 1872-1874.

42. S. H. Schwartz; X. Qin; J. A. D. Zeevaart, "Characterization of a novel carotenoid cleavage dioxygenase from plants." *The Journal of Biological Chemistry* **2001**, 276, (27), 25208-25211.
43. S. Kamoda; Y. Saburi, "Structural and enzymatical comparison of lignostilbene-alpha,beta-dioxygenase isozymes, I, II, and III, from *Pseudomonas paucimobilis* TMY1009." *Bioscience Biotechnology and Biochemistry* **1993**, 57, 931-934.
44. P. Winterhalter; M. Straubinger, "Saffron-renewed interest in an ancient spice." *Food Review. International* **2000**, 16, 39-59.
45. C. Kiefer; S. Hessel; J. M. Lampert; K. Vogt; M. O. Lederer; D. E. Breithaupt; J. von Lintig, "Identification and characterization of a mammalian enzyme catalyzing the asymmetric oxidative cleavage of provitamin A." *Journal of Biological Chemistry* **2001**, 276, (17), 14110-14116.
46. A. A. Dmitrovskii; G. N. N; S. B. Gomboeva; V. Ershov; V. Bykhovsky, "Enzymatic oxidation of beta apo-8'-carotenol to beta apo-14'-carotenal by an enzyme different from beta carotene 15-15' dioxygenase." *Biochemistry* **1997**, 62, (7), 917-923.
47. J. von Lintig; A. Wyss, "Molecular analysis of vitamin a formation: Cloning and characterization of beta-carotene 15,15 '-dioxygenases." *Archives of Biochemistry and Biophysics* **2001**, 385, (1), 47-52.
48. A. Wyss; G. Wirtz; W. D. Woggon; R. Brugger; M. Wyss; A. Friedlein; H. Bachmann; W. Hunziker, "Cloning and expression of beta,beta-carotene 15,15 '-dioxygenase." *Biochemical and Biophysical Research Communications* **2000**, 271, (2), 334-336.
49. A. Wyss; G. M. Wirtz; W. D. Woggon; R. Brugger; M. Wyss; A. Friedlein; G. Riss; H. Bachmann; W. Hunziker, "Expression pattern and localization of beta,beta-carotene 15,15 '-dioxygenase in different tissues." *Biochemical Journal* **2001**, 354, 521-529.
50. J. von Lintig; K. Vogt, "Filling the gap in vitamin A research - Molecular identification of an enzyme cleaving beta-carotene to retinal." *Journal of Biological Chemistry* **2000**, 275, (16), 11915-11920.

51. A. Lindqvist; S. Anderson, "Biochemical properties of purified recombinant human beta carotene 15-15'-monooxygenase." *Journal of Biological Chemistry* **2002**, 277, (26), 23942-23948.
52. D. Sklan, "Carotene-cleavage activity in chick intestinal mucosa cytosol: association with a high-molecular-weight lipid-protein aggregate fraction and partial characterization of the activity." *British Journal of Nutrition* **1983**, 50, 417-425.
53. J. Paik; A. During; E. H. Jarrison; C. L. Mendelsohn; K. Lai; W. Blaner, "Expression and characterization of a murine enzyme able to cleave beta carotene." *The Journal of biological chemistry* **2001**, 276, 32160-32168.
54. von Lintig; A. Dreher; C. Kiefer; M. F. Wernet; K. Vogt, "Analysis of the blind *drosophila* mutant ninaB identifies the gene encoding the key enzyme for vitamin A formation in vivo." *Proceedings of the National Academy of Sciences of the United States of America* **2001**, 98, 1130-1135.
55. M. T. Redmond; S. Gentleman; T. Duncan; S. Yu; B. Wiggert; E. Gantt; S. Cuninghame, "Identification, expression and substrate specificity of a mammalian beta-carotene 15-15'-dioxygenase." *Journal of Biological Chemistry* **2001**, 276, (9), 6560-6565.
56. F. Takuzo, *Oxygenases and model systems*. ed.; Kluwer Academic Publishers: Norwell, Ma, 1997; 'Vol.' p.
57. E. I. Solomon; T. C. Brunold; M. I. Davis; J. N. Kemsley; S. K. Lee; N. Lehnert; F. Neese; A. J. Skulan; Y. S. Yang; J. Zhou, "Geometric and electronic structure/function correlations in non-heme iron enzymes." *Chemical Reviews* **2000**, 100, (1), 235-349.
58. L. Que; R. Y. N. Ho, "Dioxygen activation by enzymes with mononuclear non-heme iron active sites." *Journal of Chemical Reviews* **1996**, 96, (7), 2607-2624.
59. E. I. Solomon; T. C. Brunold; M. I. Davis; J. N. Kemsley; L. Sang-Kyu; N. Lehnert; F. Neese; A. J. Skulan; Y. Yang; J. Zhou, "Geometric and electronic structure/function correlations in non-heme iron enzymes." *Journal Chemical Reviews* **2000**, 100, 235-349.

60. T. Wilson, "Comments on the mechanism of chemi- and bioluminescence." *Photochemistry and Photobiology* **1995**, 62, 601-606.
61. M. G. Leuenberger; C. Engeloch-Jarret; W.-D. Woggon, "The reaction mechanism of the enzyme catalyzed central cleavage of beta-carotene to retinal." *Angewandte Chemie International Edition* **2001**, 40, 2614-2617.
62. H. Singh; H. R. Cama, "Enzymatic cleavage of carotenoids." *Biochimica et Biophysica Acta* **1974**, 370, 49-61.

Chapter 2

Expression of BCDOX

2.1 Expression of BCDOX using a T5 promoter

The usefulness of retinoids in the visual process has been well understood for many decades.¹⁻³ Also, their vital role in cellular and genetic control has been well documented recently.⁴⁻⁷ However, what is not well understood is the biogenesis of retinoids in mammals. Retinoids are biosynthesized by oxidative cleavage of β -carotene by β -carotene-15-15' dioxygenase (BCDOX) (EC 1.13.11.21) to yield retinal.⁸⁻¹³ And as discussed previously, we are interested in understanding the mechanism and function of BCDOX. The first step towards the understanding of BCDOX function, regulation and mechanism would be the expression of the protein in large quantities.

Professor Lintig (University of Freiburg) graciously provided us with vector (pQE30) that contained the gene for BCDOX from mice (AJ278064).¹⁴ The expression of this BCDOX gene is under the control of a Phage T5 promoter transcription-translation system, which is recognized by the *E. coli* Polymerase. This system contains two *lac* operator sequences, which increase *lac*

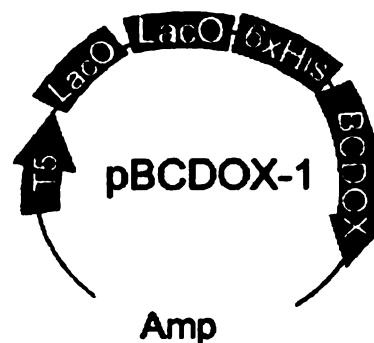


Figure 2-1. Gene of mice β -carotene-15-15' dioxygenase cloned in a pQE30 vector (pBCDOX-1).

repressor binding and ensure efficient repression. Also, in the N terminal sequence of the protein a His tag (6 His in a row) is included to facilitate BCDOX purification (Figure 2-1). Expression of mice BCDOX has been previously described¹⁵ in various systems and species such as *drosophila*,^{13,16} chicken,¹¹ and human^{17,18} but no report has been made using a pQE30 vector. Qiagen¹⁹ suggests the use of either JM109 or XL1Blue *E. coli* host to overexpress proteins cloned into pQE30 plasmids.²⁰ Thus pQE30-mice-BCDOX (pBCDOX-1) was transformed into both XL1Blue and JM109. Initial attempts to express pBCDOX-1 using different IPTG concentrations and different temperatures proved unsuccessful. To accelerate the screening of optimal conditions to express BCDOX, the supernatant was analyzed using Anti-His HRP-Conjugate. Anti-His HRP-Conjugate allows for the identification of recombinant proteins that contained a 6x Histidine tag. This Anti-His HRP Conjugate²¹ consists of a mouse monoclonal IgG Anti-His antibody coupled to horseradish peroxidase. In the presence of peroxide and 4-chloro-1-naphthol, the horseradish peroxidase produces a colored precipitate.²²⁻²⁴ Initially four different *E. coli* strains (XL1-Blue, JM109,

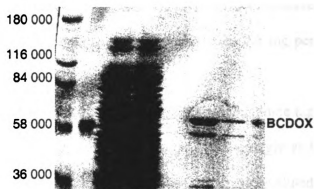


Figure 2-2. Purification of recombinant β -carotene-15-15'-dioxygenase. Purification was performed by the use of an NTA resin. 1. Marker. 2. BSA (60KDa). 3. Supernatant. 4. Run throw. 5. Wash. 6, 7. Eluted BCDOX.

Tuner(DE3)pLacI, and BL21(DE3)pLysS), and 1 mM IPTG were used to induce protein expression at 28 °C for three or six hours. But again no recombinant protein was found either in the soluble fraction or the pellet. When different amounts of IPTG were used to express the recombinant protein, no improvement in the expression level of the protein was observed. It is known that low or no expressions of proteins occurs when the expressed protein is toxic to the bacteria.

To address this problem, the most common approach is the use of low concentrations of IPTG for the expression of the protein. But when this fails, the addition of glucose to the media is recommended to increase plasmid stability.²⁵ When glucose is available as an energy source, it is used preferentially over other sugars. In more detail, glucose causes catabolite repression by reducing the level of cyclic AMP. The protein responsible for the catabolite activation (CAP) is active only in the presence of cyclic AMP. The presence of glucose in the media decreases the basal noise, increasing the plasmid stability. When 1% of glucose was used, the recombinant protein was obtained in low yield, about 0.7 mg per liter of culture.

Purification of the enzyme was achieved by the use of the 6x histidine tag present in the protein. Proteins that contain a histidine tag bind strongly and selectively to a NTA resin (Contains Nickel, resin form Qiagen) and can be eluted by the addition of imidazole or a change in the pH.

2.2 β -Carotene Dioxygenase activity

The activity of BCDOX was analyzed using an assay previously developed by Olson.²⁶ This assay consists of monitoring the oxidation of β -carotene to retinal.

In particular, about 20 μ g of BCDOX (purified recombinant protein) was incubated with β -carotene at 30 °C (For a more detailed description refer to Materials and Methods). The reaction was quenched at different times. Next, the carotenoids and retinoids were extracted with hexanes, and were analyzed by High Pressure Liquid Chromatography (HPLC), (Figure 2-3). A normal phase column (C18) was used as stationary phase, and hexane: ethyl acetate (95:5) as a mobile phase at a flow rate of 1 mL/min.

The enzyme was found to have an activity of 25.5 pmol retinal/mg BCDOX•min (Scheme 2-1). It has been reported that addition of cofactors to the growth media promotes correct folding of proteins, and therefore, increases the enzymatic activity.²⁷ Thus addition of FeSO_4 to the LB media was tried, and as reported, a 7x increase in the BCDOX activity (194 pmol retinal/mg BCDOX•min) and a $K_M = 3 \mu\text{M}$ was observed. However, the yield of the protein obtained did not change much (Table 2-1). These values slightly differ from the reported value (36 pmol retinal/mg BCDOX•min and a $K_M = 6 \mu\text{M}$).¹⁵

The activity of BCDOX observed seems to be relatively low compared to other enzymes. So far it is not clear whether the enzyme activity is low due to lack of correct folding. However, since the activity reported from other groups is

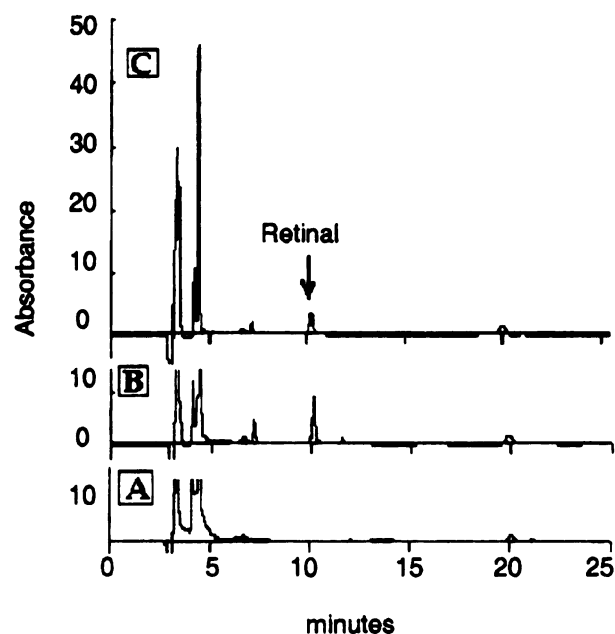
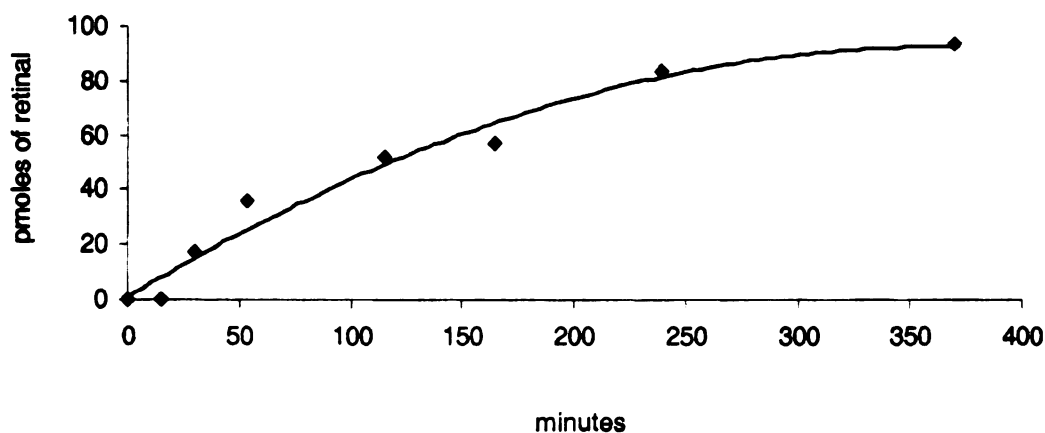


Figure 2-3. HPLC chromatographs of enzymatic reaction using BCDOX and β -carotene. The formation of retinal was monitored in an HPLC. A) Without enzyme, B) Assay using BCDOX isolated from a culture with out FeSO_4 25.5 pmol retinal /mg of BCDOX•min. C) Assay using BCDOX isolated from a culture with FeSO_4 194 pmol retinal /mg of BCDOX•min.

also low, this led us to believe that BCDOX rate of conversion of β -carotene to retinal is slow. It has been shown that BCDOX is an important source of retinoic acid *in vitro* in testis as well as in small intestine, liver, kidney and lung.²⁸ Also, the concentration of retinoic acid in the cytoplasm is very tightly controlled, thus it is possible that BCDOX is a check point in the biosynthesis of retinoids. And it is possible that retinoids work as allosteric inhibitors to control their own biosynthesis. Also, even though there is an overwhelming agreement that



Scheme 2-1. Kinetics of the recombinant β -carotene 15-15' dioxygenase. Monitoring the conversion of β -carotene to retinal.

BCDOX is a cytosolic protein, it is important to keep in mind the lipophilicity of β -carotene. Carotenoids and their derivatives are non-polar molecules with the tendency to participate in membranes and are generally insoluble in aqueous solutions. Thus, it is possible that for optimal activity of BCDOX, various lipid

Table 2-1. Comparison of protein activity shows that glucose addition is essential for BCDOX expression and addition of iron results in a higher protein activity.

<i>E. coli</i>	Protein expression	
	μg of protein per 1L	BCDOX activity pmol retinal/ mg BCDOX \cdot min
XL1-Blue-pBCDOX-1	0	0
XL1-Blue-pBCDOX-1+ glucose	736	25.5
XL1-Blue-pBCDOX-1+glucose+ FeSO_4	665	194

aggregates should be present, and recombinant protein expression does not provide the optimal conditions for oxidation of β -carotene.

Unfortunately it was found that the protein is not very stable. After protein purification, if the protein is frozen it loses activity after 48 hours. When 10% of glycerol is added to the storing buffer, it loses 63% of its activity after one week. When the recombinant protein is exposed to room temperature for more than three hours the activity decreases significantly. Instability of recombinant BCDOX has been observed by other groups. For example, Wyss and co workers observed that during a 50-fold increase in the concentration of the protein 75% of activity is lost, which they believed is due to aggregation.¹⁰

2.3 Expression of BCDOX with rare codons

The fact that all the previous efforts to improve the expression of BCDOX did not have any significant impact in the protein led us to believe that there was something else interfering with the expression. One reason for the low yield could be an insufficient tRNA pool for the codon usage in the BCDOX gene.^{29,30} Low tRNA can lead to translational stalling, premature translation termination, translation frame shifting and amino acid misincorporation.^{31,32} Evidence that this could be the reason was supported by the fact that every time BCDOX was expressed and purified, four bands co-eluted with the elution buffer. These bands could correspond to premature termination of the BCDOX due to codon bias usage. The isolated fragments could be premature termination of the protein, and since the recombinant BCDOX contained an N-terminus histidine tag, the

fragments would bind to the column as well, and co-elute with the BCDOX, Figure 2-4.

It is known that in *E. coli* some tRNA codons are underrepresented. In particular, *Arg* codons (AGA, AGG CGA), the *Ile* codon (AUA) and the *Leu* codon (CUA) all represent less than 8% of their corresponding population

of codons while the *Gly* codon (GGA), *Arg* codon (CGG) and *Pro* codon (CCC) fall to less than 2% of their respective population. An analysis of the BCDOX sequence showed: *Arg* (AGA)=2 codons, *Arg* (AGG)=10 codons, *Ile* (AUA)=5 codons and *Pro* (CCC)=12 codons.^{29,30} Figure 2-5 shows the sequence of BCDOX and rare codons (codons that are expressed in low amounts in *E. coli*) are represented in blue.

When *E. coli* is used to express proteins, the expression of affected recombinant genes could be salvaged either by replacing them with synonymous codons that are more frequently used using site-directed mutagenesis or by

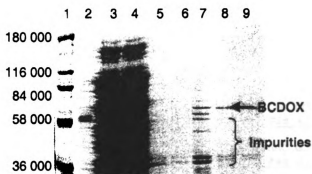


Figure 2-4. Agarose gel shows the purified recombinant mice of β -carotene-15-15' dioxygenase. The gel shows more than one protein, the shown proteins co-elute with the recombinant BCDOX. The bands with low molecular weight correspond to premature termination of proteins. 1. Marker. 2. BSA (60 KDa). 3. Supernatant. 4. Run through. 5, 6. Wash. 7-9. Eluted BCDOX.

ATG GAG ATA ATA TTT GGC CAG AAT AAG AAA GAA CAG CTG GAG CCA GTT CAG GCC
Met Glu **Ile Ile** Phe Gly Gly Asp Lys Lys Glu Gln Leu Glu Pro Val Gln Ala

AAA GTG ACA GGC AGC ATT CCA GCA TGG CTG CAG GGG ACC CTG CTC CGA AAC GGG
Lys Val Thr Gly Ser Ile Pro Ala Trp Leu Glu Gly Thr Leu **Leu** Arg Asn Gly

CCC GGG ATG CAC ACA GTG GGA GAG AGC AAG TAC AAC CAT TGG TTT GAT GGC CTG
Pro Gly Met His Thr Val Gly Glu Ser Lys Tyr Asn His Trp **Phe Asp** Gly Leu

GCC CTT CTC CAC AGT TTT TCC ATC **AGA** GAT GGG GAG GTC TTC TAC **AGG** AGC AAA
Ala Leu leu His Ser Phe Ser Ile **Arg** Asp Gly Glu Val Phe Tyr **Arg** Ser Lys

TAC CTG CAG AGT GAC ACC TAC ATC GCC AAC ATT GAG GCC AAC **AGA** ATC GTG GTG
Tyr Leu Gln Ser Asp Thr Tyr Ile Ala Asn Ile Glu Ala Asn **Arg** Ile Val Val

TCT GAG TTC GGA ACC ATG GCC TAC CCG GAC **CCC** TGC AAA AAC ATC TTT TCC AAA
Ser Glu Phe Gly Thr Met Ala Tyr Pro Asp **Pro** Cys Lys Asn Ile Phe Ser Lys

GCT TTC TCC TAC TTG TCT CAC ACC ATC **CCC** GAC TTC ACA GAC AAC TGT CTG ATC
Ala Phe Ser Tyr Leu Ser His Thr Ile **Pro** Asp Phe Thr Asp Asn Cys Leu Ile

AAC ATC ATG AAA TGT GGA GAA GAC TTC TAT GCA ACC ACG GAG ACC AAC TAC ATC
Asn Ile Met Lys Cys Gly Glu Asp Phe Tyr Ala Thr Thr Glu Thr Asn Tyr Ile

AGG AAA ATC GAC **CCC** CAG ACC **CTA** GAG ACC TTG GAG AAG GTT GAT TAC CGG AAG
Arg Lys Ile Asp **Pro** Gln Thr **Leu** Glu Thr Leu Glu Lys Val **Asp** Tyr **Arg** Lys

TAT GTG GCG GTA AAC CTG GCT ACC TCG CAC CCT CAT TAT GAC GAG GCT GGG AAT
Tyr Val Ala Val Asn Leu Ala Thr Ser His Pro His Tyr Asp Glu Ala Gly Asn

GTC CTT AAC ATG GGC ACA TCC GTC GTG GAC AAA GGG **AGG** ACA AAA TAC GTG **ATA**
Val Leu Asn Met Glu Thr Ser Val Val Asp Lys Glu **Arg** Thr Lys Tyr Val **Ile**

TTT AAG ATC CCT GCC ACA GTG CCA GAC AGC AAG AAG AAA GGG AAG AGT **CCC** GTG
Phe Lys Ile Pro Ala Thr Val Pro Asp Ser Lys Lys Lys Gly Lys Ser **Pro** Val

AAG CAC GCG GAA GTT TTC TGC TCC ATT TCC TCC CGC TCG CTG CTC TCT **CCC** AGC
Lys His Ala Glu Val Phe Cys Ser Ile Ser Ser **Arg** Ser Leu Leu Ser **Pro** Ser

TAC TAC CAC AGC TTT GGT GTC ACG GAG AAC TAT GTG GTG TTT CTG GAG CAG CCT
Tyr Tyr His Ser Phe Gly Val Thr Glu Asn Tyr Val Val Phe Leu Glu Gln Pro

TTT AAG TTG GAT ATC CTC AAG ATG GCC ACC GCA TAC ATG **AGG** GGA GTG AGC TGG
Phe Lys Leu Asp Ile Leu Lys Met Ala Tyr Ala Tyr Met **Arg** Gly Val Ser Trp

GCT TCC TGT ATG TCA TTC GAC **AGG** GAG GAC AAG ACA TAC ATT CAT ATC ATC GAC
Ala Ser Cys Met Ser Phe Asp **Arg** Glu Asp Lys Thr Tyr Ile His Ile Ile Asp

CAG **AGG** ACC AGG AAG CCT GTG CCT ACC AAG TTC TAC ACA GAT **CCC** ATG GTG GTC
Gln **Arg** Thr Arg Lys Pro Val Pro Thr Lys Phe Tyr Thr Asp **Pro** Met Val Val

Figure 2-5A. The sequence of mice of β -carotene-15-15' dioxygenase contains a total of 29 rare codons in its sequence. Arg(AGA)= 2 codons, Arg(AGG)=10 codons, Ile(AUA)=5 codons and Pro(CCC)=12 codons, sequence continues in Figure 2.5B.

TTC CAT CAT GTC AAT GCC TAC GAG GAG GAC GGC TGT GTG CTG TTT GAT GTG ATC
Phe His His Val Asn Ala Tyr Glu Glu Asp Glu Cys Val Leu Phe Asp Val Ile
 GCC TAT GAG GAC AGC AGC CTC TAT CAG CTC TTC TAC CTG GCC AAC CTG AAC AAG
Ala Tyr Glu Asp Ser Ser Leu Tyr Gln Leu Phe Tyr Leu Ala Asn Leu Ala Lys
 GAC TTC GAG GAG AAG TCC **AGG** CTG ACC TCA GTG CCT ACC CTC **AGG AGG** TTT GCT
Asp Phe Glu Glu Lys Ser **Arg** Leu Tyr Ser Val Pro Tyr Leu **Arg Arg** Phe Ala
 GTG **CCC** CTC CAT GTG GAC AAG GAT GCA GAA GTG GGC TCA AAT TTA GTC AAG GTG
Val **Pro** Leu His Val Asp Lys Asp Ala Glu Val Gly Ser Asn Leu Val Lys Val
 TCA TCT ACA ACT GCA ACA GCC CTG AAG GAG AAA GAC GGC CAT GTC TAT TGC CAG
Ser Ser Tyr Tyr Ala Tyr Ala Leu Lys Glu Lys Asp Gly His Val Tyr Cys Gln
CCC GAG GTC CTC TAC GAA GGC **CTA** GAG CTC CCT CGG **ATA** AAT TAT GCT TAC AAC
Pro Glu Val Leu Tyr Glu Gly **Leu** Glu Leu Pro Arg **Ile** Asn Tyr Ala Tyr **Asn**
 GGG AAG CCA TAT CGC TAC ATC TTT GCA GCT GAA GTA CAG TGG AGT CCA GTC CCA
Gly Lys Pro Tyr Arg Tyr Ile Phe Ala Ala Glu Val Gln Trp Ser Pro Val Pro
 ACC AAG **ATA** CTG AAA TAT GAC ATT CTC ACA AAG TCC TCC TTA AAG TGG TCT GAG
Thr Lys **Ile** Leu Lys Tyr Asp Ile Leu Thr Lys Ser Ser Leu Lys Trp Ser Glu
 GAG AGC TGC TGG CCA GCA GAG CCT CTG TTT GTT **CCC** ACG CCA GGT GCG AAG GAT
Glu Ser Cys Trp Pro Ala Glu Pro Leu Phe Val **Pro** Tyr Pro Gly Ala Lys Asp
 GAA GAT GAT GGA GTC ATT TTA TCA GCC ATC GTC TCT ACG GAT **CCC** CAA AAG CTG
Glu Asp Asp Gly Val Ile Leu Ser Ala Ile Val Ser Thr Asp **Pro** Gln Lys Leu
 CCT TTT TTA CTC ATT CTG GAT GCC AAA AGT TTT ACG GAA CTG GCT CGC GCC TCT
Pro Phe Leu Leu Ile Leu Asp Ala Lys Ser Phe Thr Glu Leu Ala Arg Ala Ser
 GTT GAT GCG GAC ATG CAC CTG GAC CTT CAT GGT TTA TTT ATC CCA GAT GCA GAC
Val Asp Ala Asp Met His Leu Asp Leu His Gly Leu Phe Ile Pro Asp Ala Asp
 TGG AAT GCA GTG AAG CAG ACT CCA GCT GAA ACG CAA GAG GTT GAA AAC TCA GAT
Trp Asn Ala Val Lys Gln Thr Pro Ala Glu Thr Gln Glu Val Glu Asn Ser Asp
 CAT **CCC** ACA GAT CCG ACA GCA CCA GAA CTG AGC CAC AGT GAA AAT GAC TTC ACA
His **Pro** Thr Asp Pro Thr Ala Pro Glu Leu Ser His Ser Glu Asn Asp Phe Thr
 GCG GGT CAT GGT GGC TCA AGT CTT **TAA**
Ala Gly His Gly Gly Ser Ser Lys ---

Figure 2-5B. The sequence of mice of β -carotene-15-15' dioxygenase contains a total of 29 rare codons in its sequence. Arg(AGA)= 2 codons, Arg(AGG)=10 codons, Ile(AUA)=5 codons and Pro(CCC)=12 codons.

supplementing the host with extra copies of the equivalent tRNA genes. Stratagene³³ has generated a strain of *E. coli* BL21-Codon Plus-RIL (BL21-RP) cells that contains a plasmid with extra copies of the tRNA genes for the codons *Arg* (AGA, AGG), *Ile* (AUA) and *Leu* (CUA), and BL21-Codon Plus-RP cells (BL21-RIL) that contains a plasmid with extra copies of the tRNA genes for the codons of the *Arg* (AGA, AGG), and *Pro* (CCC). Site directed mutagenesis of 29 different rare codons is not a simple task, and it has been reported that addition of the plasmid provides as good or better results than mutagenesis. The BL21-RIL and BL21-RP cells were generously donated by Professor Frost (MSU). Competent cells of BL21-RP and BL21-RIL were prepared, followed by transformation with pBCDOX-1. After transformation, cells from one colony were used for further experiments after the presence of both plasmids was confirmed.

Expression of BCDOX in BL21-RIL+pBCDOX-1 and BL21-RP+pBCDOX-1 was performed. The results (summarized in Table 2-2) were unexpected. First, the amount of protein obtained did not change significantly.

Table 2-2. Protein expression after addition of rare codons using XL1-Blue as host.

<i>E. coli</i>	Protein expression	
	µg of protein per 1L	BCDOX activity pmol retinal/ mg BCDOX*min
XL1-Blue-pBCDOX	612	20
XL1-Blue pACYC-RIL-pBCDOX-1	1312	6
XL1-Blue pACYC-RP-pBCDOX-1	1050	0

Second, the activity of the protein obtained was lower. As discussed previously the expression of the recombinant BCDOX seems to be very tightly controlled, and the use of a different host (BL21) might have significant impact in the control for expression of BCDOX. In more detail, the pQE30 vector, in which the BCDOX is cloned, contains a double *lac operator* for over expression of the recombinant protein. However, the plasmid does not contain the gene to express *LacI*. The BL21 cells do not have any gene that produces *LacI* protein, and thus, the expression of BCDOX is not regulated as tightly. Consequently to address this problem, the use of an *E. coli* that produces *LacI* protein would be necessary. XL1-Blue *E. coli* contains a gene that produces the *LacI* protein in its genomic DNA. For that reason XL1-Blue cells were used as host instead of BL21 cells. XL1-Blue pBCDOX competent cells were prepared and transformed with either pACYC-RIL or pACYC-RP plasmid. As expected, the expression of the protein was increased but only by a factor of two while the activity decreased showing no overall significant improvement.

Novagen has also designed a system to express proteins containing rare codons. This system contains a plasmid called pRARE, which encodes for all the rarely used codons (*Arg, Ile, Gly, Leu* and *Pro*) and it is commercially available. This plasmid is cloned in a different *E. coli* host called Rosetta Blue. This strain was particularly interesting because its genome is similar to XL1-Blue and over expresses the *LacI*^q. Therefore this was the ideal system to test if the *LacI*^q was crucial for the expression of pBCDOX-1 or not. The results obtained are summarized in Figure 2-6; all results are normalized for a better comparison. It

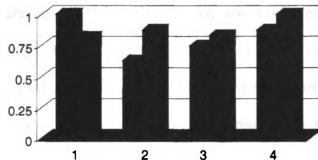


Figure 2-6. Relative activity (■) and expression of recombinant (■) BCDOX. Results using different *E. coli* hosts 1. XL1-Blue-pBCDOX-1. 2. XL1-Blue-pBCDOX-1-pRP(Arg, Pro). 3. XL1-Blue-pBCDOX-1-pRIL(Arg, Ile, Leu). 4. Rosetta-pRARE (Arg, Ile, Gly, Leu, Pro)-pBCDOX-1.

was observed that the yield of the recombinant BCDOX increased but not significantly, and the activity did not improve. The light bars show the yield of protein obtained after purification. The highest protein yield was obtained when Rosetta Blue cells were used. On the other hand, the highest activity was obtained using the original system pBCDOX-1/XL1-Blue and the lowest using pACYC-RP/pBCDOX-1. However, the differences are not significant, and could be due to the inherent variability of the system and not to a real improvement of the expression and activity of BCDOX. At this point, it was concluded that addition of tRNA's did not significantly improve either expression of the recombinant BCDOX or its activity.

2.4 Expression of BCDOX *in vivo*

The expression of BCDOX in *E. coli* was monitored *in vivo* by the colorimetric change of the oxidation of β -carotene (orange) to retinal (yellow). Lintig and co-workers, had expressed the BCDOX into a genetic background of an *E. coli* that accumulates β -carotene, which resulted in the formation of retinoids that could be monitored by a color shift.¹³ Professor Lintig graciously provided a vector that contained the genes necessary for expression of β -carotene in *E. coli*. This plasmid pORANGE¹³ (Figure 2-7) contain a gene cluster with *crtE*(a), *crtB*(b), *crtI*(c) and *crtY* (d).^{34,35} β -Carotene is derived from the condensation between the C5 skeleton of isopentyl pyrophosphate (IPP) and dimethyl allyl pyrophosphate (DMAPP). This condensation is catalyzed by *crtE*, to form geranylgeranyl pyrophosphate (GGPP). Next, the synthesis of phytoene is the first committed step in C₄₀-carotenoid biosynthesis and is catalyzed by *crtB*; it proceeds as a head-to-head condensation of two molecules of GGPP. Subsequently, four dehydrogenation steps are carried out by *crtI* in the conversion of phytoene to lycopene. Finally, lycopene is cyclized by *crtY* to form β -carotene^{36,37} (Scheme 2-2).

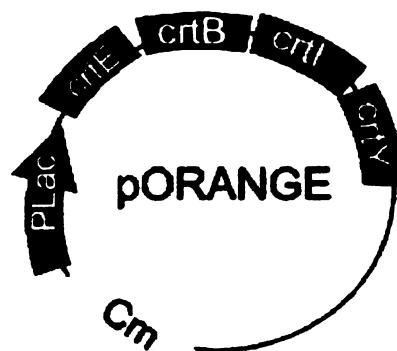
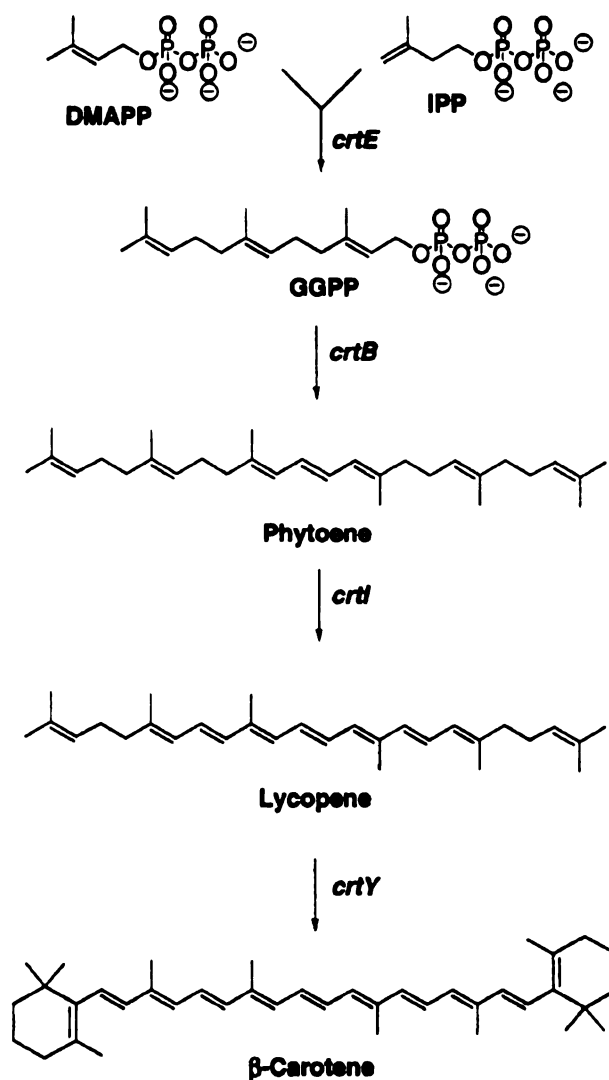


Figure 2-7. pORANGE plasmid contains a gene cluster with *crtE*, *crtB*, *crtI* and *crtY*. Those genes produce β -carotene from dimethyl allyl pyrophosphate and isopentyl pyrophosphate.



Scheme 2-2. Biosynthesis of β -carotene in *E. coli*. starts with the condensation between DMAPP and IPP.

XL1Blue cells were transformed with pORANGE. The cells obtained after transformation showed an orange coloration due to the accumulation of β -carotene (Figure 2-8). The cells were grown under red safe lights until the picture

was taken to decrease the isomerization of the produced β -carotene. The competent cells were prepared and transformed with pBCDOX-1. The transformed cells were grown and BCDOX expression was induced by addition of isopropyl- β -D-thiogalactopyranoside (IPTG). After the cells were grown for about six hours in the dark, the obtained cells were colored off white. The shift of the coloration was caused by oxidation of β -carotene (orange) to retinal (yellow) by BCDOX, as a result of BCDOX catalysis.

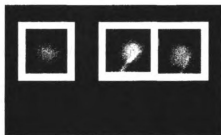


Figure 2-8. Activity of recombinant mice β -carotene-15-15' dioxygenase *in vivo*. 1. XL1-Blue transformed with pORANGE and pBCDOX-1 with out induction (addition of IPTG). Orange color due to formation of β -carotene 2. XL1-Blue transformed with pORANGE and pBCDOX-1 after addition of IPTG. The yellow coloration is due to oxidation of β -carotene to form retinal.

2.5 Expression of BCDOX using a T7 promoter

Given the fact that the main objective for this project was the understanding of BCDOX's enzymatic and mechanistic properties it was imperative to express large quantities of the enzyme. The plasmid pQE30 uses a T5 promoter (Figure 2.1), which is recognized by the *E. coli* polymerase. Nonetheless this promoter has not been widely used since it is considered a weak promoter. As a result, the low expression of the recombinant BCDOX could be due to the fact that the expression of the protein is under the control of a weak promoter. Different systems to express proteins have been developed. For example a T7-promoter (bacteriophage T7 promoter) driven system, originally developed by Studier and colleagues,³⁸⁻⁴⁰ is known to be a much stronger promoter. Also since *E. coli* RNA polymerase does not recognize the T7 promoter, there is very low transcription of the target gene in the absence of T7 polymerase. Therefore the amount of basal transcription is tightly controlled, increasing the stability of plasmids with toxic genes. Thus the BCDOX expression controlled by a T7 promoter would offer a good alternate for the expression of BCDOX. Therefore, BCDOX was to be cloned into a different plasmid under the control of a T7 promoter.

2.6 Cloning of BCDOX into a T7 promoter vector

Initially a pET-30a was chosen as the target plasmid. The MCS of pET-30a is under the control of a T7 promoter with a C-terminal histidine-tag with an intervening enterokinase cut site that should allow the removal of the his-tag in

order to study the native protein. Primers were designed using the reported sequence for mice BCDOX, in order to introduce two cut sites, *NdeI* and *XhoI*. The first step would be the extraction of the gene using PCR followed by digestion and ligation into pET-30a. Many attempts to PCR the BCDOX gene were performed. Different conditions were tried (changing the extension time, alignment temperature, concentration of different components etc.) but no PCR product was obtained. These results were puzzling for a while, and it was concluded that the only reasonable explanation was that one of the primers was not binding to the parent DNA. Therefore, a complete sequence of the gene was performed. The results showed that the reported sequence differed from the sequence in the vector in the last 51 base pairs, thus explaining why the attempted PCR provided no product. Figure 2-9 shows the differences between the reported gene and the sequenced gene.

In the course of this investigation, we decided to use a thrombin cut site to remove the His tag instead of enterokinase. The thrombin cut site is known to be more efficient for cleavage than the enterokinase 1:1000, 1:10 enterokinase:protein ratio respectively. If *XhoI* is used to clone the desired gene in the pET30a the thrombin cut site would not be present. pET-29b(+) contains a T7 promoter and a thrombin site which is known to be more efficient when it comes to enzymatic cleavage (1:1000), and there is a His-tag that would facilitate the purification using an NTA column. Thus it was chosen to be the host for mice BCDOX. Two cut sites were chosen to clone the BCDOX gene into pET-29b(+), (*NdeI*(295) and *KpnI*(238)). A set of primers were designed, the first of which

```

MEIIFGQNKKEQLEPVQAKVTGSIPAWLQGTLLRNGPGMHTVGESKYNHWF DGL
MEIIFGQNKKEQLEPVQAKVTGSIPAWLQGTLLRNGPGMHTVGESKYNHWF DGL

ALLHSFSIRDGEV FYRSKYLQSDTYIANIEANRIVVSEFGTMAYPDPCKNIFSK
ALLHSFSIRDGEV FYRSKYLQSDTYIANIEANRIVVSEFGTMAYPDPCKNIFSK

AFSYLSHTIPDFTDNCLINIMKCGEDFYATTETNYIRKIDPQTLETLEKVDYRK
AFSYLSHTIPDFTDNCLINIMKCGEDFYATTETNYIRKIDPQTLETLEKVDYRK

YVAVNLATSHPHYDEAGNVLMGTSVVDKGR TKYVIFKIPATVPDSKKKGKSPV
YVAVNLATSHPHYDEAGNVLMGTSVVDKGR TKYVIFKIPATVPDSKKKGKSPV

KHAEVFCSISSRSLSPSYHSGFVTENYVVFLEQPFKLDILKMATAYMRGVS W
KHAEVFCSISSRSLSPSYHSGFVTENYVVFLEQPFKLDILKMATAYMRGVS W

ASCMSFDREDKTYIHIIDQRTKRPVPTK FYTDPMVVFHHVNAYEEDGCVLFDVI
ASCMSFDREDKTYIHIIDQRTKRPVPTK FYTDPMVVFHHVNAYEEDGCVLFDVI

AYEDSSLYQLFYLANLNKDFEEKSRLTSVPTLRRFAVPLHVDKDAEVGSNLVKV
AYEDSSLYQLFYLANLNKDFEEKSRLTSVPTLRRFAVPLHVDKDAEVGSNLVKV

SSTTATALKEKDGHVYCQPEVLYEGLELPRINYAYNGKPYRYIFAAEVQWSPVP
SSTTATALKEKDGHVYCQPEVLYEGLELPRINYAYNGKPYRYIFAAEVQWSPVP

TKILKYDILTKSSLKWSEESCWPAEPLFVPTPGAKDEDDGVILSAIVSTDPQKL
TKILKYDILTKSSLKWSEESCWPAEPLFVPTPGAKDEDDGVILSAIVSTDPQKL

PFLILDAKSFTELARASVDADMHLDLHGLFIPDADWNAVKQTPAETQEVENSD
PFLILDAKSFTELARASVDADMHLDLHGLFIPDADWNAVKQTPAETQEVENSD

HPTDPTAPELSHSENDFTAGHGSSSL      Expected
HPTDPTAPEVPRVDLQPSLIS-----      Actual

```

Figure 2-9. Comparison of the amino acid sequence of the reported mice BCDOX (top) and the obtained BCDOX (bottom). The last seventeen amino acids differ from the reported sequence.

contained an *NdeI* cut site (5'-CGGACATATGGAGATAATATTTGGCCA GAATAAGAAAGAACAGCTGG-3') and the other a *KpnI* cut site (5'-GACCCGGGTA CCCTCTGGTGCTGTCGGATC-3'). The strategy to clone BCDOX into pET-29b(+) consisted of amplifying the gene, digesting it and ligating into the pET-29b(+) vector. The PCR was uneventful and provided the BCDOX gene with the cut

sites *NdeI* and *KpnI* (See Material and Methods for details). The next step would be the digestion of the PCR product and the vector pET-29b(+). The double digestion of the plasmid and the PCR product proved to be problematic. According to the manufacturer, for a double digestion using *KpnI* and *NdeI*, the use of NEBuffer 1 is recommended. Even though the activity of *NdeI* in NEBuffer 1 is 100%, it is not the buffer that shows the optimal activity for *NdeI*.⁴¹ A digestion test using pET-29b(+) as the standard showed that at 2 and 4 hours the digestion was not complete for *NdeI*. Only overnight (~14 hrs) the digestion showed a single band. Thus, the digestion of pET-29b(+) and BCDOX gene (from PCR) using NEBuffer 1, were performed overnight at 37 °C. After digestion, and purification of the digested vector and gene, the isolated DNA were quantified. Ligations were performed using different mole ratio between the gene and the plasmid, at different temperatures. After transformations of JM109 cells with the ligation reaction products, some colonies were obtained. Analysis of those colonies showed that no gene was cloned into pET-29b(+)-BCDOX. It has been reported that at times, residual polymerase might fill the sticky ends to generate blunt ends,⁴² thus, interfering with the proper ligation of the gene into the vector. To avoid polymerase interference, the PCR product was purified. The gene and plasmid were digested overnight at 37 °C, purified, and quantified. Ligation attempts using the conditions previously mentioned were set up, followed by transformation into JM109 which provided a few colonies. Analysis of those colonies resulted in vectors without the desired gene.

Even though the manufacturers⁴¹ suggest the use of NEBuffer 1 for the double digestion, it was possible that double digestion was not taking place; therefore cloning of the gene would not be possible. To address this problem, single digestions were performed consecutively. Also to address the re-ligation problem, two strategies were followed. The first one was based on the principle that linear DNA transforms with much lower efficiency than a circular DNA. Sequence analysis of the pET-29b(+) showed a *Bgl*III cut site between the *Nde*I and *Kpn*I, that is not present in the BCDOX gene. For that reason, digestion of the ligation product with *Bgl*III would digest only the re-ligation products and not the cloned products. After transformation, the majority of the colonies obtained should contain the cloned product and not the re-ligation product, due to the low transformation efficiency of linear DNA. Ligation reactions using the conditions previously mentioned were set up, followed by digestion with *Bgl*III for two hours. The digested ligations were transformed into JM109, and after transformation of the ligated products, no colonies were observed.

The second strategy consisted of the removal of the terminal 5'-phosphate groups in the digested plasmid DNA. During *in vitro* ligation, DNA ligase catalyzes the formation of a phosphodiester bond between adjacent nucleotides only if one nucleotide has a 5'-phosphate residue and the other carries a 3'-hydroxyl. Removal of the 5'-phosphate residues can be performed with alkaline phosphatase Calf Intestine Phosphatase (CIP).^{43,44} As previously, PCR, purification, double digestion of plasmid and BCDOX, and purification were performed. The digested vector was then treated with CIP for 2 hours at 37 °C.

This was followed by ligation and transformation into JM109. After overnight incubation, 6 colonies were obtained. Only one of them seemed to contain the BCDOX gene, Figure 2.10 (pBCDOX-2). The cloned product (Vector-BCDOX) was obtained when a 1:10 mole ratio between vector/BCDOX and incubation at 16 °C overnight were performed.

The obtained plasmid was sent for sequencing using the following primers: 5'-GATCTCGATCCCGCGAAATTAATACGAC-3' BBB0044; 5'-CCAGACCCTAGAGACCTTG GAGAAGG-3' BBB0045; 5'-CGAGGAGAAGTCCAGGCTGACC-3' BBB0046; 5'-GATCGATCTC GATCCCGCG-3' BBB0055; 5'-GCAGACTGGAATGCAGTGAAGC-3' BBB0052. The sequence showed 4-point mutations: two silent mutations and two real mutations, Cys (TGC) 102 mutated to Tyr (TAC), and Ala (GCC) 337 mutated to Thr (ACC). These plasmid obtained would be refer as pBCDOX-2.

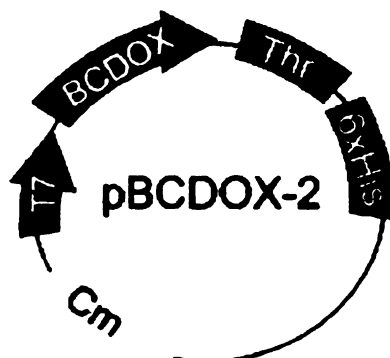


Figure 2-10. Gene of mice β -carotene-15-15' dioxygenase cloned in a pET-29b(+)=pBCDOX-2.
Thr=thrombine

2.7 Expression of pET-29b(+)-BCDOX

E. coli Rosetta(DE3)pLysS cells were chosen as the host to overexpress the pET29b(+)-BCDOX recombinant protein. This strain was chosen because it supplies the rare tRNAs, and it possesses a copy of the T7 RNA polymerase.³⁸ This gene is under the control of the *lacUV5* promoter, which allows a tight control in the expression of the recombinant protein. After the transformation of the host with pBCDOX-2, several colonies were obtained, and analysis of five of them showed the pBCDOX and the codon plus plasmid. One colony was chosen for further studies.

2.8 Optimization of pBCDOX expression

The recombinant BCDOX protein was expressed and indeed this time overexpression of the protein was obtained. Unfortunately though, all overexpressed protein was in the form of inclusion bodies. The formation of inclusion bodies was unfortunate but not surprising due to BCDOX's nature and because it has been observed that overexpression of recombinant proteins in *E. coli* can result in the formation of inclusion bodies in the cell cytoplasm or periplasm.^{45,46} It has been found that the tendency to form aggregates depends on the rate of expression and the rate of refolding. If the rate of expression exceeds the rate of refolding, co-expressed proteins come into contact with one another and aggregate prior to reaching their native state.⁴⁷ It is known that disulphide bonds and hydrophobic patches may affect the solubility of a protein and result in the formation of inclusion bodies. To improve solubility, researchers have applied a variety of

methods such as adjusting the amount of IPTG used,⁴⁸ lower the temperature,⁴⁹ co-expression of the protein of interest with chaperones and foldases,⁵⁰ the use of solubilizing fusion partners,^{51,52} different *E. coli* strains,³⁵ among others. So far, all literature reports that describe the expression and isolation of BCDOX^{11,13,53-55} agree with the fact that BCDOX is a cytosolic protein. However, there is evidence suggesting an amphiphilic character of BCDOX and that increased ionic strength and high protein concentrations are likely to promote aggregation. As previously discussed it has been observed by Wyss and co workers that at a 50-fold concentration of the protein leads to 75% loss of its activity which is attributed to aggregation.¹⁰

Due to the nature of BCDOX activity, oxidation of β -carotene, it is very possible that this protein is associated with lipid aggregates or other proteins. The fact that when BCDOX is expressed mostly as inclusion bodies supports the latter findings. Not much information is available about the overexpression of BCDOX in bacterial systems.

2.9 Attempts to express the recombinant BCDOX in a soluble form

2.9.1 Decrease the rate of expression, use of different temperatures, time and IPTG

The rate of protein folding is a first order process. By decreasing the rate of expression of the recombinant protein, one can favor correct protein folding. There are several reported methods to decrease the rate of overexpression of

recombinant proteins, a number of which were tested to express BCDOX as soluble protein. The results are discussed below. All the following experiments were analyzed by 10% polyacrilamide SDS gel electrophoresis of both the supernatant (soluble fraction) and pellet (possible inclusion bodies) of the crude expression mixture. To assure that BCDOX was not present in small amounts in the supernatant, the samples were treated with the Anti-His HRP conjugate.²¹ Treatment with Anti-His HRP conjugate would allow for the identification of histidine tagged proteins.

The most commonly used approach to improve protein solubility and folding involves adjusting the amount of IPTG used⁴⁸ and the temperature at which the protein is expressed.⁴⁹ The net effect when the temperature or the IPTG concentration is decreased is to reduce the rate at which recombinant proteins are expressed. By slowing down the rate for the protein expression proper folding is promoted. Therefore, the expression of BCDOX was performed at different temperatures (30, 28, 16 and 4 °C) and different concentrations of IPTG (0, 0.4, 1 and 10 mM). In addition all other possible permutations were screened. The cells were lysed and the supernatant and the pellets were analyzed. Unfortunately, the results showed again that in all cases, the protein was in the form of inclusion bodies and no trace of soluble protein was found using Anti-His HRP conjugate. As expected, low temperatures and low IPTG concentration decreased the amount of inclusion bodies formed (Table 2-3). Because T7 is a strong promoter and it is under the control of the lac operon, the expression of BCDOX is activated by addition of IPTG.

Table 2-3. Expression of recombinant pBCDOX-2 in *Rosetta(DE3)pLysS*, induced at different concentrations of IPTG and different temperatures. The results are shown as **S**=Supernatant, **I.B.**= Inclusion bodies. ✓ = BCDOX present, x= No BCDOX present.

IPTG mM	Temperature							
	30 °C		28 °C		26 °C		16 °C	
	S	I.B	S	I.B.	S	I.B.	S	I.B
0.0	x	x	x	x	x	x	x	x
0.4	x	✓	x	✓	x	✓	x	✓
0.1	x	✓	x	✓	x	✓	x	✓
10	x	✓	x	✓	x	✓	x	✓

A variation to the addition of IPTG was performed. Four different portions were added over a period of eight hours and each sample was analyzed. After SDS page analysis of the supernatant and the pellet, no change in the amount of soluble recombinant protein BCDOX was observed. (Table 2-4). Again all supernatants were treated with histidine tag antibody and no protein was

Table 2-4. Expression of recombinant pBCDOX-2 in *Rosetta(DE3)pLysS*, induced by addition of different portions IPTG of OD₆₀₀~ 1.0 for overnight at 26 °C. The results are shown as **S**=Supernatant, **I.B.**= Inclusion bodies. ✓ = BCDOX present, x= No BCDOX present.

IPTG	S	I.B.
0 portions	x	x
1 portion	x	✓
4 portions	x	✓
8 portions	x	✓

observed. After these trials it was concluded that the amount of IPTG used to induce protein expression did not have an impact in the folding of the protein. The expression of the proteins at 26 °C was monitored every 1, 3, 5 and 8 hours. When the cells were lysed analysis of the soluble fraction and pellet showed that BCDOX was expressed only as inclusion bodies. This experiment was repeated at 16 °C and 28 °C and the same results were obtained. Thus it could be concluded that different temperatures and different induction times did not have any positive influence in the formation of soluble proteins.

2.9.2 Different *E. coli* strain

It has been shown that in a bacterial host, the solubility of proteins changes with different *E. coli* strains. Because of this reason, different *E. coli* strains were tested for overexpression of the recombinant BCDOX. The strains chosen for overexpression were: A) Tuner(DE3)pLacI. The difference between these cells and Rosseta is that Tuner cells do not have the *LacY* mutation, which allows the uniform absorption of IPTG. B) ER2566 these cells are derived from BL21 and have a copy of the T7 RNA gene into *lac* and BL21(DE3) pLysS. Competent cells for Tuner(DE3)pLacI, BL21(DE3)pLysS., Rosseta and ER2566 were prepared. Each strain was transformed with pBCDOX-2 and the expression of the recombinant protein was tested. Because the concentration of IPTG proved to be irrelevant in the Rosseta strain the next experiments were performed at constant IPTG concentration of 0.4 mM, added at OD₆₀₀~ 1.0, and 26 °C. Aliquots of 6 mL were taken at 1, 3, 5 and 8 hours after induction. The results showed (Table

2-5) that the protein was again over expressed in the form of inclusion bodies, with the concentration of inclusion bodies increasing with time.

Treatment of all supernatants with Histidine-tag antibody resulted in negative results. Since no recombinant protein was identified in the soluble form the experiment was repeated at expression temperatures of 16 °C and 9.5 °C. Aliquots were taken at 4, 8, 19, 30 and 42 hours after induction. The expression of the recombinant BCDOX was pretty slow (as expected) and when the overexpression of BCDOX was achieved, the protein was formed as inclusion bodies, as in the previous cases. No protein was found in the supernatant when

Table 2-5. Expression of recombinant pBCDOX-2 using different host Rosetta(DE3)pLysS, Tuner(DE3)pLacI, BL21(DE3)pLysS ER2566 induced by addition of IPTG at different portions OD₆₀₀~ 1.0 for overnight at 26 °C. The results are shown as S=Supernatant, I.B.= Inclusion bodies. ✓= BCDOX present, x= No BCDOX present.

<i>E. coli</i> host	IPTG (mM)	S.	I.B.
Rosetta(DE3)pLysS- pBCDOX-2	0.0	x	x
	0.01	x	✓
	0.10	x	✓
	1.00	x	✓
Tuner(DE3)pLacI- pBCDOX-2	0.0	x	x
	0.01	x	✓
	0.10	x	✓
	1.00	x	✓
BL21(DE3)pLysS- pBCDOX-2	0.0	x	x
	0.01	x	✓
	0.10	x	✓
	1.00	x	✓
ER2566-pBCDOX-2	0.0	x	x
	0.01	x	✓
	0.10	x	✓
	1.00	x	✓

each supernatant was treated with the Anti-His HRP conjugate. At this point it was concluded that the use of different *E. coli* strains did not have a significant influence in the solubility of BCDOX. Also it has been reported that addition of the substrate in the culture can improve the solubility of a protein. The native substrate for BCDOX is β -carotene, and as previously discussed, the pORANGE vector can produce β -carotene *in vivo*. Thus an ER2566 *E. coli* strain was transformed with pORANGE. After expression of the recombinant protein no improvement in the solubility was observed (Table 2-6).

Table 2-6. Expression of recombinant BCDOX in the presence of its native substrate β -carotene. ER2566 was chosen as the host and it was transformed with pORANGE and pBCDOX-2. pORANGE is responsible for the formation of β -carotene. The protein was express under red safe light. The results are shown as S=Supernatant, I.B.= Inclusion bodies. \checkmark = BCDOX present, \times = No BCDOX present.

<i>E. coli</i> host	IPTG (mM)	S.	I.B.
ER2566-pBCDOX-2 pORANGE	0.0	\times	\times
	0.01	\times	\checkmark
	0.10	\times	\checkmark
	1.00	\times	\checkmark

2.9.3 Over expression of chaperones

In *E. coli* cells there is a complex protein machinery to help proteins fold after their biosynthesis or a heat shock.⁵⁶ Those complex proteins are called chaperone proteins. Molecular chaperones possess no catalytic activity. Their function is to prevent polypeptide chains from forming incorrect associations with itself or another protein.⁴⁷ For up regulation of chaperones in bacterial systems

two techniques are used, heat shock and addition of organic solvents, especially ethanol. When *E. coli* is subjected to a variety of stresses, including temperature up shift, exposure to organic solvents, and the accumulation of misfolded proteins, the synthesis of heat shock proteins (HSPs) is upregulated in order to repair cellular damage. It has been reported that solubility of proteins enhance using a *heat shock* prior to protein induction and that the over expression of the recombinant protein is not affected. It is believed that increase in the solubility is due to the fact that some heat-shock proteins or chaperones such as *GroEL* and *DnaK* are up regulated when the *E. coli* is heated at 42 °C. Two strains used in these experiments were: Rosetta(DE3)pLysS and Tuner(DE3)pLacI. 100 mL of cells were at 37 °C until OD₆₀₀~ 0.1 and then warmed to a 42 °C until OD₆₀₀~ 1.0. At this point expression was induced with 0.4 mM IPTG at 16 °C. Aliquots of 6 mL were taken at 3, 6, 9 and 16 hours. The cells were collected and analyzed as previously described. The results (Table 2-7) showed that there was not enhancement in the solubility of BCDOX obtained.

Since increasing the temperature did not provide the expected results, instead, to promote the expression of chaperones addition of organic solvents was performed. Ethanol is one of the most powerful elicitors of the heat-shock response in *E. coli*, and its addition to the media has been shown to increase the solubility of a variety of proteins. The amount of ethanol used to enhance solubility was 3%, and was added right before inoculation. The cells with ethanol grew much slower than the ones without. The expression was induced and the cells were collected and analyzed as previously. The results showed (Table 2-7

hat there was almost no expression of BCDOX in the pellet or in the supernatant. In addition, the His-tag detection assay did not detect any His-tagged protein in the supernatant solution. All attempts to express BCDOX as a soluble protein have failed. And as previously discussed the experiments were performed in a small scale (100 mL of culture) and the supernatants were treated with the Anti-His HRP conjugate to test for any traces of BCDOX in the supernatant. No recombinant protein was identified from any of the previous studies. To finally confirm that no protein was expressed in soluble form, a large-scale expression

Table 2-7. Overexpression of chaperones to improve the solubility of the recombinant BCDOX. **A.** Heat shock. **B.** Addition of ethanol. The results are shown as **S**=Supernatant, **I.B.**= Inclusion bodies. \checkmark = BCDOX present, \times = No BCDOX present.

<i>E. coli</i> host	Time	A		B	
		S	I.B	S	I.B
Rosetta(DE3)pLysS-pBCDOX-2	3	\times	\times	\times	\times
	6	\times	\checkmark	\times	\times
	9	\times	\checkmark	\times	\checkmark
	16	\times	\checkmark	\times	\checkmark
Tuner(DE3)pLacI-pBCDOX-2	3	\times	\times	\times	\times
	6	\times	\checkmark	\times	\times
	9	\times	\checkmark	\times	\checkmark
	16	\times	\checkmark	\times	\checkmark

was performed. Three liters of culture were grown and induced (for details see Materials and Methods). The supernatant was purified using an NTA column, under the same conditions as the recombinant protein cloned in pQE30. The purified supernatant was not as pure as expected. The purified supernatant showed a band corresponding to the size of BCDOX, but more than one protein were present in the eluted fraction. Whether the presence of those proteins was due to unspecific binding or associated proteins to BCDOX is not well understood. The purified supernatants showed no BCDOX activity (eluted fractions).

2.10 Activity *in vivo*

Finally to test whether small amounts of the recombinant BCDOX were expressed in an active form in *E. coli*., BCDOX activity was monitored *in vivo* by following the colorimetric change of the oxidation of β -carotene (orange) to retinal (yellow). The source of β -carotene *in vivo* would be the pORANGE vector as previously described and the host chosen was ER2566 because it contains a copy of the T7 DNA polymerase in its genome. After transformation of ER2566 with pORANGE the competent cells were prepared and transformed with pBCDOX-2. Unfortunately this test did not provide the expected results, because when the ER2566-pORANGE was expressed the phenotype obtained was yellowish instead of orange (Table 2-8). This experiment was repeated several times with a great variability of different shades of yellow-orange colors obtained. There is no obvious reason why the ER2566 *E. coli* strain does not work as well

as the XL1-Blue strain. The precursor of β -carotene is isoprene, and the genotype of ER2566 does not suggest any deficiency in the production of isoprene (it should be at least as good as XL1-Blue). The experiments were repeated several times and the results were not consistent. Variability in the expression of carotenoids in *E. coli* has been previously observed and no obvious correlation between the genotype and carotenoid expression has been found. Sandman has suggested that this difference is due to the different accumulation of carotenoids in the membrane.³⁶

Table 2-8. Attempts to test the recombinant mice β -carotene-15-15' dioxygenase pBCDOX-2 *in vivo*. ER2655 was transformed with pORANGE and pBCDOX-2. The yellow coloration should be due to oxidation of β -carotene to form retinal. However in this case it was impossible to establish a difference between the coloration of the *E. coli* that produced β -carotene and the one that should produce retinal. Y=yellow, O=orange.

Temperature	16 °C		26 °C		26 °C	
	IPTG	Color	IPTG	Color	IPTG	Color
ER2566/pBCDOX-2	x	Y	x	Y	x	Y
ER2566/pBCDOX-2	✓	Y	✓	Y	✓	Y
ER2566 /pORANGE	x	Y	x	Y	x	Y
ER2566 /pORANGE	✓	Y	✓	Y	✓	Y
ER2566/pBCDOX-2/pORANGE	x	Y	x	O	x	Y
ER2566/pBCDOX-2/pORANGE	✓	Y	✓	Y	✓	Y

2.11 Expression of the native BCDOX under the control of a T7 promoter

2.11.1 Mutation of BCDOX to add a stop codon to avoid translation of the 6xHis tag and thrombine cut site

The BCDOX was cloned into a T7 promoter-containing pET-29b(+) to increase the expression of the BCDOX protein. As expected the T7 promoter increased the expression of BCDOX. However, inclusion bodies were obtained despite all the attempts to express BCDOX in a soluble form. Improper folding of the recombinant proteins causes the formation of inclusion bodies. The lack of folding could be caused by the extra amino acids that formed the thrombine cut site and the histidine tag at the C- terminus in the recombinant protein. To eliminate the possibility that the extra amino acids were interfering with the correct folding of the recombinant protein, a stop codon was added to the BCDOX sequence (before the 6x histidine and thrombine). The presence of the stop codon would stop the translation of the extra amino acids, therefore, express only the native protein. The first approach to add the stop codon was by using site directed mutagenesis. Site directed mutagenesis is a technique used to introduce specific mutations into a DNA sequence. Site directed mutagenesis uses chemically synthesized oligonucleotides that contain the desired mutation, which serves as a primer for DNA synthesis using a DNA polymerase.

A set of two primers that contained the stop codon were designed; Primer 1 BBB0078 5'-GCACCAGAGGGTACCTAGGTG-3' and primer 2 BBB0079, 5'-CACCTAGGTACCCTCTGGTGC-3'. After the PCR was performed no product was

observed. To address this problem, the extension time was increased, (the vector is 7 Kbp long)⁵⁷ but no improvement was obtained. A gradient PCR to scan for the optimal temperature (from 65.5 to 45 °C) did not change the results. When site-directed mutagenesis is used, the amplification of the vector is linear and not exponential like in PCR. Therefore, increments in the concentration of parent DNA were increased up to 100 ng and 300 ng, but again no amplified DNA were observed. It is possible that the lack of product is due to the fact that the pBCDOX size was too large. Even though the amplification of long DNA (5-45 Kbp) has been reported⁵⁷⁻⁶⁰ it was accomplished by changing parameters such as pH, glycerol, and DMSO concentration, decreasing denaturing times, salt concentrations, etc. Screening for all those parameters would be costly and time consuming. There are other commercially available options from Stratagene and Invitrogene, but they are costly and not 100% efficient. Thus, a different experimental design was approached.

2.11.2 Cloning of BCDOX

The next approach was to use a conventional cloning strategy. First, mutation of the BCDOX gene to add a stop codon using normal PCR, digestion and then ligation into a pET-29b(+) vector which contains the T7 promoter was attempted. The same cut sites as before were chosen (*NdeI* and *KpnI*) since they are present in the BCDOX gene. The primers used were 5'-CACCAGGGTACCCTATGGTGCAGT-3' and the other that binds to the beginning of the BCODX gene 5'-GGA GATATACATATGGAGATAATA-3'. The PCR was simple, and

provided the BCDOX gene with the *NdeI* and *KpnI* cut sites and stop codon. As mentioned previously double digestion using *KpnI* and *NdeI*, proved problematic. Thus, this time a single digestion followed by another single digestion using the optimal conditions for each enzyme was performed.⁴¹ After the PCR reaction the DNA was purified, followed by a single digestion of the vector and PCR product using *KpnI*. The DNA was purified followed by another single digestion using *NdeI*. DNA purification and quantification was performed before ligations. The ligation reactions were performed at 16 °C at a ratio of 1:5, 1:10 and 1:15 moles of vector/BCDOX. The ligation reactions were transformed into JM109. The gene obtained from one of the transformed colonies was sent for sequencing using the following primers 5'-CAGACCCTAGAGACCTTGGAG-3' BBB0080; 5'-CGAGGAGAAGTCCAGG CTGACC-3' BBB0046; 5'-GATCGATCTCGATCCCGCG-3' BBB0055; 5'-GCAGACTG GAATGCAGTGAAGC-3' BBB0052. The sequence confirmed the presence of the stop codon, and no other mutation was found in pBCDOX-3 (Figure 2-13).

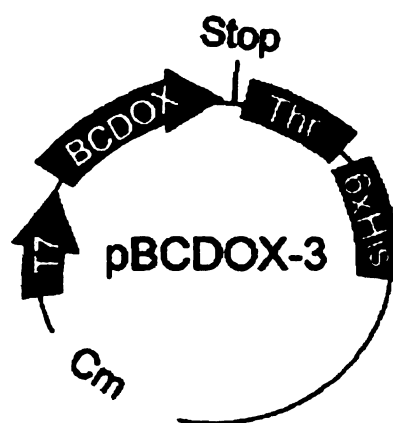


Figure 2-13. Gene of mice β -carotene-15-15' dioxygenase cloned in a pET-29b(+). A stop codon was added before the thrombin (thr) cut site and the histidine tag (pBCDOX-3).

2.11.3 Expression of pBCDOX-3; native BCDOX

For over expression of the recombinant BCDOX the plasmid was transformed into an *E. coli* Rosetta(DE3)pLysS. The first attempt to express the mutated BCDOX was performed at 16 °C and 28 °C,^{61,62} and different concentrations of IPTG as well as different induction times were screened. The cells were collected and lysed and both the supernatant and the pellet were analyzed using SDS page gel. At 0.1 mM IPTG no expression of recombinant BCDOX was observed. When protein expression was monitored at different times, the overexpressed protein was only found as inclusion bodies (Table 2-9).

Table 2-9. Expression of recombinant pBCDOX-3 in Rosetta(DE3)pLysS, induced at different concentrations of IPTG and temperatures. The results are shown as S=Supernatant, I.B.= Inclusion bodies. ✓ = BCDOX present, x=

IPTG mM	Time hours	16 °C		28 °C	
		S	I.B	S	I.B
0.00	3	x	x	x	x
	6	x	x	x	x
	9	x	x	x	x
	16	x	x	x	x
0.01	3	x	x	x	x
	6	x	x	x	x
	9	x	✓	x	✓
	16	x	✓	x	✓
0.10	3	x	x	x	x
	6	x	x	x	x
	9	x	✓	x	✓
	16	x	✓	x	✓
1.00	3	x	x	x	x
	6	x	x	x	x
	9	x	✓	x	✓
	16	x	✓	x	✓

Unfortunately the results obtained were the same as the recombinant protein with the thrombine cut site and the histidine tag. Thus it was concluded that the extra amino acids in the recombinant BCDOX were not responsible for the formation of inclusion bodies. It is possible that the amphiphilic character of BCDOX, as well as, the fact that the active enzyme could be a tetramer might decrease the rate of folding, therefore enhance formation of the inclusion bodies.

2.12 Isolation of inclusion bodies

The formation of inclusion bodies could offer several advantages for the production of recombinant proteins if they can be refolded. The proteins produced in inclusion bodies might be unstable in the cytoplasm of *E. coli* due to proteolysis and may be toxic for the host cell in the native conformation. It has been seen that under appropriate conditions, the recombinant proteins deposited in inclusion bodies amount to about 50% or more of the total cellular protein. Since inclusion bodies have a relatively high density they can be isolated from the cellular proteins, with a purity up to 90%.⁶³ A number of methods have been reported for solubilizing and refolding proteins from inclusion bodies. Denaturants such as urea and guanidine hydrochloride, and the use of highly alkaline pHs, strong thiol-containing reductants (dithiothreitol, dithioerythritol mercaptoethanol), chelating reagents (EDTA), have been reported to solubilize a significant percentage of insoluble proteins. Renaturation is usually accomplished by the removal of excess denaturants by either dilution or a buffer-exchange step

(dialysis, diafiltration, gel-filtration chromatography or immobilization onto a solid support).

As mentioned previously, attempts to produce a soluble BCDOX have been unsuccessful. Therefore, optimization of the expression and isolation of inclusion bodies was pursued. Isolation of the inclusion bodies was accomplished by over expression of BCDOX in *Rosetta(DE3)pLysS*. The soluble proteins were separated and the pellet was washed with a buffer containing 0.5% Triton X-100, followed by a wash with buffer containing 2 M urea and 0.5% Triton X-100. The pellet obtained was resuspended into a buffer containing 8 M urea and DTT, or 6 M guanidine hydrochloride. It was observed that smaller concentrations of urea (<6 M or <4 M) or guanidine did not solubilize the inclusion bodies. It was also noticed that addition of DTT increased the solubility. The inclusion bodies were purified almost with 90% purity and around 8 mg per liter (Figure 2-14).

After the BCDOX was purified (from inclusion bodies) and solubilized, refolding of the protein was the next step. Solubilization of inclusion bodies usually occurs using strong chaotropes, such as urea, and the protein is refolded to

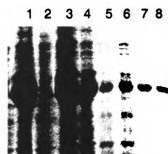


Figure 2-14. 10% agarose gel. Isolation of inclusion bodies produced by expression of recombinant mice β -carotene-15-15' dioxygenase cloned in a pBCDOX-1 in *Rosetta(DE3)pLysS*. 1. Cell pellet 2. Supernatant. 3,4. Inclusion bodies. 5, 6, 7. Inclusion bodies washed with Triton X-100. 8. BCDOX marker.

its correct structure as the denaturant concentration is reduced. In general, refolding is initiated by diluting the denatured protein into a refolding buffer solution (no chaotropic agent present). Once the denaturant concentration is reduced the protein refolds at rate equivalent to first order kinetics. The unfolded protein quickly folds to an intermediate form, and there is a slow transition from the intermediate to the native form (limiting step) and may take minutes to days to complete. At this point nonproductive aggregation reactions can occur, resulting in the formation of stable aggregates.⁴⁷ Therefore it is really important to avoid this pathway, which can be achieved by having very low concentrations of protein. In the case of disulfide-bonded proteins, renaturation buffers must promote disulfide bond formation. The options available for refolding include dilution, dialysis, diafiltration, gel filtration, and detergents among others.⁶¹ With those facts in mind, design of a system to refolded BCDOX was performed. As mentioned previously, there are no reports concerning BCDOX (or a similar protein) refolding, thus, preliminary refolding experiments were performed. The refolding method chosen was a 100-fold dilution, basically because of its simplicity and the fact that the amounts of protein used are low. Glutathione was used to aid in disulfide formation. Aliquots of the solution were withdrawn and BCDOX activity was assayed, (~0.16 mg of BCDOX were used per assay). None of the samples showed activity. Also dialysis of BCDOX protein was performed but no active protein was obtained.

As mentioned previously, this is not the only time in which bacterial expression is problematic. In addition it has been observed that the BCDOX

protein is not stable. Andersson, Zack and Cunningham reported that expression of human BCDOX in bacteria was problematic, and even though the expression in *E. coli* has been reported, there was no reference to the efficiency of the BCDOX expressed.^{15,17,18} Andersson, reported that to overcome the instability of the recombinant BCDOX protein they used insect cells to overexpress BCDOX.¹⁸ Also baby hamster HIK cells have been used to express BCDOX but whether they chose this system because they had problems with the bacterial system and then they had to switch to mammalian cells is unknown.

Progress in this project was hampered by the lack of expressing large amounts of soluble BCDOX protein. Unfortunately the use of mammalian cells to express BCDOX in our group was not possible. Thus at this point we decided not to continue to attempt to express BCDOX.

2.13 Material and methods

A. Different *E. coli* used

<i>E. coli</i> strain	Genotype
JM109	<i>c14'(McrA')recA1 endA1 gyrA96 thi-1 nsdR17 (r_k-m_k+) supE44 relA1 Δ(lac-proAB)[F' traD36 proAB lac^R Z</i>
XL1Blue	<i>recA1 endA1 gyrA96 thi-1 nsdR17 (r_k-m_k+) supE44 relA1 [F' proAB lac^R Z ΔM15 Tn10 tet^r]</i>
BL21(DE3)pLysS	<i>E. coli B F' ompT hsdS(r_B⁻ m_B⁻) dcm⁺ gal (srl-recA)306::Tn10(Tc^R)(DE3)pLysS(Cm^R)</i>
Tuner(DE3)pLacI	<i>coli B F' ompT hsdS(r_B⁻ m_B⁻) dcm⁺ gal lacY1 (DE3)pLacI(Cm^R)</i>
BL21-Codon Plus-RIL.	<i>E. coli B F' ompT hsdS(r_B⁻ m_B⁻) dcm⁺ Tet^r gal endA Hte[argU ileY leuW Cam^r]</i>
BL21-Codon Plus-RP	<i>E. coli B F' ompT hsdS(r_B⁻ m_B⁻) dcm⁺ Tet^r gal endA Hte[argU proL Cam^r]</i>
Rosetta Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 (r_{k12}-m_{k12}+) supE44 relA1 lac pRARE [F' proA⁺ B⁺ lac^R Z ΔM15::Tn10 Tet^r]</i>

B. Typical preparation of competent cells

The *E. coli* strain of interest was grown (37 °C until OD₆₀₀ of 0.4-0.6, the media contained LB, the corresponding antibiotic). After about three h The cells were harvested by centrifugation at 5,000 RPM for 10 min at 4 °C. The cells were re-suspended (100 mL of 0.9% NaCl) and then centrifugated at 5,000 RPM for 10 min at 4 °C. The cells were re-suspended (50 mL of 100 mM CaCl₂) and incubated for 30 min at 0 °C. The cells were centrifuged at 5,000 RPM for 10 min at 4 °C, and the cells were re-suspended (4 mL of 100 mM CaCl₂, 15% glycerol). The cells in suspension were aliquoted (0.1 mL) and were quickly frozen with liquid nitrogen.

C. Typical transformation of competent cells

Sterile conditions were maintained throughout the entire protocol. The competent cells were incubated at 0 °C for 5 min and the plasmid DNA (10 ng/μL) was added. The mixture was incubated for 30 min and the cells were heat shocked at 42 °C for 2 min at 0 °C. LB (antibiotic if necessary) was added and the cells were incubated at 37 °C for 1 h. Then the cells were transferred to a plate containing LB/antibiotic and were grown overnight at 37 °C.

D. Detection of 6xHis-tagged proteins (dot blot) using Anti-His HRP conjugates Qiagen® Kit

A protein solution (1-100 ng/μL) was applied directly onto a membrane. The membrane was washed twice for 10 min with TBS buffer (10 mM Tris-Cl, pH 8.0, 150 mM NaCl). The membrane was then incubated for 1 h in blocking buffer at room temperature. The membrane was washed twice for 10 min with TBS-Tween/Triton buffer (20 mM Tris-Cl, pH 8.0, 500 mM NaCl, 0.05% (v/v) Tween 20 and 0.2% (v/v) Triton X-100). The membrane was washed for 10 min with TBS buffer. Then the membrane was incubated in Anti-His HRP conjugate solution (1/1500 dilution of conjugate stock solution in blocking buffer). This was followed by a membrane washed (twice for 10 min) with TBS-Tween/Triton buffer and was washed for 10 min with TBS buffer. The membrane was stained with a solution containing 4-chloro-1-naphthol (18 mg), Tris (30 mL of solution (0.9% NaCl (w/v) in 0.1 M Tris-Cl pH 8.0) and methanol (6 mL). This was

followed by addition of hydrogen peroxide (6 mL, 30% solution) to start the catalysis of the to horseradish peroxidase. Then the membrane was rinsed twice with water and dried. The membranes were scanned as soon as possible because the colors fade with time.

E. Expression of BCDOX

The target gene (pBCDOX) was transformed into XL1-Blue cells according to standard protocols and plated on Ampicillin resistant LB plates. A single colony from the plate was inoculated (200 mL LB/carbenicillin broth (Carbenicillin 100 µg/mL)) and grown at 37 °C, while shaking, overnight. The overnight culture was inoculated (1 L of LB/1% glucose / Carbenicillin medium) and grown at 37 °C until $OD_{600} \approx 1.0$ (~ 3 h on average). The expression was induced by the addition of IPTG (1 mM) and $FeSO_4$ (55 µM final concentration). The culture was incubated at 26 °C for 6 h. The cells were harvested by centrifugation (9000 rpm, 30 min) and frozen at -20 °C overnight. The cell mass was thawed and re-suspended in Buffer A (20 mL for 1 L expression). The cells were lysed for 30 min at 0 °C with lysozyme (200 µL of 1 mg/mL), followed by sonication for 10 min at 4 °C. Then $MgSO_4$ (300 µL 1 mM) and benzoase (30 µL, of a 25 U/mL enzyme, Novagen) were added followed by incubation on ice for 30 min. The mixture was spun down (9000 rpm, 30 min) and the supernatant, that contained the soluble BCDOX was further purified. Some of the crude was saved for BCDOX activity assay. **Buffer A** (50 mM NaH_2PO_4 , 300 mM NaCl, 10 mM imidazole, pH=8)

F. Purification using a Histidine Purification Tag

The lysate that contained the histidine tag BCDOX protein, was applied to a nickel resin for purification (Novagen). The lysate was allowed to flow through the column, after which time it was washed (2x 10 mL) with the wash buffer (50 mM NaH_2PO_4 , 300 mM NaCl , 20 mM imidazole, pH=8). The protein was eluted with the elution buffer (50 mM NaH_2PO_4 , 300 mM NaCl , 250 mM imidazole, pH=8), and the fractions were then analyzed by SDS-PAGE. The fractions containing protein were then concentrated with Amicon Centriplus YM-10 filters, and buffer exchanged to a Tris- KOH buffer (Tris- KOH 10 mM buffer pH 8.0, 6 mM sodium taurocholate, 0.5 mM DTT 10 mM buffer pH 8.0, 6 mM sodium taurocholate, 0.5 mM DTT). Protein concentration was calculated via BCA protein assay.

G. Ni^{2+} Resin Regeneration

Resin is regenerated and recycled after each use by first stripping the column of all Ni^{2+} ions and anything bound to them. The column is washed with three column lengths of the strip buffer (100 mM EDTA, 0.5 M NaCl , 20 mM Tris-HCl, pH=7.9), followed by three column lengths of sterile distilled water. The column is then recharged with five column lengths of a fresh, clean Ni^{2+} source utilizing the charge buffer (50 mM NiSO_4). Finally, the column is equilibrated with three column lengths of binding buffer (50 mM NaH_2PO_4 , 300

mM NaCl, 10 mM imidazole, pH=8) to wash away any unbound Ni²⁺ ions. The column can be recycled as long as it retains a blue/green color after the cleaning.

H. Extended Ni²⁺ Resin Regeneration

When the flow rate of the column slows dramatically, or does not turn a strong blue/green after being washed with charge buffer, the following wash should be done (for a 2.5 mL column bed)

1. 5 mL of 6 M guanidine HCl, 0.2 M acetic acid
2. 5 mL sterile distilled water
3. 2.5 mL 2% SDS
4. 2.5 mL 25% EtOH
5. 2.5 mL 50% EtOH
6. 2.5 mL 75% EtOH
7. 12.5 mL 100% EtOH
8. 2.5 mL 75% EtOH
9. 2.5 mL 50% EtOH
10. 2.5 mL 25% EtOH
11. 2.5 mL sterile distilled water
12. 12.5 100 mM EDTA, pH=8.0
13. 7.5 mL sterile distilled water
14. 7.5 20% EtOH (for prolonged storage)

Note that the SDS solutions and the SDS wash steps cannot be done in the cold room as the detergent precipitates out of solution. Also, the column should be stored in 20% EtOH when not in use for long periods of time because bacteria can grow in the agarose resin.

I. Preparation of β -carotene micelles

A β -carotene solution was prepared by combining a β -carotene stock solution (875 mM in hexanes, 81 μ L) taurocholate (100 μ L) and Tween (3340 μ L, 4% w/v acetone). Then, the solution was dried and dissolved in buffer B (Buffer B,: 10 mL of Tris-KOH 10 mM buffer pH 8.0, 6 mM sodium taurocholate, 0.5 mM DTT). The β -carotene used for the assay must be purified using a pipette silica column eluted with hexanes. The eluted solution was concentrated under nitrogen, and redissolved in hexanes.

J. Assay for BCDOX, conversion of β -carotene to all *trans*-retinal

The final reaction mixture for the cleavage reaction contained β -carotene (7.1 μ M), Tris-KOH buffer (8 mM pH 8.0), sodium taurocholate (4.05 mM), DTT (0.34 mM), Tween 40 (0.3%), α -tocopherol (50 μ M), FeSO₄ (1.38 mM), and BCDOX (roughly 20 μ g of protein).

The assay was started by adding the protein to solution A (solution A: 540 μ L Tris-KOH buffer pH 8.0, 6 mM sodium taurocholate, 0.5 mM DTT, and 20 μ L of FeSO₄). The solution was pre-incubated for 5 min at 30 °C. Then the β -carotene micelles solution (160 μ L) was added and the mixture was then

incubated for 2-3 h at 30 °C. The reaction was stopped by adding formaldehyde ((200 μ L, 46% in water) and was incubated for another 10 min. The carotenoids and products were extracted sequentially with chloroform:methanol (1.5 mL, of a 1:2, v/v, 0.01% pyrogallol) and hexane (1.5 mL) and with chloroform:hexane (0.5 mL:1.5 mL). After addition of each organic solvent, the reaction mixture was mixed using a vortex for 40 seconds. The combined extracts were then dried under a stream of nitrogen gas, and dissolved in isooctane: toluene (200 μ L of (8/2, v/v) containing 0.01% butylhydroxytoluene).

K. HPLC analysis of the products obtained by a BCDOX activity

The products were analyzed using a normal phase column (C18), hexane: ethyl acetate (95:5) was used as a mobile phase at a flow rate of 1 mL/min. β -carotene was monitored at 450 nm, retinal at 370 and 325nm; and retinol at 325 nm. β -Carotene RT=10.5, retinal RT=3.0.

L. Calculation of Michelis Menten (K_M) constant (steady state).

Calculation of K_M and V_{MAX} for the isolated recombinant enzyme was performed. The formation of retinal was monitored, the enzymatic reaction was quenched at different times. Analysis of the data using a double reciprocal plot provided the values for K_M (μ M) and V_{MAX} (pmol retinal/min•mg), excess of substrate to mimic a steady state conditions was used.

M. Plasmid Purification:

Bacteria from a single colony was grown (500 mL, LB, containing appropriate antibiotics) and purified with via Qiagen® column purification. Upon completion of the Qiagen® purification, the recovered DNA was resuspended in (400 µL) sterile water, and analyzed by UV-visible for concentration determination and purity. The concentration of DNA obtained is µg/µL.

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

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

= 1.8, pure

>1.8, RNA contamination

< 1.8, protein contamination

N. Sample preparation for sequencing DNA

In a sterilized eppendorf tube (0.5 mL), DNA (2 µg) and primer (30 pmol) were mixed. When sequencing is performed, the largest amount of base pairs that are sequenced with accuracy is about 400 bp. Therefore, to sequence the complete gene, five different primers were used.

1. 5'-GATCTCGATCCCGCGAAATTAATACGAC-3'
2. 5'-CCAGACCCTAGAGACCTTGGAGAAGG-3'
3. 5'-CGAGGAGAAGTCCAGGCTGACC-3'
4. 5'-GATCGATCTCGATCCCGCG-3'
5. 5'-GCAGACTGGAATGCAGTGAAGC-3'

The DNA must be purified by a Qiagen[®] column and resuspended in sterile water prior to analysis. Failure to do this will result in poor sequence data. In addition, the sequence sample should be diluted with sterile water, not buffer.

O. Melting temperature calculation for primers:

Primers should ideally have melting points ≥ 78 °C, and end in at least one, if not more, GC base pair

$$T_m = 81.5\text{ °C} + (0.41)(\% \text{ GC}) - 675 / N - \% \text{ mismatch}$$

where N = primer length

P. Optimal PCR conditions

A small-scale gradient PCR was performed to optimize the extension temperature

PCR recipe	
Template DNA	100 ng
Primer 1	1 mM
Primer 2	1 mM
DNTP	200 μ M
Deep vent	2 U
10 x deep vent buffer	5 μ L (1x)
MgSO ₄	5 μ M
H ₂ O	50-Rxn μ L

PCR program		
1 X	94 °C	5 min
30 X	94 °C	1 min
	48 °C	3 min
	72 °C	3 min
1 X	72 °C	10 min
1 X	25 °C	10 min

Q. PCR Primers

	Primers Sequence	PCR Template
BCDOX-2	5'-CGGACATATGGAGATAATATTTGGCCAGA ATAAGAAAGAACAGCTGG-3' 5' -GACCCGGGTACCCTCTGGTGCTGTCGGATC-3'	pBCDOX-1
BCDOX-3	5'-GCACCAGAGGGTACCTAGGTG-3 5' CACCTAGGTACCCTCTGGTG-3'	pBCDOX-1

R. Purification of DNA from agarose gel

GENCLEAN Turbo Qiagen® Protocol: The desired band was excised from the agarose gel. The gel was cut into smaller pieces and transferred into an eppendorf tube. Then a GENCLEAN Turbo solution (100 µL per 100 µg of agarose) was added and the mixture was melted at 55 °C for 5 min. The melted solution was transferred into a GENCLEAN Turbo cartridge. The filter was spun for 5 seconds. The filter was washed twice by addition of GENCLEAN Turbo Wash (500 µL). The DNA was eluted by addition of GENCLEAN Turbo solution (30 µL). The DNA obtained was used directly in the enzyme digestions.

S. Digestions and ligation reactions

As mentioned previously, the double digestion did not afford the desired product. Therefore, subsequent single digestions were performed. Same protocols were followed for the gene (PCR product) and the pET-29b(+). The

first digestion was performed using *KpnI* followed by purification using GENECLAN. The second digestion was performed using *NdeI*.

The DNA was purified following the same procedure and quantified before the ligations were performed

Digestions				
	BCDOX		pET-29(+)	
KpnI	1 μ L	-	1 μ L	-
NdeI	-	1 μ L	-	1 μ L
DNA 30 μ L	30 μ L	30 μ L	30 μ L	30 μ L
Buffer 1 or 2 (10X)	1 X	1X	1 X	1X
H ₂ O	6 μ L	6 μ L	6 μ L	6 μ L

- Incubation at 37 °C for 2 h.

T. Calf Intestine Phosphatase treatment

CIP treatment		
	Digested pET-29b(+)	Digested BCDOX gene
DNA obtained from DNA purification	30 μ L	30 μ L
CIP 1 μ l	1U	1U
Buffer (10X)	1 X	1 X
H ₂ O	4 μ L	4 μ L

- Incubation at 37 °C for 2 h

U. Ligations.

Ligations	
BCDOX	4.5x10 ⁻¹³ moles
pET-29b(+)	4.5x10 ⁻¹⁴ moles
T4 ligase	1 μ l
Buffer (10X)	1 μ l
H ₂ O	2 μ l

- Incubation at 16 °C for 24 h
- JM109 *E. coli* were transformed with 5 μ L of the ligation reactions

V. Expression of pET-29b(+)

For over expression of pET-29b(+)-BCDOX (pBCDOX-1 or -2) the plasmid was transformed into an *E. coli* Rosetta(DE3)pLysS following the previously mentioned protocol. This strain was chosen because it supplies rare tRNAs, it carries a copy of the T7 lysozyme gene (for tight control), and it possesses a copy of the T7 RNA polymerase gene under the control of the *lacUV5* promoter, which allows a tight control for the over expression.

A single colony from the plate was inoculated (200 mL LB/Ampicillin-Cloramphenicol broth) and grown at 37 °C while shaking overnight. The overnight culture was re-inoculated (1 L of LB/1% glucose / Ampicillin-Cloramphenicol medium), the culture was grown at 37 °C until $OD_{600} \approx 1.0$ (~3 h on average). The expression was induced by addition of IPTG (0.4 mM) and $FeSO_4$ (55 μ M final concentration). The culture was incubated at various temperatures 4-32 °C for 6 h or overnight. The cells were harvested by centrifugation (9000 rpm, 30 min) and frozen at -20 °C overnight. The cell mass was thawed and re-suspended in **Buffer A** (20 mL for 1 L expression). The cells were lysed for 30 min at 0 °C with lysozyme (200 μ L of 1 mg/mL), followed by sonication for 10 min at 4 °C. At this point $MgSO_4$ (300 mL, 1 mM) and benzonase (30 μ L, 25 U/mL, Novagen) were added to reduce viscosity and the mixture was incubated on ice for 30 min. The mixture was spun down (9000 rpm, 30 min) and the supernatant and the insoluble pellet was analyzed by SDS PAGE gel. In all cases the over expressed BCDOX was found in the insoluble fraction as inclusion bodies.

W. Purification of BCDOX as inclusion bodies

Isolation of the inclusion bodies was accomplished by over expression of pET-29b(+)-BCDOX in Rosetta(DE3)pLysS induction by addition of IPTG (0.4 mM), OD₆₀₀~1.0 for 8 h at 26 °C. The soluble proteins were separated and the pellet was washed with a buffer containing 0.5% Triton X-100, followed by a buffer containing 2 M urea and 0.5% Triton X-100. The pellet obtained was finally resuspended into a buffer containing 8 M urea and DTT, or 6 M guanidine hydrochloride. It was observed that a smaller concentration of urea <6 M or <4 M guanidine did not solubilize the inclusion bodies. It was also noticed that addition of DTT increased the solubility. The inclusion bodies were purified almost with 90% purity and at around 8 mg per liter.

X. Refolding experiments

A 100 fold dilution was performed by adding **Buffer A** (100 mL to 1 mL of protein solution). The 100 dilution was performed using **Buffer A** (Buffer A: Tris-KOH 10 mM buffer pH 8.0, 6mM sodium taurocholate, 0.5 mM DTT. buffer and stirred for 24 h. 20 mL were concentrated ten times using Amicon Centriplus YM-30 filters, and the activity test showed an inactive protein.

Y. Expression of BCDOX *in vivo*

The expression of BCDOX in *E. coli* was monitored *in vivo* by the colorimetric change of the oxidation of β -carotene (orange) to retinal (yellow). The resultant *E. coli* cells had an orange coloration due to the accumulation of β -

carotene in the cells. In detail an *E. coli* XL1-Blue or ER2566 was transformed with pORANGE (plasmid that contains the genes necessary for the synthesis of β -carotene). The transformed cells were grown and they showed an orange phenotype. These cells were converted into competent cells and then transformed with the pBCDOX-1 or pBCDOX-2 plasmids. These transformed cells (cells that contained pORANGE and pBCDOX-1) were grown under LB/chloramphenicol, carbenicillin and 1% glucose (5 mL) to an OD of 1.00. Expression of BCDOX was then induced by addition of IPTG (1 mM final concentration). This culture was grown for 6 more h. The cells were spun down at 10000 RPM and rinsed with water. The shift of the coloration was caused by oxidation of β -carotene (orange) to retinal (yellow) by BCDOX. Where the cells had a yellow coloration the activity of BCDOX was confirmed. The same protocol was attempted using ER2566 *E. coli* instead of XL1-Blue but the transformed cells did not accumulate β -carotene.

2.14 References

1. R. A. Morton, *Photochemistry of Vision*. ed.; Springer-Verlag: New York, 1972; 'Vol.' p 33.
2. L. J. Gudas, *In The retinoids: Biology chemistry and medicine*. ed.; New York, 1994; 'Vol.' p 443-520.
3. A. C. Ross, "Symposium-retinoids-cellular-metabolism and activation-introduction." *Journal of Nutrition* **1993**, *123*, (2), 344-345.
4. P. Palozza, "Regulation of cell cycle progression and apoptosis by beta-carotene in undifferentiated and differentiated HL-60 leukemia cells: Possible involvement of a redox mechanism." *International Journal of Cancer* **2002**, *97*, 593-600.
5. L. J. Gudas, "Retinoids, retinoid-responsive genes, cell-differentiation, and cancer." *Cell Growth and Differentiation* **1992**, *3*, (9), 655-662.
6. A. C. Ross, "Cellular-metabolism and activation of retinoids-roles of cellular retinoid-binding proteins." *FASEB Journal* **1993**, *7*, (2), 317-327.
7. L. Stryer, "Vision: From photon to perception." *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, (2), 557-559.
8. H. O. Olson J. A., "The enzymatic cleavage of β -carotene into vitamin A by soluble exymes of rat liver and intestine." *Proceedings of the National Academy of Sciences of the United States of America* **1965**, *54*, 1364-1370.
9. D. S. Goodman; H. S. Huang, "Biosynthesis of vitamin A with rat intestinal enzymes." *Science* **1965**, *149*, 879-880.
10. A. Wyss; G. M. Wirtz; W. D. Woggon; R. Brugger; M. Wyss; A. Friedlein; G. Riss; H. Bachmann; W. Hunziker, "Expression pattern and localization of beta,beta-carotene 15,15 '-dioxygenase in different tissues." *Biochemical Journal* **2001**, *354*, 521-529.

11. A. Wyss; G. Wirtz; W. D. Woggon; R. Brugger; M. Wyss; A. Friedlein; H. Bachmann; W. Hunziker, "Cloning and expression of beta,beta-carotene 15,15 '-dioxygenase." *Biochemical and Biophysical Research Communications* **2000**, 271, (2), 334-336.
12. A. Takeda; T. Morinobu; K. Takitani; M. Kimura; H. Tamai, "Measurement of retinoids and beta-carotene 15,15 '-dioxygenase activity in HR-1 hairless mouse skin with UV exposure." *Journal of Nutritional Science and Vitaminology* **2003**, 49, (1), 69-72.
13. J. von Lintig; K. Vogt, "Filling the gap in vitamin A research - Molecular identification of an enzyme cleaving beta-carotene to retinal." *Journal of Biological Chemistry* **2000**, 275, (16), 11915-11920.
14. Qiagen, "Protein Expression System: QIAexpress Expression System." *The QIAexpressionist* **2000**, 1.
15. M. T. Redmond; S. Gentleman; T. Duncan; S. Yu; B. Wiggert; E. Gantt; S. Cuningham, "Identification, expression and substrate specificity of a mammalian beta-carotene 15-15'-dioxygenase." *Journal of Biological Chemistry* **2001**, 276, (9), 6560-6565.
16. von Lintig; A. Dreher; C. Kiefer; M. F. Wernet; K. Vogt, "Analysis of the blind *drosophila* mutant ninaB identifies the gene encoding the key enzyme for vitamin A formation in vivo." *Proceedings of the National Academy of Sciences of the United States of America* **2001**, 98, 1130-1135.
17. W. Yan; G.-F. Jang; F. Haeseleer; N. Esumi; J. Chang; K. Palcewski; D. Zack, "Cloning and characterization of a human beta-carotene." *Genomics* **2001**, 72, 193-202.
18. A. Lindqvist; S. Anderson, "Biochemical properties of purified recombinant human beta carotene 15-15'-monooxygenase." *Journal of Biological Chemistry* **2002**, 277, (26), 23942-23948.
19. Qiagen, "*Escherichia coli* host strains." *The QIA expressionist* **2000**, 15-16.
20. Pal-Bhowmick; Ipsita; K. Sadagopan; Vora; K. Hardeep; Sehgal; Alfica; Sharma; Shobhona; G. K. Jarori, "Cloning, over-expression, purification

- and characterization of *Plasmodium falciparum* enolase." *European Journal of Biochemistry* **2004**, 271, 4845-4854.
21. Qiagen, "Monoclonal anti-His antibodies for sensitive detection of 6xHis-tagged proteins." *QIAGEN News* **1997**, (3), 1.
 22. Qiagen, "QIAexpress Detection and Assay Handbook." **2000**, 11, 17-27.
 23. E. Pagge; v. Strandmann; C. Zoidl; H. Nakhein; R. Holewa; P. Lorenz; L. Hipab-Klein; G. U. Ryffel, "Highly specific and sensitive detection of 6xhist-tagged proteins using MRGS-His antibody." *Qiagen News* **1996**, (1), 9.
 24. P. L. Zamorano; V. B. Mahesh; L. M. D. Sevilla; L. P. Bhat; G. K. Bhat; D. W. Brann, "Expression and localization of the leptin receptor in endocrine and neuroendocrine tissues of te rat." *Neuroendocrinology* **1997**, 65, 223.
 25. Y. Zhang; L. Taiming; J. Liu, "Low temperature and glucose enhanced T7 RNA polymerase-based plasmid stability for increasing expression of glucagon-like peptide-2 in *Escherichia coli*." *Protein Expression and purification* **2003**, 29, (1), 132-139.
 26. J. A. Olson, "Provitamin-a function of carotenoids-the Conversion of beta-carotene into vitamin-A." *Journal of Nutrition* **1989**, 119, (1), 105-108.
 27. K. W. Walker; H. W. Gilbert, "Effect of redox environment on the *in vitro* and *in vivo* folding of RTEM- 1 beta-lactamase and *Escherichia coli* alkaline phosphatase." *The Journal of Biological Chemistry* **1994**, 269, (45), 28487-28493.
 28. J. L. Napoli; K. R. Race, "Biogenesis of retinoic acid from beta-carotene. Differences between the metabolism of beta-carotene and retinal." *The Journal of Biological Chemistry* **1988**, 263, 17372-17377.
 29. T. Ikemura, "Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes." *Journal of Molecular Biology* **1981**, 146, (1), 1-21.

30. H. Dong; I. Nilsson; C. G. Kurland, "Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates." *Journal of Molecular Biology* **1996**, 260, 649-663.
31. C. Kurland; J. Gallant, "Errors of heterologous protein expression." *Current Opinion in biotechnology* **1996**, 7, 489-493.
32. E. Goldman; A. H. Rosenberg; G. Zubay; F. W. Studier, "Consecutive low-usage leucine codons block translation only when near the 5' end of a message in *Escherichia coli*." *Journal of Molecular Biology* **1995**, 245, 467-473.
33. Strategene.
34. N. Misawa; Y. Stomi; K. Kondo; A. Yokoyama; S. Kajiwarra; T. Saito; T. Ohtani; W. Miki, "Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway proposed at the gene level." *Journal of Bacteriology* **1995**, 1995, 6575-6584.
35. E. T. Wurtzel; G. Valdez; P. Matthews, "Variation in expression of carotenoid genes in transfected *Escherichia coli* strains." *Bioresearch Journal* **1997**, 1, 1-11.
36. G. Sandmann, "Carotenoid biosynthesis and biotechnological application." *Archives of Biochemistry and Biophysics* **2001**, 385, 4-12.
37. M. Albrecht; N. Misawa; G. Sandmann, "Metabolic engineering of the terpenoid biosynthetic pathway of *Escherichia coli* for production of the carotenoids carotene and zeaxanthin." *Biotechnology Letters* **1999**, 21, 791-795.
38. F. W. Studier; B. A. Moffatt, "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." *Journal of Molecular Biology* **1986**, 189, 113-130.
39. A. H. Rosenberg; B. N. Lade; D. Churi; S. Lin; J. J. Dunn; F. W. Studier, "Vectors for selective expression of cloned DNAs by T7 RNA polymerase." *Gene* **1987**, 56, (1), 125-135.

40. F. W. Studier; A. H. Rosenberg; J. J. Dunn; J. W. Dubendorff, "Use of T7 RNA polymerase to direct expression of cloned genes." *Methods in Enzymology* **1990**, 185, 60-89.
41. *New England Biolabs Catalog and technical reference*. ed.; 2000-2001; 'Vol.' p 204.
42. Sambrook; Russel, *Molecular cloning*. 3rd. edition ed.; 1999; 'Vol.' p 159.
43. O. Seeburg; J. Shine; J. A. Martial; J. D. Baxter; H. M. Goodman, "Nucleotide sequence and amplification in bacteria of structural gene for rat growth hormone." *Nature* **1977**, 270, 486-494.
44. A. Ulrich; J. Shine; J. Chirgwin; R. Pictet; E. Tscher; W. J. Tutter; H. M. Goodman, "Rat insulin genes: construction of plasmids containing the coding sequences." *Science* **1977**, 196, 1313-1319.
45. F. A. O. Marston, "The purification of eukaryotic polypeptides synthesized in *Escherichia coli*." *Journal of Biochemistry* **1986**, 240, 1-12.
46. J. F. Kane; D. L. Hartley, "Formation of recombinant protein inclusion bodies in *Escherichia coli*." *Trends in Biotechnology* **1988**, 6, 95-101.
47. A. D. Guise; S. M. West; J. B. Chadhuri, *Molecular Biotechnology* **1996**, 653-664.
48. Blackwell; Horgan, "A novel strategy for production of a highly expressed recombinant protein in an active form." *FEBS Letters* **1990**, 295, 10-15.
49. N. Burton; B. Cavallini; M. Kanno; V. Moncollin; J. M. Egly, "Expression in *Escherichia coli*: purification and properties of the yeast general transcription factor TFIID." *Protein Expression and Purification* **1991**, 2, 432-441.
50. L. Lobel; S. Pollak; J. Klein; J. W. Lustbader, "High-level bacterial expression of a natively folded, soluble extracellular domain fusion protein of the human luteinizing hormone/chorionic gonadotropin receptor in the cytoplasm of *Escherichia coli*." *Endocrine* **2002**, 14, (2), 205-212.

51. W. A. Prinz; F. Åslund; A. Holmgren; J. Beckwith, "The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm." *Journal of Biological Chemistry* **1997**, 272, (25), 15661-15667.
52. E. L. Vallie; E. A. Diblasio; S. Kovacic; K. L. Grant; R. F. Schendel; J. M. McCoy, "A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *Escherichia coli* cytoplasm." *Bio/Technology* **1993**, 11, 187-193.
53. W. D. Woggon; R. M. Wenger, "Supramolecular chemistry and molecular recognition." *Chimia* **2000**, 54, (1-2), 9-12.
54. W. D. Woggon, "The central cleavage of beta,beta-carotene-A supramolecular mimic of enzymatic catalysis." *Chimia* **2000**, 54, (10), 564-568.
55. J. Paik; A. During; E. H. Jarrison; C. L. Mendelsohn; K. Lai; W. Blaner, "Expression and characterization of a murine enzyme able to cleave beta-carotene." *The Journal of Biological Chemistry* **2001**, 276, 32160-32168.
56. M. M. Altamirano; C. Garcia; L. D. Possani; A. R. Fersht, "Oxidative refolding chromatography: folding of the scorpion toxin Cn5." *Nature Biotechnology* **1999**, 17, 187-191.
57. Sambrook; Russel, "Molecular Cloning." **1999**, (8).
58. *PCR Metods applications*. ed.; 1992; 'Vol.' p 51-59.
59. M. R. Ponce; J. L. Micol, "PCR amplification of long DNA fragments." *Nucleic Acid Research* **1992**, 20, (3), 623.
60. R. Peek; R. W. L. M. Niessen; J. G. G. Schoenmakers; N. H. Lubsen, "DNA methylation as a regulatory mechanism in rat-crystallin gene expression." *Nucleic Acid Research* **1991**, 19, (1), 77-83.
61. E. D. Clark, "Protein refolding for industrial processes." *Current Opinion in Biotechnology* **2001**, 12, 202-207.

62. F. Baneyx, "Recombinant protein expression in *Escherichia coli*." *Current Opinion in Biotechnology* **1999**, *10*, 411-421.
63. T. G. Jeffrey; F. Baneyx, "Divergent effects of chaperone overexpression and ethanol supplementation on inclusion body formation in recombinant *Escherichia coli*." *Protein Expression and Purification* **1997**, *11*, 289-296.

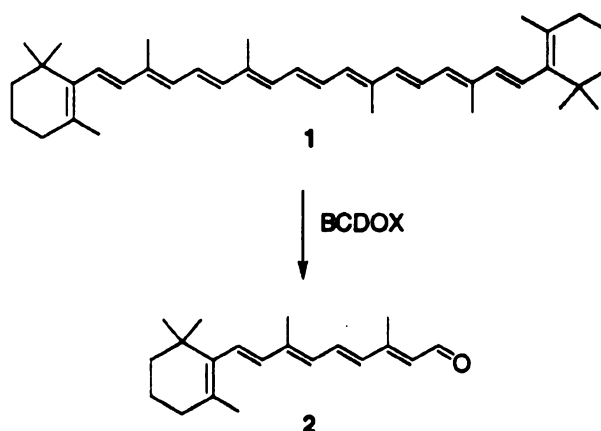
Chapter 3

Studies on BCDOX

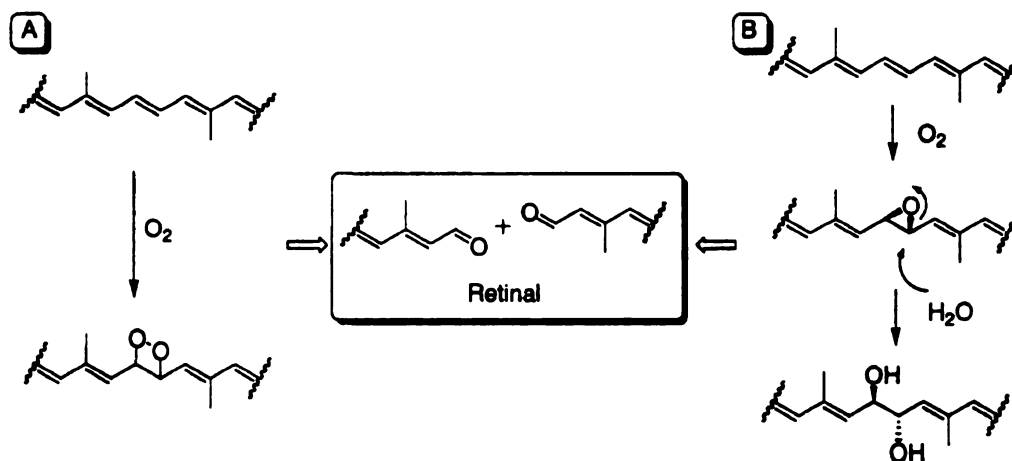
3.1 Introduction

Oxidative cleavage of β -carotene is an unprecedented process for a dioxygenase enzyme. Typically, dioxygenases require a hydrophilic handle such as a hydroxyl group to proceed.^{1,2} Enzyme/substrate interactions that lead to stereospecific catalytic action through hydrogen bonding and dipolar communication are common. However, the enzymatic mechanism of BCDOX is intriguing in that it uses a molecule of oxygen to cleave an electronically undistinguished double bond with a high degree of selectivity (Scheme 3-1). As mentioned previously,

the postulated mechanisms suggest either the involvement of a monooxygenase³ (Scheme 3-2A) or a dioxygenase⁴ (Scheme 3-2B). Despite the fact that it is known that the 15-15' double bond is selectively oxidized, the



Scheme 3-1. The enzymatic mechanism of BCDOX is intriguing in that it seemingly utilizes a molecule of oxygen to cleave an electronically undistinguished double bond with a high degree of selectivity.



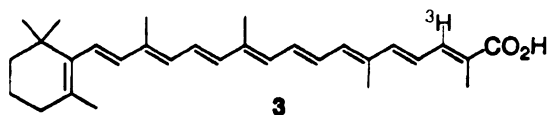
Scheme 3-2. Postulated mechanisms. A) Dioxygenase, suggest the involvement of a dioxetane intermediate. B) Monooxygenase, suggest the involvement of an epoxide intermediate.

mode of activation for inactivated double bonds is unknown. More importantly, the oxidation of a somewhat featureless hydrocarbon with such high regiospecificity is not well understood. Therefore, understanding the enzymatic action on hydrocarbon substrates could offer insight into nature's methodology for utilizing other molecular interactions for recognition and specificity.

Even though the sequence for BCDOX is known, there is no 3D structure available. Without any 3D structure there is no insight into the selective mechanism for the oxidation of the double bond. Homology studies suggest that different BCDOX proteins share overall sequence homology with a distinct pattern of highly identical conserved stretches. As mentioned previously, the highly conserved regions are located in the C-terminal part of the protein. Whether the conserved amino acids belong to the binding site of the iron or β -

carotene is not well understood. An alignment shows that around 8% of highly hydrophobic amino acids are conserved in the enzyme, suggesting that these amino acids participate in the formation of the substrate-binding pocket.⁵ Binding of the iron cation could be aided by 4 histidines and 6 acidic residues that are absolutely conserved. In addition, as mentioned previously, a region contains a very conserved arginine, a couple of histidines and a carboxylate and it is possible that this amino acids might be involved in the binding site for iron cation.⁶

Our interest in the mechanistic aspects of BCDOX has led us to synthesize the tritiated photoaffinity label [10'-³H]-8'-*apo*- β -carotenoic acid **3**. The choice of **3** as the photoaffinity probe stems from the fact that *apo*- β -carotenoic acids are bad substrates for BCDOX, and thus will bind the enzyme, its oxidation to form retinal is very slow.⁷⁻⁹ On the other hand, α,β -unsaturated polyene carboxylic acids are capable of photo-activation and can be used as photoaffinity probes. This was demonstrated by Rando and coworkers¹⁰ during the isolation of retinoic acid binding proteins in which radiolabeled retinoic acid was used as the photoaffinity labeling agent.¹¹ Photoaffinity labeling uses light induced chemical activation to form a covalent bond between a biological host (an enzyme or protein of interest) with a photoreactive substrate.¹²⁻¹⁴

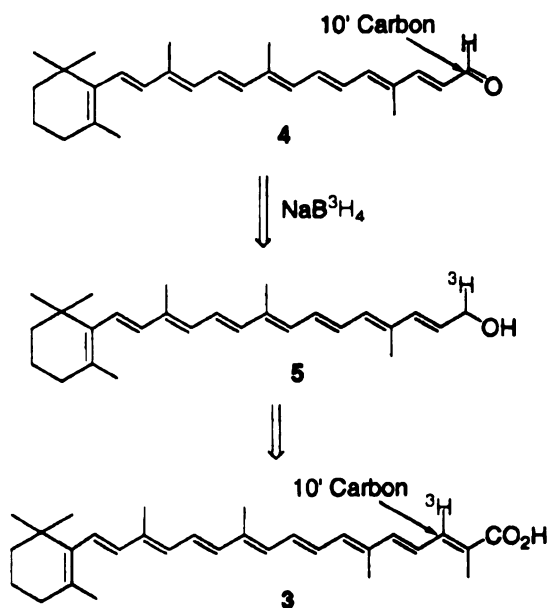


all-*trans*-[10'-³H]-8'-*apo*- β -carotenoic acid

3.2 Synthesis of all-*trans*-[10'-³H]-8'-apo-β-carotenoic acid

Synthesis of radiolabeled carotenoids are challenging for two reasons. First, the polyolefinic nature of these compounds leads to their less desirable stability, and thus introduction of radiolabels is generally limited to mild reactions. Although there is not a large body of literature of the synthesis of radiolabeled carotenoids (mostly biosynthesized by feeding of labeled precursors to various organism),^{15,16} related compounds such as retinoids have been synthesized by reduction of intermediate carbonyl group with Na/LiB³H₄, or semi-hydrogenation of intermediate acetylenes with ³H₂.¹⁷⁻²⁰ Secondly, their sensitivity to light forces the

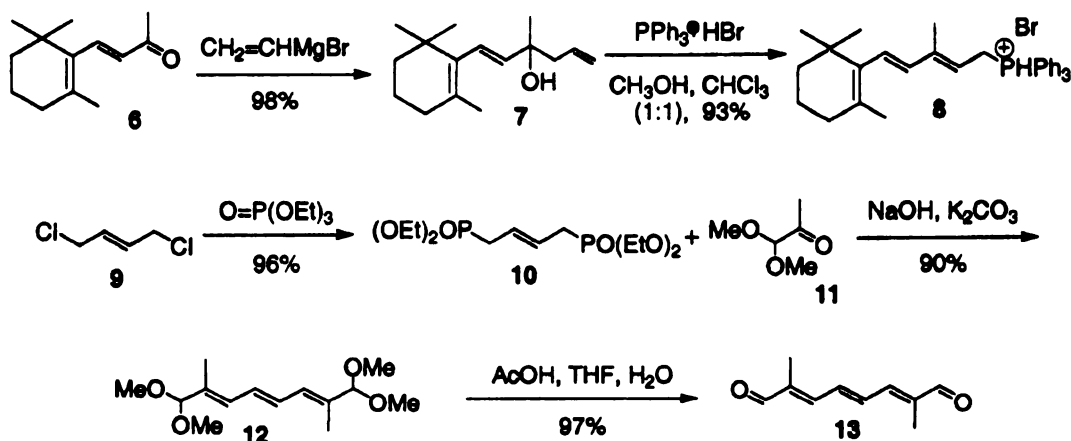
need to work in darkrooms with minimal red lighting and increases the difficulty of routine chemical manipulations. In the synthesis of any radiolabeled compound, it is preferable to incorporate the label as close to the final step as possible. This reduces the need for handling and purification of the radiolabeled intermediates, and increases the radioactive yield of the labeled compound. Although



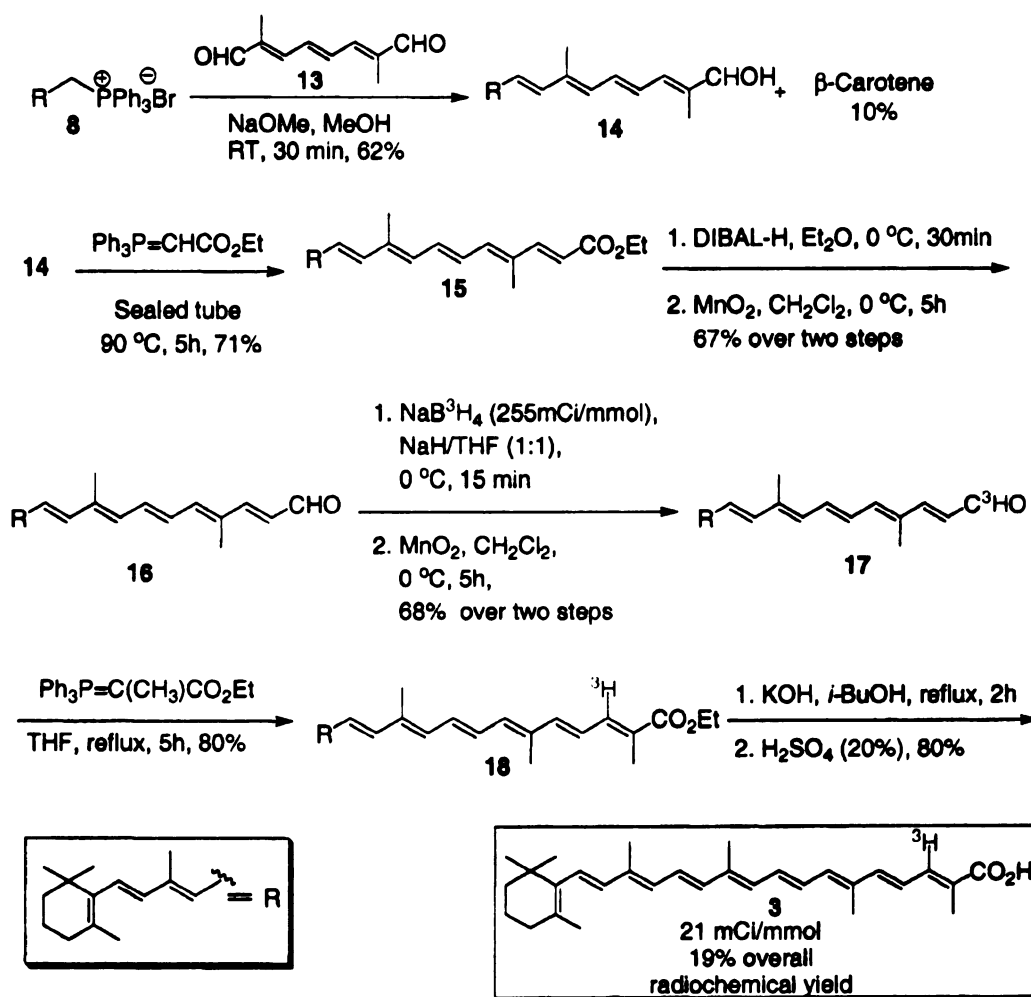
Scheme 3-3. 10' carbon was chosen as the site to introduce the radiolabel hydrogen. NaB³H₄ was chosen to be the source of tritium and thus required a reductive step to incorporate the label.

there is a range of positions present in 8'-*apo*- β -carotenoic acid **3** that could be considered as appropriate to incorporate tritium, carbon 10' was chosen as the site to introduce the radiolabel. NaB^3H_4 was chosen to be the source of tritium and thus required a reductive step to incorporate the label (Scheme 3-3). The subsequent steps could be performed in high yields, without the need for extensive manipulation and purification.

The synthesis of all-*trans*-[10'- ^3H]-8'-*apo*- β -carotenoic acid was accomplished as follows²¹ (Scheme 3-4). Construction of the C_{15} -ylide²² **8** starts with a Grignard reaction between β -ionone (**6**) and vinyl magnesium bromide to yield the α,β unsaturated alcohol **7** in quantitative yield. The alcohol **7** was reacted with triphenylphosphine hydrobromide to afford the C_{15} -ylide **8** in 93% yield. Construction of the 2,7-dimethyl-2,4,6-octatrienedial^{23,24} (**12**) was accomplished by reacting *trans*-1,4-dichloro-2-butene (**9**) with triethyl phosphate



Scheme 3.4. Synthesis of all-*trans*-[10'- ^3H]-8'-*apo*- β -carotenoic acid. Part I.



Scheme 3-5. Synthesis of all-*trans*-[10'-³H]-8'apo- β -carotenoic acid. Part II.

to yield 2-butenyl-1,4-bisphosphonic acid tetraethyl ester (**10**) in quantitative yield. Reaction of **9** with pyruvic aldehyde dimethyl acetal (**10**) afforded the protected aldehyde **11** in 90% yield. Deprotection of the acetal groups gave 2,7-dimethyl-2,4,6-octatrienedial (**12**) in 97% yield, Scheme 3-4.

Wittig olefination of **13** with ylide **8** was initiated with sodium methoxide in methanol, which delivered the C₂₅ aldehyde **14** in good yields with 97% *E*

stereoselectivity of the newly formed double bond. Treatment of aldehyde **14** with (carbethoxymethylene)triphenylphosphorane led to **15** (97% *E*) in 71% yield. The ethyl ester **15** was reduced with DIBAL, and the resulting allylic alcohol was oxidized with MnO_2 ²⁵ to provide C_{27} -aldehyde **16**. The tritium label was incorporated by the reduction of aldehyde **16** with NaB^3H_4 at 0 °C, which was immediately reoxidized with MnO_2 to yield the radiolabeled C_{27} -aldehyde **17**. Wittig olefination of **17** with (carbethoxyethylidene)triphenylphosphorane proceeded well to deliver ester **18** as the sole product (*Z* isomer was not detected by NMR spectroscopy). Hydrolysis of the ethyl ester **18** provided the radiolabeled acid **3** (21 mCi/mmol) with a 19% overall radiochemical yield (Scheme 3.5).

3.3 Enzymatic studies using of all-*trans*-8'-*apo*- β -carotenoic acid and BCDOX

The obtained 8'-*apo*- β -carotenoic acid was used to perform enzymatic studies with BCDOX. The initial studies were performed with non-labeled 8'-*apo*- β -carotenoic acid. A typical assay contained an initial concentration of 200 μM of 8'-*apo*-carotenoic acid, and 20 μg of BCDOX. The enzymatic reaction was stopped at different times and HPLC was used to monitor the formation of retinal. A normal phase column (C5), hexane: ethyl acetate (95:5) was used as a mobile phase at a flow rate of 1 ml/min. The results showed that **3** was a substrate

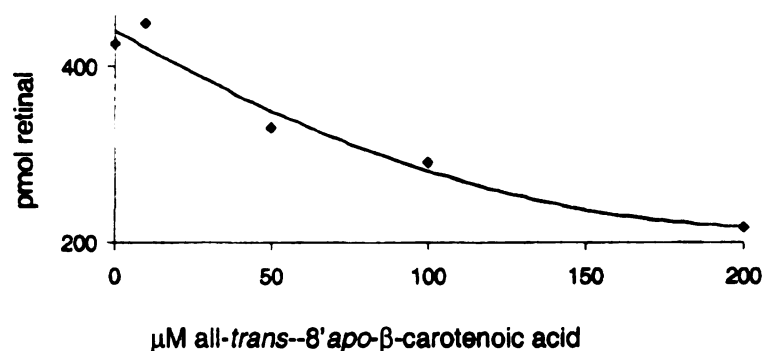


Figure 3-1. Inhibition studies using all-*trans*-[10'-³H]-8'-*apo*-β-carotenoic acid. This graph shows that in presence of β-carotene, the more all-*trans*-8'-*apo*-β-carotenoic acid is present, the less amount of retinal is formed.

with an activity of 2.1 pmol retinal/mg of BCDOX•min. When only β-carotene is used as a substrate, the obtained V_{MAX} was 194 pmol retinal/mg of BCDOX•min. These results agree with previously reported data in which *apo*-carotenals and *apo*-carotenoic acids are substrates for BCDOX, but its oxidative cleavage at the 15-15' bond is not as efficient as β-carotene.⁷⁻⁹ Therefore, **3** can be considered as a slow substrate for BCDOX. Inhibition studies with **3** were performed to investigate if the substrate was binding the same active site as β-carotene. Different concentrations of *apo*-carotenoic acid were tested. First, the BCDOX was incubated with the *apo*-carotenoic acid for 15 minutes, and then β-carotene was added. The outcome showed as expected that 8'-*apo*-carotenoic acid inhibits the oxidation of β-carotene to retinal (Figure 3-1). The fact that the 8'-*apo*-

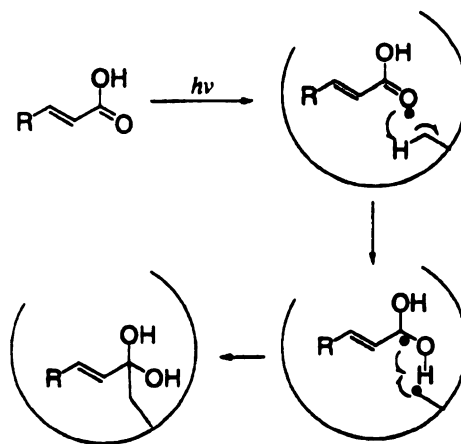
carotenoic acid binds the site of oxidation in BCDOX and is a slow substrate for the oxidation of BCDOX satisfies one of the requirements for the development of a photoaffinity probe for BCDOX.

3.4 Failed photolabeling attempts using all-*trans*-8'-apo- β -carotenoic acid and BCDOX

Preliminary crosslinking studies using [10'-³H]-8'-apo- β -carotenoic acid were performed to explore the viability of this method. All the experiments were performed under red safe-light illumination. A mixture of the protein and [10'-³H]-8'-apo- β -carotenoic acid was incubated for 15 minutes and then irradiated with a long-wave UV light (360 nm) at 0 °C for 30 minutes. For controls, three more samples were prepared. A sample containing the same concentration of [10'-³H]-8'-apo- β -carotenoic acid and β -carotene, a sample with only β -carotene, and a third sample with [10'-³H]-8'-apo- β -carotenoic acid that was not subjected to irradiation. Following irradiation, each sample was subjected to electrophoresis to remove [10'-³H]-8'-apo- β -carotenoic acid that did not cross-linked to BCDOX. The band that corresponded to BCDOX was separated and the radioactivity present in each sample was measured.

The same radioactivity was found in all the samples, suggesting that no cross-linking between the BCDOX and the acid had occurred. In the first attempt, the buffer used contained α -tocopherol but its presence could interfere with the mechanism of generation of the photolabel active species of α,β unsaturated acids. It has been postulated that α,β unsaturated acids cross-link via a di-radical

intermediate, which first abstract a hydrogen radical from the active site and then the two newly created radicals combine rapidly (Scheme 3-6). It is known that α -tocopherol²⁶ is a radical quencher, therefore, the experiment was repeated with out α -tocopherol, but the results obtained were the same. It is important to notice that the final concentration of the enzyme was only 1.2



Scheme 3-6. The proposed mechanism for cross linking of α,β -unsaturated acids involves a diradical intermediate

μM and attempts to concentrate the sample further resulted in loss of activity. Whether the loss of activity was due to mechanical manipulation or protein aggregation was not clear. Consequently, it is possible that low concentration of protein resulted in no crosslinking of the protein with the $[10'\text{-}^3\text{H}]\text{-8'-apo-}\beta\text{-carotenoic acid}$. As mentioned previously α,β unsaturated acids have been used extensively in photoaffinity labeling, however, their efficiency to cross-link to a neighboring amino acid is not very high.¹³ Further optimization should be performed to determine whether it is the α,β unsaturated acid photo labeling group that is hampering the cross linking process, or if it is the low concentration of the protein.

As discussed previously in Chapter 2, attempts to express BCDOX in large quantities in bacteria were not successful. The lack of large amounts of protein has hampered further photoaffinity labeling studies.

3.5 Attempts to develop a rapid and efficient assay for BCDOX

As a consequence of the tedious extraction and the need for HPLC analysis required for each BCDOX assay, the collection of kinetic data and mechanistic studies have been difficult. There are two major problems with the assays performed using HPLC to identify the products. First, for every assay the carotenoids and retinoids have to be extracted from the aqueous solution and then analyzed. Second, when the metabolites are extracted, non-selective oxidation could occur resulting in false positives. Thus, an attempt to develop a new sensitive and efficient assay for BCDOX was performed. The assay designed exploits the difference in polarity of β -carotene and retinal. A partition assay consists of the distribution of a known substance between two immiscible solvents (partition coefficient K , is defined as the ratio of equilibrium concentration of the substance in the octanol rich phase (A_{S1}) to the water rich phase (A_{S2}) $K=A_{S1}/A_{S2}$; $\log P$ is the octanol-water partition coefficient). The assay will contain the immiscible solvents water and isooctane. Due to differences in their polarity the β -carotene and retinal will have different distributions in those solvents (Figure 3-2). The partition coefficients ($\log P$) of β -carotene, retinal, retinol and retinoic acid are 17.62, 7.60, 5.68 and 6.30, respectively.²⁷ Thus, β -carotene has a strong affinity (solubility) for isooctane and low affinity for water. However, solubility of retinoids in water is also poor. To overcome this limitation, different co-solvents were used (acetonitrile, dimethyl sulfoxide, 2-chloroethanol, 4-chlorobutyronitrile and ethanolamine). These solvents are soluble in water and

insoluble in isooctane. They should increase the solubility of the desired metabolites into the aqueous phase while at the same time not increase the solubility of β -carotene into the aqueous phase. BCDOX oxidizes β -carotene into retinal, thus, the partition assay would extract only retinal into the aqueous phase, and β -carotene into the organic phase

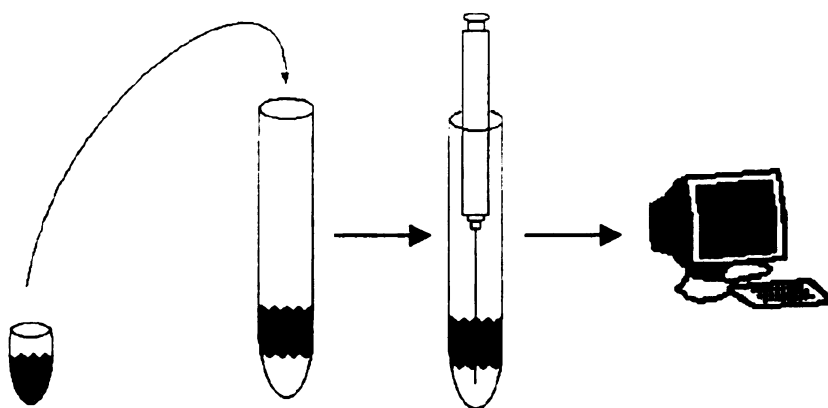
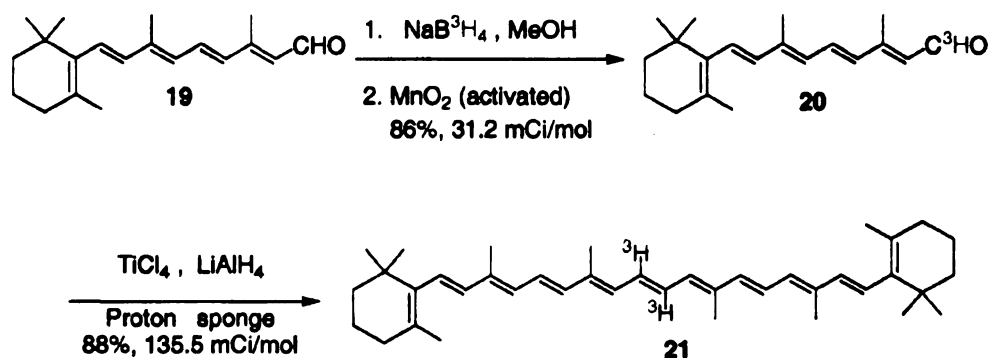


Figure 3-2. Development of a facile assay to test BCDOX activity. Partition assay consists of the distribution of a known substance between two immiscible solvents. The assay will contain the immiscible solvents water and isooctane, due to the differences in polarity of β -carotene and retinal they will have different distribution in those solvents. β -Carotene has a strong affinity (solubility) for isooctane and low affinity for water. A co-solvent is added in the aqueous phase to increase the solubility of retinoids in water.

To increase the sensitivity and efficiency of this assay, radioactive materials were used. A small aliquot from the enzymatic assay was added to a tube containing both aqueous and organic phases. The β -carotene and the retinal would then distribute differently into the two phases and finally each phase would be quantified by measuring the radioactivity. To establish the feasibility of this assay first *trans*-[15,15'- ^3H]- β -carotene (**21**), all-*trans*-[15- ^3H]-retinal (**20**) and all-*trans*-[15- ^3H]-retinol were prepared.

Radioactive all-*trans*-[15,15'- ^3H]- β -carotene (**21**) and all-*trans*-[15- ^3H]-retinal (**20**) were synthesized as shown in Scheme 3-7. The all-*trans*-retinal (**20**) was reduced with NaB^3H_4 to produce all-*trans*-[15- ^3H]-retinol. The oxidation of the labeled retinol with activated MnO_2 ²⁵ yielded all-*trans*-[15- ^3H]-retinal (**20**). McMurry coupling²⁸⁻³⁰ of **20** afforded the radioactive all-*trans*-[15,15'- ^3H]- β -carotene (**21**).



Scheme 3-7. Synthesis of ^3H labeled β -carotene and retinal.

The first step towards the development of the partition assay was to develop a system that extracts all-*trans*-[15-³H] retinal into the aqueous phase without contamination of all-*trans*-[15,15'-³H]- β -carotene. After a series of different extractions it was established that the optimum volume to work with was 1 mL of isooctane (organic phase), 1 mL of the aqueous phase (different solvents) and 0.1 mL of water or buffer. The optimal extractions of β -carotene, retinal and retinol with different solvents in the aqueous phase are shown in Figure 3-3. The graph shows that acetonitrile (Figure 3-3) is the best solvent to use in the aqueous phase. Acetonitrile hardly extracts β -carotene in the aqueous phase (1.8%) while extracting 43.5% of retinal in the aqueous phase. Unfortunately when the exact assay conditions were used, due to the presence of Tween (detergent) the

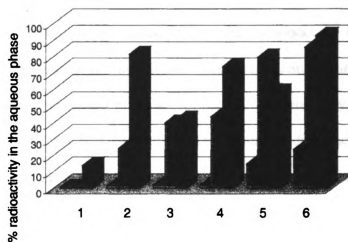


Figure 3-3. Extraction of β -carotene (), retinal (•), retinol (◐).

1. Water, 2. DMSO, 3. Ethanolamine, 4. Acetonitrile, 5. 2-Chloroethanol, 6. 4-Chlorobutyronitrile.

extraction of retinal in the aqueous phase was decreased to only 18.3%. To improve the percentage of retinal soluble in the aqueous phase, the assay was quenched with hydroxylamine. Hydroxylamine would form a Schiff Base with the aldehyde, thus increasing its solubility. This indeed increased the percentage of radioactivity in the aqueous phase to about 40%. However, this result is not very accurate because addition of hydroxylamine resulted in the formation of three phases. Thus an accurate measurement to quantify the concentration of retinal was not possible.

The assay was tested using BCDOX and labeled [15,15'-³H] β -carotene. As discussed in Chapter 2 the V_{MAX} for the BCDOX used was 194pmol retinal/mg•min. The final mixture contained β -carotene 7.1 μ M, and of [15,15'-³H]- β -carotene (1 μ L, 0.4 mM, 136 mCi/mmol). The assay was performed as previously explained in Chapter 2. After the assay was quenched with formaldehyde, 100 μ L were added into 1 mL of isooctane and 1 mL of acetonitrile. The amount of tritium in each phase was quantified, and only about 2 % of radioactivity was found in the aqueous phase. Unfortunately, the difference between 2.0 and 1.8% was not sufficient to confirm a difference between the retinal formed by the enzyme and the β -carotene extracted by the aqueous phase. As mentioned previously, attempts to concentrate the protein were unsuccessful. Thus further optimizations should be performed to optimize the extraction of retinal to the aqueous phase. The activity of BCDOX is pretty low compared to most enzymes. It is possible that such low activity of BCDOX originates from the fact that the product obtained is retinal. As mentioned previously retinal is the

major source of retinoic acid *in vivo*. The concentration of retinoic acid is tightly controlled³¹⁻³⁴ thus it is possible that BCDOX is the first step to control the amount of retinal available in the cells. BCDOX is expressed in high amounts in duodenal vile, lung, kidney, intestine, brain, stomach and testis.³⁵⁻³⁸ All these tissues require retinoids as cofactors to regulate a variety of functions.

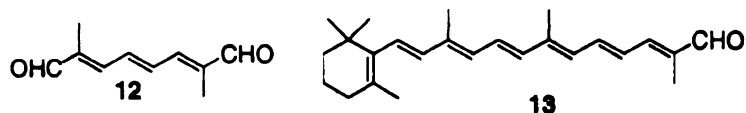
Due to the fact that expression of large quantities of BCDOX was not feasible using bacterial systems, progress in the development of a feasible assay was slowed down. Until a successful system to express BCDOX in large scale is established, a successful development of an efficient assay as well as a good photolabile probe cannot be achieved.

3.6 Materials and Methods

A. Synthesis of Compounds

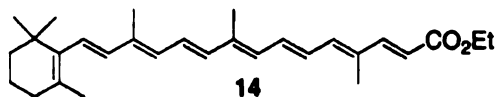
All reactions were carried under an atmosphere of nitrogen and removal of solvents was performed under reduced pressure with a Buchi rotatory evaporator. THF and Et₂O were freshly distilled from sodium/benzophenone, and CH₂Cl₂ was distilled over CaH₂ under a nitrogen atmosphere. Radioactive NaB³H₄ was purchased from American Radiolabeled Chemicals, Inc (St. Louis, MO). Analytical TLC was carried out using Merck 250 mm Silica gel 60 F₂₅₄ and spots were visualized under UV light. Column chromatography was conducted using Silicycle silica gel (230-400 mesh). 300 MHz ¹H-NMR and 75 MHz ¹³C-NMR spectra were recorded on a Varian Gemini-300 or 500 instrument, and the residual protic solvent (CDCl₃ or DMSO-d₆) was used as internal reference. UV-visible spectra were recorded on a Perkin-Elmer Lambda 40 spectrometer. A Wallac WinSpectral 1414 liquid scintillation counter was used for quantification of radioactivity. In a typical measurement 1-10 mL of sample was added to 3 mL of Optipase 'HiSafe' 3 liquid scintillation cocktail (Wallac) and the solution was counted for 1 min. All the reactions were carried out in a darkroom under minimal photographic safety red lights.

1. Synthesis of 2,7,11-Trimethyl-14-(2,6,6-trimethyl-cyclohex-1-enyl)trideca-2,4,6,8,10,12 hexaenal (**13**)



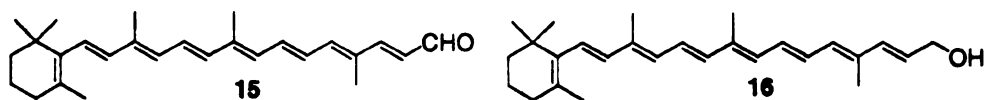
Sodium methoxide (121 mg, 2.250 mmol, 1.5 eq.) was added to a solution of **12**²⁴ (817 mg 1.5 mmol) in dry methanol (8 mL) at 0 °C, and stirred until the phosphonium salt was completely converted to the deep red phosphorane. A solution of 7-dimethyl-2,4,6-trienedial²³ (164 mg, 1 mmol) in dry methanol (1 mL) was added and the mixture was stirred for 30 minutes. The reaction mixture was diluted with water (10 mL) and extracted with diethyl ether (3 x 5 mL). The organic layer was dried over Na₂SO₄ and the solvent was removed in *vacuo* to give the C₂₅-aldehyde **13** (217 mg, 6 2%, 97:3 E:Z), and β-carotene **1** (57 mg, 10%), which was purified by flash silica chromatography (2% ethyl acetate in hexane). [¹H NMR (300 MHz CDCl₃): δ 9.4(1H, s), 7.03(1H, d, *J*=12 Hz), 6.96(1H, dd, *J*=14.4, 12.3 Hz), 6.77(1H, dd, *J*=14.4, 11.4 Hz), 6.70(1H, d, *J*=12.3 Hz), 6.50(1H, d, *J*=11.7 Hz), 6.41(1H, dd, *J*=14.4, 11.7 Hz), 6.12(3H, m), 2.00(3H, s, CH₃), 1.97(3H, s, CH₃), 1.86(3H, s, CH₃). ¹³C NMR (300 MHz CDCl₃): δ 194.38, 148.94, 141.72, 137.88, 137.71, 137.42, 136.66, 136.22, 130.73, 130.30, 129.74, 127.74, 127.74, 127.63, 127.16, 39.51, 34.17, 33.03, 28.89, 26.81, 21.70, 19.13, 12.97, 12.77, 9.51. UV: λ_{max} (hexanes) 414 nm

2. Synthesis of 4,9,13-trimethyl-15-(2,6,6-trimethyl-cyclohex-1-enyl)-pentadeca-2,4,6,8,10,12,14-heptaenoic acid ethyl ester **14**



A solution of **13** (110 mg, 0.31 mmol) and (carbethoxymethylene)-triphenylphosphorane (323 mg, 0.93 mmol) in dry THF (4 mL) was heated at 95 °C for 5 hours in a sealed tube. The reaction mixture was cooled to room temperature and diluted with water (10 mL) and Et₂O (10 mL). Phases were separated and the aqueous phase was extracted with Et₂O (2 x 10 mL). The combined organics were dried over Na₂SO₄, filtered and concentrated under reduced pressure. Crude product was purified by flash silica chromatography (5% ethyl acetate in hexane) to yield **14** (93 mg, 71%, >97% E). The coupling constant for the newly formed double bond confirms the E stereochemistry. The Z isomer was not detected by NMR spectroscopy. [¹H NMR (300 MHz CDCl₃): δ 7.35(1H, d, *J*=15.6 Hz), 6.79(1H, dd, *J*=14.7, 15.6 Hz), 6.78(1H, d, *J*=12.9 Hz), 6.65(1H, dd, *J*=13.5, 11.7 Hz), 6.45(1H, d, *J*=11.1 Hz), 6.33(1H, d, *J*=14.7 Hz), 6.23(1H, d, *J*=11.7 Hz), 5.96(3H, m), 5.84(1H, d, *J*=15.6 Hz), 4.20(2H, q, CH₂, *J*=7.2 Hz), 2.01(2H, m, CH₂), 1.97(3H, s, CH₃), 1.96(3H, s, CH₃), 1.89(3H, s, CH₃), 1.69(3H, s, CH₃), 1.60(2H, m, CH₂), 1.45(2H, m, CH₂), 1.28(3H, t, OCH₂CH₃, *J*=7.2 Hz), 1.00(6H, s, 2 x CH₃). ¹³C NMR (300 MHz CDCl₃): δ 167.51, 148.72, 139.25, 138.93, 137.79, 137.58, 136.91, 136.73, 133.74, 133.36, 131.57, 130.56, 129.54, 128.60, 127.16, 126.28, 116.23, 60.17, 39.55, 34.20, 33.07, 28.92, 21.75, 19.18, 14.31, 12.90, 12.76, 12.53. UV: λ_{max} (hexanes) 423 nm]

3. Synthesis of 4,9,13-trimethyl-15-(2,6,6-trimethyl-cyclohex-1-enyl)-pentadeca-2,4,6,8,10,12,14-heptaenal **15**

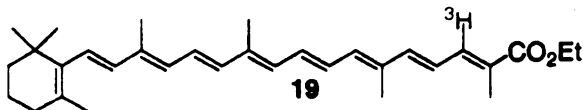


To a cold (0 °C) solution of **14** (90 mg 0.115 mmol) in dry Et₂O (6 mL) was added 1 M solution of DIBAL in cyclohexane (0.43 mL, 0.43 mmol) and stirred for 30 minutes. Saturated sodium-potassium tartrate solution (3 mL) and glycerol (5 drops) were added. The reaction mixture was stirred over night at room temperature; phases were separated and the aqueous phase was extracted with Et₂O (3 x 10 mL). The combined organic layers were washed with brine (3 mL), dried over Na₂SO₄, filtered and concentrated in *vacuo* to give the corresponding allylic alcohol, **16** which was pure enough to be used without further purification (80 mg, 90%). **Spectra for 16** [¹H NMR (300 MHz CDCl₃): δ 6.67(3H, m), 6.22(7H, m), 5.84(1H, dt, *J*=12.0 *J*=6.0 Hz), 4.22(2H, t, *J*=6.0 Hz), 2.00(2H, m, CH₂), 1.94(6H, s, 2 x CH₃), 1.89(3H, s, CH₃), 1.70(3H, s, CH₃), 1.60(2H, m, CH₂), 1.45(2H, m, CH₂), 1.00(6H, s, 2 x CH₃). ¹³C NMR (300 MHz CDCl₃): δ 137.82, 137.68, 137.06, 136.68, 136.27, 136.09, 134.62, 132.32, 131.96, 130.68, 130.35, 129.35, 127.28, 126.66, 125.14, 63.91, 39.55, 34.21, 33.05, 28.92, 21.73, 19.20, 12.82, 12.78, 12.71. UV: λ_{max} (hexanes) 400, 422 nm]

MnO₂ (100 mg, 1.16 mmol, 20 eq.) was added to a stirred solution of the aforementioned allylic alcohol **16** (22 mg 0.6 mmol) in CH₂Cl₂ (3 mL) at 0 °C and stirred for 5 hours and stirred for 5 hours. The reaction mixture was filtered through a pad of celite and washed with CH₂Cl₂ (4 x 0.5 mL). Solvent was

dried over Na_2SO_4 . Removal of solvent under reduced pressure gave the corresponding allylic alcohol **17** (13 mg, 93%). The allylic alcohol **17** was used without further purification. After removal of solvent under reduced pressure, the crude residue was dissolved in CH_2Cl_2 (2 mL), freshly prepared MnO_2 (58 mg, 20 eq. 0.68 mmol) was added at 0 °C, and the reaction was stirred for 5 h. The suspension was filtered through a pad of celite, washed with CH_2Cl_2 (1 x 2 mL), and concentrated under reduced pressure followed by chromatographic purification over silica gel (5% ethyl acetate in hexanes) to yield pure aldehyde **18** (9.6 mg 73%). The latter yield refers to product obtained from non-radioactive syntheses. The radioactive material was used without purification following extraction and removal of CH_2Cl_2 in the next step and the amount isolated was assumed to be the same as for the non-radioactive reaction.

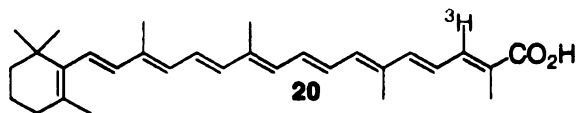
5. Synthesis of 2,6,11,15-Tetramethyl-17-(2,6,6-trimethyl-cyclohex-1-enyl)-pentadeca-2,4,6,8,10,12,14,16-octaenoic acid ethyl ester **19**



(Carbethoxyethylidene)triphenylphosphorane (18 mg 0.05 mmol) was added to a solution of **18** (9.6 mg, 0.025 mmol) in THF (3.5 mL) and the reaction was refluxed for 5 h. The reaction mixture was diluted with water (4 mL) and extracted with Et_2O (3 x 2 mL). The combined ether extracts were dried over Na_2SO_4 . The solvent was removed under reduced pressure and the crude product

was purified by flushing it through a small pad of silica in a Pasteur pipette (5% ethyl acetate in hexane) to yield ethyl ester **19** (9.4 mg, 80%). During the preparation of the radioactive material, the crude product was also flushed through a small pad of silica, however, it was not isolated and characterized, but was used directly in the next step. [^1H NMR (300 MHz CDCl_3): δ 7.25(1H, d, $J=11.6$ Hz), 6.59(4H, m), 6.40(2H, m), 6.25(1H, d, $J=11.1$ Hz), 4.21(2H, q, $J=7.2$ Hz), 2.00(2H, m, CH_2), 1.97(3H, s, CH_3), 1.95(3H, s, CH_3), 1.69(6H, s, 2 x CH_3), 1.60(2H, m, CH_2), 1.54(3H, s, CH_3), 1.45(2H, m, CH_2), 1.28(3H, t, OCH_2CH_3 , $J=7.2$ Hz), 1.00(6H, s, 2 x CH_3). ^{13}C NMR (300 MHz CDCl_3): δ 168.52, 143.84, 138.74, 137.83, 137.80, 137.65, 136.94, 136.53, 135.71, 135.39, 131.99, 131.91, 130.67, 129.48, 129.32, 126.15, 125.72, 123.14, 60.46, 39.58, 34.22, 33.07, 29.65, 28.92, 21.73, 19.19, 14.32, 12.86, 12.81, 12.75, 12.68. UV: λ_{max} (hexanes) 440 nm]

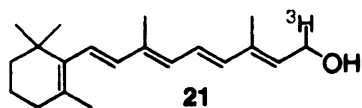
6. Synthesis of [$10'$ - ^3H]-8'-apo- β -carotenoic acid **20**



KOH (25%, 2 mL) was added to a solution of **19** (9.4 mg, 0.02 mmol), in *iso*-butyl alcohol (3 mL). The resulting mixture was heated at 95 °C for 2 h and cooled to room temperature. H_2SO_4 (20%, 2 mL) was added and the reaction mixture was again heated to 70 °C for 30 min. The organic layer was separated and washed with hot water (3 x 1 mL). Dichloromethane (5 mL) was added to the extracted *iso*-butyl alcohol resulting in two layers: the desired acid was extracted

into CH₂Cl₂. The CH₂Cl₂ layer was separated and dried over Na₂SO₄, and concentrated under reduced pressure. The crude product was purified through a small pad of silica in a Pasteur pipette (5-15% ethyl acetate in hexane) to yield pure acid (7 mg, 80% 21 mCi/mmol, 0.34 mCi, 19% of total radionuclei incorporated) based on comparative TLC, and UV with authentic cold material, and all spectroscopic data were consistent with reported values.

7. Synthesis [15-³H]-retinol **21**

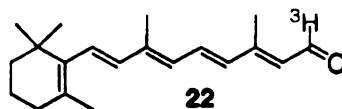


Retinal (15 0.04 mmol) was added to a cold (0 °C) solution of NaB³H₄ (1.8 mCi 265 mCi/mmol) in THF/MeOH (4 mL, 1:1). The reaction mixture was stirred for 20 min at which time unlabeled NaBH₄ (1.5 mg, 0.04 mmol) was added) to the reaction mixture (stirred for 15 min). The reaction was quenched by addition of saturated NH₄Cl and extracted with Et₂O (3 x 2 mL). Phases were separated and the organic layer was washed with sat. NaCl, and dried over Na₂SO₄. Removal of solvent under reduced pressure gave the corresponding retinol **21** (87%). 15-[³H]-retinol, 88 mCi/mmol. [¹H NMR (300 MHz CDCl₃): δ 6.60(1H, dd, *J*=15.38, 14.83 Hz), 6.27(1H, d, *J*=15.38 Hz), 6.15(1H, d *J*=13.19 Hz), 6.10 (2H, m), 5.67(1H, t, *J*=7.14), 4.29(3H, d, *J*=6.59), 2.30(3H, s, CH₃), 1.97(3H, s, CH₃), 1.67(3H, s, CH₃), 1.85(3H, s, 2 x CH₃). ¹³C NMR (300 MHz CDCl₃): δ 137.82, 137.06, 136.27, 134.62, 132.32, 130.68, 129.35, 127.28,

126.66, 125.14, 64.51, 39.55, 34.21, 28.92, 21.73, 19.20, 12.82, 12.78, 12.71.

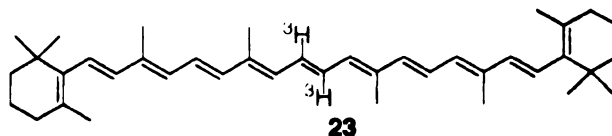
UV: λ_{max} (hexanes) 325 nm]

8. Synthesis [15- ^3H]-retinal **22**



The retinol **21** was used without further purification. After removal of solvent under reduced pressure, the crude residue was dissolved in CH_2Cl_2 (2 mL), freshly prepared MnO_2 (58 mg, 20 eq., 0.68 mmol) was added at 0 °C, and the reaction was stirred for 5 h. The suspension was filtered through a pad of celite, washed with CH_2Cl_2 (1 x 2 mL), and concentrated under reduced pressure followed by chromatographic purification over silica gel (5% ethyl acetate in hexanes) to yield pure retinal **22** (50 mg, 86% yield). The radioactive material was used without purification following extraction and removal of CH_2Cl_2 in the next step and the amount isolated was assumed to be the same as for the non-radioactive reaction. 15- ^3H -retinal, 31.2 mCi/mmol. [^1H NMR (500 MHz CDCl_3): δ 10.08(1H, d, $J=8.34$ Hz), 7.12(1H, dd, $J=15.34$, 13.19 Hz), 6.37(1H, d, $J=15.34$ Hz), 6.34(1H, d, $J=13.19$ Hz), 6.18 (1H, d, $J=15.00$ Hz), 6.16 (1H, d, $J=17.01$ Hz), 5.95(1H, d, $J=8.39$), 2.30(3H, s, CH_3), 2.00(3H, s, CH_3), 1.67(3H, s, CH_3), 1.10(3H, s, 2 x CH_3). ^{13}C NMR (300 MHz CDCl_3): δ 190.5, 154.6, 140.8, 137.6, 137.0, 134.3, 132.2, 130.2, 129.8, 129.4, 129.0, 39.6, 34.1, 33.3, 29.0, 21.7, 19.0, 13.0. UV: λ_{max} (hexanes) 369 nm]

9. Synthesis [15-³H]-β-carotene **23**



In the dark, LiAlH₄ (1 eq, 9 mmol) was added to a stirred slurry of TiCl₃ under nitrogen in dry THF at 0 °C. The reaction was exothermic and a rapid change in color was observed. Retinal (15 mg, 5 mmol) was added dropwise and it was warmed to RT. Then the reaction was refluxed for four h. The reaction was quenched slowly with NaHCO₃ and extracted with hexanes 3x. The hexane extract was dried under anhydrous Na₂SO₄ and concentrated under *vacuo*. A pipette silica column eluted with hexanes afforded 12 mg of β-carotene in a giving a 80% yield 15-[³H]- β,β carotene, 136.5 mCi/mmol. [¹H NMR (300 MHz CDCl₃): δ 7.25(1H, d, *J*=11.6 Hz), 6.59(4H, m), 6.40(2H, m), 6.25(1H, d, *J*=11.1 Hz), 4.21(2H, q, *J*=7.2 Hz), 2.00(2H, m, CH₂), 1.97(3H, s, CH₃), 1.95(3H, s, CH₃), 1.69(6H, s, 2 x CH₃), 1.60(2H, m, CH₂), 1.54(3H, s, CH₃), 1.45(2H, m, CH₂), 1.28(3H, t, OCH₂CH₃, *J*= 7.2 Hz), 1.00(6H, s, 2 x CH₃). ¹³C NMR (300 MHz CDCl₃): δ 168.52, 143.84, 138.74, 137.83, 137.80, 137.65, 136.94, 136.53, 135.71, 135.39, 131.99, 131.91, 130.67, 129.48, 129.32, 126.15, 125.72, 123.14, 60.46, 39.58, 34.22, 33.07, 29.65, 28.92, 21.73, 19.19, 14.32, 12.86, 12.81, 12.75, 12.68. UV: λ_{max} (hexanes) 440 nm]

Assay studies

All the reactions were carried out in a darkroom under minimal photographic safety red lights.

B.1 Preparation of β -carotene micelles

Five different stocks of β -carotene were prepared. β -carotene (81 μ L of a solution 875 mM) was combined with taurocholate (100 μ L) and of Tween 40 (3340 μ L, 4% w/v acetone). The solution was dried under a stream of nitrogen and redissolved in Tris- KOH 10 mM buffer (10 mL, pH 8.0, 6 mM sodium taurocholate, 0.5 mM DTT).

B.2 Preparation of 8' *apo*-carotenoic acid micelles

Five different stocks of 8' *apo*-carotenoic acid micelles were prepared (A-E); by mixing 8' *apo*-carotenoic acid (41, 42, 47, 54.7 and 82 μ L) taurocholate (5.56 mM, 10 μ L) and Tween 40 (0.5 M and 334 μ L, 4% w/v acetone). The solutions were dried under a stream of nitrogen and dissolved in Tris- KOH buffer (1 mL, pH 8.0 6 mM sodium taurocholate, 0.5 mM DTT).

B.3 Inhibition studies using all-*trans*-8' *apo*- β -carotenoic acid

The final reaction mixture for the cleavage reaction contained 8' *apo*-carotenoic acid (A-E, 1, 10, 50, 100, and 200 μ M) β -carotene (7.1 μ M), (0.068 mM), Tris-KOH buffer pH 8.0, (sodium taurocholate(4.05 mM), DTT (0.34 mM),

Tween 40 (0.9%), α -tocopherol (50 μ M), FeSO_4 (1.38 mM), enzyme (20 μ g of BCDOX).

The assay was started by adding the protein to solution A (solution A: 540 μ L Tris- KOH buffer pH 8.0, 6 mM sodium taurocholate, 0.5 mM DTT, and 20 μ L of FeSO_4). The solution was pre-incubated for 5 min at 30 °C. Then 8'-*apo*-carotenoic acid (A-E, 1, 10, 50, 100, and 200 μ M) was added and the mixture was incubated for 10 min at 30 °C. Then the β -carotene micelles solution (160 μ L) was added and the mixture was then incubated for 2-3 h at 30 °C. The reaction was stopped by addition of formaldehyde ((200 μ L, 46% in water) and was incubated for another 10 min. The carotenoids and products were extracted sequentially with chloroform:methanol (1.5 mL, of a 1:2, v/v, 0.01% pyrogallol) and hexane (1.5 mL) and with chloroform:hexane (0.5 mL:1.5 mL). After addition of each organic solvent, the reaction mixture was mixed using a vortex for 40 seconds. The combined extracts were then dried under a stream of nitrogen gas, and dissolved in isooctane: toluene (200 μ L of (8/2, v/v) containing 0.01% butylhydroxytoluene).

The formation of retinal at a different concentration of *apo*-carotenoic acid was monitored. Analysis of the results suggests that **3** 'inhibits' the oxidation of β -carotene to retinal, thus suggesting that it binds to the active site.

B.4 Radioactive extractions

Each one of the extractions was repeated six times. Test tubes containing an aqueous phase and the organic phase were prepared. A variety of solvents were used. The best system will be used as an example

Organic solvent	Aqueous co-solvent	Water or buffer
isooctane 1 mL	acetonitrile 0.9 mL	0.1 mL

Then the radioactive substrate was added

- For studies involving carotene extraction 1 μL of ^3H - β -carotene 136 mCi/mmol was used.
- For studies involving carotene extraction 1 μL of 15- ^3H -retinal 31.2 mCi/mmol was used.
- For studies involving carotene extraction 1 μL of 15- ^3H -retinol, 88 mCi/mmol was used.

Each tube was covered with parafilm and vortex for 40 sec (speed 4). The tubes were centrifuged 5 min at 2000 RPM. An aliquot (100 μL) of each were transferred out of each phase and were analyzed for its content of radioactivity. To each aliquot (100 μL) of solvent (1 mL) of Opti Phase "Hi Safe 3" cocktail was added to measure radioactivity. To determine the radioactivity in DPM the sample plus the cocktail were put in a Liquid Scintillation Counter, Protocol: 3Montse. The percentage of radioactivity in each phase was calculated.

Protocol: 3Montse

- **Wallac 1414 WinSpectral v1.40 S/N4140295**
- **Counting mode:** DPM
- **Quench index:** SQP(E)
- **Isotope:** ^3H , ^3H , 12.43 y
- **Counting time:** 60 min
- **Repeats:** 1
- **Cycles:** 1
- **Replicates:** 1
- **2 sigma %:** 0.01
- **Minimum cpm:** 0.00 checking time: 10
- **Sp. Library of isotope ^3H :** Wallac
- **Vial type:** clear (for 10 mL vials),
- **Liquid system:** HiSafe
- **Advanced modes:** Chemlum
- **Output to display:** POS, RACKPOS, CPM, CTIME, SQPI, DPM1
- **Addition to display:** Listing, Spectrum
- **Header:** ^3H
- **Spectrum:** Beta
- **Window 1:** 1-1024/Beta
- **Window 2:** 1-1024/Beta
- **Window 3:** 1-1024/Beta
- **Window 4:** 1-1024/Beta
- **Window 5:** 1-1024/Beta
- **Window 6:** 1-1024/Beta

Example:

EXTRACTION

Retinol

PROTOCOL

3 Montse

ID

P03AS079

DATE

DEC-04-2000

ORGANIC PHASE

Isooctane 1 ml

AQUEOUS PHASE

2-Cl Ethanol 0.9 ml

Buffer 0.1 ml

RETINOL

	ORGANIC	AQUEOUS
	<i>Isooctane</i>	<i>2-Cl EtOH : Buffer</i>
	DPM	DPM
1	5562.3	7215
2	6654.7	8520.2
3	4852.1	7150.2
4	4787.4	8028.8
5	6099.9	8269

RETINOL

	ORGANIC	AQUEOUS
	<i>Isooctane</i>	<i>2-Cl EtOH : Buffer</i>
	%	%
1	43.53	56.47
2	43.85	56.15
3	40.43	59.57
4	37.35	62.65
5	42.45	57.55

Average	41.52	58.48		
Standard Deviation	2.689	2.689		
Variance	7.228	7.228		
Confidence Intervals 95%	41.52	3.09	58.48	3.09
Confidence Intervals 99%	41.52	4.85	58.48	4.85

B.5 Assay of BCDOX activity using ^3H - β -Carotene

Preparation of β -carotene micelles

The final reaction mixture for the cleavage reaction contained β -carotene (7.1 μM , ^3H - β -Carotene (1 μL of (136 mCi/mmol), Tris-KOH buffer (8 mM pH 8.0), sodium taurocholate (4.05 mM), DTT (0.34 mM), Tween 40 (0.3%), α -tocopherol (50 μM), FeSO_4 (1.38 mM), and BCDOX (roughly 20 μg of protein).

The assay was started by adding the protein to solution A (solution A: 540 μL Tris- KOH buffer pH 8.0, 6 mM sodium taurocholate, 0.5 mM DTT, and 20 μL of FeSO_4). The solution was pre-incubated for 5 min at 30 $^\circ\text{C}$. Then the β -carotene micelles solution (160 μL) was added and the mixture was then incubated for 2-3 h at 30 $^\circ\text{C}$. The reaction was stopped by addition of formaldehyde ((200 μL , 46% in water) and was incubated for another 10 min. The carotenoids and products were extracted sequentially with chloroform:methanol (1.5 mL, of a 1:2, v/v, 0.01% pyrogallol) and hexane (1.5 mL) and with chloroform:hexane (0.5 mL:1.5 mL). After addition of each organic solvent, the reaction mixture was mixed using a vortex for 40 seconds. The combined extracts were then dried under a stream of nitrogen gas, and dissolved in isooctane: toluene (200 μL of (8/2, v/v) containing 0.01% butylhydroxytoluene). Then the radioactivity was measured (100 μL of the assay mixture were transferred to a tube containing 1mL of isooctane and 1 mL of acetonitrile) each test was repeated six times. Each tube was covered with parafilm and vortex for

40 sec (speed 4). The tubes were centrifuged 5 min at 2000 RPM. 100 μ L of each were aliquoted out of each phase and were analyzed for their content of radioactivity. To each (100 μ L) phase (1 mL) of Opti Phase "Hi Safe 3" cocktail was added to measure radioactivity.

3.6 References

1. E. I. Solomon; T. C. Brunold; M. I. Davis; J. N. Kemsley; S. K. Lee; N. Lehnert; F. Neese; A. J. Skulan; Y. S. Yang; J. Zhou, "Geometric and electronic structure/function correlations in non-heme iron enzymes." *Chemical Reviews* **2000**, *100*, (1), 235-349.
2. F. Takuzo, *Oxygenases and model systems*. ed.; Kluwer Academic Publishers: Norwell, Ma, 1997; 'Vol.' 1, p 1.
3. M. G. Leuenberger; C. Engeloch-Jarret; W.-D. Woggon, "The reaction mechanism of the enzyme catalyzed central cleavage of beta-carotene to retinal." *Angewandte Chemie International Edition* **2001**, *40*, 2614-2617.
4. H. O. Olson J. A., "The enzymatic cleavage of β -carotene into vitamin A by soluble exymes of rat liver and intestine." *Proceedings of the National Academy of Sciences of the United States of America* **1965**, *54*, 1364-1370.
5. A. Lindqvist; S. Anderson, "Biochemical properties of purified recombinant human beta carotene 15-15'-monooxygenase." *Journal of Biological Chemistry* **2002**, *277*, (26), 23942-23948.
6. W. Yan; G.-F. Jang; F. Haeseleer; N. Esumi; J. Chang; K. Palcewski; D. Zack, "Cloning and characterization of a human beta-carotene." *Genomics* **2001**, *72*, 193-202.
7. M. R. Lakshman; J. L. Pope; A. O. J, "The specificity of a partially purified carotenoid cleavage enzyme of rabbit intestine." *Biochemical and Biophysical Research Communications* **1968**, *33*, (2), 347-352.
8. R. C. Mordi; J. C. Walton; G. W. Burton; L. Hughes; K. Ingold; D. Lindsay; D. Moffatt, "Oxidative degradation of β -carotene and β -apo-8'-carotenal." *Tetrahedron* **1993**, *49*, 911-928.
9. H. Singh; H. R. Cama, "Enzymatic cleavage of carotenoids." *Biochimica et Biophysica Acta* **1974**, *370*, 49-61.

10. P. S. Bernstein; S.-Y. Choi; Y.-C. Ho; R. R. Rando, "Photoaffinity Labeling of Retinoic Acid-Binding Proteins." *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 654-658.
11. G. Chen; A. Radomska-Pandya, "Direct photoaffinity labeling of cellular retinoic acid binding protein I with all trans retinoic acid: Identification of amino acids in the ligand binding site." *Biochemistry*, **2000**, *39*, 12568-12574.
12. W. B. Jackoby; M. Wilchek, "Affinity Labeling." *Methods in Enzymology* **1977**, *66*, 69-114.
13. F. Kotzyba-Hibert; I. Kapfer; M. Goeldner, "Recent trends in photoaffinity labeling." *Angewandte Chemie International Edition* **1995**, *34*, 1296-1312.
14. D. Schuster; *a. et.*, "Photoaffinity Labeling." *Photochemistry and Photobiology* **1989**, *49*, 785-804.
15. A. E. Purcell; W. M. Walter, "Preparation of ¹⁴C- β -Carotene." *Methods in Enzymology* **1971**, *18*, (3), 701-706.
16. A. A. Liebman; W. Burger; S. C. Choudhry; J. Cupano, "Synthesis of carotenoids specifically labeled with isotopic carbon and tritium." *Methods in Enzymology* **1992**, *213*, 42-49.
17. P. Tosukhowong; T. Supasiri, *Journal of Labelled Compounds and Radiopharmaceuticals* **1985**, *22*, 843-850.
18. S. W. Rhee; J. I. Degraw; H. H. Kaegi, *Journal of Labelled Compounds and Radiopharmaceuticals* **1985**, *22*, 843-850.
19. H. H. Kaegi; J. E. Bupp; J. I. Degraw, *Journal of Labelled Compounds and Radiopharmaceuticals* **1982**, *19*, 745-752.
20. P. L. Chien; M. S. Sung; D. B. Bailey, *Journal of Labelled Compounds and Radiopharmaceuticals* **1979**, *16*, 791-797.

21. P. V. Reddy; M. Rabago-Smith; B. Borhan, "Synthesis of all-trans-[10 ³H]-8 ³H]-apo-beta-carotenoic acid." *Journal of Labelled Compounds and Radiopharmaceuticals* **2002**, *45*, (1), 79-89.
22. G. Britton; S. Liaen-Jenses; H. Pfander, *Carotenoids: Synthesis*. ed.; Verlag: Boston, 1995; 'Vol.' 2, p.
23. J. H. Babler 107,030, 1992.
24. P. Horst, "Synthesen in der Vitamin-A-Reihe." *Angewandte Chemie* **1960**, *72*, 811-819.
25. I. M. Goldman, "Activation of manganese dioxide by azeotropic removal of water." *Journal of Organic Chemistry* **1969**, *34*, 1979-1981.
26. R. Yamauchi; N. Miyake; K. Kato; Y. Ueno, "Reaction of alpha-tocopherol with alkyl and alkylperoxyl radicals of methyl linoleate." *Lipids* **1993**, *12*, 201.
27. C. Hansch; A. Leo, *In Exploring QSAR. Fundamentals and applications in chemistry and biology*. ed.; American Chemical Society: Washington, D. C, 1995; 'Vol.' p 125-168.
28. Y. Handa; J. Inanaga, "A highly stereoselective pinacolization of aromatic and alpha-beta unsaturated aldehydes mediated by titanium(III)-magnesium(II) complex." *Tetrahedron Letters* **1987**, *28*, (46), 5717-5718.
29. J. E. M. Murry; M. P. Fleming, "A new method for the reductive coupling of carbonyls to olefins. Synthesis of beta-carotene." *Journal of the American Chemical Society* **1974**, *94*, (17), 4708-4709.
30. J. E. M. Murry; M. P. Fleming; K. L. Kees; L. R. Krepski, "Titanium-Induced reductive coupling of carbonyls to olefins." *Journal of Organic Chemistry* **1978**, *43*, (17), 3255-3265.
31. D. E. Ong, "A novel retinol-binding protein from rat." *The Journal of Biological Chemistry* **1984**, *259*, (10), 1476-1482.

32. J. E. Smith; D. S. Goodman, "Retinol-binding protein and the regulation of vitamin A transport." *Federal Proceedings* **1979**, 38, 2504-2509.
33. S. Futterman, and Saari, J. C., "Occurrence of 11-cis-retinal-binding protein restricted to the retina." *Investigative Ophthalmology and Visual Science* **1977**, 16, 768-771.
34. Y. L. W. Lai, B., Liu, Y.P., and Chader, G. J., "Interphotoreceptor retinol-binding proteins: possible transport vehicles between compartments of the retina." *Nature* **1982**, (298), 848-859.
35. G. Wolf, "The enzymatic cleavage of beta-carotene: still controversial." *Journal of Nutritional Reviews* **1995**, 53, 134-137.
36. C. Duszka; P. Grolier; E. M. Azim; M. C. AlexandreGouabau; P. Borel; V. AzaisBraesco, "Rat intestinal beta-carotene dioxygenase activity is located primarily in the cytosol of mature jejunal enterocytes." *Journal of Nutrition* **1996**, 126, (10), 2550-2556.
37. A. During; A. Nagao; C. Hoshino; J. Terao, "Assay of beta-carotene 15,15'-dioxygenase activity by reverse-phase high-pressure liquid chromatography." *Analytical Biochemistry* **1996**, 241, (2), 199-205.
38. T. V. Vliet; R. V. Schaik; H. V. D. Berg; W. H. P. Schreurs, "Effect of vitamin A and beta-carotene intake on dioxygenase activity in rat intestine." *Annals of the New York Academy of Sciences* **1993**, 691, 220-222.

Chapter 4

Introduction

4.1 Vision

Vision is one of the most intriguing and fascinating processes in nature. Vision is the act or sense of light which involves the capture of light by the eye, followed by its transmission to the brain as electrical pulses.¹ These pulses are recognized as images and colors in the brain. The eye is composed by specialized components; the cornea, lens and the vitreous fluid that focus the light waves onto the retina (Figure 4-1).³ The retina is located in the inner part of the eye, and contains the photoreceptors. The photoreceptors, the rod and the cone cells (Figure 4-2) are responsible for capturing the image and sending the signal to the brain via the optic nerve.⁴ An eye contains approximately 100 million rod cells and about 6 million cone cells.⁵ The cone cells are responsible for color (photopic) vision and the rod cells are responsible for dim (scotopic) vision. Cone cells are both quicker to adapt to dark environments (~10 min) and sensitive to a broader range of wavelengths. On the other hand, rod cells take longer to adapt to dark

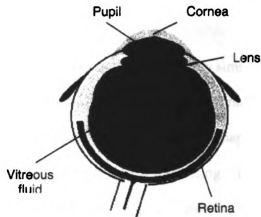


Figure 4-1. The Human Eye.

environments (20 – 30 min), but once adapted they are 500 times more sensitive than cone cells.

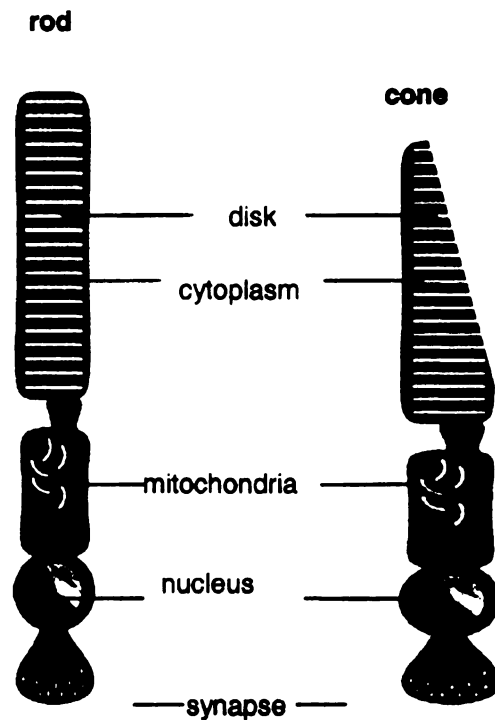


Figure 4-2. The rod and cone cells in retina.

4.1.1 Process of vision

Rod cells are more abundant than cone cells but both cells are very similar and contain a nucleus. Both cells differ from normal cells because they contain two unique features: a synaptic ending and an outer segment.⁶ The synaptic ending is where the neural signals pass through to the brain for processing. The outer segment is filled with disks made of a phospholipid bilayer. Spanning the bilayers of these disks is a membrane bound protein called opsin, which is the one responsible for the visual signal transduction (Figure 4-3). There are four

different opsins. In the rod cells the opsin protein is bound to a chromophore (11-*cis*-retinal) as a Protonated Schiff Base (PSB). This complex is called rhodopsin.

Rhodopsin (Rh) is a hydrophobic *trans*-membrane 7- α -helical protein from the family of G-protein coupled receptors (GPCR). GPCRs are membrane-bound proteins that activate a wide variety of biological functions including signal transduction and regulatory control.⁷ In the 1930s Wald discovered that

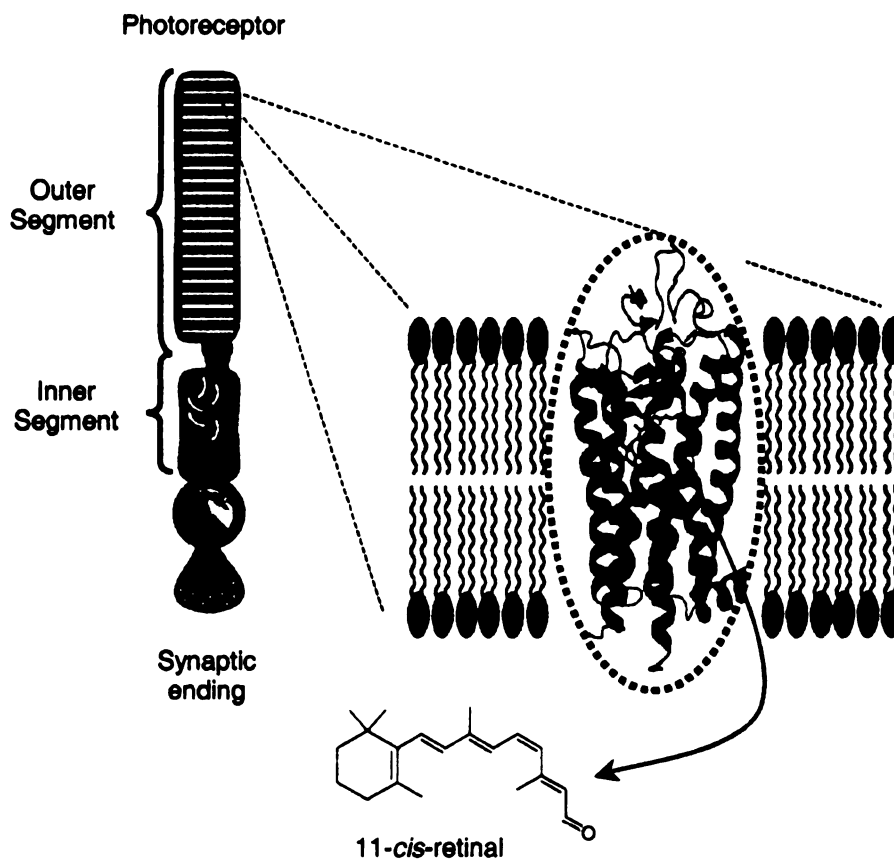


Figure 4-3. Schematic representation of rhodopsin in a rod cell.

rhodopsin is formed by an apoprotein called opsin (Op) and a chromophore, a retinoid derivative,⁸ which later on was identified as the aldehydic form, retinal.⁹ In the 50's it was suggested by Collins¹⁰ and Morton,¹¹ that the retinal was bound to the protein as a protonated Schiff base (PSB) via a Lys residue (Figure 4-4). In humans and

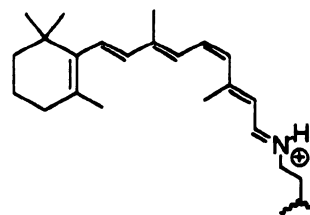


Figure 4-4. A Schiff base is formed between 11-*cis*-retinal and a lysine residue in Rhodopsin.

most animals the chromophore is 11-*cis*-retinal. Each rod cell can contain up to 1700 disks and each disk can contain up to 1.5 million molecules of rhodopsin.¹² At one time, one uses only about 10% of the rhodopsin complexes, thus allowing vision to be a continuous process.

The visual process begins in the dark state where the chromophore, 11-*cis*-retinal, is bound as a protonated Schiff base via Lys 296 (inactive state). When a photon is absorbed it results in the isomerization of 11-*cis*-retinal to all-*trans*-retinal. The isomerization of retinal results in a conformational change that catalyzes the activation of a G-protein (transducin),¹³ which activates a cyclic GMP phosphodiesterase¹⁴⁻¹⁷ leading to a decline in the free GMP that results in membrane hyperpolarization.¹⁸⁻²⁰

The isomerization can be monitored by low temperature UV-vis and other spectroscopic methods.²¹⁻²⁴ This isomerization goes through a series of discrete intermediates (Figure 4-5). The first observed intermediate, formed in less than 200 femtoseconds after irradiation, is known as photo rhodopsin, and is considered to be a highly distorted all-*trans*-PSB chromophore. Formation of the

bathorhodopsin leads to a more stable distorted *trans* intermediate.^{25,26} Then in nano seconds, the lumi rhodopsin is formed followed by meta I and meta II states. The meta II state is the only one that forms an unprotonated Schiff base.^{27,28} Finally, the imine bond is protonated and hydrolyzed to afford the free all-*trans*-retinal and the opsin protein.

The initial stages of the transduction pathway are controlled by steric factors,²⁵ while the latter stages primarily involve proton transfer processes.²⁸ The initiation of the visual transduction occurs at the meta II state, and the all-*trans*-retinal is re-isomerized to the 11-*cis*-retinal in the retinal pigment epithelium.

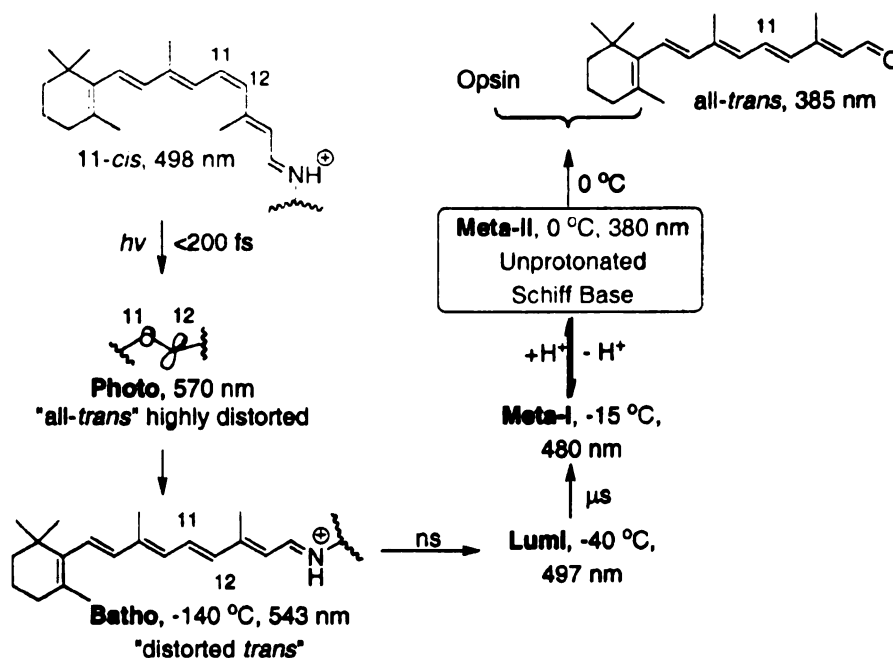


Figure 4-5. Schematic representation of the intermediates observed in the isomerization of 11-*cis*-retinal to all-*trans*-retinal in rhodopsin.

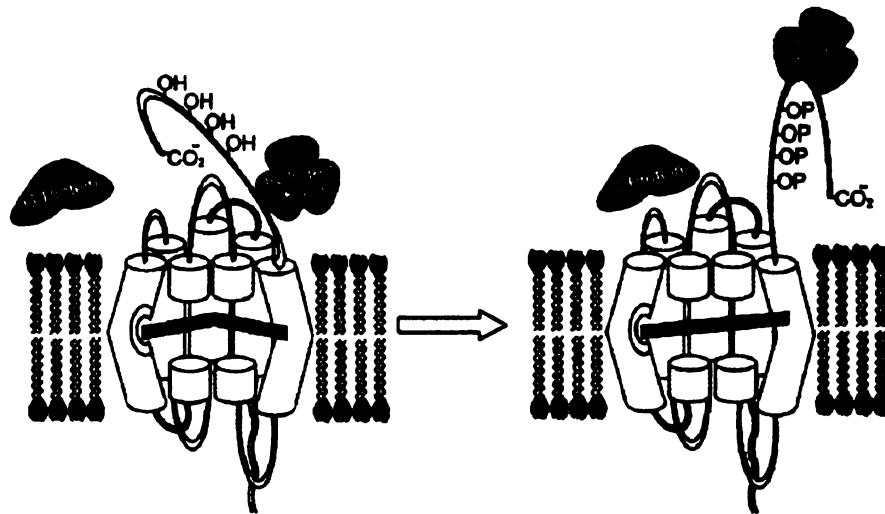


Figure 4-6. The isomerization of 11-*cis*-retinal to the all-*trans* form causes a change in the protein conformation, which allows for binding and activation of the G-protein transducin.

The isomerization of 11-*cis*-retinal to the all-*trans* form causes a change in the protein conformation (Figure 4-6), which allows for binding and activation of the G-protein transducin.²¹ Transducin is formed by three different subunits α , β and γ , and a GDP molecule is bound to the α component. When the rhodopsin conformation changes (R^*), the affinity of GTP to transducin decreases, and the affinity to GDP increases, resulting in the exchange of a GTP for a GDP molecule. Binding of GTP results in the dissociation of the α -subunit (T_α -GTP). The T_α -GTP subsequently binds to the γ -subunit of PDE (Figure 4-7). Activation of PDE results in the hydrolysis of cyclic GMP to GMP. This results in the decline in free intracellular cyclic GMP. Hydrolysis of cyclic GMP to GMP forces the Ca^{2+} ion channel in the rod/cone outer segment to close and halts the flow of Ca^{2+} into the cell. Thus a change in Ca^{2+} concentration occurs (from 0.5

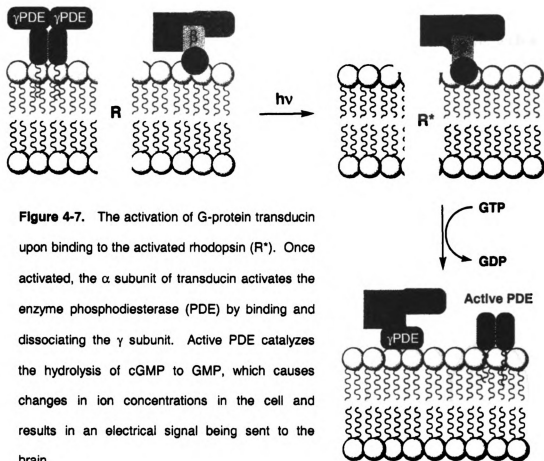


Figure 4-7. The activation of G-protein transducin upon binding to the activated rhodopsin (R^*). Once activated, the α subunit of transducin activates the enzyme phosphodiesterase (PDE) by binding and dissociating the γ subunit. Active PDE catalyzes the hydrolysis of cGMP to GMP, which causes changes in ion concentrations in the cell and results in an electrical signal being sent to the brain.

μM to $0.1 \mu\text{M}$).²⁹ The hyperpolarization of the membrane results in an electric signal that it is sent to the brain and results in a visual image.^{12,30} The visual cycle stops when rhodopsin is phosphorylated by a protein called arrestin.^{31,32} Finally, the phosphorylated opsin binds a new molecule of 11-*cis*-retinal, to restart the visual cycle.

In summary as shown in Figure 4-8, one photon of light causes the isomerization of the chromophore, leading to the activation of the photoreceptor. Thus a single isomerization of a double bond triggers a cascade of events that lead to vision.^{14,17,33,34} Each photon activates about 100 PDE molecules. This

activation leads to the hydrolysis of roughly 100,000 cyclic GMPs.^{35,36} The visual transduction is one of the most efficient and sensitive processes known, with a calculated quantum yield of ~ 0.67 .³⁷⁻⁴¹

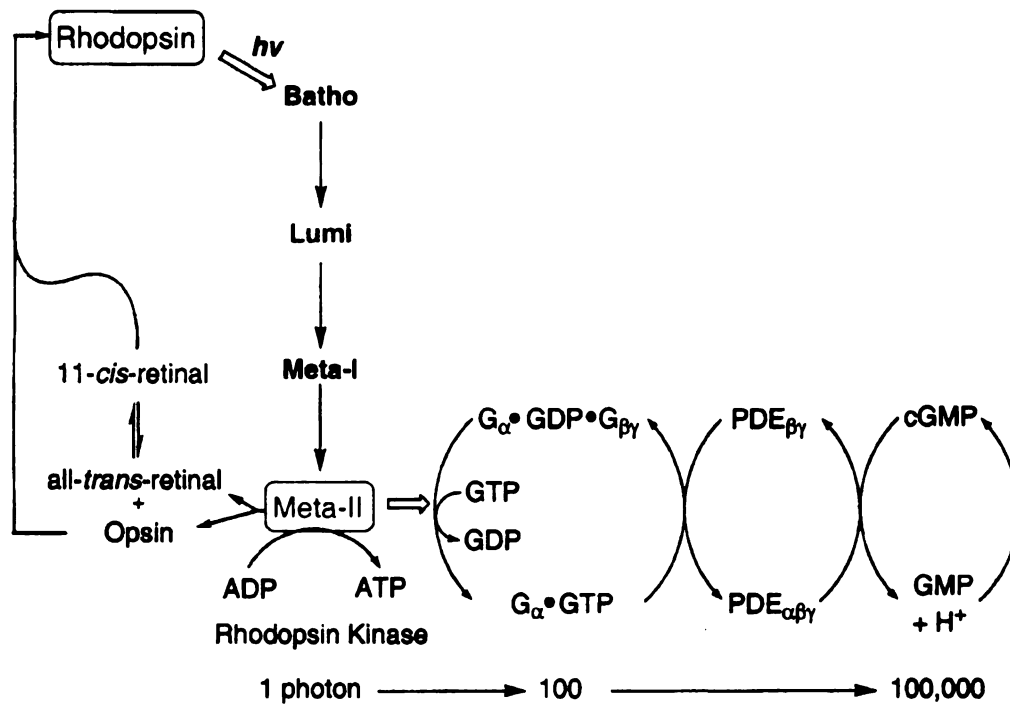


Figure 4-8. A summary of the visual transduction process. One photon of light causes the isomerization of the chromophore, leading to the activation of the photoreceptor. A single isomerization of a double bond triggers a cascade of events, which leads to vision.

4.1.2 Color vision

A very important aspect of vision is the fact that not only objects can be identified, but also different colors can be perceived. Color vision is defined as

the capacity of detecting a variety of different wavelengths. In humans this process is possible due to the existence of different receptors. The eye contains three different colored cone opsins, red, blue and green and a rod opsin. Each of these opsins binds the same chromophore, 11-*cis*-retinal as a protonated Schiff base via a Lys residue. Even though they bind the same chromophore, interestingly each one of them results in different wavelength absorption. The human rod rhodopsin absorbs at 500 nm, the green cone opsin at 530 nm, the red cone opsin 560 nm and the blue cone opsin 420 nm.^{42,43} As shown in Figure 4-9 the absorption of the four rhodopsins cover the whole visible spectrum.

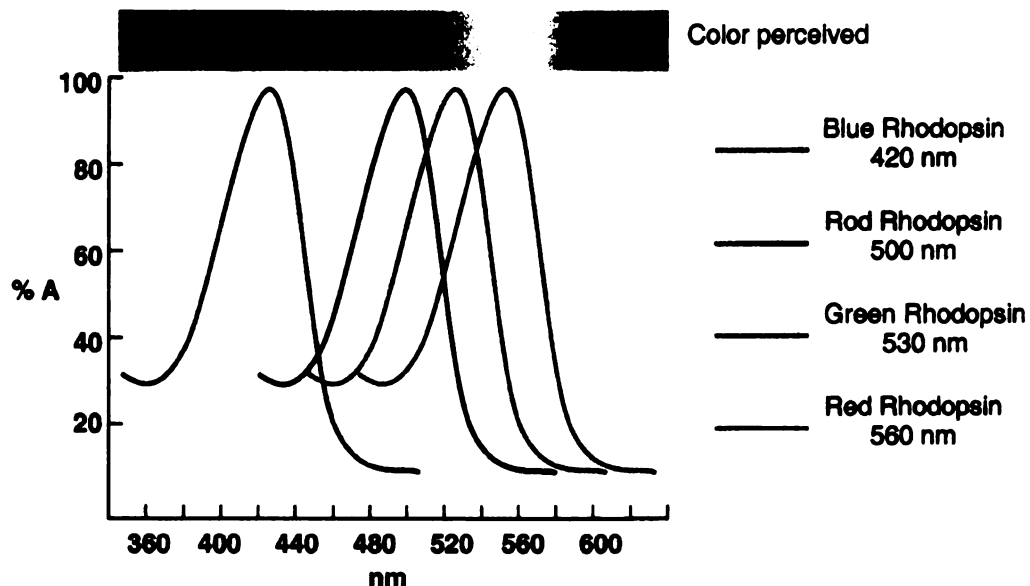
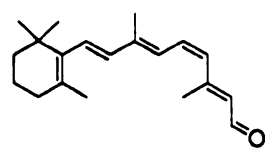
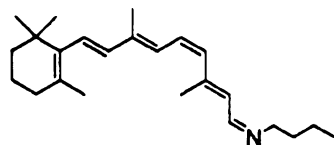


Figure 4.9. Each of the four rhodopsin complexes (blue, green, red and rod) absorb at a different wavelength, comprising the entire visible spectrum.



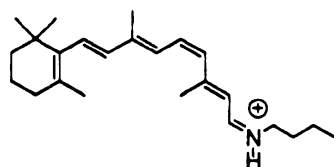
$\lambda_{\text{max}} \sim 380 \text{ nm}$

11-*cis*-retinal



$\lambda_{\text{max}} \sim 365 \text{ nm}$

11-*cis*-retinal
Schiff base



$\lambda_{\text{max}} \sim 440 \text{ nm}$

11-*cis*-retinal
protonated Schiff base

Figure 4-10. UV-vis absorbances of 11-*cis*-retinal as a free aldehyde (380 nm), a *n*-butylamine Schiff base (365 nm) and a protonated Schiff base (440 nm) when ethanol is used as solvent.

The maximal absorption of the free 11-*cis*-retinal is ~ 380 . When a Schiff base is formed with *n*-butylamine, the spectrum blue shifts to ~ 365 nm. Protonation of the Schiff base causes a large red shift to ~ 440 nm (Figure 4-10).^{44,45} Changes in the concentration and solvent conditions can shift the wavelength up to ~ 500 .⁴⁶ From the comparison of the maximal absorption of the protonated Schiff base (~ 500) and the absorptions of the opsins (~ 530 , 560), it is evident that the protonation of the Schiff base is not the only factor involved in the

red shifting of the wavelength. The difference in the absorbance observed for the visual pigments and the PSB of 11-*cis*-retinal in solution is known as the *opsin shift*. It is assumed that the *opsin shift* is due to interactions between the protein and the retinylidene chromophore. Rhodopsin is a large, hydrophobic, membrane-bound protein consisting of 348 amino acids and two oligosaccharide chains (~39,000 Da). Many attempts to solve the 3D structure of rhodopsin were performed.^{47,48} It was not until 2000 that the first crystal structure of a rod rhodopsin was published.⁴⁹ The structures of the cone opsins have not been solved up to date. Different theories that attempt to explain what causes the *opsin shift* have been proposed. Most of the proposed theories explain the wavelength regulation by interfering with the degree of conjugation of the chromophore.

As mentioned previously, Morton and Goddwin were the first to propose that the chromophore 11-*cis*-retinal binds as a Schiff base. But it was not until about ten years later that Akhtar⁵⁰ (1968) and Bownds (1967)⁵¹ independently provided evidence for the imine bond formed between the opsin and a lysine residue. The lysine identified to form the imine bond via proteolytic fragmentation was Lys 296.¹⁰ In an attempt to explain the wavelength absorption of rhodopsin, in 1958 Hubbard

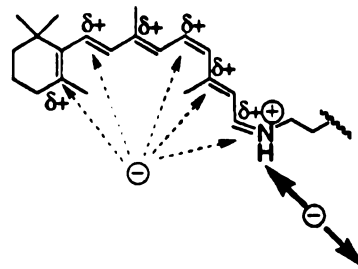


Figure 4-11. Distance of the counter anion to the protonated Schiff base and positioning of charges or dipoles along the backbone of the polyene may modulate the maximal wavelength of the chromophore.

suggested that the Schiff base must be in the protonated form, and interactions with the protein would cause a delocalization of the positive charge throughout the polyene backbone (Figure 4-11).⁵² Kroopf and Hubbard suggested that wavelength regulation in rhodopsin may result from electrostatic interactions between the SB and charged amino acids within the binding pocket.⁵³ In 1974 the protonated nature of the Schiff base was confirmed by Callender through Raman studies.⁵⁴ In 1979 Honig introduced the idea that a negatively charged amino acid might be used as a counter anion to stabilize the PSB, and that a second negatively charged amino acid could be placed at another position along the polyene backbone.⁵⁵ Later, NMR studies carried out by Nakanishi and others verified the existence of a second negative charged amino acid near carbon 12 of retinal.⁵⁶⁻⁵⁸ In 1989 Oprian⁵⁹ and Sakmar⁶⁰ identified the counter anion as Glu113. When the Glu113 was mutated to Gln the protein existed as a mixture of two different species, one with an absorbance at 490 nm, and the other at 380 nm. These two species existed in a pH-dependent equilibrium, thus they were identified as the protonated form and unprotonated form of the Schiff base. The presence of the counter anion is essential for the formation of a stable PSB, possibly due to the unfavorable positive charge formed by the PSB. When high concentrations of halide were used, the protonation was recovered.⁶¹ It was suggested that the halide functioned as a counter anion, and therefore, stabilized the PSB. Addition of halide to the wild type rhodopsin did not change the chromophore maximal absorbance. In 1972 Blatz proposed the hypothesis that varying the distance between the counter anion and the protonated Schiff base changes the maximum

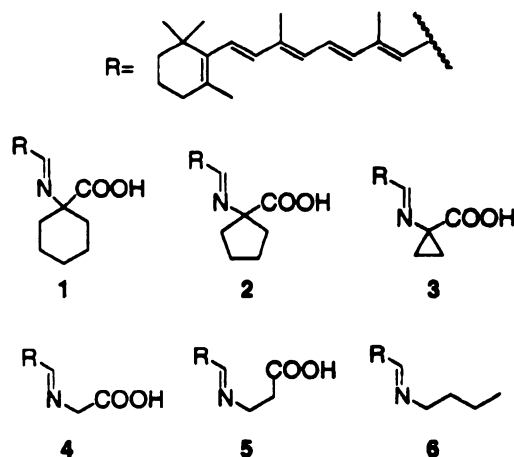
wavelength obtained.⁶²

When the counter anion is close to the positively charged nitrogen, the charge is localized at the protonated nitrogen. But when the counter anion is further away, the positive charge exists as a resonance over the polyene, causing a red shift in the λ_{\max} . A mutation of Glu113 for Asp (shorter counter anion) causes a 5 nm red shift in the

λ_{\max} (505 nm) supporting Blatz theory.⁶⁰ Sheves and co-workers prepared a series of model compounds to probe the effect of varying the

distance and the angle between the counter anion and the PSB (Figure 4-12).⁶³

The results showed that the largest shift obtained in the λ_{\max} of the PSB species was 10 nm, the larger the distance between the counter anion, the larger the red shift observed. This does not account for the *opsin shift* observed in the opsin proteins. However the pK_a varied significantly, demonstrating that positioning of



Compound	λ_{\max} (nm)		
	SB	PSB	pK_a
1	340	450	12.3 ± 0.1
2	386	448	9.5 ± 0.1
3	386	450	8.0 ± 0.1
4	380	444	8.5 ± 0.1
5	375	442	10.0 ± 0.1
6	360	440	7.5 ± 0.1

Figure 4-12. Absorption maxima and apparent pK_a values of Schiff base (SB) and protonated SB (PSB) of retinal with various amino acids. The relative position of the carboxylic acid dictates the λ_{\max} and the pK_a values.

the counter anion has a significant influence in the stability of the PSB. Hoing proposed that polar groups localized in the binding cavity could help to stabilize the PSB formed in the protein.⁶⁴

Along the same lines, Nathans suggested that regulation of the wavelength could be accomplished by the positioning of a negatively charged amino acid. Positioning of the negatively charged amino acid would result in a different distribution of the conjugation of the positive charge along the polyene (Figure 4-11).⁶⁵

A different hypothesis suggests that the wavelength is regulated by twisting the planes about the single bonds in the molecule to achieve different levels of conjugation (Figure 4-13).⁶⁶ Although retinal would prefer to adopt a planar conformation to maximize p-orbital overlap, it is known that 11-*cis*-retinal is not a

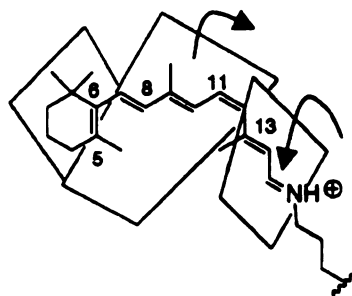


Figure 4-13. Twisting of the single bonds will reduce the degree of π orbital overlap, and thus will lead to different maximal wavelengths.

planar molecule. The chromophore has two strong steric interactions, the first one between the C8-H and the C5-CH₃ and the second one between C10-H and the C13-CH₃ (Figure 4-14). This leads to a twist of the C6-C7 and C12-C13 single bonds to alleviate the steric hindrance.^{67,68} It is possible that when the chromophore binds the protein, it is forced to a more planar conformation. Thus maximizing the p-orbital overlap.

Therefore, a red shift in the absorbance would be observed. And on the contrary, when a highly twisted chromophore exists, a blue shift in the absorbance would be observed. Different degrees of twisting would result in different conjugation, which would cause different maximal absorbances.⁶⁶ Theoretical studies provide evidence that twisting of the chromophore alone does not account for the large observed opsin shifts.⁶⁹

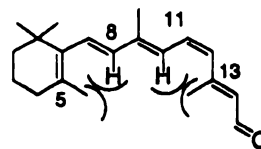


Figure 4-14. 11-*cis*-retinal is not a planar molecule due to steric interactions. The first steric interaction is between the C8-H and the C5-CH₃ and the second one between C10-H and the C13-CH₃. This leads to a twist of the C6-C7 and C12-C13 single bonds to alleviate the steric hindrance.

Many different studies such as NMR, cryoelectron microscopy, photoaffinity labeling, etc, have been performed in an attempt to prove the theories discussed previously.⁷⁰⁻⁷⁴ Some of those theories have been partially confirmed by model compounds.⁶⁶ However the real breakthrough occurred in 2000 when Palczewski *et. al.* reported the crystal structure of bovine rod rhodopsin.⁷⁵ In the binding pocket it was confirmed that the 11-*cis*-retinal was bound to Lys296 as a Schiff base, and the counter anion, Glu113, was close to the nitrogen (about 3.5 Å). A second Glu181 was found, and could support Honig's theory that the shift is modulated by a two charge model.⁷⁶

In 1979 Rafferty proposed that another way to red shift the wavelength would be via excitonic coupling which could occur between the Trp265 and the chromophore.⁷⁷ There is an important interaction between the C13-CH₃ of the

chromophore and Trp265. When a rhodopsin is prepared using a retinal that does not contain a C13 methyl group, the activity is only partially recovered.⁷⁸ This supports the proposed theory that different twisting in the chromophore (to optimize p-orbital overlap) would cause a shift of the wavelength.

As mentioned previously, a large number of studies were performed to understand the modes of wavelength regulation. Since all four rhodopsin proteins are membrane proteins, their crystallization has been hampered for many years.

The human green and red cone proteins have a 96% homology sequence and differ only in 15 amino acids.⁶⁵ Interestingly, even though their sequence is so similar, their wavelength absorption is significantly different, 570 nm and 530 nm for red and green rhodopsin, respectively. Oprian *et. al.* have performed studies using red rhodopsin, and they were able to shift the wavelength from 570 nm to 530 nm.⁷⁹ This was accomplished by mutating only seven amino acids S116Y::S180A::I230T::A233S::Y277F::T285A::Y309F. Mutation of three of those amino acids would account for most of the shift.^{80,81} All the substitutions involve the conversion of a hydrophobic amino acid to a polar residue near the ionone region, namely F277Y, A285T, and A180S. The same trend was observed when site directed mutagenesis was performed in rod rhodopsin to mutate hydrophobic residues to polar in analogous positions. The rod mutant F261Y::A269T::A164S has a very large red shift, from 500 nm to 700 nm.⁸²

Studies in blue rhodopsin have been more difficult due to the difficulty to express and obtain large protein quantities. Blue rhodopsin shares only 46% of homology with the other pigments.^{65,83} Because of this, most of the studies have

been performed using mutagenesis of other opsins to promote a blue shift in the wavelengths. Sakmar found that by mutating nine residues in rod rhodopsin M86L::G90S::A117G::E122L::A124T::W265Y::A292S::A295S::A299C a shift from 500 to 438 nm is observed.⁸⁴ Also Farrens observed that the a triple mutation in rod rhodopsin T118A::E122D::A292S, shows an absorption at 453 nm (47 nm blue shifted).⁸⁵ The single mutation W265Y produces a 15 nm shift, and the single mutations G90S and A292S causes a 11 nm, and 10 nm blue shift, respectively.⁸⁶ Mutation of corresponding amino acids in the blue pigment seem to force red shifting. Blue opsin mutant Y262W (β -ionone ring vicinity) generates a 10 nm red shift, whereas S289A (vicinity of SB) results in no change, and S87G (vicinity of SB) actually blue shifts 10 nm.

4.2 Designing of a rhodopsin surrogate

As previously discussed, several theories have been proposed to explain the mechanism of wavelength regulation. In the last fifty years, our understanding of the mode of wavelength regulation increased greatly. However, as mentioned previously, rhodopsins are membrane-bound proteins and techniques such as crystallization, mutagenesis, etc. are very challenging when with this group of proteins. Therefore, most of the structural information obtained on membrane bound proteins is the result of theoretical analysis and low resolution structural data coupled with results of site directed mutagenesis. This has hampered advances in the understanding of the mode of action of cone and rod rhodopsins.

We are interested in exploring the nature of protein/substrate interactions at the molecular level since these interactions are at the heart of biochemical events that regulate biological systems. As a model for such studies we have chosen the case of wavelength regulation in rhodopsin, i.e., the mechanism by which we can see colors as a powerful example of the consequences of protein/substrate interactions. Thus, we decided on engineering a protein that mimics rhodopsin, but it does not hold the same difficulties as rhodopsin, especially in regards of crystallographic studies and site directed mutagenesis. The engineered rhodopsin mimic would allow us to test the proposed theories more precisely and efficiently. In more detail, a rhodopsin mimic would provide information about the important factors that determine retinal binding as a PSB and how wavelength modulation occurs. In the last decade, protein design has been widely used, and has lead to an increase in the understanding of protein folding and structure/function relationships.^{31,87-99} Development of such proteins offers a greater understanding of the natural proteins. It also represents the first step toward a new generation of novel macromolecules that will have practical applications in industry and biomedicine.¹⁰⁰⁻¹⁰² After designing a protein that mimics the PSB formation with retinal we will systematically introduce charges, sterics, and dipoles at varying locations with respect to the bound chromophore with hopes of deconvoluting and understanding the possible reasons for the altering opsin shifts.

A suitable protein mimic should have the following characteristics. The protein should have a robust structure to tolerate multiple mutations. Different

amino acids will be mutated, first to accomplish the binding of retinal as a PSB, and second to probe amino acids along the polyene and PSB region to probe the effects in the absorption wavelength. The 3-D structure should be solved, mainly because the design of the rhodopsin mimic would be based on the use of rational mutagenesis. The protein cavity should be large, hydrophobic, and protected from the solvent.

Also it is important that the protein can be expressed in large scale as a water soluble protein and should be small in size. Preference for binding retinoid structures will be advantageous in the engineering process. Two families of proteins were considered as the platform protein for our engineering studies. The first family were the Retinoid Nuclear Receptors.¹⁰³ The retinoid receptors have been classified into two subfamilies,¹⁰⁴⁻¹⁰⁷ the Retinoic Acid Receptors (RAR) that bind to all-*trans*-retinoic acid and the Retinoid X Receptors (RXR) that bind to 9-*cis* retinoic acids. The partial crystal structure of these proteins has been reported, but unfortunately these proteins tend to form dimers and are membrane associated. The second family is the Intracellular Lipid Binding Proteins (iLBP) and it is believed that these proteins are involved in the transport and metabolism of hydrophobic ligands in the cell.^{108,109} In particular a very interesting subgroup of proteins (subcategory of the iLBP) are the retinoid binding proteins (RBPs). The RBP proteins bind retinoids, which are very important in the control of a wide variety of cellular processes like transduction,¹¹⁰ gene regulatory control,¹¹¹ cancer,¹¹² and disease prevention.^{113,114} The concentration of retinoids in the cell is tightly controlled by the retinoid binding proteins, Cellular Retinol Binding

Proteins (CRBP) and Cellular Retinoic Acid Binding Proteins (CRABP). CRBP control the concentration of retinol (retinoids in the alcoholic form) and CRABP control the concentration of retinal (retinoid in the acid form).¹¹⁵ Retinoic Acid Binding Proteins have been widely studied. There are two forms of CRABP proteins present in most species, CRABPI and CRABPII, which are highly conserved within vertebrates.^{116,117} Both proteins that bind all-*trans*-retinoic acid through a series of electrostatic interactions, are small (15 KDa) and cytosolic proteins.¹¹⁸ NMR, crystallography and over-expression of both proteins have been previously performed.^{115,119-122} In particular, it has been reported that CRABPII can tolerate mutations of some amino acids in its structure while maintaining its general β -barrel structure and satisfactory expression levels.¹²³ The crystal structure of both the apo-CRABPII and CRABPII bound to retinoic acid has been previously reported.^{119,124} Professor Yan (Department of Biochemistry, Michigan State University) graciously donated to us a pET-17b vector containing the human CRABPII gene (Figure 4-15).



Figure 4-15. The crystal structure of CRABPII bound to all-*trans*-retinoic acid.

4.3 Wild Type CRABP II binding properties

Cellular Retinoic Acid Binding Protein II (CRABP II) was chosen as the template for the engineered rhodopsin mimic. CRABP II is a small (137 amino acids) cytosolic, β -barrel protein with a large ($\sim 600 \text{ \AA}^3$) binding cavity made up of two five-stranded β -sheets and a helix-turn-helix motif. All-*trans*-retinoic acid, the natural substrate, is bound through a series of electrostatic interactions between Arg132 and Tyr134 and a hydrogen bond between a water molecule Arg111 and Thr54 (Figure 4-16).¹²⁵

The binding of all-*trans*-retinoic acid to CRABP II is stoichiometric, with a dissociation constant of $\sim 2 \text{ nM}$.¹¹⁸ Due to the structural homology between retinoids we expected that retinal would occupy roughly the same place within the



Figure 4-16. Wild type CRABP II bound to all-*trans*-retinoic acid. The ligand carboxylate is directly interacting with Arg132 and Tyr134, while it is forming a hydrogen bonds with Arg111.

CRABP_{II} binding pocket.

This assumption, verified by modeling studies, was critical for the rational re-designing of the binding pocket of CRABP_{II} in order to achieve a protonated Schiff base forming protein. High selectivity towards binding of retinoic acid vs other retinoids has been reported.

This observation was verified (Figure 4-17), since the aldehydic analog of retinoic acid, all-*trans*-retinal, exhibits a poor

dissociation constant for CRABP_{II}, $K_d = 6600 \pm 360$ nM, as determined by fluorescence quenching.^{118,126}

Detailed description of all the spectroscopic assays used for the evaluation of substrate binding to CRABP_{II} is provided in Chrysoula Vasileiou's Thesis.² Because CRABP_{II} binds all-*trans*-retinoic acid as the native substrate, the engineering of the rhodopsin mimic would be performed using all-*trans*-retinal instead of the rhodopsin substrate 11-*cis*-retinal. Also the use of all-

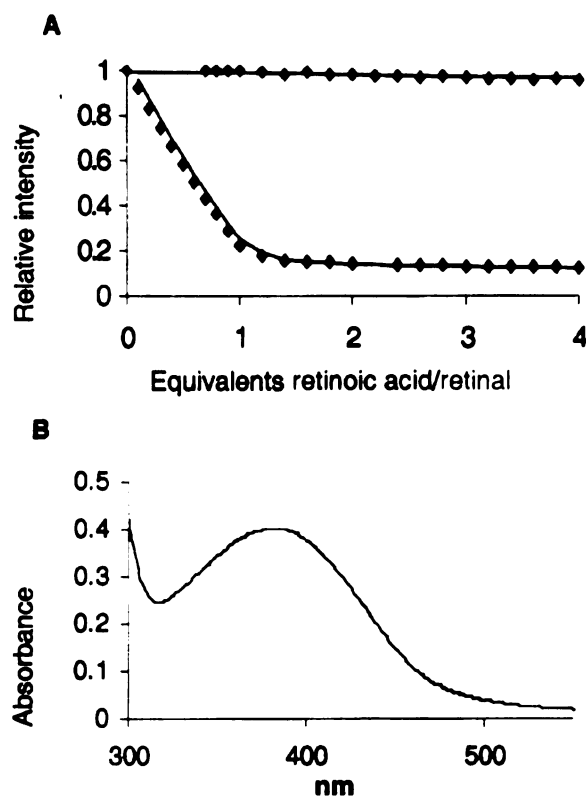


Figure 4-17. A. Fluorescence titration curves for the retinoic acid and retinal K_d determination with wild-type CRABP_{II}. Retinoic acid shows a stoichiometric binding relationship ($K_d = 2.0 \pm 1.2$ nM). Retinal shows lower binding affinity, $K_d = 6600 \pm 360$ nM.; B. UV-vis of retinal incubated with wild-type CRABP_{II} portrays a maximal absorbance at 377 nm.²

trans-retinal was preferred because of the sensitive nature of 11-*cis*-retinal. The work described next has been performed by Chrysoula Vasileiou and Rachel M. Crist (detailed explanation can be found in their Theses).²

When retinal was incubated with WT CRABP_{II} a λ_{max} of 377 nm was observed, a 3 nm shift as compared to the retinal absorbance in phosphate buffer (380 nm, Figure 4-17). There are 13 Lys residues in WT CRABP_{II} and all of them are located on the surface of the protein. Therefore, none of those Lys residues are nucleophilic enough to form a Schiff base (SB) with retinal at pH 7.3 (pK_a of Lys side chain in solution is about 10.7). MALDI-TOF analysis verified that no covalent bond species between retinal and the protein is formed. Thus the first step in the engineering of CRABP_{II} is the positioning of a nucleophilic Lys in the cavity. Assuming that the retinal binding site in the cavity will be similar to that of the retinoic acid, mutation of Arg132 directly interacting with retinoic acid (carboxylate) to a Lys residue seemed to allow formation of a SB. A minimized model of the CRABP_{II}-R132K mutant showed that the nitrogen atom of Lys132 resides $\sim 3 \text{ \AA}$ away from the retinal's carbonyl carbon, approaching from the side of the chromophore (dihedral angle $\sim 135^\circ$). This orientation is close to the optimal angle for nucleophilic attack to the carbonyl ($\sim 107^\circ$), as defined by Bürgi and Dunitz.¹²⁷ The position 134 was also considered as a possible site for the introduction of a Lys residue, however, a series of proteins containing the Y134K mutation did not yield the desired PSB formation (for details refer to Rachael M. Crist's Thesis). Mutagenesis and expression of CRABP_{II}-R132K was performed using standard molecular biology protocols. This mutation resulted in a

significant decrease in retinoic acid affinity, from 2 nM to 65 nM, as shown in Figure 4-18. At the same time the retinal binding, as determined from fluorescence quenching, increased dramatically with the dissociation constant K_d changing from 6600 nM to 280 nM. The UV spectrum of the CRABP_{II}-R132K mutant showed that the UV trace changed with time, shifting from ~392 nm, upon addition of retinal, to ~376 nm 15 min later. Second derivative analysis of the obtained spectra revealed the presence of a mixture of peaks, with a maximal at ~417 nm and ~376 nm, at a relative ratio of 1.2 to 1, respectively. The UV λ_{max}

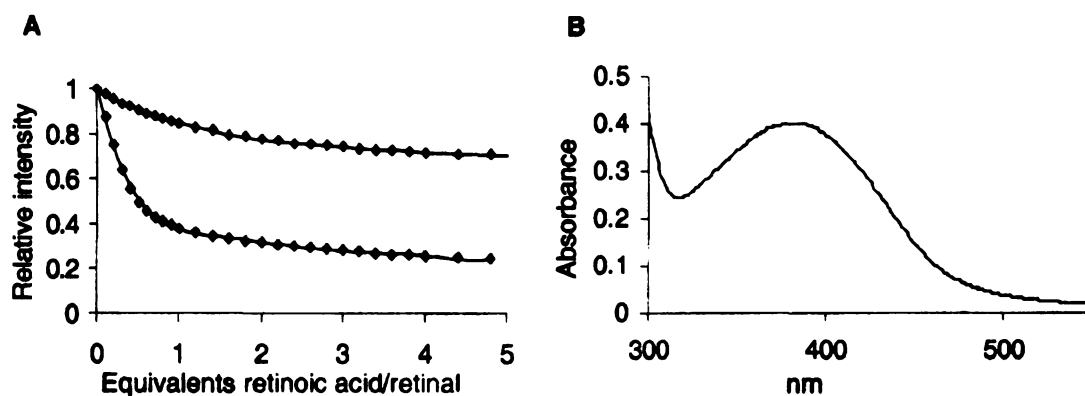


Figure 4-18. A. Fluorescence titration curves for the retinoic acid and retinal K_d determination with the single mutant CRABP_{II}-R132K. The mutation shows a drop in retinoic acid binding affinity ($K_d=65\pm14$ nM), which is concomitant with a large increase in retinal binding affinity, $K_d=280\pm17$ nM; B. UV-vis of retinal incubated with CRABP_{II} R132K. The spectrum changes with time, from 392 nm (red trace, 1 min) to 376 nm (magenta trace, 15 min).²

observed for all-*trans*-retinal bound to the R132K mutant (~417 nm) is ~40 nm red shifted as compared to the λ_{max} of free retinal in solution (~380 nm). This could be an indication of PSB formation between the engineered Lys residue and the retinal chromophore. However, the red shifted species is not stable with time and the system equilibrates into a state where the peak at 376 nm is maintained. Indication of Schiff base formation was provided by the MALDI-TOF experiments. In particular, an adduct corresponding to the SB and reductively aminated SB form is detected through mass spec analysis. The same experiments with the WT-CRABP II had failed to detect any mass adduct. These results suggested that the engineered Lys residue does yield a SB.

The next step to optimize the rhodopsin mimic was to increase the hydrophobicity in the cavity, especially to increase the nucleophilicity of the Lys132. The two candidate residues to mutate were Ser12 and Tyr134, considering that Ser12 is further away from the Lys residue of interest (5.01 Å), Tyr134 (3.39 Å away) was the immediate target for mutagenesis. Keeping these observations under consideration, Tyr134 was mutated into a Phe residue (CRABP II-R132K:Y134F). As shown in Figure 2-19, the second mutation resulted in a significant increase of the retinal binding affinity, with the calculated K_d value at 120 ± 5 nM. In addition, the retinoic acid affinity has further declined ($K_d=100 \pm 7$ nM), consistent with the removal of one of the critical residues responsible for binding. The UV titration of the new double mutant R132K:Y134F with all-*trans*-retinal revealed the presence of a peak at 396 nm, representing a ~15 nm red shifting as compared to the free retinal spectrum (λ_{max}

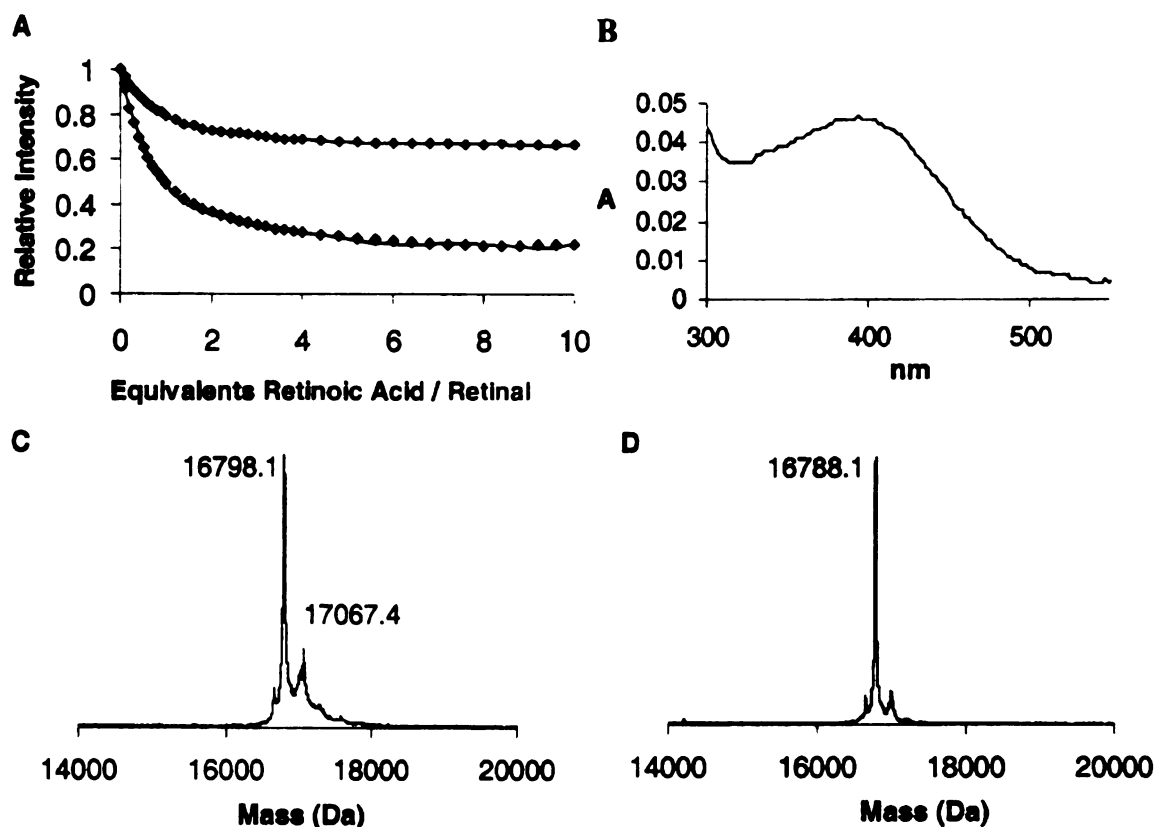


Figure 4-19. A. Fluorescence titration curves for the retinoic acid and retinal K_d determination with the double mutant CRABP II R132K::Y134F. The retinoic acid binding affinity has dropped ($K_d = 100 \pm 7.1$ nM), while the retinal binding affinity increased ($K_d = 120 \pm 4.9$ nM). B. UV-vis of retinal incubated with CRABP II R132K::Y134F portrays a maximal absorbance at 396 nm. C. MALDI-TOF of the retinal incubated double mutant shows an $M^+ = 16798.1$ Da corresponding to the protein mass, calculated to be 16785.2 Da. A second peak, $M^+ = 17067.4$ Da, 269.3 mass units higher than the protein peak, represents a covalently linked retinal molecule (calculated mass difference 266). D. Reductive amination of retinal with CRABP II R132K::Y134F results in loss of the covalent complex and only the protein mass $M^+ = 16788.1$ Da can be detected. The small peak to the right of the major protein peak is due to matrix addition to the protein, and does not correspond to the calculated retinal addition molecular weight.

~380 nm). However, the MALDI-TOF based binding assays did not support the presence of a stable bond between the CRABPII-R132K:Y134F mutant and all-*trans*-retinal (Figure 4-19C). The crystal structures obtained by our collaborators (Soheila Vaezeslami, and Professor James Geiger, Michigan State University) also support the above conclusion. The structures of CRABPII-R132K:Y134F bound to both all-*trans*-retinoic acid and all-*trans*-retinal were solved (Figure 4-20). The bound retinoic acid is located at exactly the same position as in WT-CRABPII. Two of the residues originally implicated in direct interactions with the carboxylic acid have been removed (Arg132, Tyr134), however the interaction with the water bound to Arg111 and Thr54 is maintained (water #1, 2.7 Å). Double mutant R132K:Y134F failed to produce evidence for SB formation with retinal. A variety of different mutants were prepared in the process of designing a rhodopsin surrogate. In the process it was observed that replacing Arg111 with a hydrophobic residue was a critical mutation for the formation of a PSB. It is believed that the removal of the guanidine group results in the loss of the water molecule that forms a hydrogen bond between the retinal and the Arg111. It was suggested that by removing the interaction between the retinal and the water, the carbonyl could rotate to a position in which favors the nucleophilic attack of Lys (optimal Bürgi and Dunitz angle of 107°, Figure 4-21).

Figure 4-21 shows the model structures of the CRABPII-R132K:R111L mutant with all-*trans*-retinal bound either non-covalently (A) or as a Schiff base (B). Lys132 is in close proximity to Tyr134 (2.8 Å), suggesting that Tyr134 can activate the primary amine and facilitate the nucleophilic attack. The necessary



Figure 4-20. **A.** crystal structure of CRABPII-R132K::Y134F with all-*trans*-retinal. Two independent conformations of both retinal and Ly132 are found. The chromophore exists in the free form, no SB formations is observed. **B.** Overlay of the crystal structures of CRABPII-R132K::Y134F bound to all-*trans*-retinal and all-*trans*-retinoic acid (represented by grey carbons and green carbons, respectively). **C.** Overlay of the WT-CRABPII and CRABPII-R132K::Y134F mutant bound to all-*trans*-retinoic acid (represented by orange carbons and green carbons, respectively).

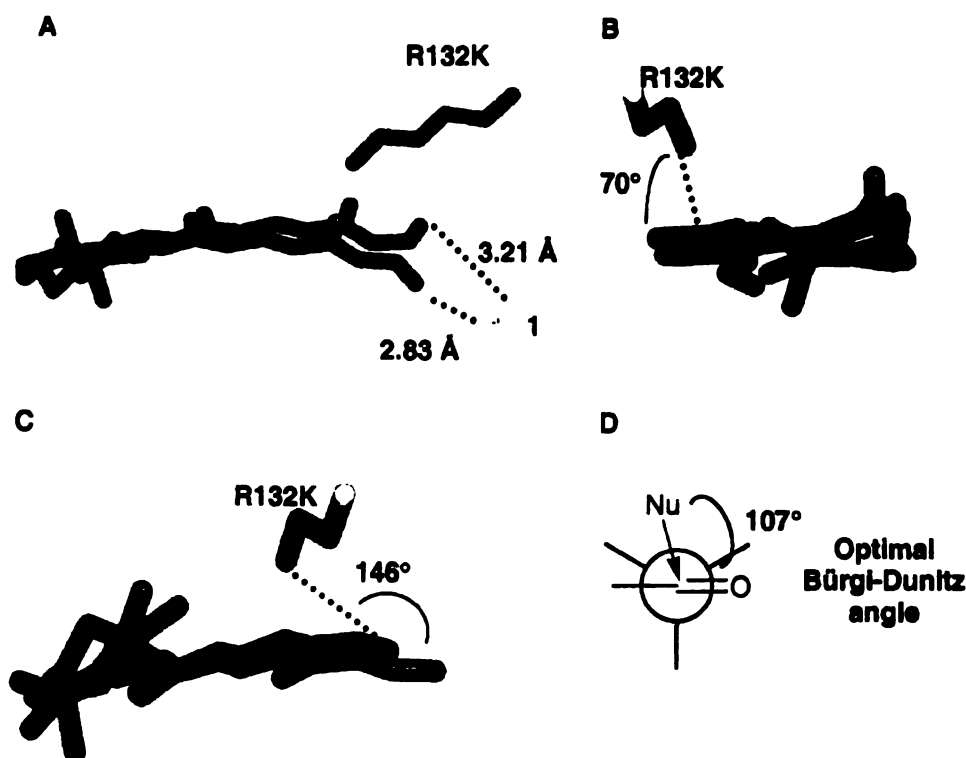


Figure 4-21. A. Closer look at the conformation of bound retinal to the R132K:Y134F mutant. The two conformations are represented by different colors. B. Lys132 can attack the 1st conformation of the carbonyl from a ~70° angle. For optimal attack a rotation of ~30° is required. C. Lys132 can attack the 2nd conformation of the carbonyl from a ~146° angle. For optimal attack a rotation of ~20° is required; D. Bürgi-Dunitz angle for optimal nucleophilic attack on a carbonyl.

rotation of the carbonyl can occur since the water molecule associated with Arg111 has been removed. This could explain why the Schiff base formation appears to be more successful when the R132K:R111L mutant is used. After the SB is formed, the phenolic OH of Tyr134 can be close to the imine nitrogen on a

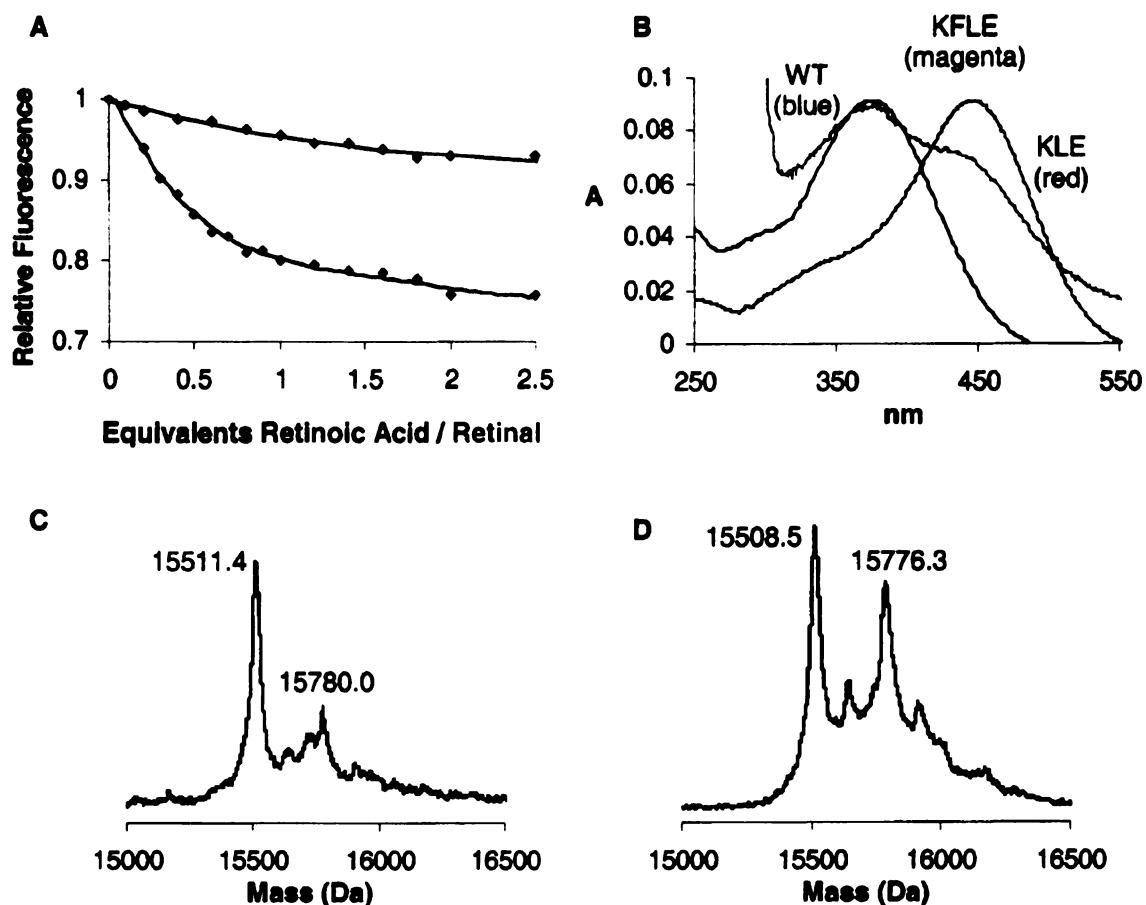


Figure 4-22. A. Fluorescence titrations of retinoic acid ($K_d = 426 \pm 47$ nM) and retinal ($K_d = 1.4 \pm 4.9$ nM) with the CRABP II mutant R132K::R111L::L121E; B. UV/Vis spectrum of retinal bound to WT CRABP II (blue trace) and the R132K::Y134F::R111L::L121E (magenta trace, KFLE), R132K::R111L::L121E (red trace, KLE) mutant. Retinal bound with wild-type absorbs at 377 nm. The UV of retinal with the KLE mutant reveals a 449 nm peak, indicating formation of PSB. The UV for the KFLE mutant is a mixture of two peaks at 378 nm and 438 nm; C. MALDI-TOF of the incubation of retinal with R132K::R111L::L121E. In addition to the apo protein peak (15511.4 Da, calculated value 15508.2 Da) a second peak at 15780.0 Da is found, that corresponds to the covalent complex of the mutant with retinal; D. Reductive amination of retinal with the R132K::R111L::L121E mutant reveals a mixture of the apo-protein (15508.5 Da) as well as the covalent complex of the protein with retinal (15776.3 Da).²

trajectory that can favor stabilization of a PSB. However, since the system is not stable, positioning of a counter anion to further stabilize the protonated imine is essential.

Research on the bovine rhodopsin suggests that the position of the counter anion around the imine is critical for the PSB stabilization.^{59-61,65,128} This was further proved by the recent crystal structure of the bovine rhodopsin which reveals the presence of a Glu113 4.2 Å away from the Schiff base nitrogen.⁷⁵ Thus, the same principle was used and the model of the covalently bound retinal suggested that the best position for introducing a counter anion is residue 121. In the course of designing a rhodopsin mimic, it was found that Tyr134 participates in the activation of both retinal and the PSB counter anion (Glu121). The



Figure 4-23. Crystal structure of the binding cavity of CRABPII-R132K:R111L:L121E bound to all-*trans*-retinal. The retinal is confirmed to be bound as a PSB via Lys132. The important distances around the imine are highlighted.

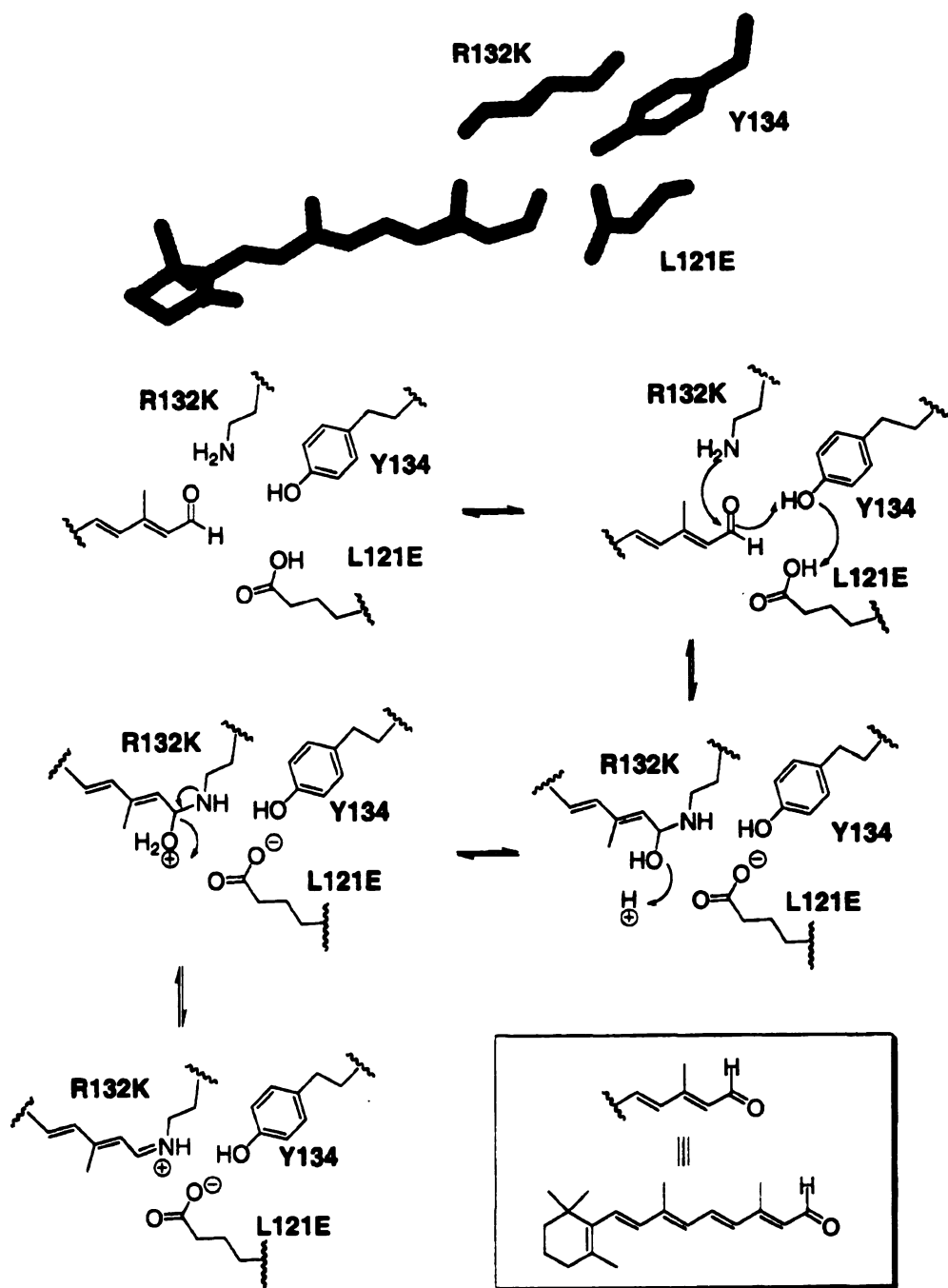


Figure 4-24. Proposed mechanism for the Schiff base formation between triple mutant CRABPII-R132K: R111L:L121E and all-*trans*-retinal.

spectroscopic results for the CRABP II triple mutant R132K: R111L:L121E are presented in Figure 4-22. The protein has a remarkable affinity for retinal as evident from the extremely low dissociation constant ($K_d=1.36\pm4.9$ nM) obtained by fluorescence quenching. At the same time, the UV-vis trace in the presence of all-*trans*-retinal reveals the presence of a single peak at 449 nm, a 72 nm red shift as compared to the trace of retinal in the presence of WT-CRABP II . The crystal structure of CRABP II -R132K:R111L::L121E bound to all-*trans*-retinal was obtained by our collaborators (Soheila Vaezslami, and Professor James Geiger, Michigan State University) is shown in Figure 4-23.

Based on the spectroscopic and crystallographic data gathered, the following mechanism was proposed for the Schiff base formation between all-*trans*-retinal and the CRABP II -R132K:R111L:L121E triple mutant. As shown in Figure 4-24, Tyr134 activates the bound aldehyde for nucleophilic attack by Lys132, followed by elimination of water and Schiff base formation. The imine remains protonated due to the close proximity of Glu121 resulting in a Glu121–Schiff base salt bridge.

4.4 References

1. H. Shichi, *Biochemistry of vision*. ed.; Academic Press, Inc: 1983.; 'Vol.' p.
2. C. Vasileou. Protein Design:Reengineering Cellular Retinoic Acid Binding Protein II into a Retinal Binding Protein. Michigan State University, East Lansing, 2005.
3. J. M. Tiffany; P. Noorjahan; A. A. Abdel-Latif; J. J. Harding; R. Mayne; R. G. Brewton; Z. X. Ren, *Biochemistry of the Eye*. ed.; Chapman and Hall: London, UK, 1997; 'Vol.' p.
4. F. Crescitelli, *Handbook of Sensory Physiology*,. ed.; Springer-Verlag.: New York , NY, 1972,; 'Vol.' Vol. VII, p 245-363.
5. B. Wissinger; L. T. Sharpe, "New aspects of an old theme: The genetic basis of human color vision." *American Journal of Human Genetics* **1998**, *63*, (1257-1262).
6. E. N. Pugh, "Rods are rods and cones are cones, and (never) the twain shall meet." *Neuron* **2001**, *32*, 375-380.
7. R. Hoffmann, *The Same and Not the Same*. ed.; Columbia University Press:: New York, NY,, 1995; 'Vol.' p.
8. G. Wald, "Pigments of the retina." *Journal of General Physiology* **1935**, *19*, 781-795.
9. R. A. Morton; T. W. Goodwin, "Preparation of retinene in vitro." *Nature* **1944**, *153*, 405-406.
10. F. D. Collins, "Rhodopsin and indicator yellow." *Nature* **1953**, *171*, 469-472.
11. R. A. Morton; G. A. J. Pitt, "Rhodopsin. IX. pH and the hydrolysis of indicator yellow." *Biochemical Journal* **1955**, *59*, 128-134.

12. R. R. Rando, "Polyenes and vision." *Journal of Chemical Biology* **1996**, *3*, 255-262.
13. J. Nathans, "Rhodopsin: Structure, function and genitics." *Biochemical Journal* **1992**, *31*, (21), 4923-4931.
14. G. L. Wheeler; M. W. Bitensky, "A light-activated GTPase in vertebrate photoreceptors: Regulation of light-activated cyclic GMP phosphodiesterase." *Proceedings of the National Academy of Sciences of the United States of America* **1977**, *74*, 4238-4242.
15. P. A. Liebman; E. N. Pugh, "The control of phosphodiesterase in rod disk membranes: kinetics, possible mechanisms and significance for vision." *Vision Research* **1979**, *19*, 375-380.
16. B. Fung; L. Stryer, "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* **1980**, *77*, 2500-2504.
17. B. Fung; J. B. Hurley; L. Stryer, "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* **1981**, *78*, 152-156.
18. E. E. Fesenko; S. S. Kolesnikov; A. L. Lyubarsky, "Induction by cyclic GMP of cationic conductance in plasma membrane of retinal rod outer segment." *Nature* **1985**, *313*, 310-313.
19. L. Haynes; K.-W. Yau, "Cyclic GMP-sensitive conductance in outer segment membrane of catfish cones." *Nature* **1985**, *317*, 61-64.
20. T. P. Sakmar, "Rhodopsin: A prototypical G protein-coupled receptor." *Progress in Nucleic Acid Research* **1998**, *59*, 1-34.
21. W. J. Degrip; D. Gray; J. Gillespie; P. H. M. Bovee; E. M. M. Vandenberg; J. Lugtenburg; K. J. Rothschild, "Photoexcitation of rhodopsin -Conformation changes in the chromophore, protein and associated lipids as determined by ftir difference spectroscopy." *Photochemistry and Photobiology* **1988**, *48*, 497-504.

22. A. Cooper, "Energy uptake in the 1st step of visual excitation." *Nature* **1979**, 282, 531-533.
23. R. R. Birge, "Photophysics of light transduction in rhodopsin and bacteriorhodopsin." *Annual Review in Biophysics and Biology* **1981**, 10, 315-354.
24. R. Mathies, "Resonance raman-spectroscopy of rhodopsin and bacteriorhodopsin isotopic analogs." *Methods in Enzymology* **1982**, 88, 633-643.
25. T. Shieh; M. Han; T. P. Sakmar; S. O. Smith, " The steric trigger in rhodopsin activation." *Journal of Molecular Biology* **1997**, 269, 373-384.
26. B. Yan; K. Nakanishi; J. L. Spudich, "Mechanism of activation of sensory rhodopsin-I - evidence for a steric trigger." *Proceedings of the National Academy of Sciences of the United States of America* **1991**, 88, 9412-9416.
27. I. Szundi; T. L. Mah; J. W. Lewis; S. Jager; O. P. Ernst; K. P. Hofmann; D. S. Kliger, "Proton transfer reactions linked to rhodopsin activation." *Biochemistry* **1998**, 37, 14237-14244.
28. S. Arnis; K. P. Hofmann, "2 Different forms of metarhodopsin-II-Schiff-base deprotonation precedes proton uptake and signaling state." *Proceedings of the National Academy of Sciences of the United States of America* **1993**, 90, 7849-7853.
29. P. P. M. Schnetkamp, "How does the retinal rod Na-Ca + K exchanger regulate cytosolic free Ca²⁺?" *Journal of Biological Chemistry* **1995**, 270, 13231-13239.
30. R. R. Rando, "The chemistry of vitamin A and vision." *Angewandte Chemie International Edition* **1990**, 29, 461-480.
31. D. A. Moffet; M. A. Case; J. C. House; K. Vogel; R. D. Williams; T. G. Spiro; G. L. McLendon; M. H. Hecht, "Carbon monoxide binding by de novo heme proteins derived from designed combinatorial libraries." *Journal of the American Chemical Society* **2001**, 123, (10), 2109-2115.

32. R. B. Penn; R. M. Pascual; Y. M. Kim; S. J. Mundell; V. P. Krymskaya; R. A. Panettieri; J. L. Benovic, " Arrestin specificity for G protein-coupled receptors in human airway smooth muscle." *Journal of Biological Chemistry* **2001**, 276, 32648-32656.
33. G. A. Schick; T. M. Cooper; R. A. Holloway; L. P. Murray; R. R. Birge, "Energy-storage in the primary photochemical events of rhodopsin and isorhodopsin." *Biochemistry* **1987**, 26, (9), 2556-2562.
34. A. Cooper; C. A. Converse, "Energetics of primary processes in visual excitation -photocalorimetry of rhodopsin in rod outer segment membranes." *Biochemistry* **1976**, 15, (14), 2970-2978.
35. N. Miki; J. J. Keirns; F. R. Marcus; J. Freeman; M. W. Bitensky, "Regulation of cyclic nucleotide concentrations in photoreceptors-ATP-dependent stimulation of cyclic nucleotide phosphodiesterase by light." *Proceedings of the National Academy of Sciences of the United States of America* **1973**, 70, (12), 3820-3824.
36. T. D. Lamb, "Gain and kinetics of activation in the G-protein cascade of phototransduction." *Proceedings of the National Academy of Sciences of the United States of America* **1996**, 93, (2), 566-570.
37. S. Hecht; S. Shlaer; M. H. Pirenne, "Energy, quanta, and vision." *Journal of General Physiology* **1942**, 25, 819-840.
38. Schneider, "Spectral variation in the photosensitivity of visual purple." *Proceedings of the Royal Society A* **1939**, 170, 102-112.
39. J. E. Kim; M. J. Tauber; R. A. Mathies, "Wavelength dependent cis-trans isomerization in vision." *Biochemistry* **2001**, 40, (46), 13774-13778.
40. H. J. A. Dartnall, "The photosensitivities of visual pigments in the presence of hydroxylamine." *Vision Research* **1968**, 8, 339-358.
41. H. J. A. Dartnall; C. F. Goodeve; R. J. Lythgoe, "The quantitative analysis of the photochemical bleaching of visual purple solutions in monochromatic light." *Proceedings of the Royal Society A* **1936**, 156, 158-170.

42. S. L. Merbs; J. Nathans, "Absorption-Spectra of Human Cone Pigments." *Nature* **1992**, *356*, 433-435.
43. D. D. Oprian; A. B. Asenjo; N. Lee; S. L. Pelletier, "Design, chemical synthesis, and expression of genes for the 3 human color-vision pigments." *Biochemistry* **1991**, *30*, 11367-11372.
44. P. E. Blatz; N. Baumgartner; B. Balasubramanian; P. Balasubramanian; F. Stedman, "Wavelength regulation in visual pigment chromophore. Large induced bathochromic shifts in retinol and related polyenes." *Photochemistry and Photobiology* **1971**, *14*, 531-549.
45. G. A. J. Pitt; F. D. Collins; R. A. Morton; P. Stok, "Rhodopsin. VIII. N-Retinyldenemethylamine, an indicator yellow analog." *Biochemical Journal* **1955**, *59*, 122-128.
46. P. E. Blatz; J. H. Mohler; W. Ahmed, "Spectroscopic observation of solvent interaction with selected retinal Schiff-bases." *Photochemistry and Photobiology* **1991**, *54*, 255-264.
47. P. A. Hargrave; J. H. McDowell; D. R. Curtis; J. K. Wang; E. Juszczak; S. L. Fong; J. K. M. Rao; R. Argos, "The structure of bovine rhodopsin." *Biophysics of Structure and Mechanism* **1983**, *9*, 235-244.
48. P. A. Hargrave, "The amino-terminal tryptic peptide of bovine rhodopsin. A glycopeptide containing two sites of oligosaccharide attachment." *Biochimica and Biophysics Acta* **1977**, *492*, 83-94.
49. K. Palczewski; T. Kumasaka; T. Hori; C. A. Behnke; H. Motoshima; B. A. Fox; I. Le Trong; D. C. Teller; T. Okada; R. E. Stenkamp; M. Yamamoto; M. Miyano, "Crystal structure of rhodopsin: A G protein-coupled receptor." *Science* **2000**, *289*, 739-745.
50. M. Akhtar; D. Blosse; P. Dewhurst, "Studies on vision. The nature of the retinal-opsin linkage." *Biochemical Journal* **1968**, *110*, 693-702.
51. D. Bownds, "Site of attachment of retinal in rhodopsin." *Nature* **1967**, *216*, 1178-1181.

52. R. E. Hubbard *Proceedings of the Natural and Physics Laboratory Symposium*; London, 1958; p[^]pp.
53. A. Kropf; R. Hubbard, "The mechanism of bleaching rhodopsin." *Annals of the New York Academy of Sciences* **1958**, *74*, 266-280.
54. A. R. Oseroff; R. Callender, "Rh "Resonance Raman-Spectroscopy of Rhodopsin in Retinal Disk Membranes"." *Biochemistry* **1974**, *13*, 4243-4248.
55. B. Honig; U. Dinur; K. Nakanishi; V. Balogh-Nair; M. A. Gawinowicz; M. Arnaboldi; M. Motto, "An external point-charge model for wavelength regulation in visual pigments." *Journal of the American Chemical Society* **1979**, *101*, 7084-7086.
56. K. Nakanishi, "Bioorganic studies with rhodopsin." *Pure Applied Chemistry* **1985**, *57*, 769-776.
57. L. C. P. J. Mollevanger; A. P. M. Kentgens; J. A. Pardoën; J. M. L. Courtin; W. S. Veeman; J. Lugtenburg; W. J. Degrip, "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* **1987**, *163*, 9-14.
58. M. Arnaboldi; M. G. Motto; K. Tsujimoto; V. Balogh-nair; K. Nakanishi, "Hydro-retinals and hydro-rhodopsins." *Journal of the American Chemical Society* **1979**, *101*, 7082-7084.
59. E. A. Zhukovsky; D. D. Oprian, "Effect of carboxylic side chains on the absorption maximum of visual pigments." *Science* **1989**, *246*, 928-930.
60. T. P. Sakmar; R. R. Franke; H. G. Khorana, "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* **1989**, *86*, 8309-8313.
61. J. Nathans, "Determinants of visual pigment absorbance: Role of charged amino acids in the putative transmembrane segments." *Biochemistry* **1990**, *29*, 937-942.

62. P. E. Blatz; J. H. Mohler; H. V. Navangul, ""Anion-induced wavelength regulation of absorption maxima of Schiff-bases of retinal." *Biochemistry* **1972**, *11*, 848-856.
63. Y. Gat; M. Sheves, "A mechanism for controlling the pKa of the retinal protonated Schiff base in retinal proteins. A study with model compounds." *Journal of the American Chemical Society* **1993**, *115*, 3772-3773.
64. B. Honig; A. Greenberg; U. Dinur; T. Ebrey, "Visual pigment spectra: Implications of the protonation of the retinal Schiff base." *Biochemistry* **1976**, *15*, 4593-4499.
65. J. Nathans; D. Thomas; D. S. Hogness, "Molecular-genetics of human color-vision - the genes encoding blue, green, and red pigments." *Science* **1986**, *232*, 193-202.
66. P. E. Blatz; P. A. Liebman, "Wavelength regulation in visual pigments." *Experimental Eye Research* **1973**, *17*, 573-580.
67. P. J. E. Verdegem; P. H. M. Bovee-Geurts; W. J. d. Grip; J. Lugtenburg; H. J. M. d. Groot, "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* **1999**, *38*, 11316-11324.
68. Q. Tan; J. H. Lou; B. Borhan; E. Karnaukhova; N. Berova; K. Nakanishi, "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* **1997**, *36*, 2089-2093.
69. B. Honig; B. Hudson; B. D. Sykes; M. Karplus, "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* **1971**, *68*, 1289-1293.
70. G. F. X. Schertler; P. A. Hargrave; P. C. Edwards; V. M. Unger, "The arrangement of alpha helices in rhodopsin." *Investigative Ophthalmology and Visual. Science* **1997**, *38*, 3379-3379.

71. V. M. Unger; P. A. Hargrave; J. M. Baldwin; G. F. X. Schertler, "Arrangement of rhodopsin transmembrane alpha-helices." *Nature* **1997**, *389*, 203-206.
72. G. F. X. Schertler; P. A. Hargrave; V. M. Unger, "Three dimensional structure of rhodopsin obtained by electron cryo-microscopy." *Investigative Ophthalmology and Visual. Science* **1996**, *37*, 3701-3701.
73. J. M. Baldwin; G. F. X. Schertler; V. M. Unger, "An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors." *Journal of Molecular Biology* **1997**, *272*, 144-164.
74. G. F. X. Schertler; C. Villa; R. Henderson, "Projection structure of rhodopsin." *Nature* **1993**, *362*, 770-772.
75. K. Palczewski; T. Kumasaka; T. Hor; C. A. Behnke; H. Motoshima; B. A. Fox; I. L. Trong; D. C. Teller; T. Okada; R. E. Stenkamp; M. Yamamoto; M. Miyano, "Crystal structure of rhodopsin: A G protein-coupled receptor." *Science* **2000**, *289*, 739-745.
76. K. Nakanishi; V. Balogh-Nair; M. Arnaboldi; K. Tsujimoto; B. Honig, "An external point charge model for bacteriorhodopsin to account for its purple color." *Journal of the American Chemical Society* **1980**, *102*, 7954-7947.
77. C. N. Rafferty, "Light-induced perturbation of aromatic residues in bovine rhodopsin and bacteriorhodopsin." *Photochemistry and Photobiology* **1979**, *29*, 109-120.
78. T. Ebrey; M. Tsuda; G. Sassenrath; J. L. West; W. H. Waddell, "Light activation of bovine rod phosphodiesterase by non-physiological visual pigments." *FEBS Letters* **1980**, *116*, 217-219.
79. A. B. Asenjo; J. Rim; D. D. Oprian, "Molecular determinants of human red/green color discrimination." *Neuron* **1994**, *12*, 1131-1138.
80. M. Neitz; J. Neitz; G. H. Jacobs, "Spectral tuning of pigments underlying red-green color vision." *Science* **1991**, *252*, 971-974.

81. R. Yokoyama; S. Yokoyama, "Convergent evolution of the red- 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* **1990**, *87*, 9315-9318.
82. T. Chan; M. Lee; T. P. Sakmar, "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* **1992**, *267*, 9478-9480.
83. T. Yoshizawa, "The road to color-vision-structure, evolution and function of chicken and gecko visual pigments." *Photochemistry and Photobiology* **1992**, *56*, 859-867.
84. S. W. Lin; G. G. Kochendoerfer; K. S. Carool; D. Wang; R. A. Mathies; T. P. Sakmar, "Mechanisms of spectral tuning in blue cone visual pigments." *Journal of Biological Chemistry* **1998**, *273*, 24583-24591.
85. J. M. Janz; D. L. Farrens, "Engineering a functional blue-wavelength-shifted rhodopsin mutant." *Biochemistry* **2001**, *40*, 7219-7227.
86. J. I. Fasick; N. N. Lee; D. D. Oprian, "Spectral tuning in the human blue cone pigment." *Biochemistry* **1999**, *38*, 11593-11596.
87. M. H. Hecht; M. W. West; S. Roy; N. R. Rojas; S. Kamtekar; J. McLean; K. J. Helmer; C. T. Simons; J. R. Beasley, "Design or libraries of de novo proteins by binary patterning of polar and nonpolar amino acids." *Protein Engineering* **1997**, *10*, 20-20.
88. C. Kiefer; S. Hessel; J. M. Lampert; K. Vogt; M. O. Lederer; D. E. Breithaupt; J. von Lintig, "Identification and characterization of a mammalian enzyme catalyzing the asymmetric oxidative cleavage of provitamin A." *Journal of Biological Chemistry* **2001**, *276*, (17), 14110-14116.
89. M. H. Hecht, "Structures and functions of de novo proteins from designed combinatorial libraries." *Abstracts of Papers of the American Chemical Society* **2003**, *226*, U271-U271.

90. M. H. Hecht, "De novo proteins from rationally designed combinatorial libraries." *Abstracts of Papers of the American Chemical Society* **1999**, 217, U603-U603.
91. M. H. Hecht, "De-Novo Design of Beta-Sheet Proteins." *Proceedings of the National Academy of Sciences of the United States of America* **1994**, 91, (19), 8729-8730.
92. M. Roth; A. Jeltsch, "Changing the target base specificity of the EcoRV DNA methyltransferase by rational de novo protein-design." *Nucleic Acids Research* **2001**, 29, (15), 3137-3144.
93. U. T. Bornscheuer; M. Pohl, "Improved biocatalysts by directed evolution and rational protein design." *Current Opinion in Chemical Biology* **2001**, 5, (2), 137-143.
94. T. Lanio; A. Jeltsch; A. Pingoud, "On the possibilities and limitations of rational protein design to expand the specificity of restriction enzymes: a case study employing EcoRV as the target." *Protein Engineering* **2000**, 13, (4), 275-281.
95. J. P. Caradonna; A. L. Pinto; C. D. Coldren; Y. Wang, "Incorporation of metal-based active sites into host proteins by rational protein design methods." *Journal of Inorganic Biochemistry* **1999**, 74, (1-4), 16-16.
96. H. W. Hellinga, "Construction of a blue copper analogue through iterative rational protein design cycles demonstrates principles of molecular recognition in metal center formation." *Journal of the American Chemical Society* **1998**, 120, (39), 10055-10066.
97. H. W. Hellinga, "Rational protein design: Combining theory and experiment." *Proceedings of the National Academy of Sciences of the United States of America* **1997**, 94, (19), 10015-10017.
98. A. L. Pinto; H. W. Hellinga; J. P. Caradonna, "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* **1997**, 94, (11), 5562-5567.

99. G. Kakefuda; J. G. Kwagh; K. H. Ott; G. Stockton; V. Sidorov, "Rational, protein design of herbicide resistant acetohydroxyacid synthase." *Plant Physiology* **1996**, *111*, (2), 49-49.
100. M. A. Dwyer; L. L. Looger; H. W. Hellinga, "Computational design of a Zn²⁺ receptor that controls bacterial gene expression." *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 11255-11260.
101. A. J. Chirino; M. L. Ary; S. A. Marshall, "Minimizing the immunogenicity of protein therapeutics." *Drug Discovery Today* **2004**, *9*, 82-90.
102. S. A. Marshall; G. A. Lazar; A. J. Chirino; J. R. Desjarlais, "Rational design and engineering of therapeutic proteins." *Drug Discovery Today* **2003**, *8*, 212-221.
103. J. Bastien; C. Rochette-Egly, "Nuclear retinoid receptors and the transcription of retinoid-target genes." *Gene* **2004**, *328*, 1-16.
104. S. M. Lipkin; T. L. Grider; R. A. Heyman; C. K. Glass; F. H. Gage, "Constitutive retinoid receptors expressed from adenovirus vectors that specifically activate chromosomal target genes required for differentiation of promyelocytic leukemia and teratocarcinoma cells." *Journal of Virology* **1996**, *70*, (10), 7182-7189.
105. M. F. Boehm; L. Zhang; B. A. Badea; S. K. White; D. E. Mais; E. Berger; C. M. Suto; M. E. Goldman; R. A. Heyman, "Synthesis and Structure-Activity-Relationships of Novel Retinoid-X Receptor-Selective Retinoids." *Journal of Medicinal Chemistry* **1994**, *37*, (18), 2930-2941.
106. D. J. Mangelsdorf; R. M. Evans; R. A. Heyman, "RXR - Retinoid Receptor or Cofactor." *Calcified Tissue International* **1994**, *54*, (4), 331-331.
107. D. J. Mangelsdorf; C. Thummel; M. Beato; P. Herrlich; G. Schutz; K. Umesono; B. Blumberg; P. Kastner; M. Mark; P. Chambon; R. M. Evans, "The Nuclear Receptor Superfamily - the 2nd Decade." *Cell* **1995**, *83*, (6), 835-839.

108. P. Krieg; S. Feil; G. Furstenberger; G. T. Bowden, "Tumor-specific verexpression of a novel keratinocyte lipid-binding protein - identification and characterization of a cloned sequence activated during multistage carcinogenesis in mouse skin." *Journal of Biological Chemistry* **1993**, 268, 17362-17369.
109. M. Buelte; L. Banaszak; D. A. Bernlohr, "Regulation of adipocyte lipid-binding protein-phosphorylation by fatty-acids." *FASEB Journal* **1991**, 5, A417-A417.
110. R. A. Morton, *Photochemistry of Vision*. ed.; Springer-Verlag: New York, 1972; 'Vol.' p 33.
111. C. Merlet-Benichou; J. Vilar; M. Lelievre-Pegorier; T. Gilbert, "Role of retinoids in renal development: pathophysiological implication." *Current Opinion in Nephrology and Hypertension* **1999**, 8, (1), 39-43.
112. P. Palozza; e. al., "Regulation of cell cycle progression and apoptosis by b-carotene in undifferentiated and differentiated HL-60 leukemia cells: Possible involvement of a redox machanism." *International Journal of Cancer* **2002**, 97, 593-600.
113. L. J. Gudas, *In The retinoids: Biology chemistry and medicine*. ed.; New York, 1994; 'Vol.' p 443-520.
114. A. C. Ross, "Symposium-retinoids-cellular-metabolism and activation-introduction." *Journal of Nutrition* **1993**, 123, (2), 344-345.
115. B. N. Chaudhuri; G. J. Kleywegt; I. Broutin-L'Hermite; T. Bergfors; H. Senn; P. L. Motte; O. Partouche; T. A. Jones, "Structures of cellular retinoic acid binding proteins I and II in complex with synthetic retinoids." *Acta Crystallography D* **1999**, 55, 1850-1857.
116. J. S. Bailey; C. H. Siu, "Purification and partial characterization of a novel binding-protein for retinoic acid from neonatal rat." *Journal of Biological Chemistry* **1988**, 263, 9326-9332.
117. D. E. Ong; F. Chytil, "Cellular retinol-binding protein from rat-liver-purification and characterization." *Journal of Biological Chemistry* **1978**, 253, 828-832.

118. L. C. Wang; Y. Li; H. G. Yan, "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* **1997**, 272, 1541-1547.
119. G. J. Kleywegt; T. Bergfors; H. Senn; P. L. Motte; B. Gsell; K. Shudo; T. A. Jones, "Crystal structures of cellular retinoic acid binding proteins I and II in complex with all-trans retinoic acid and a synthetic retinoid." *Structure* **1994**, 2, 1241-1258.
120. X. Chen; M. Tordova; G. L. Gilliland; L. C. Wang; Y. L. G. Yan; X. H. Ji, "Crystal structure of apo-cellular retinoic acid-binding protein type II (R111M) suggests a mechanism of ligand entry." *Journal of Molecular Biology* **1998**, 278, 641-653.
121. J. Rizo; Z. P. Liu; L. M. Gierasch, "Structure and folding of cellular retinoic acid-binding protein, a predominantly beta-sheet protein." *Biophysical Journal* **1994**, 66, A397-A397.
122. J. Rizo; Z. P. Liu; L. M. Gierasch, "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* **1994**, 4, 741-760.
123. L. C. Wang; H. G. Yan, "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* **1998**, 37, 13021-13032.
124. A. C. M. Young; G. Scapin; A. Kromminga; S. B. Patel; J. H. Veerkamp; J. C. Sacchettini, "Structural studies on human muscle fatty-acid-binding protein at 1.4-Angstrom resolution-binding interactions with 3 C18 fatty-acids." *Structure* **1994**, 2, 523-534.
125. G. J. Kleywegt; T. Bergfors; H. Senn; P. Le Motte; B. Gsell; K. Shudo; T. A. Jones, "Crystal structures of cellular retinoic acid binding proteins I and II in complex with all-trans retinoic acid and a synthetic retinoid." *Structure* **1994**, 2, 1241-1258.
126. A. W. Norris; L. Cheng; V. Giguere; M. Rosenberger; E. Li, "Measurement of subnanomolar retinoic acid-binding affinities for cellular retinoic acid-binding proteins by fluorometric titration." *Biochimia*

Biophysics Acta-Protein Structure and Molecular Enzymology **1994**, *1209*, 10-18.

127. M. M. Midland; Y. C. Kwon, "Stereochemistry of hydroboration of a-chiral olefins and reduction of a-chiral ketones. An unusual anti-Cram selectivity with dialkylboranes." *Journal of the American Chemical Society* **1983**, *105*, 3725-3727.
128. A. Terakita; M. Koyanagi; H. Tsukamoto; T. Yamashita; T. Miyata; Y. Shichida, "Counterion displacement in the molecular evolution of the rhodopsin family." *Nature Structural Molecular Biology* **2004**, *11*, 284-289.

**STUDIES ON 15-15'- β -CAROTENE DIOXYGENASE AND REENGINEERING
CELLULAR RETINOIC ACID BINDING PROTEIN II INTO A RETINAL
BINDING PROTEIN AND ITS INTERACTION WITH RETINAL MIMICS**

VOLUME II

By

Montserrat Rabago-Smith

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Chemistry

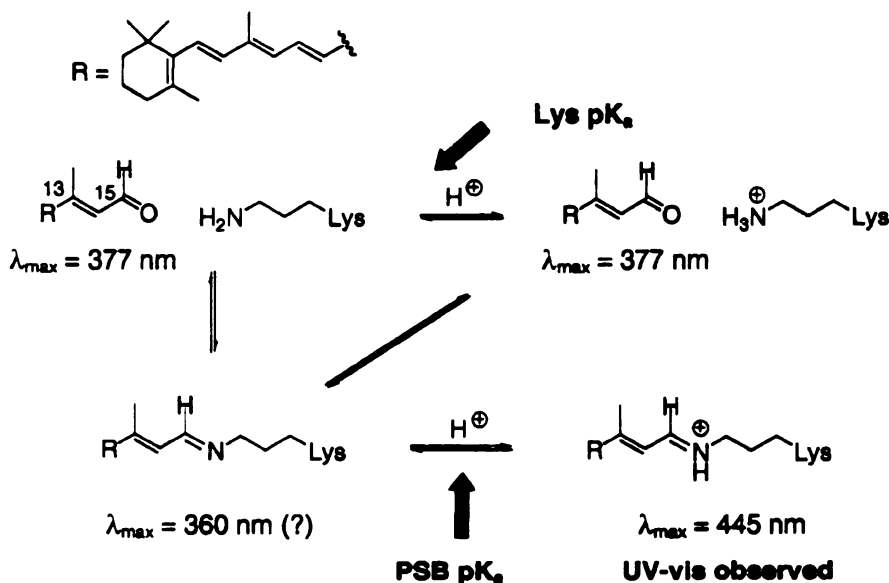
2006

Chapter 5

Attempts to label Lys132 to determine the pK_a of CRABPII mutants

5.1 Introduction

In the process of designing a rhodopsin mimic, it was apparent that the environment around the lysine would have a great influence in the nucleophilicity of the lysine, as well as the stability of the Schiff base. For example, was found that Tyr134 participates in the activation of the bound aldehyde followed by nucleophilic attack by Lys132. The effect of Tyr134 is evident when one compares the UV traces of the R132K::R111L::L121E triple mutant ($\lambda_{\max}=449$ nm) and that of the R132K::Y134F::R111L::L121E tetra mutant ($\lambda_{\max}=378$ nm)



Scheme 5-1. In the formation of a protonated Schiff base various equilibrium are involved.

and $\lambda_{\text{max}}=449$ nm, figure 5-

1). Also, replacing Arg111 with a hydrophobic residue was a critical mutation for PSB formation, we believe that the removal of the guanidine moiety removes a water molecule that forms a H-bond with the retinal molecule, thus it positions it for an optimal nucleophilic attack. A good example for the stabilization of the PSB is the presence of the counter

anion Glu121 (L121E), which stabilizes the protonation of the Schiff, base formed. The effect of the Glu121 is can be observed when one compares the UV traces of the R132K::R111L::L121E triple mutant ($\lambda_{\text{max}}=449$ nm) and that of the R132K::R111L::L121Q triple mutant ($\lambda_{\text{max}}=376$ nm). We have realized that it is important to tune the electronic environment to increase the nucleophilicity of Lys, while at the same time retain a polar enough cavity to support PSB formation. It was envisioned that determining the pK_a values in an accurate and quantitative manner for both the Lys residue of good and bad proteins and the protonated Schiff base could be instructive for further designs. The best mutant,

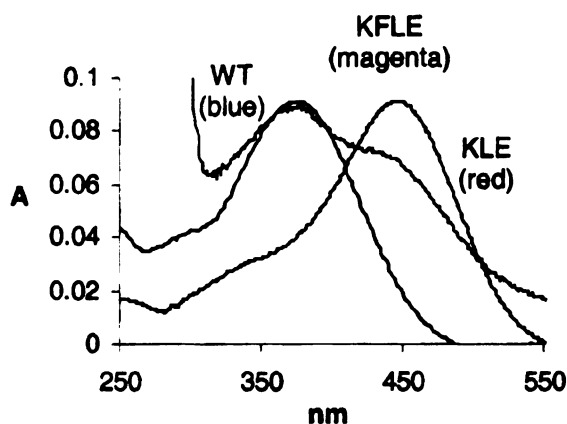


Figure 5-1. UV/Vis spectrum of retinal bound to WT CRABP II (blue trace) and the R132K::Y134F::R111L::L121E (magenta trace, KFLE), R132K::R111L::L121E (red trace, KLE) mutant. Retinal bound with wild-type absorbs at 377 nm. The UV of retinal with the KLE mutant reveals a 449 nm peak, indicating formation of PSB. The UV for the KFLE mutant is a mixture of two peaks at 378 nm and 438 nm.

triple mutant R132K::R111L::L121E exhibits stoichiometric binding in the formation of the protonated Schiff base at pH=7.5. Scheme 5-1 shows the pH dependent equilibrium between the non-bonded and the retinal-bound forms of Lys132. Calculation of the pK_a values for the PSB of various proteins using UV-vis spectroscopy has been possible. This was accomplished by a pH titration between the PSB ($\lambda_{max}=445$ nm) and the SB ($\lambda_{max}=377$ nm). It is important to mention that measurement of this pK_a is only possible when the equilibrium of the protein favors the PSB. As shown in Scheme 5-1 UV-vis spectroscopy does not differentiate between the unbound retinal and the SB between retinal and the rhodopsin mimic. Because of this, the measurement of the pK_a of Lys132 using UV-vis assays cannot be measured directly. The apparent pK_a value for Lys132 could be estimated by first incubating the protein at different pHs, followed by the addition of retinal. Second, addition of acid to push the equilibrium towards the PSB. Then followed by plotting the amount of red shifting present at 441 nm (A_{441}) versus the different pH values. Following this method, the calculated value for the R132K::R111L::L121E triple mutant was ~5.4 (Figure 5-2).² In a similar manner, the value calculated for the R132K::R111L::L121E::T54V mutant was even lower (~4.9).² These values compare well with the reported pK_a for the active site Lys residue for rhodopsin (~6.0).² Because we are monitoring an overall change of the UV-vis absorbance that it is indirectly connected to the calculated pK_a , therefore the validity of these data has been questioned. Importantly we cannot be confident that the change we observe after a change in pH is due to the pK_a of the nucleophilic Lysine and not due to other amino acid

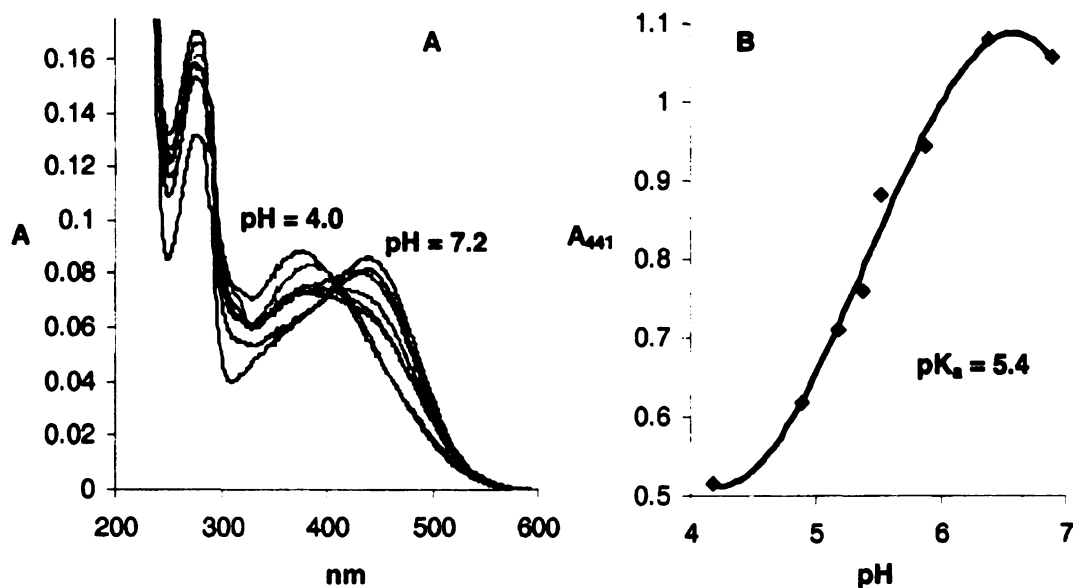


Figure 5-2. A. UV titration of R132K::R111L::L121E incubated at different pHs (4-7.5) upon addition of all-*trans*-retinal. B. Graph of the absorption change at 441 nm (A_{441}) with different pHs.

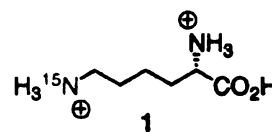
residues that we have identified as essential for PSB formation. For example a protonation of Glu121 would probably result in the loss of the UV red shifting (PSB) but would necessarily prevent the SB formation. Because of the fact that we cannot decisively distinguish between the UV-vis absorbance of free retinal and that of a SB, a different method should be used to determine the pK_a of the Lys132. A direct measurement of the protonation state of the engineered Lys could be performed using NMR spectroscopy.

NMR spectroscopy is one of the most common methods for the study of biomolecules in solution. CRABPII has been extensively studied by NMR, and the solution structures of both apo and holo protein have been solved. NMR

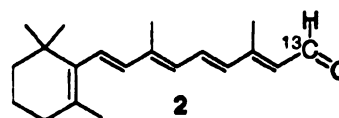
studies have been performed to determine the pK_a of different essential amino acids in proteins.⁴ NMR could provide direct information about the electronic environment of the retinal and the protein-imine formed. Measurement of a lysine residue has been previously described for the case of 4-oxalocrotonate tautomerase.^{4,5}

Use of NMR for measurement of the pK_a would entail labeling of the Lys of interest with ¹⁵N, a change in the pH of the solution will lead to a change of the ¹⁵N-NMR chemical shift of the specifically labeled lysine from unprotonated to protonated state, affording directly the pK_a.⁶ Therefore, we could use the same methodology to estimate an accurate pK_a value for the free Lys and the PSB nitrogen in the rhodopsin surrogate. Also, it has

been shown that a free aldehyde, a Schiff base and a protonated Schiff base have very different ¹H-NMR,



¹³C-NMR and ¹⁵N-NMR signals.⁷ The pK_a for the protonated Schiff base in the rhodopsin family



proteins has been estimated by using labeled ¹³C-retinal and ¹⁵N-lysine.⁸ The ¹³C-NMR chemical shifts for several iminium salts have been previously reported, with the imine (~160 ppm) and the iminium (~180 ppm), and free aldehyde (~200 ppm) are significantly different chemical shifts.⁹⁻¹¹ Along the same lines the use of ¹⁵N NMR for determining the pK_a values has been previously reported.⁶ In the study of bacteriorhodopsin, addition of a ¹³C-retinal has been used to perform ¹³C-NMR studies. The results revealed a shifting of the retinal C-15 (the aldehydic carbon) of ~30 ppm upon addition to bacteriorhodopsin (the shift is due

to the formation of protonated imine). The C-13 was also altered by a downfield ~11 ppm shift.^{7,12-15} In addition, solid state NMR experiments showed that the aldehyde proton in ¹H-NMR shifts from 9.7 ppm (free aldehyde) to 8.9 ppm (protonated Schiff base).^{16,17} As mentioned previously, if the side chain of Lys132 is labeled with ¹⁵N or ¹³C, a change in the protonation state of N would directly translate into a change of the ¹⁵N-NMR/¹³C-NMR chemical shifts. Thus by following the signal at different pHs a value for the pK_a of Lys could be calculated. Thus, we proposed to perform experiments that would involve the use of ¹³C-labeled retinal (1) and/or ¹³C and ¹⁵N labeled lysine (2) (for synthesis of ¹³C-retinal refer to Rachael Crist's Thesis). A change in the pH of the solution translates into a change in the protonation state of the imine nitrogen. The change in the imine protonation would cause a change in the NMR signal (from the unprotonated to the protonated state). Therefore it would be possible to monitor the shift in the NMR signal to calculate the pK_a. Since the ¹⁵N chemical shift changes will be used to follow the protonation state of the amine, the ε-amino group of Lys132 must be specifically labeled. Thus two different approaches were designed to accomplish this goal. The first one involves the non specific labeling of Lys132 with ε-¹³C-lysine. And the second involves the specific labeling of Lys132 with ε-¹³C-lysine via Native Chemical Ligation.

5.2 Non selective labeling of CRABPII

5.2.1 Use of ^{15}N -NMR to determine the pK_a of Lys

The first attempt to determine pK_a values using NMR was by preparation of a fully labeled $\epsilon\text{-}^{15}\text{N}$ -Lys protein. This proteins was prepared by addition of ^{15}N - ϵ -Lys (commercially available and the price is accessible) to the culture, and a uniformly ^{15}N labeled protein was obtained. MALDI-TOF and electrospray ionization spectra of the labeled double mutant verified the incorporation of the labeled Lys (For details refer to Chrysoula Vasileiou's Thesis).² Unfortunately, when the ^{15}N -NMR experiments were performed, the ^{15}N signal could not be identified. The signal for ^{15}N -lysine could not be identified, even when up to 20 mg of protein were used ($> 1\text{ mM}$). The major problem when ^{15}N NMR is used is that the natural sensitivity for ^{15}N is 0.001104 as compared to ^1H (1.0) or ^{13}C (0.0159). The natural sensitivity of each nuclei is determined by the gyromagnetic ratio and the energy difference ($\Delta E = \gamma \hbar B_0$) between the spin states. The larger the energy differences the more nuclei are present in the lower spin state and hence are available to absorb energy.¹⁸ When ^1H - ^{15}N HMQC spectrums was performed, still no signal was identified. Besides the fact that the natural sensitivity for N is too low, the failure of the ^1H - ^{15}N HMQC experiment could be due to the fact that the amine protons are exchangeable protons and were exchanging with the D_2O present in the sample. Therefore, to solve this problem, instead of using ^{15}N -NMR we decided to use ^{13}C -NMR. The first step was to accomplish the synthesis of $[6\text{-}^{13}\text{C}]$ -L-lysine, express the protein in the presence

of the labeled lysine and finally use the ^{13}C -Lys labeled protein for NMR experiments for the determination of the Lys pK_a value.

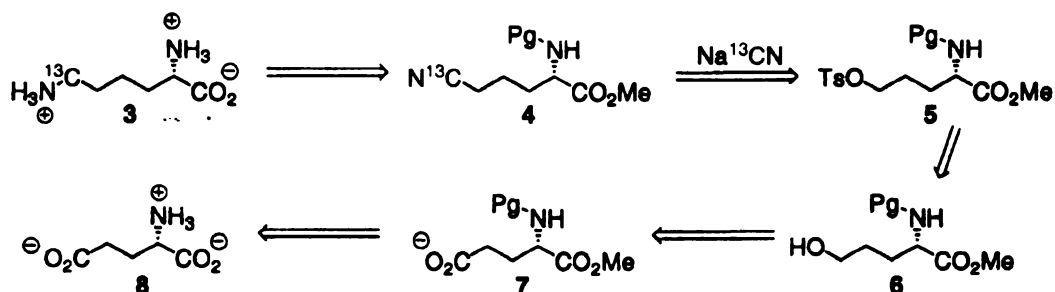
5.2.2 Use of ^{13}C -NMR

Because of the fact that the performed ^{15}N -NMR experiments resulted in the lack of signal, we decided to perform a ^1H - ^{13}C HMQC instead using ^{13}C -C15 labeled retinal, (for its preparation refer to Rachel Crist Thesis). The first attempt to take the NMR of this protein was using a protein concentration of about 1.5 mM. No ^{13}C signal could be identified in the ^{13}C -NMR spectrum of this sample, suggesting that the labeled retinal had precipitated. When the ^1H - ^{13}C HMQC spectrum was recorded no signal was identified for either the aldehydic proton or the PSB proton. The experiment was repeated with the final concentration of the protein at 0.5 mM. This time, when the ^1H - ^{13}C HMQC spectrum was recorded a signal at about 8.8 ppm, which corresponds to the PSB was identified. (24 hours in the 600 NMR instrument). The lack of sensitivity is mainly attributed to the fact that the retinal concentration was about 0.25 mM and protein instability. However, as mentioned previously, when the concentration is increased, the retinal precipitates. The next step would be the titration of the NMR signal at different pHs. When HCl was added into the solution, the protein-retinal mixture became unstable, and no signal could be identified in the ^1H - ^{13}C HMQC spectrum even after 24 of hours acquisition. After 24 hours the NMR tube was colorless and cloudy. It seems that the problem was again that when the pH was changed in the presence of the protein complex (retinal-CRABPII mutant PSB) it precipitated

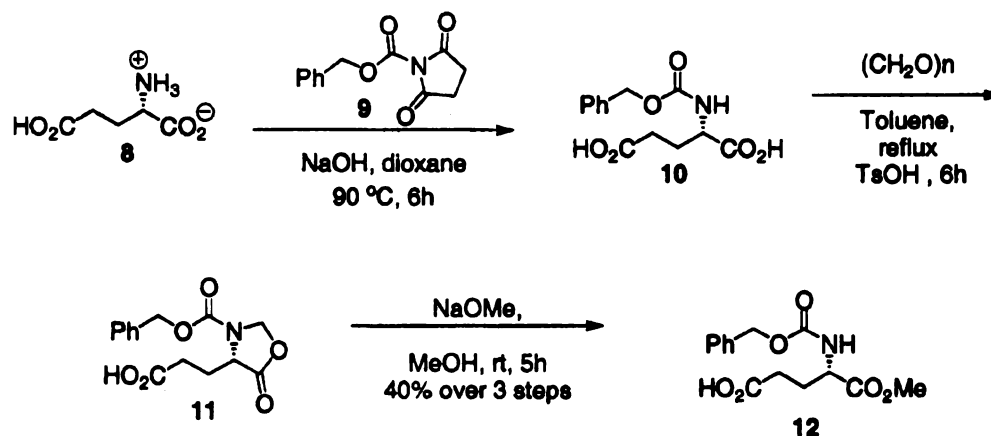
out of solution. We envision that this problem could be avoided if the protein itself is labeled with ^{13}C , because, to calculate the pK_a of the Lys addition of retinal would not be necessary. Therefore, the protein could be grown in media supplemented with $[6\text{-}^{13}\text{C}]\text{-L-lysine}$ instead of ^{15}N . Therefore, the synthesis of $[6\text{-}^{13}\text{C}]\text{-L-Lys}$ was performed.

5.2.3 Synthesis of $6\text{-}^{13}\text{C}$ -lysine

The preparation of $[6\text{-}^{13}\text{C}]\text{-L-lysine}$ would be accomplished by using $[^{13}\text{C}]\text{-sodium cyanide}$ as the source of isotopic label. The retro synthetic strategy is shown in Scheme 5-2, $[6\text{-}^{13}\text{C}]\text{-L-Lysine}$ (**3**) could be obtained from the reduction of nitrile **4**. Nitrile **4** would be derived from the treatment of tosyl **5** with cyanide. The tosyl **5** could be obtained from a chemoselective reduction of **7** and tosylation. The α -methyl ester **7** could be obtained from L-glutamic acid (**8**) after a series of transformations (Scheme 5-2).



Scheme 5-2. Retrosynthetic analysis for the preparation of $[6\text{-}^{13}\text{C}]\text{-L-Lysine}$.



Scheme 5-3. Synthesis of N-CBZ-Glutamic acid α -methyl ester **12**.

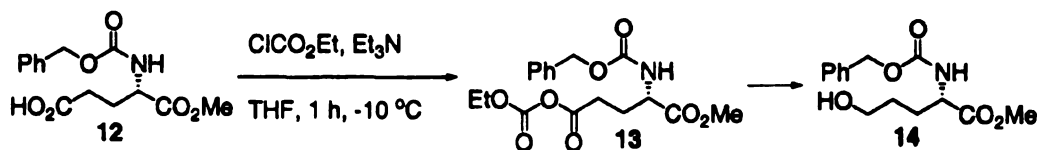
The synthesis of lysine started by treatment of L-glutamic acid (**8**) with N-benzyloxycarbonylsuccinimide (**9**) to afford the N-CBZ protected amino acid (**10**). A selective protection of the α -carboxylic acid was accomplished following Scholtz and Barlett's work.¹⁹⁻²¹ The protected glutamic acid was used without further purification and was treated with paraformaldehyde to form N-CBZ-5-oxo-4-oxazolidinepropanoic acid (**11**) followed by a selective ring opening after treatment with sodium methoxide in methanol, to yield the N-benzyloxycarbonyl-L-glutamic acid α -methyl ester (**12**) (Scheme 5-3). Attempts to purify **12** using column chromatography were unsuccessful. Thus, **12** was purified by recrystallization from dicyclohexylammonium as described in the literature.²²

Compound **12** was obtained in a 40% overall yield from glutamic acid (**8**). The overall yield for **12** was low due to the last treatment of the oxazolidine with methoxide. Even when the reaction was performed using absolute methanol and molecular sieves, no real change in the overall yield was obtained. All these reactions can be scaled up to about 20 grams of starting material without problem or any change in the yields.

The chemoselective reduction of the carboxylic acid of **12** to the alcohol was next. The most common strategy to reduce carboxylic acids in presence of esters is via the preparation of the mixed anhydride **13** follow by reduction using

Table 5-1. Attempts to selectively reduce N-CBZ-Glutamic acid α -methyl ester.

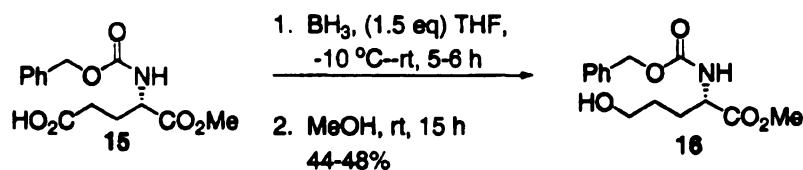
^aDepends on scale.



Entry	Reaction conditions	Yield of 14
1	20 eq. NaBH ₄ , H ₂ O:THF (10:2) 0°C → RT, 2hrs	12-60% ^a
2	3 eq. NaBH ₄ , H ₂ O:THF (10:2) 0°C → RT, 2hrs	16%
3	3 eq. NaBH ₄ /aluminaTHF, RT, 4hrs	80% (12)
4	3 eq. NaBH ₄ :THF 0°C → RT, 2hrs	68% (12)
5	3 eq. NaBH ₄ , H ₂ O:THF (4:1) 0°C → RT, 2hrs	60%
6	3 eq. NaBH ₄ , in THF, MeOH added drop wise 0°C → RT, 1hrs	25-89% ^a
7	3 eq. NaBH ₄ , in H ₂ O: 0°C → RT, 1hrs	25%

NaBH_4 (Table 5-1).²³⁻²⁷ The first attempt was using 20 equivalent of the reducing agent and addition of water (see details in experimental). The reaction works well in small scale (30-100 mg) giving about 60% yield of the desired product **14**. However in a larger scale the yield decreased drastically to 16%. When a large-scale reaction was performed only the alcohol could be isolated after column. When NaBH_4 supported in alumina was used no reduction took place, only starting material was recovered. Reduction using only NaBH_4 provided the same result. By increasing the amount of water the reaction seemed to proceed better, although no transformation takes place when water is used as the sole solvent. Conditions 4, 5 and 7 (Table 5-1) were only tried in a small scale (≤ 100 mg). Since entry 6 was promising in small scale, (89% yield), the reaction was repeated on a larger scale but again the yield decreased significantly (25%). All the previous reactions were repeated using recrystallized NaBH_4 , but no improvement in the final yield was observed.

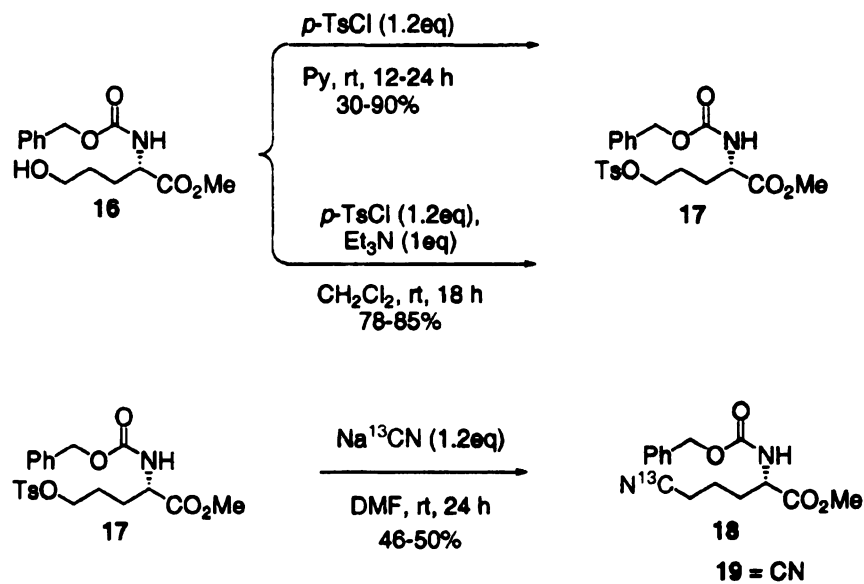
The chemoselective reduction of carboxylic acids using BH_3 has also been reported. When BH_3 (1 M in THF) was used to reduce **4**, the desired alcohol was obtained. The yield for the desired alcohol varied from 44 to 48% (Scheme 5-4). and was reproducible regardless of the scale performed.



Scheme 5-4. Selective reduction of N-CBZ-Glutamic acid α methyl ester.

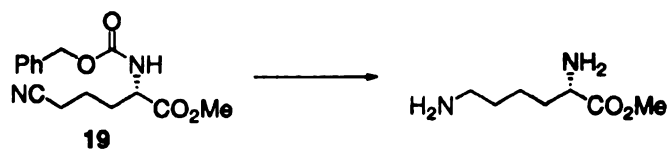
Tosylation of the alcohol **16** was accomplished by treatment of the alcohol (**16**) with *p*-toluene sulfonyl chloride in triethylamine. (The use of pyridine as a base did not provide consistent results, 30–90%, Scheme 5-5). The next step would be the introduction of the ^{13}C using Na^{13}CN . The transformation of the tosyl (**17**) to the nitrile (**18**) was uneventful yielding nitrile (**18**) in 46–50% yield (when no label NaCN was used the same yield was obtained).

The next step toward the synthesis of ^{13}C -labeled Lys would be the hydrogenation of the nitrile to an amine. Hydrogenation of nitriles has been previously reported and a variety of different catalyst have been used, including



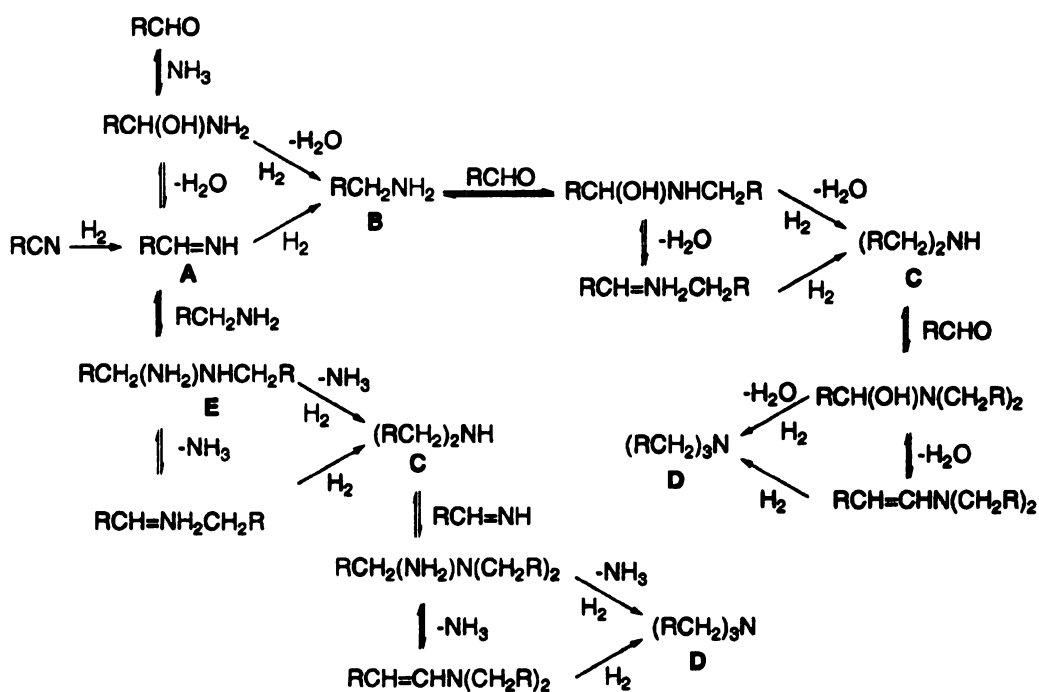
Scheme 5-5. Tosylation of alcohol using triethyl amine provides more consistent results. Introduction of the label ^{13}C occurred in a 50% yield using Na^{13}CN .

Table 5-2. Attempts to hydrogenate nitrile **19**.



Entry	Conditions	Product
1	PtO_2 , isopropanol HCl (50:1) 20 psi, RT, 3 hrs	SM
2	PtO_2 , isopropanol HCl (50:1) 50 psi, RT, 3 hrs	SM
3	$\text{Pd}(\text{OH})_2$ 20% wt isopropanol HCl (50:1) 100 psi, RT, 3 hrs	SM
4	$\text{Pd}(\text{OH})_2$ 20% wt isopropanol HCl (50:1) 300 psi, RT, 3 hrs	SM
5	$\text{Pd}(\text{OH})_2$ 20% wt isopropanol HCl (50:1) 500 psi, RT, overnight	Complex mixture

Pd/C, Pt/C, Ru/C, Rh/C Raney Ni, Raney Co, Adams catalyst (PtO_2), and $\text{Pt}(\text{OH})_2$.²⁸⁻³⁵ The reaction have been performed at different pressures (1-500 psi). The first attempt to reduce the nitrile was performed using PtO_2 (Table 5-2) at 20 psi and only starting material was recovered. The pressure was increased to 50 psi without any change in the course of the reaction. When $\text{Pt}(\text{OH})_2$ was used, again, only starting material was observed, and increasing the pressure to 300 psi did not improve the results. When both the pressure and the reaction time were increased, decomposition was observed.



Scheme 5-6. Proposed mechanism for the hydrogenation of nitriles.¹

It has been reported that the reduction of nitriles can afford primary (B), secondary (C) and/or tertiary (D) amines via an intermediate imine (A). In the proposed mechanism for reduction of nitriles, first, nitriles are hydrogenated to the

imine (A) and then to the primary amine (B). The primary amine could react with the intermediate imine and form the secondary amine via a *gem*-diamine intermediate (E) (Scheme 5-6).¹ This explains the formation of different products. It has been suggested that to avoid the formation of side products addition of acid is required. However, other researchers have observed the contrary; an increase in the rate of the hydrolysis instead the reduction in the presence of acid. At this point the 6-nitrile-1-hexyne was used as a model until the reductions were optimized. After quenching the reaction, the same decomposition was observed. Different catalysts were tried (PtO₂), Pt(OH)₂ and Pt/C, but the same result was observed.

Reductions of the nitrile under atmospheric pressure are not very common, a few examples can be found in the literature.^{33,36} First the search for a good reduction protocol to reduce the nitrile to primary amine was performed using the model compound (6-nitrile-1-hexyne). The first attempt was using Raney nickel and hydrazinium monoformate³⁶ as the hydrogen donor at room temperature resulting in 30–80% of recovered starting material, while the rest was a combination of many different spots in the TLC (the discrepancy in the yield of recovered starting material was due to the fact that different Raney Nickel bottles gave different results, Table 5-3). This method of reduction for nitriles has been previously used only for aromatic nitriles. Thus, it is possible that aliphatic aromatics are not activated enough to be reduced by this method. Very few protocols that reduce aliphatic nitriles at atmospheric pressure are reported in the literature. One of these methods utilizes a palladium activated, Raney-Nickel in

presence of H₂.³³ The first time this reaction was performed using 3 eq. of Raney-Ni, the crude ¹H-NMR showed the desired amine product (the reduced nitrile **21**), but also the reduced alkane **22** was obtained (Table 5-3). To allow the full reduction of the nitrile and alkyne the reaction was repeated using 6 equivalents and only the *n*-aminohexane was obtained in about 76% yield (crude).

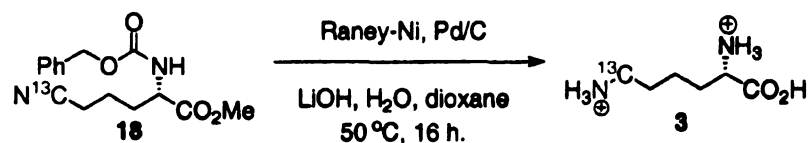
Table 5-3. Reduction of nitrile using Raney-Ni. ^aThe yield was calculated using NMR and the final total weight, no attempt to purify **22** was performed.



Entry	Reduction conditions	Results
1	Raney Ni /N ₂ H ₂ *HCOOH / MeOH, RT 4 hours	20 , 80%
2	Raney Ni /N ₂ H ₂ *HCOOH / MeOH, RT 8 hours	20 , 40%
3	Raney Ni /N ₂ H ₂ *HCOOH / MeOH, RT overnight	20 , 30%
4	Raney Ni (3 eq) / Pd/C, H ₂ 1,4 dioxane, H ₂ O, LiOH	21 20% ^a , 22 20% ^a
5	Raney Ni (6 eq) / Pd/C, H ₂ 1,4 dioxane, H ₂ O, LiOH	22 about 76% ^a

Since the above conditions were successful for the reduction of the model compound, the actual system (**18**) was treated under the same conditions. The results showed that lysine was obtained, which indicated that the CBZ group came off (as expected), but also the methyl ester was hydrolyzed. The purity of the lysine obtained was not good and at least three different spots were identified. The most common procedure to purify lysine is the use of Cation Exchange Resins. The use of a variety of resins have been reported, such as IR 120,

DOWEX N406, DOWEX MARATHON, DOWEX HCRS DOWEX N606, Dowex 50Wx8-400 among others.³⁷⁻³⁹ Due to the accessibility to a DOWEX-50Wx8 resin it was chosen to be used for the purification of lysine. Unfortunately, when the purification was performed the eluted lysine was not pure. A detailed analysis of the mechanism of nitrile reduction (Scheme 5-7) suggests that a variety of side products can be formed.¹ In particular, if a secondary amine were formed the separation of those two components using an ion exchange column would not be possible. Also it is important to keep in mind that if the side product results from the hydrolysis, separation using ion exchange resin would not be possible.



Scheme 5-7. Reduction of nitrile using Raney-Ni, Pd/C.

Finally, the method used to purify the [6-¹³C]-L-lysine was a modification of the procedure reported by Schuster in 1980.^{40,41} The lysine was dissolved in a 2 M HCl solution and was injected into an HPLC. The HPLC was fitted with a μ Bondapak NH₂ column and the amino acid were detected at 200 nm. The mobile phase used was 30% 0.01 M KH₂PO₄/70%CH₃CN/H₂O (100/16). The [6-¹³C]-L-Lysine was purified in a 85% yield from nitrile 18.

Expression of the CRABP_{II} mutants with the label Lys has not been performed. However we are confident that by following the protocol used to

express the uniformly label ^{15}N -Lys, the desired ^{13}C -Lys labeled protein would be obtained, and the ^1H - ^{13}C HMQC can be performed.

5.3 Specific labeling of Lysine

5.3.1 Chemical Ligation

Chemical synthesis has always been considered an attractive alternative to biological methods for protein production. The use of chemical synthesis provides the possibility of creating unlimited variation of polypeptide chains. Total chemical synthesis of native proteins has made a number of important contributions to biomedical research.⁴²⁻⁴⁵ In polypeptides and proteins the size and the need for correct folding of the peptides synthesized increases the complexity of the synthesis. The challenge for a chemist was always to be able to chemically synthesize active proteins by stitching various synthetic peptides together⁴⁶⁻⁴⁸ and chemical ligation seems to provide that tool. However, the original ligation chemistries gave a nonpeptide bond at the site of ligation. The first ligation chemistry consisted of a nonpeptide bond (oxime, thiazoline) at the site of ligation, and although these unnatural structures were often well tolerated, trials for the formation of a native peptide bond were persisted.⁴⁹⁻⁵¹ In the native chemical ligation method, the ligation occurs at a unique N-terminal Cys residue, independent of any additional Cys residues. The side chain of the N-terminus Cys forms a thioester-linked product that undergoes spontaneous rearrangement, via

an irreversible intramolecular nucleophilic attack of the Cys amine nitrogen, to provide the amide-linked product.

The first case of Native Chemical Ligation was reported by Muir and Dawson in 1994⁵² and since then a series of chemically and enzymatically active proteins have been made following different variations of the same central idea. According to the established protocols, proteins can be made in high yield and good purity from unprotected peptide building blocks, allowing for the incorporation of unnatural segments, essential for the study of protein structure and function.⁵³

5.3.2 Native chemical ligation

The essential feature of the Native Chemical Ligation is the formation of a thioester-linked product that undergoes spontaneous rearrangement, via intramolecular nucleophilic attack, to give the desired amide-linked product. This method requires a N-terminal Cys that is responsible of the formation of the thioester, which is followed by rearrangement to form the peptide bond. Since Lys132 that was to be labeled is only two amino acids away from Cys130, we strategize to use this method to selectively label Lys132. Recently, this approach has been extended by the use of self splicing proteins as a mediator for the formation of a native amide bond for the synthesis of proteins. Protein splicing is a post-translational process that includes the excision of an internal protein segment, called intein from the primary translational product with associated ligation of the flanking sequences, the exteins.⁵⁴⁻⁵⁶ Intein-mediated protein

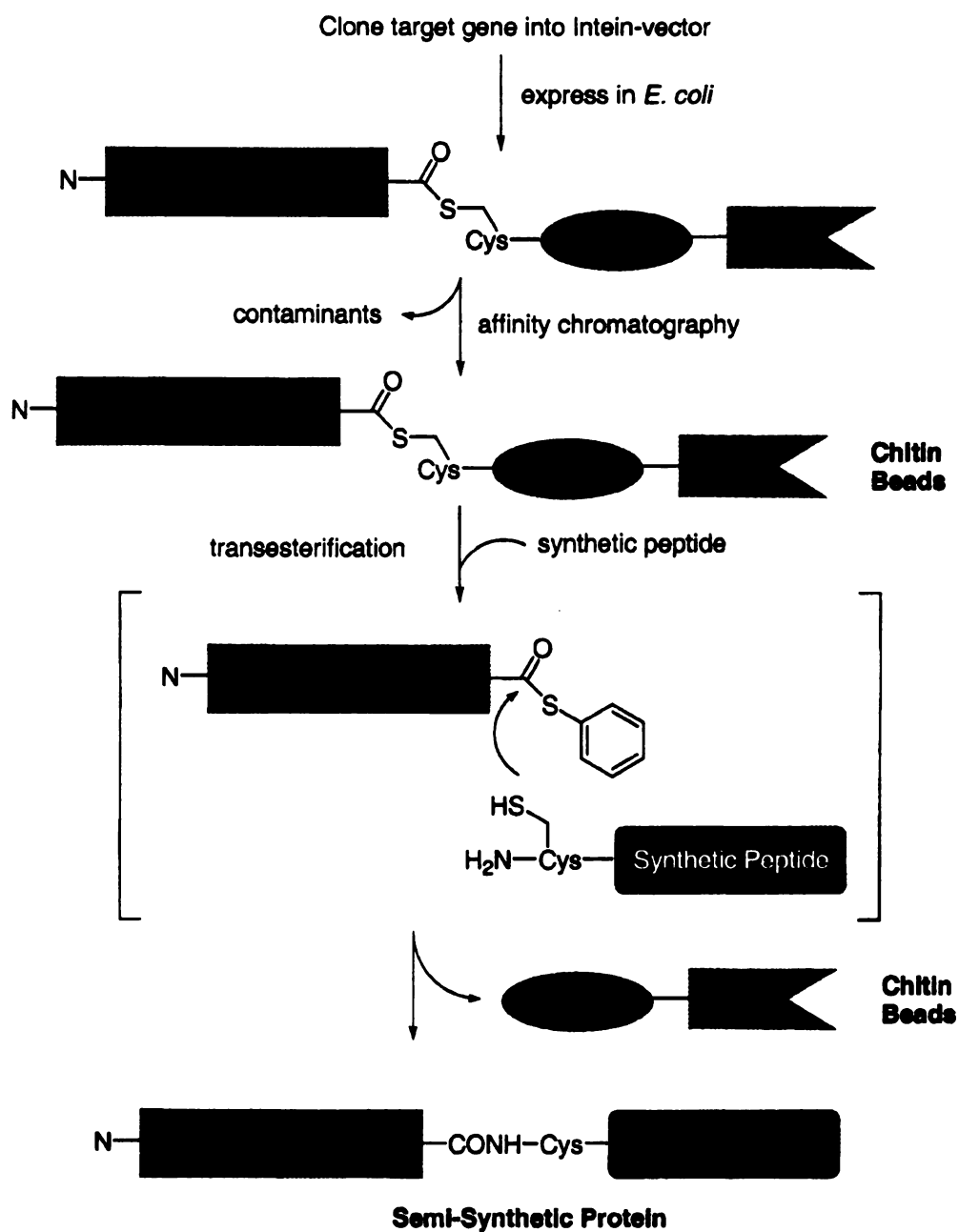


Figure 5-3. Schematic representation of the Intein-mediated native chemical ligation process.

splicing has been proposed to occur *via* intein-mediated formation of thioester-linked intermediates. This intermediate is cleaved by nucleophilic attack to form the final amide-linked spliced polypeptide. Utilizing similar features, intein-mediated splicing has been used for applying the native chemical ligation methodology (Figure 5-3).⁵⁷

The intein mediated chemical ligation, requires the gene of a target peptide to be cloned in an intein-containing vector and expressed as a recombinant protein. To facilitate its purification the protein contains an affinity tag, (a chitin binding domain (CBD)). While the recombinant protein is attached to the column, a second, synthetic peptide, that contains a Cys residue at the N-terminus, is added and a *trans*-thioesterification occurs. This process is facilitated by the intein conformation, that eventually results in elution of the semi-synthetic protein, while the intein part remains bound on the affinity column (Figure 5-4). Thus, this method allows for the incorporation of labeled amino acid or any unnatural amino acid in the part of the protein originating from the synthetic peptide.

As mentioned previously, CRABP II contains a Cys residue at position 130 that would allow for N-terminus ligation because the labeled Lys132 can be accommodated in a chemically synthesized peptide. Thus, expression of the fusion protein of intein-CRABP II (1-129) can be combined with a synthetic peptide that has the remaining 16 amino acids (His-tag protein) to complete the CRABP II sequence. After intein splicing, the method should afford the whole CRABP II protein (Figure 5-4).

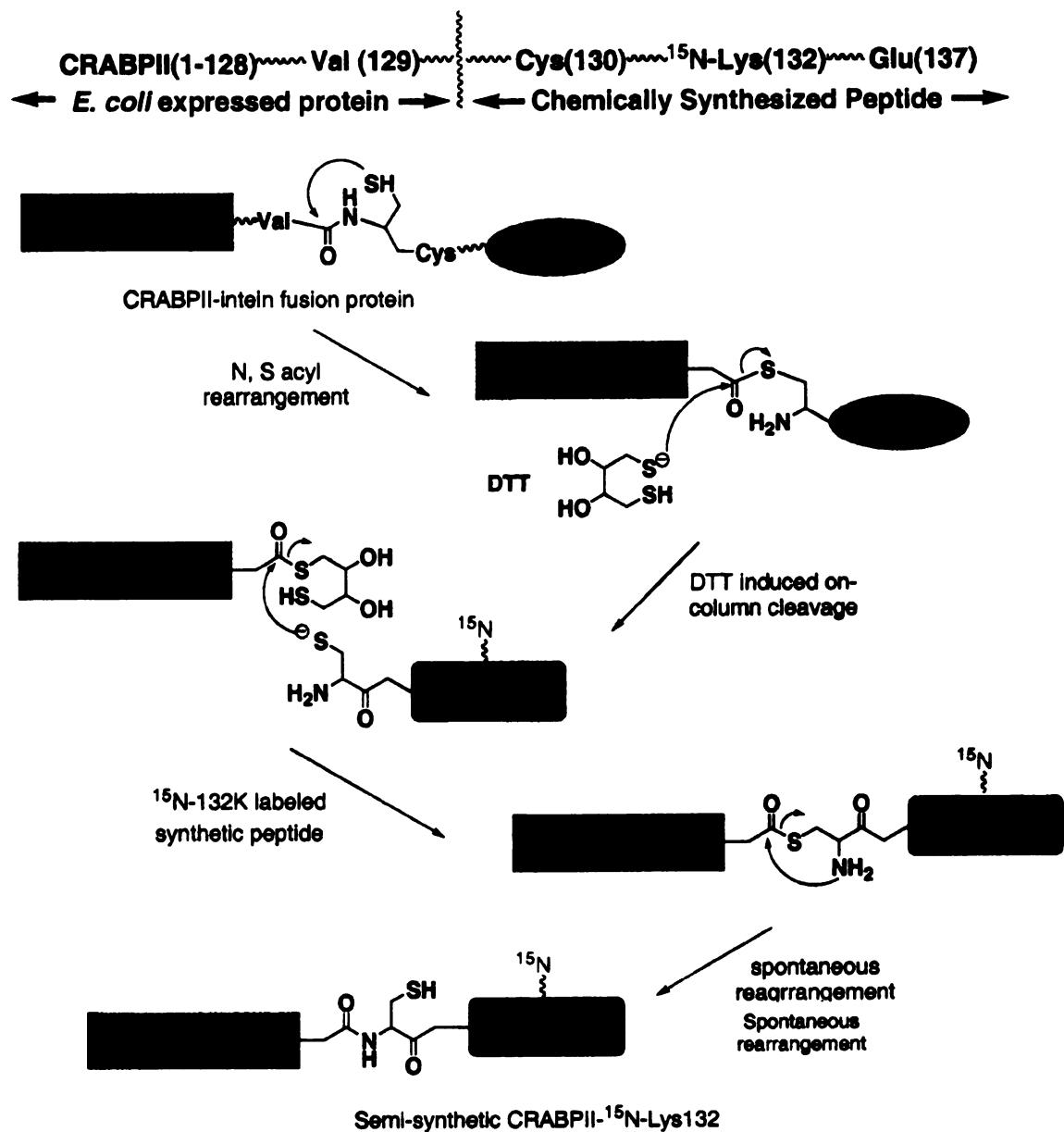


Figure 5-4. Formation of ¹⁵N-Lys132 CRABPII via Intein Mediated Chemical Ligation.

The system currently used is called IMPACTTM-CN System (Intein Mediated Purification with an Affinity Chitin-binding Tag, New England Biolabs) and utilizes the inducible self-cleavage activity of inteins to separate the target protein from the affinity tag. IMPACT systems have been used in the literature with good results for proteins up to 160 KDa.⁵⁷⁻⁷⁰ The vectors used contain an engineered intein and are designed for fusing the C-terminus of a target protein to the intein-Chitin Binding Domain tag. The Chitin binding domain is used to generate C-terminal thio-ester tagged proteins, which would be used in further protein ligations. Therefore, the desired protein will be expressed as an intein fused protein. The overexpressed fusion protein would be loaded on a chitin column and could separate from the rest of the bacterial proteins. After the recombinant protein is bonded to the column, treatment with a thio-containing molecule (thiophenol, or 2-mercaptoethanesulfonic acid (MESNA) N-methylmercaptoacetamide (NMA)) induces protein splicing. Finally, the synthetic peptide is added and the newly formed protein elutes from the column. As discussed in Chrysoula Vasileiou's Thesis, many mutations of CRABP^{II} were carried out to optimize the retinal binding as a protonated Schiff base (PSB). To date, the R132K::R111L::L121E triple mutant is the best protein obtained. This protein contains two mutations within the first 130 amino acids that are proven essential for the PSB formation. Therefore, one would expect that the gene of CRABP^{II}(1-129, R111L:L121E) would be the obvious choice for the formation of the recombinant protein. However, when our efforts to utilize the native chemical ligation method began, double mutant R132K::Y134F was the most promising

protein in hand. Because of this, the native CRABP^{II}(1-129) gene was used for the formation of the recombinant protein. We thought that if the process was successful, it could then be repeated for any desired mutant, without major modifications to the protocol.

5.3.3 Expression of CRABP^{II} fused to a mini intein

For the cloning of the CRABP^{II}(1-129) gene into the pTXB3 (a mini-intein (198 residues) derived from the *gyrA* gene of *Mycobacterium xenopi*), refer to Chrysoula Vasileiou's Thesis. CRABP^{II} (1-129) and CRABP^{II} (1-V129A) were cloned into the pTXB3 vector and were expressed as a fusion protein with a mini-intein (*Mxe* GyrA). The band expected for the CRABP^{II}(1-129)-Intein fusion protein should show at ~43 KDa (327 amino acids total) and as expected

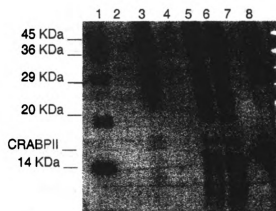


Figure 5-5. Overexpression of the fusion CRABP^{II}(1-V129A)*Mxe* GyrA in (BL21(DE3)pLysS, host 1mM IPTG, 30 °C). 1. Marker. 2.empty lane, 3. Crude. 4. Supernatant. 5.,6 Unsoluble fraction. 7. empty lane, 8. CRABP^{II}. The expression using different IPTG concentrations and different *E. coli* host gives similar results.

Table 5-4. Use of different *E. coli* hosts to test solubility of the fusion intein protein. The results are shown as: S=Supernatant, I.B.= Inclusion bodies. ✓= CRABPII fusion present, x= No CRABPII fusion present. x=No cleavage of the intein was observed. CRABPII corresponds to both CRABPII(1-129) Mxe GyrA and CRABPII(1-A129V) Mxe GyrA, both provide same results.

<i>E. coli</i> host	IPTG (mM)	16 °C		30 °C		Intein Activity
		S.	I.B.	S.	I.B.	
ER2566	0.0	x	x	x	x	x
	0.01	x	✓	x	✓	x
	0.10	x	✓	x	✓	x
	1.00	x	✓	x	✓	x
Origami	0.0	x	x	x	x	x
	0.01	x	x	x	x	x
	0.10	x	x	x	x	x
	1.00	x	x	x	x	x
Tuner(DE3)pLacI	0.0	x	x	x	x	x
	0.01	x	✓	x	✓	x
	0.10	x	✓	x	✓	x
	1.00	x	✓	x	✓	x
BL21(DE3)pLysS	0.0	x	x	x	x	x
	0.01	x	✓	x	✓	x
	0.10	x	✓	x	✓	x
	1.00	x	✓	x	✓	x

the desired protein band was overexpressed. Unfortunately, it was expressed in the form of inclusion bodies as summarized in Table 5-4, (Figure 5-5 shows an example for the expression of the fusion intein).

5.3.4 Attempts to solubilize fusion protein

When inteins are used as fusion proteins, the formation of inclusion bodies is a common problem. Literature reports to solve this kind of problem mostly

consist of a trial and error approach.⁶¹ Refolding of proteins is not commonly used due to the low yield in activity after refolding^{61,71} At this point the challenge was to express a fusion protein that showed high splicing activity. Because refolding leads to a low yield of active protein, an attempt to express the fusion protein in soluble form was performed.

The *E. coli* Origami cells have been used to improve the solubility of the fusion proteins. These host cells promote disulfide bond formation due to the fact that they express a mutated thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes.^{72,73} Protein solubility can increase due to the optimization of disulfide bond formation.^{72,74 75} The vector containing CRABPII(1-129)- *Mxe* GyrA, or CRABPII(1-V129A)- *Mxe* GyrA, were transformed into Origami cells and expression of the recombinant proteins was attempted. Unfortunately the results showed no recombinant protein. Even when different concentrations of IPTG were tried (1, 0.1 and 0.01 mM) and at different temperatures (16 °C and 32 °C) no overexpression was observed (Table 5-4). As discussed in Chrysoula Vasileiou's thesis the Tuner cells provided only inclusion bodies. Lack of overexpression has been previously reported by Walker and colleagues.⁷⁶ They suggest that the expression of mutated *trxB* and *gor* exhaust the *E. coli* metabolism, thus affecting the overexpression of the fusion protein.

5.3.5 Refolding of *Mxe* GyrA intein system

As mentioned previously the formation of inclusion bodies when using inteins is not surprising. In the literature, however not many refolding attempts

have been reported, and the reported data shows low activity after refolding (up to 50% activity is recovered).^{48,64} Formation of inclusion bodies could offer several advantages for the production of recombinant proteins if they can be refolded. The proteins produced in inclusion bodies might be unstable in the cytoplasm of *E. coli* due to proteolysis and may be toxic for the host cell in the native conformation. It has been reported that under appropriate conditions, the recombinant proteins deposited in inclusion bodies amounts to about 50% or more of the total cellular protein. And because inclusion bodies have a relatively high density⁴⁸ they can be isolated away from the cellular proteins, with a purity up to 90% under optimal condition.⁴⁹ A number of methods have been reported for solubilizing and refolding proteins from inclusion bodies. Denaturants such as urea and guanidine hydrochloride, and the use of highly alkaline pHs, strong thiol-containing reductants (dithiothreitol, dithioerythritol mercaptoethanol), chelating reagents (EDTA), have been reported to solubilize a significant percentage of insoluble proteins. Renaturation is usually accomplished by the removal of excess denaturants by either dilution or a buffer-exchange step (dialysis, diafiltration, gel-filtration chromatography or immobilization onto a solid support). Therefore, different proteins may require different conditions for folding and searching for folding conditions can be a daunting task. In general, refolding is initiated by diluting the denatured protein into a refolding buffer. Once the denaturant concentration is reduced the protein refolds at rate equivalent to first order kinetics. The unfolded protein quickly folds to an intermediate form, then there is a slow transition from the intermediate to the native form (limiting step) which

may take minutes to days to complete. At this point nonproductive aggregation reactions can occur, resulting in the formation of unstable aggregates. Therefore, it is very important to avoid this pathway, which can generally be achieved by having very low concentrations of protein.

In the course of our experiments we came across a protocol used in Muir's group on how to deal with the inclusion bodies problem (For details see Chrysoula Vasileiou Thesis). The inclusion bodies were resuspended in the presence of 6 M urea and then dialyzed. This was followed by incubation with DTT and the results revealed the presence of a new protein band, at around 15 KDa, that can be attributed to the CRABPII(1-129) protein. At the same time, there was another new protein band, of almost

equal intensity, present on the gel, at ~29 KDa, which was believed to be the cleaved intein.

However, this band corresponded to only a small amount of the protein bound on the chitin beads. Furthermore when this intein was used to carry out the actual chemical ligation cleavage conditions, that is the use of MESNA (2-mercaptoethanesulfonic acid)

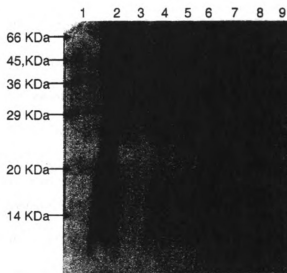


Figure 5-6. Protein Gel. 1. Marker. 2. Dialysis supernatant loaded on Chitin Beads. 3,4. Chitin beads after loading. 5,6. Chitin beads digested with 10 mM DTT. 7. 1st flow through. 8. 2nd flow through. 9. Wash.

instead of DTT to induce the protein splicing (the use of thiol reagents such as MESNA provides a reactive thioester at the C-terminus of the target protein for use in intein mediated Protein Ligation, while DTT does not), only traces of protein at ~15 KDa were present (Figure 5-6). The problem with dialysis is that the protein is exposed to a large amount of chaotrope, increasing the amount of a partially formed intermediate in the folding pathway. Accumulation of this intermediate shifts the equilibrium to the incorrect folding, thus producing inactive protein.

Therefore, a different approach to attempt to refold the fusion protein *MxeGyrA-CRABPII* (both recombinant protein were tested) was followed. This time a 100 fold dilution was performed. This technique has the advantage that the concentration of both chaotropes and protein is significantly decreased, and therefore, the equilibrium should shift to the correct folding of the protein. The disadvantage is that the final concentration of the protein is very small. After the protein was diluted 100 times and stirred for at least 24 hours, it was directly added into the chitin beads. When this small amount was treated with DTT or MESNA no cleavage of the intein protein was observed. Because it has been reported that inteins can cleave at high urea concentrations (up to 4 M), a 1/6 dilution was performed (decrease of the chaotropic from 6 M urea to 1 M urea). Right after the solution was diluted, it became cloudy suggesting that probably the protein was precipitating. Still the protein was stirred for 24 hours and the solution was loaded onto the chitin beads. After addition of MESNA no cleavage was observed. Because the inteins have been reported to be active in the presence

of 4 M urea the experiment was repeated at a 2/3 dilution, but again the cloudiness was observed and after activity was tested no cleavage was observed.

The Hampton Research FoldIt Screen is a commercially available kit that allows for determination, whether a protein of interest can be folded from inclusion bodies. FoldIt is a fractional protein-folding screen which evaluates 10 factors in 16 unique solutions: protein concentration, polar additive, detergent, pH, chaotrope, ionic strength, divalent cations, non-polar additive and polyethylene glycol. The use of FoldIt could give significant information about the feasibility of refolding the CRABPII-*Mxe* GyrA protein. After preparation of inclusion bodies, (about 70% pure, determined by running an acrylamide gel, stained with commasie blue, and comparing the CRABPII *Mxe* GyrA band intensity with the rest of the proteins), the refolding experiments were started with a final concentration of the proteins (all proteins) of 1 mg/mL. Then a 50-fold dilution was performed. The resulting solution was added to a chitin bead column and slowly eluted. Unfortunately after analysis of the chitin beads, no fused protein was observed. Therefore, a second trial was performed, but this time a 10 times dilution, was used. The solution again was to chitin beads, and the slowly eluted solution was resubmitted to the resin two more times. When DTT was added to promote cleavage a weak band at 48 KDa corresponding to uncleaved intein was observed.

The search for folding conditions can be an overwhelming task, proteins require different conditions for folding, and no prediction for the appropriate conditions for folding can be performed. The Hampton Research FoldIt Screen

allows determining with a reasonable degree of confidence whether the protein of interest can be refolded from inclusion bodies. FoldIt is a screening kit designed to be a rapid and statistically meaningful method for determining if a protein will fold *in vitro* and what factors are important for folding. If the results of FoldIt indicate that the folding from inclusion bodies is not feasible, then the results of FoldIt suggests that the search for an appropriate system to fold the protein properly, would be time consuming and expensive. The use of this kit consists in mixing the protein (solubilized in 6 M urea) with 16 different prepared solutions. Each sample was incubated at 4 °C for 10 hours, and the intein activity was determined. The results showed that none of the conditions resulted in an active protein suggesting that in the folding of this protein many different factors are involved and screening of the best conditions would not be a trivial screening. In the literature, there are a few examples for screening of soluble and active inteins, however, all of them involve engineering the protein by adding a fluorescent tag. In this case, addition of a fluorescent tag would only increase the complexity of our system.

5.3.6 Use different intein

In the intein literature it has been found that the final yield of the desired recombinant protein can depend strongly on which intein-fusion system is employed.¹⁴ There are two major contributions to low yield in the expression of fusion proteins, the low levels of soluble expression (as observed with *Mxe* GyrA) and *in vivo* cleavage. Up to now, no conclusive study on which intein works

better for which system has been performed, and therefore, the suggestion is that if a system is problematic a different intein-mediator should be tried. As previously discussed when the mini intein *Mxe* GyrA was used, inclusion bodies were formed. Other available intein is *Sce* VMA1. This system utilizes an intein (454 amino acid residues) from the *Sacharomyces cerevisiae* VMA1 gene. A chitin-binding domain is attached to the C terminal for easy purification. Because this is a larger intein we thought that the proper intein folding would not be affected with the CRABPII fusion. Therefore, the obtained fusion protein would be more soluble and more active. The pTYB3 vector, contains a larger intein (454 amino acid residues) from the *Saccharomyces cerevisiae* VMA1 gene (*Sce* VMA). Therefore, the CRABPII(1-V129A) gene was cloned into the pTYB3 vector. PCR afforded the CRABPII (1-V129A) gene with the cut sites *NcoI* and *SapI*. The gene and the vector were purified and further digested. Ligations were performed in a mole ratio of gene/vector of 10/1 and 15/1 with T4 Ligase at 16 °C for 24 hours. After transformation into JM109 cells the correct sequence was confirmed by sequencing. The gene was transformed and expressed in ER2566, Tuner cells and Origami yielding the corresponding fusion protein.

The initial results showed the over expressed fusion protein CRABPII(1-V129A)- *Sce* VMA band about 70 KDa (Figure 5-7) in the crude lysate. The soluble fraction accounted for about 60% of the protein and the inclusion bodies for about 40%. But when the soluble fraction was added to the chitin beads very small amount of protein was found to bind the beads (shown in Figure 5-8). As mentioned previously, three different strains were tested for the expression of the

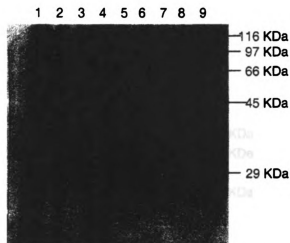


Figure 5-7. Fused protein CRABPII-(1-V129A) *Sce* VMA1 70KD expressed at 30 °C induced with 1 mM IPTG. 1. crude. 2. soluble fraction. 3. inclusion bodies. 4. Wash 1. 5. Wash 2. 6. Chitin beads. 7. Chitin beads. 8. empty lane 9. Marker.

recombinant intein protein, ER2566, Tuner cells and Origami. All the *E. coli* strains gave very similar results. When different IPTG concentrations and temperatures were used no improvement in the overexpression of the soluble protein was found. Therefore, only the results using ER2566 are discussed. The soluble fraction was added to the chitin beads, and the expected band at 70 KDa was observed, however, two strong bands close to 50 KDa were also obtained. The existence of the band at about 50 KDa suggests that the intein was cleaving *in vivo*.

The same results were obtained when the protein expression was performed at 16 °C overnight. When the beads were incubated with DTT to

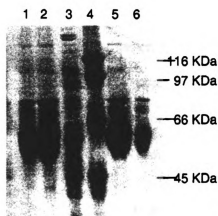


Figure 5-8. Binding of the overexpressed fusion protein to the chitin beads. A major band at 50 KDa is observed. Expression at 30 °C induced with 1 mM IPTG. Lanes 1,2,5,6. Shows the binding of the super natant to the chitin beads. 3. SceVMA1. 4. Marker.

induce intein cleavage the CRABPII(1-V129A)- Sce VMA band around 70 KDa disappeared (Figure 5-9) and at the same time the band that corresponds to the MW of CRABPII appeared in the eluted fraction (Figure 5-10). However, it can be observed that a band at about 70 KDa also appeared in the eluted fraction. This observation makes it difficult to conclude with certainty whether the absence of the apparent fusion protein was due to cleavage or due to the fact that the observed band was just a random protein, which had a relatively good binding to the chitin beads. The problem was that if the intein was cleaving *in vivo*, the band for the cleaved CRABPII could not be identified due to the fact that there are

many *E. coli* proteins that have similar molecular weight. Also it is possible that the proteases would destroy the cleaved CRABP_{II}. When a time study was performed for the expression of the fusion protein an overexpressed band at the same size as CRABP_{II} was identified; however, it was present in the insoluble fraction. CRABP_{II} is known to be a soluble protein with a very robust structure, thus it seems unlikely that this band corresponds to the cleaved CRABP_{II}. However it is important to mention that 17 amino acids were removed from

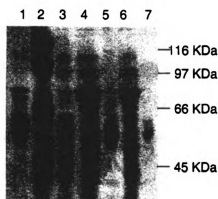


Figure 5-9. 10% acrylamide gel. Cleavage of CRABP_{II}-Sce VMA1 70 KDa. Expected bands after cleavage, Sce VMA1 50 KDa, and CRABP_{II} 20 KDa. 1. chitin beads. 2. Marker. 3. Chitin beads after cleavage with DTT. 4. Elution after cleavage with MESNA. 5. Chitin beads after cleavage with DTT. 6. Elution after cleavage with DTT. 7. Sce VMA1.

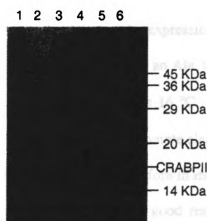


Figure 5-10. 20% acrylamide gel. Cleavage of CRABP_{II}-Sce VMA1 70 KDa. Expected bands after cleavage, Sce VMA1 50 KDa, and CRABP_{II} 20 KDa. 1. Elution after DTT cleavage. 2. empty lane. 3. Marker. 4. CRABP_{II}. 5. empty lane. 6. Elution after MESNA cleavage.

CRABP_{II}, and thus it is possible that the protein could collapse to form inclusion bodies. An attempt to isolate this band using a Fast Q column resulted in no

protein, which suggests that the protein observed is not CRABP_{II}. As mentioned previously cleavage *in vivo* is a frequent problem encountered by intein users.^{14,27} It has been reported that the cleavage *in vivo* depends mostly on the last amino acid in the protein of interest. In our case, the amino acid at position 129, was a valine. According literature reports, the last amino acid fused with the intein is important for the efficiency of the thiol induced cleavage.⁷⁷ In fact, in a comprehensive study to evaluate the efficacy of the twenty amino acids as the C-terminus residue, Val, residue 129 in the CRABP_{II} sequence, is one of the worst possible choices. In particular, in model cases such as that of the expression and purification of maltose-binding protein, if the C-terminal residue is an Ala, the on column cleavage with DTT occurs at 75-100% both at 4 °C and at 16 °C. Even though results that have been reported about which amino acids promote cleavage and which do not, they depend strongly on the protein used. Therefore in most of the reports a variety of different amino acids are tested until a good result is observed.³⁷

At this point, although significant progress has been made towards applying the intein-mediated native chemical ligation for our system, we were not able to isolate enough thiol-activated CRABP_{II} fragment to continue with the final ligation step.

5.4 Labeling of Lys with ϵ -¹⁵N-Lys or 6-¹³C-Lys

The downfall in the use of native chemical ligation is the necessity to form a fused CRABP_{II}-intein protein. As discussed previously this methodology was

tried but many different drawbacks were faced. Therefore, a different methodology was designed. This would consist of the conversion of a cysteine residue to a lysine, which potentially will allow for the selective labeling of the lysine mimic by forming a thiolysine adduct.

5.4.1 Conversion of Cys to Thio-Lys

In proteomics selective alkylation of Cys has been used. Alkylation of Cys has been accomplished using ethylenimine, iodoacetate, 2-bromoethylamine and n-(iodoethyl)trifluoroacetamide.⁷⁸⁻⁸³ In all the previous alkylation cases (except iodoacetate), the alkylated product (**23**) resembles a lysine (**24**) (Figure 5-11). The only difference between the lysine mimic and a lysine is a methylene vs. a sulfur. It could be

envisioned that a chemical transformation of a Cys into a Lys mimic using a labeled alkylating agent would provide the

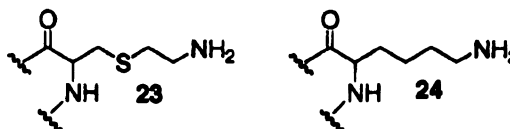


Figure 5-11. Lysine **23** vs lysine mimic (thiolysine) **24**.

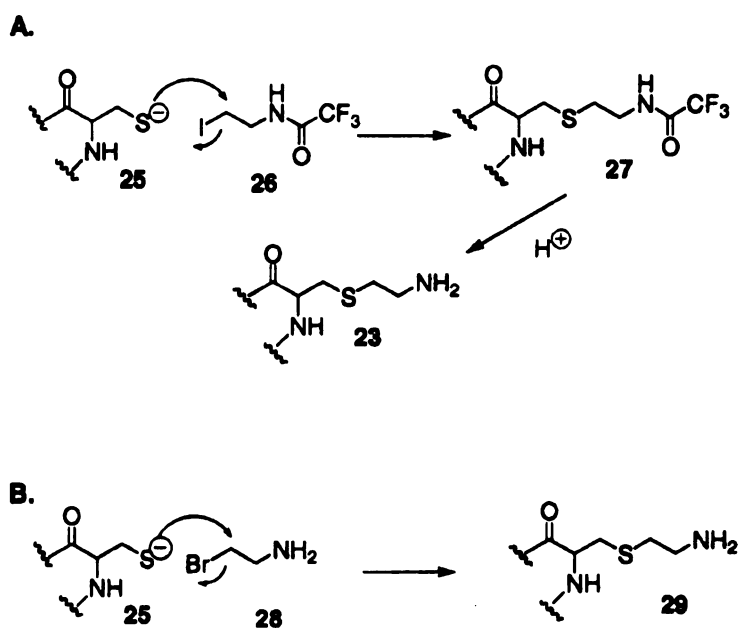
The difference is that lysine contains a CH₂ and the thiolysine a S.

labeled thiolysine. The labeled thiolysine would behave as lysine. Assuming this works the process would be very efficient, mostly because the full protein can be expressed using bacterial expression and a small chemically synthesized molecule will be added afterwards.

It has been previously reported that the lysine mimic **23** behaves as a real lysine. For example, Ubiquitin is a protein that mediates proteolysis, which would

be involved in DNA repair, cell cycle passage, gene transcription among other.⁸⁴⁻

⁸⁶ The experts in that area believed that a lysine residue was responsible for Ubiquitin's activity.⁸⁴ To prove it, Chau and colleagues prepare ubiquitin with a cysteine instead of lysine. It was observed that the cysteine containing protein had no activity. But after treatment with *n*-(iodoethyl) trifluoroacetamide the protein recovered its activity. The alkylating agent *n*-(iodoethyl)trifluoroacetamide works via an S_N2 displacement of the iodide by the sulfur, followed by a hydrolysis of the trifluoroacetamide protecting group under acidic conditions (Scheme 5-8). The fluoroacetamide group in **26** helps to avoid



Scheme 5-8. Conversion of Cys to Thiolysine via alkylation of cysteine **25**. Alkylating agent with Iodoethyl-trifluoroacetamide **26** (A), or bromo ethylamine **28**(B) to form the thiolysine **29**.

the polymerization of the reagent. It has been observed that in the case of ethylenimine and 2-bromoethylamine sometimes the alkylation is not accomplished due to polymerization of the alkylating agent.

Even though it has been reported that **23** could be a potential lysine mimic, in order to selectively label CRABPII only one Cys should be present in the protein structure. In the wild type CRABPII sequence three cysteines are present. Two of those cysteines are located close to the surface and far away from the PSB (Cys81 and Cys95) and one (Cys130) inside of the pocket close to the active site Lys residue (Figure 5-12). Because Cys81 and Cys95 are on the surface, therefore

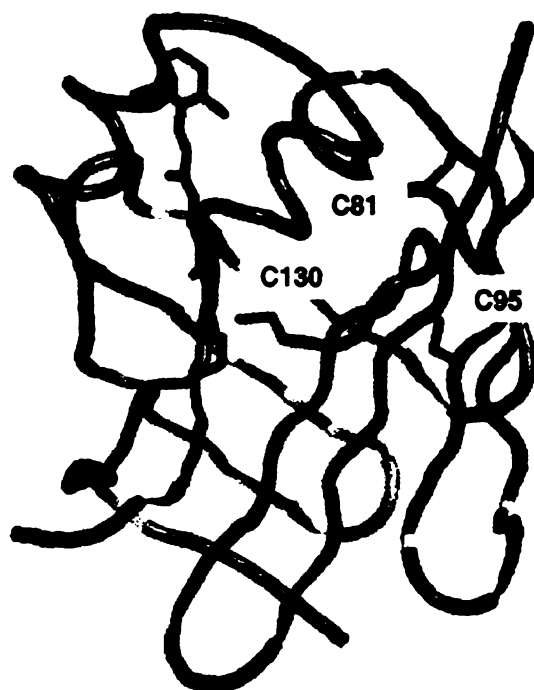


Figure 5-12. In the CRABPII sequence there are three different Cys (C89, C95 and C130). Two of those Cys are close to the surface and far from the PSB.

most likely the rate of alkylation compared to Cys130 and L132C would be slower and potentially not a problem. But if they do they could be mutated to alanine. On the other hand Cys130 is in close proximity with Lys132 and this could pose some difficulties. Both Cys130 and Lys132 are buried deep in the binding pocket, the environment surrounding them is similar. Thus their rate of alkylation would be very similar. If Cys130 is alkylated first it could interfere with the formation of thiolysine 132. But if this indeed pose a major problem it could be easily avoided by mutating C130A.

The triple mutant shows a shift of 447 nm in the UV as well as a 1:1 molar ratio in the formation of protonated Schiff base of retinal and protein. Thus this protein was used as a template to test this methodology. The first step then, would be to prepare the R132C::R111L::L121E mutant. This mutant would retain the C130, and C132 would be converted into the active thiolysine 132. If the alkylation is successful the thiolysine would be formed. Thiolysine 132 would act as a Lys and therefore after addition of retinal we would expect that the thiolysine would form a PSB with retinal. The formation of PSB would cause a red shift wavelength. To prepare the R132C::R111L::L121E mutant standard mutagenesis protocols were used (for details see the Materials and Methods section). The primers used were 5'- *GACGTTGTGTGCACCTGCGTCTACGTCCGAGAG*-3' and 5'- *CTCTCGGACGTAGACGCAGGTGCACACAACGTC*-3'. The PCR product was transformed into XL1-Blue after digestion with Dpn1. The plasmid DNA was isolated and sent for sequence to confirm the mutation. The DNA was transformed into BL21(DE3)pLysS and the protein was expressed by addition of

IPTG and purified using the FPLC (for details see Materials and Methods). The protein was expressed using the optimal conditions developed for CRABPII-mutants. The expression was uneventful and provided about 25 mg/L of protein. When this mutant was titrated using retinal, as expected, no shift in the wavelength was observed (Figure 5-13). The ϵ for R132C::R111L::L121E was calculated to be 20,070.

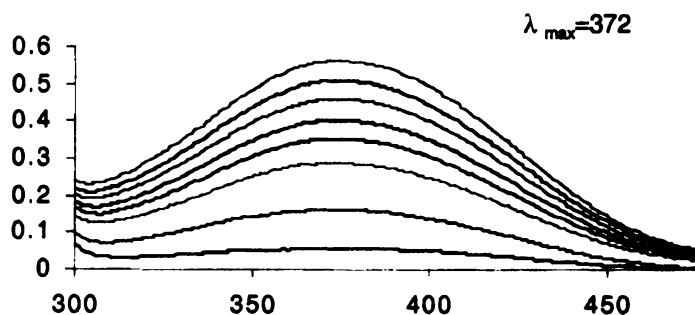


Figure 5-13. Titration of R132C::R111L::L121E with all-*trans*-retinal. Only a peak at 372 nm is present, no PSB formation was observed. Each line corresponds to addition of 0.1 equivalent of retinal.

To convert the cysteine to the lysine mimic the first alkylating agent used was *n*-(iodoethyl)trifluoroacetamide. This alkylating agent was chosen for two major reasons.

The first one was to avoid polymerization, due to the presence of the protecting group on the amine. The second one is because the protecting group is easily removed. According to the literature a variety of conditions have been used for alkylation of cysteins using *n*-(iodoethyl)trifluoroacetamide. Most of these

used basic conditions (pH 9-10) and high temperatures (50-90 °C); the proteins are also denatured to assure that all Cys are alkylated.

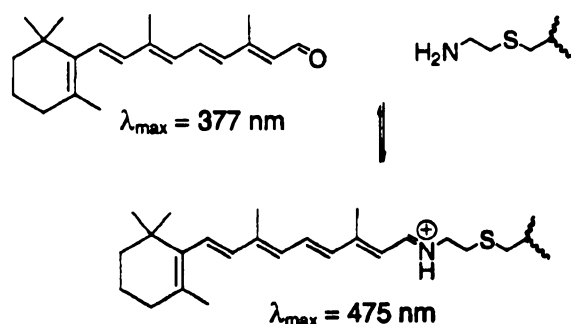
To convert the cysteine to the lysine mimic the first alkylating agent used was *n*-(iodoethyl)trifluoroacetamide. This alkylating agent was chosen for two major reasons.

The first one was to avoid polymerization, due to the presence of the protecting group on the amine. The second one is because the protecting group is easily removed. According to the literature a variety of conditions have been used for alkylation of cysteins using *n*-(iodoethyl)trifluoroacetamide. Most of these used basic conditions (pH 9 - 10) and high temperatures (50 - 90 °C); the proteins are also denatured to assure that all Cys are alkylated.

Because we wish to have some selectivity among the different Cys, and keep a folded protein, relatively mild conditions were used. Thus the first attempt was using 100 equivalent of **26** at pH 8.5 at 50 °C for 6 hours, and this was followed by a 10 x dilution with buffer at pH 4.5. The reaction was left overnight in order to remove the protecting group. The protein was concentrated (for details see Materials and Methods) and the UV-vis was recorded. The UV of the protein showed a very big signal around 250 nm and when the protein obtained was titrated with retinal the results obtained were inconsistent. Thus, the next time the method was modified. First 100 equivalents of **26** were added to about a 1.5 mL of a 0.5 mg/mL protein solution (two tubes) at pH 8.5 at 50 °C for 6 hours. This solution was then concentrated using mini filters, followed by a buffer exchange to pH 7.5 (phosphate buffer), and then, the pH was adjusted to 4.0 (HCl 0.5 M).

After an overnight incubation the buffers were exchanged one more time to pH 7.5.

The resulting protein did not have the extra peak at 250 nm, and therefore it was titrated with retinal. The results were promising, since after addition of 0.1 equivalents of retinal a shift to 457 nm is observed (Figure 5-14, Scheme 5-9). The peak at 457 nm indicates the formation of the PSB between the retinal and the thiolysine (lysine mimic).



Scheme 5-9. The thio-Lys acts as Lysine and it further forms a PSB when treated with retinal.

Interestingly this peak is 8 nm red shifted from the λ_{max} observed in the KLE mutant. Whether this shift is caused by the sulfur in the thio-lysine or not is not clear. This observation proved the feasibility of this method. Unfortunately the protein is active only at low equivalents of retinal.

There are many possible reasons why the overall ratio of equivalents of retinal vs. protein is low. Most probably the problem was due to low alkylation yield in the cysteine alkylation. Other problems could be over alkylation, decomposition of the alkylating agent or denaturation of the protein.

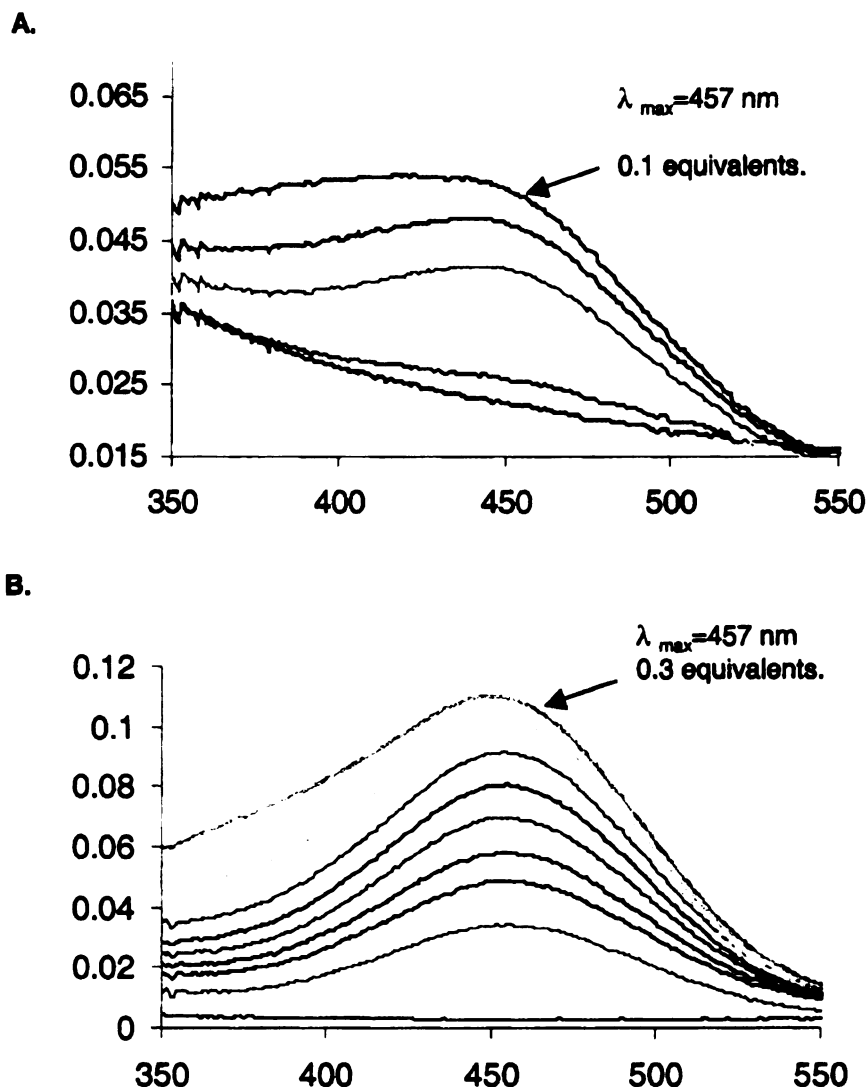


Figure 5-14. Titration of the alkylated mutant R132C::R111L::L121E treated with iodoethyl-trifluoroacetamide **26**, (alkylation with bromo ethylamine **28** showed similar results, therefore it is not shown). **(A)** Titration of the alkylated triple mutant with 100 equivalents of the alkylating agents. The λ_{\max} observed at 457 nm corresponds to the formation of the PSB of the thiolysine with retinal after addition of 0.1 equivalents of retinal only the 380 increased. **(B)** Titration of the alkylated triple mutant with 300 equivalents of the alkylating agents. The λ_{\max} observed at 457 corresponds to the formation of the PSB of the thiolysine with retinal after addition of 0.3 equivalents of retinal only the 380 nm absorption increased.

To test the alkylation, MALDI-TOF analysis was used (Figure 5-15). After incubation of alkylating agent with the protein an S_N2 displacement would occur. The calculated mass protein for the triple mutant was $M^+=15482.7$ Da. After alkylation occurs under basic conditions an $M^+=15484.8$ Da corresponding to the mass $[M+141]^+=15623.7$ Da, should be observed due to the addition of $CH_2CH_2NH_2C(O)CF_3$ (141). Finally after acidic hydrolysis an $[M+44]^+=15484.8$ Da would correspond to a covalently linked $CH_2CH_2NH_2$ $M^+=15523.7$ Da. The results are shown in Figure 5-15. Figure 5-15A shows the CRABP II R132C::R111L::L121E $M^+=15484.8$ Da corresponding to the protein mass, calculated to be 15482.7 Da. Figure 5-15B shows an $M^+=15484.8$ Da corresponding to the protein mass. A second peak, $M^+=15528.64$ Da, 41 mass units higher than the protein peak, represents a covalently linked $CH_2CH_2NH_2$ (calculated mass difference 44) and a third peak $M^+=15622.7$ Da, 140 mass units higher than the protein peak, could be three covalently linked $CH_2CH_2NH_2$ (calculated mass difference 44×3 (132)) or one $CH_2CH_2NH_2C(O)CF_3$ (141). Figure 5-15C shows $M^+=15484.8$ Da corresponding to the protein mass, second peak, $M^+=15528.64$ Da, 41 mass units higher than the protein peak, represents a covalently linked $CH_2CH_2NH_2$ (calculated mass difference 44). Interestingly, the results suggested that hydrolysis of the protecting group was occurring even at basic pH.

The samples that were not treated with acid were titrated with retinal. The results shows a shift in the wavelength ($\lambda_{max}=457$ nm). The shift in the UV

supports the observed formation of the thiolysine $[M+41]^+$. The presence of M^+ mass suggests that alkylation is not complete. The $[M+140]^+$ suggests that

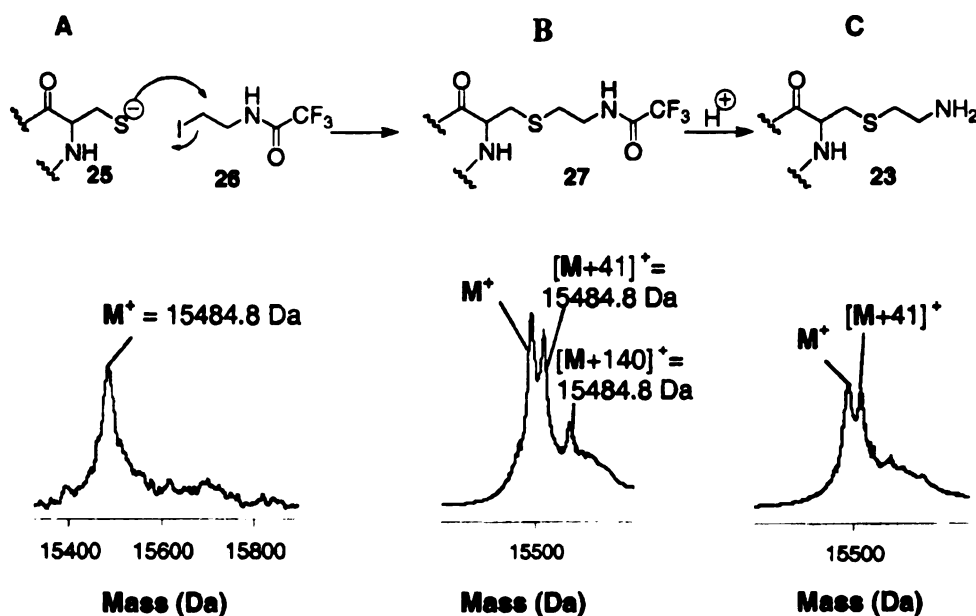


Figure 5-15. MALDI-TOF of the alkylated CRABPII R132C::R111L::L121E. **A.** CRABPII R132C::R111L::L121E $M^+ = 15484.8$ Da corresponding to the protein mass, calculated to be 15482.7 Da. **B.** $M^+ = 15484.8$ Da corresponding to the protein mass, a second peak, $M^+ = 15528.64$ Da, 41 mass units higher than the protein peak, represents a covalently linked $\text{CH}_2\text{CH}_2\text{NH}_2$ (calculated mass difference 44) and a third peak $M^+ = 15622.7$ Da, 140 mass units higher than the protein peak, could be three covalently linked $\text{CH}_2\text{CH}_2\text{NH}_2$ (calculated mass difference 44×3 (132)) or one $\text{CH}_2\text{CH}_2\text{NH}_2\text{C}(\text{O})\text{CF}_3$ (141). **C.** $M^+ = 15484.8$ Da corresponding to the protein mass, second peak, $M^+ = 15528.64$ Da, 41 mass units higher than the protein peak, represents a covalently linked $\text{CH}_2\text{CH}_2\text{NH}_2$ (calculated mass difference 44).

multiple alkylation occurs (Figure 5-16). The rest of the experiments performed were done only in basic solutions.

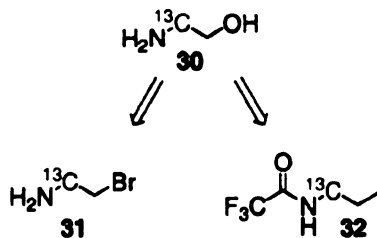
Different temperatures were tested for alkylation, and each sample was treated with retinal. It was found that 50 °C was the temperature at which the ratio of retinal vs. protein (to form PSB) was optimal. When the pH was changed, no improvement was observed. When the equivalents of the alkylating agent **26** were increased to 300 equivalents at pH 8.6, with 50 °C overnight incubation, alkylation was improved, the UV-vis titrations shown a 1:0.3 equivalents binding of retinal vs. protein.

It has been reported that alkylation using 2-bromo ethylamine (**28**) results in low alkylation yields, mainly due to the polymerization of the alkylating agent. Interestingly when alkylation using **28** was performed the same results as with **26** were observed. Even though the protein obtained was only active at low equivalents we believed that only the labeled ^{13}C thiolysine would show a distinguishable NMR signal. Thus we could monitor the change in the protonation state of the labeled thiolysine, which will allow for the calculation of the thiolysine pK_a .

5.4.2 Synthesis of the alkylating agent to convert Cys to Thio-Lys

The next step was the synthesis of the ^{13}C -labeled alkylating agent. As mentioned previously, formation of the thiolysine was attempted to insert the ^{13}C label at the ϵ -position in the lysine mimic in order to use this labeled protein for ^{13}C -NMR experiments. However, the 2- ^{13}C -Iodo-2-(trifluoroacetyl amino)ethane is

not commercially available, thus it must be synthesized. For the synthesis of 2-¹³C-1-iodo-2-(trifluoroacetyl-amino)ethane we



decided to use the commercially available N-¹³C-ethanolamine as the source of the label ¹³C (Scheme 5-10). The amine could be selectively protected as the trifluoroacetamide

Scheme 5-10. The label ethanolamine (30) could be used to prepare iodoethyl-trifluoroacetamide (32), and bromo ethylamine (31).

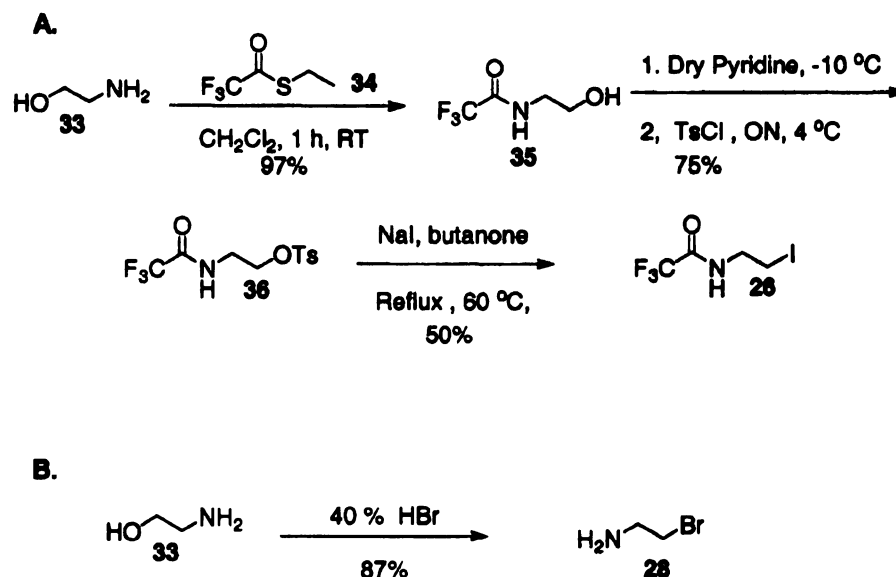
and the alcohol in the ethanolamine could be converted to the iodo compound.

The first attempt to selectively protect the amine was done using trifluoroacetic anhydride. A di-protected product was obtained in 50% when 1 equivalent was used. When two equivalents were used, 100% of the diprotected product was obtained. This reaction was done in DCM and the ethanolamine was not very soluble in it, thus it is possible that the monoprotected product was more soluble in the DCM than the ethanolamine giving the diprotected product as the major product. Thus the reaction was diluted (to increase the solubility of ethanolamine in DCM) but again the same result was observed. Slow addition of the trifluoroacetic acid anhydride did not improve the results. Also lower temperatures were tried but a selective protection was not observed. Because a selective protection was not accomplished, a selective deprotection was attempted. Diprotected ethanolamine (30 mg) was stirred overnight with wet silica and a selective deprotection of the alcohol was observed. However when a larger scale

was tried the reaction took much longer and at the same time the recovery of the product was low (in a 1 g scale reaction only 20% of the mono protected alcohol could be recovered). When a biphasic solution ($\text{CH}_2\text{Cl}_2:\text{H}_2\text{O}$ pH 4) was stirred overnight the same result was observed. The first attempt was by mesylation of the alcohol followed by treatment *in situ* with NaI. This reaction never gave the desired product, mostly starting material was observed but the recovery yield was low. The next attempt was the tosylation of the alcohol to make it a good leaving group, but unfortunately the product was not obtained. Also treatment of the diprotected ethanolamine with I_2 was tried, since the trifluoroacetate is a relatively good leaving group. However, again no desired product was obtained. In the literature,⁸⁷ *s*-ethyl trifluorothioacetate (**34**) has been used to selectively protect amines in the presence of alcohols. This reaction worked very well giving a 97% yield of **35**. Next the alcohol was converted to the tosyl by treatment with *p*-toluene sulfonyl chloride. This reaction gave tosyl **36** in 75% yield. The tosylated compound **36** was refluxed with NaI in butanone to give the desired iodo **26** in 50% yield. All the latter yields correspond to unlabelled reactions (Scheme 5-11A).

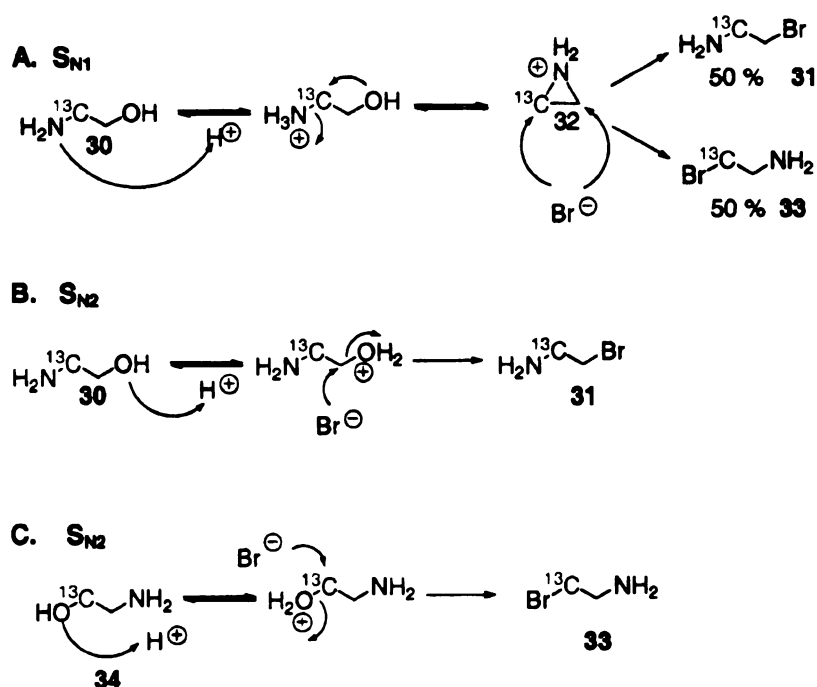
Since we envisioned to use labeled ethanolamine as the source of ^{13}C , we would prefer to have less steps involved in the synthesis of the label thiolysis alkylating agent.

Synthesis of 2-bromo ethyl amine (**28**) from ethanolamine has been previously reported.⁸⁸⁻⁹¹ When the ethanolamine was refluxed in the presence of HBr a 87% yield of the desired bromo ethyl amine (**28**) was obtained.



Scheme 5-11. Synthesis of alkylating agents using ethanol amine as the starting material. **A.** Synthesis of iodoethyl-trifluoroacetamide. **B.** Synthesis of bromo ethyl amine.

The mechanism of bromination of ethanolamine is not clear.⁹² Reports that propose S_N1 and an S_N2 mechanism have been published, however none of them provides conclusive results. The major concern was the fact that if the bromination of the alcohol occurred via the formation of the intermediate aziridine a mixture of **31** and **32** would be observed (Scheme 5-12A). If **30** is used as the starting material, the ^{13}C would be scrambled, and 50% of each isomer would be obtained. Because the synthesis **31** from ethanol amine via bromination could offer the label product after only one reaction this reaction was attempted first regardless of the possibility of scrambling of the label ^{13}C . Therefore, the reaction



Scheme 5-12. Possible mechanisms involved in the bromination of the label ethanolamine (**34**). **A.** S_N1 type mechanism in which aziridine **32** is the intermediate. After bromination, then two products would be obtained in 50%, **31** and **33**. **B.** S_N2 type mechanism in which the bromide attacks displaces the amine directly to afford only **31**. The reaction is observed to go via an S_N2 type mechanism in which the bromide attacks displaces the amine directly to afford only **34**.

using **30** was performed on small scale. The results showed that only one labeled isomer was obtained. Initially this was very encouraging because it suggests that only the desired product **31** was observed. But a closer look at the NMR obtained showed that the product obtained was **33** not **31**.

In order to obtain 100% of this isomer the only possible explanation was that the label in the initial ethanol amine was in the opposite site (**34**), the ^{13}C was next to the O instead of the N. Thus the starting material was analyzed, by NMR revealing that indeed the structure for the molecule was **34** not **30** (Scheme 5-12A, 12B). The labeled ethanolamine was purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA) after the company was contacted, they accepted that they had a mistake in their catalog, and that the desired **31** was not commercially available.

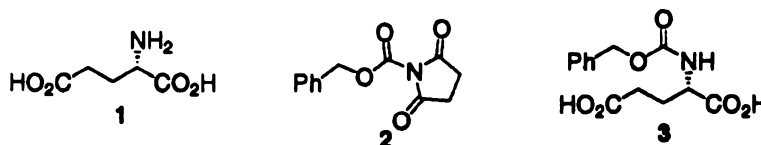
Due to the lack of time the synthesis of **31** using a different source of ^{13}C label was not performed. Even though that advancement in formation of the thiolysine was accomplished the final objective of using the labeled thiolysine to measure the pKa could not be completed. However it is important to mention that for the first time, we can finally prove that the bromination of **31** occurs via an $\text{S}_{\text{N}}2$ type mechanism and not the proposed $\text{S}_{\text{N}}1$ type mechanism previously proposed.

5.5 Materials and Methods

A. Synthesis of Compounds

All reactions were carried under an atmosphere of nitrogen and removal of solvents was performed under reduced pressure with a Buchi rotatory evaporator. THF and Et₂O were freshly distilled from sodium/benzophenone, and CH₂Cl₂ was distilled over CaH₂ under a nitrogen atmosphere. Labeled H₂NCH₂¹³CH₂OH was purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA). Analytical TLC was carried out using Merck 250 mm Silica gel 60 F₂₅₄ and spots were visualized under UV light. Column chromatography was conducted using Silicycle silica gel (230–400 mesh). 300 MHz ¹H-NMR and 75 MHz ¹³C-NMR spectra were recorded on a Varian Gemini-300 or 500 instrument, and the residual protic solvent (CDCl₃, D₂O or DMSO-d₆) was used as internal reference. UV-visible spectra were recorded on a Perkin-Elmer Lambda 40 spectrometer.

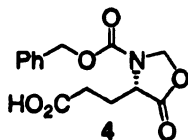
1. Synthesis of N-CBZ-Glutamic acid (3)



A solution of L-glutamic acid **1** (0.59 g, 4 mmol) in a mixture of dioxane (20 mL) and 0.4 M NaOH (20 mL) was treated with N-(benzyloxycarbonyloxy)succinimide **2** (1.2 g 5 mmol) at 90 °C for 1 h. The dioxane was evaporated under

reduced pressure and the product was partitioned between EtOAc and water. The aqueous layer was acidified with HCl and extracted with EtOAc. Evaporation of the organic layer yielded N-benzoyloxycarbonyl)-L-glutamic acid **3** as a partially crystallized gum (1.3 g crude 99%) which was used with no additional purification. [^1H NMR (300 MHz CDCl_3): δ 9.89(2H, b, COOH), 7.23(5H, s, ArH), 5.68(1H, d, $J=7.41$ Hz), 5.06(2H, s, PhCH_2), 4.41(1H, dd, $J=6.87, 7.41$ Hz), 2.64(3H, s, CH_3O), 2.45(2H, t, $J=6.04$ Hz CH_2), 2.17(1H, m, CH_2), 2.06 (1H, m, CH_2). ^{13}C NMR (300 MHz CDCl_3): δ 177.94, 172.7, 135.8, 128.51, 128.29, 128.07, 127.66, 127.05, 60.49, 25.24, 14.121. m/z (E.I): 281(M^+), 262.9, 235.9, 192.0 (100)].

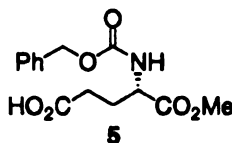
2. Synthesis of N-CBZ-Glutamic acid oxazolidinone (**4**)



A mixture of the diacid (0.4 g 1.42 mmol), paraformaldehyde (0.12 g 2.66 mmol) and *p*-toluene sulfonic acid $\cdot\text{H}_2\text{O}$ (40 mg) in toluene (100 ml) was refluxed in a Dean-Stark apparatus for 2 h. The cooled mixture was diluted with EtOAc until it became clear, and then was washed with a solution of NaHCO_3 (16 mg) in a few mL of water. Evaporation of the organic layer afforded the oxizolidinone **4**, as an oil (0.32 g in 69% yield), which was used with no additional purification. [^1H NMR (300 MHz CDCl_3): δ 9.63(1H, bs, COOH), 7.34(5H, s, ArH), 5.5(1H, br, s, 1H, H-2), 5.19(1H, d, $J=4.5$ Hz, H-2), 5.16(2H, s, PhCH_2), 4.37(1H, t, $J=5.7$ Hz), 2.46(2H, m), 2.27(1H, m), 2.16(1H, m). ^{13}C NMR (300 MHz CDCl_3): δ

177.23, 171.67, 153.00, 135.16, 128.57, 128.51, 128.42, 128.21, 68.04, 53.78, 29.95, 25.58. m/z (E.I): 293(M^+), 274.9(100), 247.9, 235, 204, 185.8

3. Synthesis of N-CBZ-Glutamic acid a methyl ester (**5**).

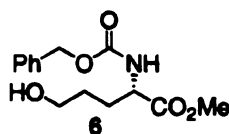


A solution of oxazolidinone **4** (0.32 g) in MeOH (10 mL) was treated with a NaOMe in MeOH (1.4 M, of 3.5 mL). The solution was kept at room temperature for 40 min and then neutralized with 1 N HCl (6 mL). The MeOH was evaporated and the residue was partitioned between ethylacetate and HCl (0.1 N). The organic layer was evaporated under reduced pressure to obtain the monoacid as oil (0.27 g). The oil was re-dissolved in EtOAc, N, N-dicyclohexylamine (360 μ l, 1 eq) was added and the solution was allowed to evaporate passively until crystals of the dicyclohexylammonium salt were formed.

A solution of the previous crystals in 0.2 N HCl was extracted with EtOAc and the organic layer was evaporated to dryness to obtain the free monoacid as an oil (340 mg, 40% yield after three steps). The oil solidified to form white crystals of **5**. [^1H NMR (300 MHz CDCl_3): δ 10.72(1H, bs, COOH), 7.32(5H, s, ArH), 5.56(1H, d, $J=7.99$ Hz), 5.07(2H, s, PhCH_2), 4.41(1H, dd, $J=5.2$ Hz), 3.71(3H, s, CH_3O), 2.40(2H, m), 2.16(1H, m), 1.97 (1H, m). ^{13}C NMR (300 MHz CDCl_3): δ 177.91, 172.34, 56.02, 136.00, 128.475, 128.17, 128.03, 67.10, 53.09, 52.53, 29.79, 27.29. m/z (E.I): 295(M^+), 277.0, 235.9. IR: 3331.49(br), 2955.32,

2361.17, 2339.95, 1734.23, 1701.13. TLC (EtOAc:Hex, 1:1). $R_f=0.27$ (br). $[\alpha]_D^{25}=+7.5$ (c 2.0, CHCl_3)

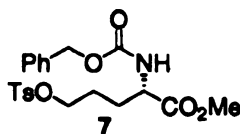
4. Synthesis of N-CBZ-2-Amino-5-hydroxypentanoic Acid Methyl Ester (**6**).



The acid (0.5 g 1.69 mmol) of **5** were dissolved in dry THF (10 mL). The solution was cooled down to $-10\text{ }^{\circ}\text{C}$. $\text{BH}_3\text{-THF}$ (1 M, 2.5 mL, 2.5 mmol) were added drop wise. After complete addition, the reaction was warmed slowly after 3 h the and when the reaction was completed MeOH (10 mL) was added and it was stirred overnight. The reaction was neutralized (with 10% HCl) and the solvents were evaporated. The residue was partitioned in ethylacetate and aqueous sodium hydrogencarbonate. The organic phase was dried with anhydrous sodium sulfate. A column 1:1 hexanes:ethylacetate gave the pure alcohol **6** in 48% yield (230 mg, 0.81 mmol) [Note a brand new bottle of 1M $\text{BH}_3\text{-THF}$ was used]. $^1\text{H NMR}$ (300 MHz CDCl_3): δ 7.32(5H, s, ArH), 5.61(1H, d, $J=7.14$ Hz), 5.07 (2H, s, PhCH_2), 4.37(1H, dd, $J=4.95, 7.14$ Hz), 3.70(3H, s, CH_3O), 3.61(1H, t, $J=10.77$ Hz), 2.15(1H, br s, OH), 1.95-1.85(1H, m), 1.8-1.68(1H, m), 1.65-1.5 (2H, m). $^{13}\text{C NMR}$ (300 MHz CDCl_3): δ 172.91, 156.02, 136.13, 128.45, 128.11, 128.02, 66.94, 61.80, 53.50, 52.33, 29.19, 28.05. m/z (E.I): 281.0(M^+), 263.0, 248(100). IR: 3339.49(br), 3065.28, 3034.41, 2953.39, 2878.16, 17030, 1705.29, 1533.6, 1215.31. TLC (EtOAc:Hex, 1:1). $R_f=0.15$. $[\alpha]_D^{25}=+6.4$ (c 2.0, CHCl_3)

5. Synthesis of N-CBZ-2-Amino-5-(*p*-toluenesulfonyloxy)pentanoic Acid

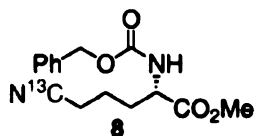
Methyl Ester (7)



(*s*)-N-CBZ-2-amino-5-hydroxypentanoic acid α -methyl ester (400 mg, 1.4 mmol) was dissolved in dry DCM. Triethylamine (1 eq.) and *p*-toluenesulfonyl chloride (1.4 mmol) were added to the mixture and it was stirred overnight. After the reaction was completed, the resulting solution was then quenched with distilled water, acidified (2 M HCl). The organic layer was separated and the aqueous phase was extracted with DCM and dried with anhydrous sodium carbonate. The solvent was removed under vacuum and the product purified by flash chromatography. Elution with 20% ethyl acetate in light petroleum gave (490 mg) the tosylate as a colourless oil in a 80% yield. [¹H NMR (300 MHz CDCl₃): δ 7.74(2H, d, J =8.24, ArH), 7.31(5H, s, ArH), 7.30(2H, d, J =8.24 Hz, ArH), 5.30(1H, d, J =7.69 Hz), 5.06(2H, s, PhCH₂), 4.30(1H, dd, J =7.69, 5.59 Hz), 3.99(1H, t, J =5.59 Hz), 3.70(3H, s, CH₃O), 2.41 (3H, s, CH₃Ar), 2.00-1.80(1H, m), 1.72-1.60(3H, m). ¹³C NMR (300 MHz CDCl₃): δ 172.29, 156.00, 144.82, 136.02, 132.80, 129.83, 128.48, 128.18, 128.02, 127.81, 69.52, 67.12, 53.12, 52.48, 28.77, 24.85, 21.56. m/z (E.I): 434.9(M⁺), 375.9, 331.9, 299.9, 204.0, 16.1, 127.9, 108.0, 90.9(100). IR: 3373.93, 2955.32, 2878.16, 1722.65, 1527.82,

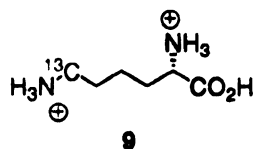
1358.06, 1176.73. TLC (EtOAc:Hex, 20:80). $R_f=0.15$. $[\alpha]_D^{20}=+11.6$ (c 2.0, CHCl_3)

6. Synthesis of [6]-(S)- N-CBZ-2-Amino-5-cyanopentanoic acid methyl ester (**8**)



The tosylate **7** (0.7g, 1.6 mmol) was dissolved in DMF (5mL) and carefully added to sodium ^{13}C -cyanide (0.17g, 4.4 mmol) in DMF (4mL) at room temperature under nitrogen atmosphere. After 24 hours, the reaction was quenched by the addition of aqueous sodium hydrogen carbonate, with the resulting mixture extracted using diethyl ether. The solvent was dried, removed under vacuum and the product purified by flash chromatography. Elution with 20% ethylacetate in hexanes gave nitrile **8** in a 48% yield (0.768 mmol, 220 mg)). ^1H NMR (300 MHz CDCl_3): δ 7.33(5H, s, ArH), 5.39(1H, d, $J=7.69$ Hz), 5.09(2H, s, PhCH_2), 4.38(1H, dd, $J=7.4$ 7.69 Hz), 3.74(3H, s, CH_3O), 2.36(1H, t, $J=6.31$ Hz), 2.40-1.90(1H, m), 1.80-1.60(3H, m). ^{13}C NMR (300 MHz CDCl_3): δ 172.15, 155.86, 135.95, 128.53, 128.26, 128.10, 118.97, 67.14, 52.92, 52.64, 31.75, 21.39, 16.69. m/z (E.I): 290.4(M^+), 231.1(100). IR: 3340.70, 2878.16, 2247.36, 1718.79. TLC (EtOAc:Hex, 20:80). $R_f=0.1$.

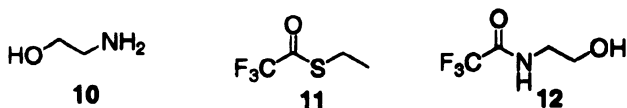
7. Synthesis of ^{13}C -5-Lysine



Nitrile **9** (114 mg 0.393 mmol) was dissolved in a mixture of dioxane (5 mL) and water (1.3 mL). Raney -nickel (120 mg) as a 50% suspension in water and 10% palladium on charcoal (35 mg) were added together with lithium hydroxide monohydrate (35 mg) and the reaction mixture was stirred under hydrogen atmosphere at 50 °C for 20 h. The catalyst were filtered off, the solvents were removed *in vacuo* and the solution was dissolved in a 2 M HCl solution.

Purification with the HPLC fitted with a μ Bondapak NH_2 column was. The amino acid was detected at 200 nm. The mobile phase was 30% 0.01 M KH_2PO_4 /70% $\text{CH}_3\text{CN}/\text{H}_2\text{O}$ (100/16). The [$6\text{-}^{13}\text{C}$]-L-Lysine was purified in a 85% yield.

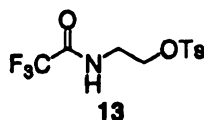
8. Synthesis of (12)



2-aminoethanol **10** (600 μL , 600 mg, 9.8 mmol) was dissolved in DCM (3 mL) and stirred as S-ethyl trifluorothioacetate **11** (1.26 mL) was added drop wise. The mixture was stirred vigorously for 1 h. The solvent was boiled off (in a fume hood), and several portions of DCM were added and evaporated in a N_2 stream to remove the last of the ethane thiol. The remainder of the solvent was removed *in vacuo*, leaving a yellowish crystalline oil 1.5 g of **12** in 97% yield. ^1H NMR (300

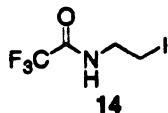
MHz CDCl₃): δ 7.68(1H, bs, NH), 3.79 (2H, t, $J=4.94$, CH₂), 3.53 (2H, q, $J=5.43$, CH₂), 1.8(1H, bs, OH). **¹³C NMR (300 MHz CDCl₃)**: δ 158.04 (d, $J=37.8$ Hz CCF₃), 117.70 (t, $J=287.45$ Hz CF₃), 60.59, 41.97. IR: 3310.70(bs), 2951.46, 1716.86, 1560.61, 912.45.

9. Synthesis of (13)



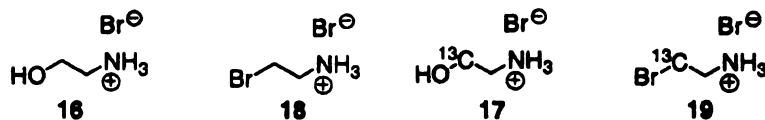
A solution of the alcohol (700 mg, 4 mmol) in dry pyridine (3 mL) was cooled in an ice-salt bath with stirring, and reacted with *p*-toluenesulfonyl chloride (1.715 g, 9 mmol). After 2 h, the solution was transferred to a refrigerator (4 °C) and kept overnight. The suspension was poured into ice-water (80 mL) and mixed vigorously. After a few minutes of stirring in the cold room, according to the reference the product should solidify and then collected on a buchner funnel. However, only an oily substance insoluble in water was observed, thus I extracted added DCM and extracted the aqueous layer 4 times. The organic extracts were combined and dried. After evaporation of the solvent an oil was obtained and it was dried *in vacuo* to leave 1 g of the tosylated product in a 75% yield. [**¹H NMR (300 MHz CDCl₃)**: δ 7.76(2H, d, $J=8.24$ Hz), 7.34(2H, d, $J=8.24$ Hz), 4.13(2H, t, $J=4.94$ Hz, CH₂), 3.60(2H, q, $J=5.45$ Hz, CH₂), 2.44(3H, s, CH₃). **¹³C NMR (300 MHz CDCl₃)**: δ 157.85(d, $J=37.8$ Hz CCF₃), 145.56, 131.95, 130.07, 128.21, 127.84, 117.45 (t, $J=287.45$ Hz CF₃), 67.62, 38.97, 21.58. IR: 3343.06(bs), 2959.18, 1734.23, 1558.68. *m/z* (E.I): 311.1(M⁺), 91.7(100)

10. Synthesis of 14



The tosylate **13** (1 g) was dissolved in butanone (10 mL) of, this solution was added to a NaI solution (0.839 g in 10 mL of butanone). The solution was boiled (about 80 °C) under reflux for 3 hours. The suspension was filtered and the precipitate washed with acetone. The combined filtrates were concentrated *in vacuo*. The resultant residue was partitioned between ethyl ether and sodium hydrogen carbonate. The aqueous phase was washed 2x with ether and the ether extracts were combined and washed with sat NaCl, and dried. The solvent was evaporated to provide a 50% yield. ^1H NMR (300 MHz CDCl_3): δ 3.72(2H, q, $J=6.22$ Hz, CH_2), 3.29(2H, t, $J=6.31$ Hz, CH_2), 1.54(2H, s). ^{13}C NMR (300 MHz CDCl_3): δ 157.25(d, $J=37.8$ Hz CCF_3), 117.45 (t, $J=287.45$ Hz CF_3), 41.92, 1.91. IR: 3304.00(bs), 3100.00, 2361.17, 1705.29, 1558.75. m/z (E.I): 266.6(M^+), 140.0(100).

11. Synthesis of 18 or 19



In a hood, concentrated HBr (40%) (2 eq. 0.072 mmol) was added dropwise to ethanol amine (**16** or **17**) (5 mg, 0.036 mmol) in a distillation flask. The solution was heated to boiling and then it was refluxed over night. The solution was cooled for 20 min. and the water was evaporated using vacuum distillation to afford bromo ethylamine salt **19** (6.3 mg 87%). ^1H NMR (300

MHz H₂O): δ 3.58(2H, dt, $J=120, 5.5$ Hz, Br¹³CH₂), 3.31(2H, t, $J=4.7$ Hz, CH₂NH₃⁺). **¹³C NMR (300 MHz H₂O):** δ 57.66, 27.94.

Protein Mutation, Bacterial Expression and Purification

A. Protein Mutation

Mutagenesis of all CRABPII proteins was accomplished by utilizing Stratagene's QuikChange[®] protocol. This method entails the use of complimentary mutation encoding primers. Following a typical PCR process the reaction mixture will contain circular parental plasmid DNA (unmutated) and linear mutated DNA (Figure 2-32). The original template DNA is different from the mutated DNA, not only in that it is circular, but also it has methylated adenine residues. As the DNA is copied in a bacterial host, some of the adenine bases become methylated. The restriction enzyme *DpnI* recognizes the specific sequence GATC (blunt cut between A and T), but will only cut where methylated adenine residues exist. The plasmids used in our studies contain several cut sites (between 10-20, depending on the vector used), thus incubation of the PCR product with the *DpnI* restriction enzyme effectively destroys the parental plasmid and leaves only the linear, mutated DNA. The desired product DNA can then be incorporated into the host strain of choice (JM109 or XL1-Blue *E. coli*).

B. Plasmid Purification

Following successful transformation into either JM109 or XL1-Blue competent cells (Tet resistant), a colony was grown up in 500 mL LB (containing

appropriate antibiotics) and purified with via Qiagen® column purification. Upon completion of the Qiagen® purification, the recovered DNA was resuspended in approximately 400 µL sterile water, and analyzed by UV-vis for concentration determination and purity.

Concentration of DNA sample (µg / µL) =

$$\text{Abs}_{260} \times (50 \text{ µg} / 1000 \text{ µL}) \times (\text{volume of DNA used} / \text{total volume UV sample})$$

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

= 1.8, pure

>1.8, RNA contamination

< 1.8, protein contamination

C. Sample preparation for sequencing DNA

In a sterilized eppendorf tube (0.5 mL), DNA (2 µg) and primer (30 pmol) were mixed. When sequencing is performed, the largest amount of base pairs that are sequenced with accuracy is about 400 bp. Therefore, to sequence the complete gene, five different primers were used.

1. 5'-GCAGACTGGAATGCAGTGAAGC-3' BBB0052

The DNA must be purified by a Qiagen® column and resuspended in sterile water prior to analysis. Failure to do this will result in poor sequence data. In addition, the sequence sample should be diluted with sterile water, not buffer.

D. Melting temperature calculation for primers

Primers should ideally have melting points ≥ 78 °C, and end in at least one, if not more, GC base pair

$$T_m = 81.5\text{ °C} + (0.41)(\% \text{ GC}) - 675 / N - \% \text{ Mismatch}$$

where N = primer length

E. Site directed Mutagenesis

E.1 Typical PCR Conditions for Stratagene's QuikChange Site-Directed Mutagenesis Protocol:

PCR recipe	
Template DNA	10 ng
Primer 1	15 pMol
Primer 2	15 pMol
DNTP	200 μ M
2 U	
10 x <i>Pfu</i> buffer	5 μ L (1x)
H ₂ O	50-Rxn μ L

PCR program		
1 X	94 °C	5 min
30 X	94 °C	1 min
	48 °C	1 min
	72 °C	1 min
1 X	72 °C	10 min
1 X	25 °C	10 min

E.2. PCR Primers

Primers Sequence		PCR Template
CRABP II-CLE	5'-CGGACATATGGAGATAATATTTGGCCAGA	CRABP II-KLE
	ATAAGAAAGAACAGCTGG-3'	
	5' -GACCCGGGTACCCTCTGGTGCTGTCGGATC-3'	

Following the PCR reaction and agarose gel verification, 1 μL of *DpnI* is added to the reaction tube and allowed to incubate at 37 °C for 1 h, after which time 1 – 5 μL of reaction is transformed into 100 μL of either JM109 or XL1-Blue competent cells.

F. Protein Characterization and Binding Assays

F.1. Calculating Extinction Coefficients of CRABP II mutants

The absorption extinction coefficients (ϵ) for the various CRABP II mutants were determined according to the method first described by Gill and von Hippel.⁹³ The ϵ_{den} (ϵ of denatured proteins in 6 M Guanidine HCl) is dependant primarily on the number of Tyr, Trp and Cys residues present and can be calculated according to equation:

$$\epsilon_{\text{den}} = a\epsilon_{\text{tyr}} + b\epsilon_{\text{trp}} + c\epsilon_{\text{cys}}$$

where a, b, and c represent the number of respective amino acids per molecule of protein, and their ϵ 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 CRABP II 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 NaH_2PO_4 , 16 mM Na_2HPO_4 , 150 mM NaCl, pH=7.3) and the other in a denaturing buffer (6 M guanidine HCl, 4 mM NaH_2PO_4 , 16 mM Na_2HPO_4 , 150 mM NaCl), and the Abs_{280} for each sample is measured.

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

$$\text{Abs}_{\text{den}} + \epsilon_{\text{den}} = C_{\text{den}}$$

$$\text{Abs}_{\text{nat}} + \epsilon_{\text{nat}} = C_{\text{nat}}$$

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

$$\epsilon_{\text{nat}} = \frac{(\text{Abs}_{\text{nat}})(\epsilon_{\text{den}})}{(\text{Abs}_{\text{den}})}$$

the two equations and solve for the extinction coefficient by:

G. Typical preparation of competent cells

The *E. coli* strain of interest was grown (37 °C until OD₆₀₀ of 0.4-0.6, the media contained LB, the corresponding antibiotic). After about three h The cells were harvested by centrifugation at 5,000 RPM for 10 min at 4 °C. The cells were re-suspended (100 mL of 0.9% NaCl) and then centrifugated at 5,000 RPM for 10 min at 4 °C. The cells were re-suspended (50 mL of 100 mM CaCl₂) and incubated for 30 min at 0 °C. The cells were centrifuged at 5,000 RPM for 10 min at 4 °C, and the cells were re-suspended (4 mL of 100 mM CaCl₂, 15% glycerol). The cells in suspension were aliquoted (0.1 mL) and were quickly frozen with liquid nitrogen.

H. Typical expression of CRABPII mutants

The DNA obtained from the site directed mutagenesis protocol was transformed into *BL21(DE3)pLysS* cells using the standard protocol. The cells

were plated on an ampicillin-chloramphenicol LB plates. The number of colonies obtained varied from 50-1000. A single colony was inoculated (100 or 200 mL of LB ampicillin-chloramphenicol) and grown at 37 °C overnight. This culture was used to inoculate a fresh culture (30 ml per liter, LB with ampicillin and chloramphenicol) this culture was grown ($OD_{600} \approx 0.6-1.0$, approximately 3 h). Expression was induced by addition of IPTG (1 mM) and the culture was incubated at 30 °C for about 6 h. The cells were harvested by centrifugation (6000 rpm, 30 min) and frozen at -20 °C overnight. The cell mass was thawed and resuspended in Tris HCl (pH=8.0, 10 mM, 100 mL for 4 L expression). The cells were sonicated (probe sonicator, 60% power, 3x1 min), the mixture was spun down (20 min, 4 °C, 7000 RPM) and the supernatant was collected. At this point $MgSO_4$ (1mM, 300 μ L) and benzonase (20 μ L) were added to reduce viscosity. The mixture was incubated on ice for 30 min. Then the cells were collected (6000 RPM, 30 min) and the supernatant was collected to be loaded on the Fast-Q column.

I. Fast-Q column purification.

A detailed description of the purification of CRABP_{II} mutants is provided in Chapter 2, of Chrysoula Vasileiou's Dissertation. In short, the crude CRABP_{II} expression was purified by affinity chromatography using Q Sepharose, Fast Flow resin. This quaternary ammonium based resin is a strong anion exchanger. After loading the crude protein, the column was washed with 10 mL of 10 mM Tris HCl, pH=8.0. CRABP_{II}-mutants were eluted with 100 mL Buffer of 10 mM Tris

HCl, 100 mM NaCl, pH=8.0. The collected fraction was concentrated and desalted by a stirred Ultrafiltration Cell (Millipore). The concentrated sample was dissolved in 10 mM Tris HCl, pH=8.0 to a final volume of 50 mL in order to be further purified by FPLC.

J. FPLC protocol

A detailed description of the purification of CRABP_{II} mutants is provided in Chapter 2, of Chrysoula Vasileiou's Dissertation. In short, the 50 mL obtained from the Fast Q were applied into FPLC (BioLogics Duo Flow, BioRad). The CRABP_{II} proteins elute at ~4% NaCl. The purity of the protein analyzed by SDS-PAGE electrophoresis, is greater than 90%.

K. UV-vis titrations of CRABP_{II} with all-*trans*-retinal

A stock solution of the all-*trans*-retinal was prepared in 95% spectroscopy grade EtOH. All-*trans*-retinal, $\epsilon=48,000\text{M}^{-1}\text{cm}^{-1}$, $\lambda_{\text{max}}=380\text{ nm}$. For a 5 μM protein solution in a buffer 4 mM NaH₂PO₄, 16 mM Na₂HPO₄, 150 mM NaCl, pH=7.3.

Additions from 0.1 – 1.0 equivalents at 0.1 equivalent increments were added and spectra recorded at room temperature from 200 - 700 nm (Cary WinUV, Varian). The 2nd derivative of the spectra was calculated by using the corresponding software provided with the UV instrument. The determination of the λ_{max} of a UV peak using mathematical equations has been established as a valid method, and has been used in a variety of applications.^{94,95} As previously

mentioned, all the UV data discussed and shown in tables refer to spectra recorded in the presence of 0.1 – 0.2 equivalents of chromophore, unless otherwise noted.

L. Fluorescence and MALDI-TOF.

A detailed description of the determination of binding constant for the CRABP_{II} mutants using retinal, is provided in Chapter 2, of Chrysoula Vasileiou's Dissertation.

M. Molecular Modeling:

Computational results were obtained using software programs from Accelrys. Dynamics and minimizations were done using the Discover 3[®] program, using the CVFF forcefield from within the InsightII 2000 molecular modeling system. Protein figures were generated using PyMol Molecular Graphics System, version 0.93, copyright by DeLano Scientific LLC.

N. Intein Mediated Native Chemical Ligation

PCR	
CRABP _{II} (1-129) (50 ng)	5 µl
BBB24 (NcoI)	4 µl
BBB58 (SapI)	7 µl
MgSO ₄ (500 µM)	1 µl
<i>Deep vent</i> 1 U	1 µl
DNTP (200 µM)	10 mM
Buffer 10x	10 µl
H ₂ O	70 µl

Program		} 30 cycles
94 °C	5 min	
94 °C	1 min	
55 °C	1 min	
72 °C	105 sec	
72 °C	10 min	
23 °C	10 min	

O. QIAquick gel extraction kit protocol

The DNA fragment from the agarose gel was extracted with a clean sharp scalpel. The silica gel was weight and sliced in a colorless tube. Buffer QG was added (3 volumesto 1 volume of gel, 100 mg-100 µl). The mixture was incubated at 50 °C for 10 min until the gels slice were dissolved. After the gel slice was dissolved completely the solution was placed in a QIAquick spin column in a provided 2ml collection tube. To bind DNA, apply the sample to the QIAquick column, and centrifuge for 1 min. The flow-through was discarded and washed with Buffer PE (0.5 ml) and the centrifugation was repeated for 1 min Buffer PE. The DNA was eluted with Buffer EB (50 µl) and centrifuged for 1 min at maximum speed. The average eluant volume is 48 µl from 50 µl elution buffer volume.

P. Digestions

Digestions	
CRABP II (1-129) (from PCR)	48 µl
NcoI	1 µl
SapI	1 µl
Buffer NEB 4 (10X)	6 µl
H ₂ O	4 µl

- o Incubation at 37 °C for 2 hours

- **NEBuffer 4** (20mM Tris-acetate, 10 mM magnesium acetate, 50mM potassium acetate, 1mM dithiothreitol (DTT) and pH 7.9).

Q. Ligations

Ligations		
	1/10	1/15
CRABPII (1-129)	4.5×10^{-13} moles	6.8×10^{-13} moles
pTYB3	4.5×10^{-14} moles	4.5×10^{-14} moles
T4 ligase	1 μ l	1 μ l
Buffer (10X)	1X	1 X
H ₂ O		

- Incubation at 16 °C for 24 hours

R. Intein-CRABPII fused protein expression

Transformation of the target gene (pTYB3-CRABPII) into *E. coli* (ER2566, BL21(DE3)pLysS, Tuner(DE3)pLacI) according to previous protocol affords the expression system. A single colony was inoculated in 100 ml LB/(ampicilin ER2566) (Ampicillin and chloramphenicol in BL21(DE3)pLysS) and grow at 37 °C overnight. A fresh LB/antibiotic resistant medium was inoculated with 30 ml (for 1 L) of the overnight culture and it was grown at 37 °C until OD₆₀₀ of about 1.0. Expression was induced by addition of IPTG (1 mM final concentration) and incubated at 30 °C for 6 h or 16 °C overnight. Harvested cells (centrifugation 7000 RPM for 20 min) and resuspended in Buffer A. The cells were lysed by sonication for 1 min 3 times at 4 °C. After lysis (or not) Triton X-100 was added (final concentration of 1%) and then it was incubated at room temperature for 30 minutes. The insoluble fraction was collected by

centrifugation 7,000 RPM (1 h at 4 °C). The supernatant and insoluble fractions were analyzed by acrylamide gel.

S. Chitin binding and cleavage analysis

About 1 mL chitin resin was used for 2 mL of supernatant or solubilized inclusion bodies. The solution was soaked with the resin for about 2 h. The column was empty with a flow through at rate of 0.3 ml/min. The eluant was collected and resubmitted to the column keeping the same flow rate. The column was washed with 100 mL of 25 mM Tris, 200 mM NaCl, pH 7.5. Cleavage of the intein protein was induced with 1 mM DTT or 10 mM MESNA (fresh solution should be used) at RT overnight.

Expression of CRABPII mutants.

The CRABPII mutants were analyzed by sequencing. Followed by expression of the mutant protein. The methods used were previously described.

T. Refolding experiments

Preparation of the inclusion bodies was achieved by following the protocol of expression of fusion intein-CRABPII proteins. However, no TritonX-100 was added and the pellet contained the inclusion bodies. The refolding experiments were done in small scale (1.5 mL microcentrifuge tubes) and sixteen different conditions were tested. 950 μ L of FoldIt reagent (1-16) were added. Approximately 0.1 mg of the protein was added to each tube. The tubes were placed on a rocker and incubated at 4 °C overnight. The samples were diluted (50

times or 10 times). Each one of the refolding conditions was tested for intein cleavage (following the previously shown protocols).

The content of each one of the refolding systems is shown in the following table.

Solution 1 55 mM Tris pH 8.2 264 mM NaCl 11 mM KCl 0.055% PEG 3350 1.1 mM EDTA	Solution 2 55 mM MES pH 6.5 10.56 mM NaCl 0.44 mM KCl 550 mM Guanidine HCl 2.2 mM MgCl 2.2 mM CaCl	Solution 3 55 mM MES pH 6.5 10.56 mM NaCl 0.44 mM KCl 0.055% PEG 3350 550 mM Guanidine HCl 1.1 mM EDTA 440 mM Sucrose 550 mM L-Arginine	Solution 4 55 mM Tris pH 8.2 264 mM NaCl 11 mM KCl 2.2 mM MgCl 2.2 mM CaCl 440 mM Sucrose 550 mM L-Arginine
Solution 5 55 mM MES pH 6.5 264 mM NaCl 11 mM KCl 2.2 mM MgCl 2.2 mM CaCl 440 mM Sucrose	Solution 6 55 mM Tris pH 8.2 10.56 mM NaCl 0.44 mM KCl 0.055% PEG 3350 550 mM Guanidine HCl 1.1 mM EDTA 440 mM Sucrose	Solution 7 55 mM Tris pH 8.2 10.56 mM NaCl 0.44 mM KCl 550 mM Guanidine HCl 2.2 mM MgCl 2.2 mM CaCl 550 mM L-Arginine	Solution 8 55 mM MES pH 6.5 264 mM NaCl 11 mM KCl 0.055% PEG 3350 1.1 mM EDTA 550 mM L-Arginine
Solution 9 55 mM MES pH 6.5 264 mM NaCl 11 mM KCl 0.055% PEG 3350 550 mM Guanidine HCl 2.2 mM MgCl 2.2 mM CaCl 440 mM Sucrose	I. Solution 9 55 mM Tris pH 8.2 10.56 mM NaCl 0.44 mM KCl 1.1 mM EDTA 440 mM Sucrose	Solution 11 55 mM Tris pH 8.2 10.56 mM NaCl 0.44 mM KCl 0.055% PEG 3350 2.2 mM MgCl 2.2 mM CaCl 550 mM L-Arginine	Solution 12 55 mM MES pH 6.5 264 mM NaCl 11 mM KCl 550 mM Guanidine HCl 1.1 mM EDTA 550 mM L-Arginine
Solution 13 55 mM Tris pH 8.2 264 mM NaCl 11 mM KCl 550 mM Guanidine HCl 1.1 mM EDTA	Solution 14 55 mM MES pH 6.5 10.56 mM NaCl 0.44 mM KCl 0.055% PEG 3350 2.2 mM MgCl 2.2 mM CaCl	Solution 15 55 mM MES pH 6.5 10.56 mM NaCl 0.44 mM KCl 1.1 mM EDTA 440 mM Sucrose 550 mM L-Arginine	Solution 16 55 mM Tris pH 8.2 264 mM NaCl 11 mM KCl 0.055% PEG 3350 550 mM Guanidine HCl 2.2 mM MgCl 2.2 mM CaCl 440 mM Sucrose 550 mM L-Arginine

T. Conversion of Cys to Lys

To a protein (1-10 mg/mL in 25 mM Tris, 200mM NaCl pH 8.00) the alkylating reagent was added in a 10 and 100 fold molar excess over the (number of sulfhydryl groups present in the mutant CRABPII). This calculation should include the sulfhydryl groups contributed by any DTT added. The reaction was incubated at 37 °C for 4 h. A buffer exchange was performed via filtration. (4 mM NaH₂PO₄, 16 mM Na₂HPO₄, 150 mM NaCl, pH=7.3). The obtained concentrated solution was tested for retinal binding (the equivalents of retinal were calculated based on the total amount of protein present). Retinal binding was performed as previously described. In cases where the MALDI-TOF was acquired the same protocol as previously discussed was used.

Note. The stock solution of alkylating agents should be 750 mM in MeOH, and 10 mL were added to a 1 mL of protein solution. The final concentration of methanol in the reaction solution should always be below 10%.

5.6 References

1. S. Gomez; J. A. Peters; T. Maschmeyer, "The reductive amination of aldehydes and ketones and the hydrogenation of nitriles: Mechanistic aspects and selectivity control." *Advanced Synthesis and Catalysis* **2002**, *344*, (10), 1037-1057.
2. C. Vasileou. Protein Design: Reengineering Cellular Retinoic Acid Binding Protein II into a Retinal Binding Protein. Michigan State University, East Lansing, 2005.
3. L. C. Wang; A. Y.; F. Li; J. L. Markley; H. G. Yan, "NMR solution structure of type II human cellular retinoic acid binding protein: Implications for ligand binding." *Biochemistry* **1998**, *37*, 12727-12736.
4. R. M. Czerwinski; T. K. Harris; W. H. Johnson; P. M. Legler; J. T. Stivers; A. S. Mildvan; C. P. Whitman, "Effects of mutations of the active site arginine residues in 4-oxalocrotonate tautomerase on the pK(a) values of active site residues and on the pH dependence of catalysis." *Biochemistry* **1999**, *38*, 12358-12366.
5. R. M. Czerwinski; T. K. Harris; M. A. Massiah; A. S. Mildvan; C. P. Whitman, "The structural basis for the perturbed pKa of the catalytic base in 4-oxalocrotonate tautomerase: Kinetic and structural effects of mutations of Phe-50." *Biochemistry* **2001**, *40*, (1984-1995).
6. A. F. L. Creemers; C. H. W. Klaassen; P. H. M. Bovee-Geurts; R. Kelle; U. Kragl; J. Raap; W. J. d. Grip; J. Lugtenburg; H. J. M. d. Groot, "Solid state N-15 NMR evidence for a complex Schiff base counterion in the visual G-protein-coupled receptor rhodopsin." *Biochemistry* **1999**, *38*, (7195-7199).
7. Y. Inoue; Y. Tokito; S. Tomonoh; R. Chujo, "C-13 Chemical-shifts of retinal isomers and their Schiff-bases as models of visual chromophores." *Bulletin of the Chemical Society of Japan* **1979**, *52*, 265-266.
8. M. Eilers; P. J. Reeves; W. Ying; H. G. Khorana; S. O. Smith, "Magic angle spinning NMR of the protonated retinylidene Schiff base nitrogen in rhodopsin: expression of 15N-lysine- and 13C-glycine-labeled opsin in a

stable cell line." *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 487-492.

9. J. Lugtenburg; A. F. Creemers; M. A. Verhoeven; A. A. C. Wijk; P. J. E. Verdegem; M. C. Monnee; F. Jansen, "Synthesis of C-13-labeled carotenoids retinoids." *Pure Applied Chemistry* **1999**, *71*, 2245-2251.
10. J. A. Pardoen; P. P. J. Mulder; E. M. M. Vanderberg; J. Lugtenburg, "Synthesis of 8-mono-C-13-retinal, 9-mono-C-13-retinal, 12-mono-C-13-retinal, 13-mono-C-13-retinal." *Canadian Journal of Chemistry -Reviews* **1985**, *63*, (1431-1435).
11. J. Lugtenburg, "The synthesis of C-13-labeled retinals." *Pure Applied Chemistry* **1985**, *1985*, (57), 753-762.
12. R. R. Birge; L. P. Murray; R. Zidovetzki; H. M. Knapp, "2-Photon, C-13 and two-dimensional H-1-NMR spectroscopic studies of retinyl Schiff-bases, protonated Schiff-bases, and Schiff-base salts- evidence for a protonation induced pi-pi-star excited-state level ordering reversal." *Journal of the American Chemical Society* **1987**, *109*, 2090-2101.
13. A. Albeck; N. Livnah; H. Gottlieb; M. Sheves, "C-13 NMR-Studies of model compounds for bacteriorhodopsin-factors affecting the retinal chromophore chemical-shifts and absorption maximum." *Journal of the American Chemical Society* **1992**, *114*, 2400-2411.
14. Y. Inoue; Y. Takahashi; Y. T. A.; R. Chujo; Y. Miyoshi, "C-13 NMR-spectra of retinal-1 and its related compounds." *Organic Magnetic Research* **1974**, *6*, 487-491.
15. Y. Inoue; Y. Tokito; R. Chujo; Y. Miyoshi, "Study of pi-electron delocalization in model compounds of visual pigment by UV and C-13 NMR-spectra." *Journal of the American Chemical Society* **1977**, *99*, 5592-5596.
16. G. S. Harbison; S. O. Smith; J. A. Pardoen; J. M. L. Courtin; J. Lugtenburg; J. Herzfeld; R. A. Mathies; R. G. Griffin, "Solid-state C-13 NMR detection of a perturbed 6-S-trans chromophore in bacteriorhodopsin." *Biochemistry* **1985**, *24*, 6955-6962.

17. S. O. Smith; H. J. M. Degroot; R. Gebhard; J. M. L. Courtin; J. Lugtenburg; J. Herzfeld; R. G. Griffin, "Structure and protein environment of the retinal chromophore in light-adapted and dark-adapted bacteriorhodopsin studied by solid-state NMR." *Biochemistry* **1989**, 28, 8897-8904.
18. J. B. Lambert; H. F. Shurvell; D. A. Lightner; R. G. Cooks, *Organic Structural Spectroscopy*. ed.; Prentice-Hall, Inc: New Jersey, 1998; 'Vol.' p 25-27.
19. L. Mou; G. Singh, "Synthesis of (s)-2-amino-8-oxodecanoic acid and apicidin A." *Tetrahedron Letters* **2001**, 42, 6603-6606.
20. J. M. Scholtz; P. A. Barlett, "A convenient differential protection strategy for functional group manipulation of aspartic and glutamic acids." *Synthesis* **1989**, 542-544.
21. P. D. Baley; J. S. Bryans, "Chiral synthesis of 5-hydroxy-(L)-pipecolic acids from (L)-glutamic acid." *Tetrahedron Letters* **1988**, 29, (18), 2231-2234.
22. R. Forch; A. Rosowsky, "Synthesis of γ -[^{15}N]-L-glutamyl derivatives of 5,10-dideazantetrahydrofolate." *Journal of Labelled Compounds and Radiopharmaceuticals* **1999**, 42, 1103-1117.
23. A. Sutherland; C. L. Willis, "Synthesis of [6- ^{13}C]-L-Lysine." *Journal of Labelled Compounds and Radiopharmaceuticals* **1996**, 38, (1), 95-102.
24. R. K. Olsen; K. Ramasamy; T. Emery, "Synthesis of N,N-Protected N-hydroxy-L-ornithine from L-glutamic acid." *Journal of Organic Chemistry* **1984**, 49, 3527-3534.
25. M. Rodriguez; M. Linares; S. Doulut; A. Heitz; J. Martinez, "A facile synthesis of chiral n-protected amino alcohols." *Tetrahedron Letters* **1991**, 32, (7), 923-926.
26. K. Soai; S. Yokoyama; K. Mochida, "Reduction of symmetric and mixed anhydrides of carboxylic acids by sodium borohydride with dropwise addition of methanol." *Synthesis* **1987**, 647-648.

27. E. R. Civitello; H. Rapoport, "Synthesis of the enantiomeric furobenzofurans, late precursors for the synthesis of (+)- and (-)-aflatoxins B1, B2, G1, and G2." *Journal of Organic Chemistry* **1994**, *59*, (14), 3775-3782.
28. H. Walter, "Catalytic reduction of nitriles and oximes." *Journal of the American Chemical Society* **1928**, *50*, 3370-3374.
29. R. Lusskin; J. Ritter, "A new reaction of nitriles. V. Preparation of N-(2-halo-1-ethyl)-amides." *Journal of the American Chemical Society* **1950**, *72*, 5577-5578.
30. F. Benson; J. Ritter, "A new reaction nitriles. III. Amides from dinitriles." *Journal of the American Chemical Society* **1949**, *71*, 4128-4129.
31. H. Ritter; P. Minieri, "A new reaction of nitriles. I Amides form alkenes and mononitriles." *Journal of the American Chemical Society* **1948**, *70*, 4045-4048.
32. W. Huber, "Hydrogenation of basic nitriles with Raney nickel." *Journal of the American Chemical Society* **1944**, *66*, 876-879.
33. B. Klenke; I. H. Gilbert, "Nitrile reduction in the presence of Boc-protected amino groups by catalytic hydrogenation over palladiu activated raney nickel." *Journal of Organic Chemistry* **2001**, *66*, 2480-2383.
34. C. Stephen; J. Duncan; L. Alexandra; R. K. de K.; M. Williams; R. V. Meredith, "A generic approach for the catalytic reduction of nitriles." *Tetrahedron* **2003**, *59*, 5417-5423.
35. H. Ritter; P. Minieri, "A new reaction of nitriles. I. Amides form alkenes and mononitriles." *Journal of the American Chemical Society* **1948**, *70*, 4045-4048.
36. S. Gwoda; C. Gowda, "Application of hydrazinium monoformate as new hydrogen donor with Raney nickel: facile reduction of nitro and nitrile moieties." *Tetrahedron* **2002**, *58*, 2211-2213.

37. M. D. Jaffari; J. T. Mahar; R. L. Bachert Ion Exchange recovery of L-Lysine. 1989.
38. E. Takahashi; M. Furui; H. Seko; T. Shibatani, "D-Lysine production from L-Lysine by successive chemical racemization and microbial asymmetric degradation." *Applied Microbiology and Biotechnology* **1997**, *47*, 347-351.
39. R. N. Costilow; O. M. Fochovansky; H. A. Barker, "Isolation and identification of b-lysine as an intermediate in lysine fermentation." *Journal of Biochemical Chemistry* **1966**, *241*, (7), 1573-1580.
40. R. Schuster, "Determination of free amino acids by high performance chromatography." *Analytical Chemistry* **1980**, *52*, 617-620.
41. J. D. Aberhart; S. J. .Gould; H.-J. Lin; T. K. Thiruvengadam; B. H. Weiller, "Stereochemistry of lysine 2,3aminomutase isolated form *Clostridium subterminale* strain SB4." *Journal of the American Chemical Society* **1983**, *105*, 5461-5470.
42. A. Wlodawer; M. Miller; M. Jaskolski; B. K. Sathyanarayana; E. Baldwin; I. T. Weber; L. M. Selk; L. Clawson; J. Schneider; S. B. H. Kent, "Conserved folding in retroviral proteases crystal-structure of a synthetic HIV-1 protease." *Science* **1989**, *245*, 616-621.
43. M. Miller; J. Schneider; B. K. Sathyanarayana; M. V. Toth; G. R. Marshall; L. Clawson; L. Selk; S. B. H. Kent; A. Wlodawer, "Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3-A resolution." *Science* **1989**, *246*, 1149-1152.
44. J. Schneider; S. B. H. Kent, "Enzymatic-activity of a synthetic-99 residue protein corresponding to the putative HIV-1 protease." *Cell* **1988**, *54*, (3), 363-368.
45. M. Jaskolski; A. G. Tomasselli; T. K. Sawyer; D. G. Staples; R. L. Heinrikson; J. Schneider; S. B. H. Kent; A. Wlodawer, "Structure at 2.5-A resolution of chemically synthesized human-immunodeficiency-virus type-1 protease complexed with a hydroxyethylene-based inhibitor." *Biochemistry* **1991**, *30*, (6), 1600-1609.

46. T. M. Hackeng; P. E. Dawson, "Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology." *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 10068-10073.
47. P. E. Dawson; M. J. Churchill; M. R. Ghadiri; S. B. Kent, "Modulation of reactivity in native chemical ligation through the use of thiol." *Journal of the American Chemical Society* **1997**, *119*, (19), 4325-4329.
48. T. M. Hackeng; C. M. Mounier; C. Bon; P. E. Dawson; J. H. Griffin; S. B. H. Kent, "Total chemical synthesis of enzymatically active human type II secretory phospholipase A2." *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 7845-7850.
49. P. E. Dawson; M. J. Churchill; M. R. Ghadiri; S. B. H. Kent, "Modulation of reactivity in native chemical ligation through the use of thiol additives." *Journal of the American Chemical Society* **1997**, *119*, (19), 4325-4329.
50. T. M. Hackeng; C. M. Mounier; C. Bon; P. E. Dawson; J. H. Griffin; S. B. H. Kent, "Total chemical synthesis of enzymatically active human type II secretory phospholipase A(2)." *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, (15), 7845-7850.
51. T. M. Hackeng; J. H. Griffin; P. E. Dawson, "Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology." *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, (18), 10068-10073.
52. P. E. Dawson; T. W. Muir; I. Clarklewis; S. B. H. Kent, "Synthesis of proteins by native chemical ligation." *Science* **1994**, *266*, 776-779.
53. P. E. Dawson; S. B. H. Kent, "Synthesis of native proteins by chemical ligation." *Annual Review of Biochemistry* **2000**, *69*, 923-960.
54. S. R. Chong; F. B. Mersha; D. G. Comb; M. E. Scott; D. Landry; L. M. Vence; F. B. Perler; J. Benner; R. B. Kucera; C. A. Hirvonen; J. J. Pelletier; H. Paulus; M. Q. Xu, "Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element." *Gene* **1997**, *192*, (2), 271-281.

55. T. T. Hoang; Y. F. Ma; R. J. Stern; M. R. McNeil; H. P. Schweizer, "Construction and use of low-copy number T7 expression vectors for purification of problem proteins: purification of *Mycobacterium tuberculosis* RmlD and *Pseudomonas aeruginosa* LasI and RhlI proteins, and functional analysis of purified RhlI." *Gene* **1999**, 237, (2), 361-371.
56. S. Mathys; T. C. Evans; I. C. Chute; H. Wu; S. R. Chong; J. Benner; X. Q. Liu; M. Q. Xu, "Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation." *Gene* **1999**, 231, (1-2), 1-13.
57. G. J. Cotton; T. W. Muir, "Peptide ligation and its application to protein engineering." *Chemistry and Biology* **1999**, 6, (9), R247-R256.
58. T. Otomo; N. Ito; Y. Kyogoku; T. Yamazaki, "NMR observation of selected segments in a larger protein: Central-segment isotope labeling through intein-mediated ligation." *Biochemistry* **1999**, 38, 16040-16044.
59. K. Alexandrov; I. Heinemann; T. Durek; V. Sidorovitch; R. S. R. S. Goody; H. J. Waldmann, "Intein-mediated synthesis of geranylgeranylated Rab7 protein *in vitro*." *Journal of the American Chemical Society* **2002**, 124, (20), 5648-5649.
60. U. Arnold; M. P. Hinderaker; B. L. Nilsson; B. R. Huck; S. H. Gellman; R. T. Raines, "Protein prosthesis: A semisynthetic enzyme with a -peptide reverse turn." *Journal of the American Chemical Society* **2002**, 124, 8522-8523.
61. B. Ayers; U. K. Blaschke; J. A. Camarero; G. J. Cotton; M. Holford; T. W. Muir, "Introduction of unnatural amino acids into proteins using expressed protein ligation." *Biopolymers* **1999**, 51, (5), 343-354.
62. U. K. Blaschke; J. Silberstein; T. W. Muir, "Protein engineering by expressed protein ligation." *Methods in Enzymology* **2000**, 328, (478-496).
63. U. K. Blaschke; G. J. Cotton; T. W. Muir, "Synthesis of multi-domain proteins using expressed protein ligation: Strategies for segmental isotopic labeling of internal regions." *Tetrahedron* **2000**, 56, (48), 9461-9470.

64. J. A. Borgia; G. B. Fields, *Chemical Synthesis of Proteins. Trends Biotechnology* **2000**, *18*, 6.
65. J. A. Camarero; A. Shekhtman; E. A. Campbell; M. Chlenov; T. M. Gruber; D. A. Bryant; S. A. Darst; D. Cowburn; T. W. Muir, "Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR." *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, (13), 8536-8541.
66. H. D. Mootz; T. W. Muir, "Protein splicing triggered by a small molecule." *Journal of the American Chemical Society* **2002**, *124*, (31), 9044-9045.
67. T. Otomo; K. Teruya; U. K.; T. Yamazaki; Y. Kyogoku, "Improved segmental isotope labeling of proteins and application to a larger protein." *Journal of Biomolecular NMR* **1999**, *14*, 105-114.
68. T. J. Tolbert; C. H. Wong., "Intein-mediated synthesis of proteins containing carbohydrates and other molecular probes." *Journal of the American Chemical Society* **2000**, *122*, 5421-5428.
69. M. Q. Xu; T. C. J. Evans, "Purification of recombinant proteins from *Escherichia coli* by engineered inteins." *Methods in Molecular Biology* **2003**, *205*, 43-68.
70. G. Amitai; S. Pietrokovski, "Fine-tuning an engineered intein." *Nature Biotechnology* **1999**, *17*, (9), 854 - 855.
71. Z. Y. Li; J. H. Fan; J. M. Yuan, "Single-column purification of recombinant human neurotrophin-3 (hNT3) by using the intein mediated self-cleaving system." *Biotechnology Letters* **2002**, *24*, 1723-1727.
72. W. A. Prinz; F. Åslund; A. Holmgren; J. J. Beckwith, "The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm." *Journal of Biological Chemistry* **1997**, *272*, (25), 15661-15667.
73. E. LaVallie; E. A. Diblasio; S. Kovacic; K. L. Grant; P. F. Schendel; J. M. McCoy, "A thioredoxin gene fusion expression system that circumvents

inclusion body formation in the *Escherichia coli* cytoplasm." *Bio/Technology* **1993**, *11*, 187-193.

74. L. Lobel; S. Pollak; B. Lustbader; J. Klein; J. W. Lustbader, "Bacterial expression of a natively folded extracellular Domain fusion protein of the hFSH receptor in the cytoplasm of *Escherichia coli*." *Protein Expression and Purification* **2002**, *25*, (1), 124-133.
75. L. I. Lobel; S. Pollak; J. Klein; J. W. Lustbader, "High-level bacterial expression of a natively folded, soluble extracellular domain fusion protein of the human luteinizing hormone/chorionic gonadotropin receptor in the cytoplasm of *Escherichia coli*." *Endocrine* **2001**, *14*, (2), 205-212.
76. M. Albrecht; N. Misawa; G. Sandman, "Metabolic engineering of the terpenoid biosynthetic pathway of *Escherichia coli* for production of the carotenoids -carotene and zeaxanthin." *Biotechnology Letters* **1999**, *21*, 791-795.
77. In *IMPACTTM-CN System Instruction manual*, New England Biolabs, ed.; 'Ed.' 'Eds.' 'Vol.' p^pp.
78. L. Gregori; M. S. Poosch; G. Cousins; V. Chaw, "A uniform isopeptide-linked multiubiquitin chain is sufficient to target substrate for degradation in ubiquitin mediated proteolysis." *Journal of Biological Chemistry* **1990**, *265*, (15), 8354-8357.
79. S. Hartmann; J. Hofsteenge, "Properdin, the positive regulator of complement is highly C. mannosylated." *Journal of Biological Chemistry* **2000**, *275*, (37), 28569-28574.
80. J. Piotrowski; e. al, "Inhibition of the 26 S proteasome by polyubiquitin chains synthesized to have different lengths." *Journal of Biological Chemistry* **1997**, *272*, (38), 23712-23721.
81. C. E. Hopkins; e. al, "Chemical modification rescue assessed by mass spectrometry demonstrates that thia-lysine yields the same activity as lysine in aldolase." *Protein Science* **2002**, *11*, 1591-1599.
82. S. N. Gorlatov; T. C. Stadtman, "Human thioredoxin reductase from HeLa cells: Selective alkylation of selenocysteine in the protein inhibits enzyme

activity and reduction with NADPH influences affinity to heparin." *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 8520-8525.

83. M. A. Raftery; D. Cole, "On the aminoethylation of proteins." *Journal of Biological Chemistry* **1966**, *241*, (15), 3457-3461.
84. M. Rechsteiner, "Ubiquitin-mediated pathways for intracellular proteolysis." *Annual Reviews in Cellular Biology* **1987**, *3*, 1-30.
85. A. Hershko, "Ubiquitin-mediated protein degradation." *Journal of Biological Chemistry* **1988**, *263*, 15237-15240.
86. M. G. Goebel; e. al, "Yeast cell cycle gene CDC34 encodes a ubiquitin-conjugating enzyme." *Science* **1988**, *241*, 1331-1335.
87. J. Slama; R. R. Rando, "The synthesis of glycolipids containing a hydrophilic spacer-group." *Carbohydrate Research* **1981**, *88*, 213-221.
88. I. Bird; P. B. Farmer, "The synthesis of deuterium labelled 2-bromoethanol, acrylonitrile, acrylamide, 2-aminoethanol, 2-bromoethylamine hydrobromide and 1-bromo-2-chloroethane." *Journal of Labelled Compounds and Radiopharmaceuticals* **1988**, *27*, (2), 1999-216.
89. P. H. Bach; J. Bridges, "Synthesis of 2-Bromo[1-14C]ethanamine hydrobromide." *Journal of Labelled Compounds and Radiopharmaceuticals* **1981**, *19*, (3), 425-431.
90. V. P. Wystrach; D. W. Kaiser; F. C. Schaefer, "Preparation of ethylenimine and tarieghlenemelamine." *Journal of Organic Chemistry* **1955**, *77*, 5915-5918.
91. E. Schutte; R. J. R. Weakle; D. R. Tyler, "Radical cage effects in the photochemical degradation of polymers: Effect of radical size and mass on the cage recombination efficiency of radical cage pairs generated photochemically from the (CpCH₂CH₂N(CH₃)C(O)(CH₂)_nCH₃)₂Mo₂(CO)₆ (n = 3)." *Journal of the American Chemical Society* **2003**, *125*, 10319-10326.

92. A. S. Nagle; e. al., "Efficient snthesis of beta-amino bromides." *Tetrahedron Letters* **2000**, 41, 3011-3014.
93. S. C. Gill; P. H. Vonhippel, "Calculation of protein extinction coefficients from amino-acid sequence data." *Analytical Biochemistry* **1989**, 182, (2), 319-326.
94. D. C. Lee; J. A. Hayward; C. J. Restall; D. Chapman, "2nd-Derivative infrared spectroscopic studies of the secondary structures of bacteriorhodopsin and Ca-2+-ATPase." *Biochemistry* **1985**, 24, (16), 4364-4373.
95. M. P. Horvath; R. A. Copeland; M. W. Makinen, "The second derivative electronic absorption spectrum of cytochrome c oxidase in the Soret region." *Biophysical Journal* **1999**, 77, (3), 1694-1711.

Chapter 6

Interaction of 11-*cis* retinal and other chromophores with the protein rhodopsin mimics.

6.1 Introduction

We are interested in exploring the nature of protein–substrate interactions at the molecular level since they are at the heart of biochemical events that regulate biological systems. Wavelength regulation in rhodopsin is a remarkable example of such interaction in which different proteins elicit different properties from the same substrate, 11-*cis*-retinal **1**. Visual transduction is initiated by photo-isomerization of the 11-*cis*-retinal chromophore to all-*trans*-retinal **2**, which in turn leads to a series of conformational changes in rhodopsin, and eventually results in the enzymatic cascade responsible for vision (Figure 6-1).^{1,2} As discussed previously, we chose to use wavelength regulation as a model to study protein–substrate interactions and due to the difficulty of the study of rhodopsin proteins, we engineer a protein mimic of rhodopsin, that can act as a surrogate.

The design of the engineered rhodopsin was performed using all-*trans*-retinal **2** instead of 11-*cis*-retinal due to the difficulty in working with 11-*cis*-retinal. Other group members have

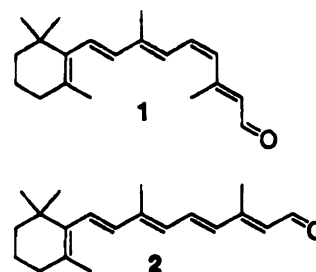


Figure 6-1. Representation of 11-*cis*-retinal and all-*trans*-retinal.

successfully developed a number of successful rhodopsin surrogates that bind all-*trans*-retinal as a PBS (See Crist and Vasilieou's Thesis). Therefore, we have shifted our interest to optimizing our system for binding of 11-*cis*-retinal. At the same time, we were intrigued in how the engineered proteins will interact with different aldehydes that mimic the all-*trans*-retinal.

As mentioned previously we were interested in deconvoluting the factors involved in wavelength regulation. Several theories have been proposed for the mechanism of wavelength regulation.^{3,4} Generally, an increase in the conjugation results in a bathochromic shift, while decreasing the conjugation yields a blue shifted pigment. Twisting about retinal's single bonds decrease de-localization of the PSB, which should result in a blue-shift.³⁻⁷ Using the triple mutant R132K::R111L::L121E we developed a series of mutants with the hope of deconvoluting some of the factors involved in wavelength regulation.

As discussed previously, 11-*cis*-retinal is bound as a Schiff base *via* Lys296, in rhodopsin and Glu113 acts as the counter anion. The binding pocket of rod rhodopsin and the residues directly surrounding the retinal molecule are shown in Figure 6-2A. As can be seen, the 11-*cis*-retinal is twisted due to sterics, which supports the suggestion that twisting around the single bonds dictates different level of conjugation, therefore, changing the observed maximal absorption. The indole ring of Trp265 is close to the C13-Me (3.8 Å) causing some steric interactions. Because deletion of this methyl group is known to cause partial constitutive activity of rhodopsin in the dark,⁸ loss of its interaction with Trp265 may be a possible mechanism of this activity. Also the position of Trp265 can

suggest possible excitonic coupling with the polyene, as originally proposed by Rafferty in 1979.⁹ The breakthrough in the crystallization of bovine rhodopsin presented the opportunity for a better understanding of how the different pigments perform spectral tuning over a wide range of wavelengths. Although the crystal structures of the cone opsins have not been solved yet, their theoretical models have been updated based upon the rhodopsin structure.¹⁰ Figure 6-2A Figure 6-2B and Figure 6-2C shows the models for the three pigments, red, green, and blue opsin.¹¹ It can be observed that in bovine rhodopsin, the environment around the

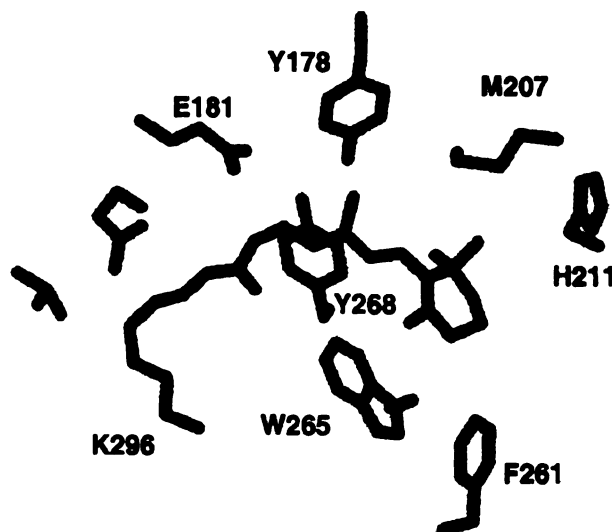
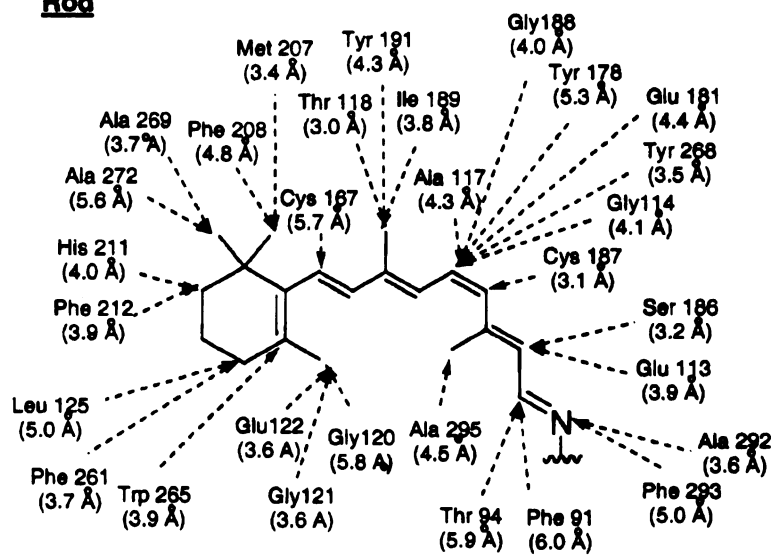


Figure 6-2A. Binding pocket of the rod rhodopsin from the crystal structure. The chromophore, 11-*cis*-retinal is shown bound as a Schiff base to Lys296. Selected important amino acids around the polyene have been labeled.

Rod



1KPN (blue)

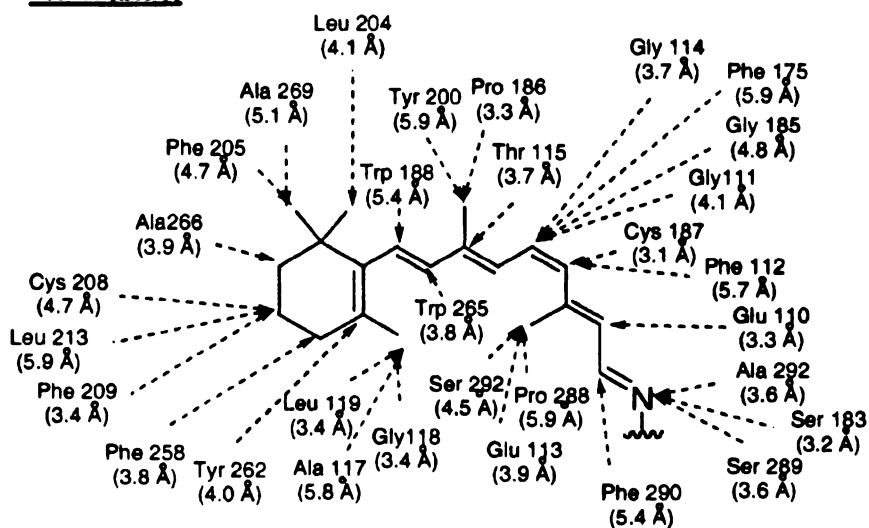
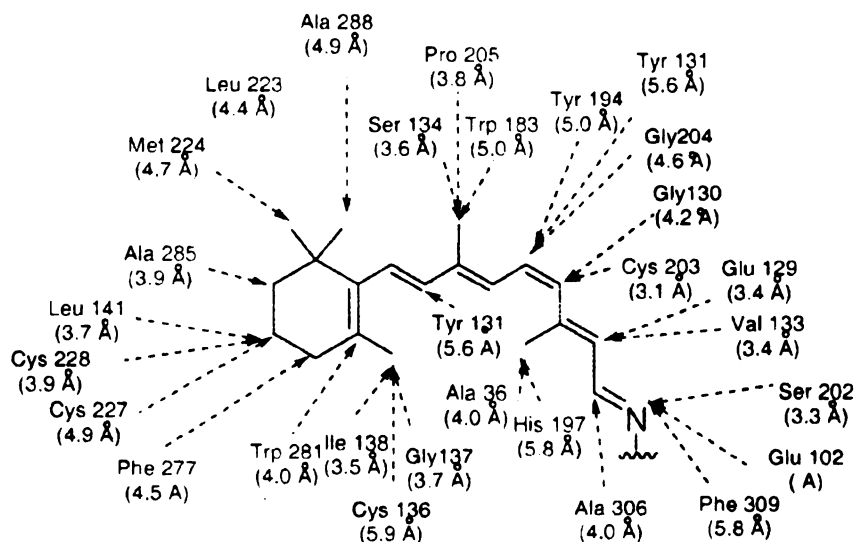


Figure 6-2B. This figure shows the amino acid residues within 6 Å of the polyene for rhodopsin (**Rod**), and the colored pigment blue (**blue**). The rhodopsin data is extrapolated from the crystal structure. The data for the blue were obtained from analysis of theoretical models (RCSB protein data bank accession no 1KPN (blue)).

1KPN (green)



1 KPX (red)

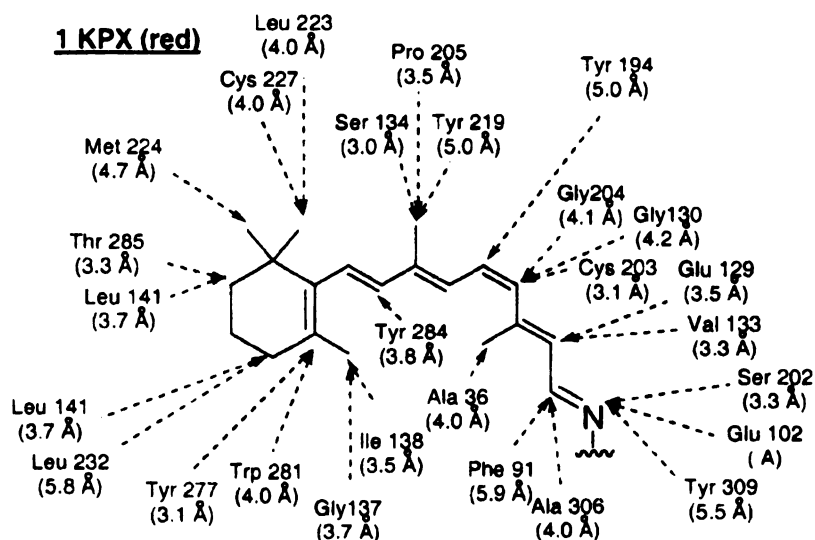


Figure 6-2C. This figure shows the amino acid residues within 6 Å of the polyene for the colored pigment green (**green**) and the colored pigment red (**red**). The rhodopsin data is extrapolated from the crystal structure. The data for the blue were obtained from analysis of theoretical models (RCSB protein data bank accession no 1KPW (**green**) and 1KPX (**red**), respectively).

retinal in all three cone pigments is strongly hydrophobic. Red and green cone pigments share nearly the same sequence (96% homologous, differ by only 15 amino acids). However, their UV-vis absorbance differs by 40 nm (570 nm and 530 nm, respectively).¹² Through site directed mutagenesis Oprian has proven by changing only seven amino acids of the red cone protein, namely S116Y::S180A::I230T::A233S::Y277F::T285A::Y309F mutant, a pigment resembling the green opsin is obtained.^{13,14} As discussed previously, hydrophobic to polar mutations at analogous positions in rod rhodopsin can also produce a red shifted pigment. Rod mutant F261Y::A169T::A164S red shifts to ~700 nm, an amazing 200 nm shift.¹⁵

Although the latter residues are important, as can be seen in Figure 1-13B and C, they are not the residues that the model places in the vicinity of the chromophore. According to the latter, the central residue forming the cavity is Trp281, analogous to Trp265 in bovine rhodopsin. The counter anion to the SB is Glu129, while a second carboxylate, Glu102, is located close to Lys312, the Lys covalently linked to retinal.

In the case of the blue cone pigment, which only bears a 46% homology with the other cone pigments,^{12,16} the central residue forming the cavity is Tyr262 (Figure 1-13A). Because of the proximity of the retinal β -ionone ring, Tyr262 is considered to be the major factor in the blue shifting of this pigment. This explanation agrees with the experimental observation that single mutant Y262W generates a 10 nm red shifted absorption spectrum.¹⁷ In addition, the second

glutamic acid present in red and green pigments is absent in this protein, which is believed to cause an additional blue shift.

Rhodopsin and the related cone pigments are members of a broader family of membrane bound light transducing proteins.¹⁸ This group of proteins contain a seven transmembrane helices motif with a binding site imbedded in the interior of the protein. All of them have a photoactive chromophore (retinal) attached as a Schiff base to a conserved Lys residue, which upon absorption of light isomerizes, triggering conformational changes translating into either light-driven ion transport or photosensory signaling.^{19,20}

Bacterial and archaea rhodopsins which also belong to this group of proteins, perform an ion pump action and significant body of work has been carried out in an attempt to elucidate their mode of action. Bacteriorhodopsin (bRh), the purple light-harvesting membrane protein of *Halobacterium halobium*, is responsible of the conversion of light energy into a proton gradient across the bacterial membrane.²¹ Bacteriorhodopsin binds all-*trans*-retinal through Lys216 as a PSB absorbing at 570 nm.³ Similar to rhodopsin, several theories have been proposed to account for the 70 nm opsin shift observed (relative to rod rhodopsin), most of them involving the electronic environment around the polyene chain.^{4,7}

The crystal structure of bRh bound to all-*trans*-retinal revealed a binding site were (Figure 6-3) a group of aromatic amino acids surround the SB bound chromophore, joined by several Met and Thr residues.⁶ The pK_a of the bRh PSB has been calculated to be >12, 5 units higher than what is typically observed for protonated SB model compounds.^{5,22,23} This has been attributed to a hydrogen

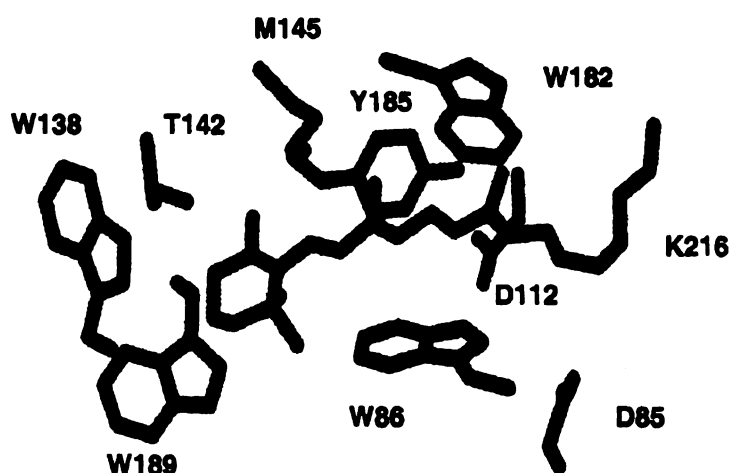


Figure 6-3. All *trans*-retinal binding site of ion pump bacteriorhodopsin (bRh).

bond between the retinal PSB and either a carboxylate or a tyrosinate residue. As can be seen in Figure 6-3, the protonated form of the SB could be stabilized by either of two Asp residues (D112, D85, 3.74 Å and 3.79 Å away from the SB nitrogen, respectively), which have also been proposed to accommodate the proton transfer across the membrane,²¹ or by the nearby Tyr85, although the distance of 4.77 Å between the phenolic oxygen and the SB nitrogen does not support the latter. Interestingly, there is a clear interaction, in the form of a hydrogen bond, between the carboxylate of Asp85 and the tyrosine oxygen, which suggests that one of them can exist in the unprotonated form, to act as a counterion. Previous studies in this area have failed to detect the presence of a tyrosinate ion, suggesting that it is a carboxylate moiety that is responsible for the PSB stabilization.^{24,25} In addition, Arg82 (Figure 6-4) has been identified as a

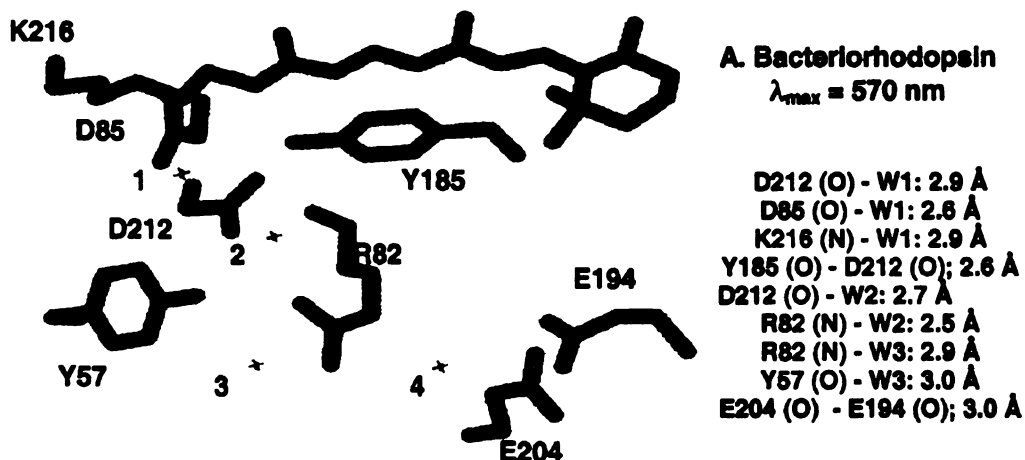


Figure 6-4. Crystal structures of the retinal binding sites of bacteriorhodopsin. The important amino acids surrounding the polyene are highlighted. Red crosses indicate ordered water molecules. Selected distances between residues are shown on the right.

principal counterion, important for the stabilization of the two carboxylates in the vicinity of the SB, while it is believed that the imine proton is strongly hydrogen bonded to a structured water molecule in the binding pocket.^{7,26} As previously mentioned, the counterion structure in bRh is quite unusual, involving Asp85, Asp212 and Arg82. At the same time, there is an ordered water molecule in close proximity to the SB, which acts as the bridge for the interaction with the two aspartates. The importance of Arg82 for the protonation of the bRh SB has been pointed out, although the residue is placed relatively far away from the imine moiety. As can be seen in Figure 6-4, in the bRh case Arg82 is pointing towards Asp212, interacting with it, as well as with Tyr57 through bridged water molecules.

Table 6-1. Spectroscopic characteristics of rhodopsin proteins of different species.

Protein	Retinal	UV of Free Retinal (nm)	UV of SB (nm)	UV of PSB (nm)	PSB pK _a
Rh	11- <i>cis</i>	380	360	500	>16
Octopus Rh	11- <i>cis</i>	380	376	475	10.6
Octopus Rh	9- <i>cis</i>	380	362	468	10.4
Gecko	11- <i>cis</i>	380	376	521	9.9
Gecko	9- <i>cis</i>	380	350	487	9.3

Vertebrates, also use rhodopsin as visual pigments, and it has been observed that in all cases the Glu113, is a conserved residue. Glu113 is considered to be the counterion responsible for stabilizing the PSB formed between retinal and the opsin protein. For invertebrates, retinal remains the chromophore present in all pigments, although for most systems it is the all-*trans*-retinal instead of the 11-*cis*-retinal.^{19,20} Table 6-1 describes the spectroscopic characteristics of Octopus and Gecko rhodopsin proteins, as compared to the bovine Rh.²⁷⁻²⁹ The opsin shift, as well as the pK_a value for the PSB can differ significantly within different systems, or in the presence of a different retinal isomer. In the presence of the 11-*cis*-retinal chromophore, the UV-vis absorbance of octopus rhodopsin is blue shifted while that of the gecko pigment red shifts around 20 nm, as compared to bovine rhodopsin. It is believed that the different placement of the PSB counterion is what determines the spectral behavior, however, there is no crystal structure available to prove the proposed explanation.

The above observations verify the suggestion that the mechanism of wavelength regulation is much more complicated than originally thought, and is probably due to a combination of the factors previously mentioned. It is therefore apparent that since no crystal structures of the cone pigments are available, an alternative approach for the study of the wavelength regulation, involving systems easier to study and manipulate, could be of great use.

6.2 Preparation of mutants to study wavelength regulation

Using CRABP_{II} we have developed a series of mutants with the hope of deconvoluting some of the factors involved in wavelength regulation. Mutagenesis of CRABP_{II} was performed using standard molecular biology protocols. When the CRABP_{II}-mutant plasmids were obtained, the protein was overexpressed in *E. coli* and purified following established protocols (for all procedures mentioned above see Materials and Methods).

We were interested to first explore the theory proposed by Blatz,³⁰ which hypothesizes that the linear distance of the iminium ion conjugation within the polyene can be modulated by placement of an overall negative dipole or point charge at the vicinity of the polyene (Figure 6-5). For example a short conjugation can be

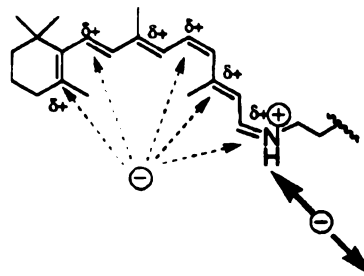


Figure 6-5. Distance of the counter anion to the protonated Schiff base and positioning of charges or dipoles along the backbone of the polyene may modulate the maximal wavelength of the chromophore.

promoted by placing a negative point charge close to the iminium nitrogen. On the other hand, positioning of the negative point charge in the vicinity of the polyene or the ionone ring extends the iminium conjugation and should yield a red-shifted pigment. To explore this point charge theory, a set of mutants were prepared and their spectroscopic properties were studied using all-*trans*-retinal (2), 11-*cis*-retinal (1), merocyanine (3) and azulene (4). Because the triple mutant has been the optimal rhodopsin surrogate obtained so far, this mutant would be used as the standard protein to test the effects of different negative point charges placed along the polyene. Figure 6-6 shows the amino acids that are in proximity to the all-*trans*-retinal in the binding site of CRABPII-R132K::R111L::L121E. The mutant

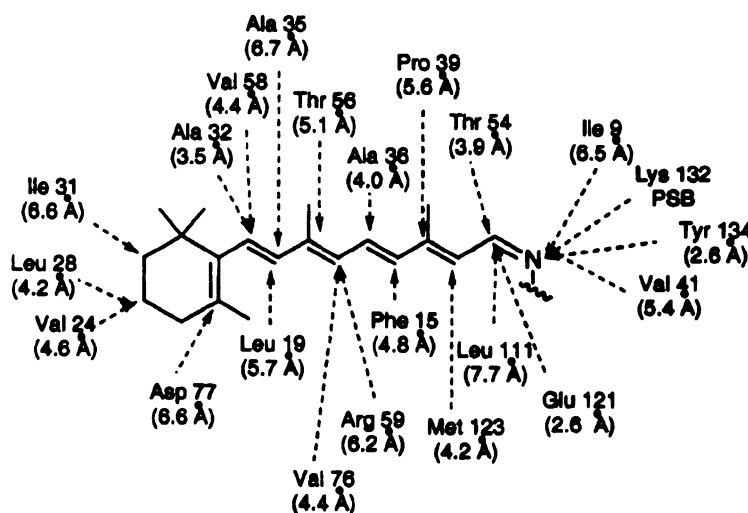
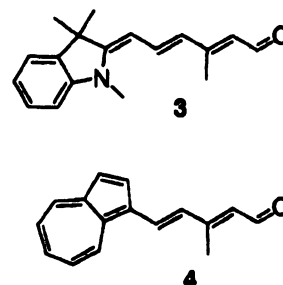


Figure 6-6. This figure represents the amino acid residues within 7 Å of the polyene for the triple mutant CRABPII-R132K::R111L::L121E. The data is extrapolated from the crystal structure.

R132K::R111L::L121E::R59E was prepared to explore the influence of a point charge in the vicinity of the polyene. Arg59 is positioned 6.2 Å away from C-10, right in the middle of the polyene. Therefore, if a Glu is present at this position a negative point charge would be present in the middle of the polyene, thus the delocalization of the positive iminium would be more extensive than the triple mutant R132K::R111L::L121E, thus leading to a red shift in the wavelength. However, when the mutant was incubated with all-*trans*-retinal the wavelength observed was blue shifted 2 nm ($\lambda_{\text{max}}=445$ nm).

The next mutant prepared was R132K::R111L::L121E::A36E. Ala 36 is positioned 4.0 Å away from C-11, and it is also located in the middle of the polyene. Because the Glu is closer to C-11 than the N of the iminium we would expect a larger delocalization of the positive charge leading to a red shifted wavelength compared to the triple mutant R132K::R111L::L121E.

Also some of the mutants that were prepared in the process of developing the triple mutant R132K::R111L::L121E were studied using 11-*cis*-retinal **1** and the retinal mimics (**3**) and (**4**). In Chrysoula Vassileou's Thesis the spectroscopic data of these mutants with all-*trans*-retinal are discussed in detail. Therefore in this chapter we would discuss the properties between the different ligands and their wavelengths observed when they interact with different mutants. In particular we were interested in studying the single mutant R132K, the double mutant R132K::R111L, and the triple mutant R132K::R111L::L121Q. These mutants lack the counter anion L121E, so the positive charge would be forced to be delocalized along the polyene to increase its stability. We would thus expect to

observe a red-shifted wavelength. Finally the double mutant R132K:L121E was also studied.

The tetra mutant R132K::R111L::L121E::R59W was prepared to test the possibility of an excitonic coupling between the polyene and the tryptophan, and as previously discussed it has been proposed that excitonic coupling between Trp265 in rod rhodopsin and the polyene can contribute significantly to the observed opsin shift.

Also the CRABP_{II}-WT and the triple mutant R132L::R111L::L121E were studied and both proteins were used as controls to assure that the observed results were due to a PSB formed between the chromophores and Lys132 and not any other Lys present in the sequence.

6.3 Studies using 11-*cis*-retinal

The major sources of 11-*cis*-retinal (1), used in rhodopsin studies, come from light induced isomerization of all-*trans*-retinal (2).³¹⁻³⁶ However, its isomerization leads to a complex mixture of isomers, which are hard to separate, and 11-*cis*-retinal is only one of the minor isomers obtained. It is important to mention that even though isomerization of all-*trans*-retinal seems to be cheaper and faster, it was already attempted in this laboratory with out success. The chemical synthesis of 11-*cis*-retinal (1) has been previously accomplished. However due to the extreme tendency of the 11-*cis*- double bond to isomeraze to the more stable all-*trans*-retinal, its synthesis is very challenging (2).³⁷⁻³⁹ It is believed that the unusual difficulty in generating the 11-*cis* double bond is a result

of the steric hindrance between 10-H and the 13-CH₃. (Figure 6-7). A few elegant syntheses have been reported in which different strategies have been used as the key steps.^{40,41} The synthesis proposed by Nakanishi⁴¹ and co-workers appeared to be one of the most promising, due to the high final yield and the mild conditions used. The key step in the proposed synthesis involves the zinc-mediated semi-hydrogenation of the 11-yne-retinoid to the 11-*cis*-retinal.

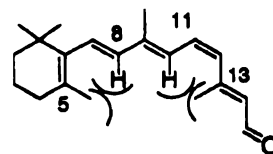
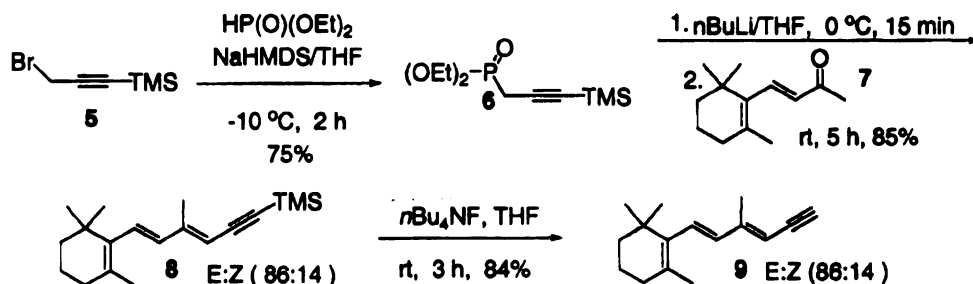


Figure 6-7. 11-*cis*-retinal is not a planar molecule due to steric interactions. The first steric interaction is between the C8-H and the C5-CH₃ and the second one between C10-H and the C13-CH₃. This leads to a twist of the C6-C7 and C12-C13 single bonds to alleviate the steric hindrance.

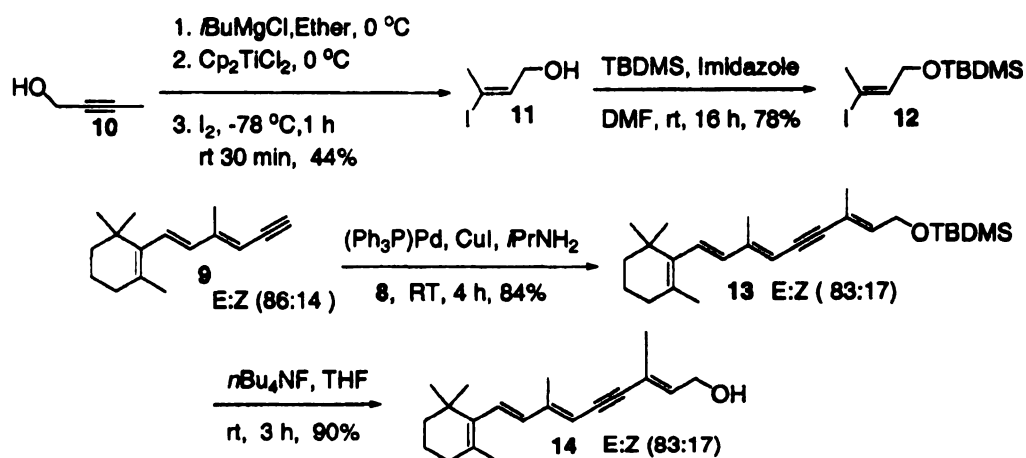
6.3.1 Synthesis using 11-*cis*-retinal

The synthesis of 11-*cis*-retinal (**1**) started by the preparation of ylide diethyl (3-trimethylsilyl-2-propynyl) phosphonate (**6**), (Scheme 6-1). The ylide was prepared from commercially available protected propargyl bromide **5** (75% yield). Subsequent HWE coupling⁴² of β -ionone (**7**) with (3-trimethylsilyl-2-propynyl)phosphonate (**6**) afforded the alkyne **8** in 85% yield, with an isomeric ratio of E:Z (86:14) at C9. Deprotection of the silyl group with tetrabutylammonium fluoride (TBAF) gave the terminal acetylene **9**, E:Z at a ratio of 86:14.



Scheme 6-1. Synthesis of acetylene (9) intermediate in the synthesis of 11-*cis*-retinal.

Hydromagnesiation⁴³ (Scheme 6-2), of prop-2-vinyl alcohol (10), followed by treatment with iodine at $-78\text{ }^{\circ}\text{C}$, afforded the vinyl iodide 11 in 44% yield after column purification as a single isomer (the column was performed in the dark). Protection of the alcohol using *tert*-butyldiethylsilyl chloride (TBDMSCl) afforded vinyl iodide 12 in 78% yield. Palladium coupling of 9 with



Scheme 6-2. Synthesis of acetylene (14) precursor of the Zn mediated reduction.

vinyl iodide **12** proceeded smoothly with complete stereochemical retention to yield retinoid **13** as a mixture of E:Z isomers (83:17) in 84% yield. Deprotection with TBAF afforded the allyl alcohol **14** in 90% as a mixture of E:Z isomers (83:17).

As previously mentioned, the acetylene is only 86% pure, it contains 14% of the isomer **17** (Figure 6-8). Therefore, to simplify the separation of isomers, after the Zn-mediated reduction, the acetylene was purified using HPLC (see Materials and Methods for detail (Figure 6-8)).

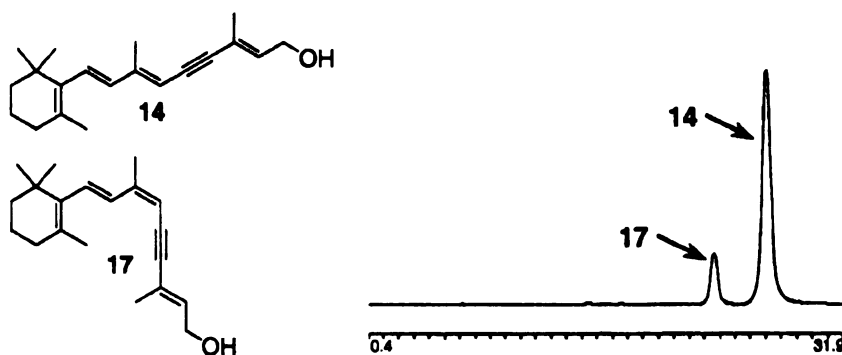
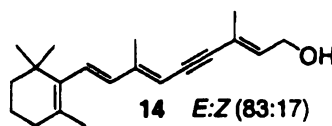

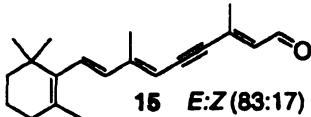
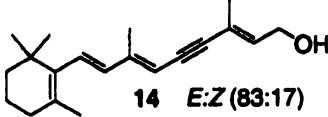
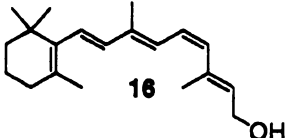
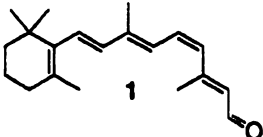


Figure 6-8. Separation of the isomers alkyne **14** and **17**. Using HPLC ZORBAX-SIL column mobile phase of 80:20 hexanes:ethyl acetate, 3 mL/min, at room temperature and a UV detection at 317 nm.

The semi-hydrogenation of acetylenes via zinc mediated in our case the reaction was more challenging than expected. The first attempt was performed following Feigel's protocol, and the MnO_2 oxidation was performed immediately because it is known that the 11-*cis*-retinol (**1**) is not very stable (Table 6-2).⁴⁴ Home made MnO_2 ⁴⁵ was prepared (from previous experiences, old MnO_2 is too

Table 6-2. Reduction of alkyne **14** to obtain 11-*cis*-retinal (**1**).

				
Trial	Conditions	Products		
1	<ul style="list-style-type: none"> •Zinc •Filtration for Zn work up, •Room temperature, 16 hours, •Isopropanol •MnO₂ oxidation. 	75%		
2	<ul style="list-style-type: none"> •Zn (dry stored) •Filtration for Zn work up, •Room temperature, 16 hours, •Methanol •No MnO₂ oxidation. 	85%		
3	<ul style="list-style-type: none"> •Purified 14 only E •Zn (dry stored) •Centrifugation for Zn work up, •Room temperature, 16 hours, •Methanol •No MnO₂ oxidation. 	~50% crude		
4	<ul style="list-style-type: none"> •Purified 14 only E •Zn (dry stored) •Centrifugation for Zn work up, •Room temperature, 16 hours, •Methanol •MnO₂ oxidation 	~10% after purification		

slow and the major product is the decomposition of retinoids). After oxidation the crude product was run through a small silica column, however, ¹H-NMR showed no reduction of the alkyne, only the expected oxidation of the alcohol to the aldehyde. The fact that reduction did not occur suggested that the Zn was not activated enough to perform the reduction.

This time only the reduction step was performed and the crude ^1H NMR (CDCl_3) showed starting material, some unidentified products, but no 11-*cis*-retinol **1** was observed. When zinc is activated the last step involves the filtration of the powder zinc. But every time the filtration was performed it was very slow and messy. And it has been reported that activated zinc deactivates in contact with air. The next time, the Zn was centrifuged instead of filtrated (see experimental for details) and MeOH was used as solvent. Gratifyingly the crude ^1H NMR (CDCl_3) showed the desired product **16**. When the oxidation using MnO_2 was performed it was not as clean as expected (several spots were formed before all starting material was consumed). The MnO_2 oxidation was performed at various temperatures, to minimize the formation of undesired products. But decreasing the temperature to $-78\text{ }^\circ\text{C}$, $-50\text{ }^\circ\text{C}$ and $-20\text{ }^\circ\text{C}$ did not show any improvement in the formation of pure 11-*cis*-retinal. It was also observed that decreasing the temperature seriously slowed down the reaction to a great extent. Therefore, oxidation at $0\text{ }^\circ\text{C}$ was adapted even though other products were formed. The crude was filtered through a celite pad and a micro column was performed to separate the polar products before HPLC purification. The ^1H NMR (CDCl_3) of the eluted fractions showed about 50% of the desired 11-*cis*-retinal and some isomers.

The HPLC isolation of 11-*cis*-retinal (**1**) has been previously reported using either normal phase or reverse phase. The first attempt to isolate 11-*cis*-retinal was made using a normal phase ZORBAX-SIL semi-prep column (4.6 mm ID x 25cm (5 μm), mobile phase 97:3 Hexanes:Ethyl acetate, 3 mL/min, at room

temperature and a UV detection at 362 nm. A small injection of the crude showed four major peaks in the HPLC but the ratios did not match with the ^1H NMR data. Isolation of each one of those peaks and analysis showed that none of those fractions was the 11-*cis*-retinal, suggesting that the retinal was isomerized in the column. To avoid the isomerization of the 11-*cis*-retinal catalyzed by the presence of acid, 0.05% of triethylamine was added to the HPLC solvent. When the reaction was repeated the separation seemed better but again no pure 11-*cis* retinal could be isolated.

When pure acetylene **14** was used in the reduction, and was followed by oxidation with freshly prepared MnO_2 , the 11-*cis*-retinal obtained was obtained in 80% yield. The 11-*cis*-retinal was the major isomer 94%. The obtained retinal was very pure, about 94% and two other different isomers, one 4% and other 2%. An attempt to further purify the retinal, via HPLC was performed. But

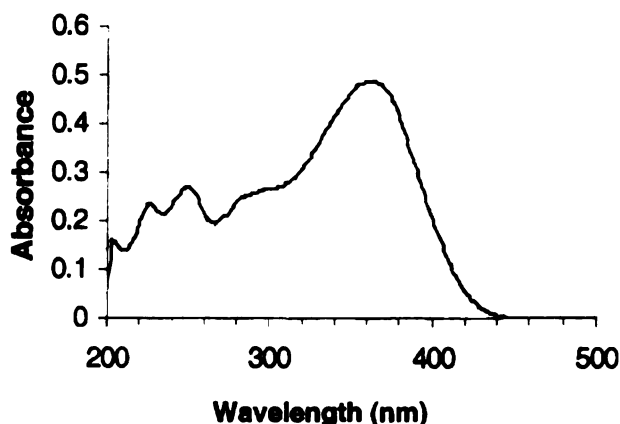


Figure 6-9. UV spectra of the isolated 11-*cis*-retinal, which matches with the reported $\lambda_{\text{max}}=362$.

unfortunately the isolated 11-*cis*-retinal was only 92% pure. No more attempts to purify the 11-*cis*-retinal were performed. Therefore, all the studies performed were done using 11-*cis*-retinal that was 94% pure. The obtained 11-*cis*-retinal matched the reported spectroscopic data (Figure 6-9).⁴¹

6.3.2 Protein substrate interactions using 11-*cis*-retinal

As discussed previously the binding pocket in CRABP II mutants is large, therefore it can be envisioned that 11-*cis*-retinal can accommodate into a variety of different configurations. The model of the covalently bound 11-*cis*-retinal in Figure 6-10A and Figure 6-10B suggests that the chromophore could exist in two different conformations. The conformation shown in Figure 6-7 has less steric interactions, therefore the following analysis will be based on this conformation. When all-*trans*-retinal binds the triple mutant R132K::R111L::L121E the Glu121 is close enough to form a strong H bonding (3.4 Å) away from the nitrogen of the iminium 11-*cis*-retinal, on the other hand the 11-*cis*-retinal model shows that Glu121 is further from the nitrogen of the iminium ion (4.2 Å). The increase in the distance of the counteranion (based in our models) is reflected in the wavelength observed. When the 11-*cis*-retinal was incubated with the rhodopsin surrogate R132K::R111L::L121E a broad absorption spectra was observed. The second derivative was used to deconvolute the individual peaks that were observed in the wavelength spectra (369 nm and another one at 464 nm). The first peak (464 nm) corresponds to the formation of the PSB and the second peak (369 nm) could be the SB or the free retinal. As shown in Figure 6-10 the counter anion is

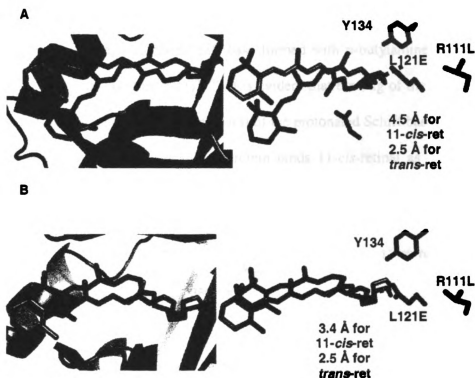


Figure 6-10. Using the crystal structure of CRABP II-R123K::R111L::L121E (blue) two possible conformations were found for the 11-*cis*-retinal. **A.** The counter anion Glu121 is 4.5 Å from the PSB and **B.** The counter anion Glu121 is 3.4 Å from the PSB.

further away, thus it could destabilize the protonation of the imine. Also, because of the different geometry it is possible that the nucleophilic Lys132 cannot access the aldehyde to form the imine. It is interesting to notice that the λ_{\max} observed is 15 nm red shifted as compared to the λ_{\max} of the all-*trans*-retinal. This could be partially explained due to the fact that the counter anion is further away, thus increasing the delocalization of the positive iminium. The red shifted peak

observed (λ_{max} of 464 nm) is blue shifted compared to the absorptions of the opsins (~500, 530 and 560).⁴⁶ But when this value is compared to the maximum wavelength of the Protonated Schiff base formed with *n*-butylamine, and 11-*cis*-retinal (~440 nm) (Figure 6-11)^{47,48} it is evident that binding of the ligand to the rhodopsin surrogate causes a 24 nm shift of the protonated Schiff base. This is the first example in which a designed protein binds 11-*cis*-retinal as a PSB. The double mutant

R132K::L121E

shows a similar behavior as the

triple mutant,

(Figure 6-12, 6-13). The pK_a

was estimated to

be 7.2, (Figure

6-14) which is

lower when

compared to the

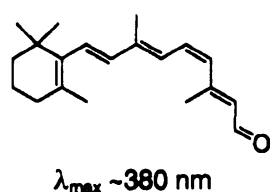
pK_a with all

trans retinal 8.0,

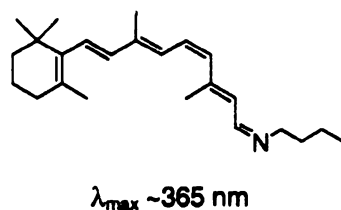
which could

explain the

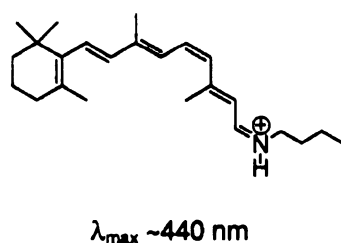
difference in the



11-*cis*-retinal



11-*cis*-retinal
Schiff base



11-*cis*-retinal
protonated Schiff base

Figure 6-11. UV-vis absorbances of 11-*cis*-retinal as a free aldehyde (380 nm), a *n*-butylamine Schiff base (365 nm) and a protonated Schiff base (440 nm), in ethanol as solvent.

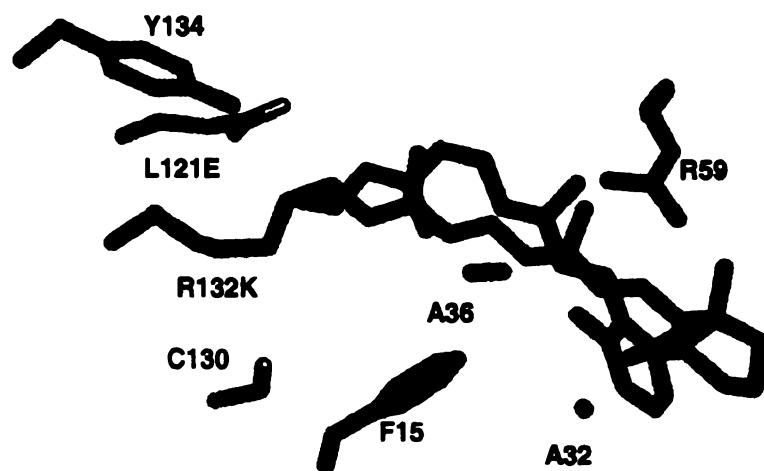


Figure 6-12. Model structures of the binding cavity of the triple mutant and 11-*cis*-retinal and the mutated amino acids are also shown.

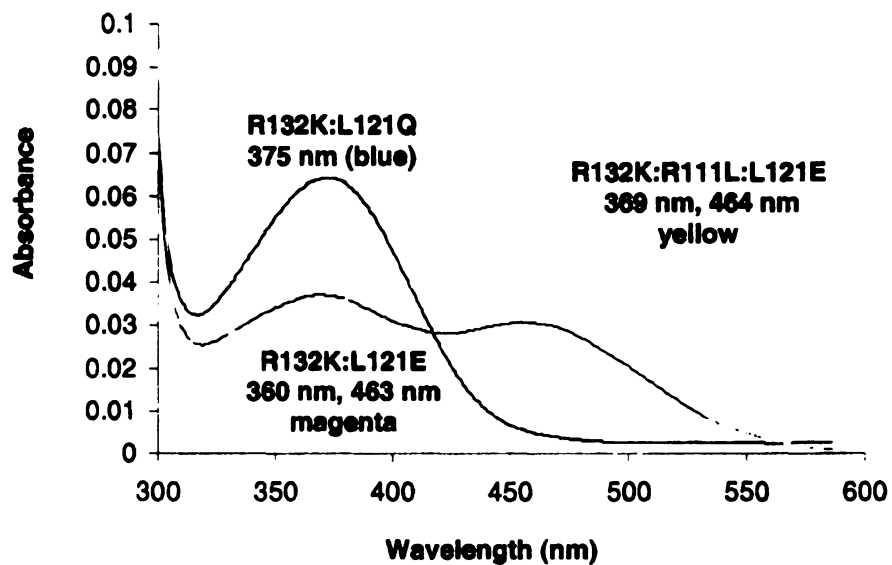


Figure 6-13. UV-vis absorbances of the ligand 11-*cis*-retinal with different rhodopsin surrogates.

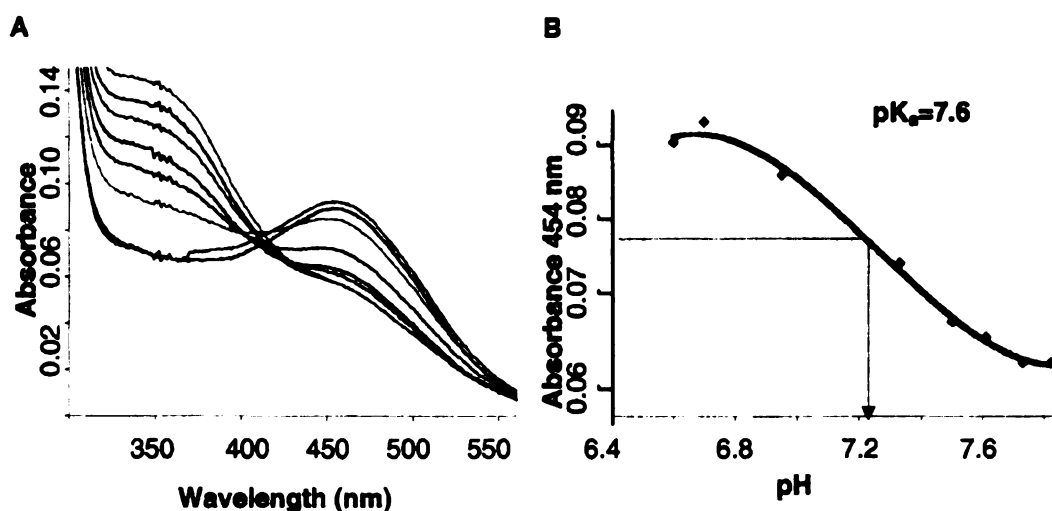


Figure 6-14. A. UV titration of R132K::L121E at different pHs. B. Graph of the absorption change at 454 nm at different pHs.

pK_a. Based on the model shown in figure 6-10 the counter anion seems to be farther in the 11-*cis*-retinal than the all-*trans*-retinal.

The absorbance of 11-*cis*-retinal using different proteins is summarized in Table 6-3. As expected the CRABPII-WT ($\lambda_{\text{max}}=387$ nm) and CRABPII-R132L::R111L::L121E ($\lambda_{\text{max}}=383$ nm) do not show any red shift. When the counter anion is placed in the opposite site (compared to L121E) the observed wavelength is blue shifted, which suggests that a counter anion at that position does not stabilize the protonation of the imine. When a second counter anion is present at position R59E (close to the polyene) the wavelength observed did not change significantly, but when the temperature of incubation was decreased, the PSB was more stabilized. However, interestingly when the counter anion is close to the polyene region (R132L::R111L::L121E::AS36E,

Table 6-3. Titration of different CRABP II mutants with 11-*cis*-retinal (1) and all-*trans*-retinal.

Protein	11-<i>cis</i>-Retinal (λ_{max}, nm)	all-<i>trans</i>-Retinal (λ_{max}, nm)
WT	387	377
R132K::R111L::L121E	369, 464	449
R132K::L121E	360, 463	457
R132K::L121Q	375	380
R132L::R111L::L121E	383	377
R132K::R111K::L121E	362, 468	365, 455
R132K::R111K::L121E::F15Y	382	373
R132K::R111K::L121E::A36E	374	394
R132K::R111K::C130D	383	377
R132K::R111K::L121E::C130D	395	440
R132K::R111K::L121E::R59E	367, 460	448

R132L::R111L::L121E::C130D) it results in the destabilization of the PSB and the observed peak corresponds to either the SB or binding of retinal.

6.4 Retinal analogs

Since CRABP II contains a large binding pocket (~600 Å) we expect that a variety of different aldehydic chromophores could fit and form a PSB or SB. Distefano and co-workers have performed studies on a related binding protein, Adipocyte Lipid Binding Protein (ALBP).⁴⁹ ALBP has a β -barrel topology with a remarkable structural similarity to CRABP II. Distefano has managed to modify

the pocket in order to incorporate a series of cofactors and organic molecules, which are poised to perform chemistry on a variety of molecules that enter the binding pocket.⁵⁰⁻⁵² As previously discussed we are especially interested in the study of protein-ligand interactions. The study of the interactions between rhodopsin surrogates and different chromophores that potentially could bind as PSB would help us to understand those interactions in more depth.

Also it is possible that the rhodopsin surrogates could be ideally suited as colorimetric protein fusion tags similarly to Green Fluorescent protein. Green Fluorescent Protein (GFP) is a protein isolated from jellyfish *Aequorea victoria*,⁵³ and it shows fluorescence a couple of hours after it is transcribed. The inherent fluorescence of GFP has been used as a tool in molecular and cellular biology. In more detail, a fusion protein of GFP and the protein of interest can be formed. The fusion protein would be fluorescent, and therefore, the fluorescence can be monitored in real-time. The use of GFP has led to a better understanding of cellular biology and processes such as signal transduction, apoptosis, RNA localization, monitoring ion channels among others.⁵⁴⁻⁵⁸ Even though GFP has been a great tool in the understanding of molecular and cell biology, they do have a few limitations. The first one is that the formation of the GFP chromophore is slow (up to 2 hours) and this could be problematic for the proteins with short half-lives.^{59,60} Also, molecular oxygen is required for the formation of the fluorescent chromophore, and the by-product formed is H₂O₂ that could have significant harmful effects in the system that produces the GFP tag.⁵³ Also some variants form homo-oligomers and inclusion bodies.⁶¹ A variety of GFPs that show

different excitation and emission spectra have been produced, in general green, cyan and yellow pigments. The blue GFP variant photobleaches very rapidly and it is highly unstable.^{58,60,62-64} The rhodopsin mimics prepared are smaller in size, (136 aa vs. 240 aa for GFP) and thus its use as a protein tag would impose less of a change to the overall size of the protein fusion tag.

As mentioned previously we thought that the engineered proteins could be used successfully as spectroscopically active fusion protein tags, that may offer an alternative to the use of GFP. Especially since their wavelengths can be tunable and they could be switchable. The use of different chromophores as ligands could change the color of the engineer proteins. Furthermore, binding of a fluorophore could produce fluorescent tags if necessary. Also we could tune the color observed using the same chromophore and different proteins. The proteins would regulate the color based on the principles observed in wavelength regulation. In particular we were interested in two groups of chromophores, azulene and merocyanine derivatives, both which show a large red-shifted absorption relative to the retinal.

The synthesis, design, and reactivity of cyanines dyes have been intensively studied in the last decades.^{48,65,66} The interest is mostly due to their wide applications. For example they are important materials for non-linear optics (NLO), laser technology, data storage, etc.⁶⁵ Depending on the charge of the polymethine unit, cyanine dyes can be classified into three main groups: cationic (cyanine and hemicyanine) dyes, anionic (oxonol) dyes and neutral (merocyanine) dyes. We were interested in the synthesis of merocyanines that contained an

indolylidene end group for two main reasons: the first was that the UV spectra of the dye would red shift upon addition to the protein. Second, if the PSB is formed with the rhodopsin surrogate the positive charge would be strongly delocalized (Figure 6-15).

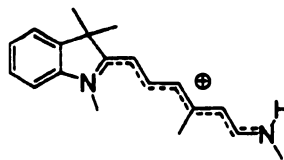


Figure 6-15. Merocyanines forms a very stable PSB due to the fact that the positive charge is stabilized all around the polyene.

Azulene has intrigued scientists since its discovery. Azulene (**18**) has a very beautiful blue color. The unusual color of azulene is attributed to its very low first excited state S_1 .⁶⁷⁻⁶⁹ Replacement of the cyclohexene ring in retinal by an azulene group results in a molecule with a red shifted wavelength. An interesting feature of azulene is that it can stabilize cations as well as anions, due to its remarkable polarizability (Figure 6-16). When an azulenenic aldehyde derivative forms a PSB the cation formed can be in resonance, the positive charge can be stabilized either at the iminium ion **21**, at the tropylium ion **19** or all along the molecule **20** (Figure 6-17). This could be helpful in the study of wavelength regulation in our rhodopsin mimics. Similar azulenes were used in the study of bacteriorhodopsin. After the retinal analogs were incubated in

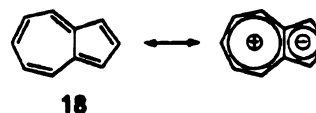


Figure 6-16. Azulene (**18**) has a remarkable polarizability. Therefore it can stabilize cations as well as anions.

the presence of the protein, they formed PSB that had maximal absorption in the in near infrared region (NIR).⁷⁰

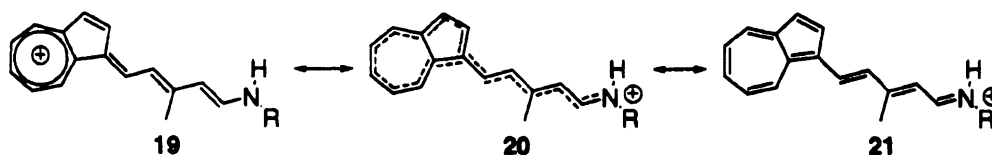
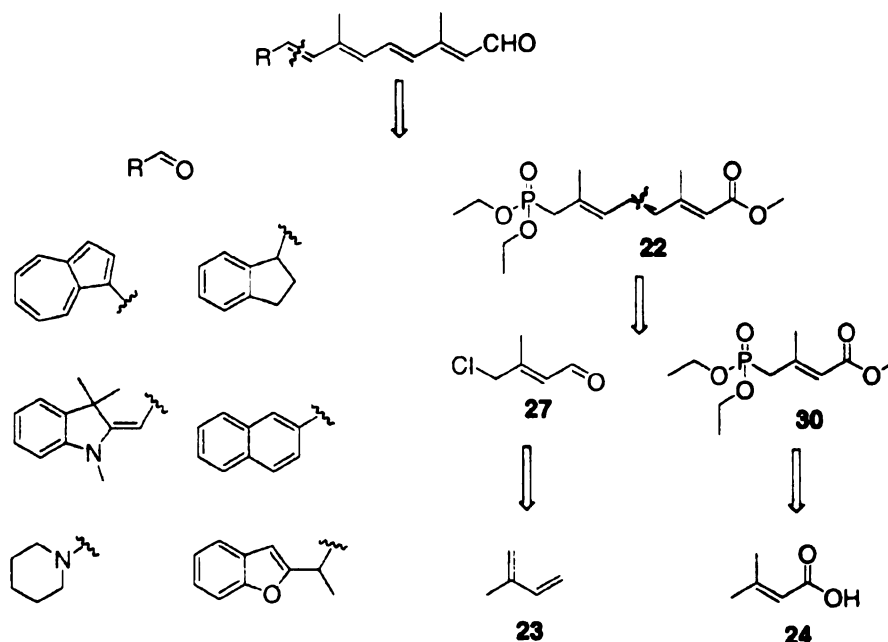


Figure 6-17. Azulenes can stabilize the positive charge either at the iminium ion 21, at the tropylium ion 19 or in resonance all along the molecule 20.

6.4.1 Synthesis of different chromophores

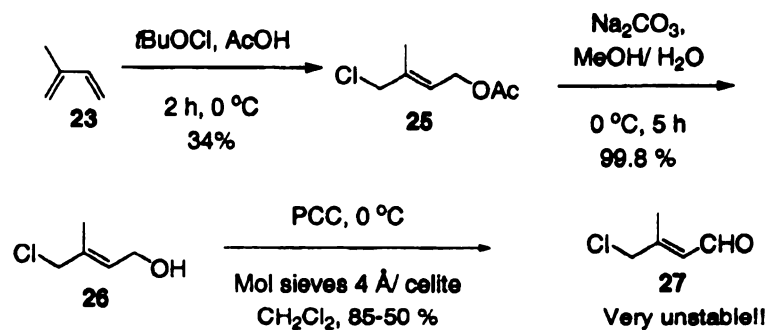
The initial approach for the synthesis of retinoid mimics consisted of the synthesis of a large ylide (**22**) that resembles the retinal polyene structure. This ylide could be coupled with a variety of different aldehydes that have different electronic and spectroscopic properties. (Scheme 6-3). The ylide would be prepared from the available isoprene (**23**) and dimethyl butyric acid (**24**).

The synthesis of **22** started by the preparation of aldehyde **27** and a ylide **30**. The synthesis aldehyde **27** started with the preparation of (*E*)-4-chloro-3-methyl-2-buten-1-ol acetate **25** (Scheme 6-4). The chloro-acetate **25** was prepared by the 1,4 acetoxychlorination of isoprene (**23**) in presence of *t*-butyl hypochlorite and acetic acid in a 34% yield. Hydrolysis under basic conditions at 0 °C afforded 99.8% of the chloro alcohol **26**. Oxidation of the chloro alcohol was achieved in 85% yield. In order for the reaction to give a good yield dry pyridinium



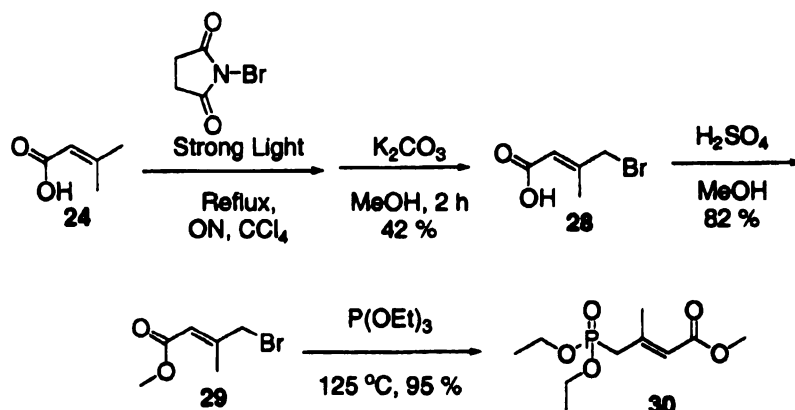
Scheme 6-3. Retrosynthesis of ylide **22**.

chlorochromate (PCC) must be used as well as molecular sieves and celite in dichloromethane (Scheme 6-4) The resulting aldehyde **27** is very unstable, and it cannot be stored for more than a couple of days.



Scheme 6-4. Synthesis of aldehyde **27**.

The ylide **30** was prepared first by the bromination of 3,3 dimethyl acrylic acid (**24**) with N-bromo succinimide (Scheme 6-5). The $^1\text{H-NMR}$ showed a mixture of *E* and *Z* isomers. The residue was picked up in MeOH and excess K_2CO_3 anhydrous was added to promote the cyclization of the *Z* isomer. The pure (*E*)-4-bromo-3-methyl-2-butene acid (**28**) was obtained in a 42% overall yield from **24**. Esterification of the acid **28** afforded 82% of the (*E*)-4-bromo-3-methyl-2-butene methyl ester **29**. An Arbuzov reaction afforded the ylide **30** in a 95% yield (Scheme 6-5).



Scheme 6-5. Synthesis of ylide **30** starting from 3,3, dimethyl acrylic acid (**24**).

Coupling of the ylide **30** and the aldehyde **27** using sodium methoxide as a base did not provide any desired product. When *n*-butyl lithium was used as base instead only a 15% of the desired chloro ester was obtained. On the other hand, when NaHMDS was used as a base, an average 40% yield was obtained as a mixture of *E*:*Z* isomers (2:1) (Table 6-4). When **32** was used to prepare ylide **33**

Table 6-4. HWE between ylide **30** and aldehyde **27** to obtained chloro ester **31**.

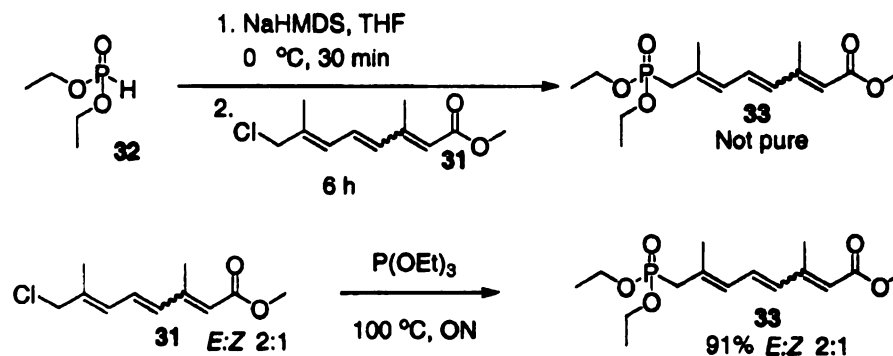


Trials	Conditions	Product yield
1	MeOH, NaOMe, 0 °C 30 min 24 hours then aldehyde 27	0 %
2	<i>n</i> -BuLi, 0 °C 30 min 6 hours then aldehyde 27	14.5 % (<i>E:Z</i>) (1.5:2)
3	NaHMDS, THF, 0 °C 30 min 4 hours. then aldehyde 27	30-45 % (<i>E:Z</i>) (2:1)

the expected ylide was obtained in a 40% yield, but the product was not pure and it could not be isolated. On the other hand, when the ylide **33** was prepared via an Arbuzov reaction the desired product was obtained in a 91% yield (Scheme 6-6).

A few different aldehydes were prepared to be coupled with the ylide. Azulene (**18**) was formylated using DMF and POCl₃ to afford **34** in 76%. The porphyrin **36** was prepared via oxidation of the alcohol **35** with pyridinium chlorochromate (PCC) in 20% yield. Indole **37** was formylated using DMF and (CO)₂Cl₂ to afford **38** in about 45% (Scheme 6-7).

The HWE condensation using ylide **33** and NaHMDS as a base provided very low or no yield of the expected ester product. Different temperatures and

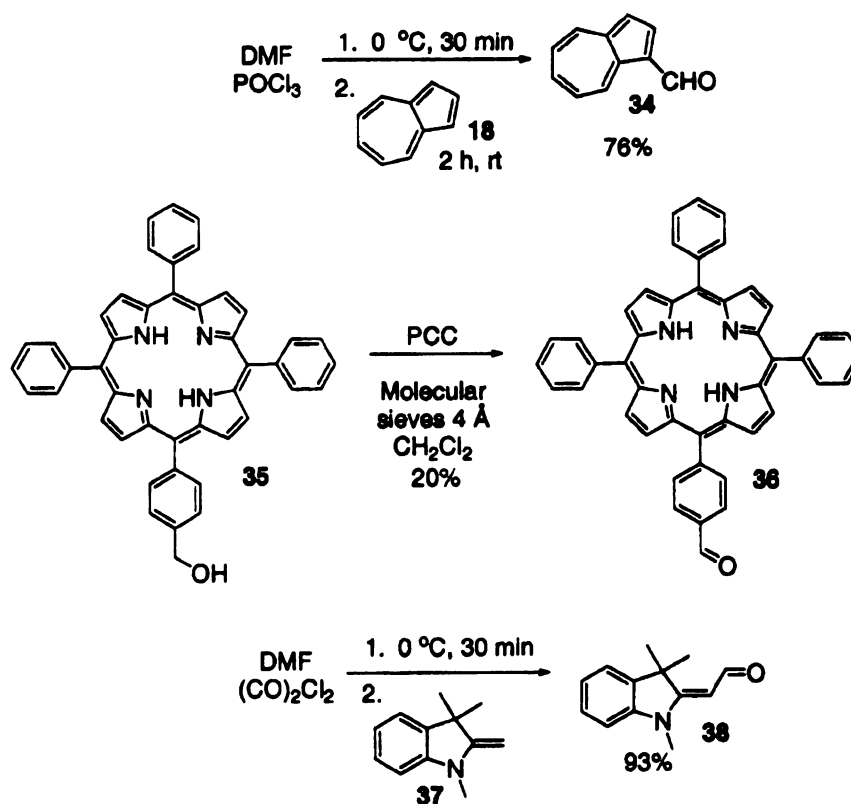


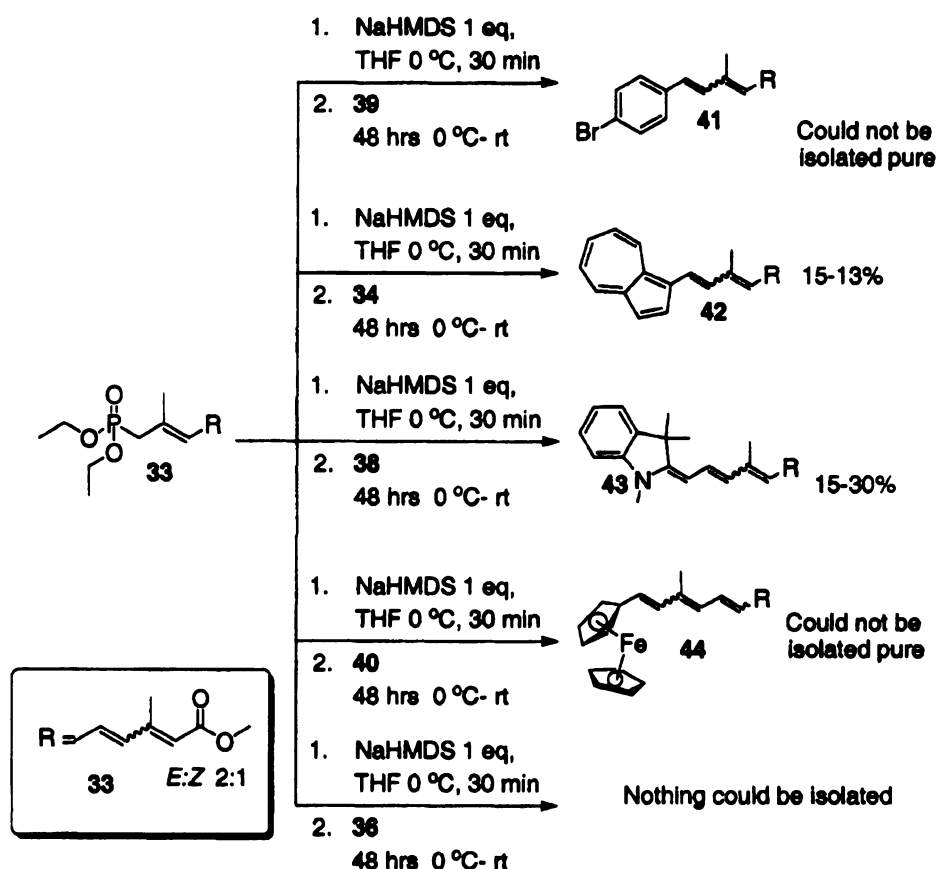
Scheme 6-6. Synthesis of ylide **33**.

bases were tried but no improvement was observed. When porphyrin **36** and **40** were used the crude $^1\text{H-NMR}$ showed some product, but nothing could be isolated. On the other hand when azulene **34** and merocyanine **38** were used **30** and 40% yields were obtained, respectively.

Unfortunately, the yield of the HWE reaction was not very consistent (Scheme 6-8, Figure 6-18). When the product ester was reduced with DIBAL no desired product could be recovered. Again, different temperatures and equivalents of the reducing agent were utilized but no improvement was observed.

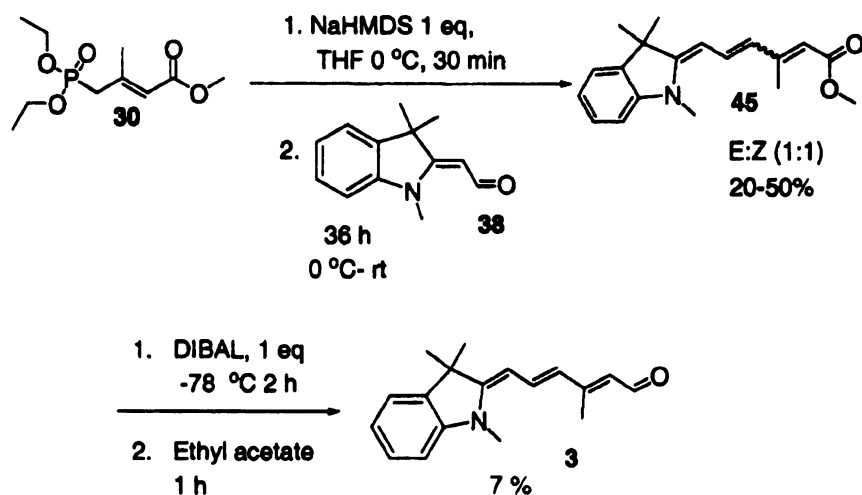
Because the use of ylide **33** did not provided consistent results a different approach for the synthesis of the chromophores was taken. We started with the synthesis of the merocyanine **4** (Scheme 6-9). This new methodology utilized a smaller ylide (**30**) instead. The synthesis of **45** via HWE using potassium *t*-butoxide and KHMDS as bases afforded no product. When NaHMDS was used instead about 36% yield of **45** was obtained. If the amount of ylide was increased to 10 equivalents the yield of **45** was increased to 50%.





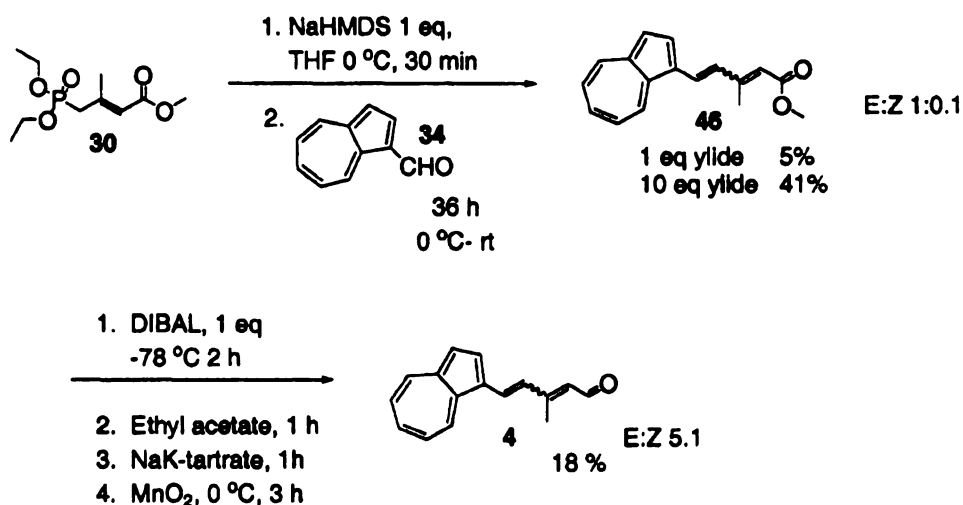
Scheme 6-8. HWE using ylide **33** and various aldehydes.

When two equivalents of DIBAL were used at $-78\text{ }^{\circ}\text{C}$ to reduce the ester to the alcohol followed by oxidation with MnO_2 a mixture of unidentified products was obtained. When only one equivalent of DIBAL was used at $-78\text{ }^{\circ}\text{C}$ and the reaction was quenched with ethyl acetate, a 7% yield of the desired aldehyde **3** was obtained (Scheme 6-9). Even though the yield was low it gave us enough product to test the binding of the cyanine with the different proteins.



Scheme 6-9. Synthesis of merocyanine **3**, an HWE condensation between ylide **30** and aldehyde **38** followed by DIBAL reduction.

An HWE reaction of **34** with ylide **35** afforded the ester **46** in a 5% yield. But when 10 equivalent of ylide was used instead, the yield was enhanced to 41%. Reduction was accomplished using 2 equivalents of DIBAL at $-78\text{ }^{\circ}\text{C}$, and quenched with ethyl acetate. The quenched reaction was stirred with Na-K tartrate for one hour. Then the crude was treated with fresh MnO_2 , and after purification using column chromatography only 18% of the desired aldehyde **4** was isolated (Scheme 6-10). Even though the yield was low it provided enough material to perform a study of this chromophore with different proteins. If the reaction is not quenched with EtOAc or stirred with NaK tartrate no product can be isolated.



Scheme 6-10. Synthesis of azulene **4**, an HWE condensation between ylide **30** and aldehyde **34** followed by DIBAL reduction.

6.4.2 Binding studies with merocyanine **3** and azulene **4**

As previously mentioned, binding of bacteriorhodopsin and merocyanine **3** has been studied. Hans-Dieter observed that merocyanine **3** bound bRh as a PSB (λ_{max} of 610 nm) with an opsin shift of 48 nm (Table 6-5) (λ_{max} of the PSB = 578 nm).⁷¹ When merocyanine **3** was incubated with the CRABPII R132L::R111L::L121E a single peak was observed with a λ_{max} of 587 nm Figure 6-18A show comparison of the all-*trans*-retinal bond structure with the modeled structure of merocyanine **3**. It was assumed that when the merocyanine forms a PSB it binds similar to the retinal. Interestingly merocyanine **3** shows a different behavior than all-*trans*-retinal. The major difference is that when all-*trans*-retinal binds the triple mutant CRABPII, binding and formation of the PSB occurs very

A



B

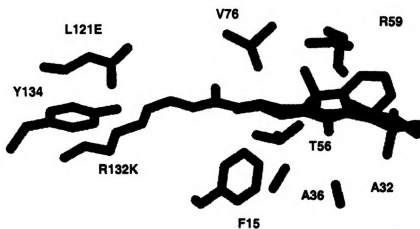


Figure 6-18. **A.** Fill in model of binding of merocyanine 3 (red) to the triple mutant CRABP II-R132K::R111L:L121E. The ionone ring is buried in the cavity. It was assumed that the merocyanine would bind in a similar manner to all-*trans*-retinal. **B** This model represents a comparison of the binding of all-*trans*-retinal (blue) and merocyanine (red) the binding site show important amino acids in the binding pocket.

Table 6-5. Spectroscopic data of merocyanine 3 and the bacteriorhodopsin (bRh).

(λ _{max} , nm)			
Aldehyde	SB	PSB	bRh
463	422	578	610

rapidly, it usually occurs in about 5 minutes. On the other hand it was observed that the merocyanine forms a PSB more slowly. When 0.2 eq. of merocyanine were added at time 0 no shift in the wavelength was observed (λ_{max}=480 nm). After 5 minutes a small peak was observed (λ_{max}=587 nm) and with time the blue shifted peak disappeared and the red shifted peak increased. The transition between one peak to the other showed a clear isosbestic point indicating the transition of one species to the another. After one hour complete transition was observed (Figure 6-20). The formation of the PSB is much slower than that of retinal. As shown in figure 6-19 all-*trans*-retinal and merocyanine 3 have a similar 3D shape, thus we could expect that they would have similar binding properties (for retinal K_d=1.36±4.9 nM CRABP II triple mutant R132K: R111L:L121E). However so far we cannot prove this because the method used to determine the dissociation constant for retinal cannot be used to determine the dissociation constant of merocyanine 3. In brief, the binding of retinal to the hydrophobic pocket can be followed by fluorescence quenching.^{72,73} The Trp residues of the protein absorb at 280 nm and fluoresce at ~340 nm. When the

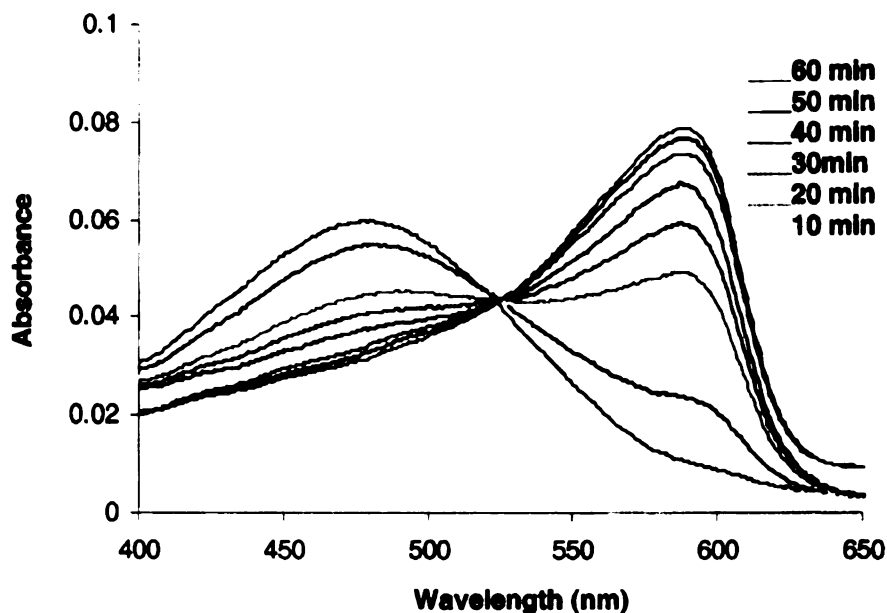


Figure 6-20. UV-vis spectrum of the binding of merocyanine **3** and CRABPII-R132K::R111L::L121E. Complete shift of the wavelength from $\lambda_{\text{max}} = 480$ to $\lambda_{\text{max}} = 587$ nm takes about 60 minutes.

chromophore is within the CRABPII cavity, in close proximity to the Trp residues, it can absorb the light emitted by the protein. The free chromophore **3** absorbs at ~ 465 nm in an ethanolic solution, therefore quenching of fluorescence will not occur. Thus it is believed that the rate limiting step is the formation of the Schiff base, such decrease in the rate of SB formation could be due to the fact that the aldehydic carbonyl of merocyanine **3** is less electrophilic when compared to that of retinal, or it could be due to steric factors. We yet have to design a method that would allow us to measure the dissociation constant accurately.

The maximal absorbance observed is red shifted by 9 nm when compared to the PSB with butylamine in ethanol ($\lambda_{\text{max}}=578$ nm), however it is not as red shifted as the maximal absorption observed when merocyanine binds to bRh. Upon incubation of chromophore 3 with CRABPII-WT and CRABPII-R132L::R111L::L121E (Lys control mutant) a small peak at 580 nm appeared after 6 h. Appearance of the signal at 580 nm was probably due to the formation of non specific PSB (a different Lys in the CRABPII sequence could have reacted) (Figure 6-21).

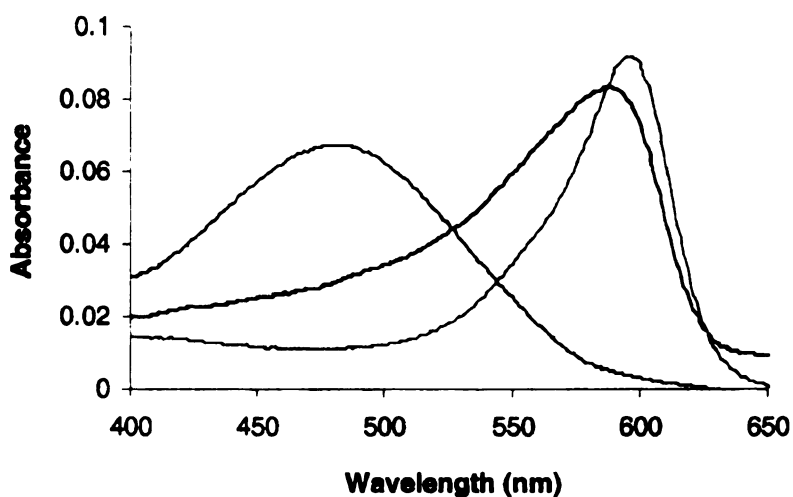


Figure 6-21. Comparison of the UV-vis spectrum obtained after incubation of merocyanine 3 and CRABPII-R132K::R111L::L121E $\lambda_{\text{max}}=587$ (blue), CRABPII-R132K::R111L::L121E::R59W $\lambda_{\text{max}}=599$ (green) and WT-CRABPII (magenta) $\lambda_{\text{max}} = 480$.

Table 6-6 summarizes the results observed between the titration of the rhodopsin surrogates and merocyanine 3. As expected the double mutant R132K::R111L shows red-shift wavelength by 12 nm (λ_{max} =599 nm) as compared to the triple mutant CRABPII-R132K::R111L::L121E, the lack of the counter anion forces the positive charged formed to be stabilized all along the polyene. However this protein takes much longer to reach equilibrium (6 hours). The triple mutant CRABPII-R132K::R111L::L121Q shows similar results, a red shift of 12 nm (λ_{max} =597 nm) as compared to the triple mutant CRABPII-R132K::R111L::L121E and also it takes a long time (10 hours) to stabilize(Figure 6-22).

Table 6-6. Titration of different CRABPII mutants with merocyanine 3 and all-*trans*-retinal.

Protein	Merocyanine (λ_{max}, nm)	all-<i>trans</i>-Retinal (λ_{max}, nm)
WT	480	377
R132K	597 (small)	375
R132K::R111L	599 (6 hours)	377
R132K:: R111L::L121E	587 (1 hour)	449
R132L::R111L::L121E	479	377
R132K::R111L::L121E::R59E	599 (1 hour)	447
R132K::R111K::L121E::R59W	596 (40 minutes)	443
R132K::R111K::L121E::A36E	590 (4 hours)	382
R132K::R111K::L121Q	597 (10 hours)	379
R132K::L121E	591 (2 hours)	457

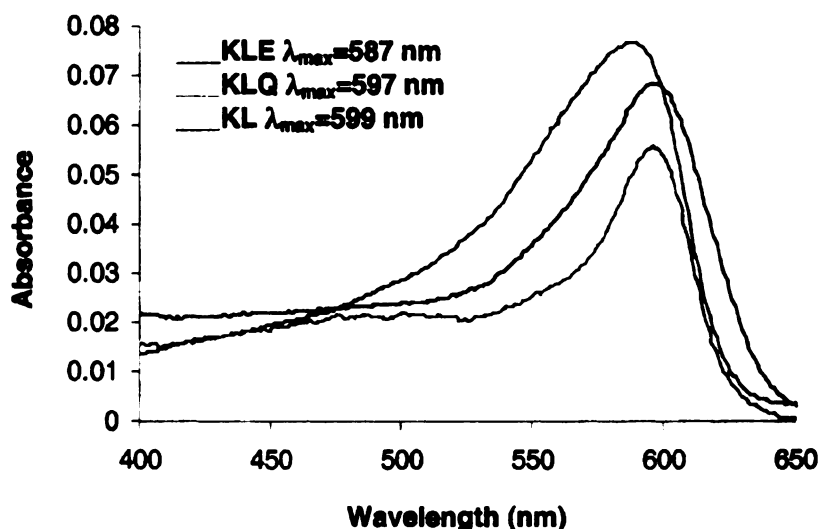


Figure 6-22. Comparison of the UV-vis spectra obtained after incubation of merocyanine 3 and CRABPII-R132K::R111L::L121E $\lambda_{\text{max}}=587$ (magenta), CRABPII-R132K::R111L::L121Q $\lambda_{\text{max}}=597$ (green) and CRABPII-R132K::R111L (blue) $\lambda_{\text{max}} = 599$.

When a second counter anion is placed in the polyene the wavelength changed as expected, the closer the second counter anion to the Lysine nitrogen (iminium) the more blue shift is observed. Arg59 is about 6.2 Å from the C-9 and the maximal wavelength observed is 599 nm, and Ala36 is about 4.0 Å from C-10 and the maximal wavelength blue shift is observed. Arg59 is about 6.2 Å from the C-9 and the maximal wavelength observed is 599 nm, and Ala36 is about 4.0 Å from C-10 and the maximal wavelength observed is 590 nm. When compared with the triple mutant CRABPII-R132K::R111L::L121E, a 12 and 3 nm

wavelength red shift is observed for CRABP II -R132K::R111L::L121E::R59E and CRABP II -R132K::R111L::L121E::A36E correspondingly.

Also as discussed previously it has been suggested that a different mode of regulation is excitonic coupling. The results observed when the tetra mutant CRABP II -R132K::R111L::L121E::R59W was incubated with merocyanine 3 support this proposed theory, a red shift of 9 nm is observed when a Trp is close to the polyene region.

Interestingly the presence of a Trp59 also increases the rate at which the PSB is formed between the rhodopsin surrogate merocyanine 3 (Figure 6-23).

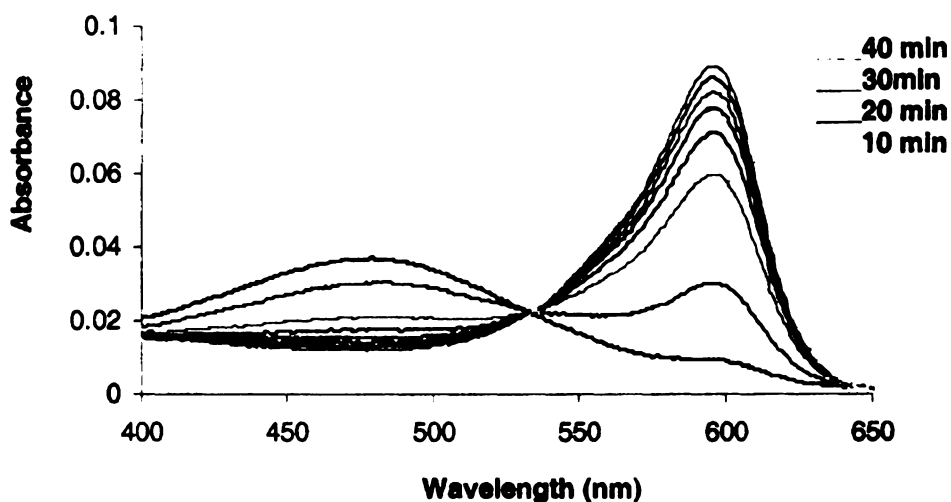


Figure 6-23. UV-vis spectrum of the titration of merocyanine 3 and CRABP II -R132K::R111L::L121E::R59W. Complete shift of the wavelength from $\lambda_{\text{max}}=480$ to $\lambda_{\text{max}}=599$ nm takes about 40 minutes.

A



B

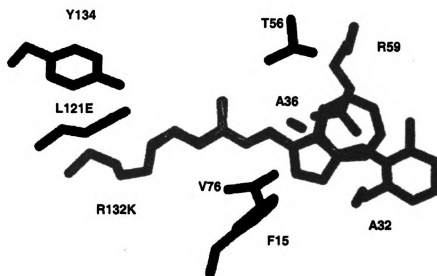


Figure 6-24. **A.** Fill in model of binding of azulene 4 (red) to the triple mutant CRABP II-R132K::R111L:L121E. The ionone ring is buried in the cavity. It was assumed that the azulene would bind in a similar manner to all-*trans*-retinal. **B** This model represents a comparison of the binding of all-*trans*-retinal (blue) and azulene (red) the binding site show important amino acids in the binding pocket.

Figure 6-24 compares the all-*trans*-retinal model with those of azulene 4. Chromophore 4 is smaller than the retinal. Incubation of CRABPII-K132E::R111L::L121E with azulene 4 reveals a wide wavelength absorption, which consists of by two different peaks. When the peaks were deconvoluted one peak had a maximal absorption of $\lambda_{\text{max}}=426$ nm and the second peak had a maximal absorption of $\lambda_{\text{max}}=527$ nm while the ratio between those peaks was (2:1) nm. No red shifting was observed when the chromophore 4 was incubated with the CRABPII-WT ($\lambda_{\text{max}}=440$ nm) and CRABPII-R132L::R111L::L121E ($\lambda_{\text{max}}=439$ nm) which suggests that the peak observed at 527 nm is due to the formation of the PSB with Lys132 (Figure 6-25). The double mutant CRABPII-

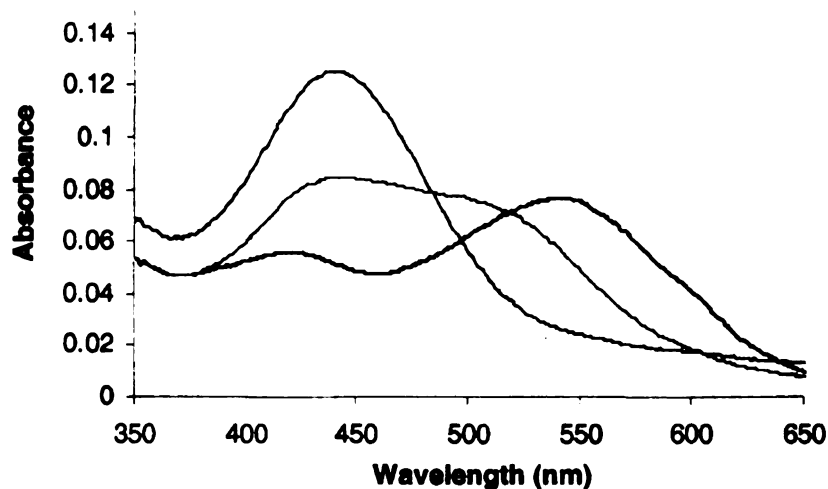


Figure 6-25. Comparison of the UV-vis spectrum obtained after incubation of azulene 4 and CRABPII- R132K::R111L::L121E $\lambda_{\text{max}}=434$, 511 (green), CRABPII- R132K::R111L: $\lambda_{\text{max}}=418$, 540 (blue) and WT-CRABPII (magenta) $\lambda_{\text{max}}=440$.

R132K::L121E shows very similar results to the triple mutant. Whether the second peak ($\lambda_{\text{max}}=426$ nm) corresponds to free chromophore or the SB is not clear. But when the pH is decreased the $\lambda_{\text{max}}=453$ nm and when the pH is basic the $\lambda_{\text{max}}=415$ predominates (Figure 6-26). This results suggest that a there is a formation of a SB and there is an equilibrium between the PSB and SB. Incubation of CRABPII-K132E::R111L::L121Q with azulene 4 showed an unexpected result. Two peaks were observed, one with a maximal absorbance of $\lambda_{\text{max}}=421$ nm and $\lambda_{\text{max}}=538$ nm and a ratio of 2:3, which suggests that the PSB is more stable than when the triple mutant that has the counter anion. The double mutant CRABPII-R132K::R111L also shows two peaks one at $\lambda_{\text{max}}=418$ nm and other at $\lambda_{\text{max}}=540$ nm with a ration of 1:2 respectively (Figure 6-25). The tetra

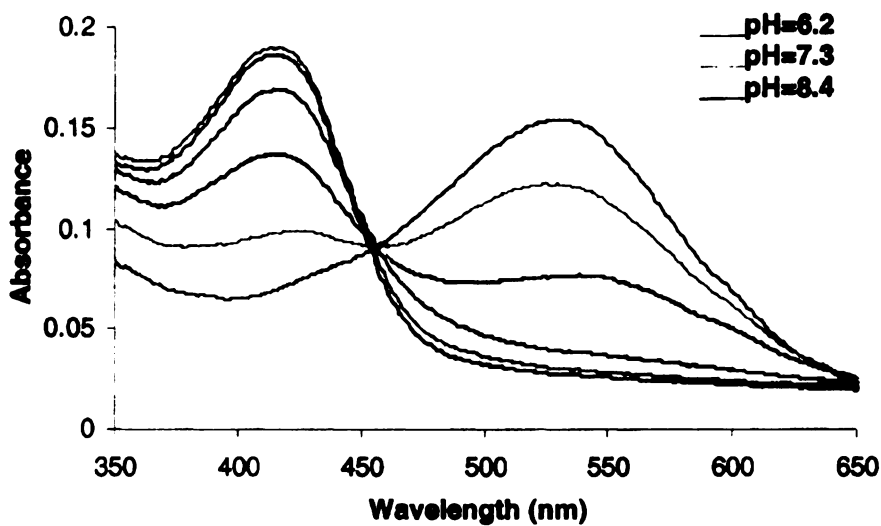


Figure 6-26. UV-vis spectrum of the titration of azulene 4 and CRABPII-R132K::R121E. At basic pH the $\lambda_{\text{max}} = 415$ species predominates, presumably the SB. And acidic pH the $\lambda_{\text{max}} = 534$ nm predominates, presumably the PSB.

mutant CRABPII-R132L::R111L::L121E::R59W shows very similar results, two peaks with a 1:2 ratio ($\lambda_{\max}=439$ nm $\lambda_{\max}=528$ nm). These observations are unexpected because it seems that the presence of a Glu121 does not promote the formation of a PSB. Table 6-7 summarizes all the different proteins treated with azulene 4.

Table 6-7. Titration of different CRABPII mutants with azulene aldehyde 4 and all-*trans*-retinal.

Protein	Azulene 4 (λ_{\max} , nm)	Azulene 4 (λ_{\max} , nm) 2 nd derivative	all- <i>trans</i> - Retinal (λ_{\max} , nm)
WT	440	438	377
R132K::R111L	418, 540 (1:2)	421, 547	377
R132K:: R111L::L121E	434, 511 (2:1)	426, 527	449
R132L::R111L::L121E	439	439	377
R132K::R111K::L121E::R59W	431, 528 (1:2)	424, 537	443
R132K::R111K::L121E::A36E	No binding	401	382
R132K::R111K::L121Q	421, 534 (2:3)	421,534	379
R132K::L121E	429, 522	429,522	457

In conclusion, different chromophores (3, 4) can form a PSB with the rhodopsin surrogates developed. It is clear that because the protein surrogates were developed to bind all-*trans*-retinal, when 11-*cis*-retinal is incubated, the formation of the PSB is not as optimal as with all-*trans*-retinal. However this is

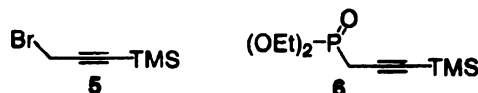
the first example in which an artificial protein binds 11-*cis*-retinal as a PSB. The binding of different chromophores was confirmed by incubation of the rhodopsin surrogates with **3** and **4**. Both chromophores form PSB with the designed proteins. Merocyanine **3** is particularly interesting because it provides a system to be used to study wavelength regulation. Also when it is incubated with CRABPII-R132K::R111L::L121E::R59W a strongly red shifted chromophore is obtained (λ_{max} =599 nm) compared to the triple mutant CRABPII-R132K::R111L::L121E.

6.5 Materials and Methods.

A. Synthesis of Compounds

All reactions were carried under an atmosphere of nitrogen and removal of solvents was performed under reduced pressure with a Buchi rotatory evaporator. THF and Et₂O were freshly distilled from sodium/benzophenone, and CH₂Cl₂ was distilled over CaH₂ under a nitrogen atmosphere. Radioactive NaB³H₄ was purchased from American Radiolabeled Chemicals, Inc (St. Louis, MO). Analytical TLC was carried out using Merck 250 mm Silica gel 60 F₂₅₄ and spots were visualized under UV light. Column chromatography was conducted using Silicycle silica gel (230-400 mesh). 300 MHz ¹H-NMR and 75 MHz ¹³C-NMR spectra were recorded on a Varian Gemini-300 or 500 instruments, and the residual protic solvent (CDCl₃ or DMSO-d₆) was used as internal reference. UV-visible spectra were recorded on a Perkin-Elmer Lambda 40 spectrometer.

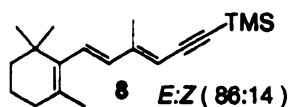
1. Synthesis of (6) diethyl(3-trimethylsilyl-2-propynyl)phosphonate



To a stirred solution of NaHMDS in THF (1 M, 10 mL, 10 mmol) at -10 °C is added a solution of diethyl phosphonate in THF (3 mL). The solution was stirred for 15 min at -10 °C, and a solution of 3 (1.7 g 0.9 mmol) in THF (3 mL) was added at -10 °C. The reaction mixture was stirred for 2.5 h when the TLC showed that the reaction was complete. The solution was diluted with water (6 mL) and

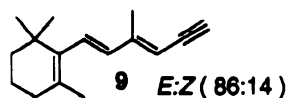
extracted with ether (2x). The organic layer was dried with Na₂SO₄. The crude product was pure enough to continue to the next reaction. [¹H NMR (300 MHz CDCl₃): δ 4.14(4H, q, J=7 Hz), δ 2.75(2H, d, J=22.2 Hz, CH₂P), δ 1.30(6H, t, J=7.0 Hz), δ 0.09(s, 9H, 3 x SiCH₃). ¹³C NMR (300 MHz CDCl₃): δ 95.59, 87.81, 63.00, 19.22(d, J=144.3 Hz, CH₂P), 16.31, -0.27.].

2. Synthesis of (8)



The ylide **6** was dissolved in anhydrous THF (1.6 g, 6.66 mmol in 30 mL). The solution was chilled to 0 °C under nitrogen, in the dark room, and it was treated with *n*-butyl lithium (1.6 M in hexanes, 4.2 mL, 6.66 mmol). The reaction was stirred at room temperature for 15 min after which time β-ionone (640 mg in 5 mL was added). The mixture was stirred at room temperature for 5 h, after which the TLC showed no starting material. The reaction was quenched with aqueous ammonium chloride and extracted 2 x with ether. The organic phases were washed with saturated NaCl and dried over Na₂SO₄. The crude product was purified by column chromatography, hexanes:ethyl acetate 95:5 to yield 85% of **8** (E:Z, 86:14). [¹H NMR (300 MHz CDCl₃): δ 6.24(1H, d, J=15.9), 6.05(1H, d, J=15.9 Hz), 5.54(1H, s), 2.04(3H, s), 1.98(2H, t, J= 6.59), 1.66(3H, s, CH₃), 1.59(2H, m), 1.43(2H, m), 0.98(6H, s, 2 x CH₃), 0.18(9H, s 3 x SiCH₃). ¹³C NMR (300 MHz CDCl₃): δ 148.55, 137.14, 135.15, 129.80, 129.74, 108.19, 103.76, 100.89, 39.24, 33.93, 32.75, 28.64, 21.39, 18.94, 14.88, -0.17].

3. Synthesis of (9)



Acetylene **8** (2.78 mmol, 0.798 g) in 10 mL of THF was treated with 10 mL of TBAF (1M, 10 mmol) and stirred at room temperature. After two hours a small aliquot was quenched and a ^1H NMR was taken. No starting material was observed. Therefore the reaction was quenched with NH_4Cl and extracted with 2 x ether. The combined organic layers were washed with saturated NaCl and dried over Na_2SO_4 . The product was purified by column chromatography, hexanes to yield 84% of **9**. [^1H NMR (300 MHz C_6D_6): δ 6.14(1H, d, $J=15.9$), 5.95(1H, d, $J=15.9$ Hz), 5.25(1H, s), 3.14(1H, s), 1.93(3H, s), 1.87(2H, t, $J=6.03$ Hz), 1.55(3H, s, CH_3), 1.49(2H, m), 1.33(2H, m), 0.88(6H, s, 2 x CH_3). ^{13}C NMR (300 MHz C_6D_6): δ 149.16, 137.07, 135.08, 130.10, 107.32, 107.26, 83.49, 82.35, 39.46, 34.13, 32.95, 28.84, 21.61, 19.15, 15.01].

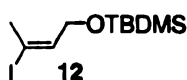
4. Synthesis of (11)



Alcohol **10** (1.5 g, 21.40 mmol) was dissolved in dry ether (30 mL) and chilled to 0 °C under nitrogen. 2.4 equivalents of *i*-BuMgCl (51.36 mmol, 2 M solution, 25.7 mL) were added drop wise and the solution was stirred for 15 min at 0 °C. 5% mol of catalytic CpTiCl_2 was added (1.07 mmol, 261.8 mg). The solution was warmed up to room temperature and stirred for 5 h. The reaction was diluted with 30 mL of dry ether and a solution of 1.65 equivalents of I_2 ,

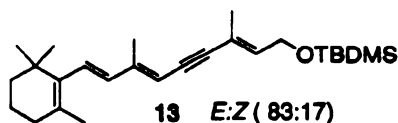
(35.31 mol, 8.82 g in 65 mL of dry ether) were added drop wise. The resulting suspension was stirred at $-78\text{ }^{\circ}\text{C}$ for 1 hour and 30 min at room temperature. To afford **11** mg in a 44% yield. [^1H NMR (300 MHz CDCl_3): δ 6.31 (1H, t, $J=7.14$ Hz), 3.99(2H, d, $J=6.59$ Hz), 2.98 (1H, s), 2.38 (3H, s, CH_3). ^{13}C NMR (300 MHz CDCl_3): δ 139.55, 98.11, 59.43, 27.87].

5. Synthesis of (12)



Alcohol **11** (9.1 mmol, 1.8 g) was dissolved in dry DMF (30 mL) followed by addition of TBDMSCl (13.6 mmol, 2.04 g) and imidazole (18.2 mmol, 1.24 g). The reaction was stirred under nitrogen at room temperature overnight. The reaction was pick up with ether and water (100 mL, and 100 mL). The organic layer was extracted 2x with ether. The combined organic layers were washed with 2x water, rinsed with saturated NaCl, and dried over Na_2SO_4 . The product was purified via column chromatography using 1% ethyl acetate in hexanes giving 78% yield of **12**. [^1H NMR (300 MHz CDCl_3): δ 6.25(1H, t, $J=6.04$), 4.08(2H, d, $J=6.59$ Hz), 2.37(s, 3H, CH_3), δ 0.86 (9H, s, 3 x CH_3), δ 0.032 (6H, s, 2 x SiCH_3). ^{13}C NMR (300 MHz CDCl_3): δ 140.32, 95.72, 60.40, 27.86, 25.65, 18.06, -5.42].

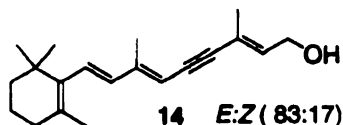
6. Synthesis of (13)



Vinyl iodide **12** (268 mg, 0.86 mmol) was dissolved in isopropyl amine (3 mL), followed by addition of tetrakis(triphenylphosphine)palladium (1% mol, 8.2

mg, 0.007 mmol) was added. The solution was stirred at room temperature for 5 min and CuI (1% mol) was added (1.4 mg, 0.007mmol). After 5 min acetylene **9** was added (149 mg, 0.7 mmol, dissolved in isopropyl amine) and the reaction was stirred at room temperature for 4 h. The reaction was quenched by removal of solvent under reduced pressure followed by addition of ether. The organic layer was extracted with aqueous NH₄Cl. The organic phase was washed with water, rinsed with saturated NaCl and dried over Na₂SO₄. The product was purified by column chromatography using 2% ethyl acetate in hexanes to afford **13** in 87% yield. ¹H NMR (300 MHz CDCl₃): δ 6.22(1H, d, *J*=16.0 Hz), 6.07(1H, d, *J*=15.9 Hz), 5.90(1H, t, *J*=6.59 Hz), 5.50 (1H, s), 4.25(2H, d, *J*=6 Hz), 1.98(2H, t, *J*=6.59), 1.66(3H, s, CH₃), 1.59 (2H, m), 1.43 (2H, m), 0.98(9H, s, 3 x CH₃), δ 0.18(6H, s, 2 x SiCH₃). ¹³C NMR (300 MHz CDCl₃): δ 147.07, 137.43, 135.99, 135.55, 129.95, 129.23, 119.29, 98.77, 60.09, 39.51, 34.18, 33.01, 28.99, 28.89, 25.91, 21.66, 19.19, 18.34, 17.72, 15.03, -5.17].

7. Synthesis of (14)



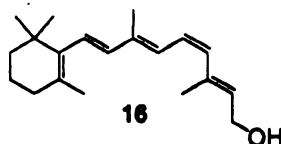
Acetylene **13** (231 mg, 0.58 mmol) was dissolved in dry THF (5 mL) and was slowly treated with TBAF (3 ml, 1M in THF, 3 mmol). The reaction was stirred at room temperature for 3 h until the TLC showed that all starting material was consumed. Water was added to quench the reaction and it was extracted 2x with ether. The organic layers were washed with saturated NaCl and dried over

anhydrous Na₂SO₄. The product **14** was purified by column chromatography (90% hexanes, 10% ethyl acetate) in a 90% yield as a mixture of (E:Z, 83:17) isomers.

Separation of isomers **14** and **17**

A ALTEX UltraShereTM-Si (C8) semi-prep column (410 mm ID x 25cm (5 μ m) was used. As a mobile phase 80:20 Hexanes:Ethyl acetate was used, and a 3 mL/min flow. The peaks were detected at 317 nm under dim red light. A solution of **14** was prepared (1 mg/100 μ L 80:20 Hx:EtOAc). 100 μ L was injected at every run and the two different fractions were collected (R_f =22.9 and R_f =26.3). About 94% of the total weight was recovered every time. Each fraction was concentrated under rotovap and dried under vacuum, and further analyzed by ¹H-NMR (C₆D₆). [¹H NMR (500 MHz C₆D₆): δ 6.28(1H, d, J =15.91 Hz), 6.16(1H, d, J =16.34 Hz), 6.02(1H, t, J =6.60 Hz), 5.68(1H, s), 3.83(2H, t, J =6.19 Hz), 2.12(3H, s, CH₃), 1.91(2H, t, J =6.19), 1.68(3H, s), 1.66(3H, s, CH₃), 1.55(2H, m), 1.43 (2H, m), 1.043 (6H, s, 2 x CH₃). ¹³C NMR (300 MHz CDCl₃): δ 147.07, 137.43, 135.99, 135.55, 129.95, 129.23, 119.29, 98.77, 60.09, 39.51, 34.18, 33.01, 28.99, 28.89, 25.91, 21.66, 19.19, 18.34, 17.72, 15.03, -5.17].

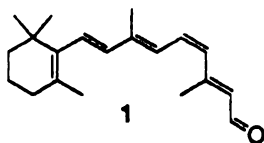
8. Synthesis of (**16**)



Activated Zn dust was prepared as described by Boland with a few modifications. Argon was bubbled through a suspension of Zn dust (1 g) in

distilled water (6 mL) for 15 min. $\text{Cu}(\text{OAc})_2$ (100 mg) was added and the flask (exothermic reaction) was sealed immediately. The mixture was stirred vigorously for 15 min. AgNO_3 (1 g) was then added and the solution was stirred for 30 min. The activated Zn was centrifuged (400 RPM, rotor 243, 90 sec) and washed successively with water, methanol, acetone and ether. The moist activated Zn (important) was transferred immediately to a flask of the reaction solvents (water 2mL and methanol 2 mL). Compound **14** (15 mg, 0.053 mmol) was added to this mixture and was then stirred at room temperature in the dark overnight. After the TLC showed no starting material, the Zn dust was filtered through celite with ether and water. The organic phases were separated, washed with saturated NaCl, and dried over anhydrous Na_2SO_4 . The solvent was removed under reduced pressure to yield **16** plus some isomers. Purification at this stage was not performed. [^1H NMR (500 MHz CDCl_3): δ 6.53(1H, d, $J=11.93$ Hz), 6.35(1H, t, $J=11.93$, 11.93 Hz), 6.41(1H, d, $J=16.13$ Hz), 6.10(1H, d, $J=15.90$), 5.87(1H, d, $J=11.71$), 5.72(1H, t, $J=6.45$), 4.26(2H, d, $J=7.00$ Hz), 2.01(2H, t, $J=6.19$), 1.93(3H, s, CH_3), 1.89 (3H, s, CH_3), 1.71 (3H, s, CH_3), δ 1.62 (2H, m), 1.47 (2H, m), 1.03(6H, s, 2 x CH_3)].

9. Synthesis of (1)



The crude reaction from **16** was dried under vacuum, dissolved in CH_2Cl_2 (3 mL) and chilled to 0 °C. MnO_2 was added (1.30 mmol, 120 mg) and the

reaction was stirred for 4 h. The reaction was passed through a celite pad to filter MnO_2 and washed with CH_2Cl_2 . The solvent was dried over anhydrous Na_2SO_4 . The crude was passed through a micro column to separate the polar impurities (Silica was stabilized with hexanes and eluted with 5% EtOAc in hexanes). The fraction containing the 11-*cis*-retinal **14** was collected and purified using HPLC, (for details see below). The ^1H -NMR spectrum matches the reported data. [^1H NMR (500 MHz CDCl_3): δ 10.07 (1H, d, $J=8.17$ Hz), 6.66 (1H, t, $J=12.15$ Hz), 6.51 (1H, t, $J=12.59$ Hz), 6.31 (1H, d, $J=16.13$ Hz), 6.12 (1H, d, $J=16.13$ Hz), 6.06 (1H, d, $J=8.17$ Hz), 5.91 (1H, t, $J=11.49$ Hz), δ 2.33 (3H, s, CH_3), 2.00 (2H, t, $J=5.29$ Hz), 1.97 (3H, s, CH_3), 1.69 (3H, s, CH_3), 1.60 (2H, m), 1.45 (2H, m), 1.02 (6H, s, 2 x CH_3)].

Purification of 11-*cis*-retinal

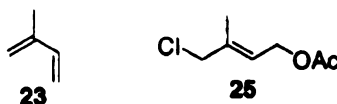
A column ALTEX UltraShereTM-Si (C8) semi-prep (410 mm ID x 25cm (5 μm) was used, and the column was rinsed with 98:2 hexanes:ethyl acetate 0.05 v/v Et_3N (to neutralize any acid present). All procedure were performed under dim red light. All samples were kept at 0 $^\circ\text{C}$ or less, and in the dark when possible. A mobile phase of 98:2 hexanes:ethyl acetate, was used, 3 mL/min, at room temperature and a UV detection at 362 nm. The crude (10-15 mg) was dissolved in 1 mL of 2% EtOAc in hexanes. 100 μL was injected at every run and three different fractions were collected ($R_f=18.39$, 17.33%; $R_f=22.64$, 27%; and $R_f=39.74$, 24%).

Each fraction was concentrated under rotovap and dried under vacuum. Further analysis of the $^1\text{H-NMR}$ (CDCl_3) of each fraction showed that fraction 2 ($R_f=22.64$) was 11-*cis*-retinal. 1.5 mg of 11-*cis*-retinal were recovered 10%.

Preparation of *tert*-butyl hypochlorite *t*-BuOCl

In a 1 L round-bottomed flask, provided with mechanical stirring and protected from light, bleach (500mL) was cooled in an ice/water bath until the temperature of the bleach was below 10 °C. A mixture containing *tert*-butanol (35.25 mL) and acetic acid (25 mL) was then added in one portion and the whole mixture was vigorously stirred for 3 min. The contents of the flask were poured in a extraction funnel and the lower aqueous layer was discarded. The organic fraction was washed with 10% solution of Na_2CO_3 (50 mL) and then with water (50 mL). The *t*-butyl hypochlorite thus formed was dried over CaCl_2 , filtered and used subsequently without further purification. About 70 g were obtained; the *tert*-butyl hypochlorite was used without further purification.⁷⁴

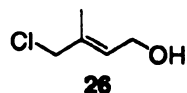
10. Synthesis of (*E*)-1-O-acetyl-4-chloro-3-methyl-2-butene (25)



Isoprene **23** (12.5 mL, 12.5 mol) was dissolved in acetic acid (72.5 mL, 12.5 mol) and cooled in an ice/salt bath. *tert*-Butyl hypochlorite (12.5 g, 12.5 mol) was added in small portions, and the solution was stirred for 1 h. The excess acid was carefully neutralized with a solution of sodium hydroxide and extracted

with ether (20 N NaOH was used). The organic phase was dried over Na₂SO₄, the solvent evaporated and the crude product was purified by column chromatography (pentane/ethyl ether 10:1) to afford **25** in 33% yield. [¹H NMR (300 MHz CDCl₃): δ 5.64(1H, t, *J*=5.49 Hz), 4.54(d, 2H *J*=6.59 Hz), 3.95(2H, s, CH₂Cl), 2.01(3H, s, CH₃), 1.77(3H, s, CH₃). ¹³C NMR (300 MHz CDCl₃): δ 170.63, 136.81, 123.62, 60.57, 50.65, 20.71, 14.38].

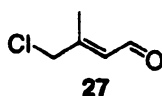
11. Synthesis of (*E*)-1-O-acetyl-4-chloro-3-methyl-2-buten-1-ol (**26**)



tert-Butyl hypochlorite (8.65 g, 0.8 mol), was added at 0 °C to a solution of isoprene (6.8 g, 0.1 mol) in acetic acid (29 mL, 0.48 mol). The reaction mixture was stirred for 1 h, and extracted with water. The combined extracts were successively washed with sat. NaHCO₃ and sat NaCl, and dried with Na₂SO₄. The filtrate was concentrated under reduced pressure, dissolved in acetic acid (30 mL) and treated with both copper sulfate (0.20 g, 1 mmol) and sulfuric acid (0.2 g, 2 mmol). The resultant mixture was stirred for 4 days at ambient temperature quenched by the addition of water and extracted with ether.⁷⁵ The separation of this isomers was impossible. 3.4 g of the crude product was dissolved in methanol (60 mL) and stirred for 5 h at 0 °C in the presence of an aqueous solution (20 mL) of sodium carbonate (3.6 g in 20 mL).⁷⁵ The mixture was filtered, concentrated and poured onto ice/water. The water was extracted with chloroform (3x), and dried with anh. Na₂SO₄. The crude product was purified by

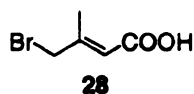
chromatography to afford the alcohol as a mixture of isomers. The crude product (5 g) was dissolved in methanol (60 mL) and stirred for 5 h at 0 °C in the presence of an aqueous solution (20 mL) of sodium carbonate (3.6 g in 20 mL).² The mixture was filtered, concentrated and poured onto ice/water. The water was extracted with chloroform (3x), and dried with anh. Na₂SO₄. The crude product was purified by column chromatography using 7:3 pentane:ethyl ether. The chloro-alcohol was isolated in (3.7 g) a 99.8%. [¹H NMR (300 MHz CDCl₃): δ 5.67(1H, t, *J*=7.14 Hz), δ 4.11(2H, d, *J*=6.04 Hz), 3.95 (2H, s, CH₂Cl), 1.72(3H, s, CH₃). ¹³C NMR (300 MHz CDCl₃): δ 134.17, 128.83, 58.77, 51.24, 14.23].

12. Synthesis of (*E*)-1-O-acetyl-4-chloro- 3-methyl-2-butenal (27)



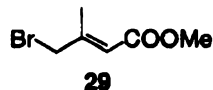
The chloro alcohol **26** (1.8 g, 13.2 mmol) in dichloromethane (20 mL) was added at 0 °C to a suspension of pyridinium chlorochromate (4.4 g, 21 mmol) in dichloromethane (40 mL). The solution was kept at 0 °C for an additional 90 min then filtered through a short alumina column. Elution with ether yield 1.5 g 4-chloro-3-methyl-2-butenal (85%) (two aldehyde peaks could be observed, in 87:13 ratio). After this material is prepared it should be used within one week to perform the next reaction, otherwise it decomposes. ¹H NMR (300 MHz CDCl₃): δ 9.95(1H, d, *J*=7.69 Hz), 6.03(2H, d, *J*=7.69 Hz), 4.04(2H, s, CH₂Cl), 2.19(3H, s, CH₃). ¹³C NMR (300 MHz CDCl₃): δ 190.82, 155.19, 49.06, 15.26. *m/z* (E.I): 118.01(M⁺), 53(100)].

13. Synthesis of (*E*)-4-bromo-3-methyl-2-butene acid (**28**)



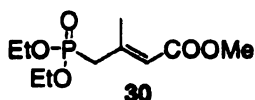
3,3 dimethyl acrylic acid (10 g, 0.1 mol) was dissolved in CCl₄ (70 mL). Then of N-bromo succinimide (NBS) (0.11 mol 20 g) was added and the solution was illuminated with a lamp and heated to a reflux for about 4 h. After 4 h the succinamide had precipitated out. The solution was stirred for extra 2 h without the lamp. To the solution, hexane (30 mL) was added and the precipitate was filtered out under reduced pressure. The filtrate was concentrated under *vacuo*. The residue was picked up in MeOH and excess K₂CO₃ anhydrous was added. Then, the solution was stirred for 2 h resulting in lactonization of the *cis*-bromo-alkene. Most of the MeOH was removed under reduced pressure and the residue was picked up with ether and water. The aqueous phase was the acidified to pH=1 with concentrated H₂SO₄ and extracted three times with ether. The ether was washed with saturated NaCl and dried over Na₂SO₄. The crude product gave from 20 to 42% yield of the pure *E*-isomer. The highest yield was obtained when fresh re-crystallized NBS was used. [¹H NMR (300 MHz CDCl₃): δ 8.10(1H, b, COOH), 5.96(1H, s), 3.94(2H, s, CH₂Br), (3H, s, CH₃). ¹³C NMR (300 MHz CDCl₃): δ 171.08, 155.27, 118.81, 37.91, 17.52].

14. Synthesis of (29)



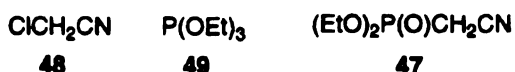
The product **28** was used without further purification and was dissolved in MeOH (25 mL) and it was acidified with concentrated H₂SO₄ (2 mL). The reaction was refluxed overnight. The reaction was picked up with ether and water. The aqueous layer was extracted three times with ether. The organic layer was washed with aqueous K₂CO₃ two times and with saturated NaCl. The product was purified using column chromatography using eluant 2% ethyl acetate 98% hexanes to afford 82% yield of the bromo ester **29**. [¹H NMR (300 MHz CDCl₃): δ 5.94(1H, s), 3.92(2H, s, CH₂Br), 3.69(3H, s, OCH₃), 2.26(3H, s, CH₃). ¹³C NMR (300 MHz CDCl₃): δ 166.15, 152.67, 118.94, 51.18, 38.10, 17.07. m/z (E.I): 191.9, 193.9 (1:1) (M⁺), 113.1(100)].

15. Synthesis of (30)



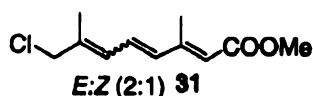
The bromo ester **29** (1 g, 5 mmol) was combined with ethyl phosphite (0.86 g, 5 mmol) in a sealed tube. The mixture was heated overnight at 125 °C. The reaction yielded 2.2 g of crude product (93% yield). The product was used without further purification. [¹H NMR (300 MHz CDCl₃): δ 5.70(1H, s), 4.02(2H, q, *J*=8.14 Hz, OCH₂CH₃), 3.60(3H, s, OCH₃), 2.60(2H, d, *J*=23.35 Hz, PCH₂), 2.22(3H, s), 1.23(6H, t, *J*=4.14 Hz, 2 x OCH₂CH₃). ¹³C NMR (300 MHz CDCl₃): δ 166.27, 149.96, 119.44, 62.13, 50.84, 38.33(d, *J*=134.33 Hz, PCH₂), 19.87, 16.27. m/z (E.I): 250.0(M⁺), 133.7(100)].

16. Synthesis of (47)



To a sealed tube one equivalent of and triethylphosphite **49** (11 g, 66 mmol) was added followed by a drop wise addition of one equivalent of chloroacetonitrile **48** (5 g, 66 mmol). The mixture was heated to 125 °C over night. The reaction was cooled down to room temperature, and solid formed was resuspended in a minimal amount of dichloromethane. The product was precipitated by addition of ether, and the crystals obtained were re-crystallized using dichloromethane and ether. The ylide **47** was obtained in a 98% yield (11.5 g as a yellowish oil). [¹H NMR (300 MHz CDCl₃): δ 4.45(4H, q, *J*=7.14 Hz, OCH₂CH₃), 2.60(2H, d, *J*=20.88 Hz, PCH₂), 1.48(6H, t, *J*=7.14 Hz, 2 x OCH₂CH₃).

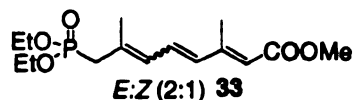
17. Synthesis of (31)



The ylide **30** (0.4 g) were dissolved in dry THF (3 mL) and NaHMDS (1.6 mL of 1 M) were added at 0 °C. The solution was stirred at 0 °C for 30 min. The aldehyde **27** (0.19 g) was dissolved in THF (5 mL) and was added slowly to the ylide solution at 0 °C. Then, the reaction was stirred for 4 h. The reaction quenched with water and the aqueous layer was extracted with ether three times. The combined organic layers were washed with 10% aqueous HCl, followed by saturated NaHCO₃, and saturated NaCl. Then the organic layer was dried Na₂SO₄ (anh) and was concentrated to give the chloro ester **31**. The compound was

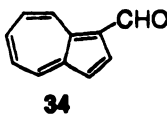
purified using chromatography gel hex:ether 92:8 and 100 mg of chloro ester 31 were obtained in a 30% yield. [^1H NMR (300 MHz CDCl_3): δ 7.75(1H, d, $J=16.0$ Hz), 6.73(1H, m), 6.23(2H, m), 5.65(1H, s), 4.06(2H, s, CH_2Cl), 3.66(3H, s, OCH_3), 2.00(3H, s, CH_3), 1.90(3H, s, CH_3). ^{13}C NMR (300 MHz CDCl_3): δ 166.27, 150.52, 136.96, 130.51, 129.68, 119.34, 117.37, 51.75, 50.91, 22.55, 20.72].

18. Synthesis of (33)



The chloro ester (0.63 g, 2.75 mmol) was combined with ethyl phosphite (0.46 g, 2.75 mmol) of in a sealed tube. The mixture was heated for 16 hours at 100 °C. After the reaction the crude gave **33** (0.78 g) in a 90% yield. The product was used without further purification. [^1H NMR (300 MHz CDCl_3): δ 7.64(1H, d, $J=16.0$ Hz), 6.76(1H, m), 6.14(2H, m), 5.66(1H, d, $J=14.10$ Hz), 4.05(2H, q, $J=8.14$ Hz, OCH_2CH_3), 3.66(3H, s, OCH_3), 2.62(2H, d, $J=23.45$ Hz, PCH_2), 2.29(3H, s, CH_3), 2.00(3H, s, CH_3), 1.23(4H, t, $J=4.14$ Hz, OCH_2CH_3)].

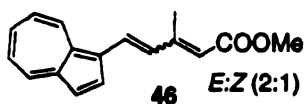
19. Synthesis of (34)



POCl_3 (394 mg, 2.57 mmol) was added to a flask and cooled down to 0 °C, then DMF (6 mL) was added drop wise under N_2 . The mixture was stirred at 0 °C for 30 min. Azulene (300 mg, 2.0 mmol) was dissolved in DMF (4 mL, anh.) and were cooled down to 0 °C, then the azulene solution was added drop

wised to the mixture. The reaction was stirred for 2 h. Water was added to quench the reaction and it was let to warm slowly (to avoid sudden reaction with water and excess POCl₃). Saturated solution of NaHCO₃ (3-5 mL) was added to basify the solution. The aqueous layer was extracted with ether five times. The combined organic layer were washed with saturated NaCl and dried over anhydrous sodium sulfate. The crude was purified using flash chromatography hexanes:ether 4:1 to afford 75.5% of the aldehyde and about 10% of SM. TLC: Hexanes:Ether 4:1, Starting Material R_f=0.64 blue color, product R_f 0.23 red color. [¹H NMR (300 MHz CDCl₃): δ 10.33(1H, s), 9.54(1H, d, *J*=9.89 Hz), 8.45(1H, d, *J*=9.6 Hz), 8.23(1H, d, *J*=9.12 Hz), 7.81(1H, t, *J*=9.9 Hz), 7.57(1H, t, *J*=9.9 Hz), 7.48(1H, t, *J*=9.9 Hz), 7.29 (1H, d, *J*=5.1 Hz). ¹³C NMR (300 MHz CDCl₃): δ 186.51, 146.02, 141.88, 139.69, 138.96, 137.44, 129.43, 128.20, 125.89, 119.00].

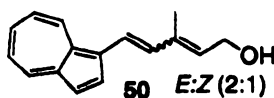
20. Synthesis of (46)



Ylide **30** (0.96 g of the 3.8 mmol) was dissolved in dry THF(10 mL) and of NaHMDS (3.84 mL, 1 M) were added slowly at 0 °C. The solution was stirred for 30 min. and then the solution was cooled to 0 °C. The aldehyde **34** (0.06 g) was dissolved in THF (5 mL) and cooled down to 0 °C and then it was added slowly to the ylide solution. The reaction was stirred for 1 h and then it was warmed slowly to 0 °C. After 6 h the reaction was stopped by addition of water and it was warmed to room temperature. The aqueous layer was extracted

three times with ether. The combined organic layers were washed with 10% aqueous HCl, saturated NaHCO₃, saturated NaCl and dried over Na₂SO₄ (anh). The organic layer was concentrated to give the ester **46**. The compound was purified via flash chromatography using hexane:ethylacetate 92:8 to yield 40 mg of product **46** were obtained in a (41%) yield as a mixture of two isomers. TLC 90:10 Hex:EtOAc, R_f=0.48 greenish and 20% of starting material **34** was recovered. [¹H NMR (300 MHz CDCl₃): δ 8.43(1H, d, *J*=9.61 Hz), 8.21(1H, d, *J*=9.06 Hz), 8.12(1H, d, *J*=4.12 Hz), 7.53(2H, d, *J*=13.18 Hz), 7.38(1H, d, *J*=4.12 Hz), 7.13(2H, q, *J*=10.44 Hz), 6.91(1H, d, *J*=15.66 Hz), 5.89 (1H, s), 3.74(3H, s, OCH₃), 2.51(3H, s, CH₃). ¹³C NMR (300 MHz CDCl₃): δ 167.82, 138.58, 137.11, 136.63, 133.69, 133.57, 129.44, 126.08, 124.67, 123.63, 119.59, 117.03, 67.96, 50.95, 13.82].

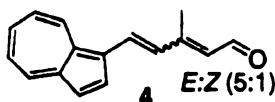
21. Synthesis of (**50**)



Ester **46** (40 mg, 0.16 mmol) were dissolved in dry THF (5 mL). The solution was cooled down to -78 °C and DIBAL-H (0.130 mL 2.5 M in hexanes) was added drop wise, and then it was stirred for 4 h. The reaction was quenched with Na,K tartrate. The solution was warmed to room temperature and stirred for 1 hour. The aqueous solution was extracted five times with ether. The organic layers were combined and dried over Na₂SO₄ (anh), then the organic layer was concentrated to give the alcohol **50**. The crude product was used directly for the next reaction. [¹H NMR (300 MHz CDCl₃): δ 8.38(1H, d, *J*=9.99 Hz), 8.17(2H,

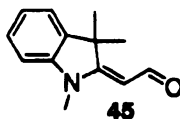
m), 7.51(2H, m), 7.36(1H, t, $J=4.54$ Hz), 7.13-7.03(3H, m), 6.91(1H, d, $J=15.66$ Hz), 5.61(1H, t, $J=6.87$ Hz), 4.43 (2H, d, $J=17.41$ Hz), 2.09(3H, s, CH₃)].

22. Synthesis of (4)



The crude mixture of **50** (6.5 mg, 0.16 mmol) was dissolved in CH₂Cl₂ (5 mL) and chilled to 0 °C. MnO₂ was added (3.2 mmol, 300 mg) and the reaction was stirred for 3 h. The reaction was passed through a celite pad to filter MnO₂ and washed with CH₂Cl₂. The solvent was dried over anhydrous Na₂SO₄. The aldehyde **4** was purified via flash chromatography using exane:EtOAc 92:8 to yield 6.5 mg of **4** 18%. [¹H NMR (300 MHz CDCl₃): δ 10.13(1H, d, $J=8.24$ Hz), 8.46(1H, d, $J=9.82$ Hz), 8.23(2H, m), 7.68(1H, d, $J=15.9$ Hz), 7.63(2H, d, $J=13.18$ Hz), 7.38(1H, d, $J=3.82$ Hz), 7.2(2H, q, $J=9.88$ Hz), 6.96(1H, d, $J=15.96$ Hz), 5.89 (1H, d, $J=8.2$ Hz), 2.47(3H, s, CH₃). ¹³C NMR (300 MHz CDCl₃): δ 167.82, 138.58, 137.11, 136.63, 133.69, 133.57, 129.44, 126.08, 124.67, 123.63, 119.59, 117.03, 67.96, 50.95, 13.82. UV: λ_{max} (EtOH) 435 nm; ε = 14,1918 M⁻¹cm⁻¹].

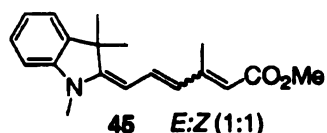
23. Synthesis of (45)



(COCl)₂ (1.175 g, 9.3 mmol) were added to a cooled flask (0 °C) containing DMF (30 mL) and CCl₄ (25 mL) and the mixture was stirred at 0 °C for 30 min. Indolin was added dropwised (2 ml, 1.958 g, 11.3 mmol) were

dissolved in DMF (10 mL) and were cooled down to 0 °C. The reaction was then warm up to room temperature and to 50-60 °C for 3 h. Cold water was added to quench the reaction and the pH was adjusted to >10. The reaction was extracted with ether five times. The organic layer was washed with saturated NaCl and dried over anhydrous sodium sulfate. The crude product was purified using flash chromatography hexanes:ether 4:1 to afford 93% of the aldehyde. [¹H NMR (500 MHz CDCl₃): δ 9.94(1H, d, *J*=9.07 Hz), 7.23(1H, t, *J*=7.42 Hz), 7.19(1H, d, *J*=6.31 Hz), 6.99(1H, d, *J*=7.41 Hz), 6.80(1H, d, *J*=6.97 Hz), 5.33(1H, d, *J*=9.16 Hz), 3.17(3H, s, NCH₃), 1.59(6H, s, 2 x CH₃). ¹³C NMR (300 MHz CDCl₃): δ 186.32, 173.49, 143.26, 139.19, 127.90, 122.30, 121.63, 117.00, 107.87, 98.77, 47.24, 29.37].

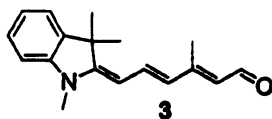
24. Synthesis of (45)



The ylide **30** (1.37 g 5.50 mmol) was dissolved in THF (20 mL, dry) and NaHMDS (5.5 mL 1 M) were added at 0 °C. Then, the solution was stirred for 30 min and it was cooled to 0 °C. Aldehyde **38** (0.11 g, 1 mmol) was dissolved in THF (5 mL) and cooled down to 0 °C and then it was added slowly to the ylide solution. The reaction was stirred for 1 h and then it was warmed slowly to room temperature. After 6 h the reaction was stopped by addition of water and the reaction was extracted three times with ether. The combined organic layers were then washed with saturated NaHCO₃, saturated NaCl and dried with Na₂SO₄.

(anh). The organic layer was concentrated to give the ester **38**. The compound was purified via flash chromatography using hexane: EtOAc 9:1 and 40 mg of product **38** were obtained in a 50% yield as a mixture of two isomers. TLC 85:15 Hex:EtOAc, $R_f=0.35$. [^1H NMR (500 MHz CDCl_3): δ 7.29-7.16(2H, m.), 6.89(1H, t, $J=7.42$ Hz), 6.66(1H, d, $J=8.06$ Hz), 6.10(1H, d, $J=15.65$ Hz), 5.69(1H, s), 5.54(1H, d, $J=15.14$ Hz), 5.41(1H, d, $J=11.96$ Hz), 3.73(3H, s, NCH_3), 3.17(3H, s, OCH_3), 1.63(3H, s, CH_3), 1.62(6H, s, CH_3). ^{13}C NMR (300 MHz CDCl_3): δ 168.10, 159.43, 154.33, 144.88, 138.69, 133.39, 127.74, 121.52, 119.79, 113.59, 111.42, 106.04, 95.96, 50.67, 45.66, 28.4, 21.09, 14.40, 13.87].

25. Synthesis of (**3**)



The ester **45** (40 mg, 0.13 mmol) was dissolved in dry THF (5 mL). The solution was cooled down to -78°C . DIBAL-H (0.130 mL of 1 M in hexane) was added drop wise. After 4 hours the reaction was quenched with Na,K tartrate and then it was warmed to room temperature and stirred for 4 h. The aqueous solution was extracted five times with ether. The organic layers were combined and dried over Na_2SO_4 (anh). The organic layer was concentrated under reduced pressure. The crude product was used directly for the next reaction. The crude mixture was dissolved in CH_2Cl_2 (5 mL) and chilled to 0°C . MnO_2 freshly homemade was added (2.6 mmol, 244 mg) and the reaction was stirred for 3 hours. The reaction was passed through a celite pad to filter MnO_2 and washed with CH_2Cl_2 . The solvent was dried over anhydrous Na_2SO_4 . The aldehyde **3** was purified via flash

chromatography using Hexane:EtOAc 85:15 to yield 1.5 mg (5% yield) starting from the ester **45**. [^1H NMR (500 MHz CDCl_3): δ 10.01(1H, d, $J=8.80$ Hz), 7.20-7.00(2H, d, $J=14.28$), 6.80-6.45(2H, m), 6.10(1H, d, $J=15.93$ Hz), 5.87(1H, d, $J=14.83$ Hz), 5.66(1H, d, $J=9.21$ Hz), 5.44(1H, d, $J=14.28$ Hz), , 3.20(3H, s, NCH_3), 3.17(3H, s, OCH_3), 2.29(3H, s, CH_3), 1.58(6H, s, 2 x CH_3). UV: λ_{max} (EtOH) 468 nm].

B. Analysis of CRABPII mutants with chromophores.

UV-vis titrations of CRABPII with chromophores

A stock solution of the chromophores was prepared in 95% spectroscopy grade EtOH. For a 5 μM protein solution in PBS (4 mM NaH_2PO_4 , 16 mM Na_2HPO_4 , 150 mM NaCl, pH=7.3).

Additions from 0.1–1.0 equivalents at 0.1 equivalent increments were added and spectra recorded at room temperature from 200-700 nm (Cary WinUV, Varian). The 2nd derivative of the spectra was calculated by using the corresponding software provided with the UV instrument. The determination of the λ_{max} of a UV peak using mathematical equations has been established as a valid method, and has been used in a variety of applications.⁷⁶⁻⁷⁸ As previously mentioned, all the UV data discussed and shown in tables refer to spectra recorded in the presence of 0.1–0.2 equivalents of chromophore, unless otherwise noted.

C. Fluorescence and MALDI-TOF

A detailed description of the determination of binding constant for the CRABPII mutants using retinal, is provided in Chapter 2, of Chrysoula Vasileiou's Dissertation.

In short, the cuvette is allowed to sit with 3 mL of a 0.01% gelatin containing PBS (4 mM NaH₂PO₄, 16 mM Na₂HPO₄, 150 mM NaCl, pH=7.3) for 30-60 min. The sample is excited at 283 nm with a slit width about 1.5 nm (this varies in order to ensure that the intensity of each sample remains below one million counts). The fluorescence is measured at the peak maximum, about 345 nm (varies with different proteins). The chromophore is added at varying equivalents from a ~1.5 mM stock (in 95% spectroscopy grade EtOH) sample maintained in the dark. A measurement is taken after each addition at the same wavelength. The results are plotted as concentration of chromophore versus relative fluorescence intensity. When the curve levels off, the titration is complete. The dissociation constant is

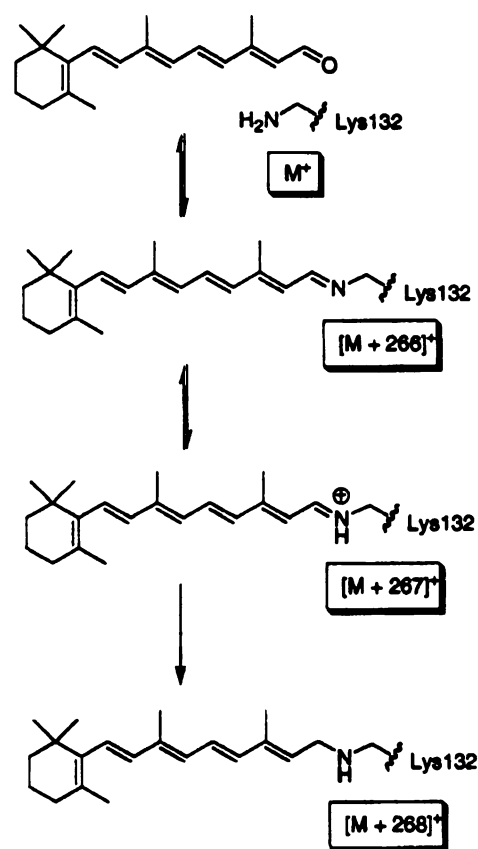


Figure 6-A. Covalent bond formation between retinal and the protein can be detected by MALDI-TOF. Schiff base or PSB formation can be seen as an $[M+266]^+$ or $[M+267]^+$ peak, respectively. Both species can be reductively aminated and trapped as an $[M+268]^+$ species.

calculated with the program SigmaPlot through non-linear least square regression analysis.⁷²

D. MALDI-TOF

A detailed description of the determination of binding constant for the

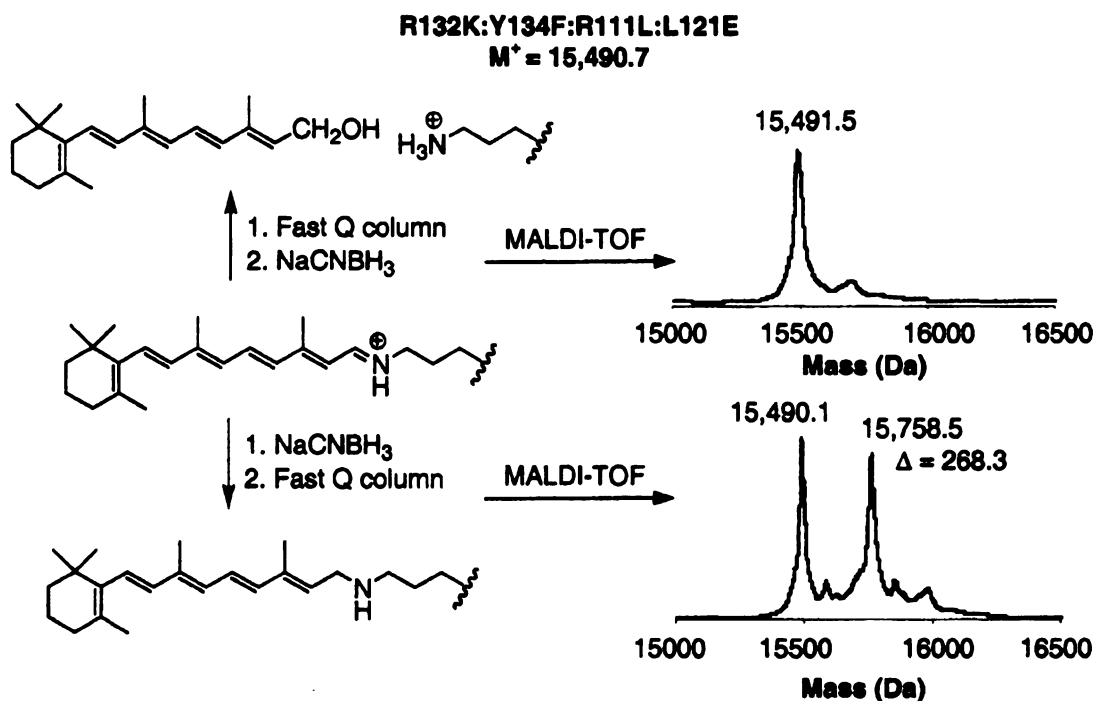


Figure 6-B. Schematic representation of the reductive amination process using R132K:Y134F:R111L:L121E mutant. If the Fast Q column is run before the reductive amination no retinal adduct is detected MALDI-TOF (top). On the contrary, when the column follows the reductive amination, the characteristic adduct of $(M+268)^+$ is observed (bottom). Fast Q column can successfully separate the free retinal from the CRABPII mutants.

CRABPII mutants using retinal, is provided in Chapter 2, of Chrysoula Vasileiou's Dissertation.

An additional binding assay performed on each CRABP_{II} mutant involves the use of two MALDI-TOF experiments. The first of these experiments is the incubation of retinal with the protein for Schiff base formation. The sample is then washed with ethanol and concentrated. The covalent bond between retinal and the protein, if formed, can be detected as an $[M+266]^+$ peak in the MALDI-TOF spectrum. In addition, reductive amination of this incubated sample with NaCNBH₃ is performed to covalently trap the retinal molecule with the protein (Figure 6-B). The necessity for both MALDI-TOF experiments arises from the fact that some Schiff bases, protonated or not, may not be stable enough to be trapped under the reductive amination reaction conditions, but will survive the work-up from the incubation long enough to be detected.

Since an excess of retinal is used for the reductive amination reaction, the non-covalently bound chromophore must be separated using a Fast Q column (quaternary ammonium resin).

The MALDI-TOF spectra for either the incubation or the reductive amination reactions can only be evaluated qualitatively and not quantitatively.

E. Molecular Modeling:

Computational results were obtained using software programs from Accelrys. Dynamics and minimizations were done using the Discover 3[®] program, using the CVFF forcefield from within the InsightII 2000 molecular modeling system.

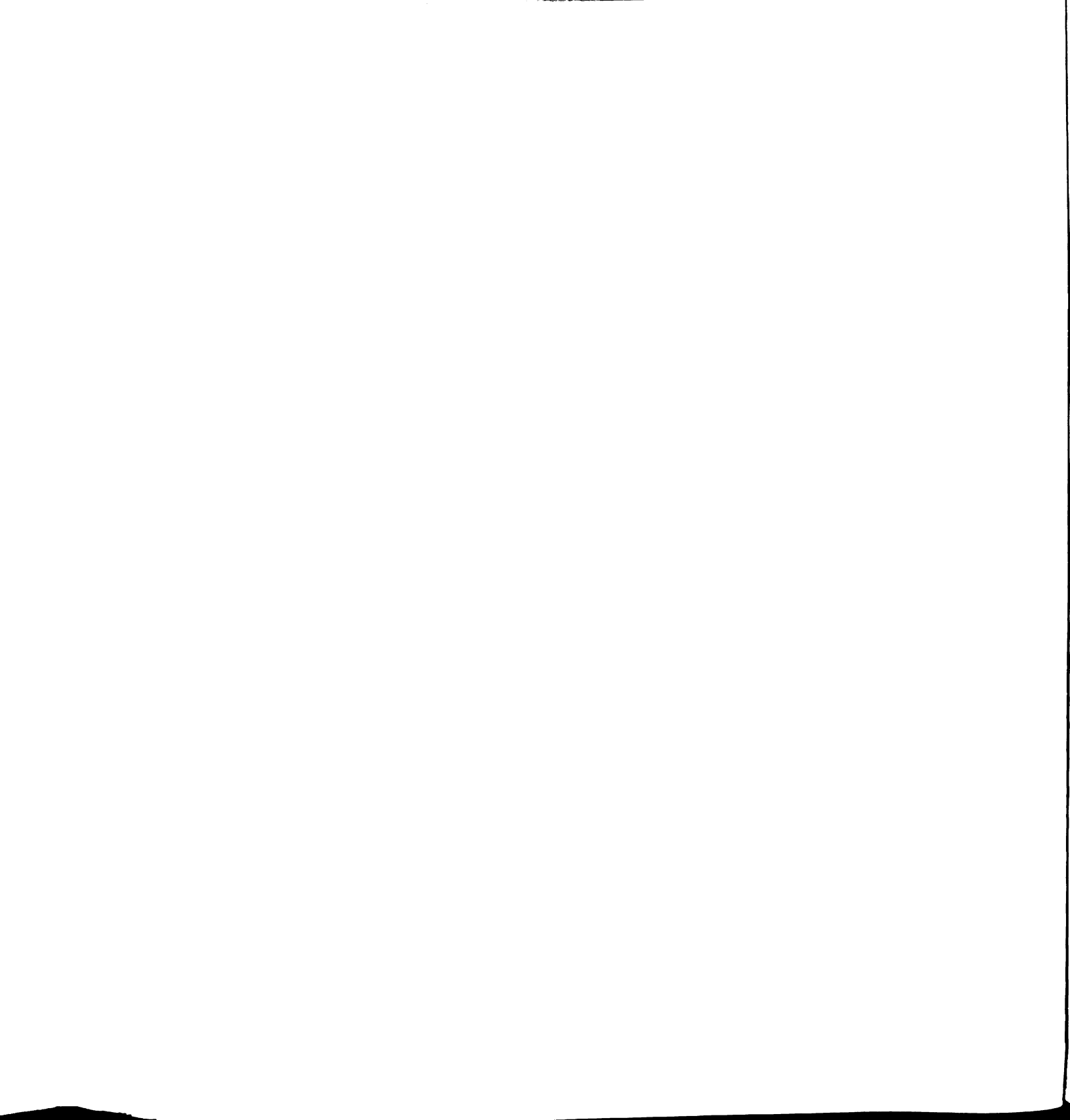
Protein figures were generated using PyMol Molecular Graphics System,
version 0.93, copyright by DeLano Scientific LLC.

6.6 References

1. D. D. Oprian; A. B. Asenjo; N. Lee; S. L. Pelletier, "Design, chemical synthesis, and expression of genes for the 3 human color-vision pigments." *Biochemistry* **1991**, *30*, 11367-11372.
2. S. L. Merbs; J. Nathans, "Absorption-spectra of human cone pigments." *Nature* **1992**, *356*, 433-435.
3. H. J. M. Degroot; G. S. Harbison; J. Herzfeld; R. G. Griffin, "Nuclear magnetic-resonance study of the Schiff-base in bacteriorhodopsin-counterion effects on the N-15 shift anisotropy." *Biochemistry* **1989**, *28*, (8), 3346-3353.
4. J. Hu; R. G. Griffin; J. Herzfeld, "Synergy in the spectral tuning of retinal pigments: Complete accounting of the opsin shift in bacteriorhodopsin." *Proceedings of the National Academy of Sciences of the United States of America* **1994**, *91*, 8880-8884.
5. A. G. Doukas; A. Pande; T. Suzuki; R. H. Callender; B. Honig; M. Ottolenghi, "On the mechanism of hydrogen-deuterium exchange in bacteriorhodopsin." *Biophysical Journal* **1981**, *33*, (2), 275-279.
6. H. Luecke; B. Schobert; H. T. Richter; J. P. Cartailler; J. K. Lanyi, "Structure of bacteriorhodopsin at 1.55 angstrom resolution." *Journal of Molecular Biology* **1999**, *291*, (4), 899-911.
7. A. Kusnetzow; D. L. Singh; C. H. Martin; I. J. Barani; R. R. Birge, "Nature of the chromophore binding site of bacteriorhodopsin: The potential role of Arg(82) as a principal counterion." *Biophysical Journal* **1999**, *76*, (5), 2370-2389.
8. T. Ebrey; M. Tsuda; G. Sassenrath; J. L. West; W. H. Waddell, "Light activation of bovine rod phosphodiesterase by non-physiological visual pigments." *FEBS Letters* **1980**, *116*, 217-219.
9. C. N. Rafferty, "Light-induced perturbation of aromatic residues in bovine rhodopsin and bacteriorhodopsin." *Photochemistry and Photobiology* **1979**, *29*, 109-120.

10. T. Okada; M. Sugihara; A. N. Bondar; M. Elstner; P. Entel; V. Buss, "The retinal conformation and its environment in rhodopsin in light of a new 2.2 angstrom crystal structure." *Journal of Molecular Biology* **2004**, 342, (2), 571-583.
11. R. E. Stenkamp; S. Filipek; C. A. G. G. Driessen; D. C. Teller; K. Palczewski, "Crystal structure of rhodopsin: a template for cone visual pigments and other G protein-coupled receptors." *Biochimica Et Biophysica Acta-Biomembranes* **2002**, 1565, (2), 168-182.
12. J. Nathans; D. Thomas; D. S. Hogness, "Molecular-genetics of human color-vision-the genes encoding blue, green, and red pigments." *Science* **1986**, 232, 193-202.
13. R. Yokoyama; S. Yokoyama, "Convergent evolution of the red- 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* **1990**, 87, 9315-9318.
14. A. B. Asenjo; J. Rim; D. D. Oprian, "Molecular determinants of human red/green color discrimination." *Neuron* **1994**, 12, 1131-1138.
15. T. Chan; M. Lee; T. P. Sakmar, "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* **1992**, 267, 9478-9480.
16. T. Yoshizawa, "The road to color-vision-structure, evolution and function of chicken and gecko visual pigments." *Photochemistry and Photobiology* **1992**, 56, 859-867.
17. J. I. Fasick; N. Lee; D. D. Oprian, "Spectral tuning in the human blue cone pigment." *Biochemistry* **1999**, 38, (36), 11593-11596.
18. J. L. Spudich; C. S. Yang; K. H. Jung; E. N. Spudich, "Retinylidene proteins: Structures and functions from archaea to humans." *Annual Review of Cell and Developmental Biology* **2000**, 16, 365-392.

19. C. Pande; A. Pande; K. T. Yue; R. Callender; T. G. Ebrey; M. Tsuda, "Resonance raman-spectroscopy of octopus rhodopsin and its photoproducts." *Biochemistry* **1987**, *26*, 4941-4947.
20. T. Kitagawa; M. Tsuda, "Resonance raman-spectra of octopus acid and alkaline metarhodopsins." *Biochimica Et Biophysica Acta* **1980**, *624*, 211-217.
21. W. Stoeckenius; R. A. Bogomolni, "Bacteriorhodopsin and Related Pigments of Halobacteria." *Annual Review of Biochemistry* **1982**, *51*, 587-616.
22. B. Ehrenberg; A. Lewis; T. K. Porta; J. F. Nagle; W. Stoeckenius, "Exchange kinetics of the Schiff-base proton in bacteriorhodopsin." *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1980**, *77*, (11), 6571-6573.
23. J. Favrot; C. Sandorfy; D. Vocelle, "Infrared study of basicity of non-aromatic imines-relation to visual pigments." *Photochemistry and Photobiology* **1978**, *28*, (2), 271-272.
24. J. Herzfeld; S. K. Dasgupta; M. R. Farrar; G. S. Harbison; A. E. McDermott; S. L. Pelletier; D. P. Raleigh; S. O. Smith; C. Winkel; J. Lugtenburg; R. G. Griffin, "Solid-state C-13 NMR-study of tyrosine protonation in dark-adapted bacteriorhodopsin." *Biochemistry* **1990**, *29*, (23), 5567-5574.
25. G. S. Harbison; S. O. Smith; D. P. Raleigh; J. E. Roberts; J. A. Pardo; S. K. D. Gupta; J. Lugtenburg; R. A. Mathies; J. Herzfeld; R. G. Griffin, "High-resolution solid-state NMR-studies of bacteriorhodopsin." *Biophysical Journal* **1986**, *49*, (2), A180-A180.
26. L. Ren; C. H. Martin; K. J. Wise; N. B. Gillespie; H. Luecke; J. K. Lanyi; J. L. Spudich; R. R. Birge, "Molecular mechanism of spectral tuning in sensory rhodopsin II." *Biochemistry* **2001**, *40*, (46), 13906-13914.
27. C. Yuan; O. Kuwata; J. Liang; S. Misra; S. P. Balashov; T. G. Ebrey, "Chloride binding regulates the Schiff base pKa in gecko P521 cone-type visual pigment." *Biochemistry* **1999**, *38*, 4649-4654.



28. M. Nakagawa; T. Iwasa; S. Kikkawa; M. Tsuda; T. G. Ebrey, "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* **1999**, *96*, 6189-6192.
29. S. Nishimur; H. Kandori; M. Nakagawa; M. Tsuda; A. Maeda, "Structural dynamics of water and the peptide backbone around the Schiff base associated with the light-activated process of octopus rhodopsin." *Biochemistry* **1997**, *36*, 864-870.
30. P. E. Blatz; J. H. Mohler; H. V. Navangul, "Anion-induced wavelength regulation of absorption maxima of Schiff-bases of retinal." *Biochemistry* **1972**, *11*, 848-856.
31. M. Sigh; H. R. Cama, "Enzymatic cleavage of carotenoids." *Biochimica Et a. Biophysics Acta* **1974**, *370*, 49-61.
32. R. C. Mordi; J. C. Walton; G. W. Burton; L. Hughes; K. U. Ingoid; D. A. Lindsay; D. J. Moffatt, "Oxidative degradation of β -carotene and β -apo-8'-carotenal." *Tetrahedron* **1993**, *49*, 911-928.
33. A. C. Ross, "Cellular metabolism and activation of retinoids: roles of cellular retinoid-binding proteins." *FASEB Journal* **1993**, *7*, 317-327.
34. T. R. Evans; S. B. Kaye, "Retinoids: present role and future potential." *British Journal of Cancer* **1999**, *80*, 1-8.
35. S. A. Kliewer; J. M. Lehmann; T. M. Willson, "Orphan nuclear receptors: shifting endocrinology into reverse." *Science* **1999**, *284*, 757-760.
36. G. M. Morris-Kay; S. J. Ward, "Retinoids and mammalian development." *International Reviews in Cytology* **1999**, *188*, 73-131.
37. D. Koch; W. Gartner, "Steric hindrance between chromophore substituents as the driving force of rhodopsin photoisomerization: 10-methyl-13-demethyl retinal containing rhodopsin." *Photochemistry and Photobiology* **1997**, *65*, 181-186.

38. G. G. Kochendoerfer; P. J. E. Verdegem; I. v. d. Hoef; J. Lugtenburg; R. A. Mathies, "Retinal analog study of the role of steric interactions in the excited state isomerization dynamics of rhodopsin." *Biochemistry* **1996**, *35*, 16230-16240.
39. R. S. H. Liu; A. E. Asato; M. Denny; D. Mead, "The nature of restrictions in the binding site of rhodopsin. A model study." *Journal of the American Chemical Society* **1984**, *106*, 8298-8300.
40. J. Uenishi; R. Kawahma; O. Yonemitsu; A. Wada; M. Ito, "Stereocontrolled synthesis of (11Z)-retinal and its analogues." *Angewandte Chemie International Edition* **1998**, *37*, (3), 320-323.
41. B. Borhan; M. L. Souto; J. M. Um; B. Zhou; K. Nakanishi, "Efficient synthesis of 11-cis-retinoids." *Chemical European Journal* **1999**, *5*, (4), 1172-1175.
42. A. W. Gibson; G. R. Humphrey; D. J. Kennedy; H. Stanley; S. H. B. Wright, "Diethyl (3-Trimethylsilyl-2-propynyl)phosphonate, a new reagent for the preparation of terminal conjugated enynes." *Synthesis* **1991**, 414-416.
43. F. Sato; Y. Kobayashi, "Hydromagnesiation reaction of propargylic alcohols: (E)-3-pentyl-2-nonene-1,4-diol from 2-octyn-1-ol." *Organic Synthesis* **1990**, *69*, 106.
44. W. Boland; N. Schroer; C. Sieler; M. Feigel, *Helvetica Chimia Acta* **1987**, *70*, 1025-1046.
45. *Journal of the Chemical Society* **1952**, 1094-1111.
46. P. E. Blatz; J. H. Mohler; W. Ahmed, "Spectroscopic observation of solvent Interaction with selected retinal Schiff-bases." *Photochemistry and Photobiology* **1991**, *54*, 255-264.
47. G. A. J. Pitt; F. D. Collins; R. A. Morton; P. Stok, "Rhodopsin. VIII. N-Retinyldenemethylamine, an indicator yellow analog." *Biochemical Journal* **1955**, *59*, 122-128.

48. P. E. Blatz; N. Baumgartner; B. Balasubramaniyan; P. Balasubramaniyan; F. Stedman, "Wavelength regulation in visual pigment chromophore. Large induced bathochromic shifts in retinol and related polyenes." *Photochemistry and Photobiology* **1971**, *14*, 531-549.
49. M. D. Distefano; H. Kuang; D. F. Qi; A. Mazhary, "The design of protein-based catalysis using semisynthetic methods." *Current Opinion in Structure Biology* **1998**, *8*, 459-465.
50. H. Kuang; D. Haring; D. F. Qi; A. Mazhary; M. D. Distefano, "Synthesis of a cationic pyridoxamine conjugation reagent and application to the mechanistic analysis of an artificial transaminase." *Bioorganic Medicinal Chemical Letters* **2000**, *10*, 2091-2095.
51. H. Kuang; M. D. Distefano, "Catalytic enantioselective reductive amination in a host-guest system based on a protein cavity." *Journal of the American Chemical Society* **1998**, *120*, 1072-1073.
52. R. R. Davies; M. D. Distefano, "A semisynthetic metalloenzyme based on a protein cavity that catalyzes the enantioselective hydrolysis of ester and amide substrates." *Journal of the American Chemical Society* **1997**, *119*, 11643-11652.
53. R. Y. Tsien, "The green fluorescent protein." *Annuals Reviews in Biochemistry* **1998**, *67*, 509-544.
54. J. M. Tavaré; L. M. Fletcher; G. I. Welsch, "Review-Using green fluorescent protein to study intracellular signalling." *Journal of Endocrinology* **2001**, *170*, 297-306.
55. N. Billinton; A. W. Knight, "Seeing the wood through the trees: A review of techniques for distinguishing green fluorescent protein from endogenous autofluorescence." *Analytical Biochemistry* **2001**, *291*, 175-197.
56. P. Roessel; A. H. Brand, "Imaging into the future: visualizing gene expression and protein interactions with fluorescent proteins." *Nature Cellular Biology* **2002**, *4*, E15-E20.

57. A. Chiesa; E. Rapizzi; V. Tosello; P. Pinton; M. d. Virgilio; K. E. Fogarty; R. Rizzuto, "Recombinant aquorin and green fluorescent protein as valuable tools in the study of cell signalling." *Biochemical Journal* **2001**, 355, 1-12.
58. B. P. Cormack; R. H. Valdivia; S. Falkow, "FACS-optimized mutants of the green fluorescent protein (GFP)." *Gene* **1996**, 173, 33-38.
59. G. J. Phillips, "Green fluorescent protein-a bright idea for the study of bacterial protein localization." *FEMS Microbiology Letters* **2001**, 204, 9-18.
60. R. Heim; D. D. Prasher; R. Y. Tsien, "Wavelength mutations and postranslational autoxidation of green fluorescent protein." *Proceedings of the National Academy of Sciences of the United States of America* **1994**, 91, 11984-11989.
61. G. S. Baird; D. A. Zacharias; R. Y. Tsien, "Biochemistry, mutagenesis , and oligomerization of DsRed, a red fluorescent protein from coral." *Proceedings of the National Academy of Sciences of the United States of America* **2000**, 97, 11984-11989.
62. C. T. Baumann; C. S. Lim; G. L. Hager, "Simultaneous visualization of the yellow and green forms of the green fluorescent protein in living cells." *Journal of Histochemistry and Cytochemistry* **1998**, 46, 1073-1076.
63. M. Chalfie; Y. Tu; G. Euskirchen; W. W. Ward; D. C. Prasher, "Green fluorescent protein as a marker for gene-expression." *Science* **1994**, 263, 802-805.
64. R. Hein; R. Y. Tsien, "Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer." *Current Biology* **1996**, 6, (178-182).
65. A. Mishra; R. K. Behera; P. K. Behera; B. K. Mishra; G. B. Behera, "Cyanines during the 1990s: A Review." *Chemical Reviews* **2000**, 100, 1973-2001.
66. L. Viteva; e. al, "Organometallics in cyanine chemistry-synthesis, reactivity and photophysical properties for some heptamethine

merocyanine dyes." *European Journal of Organic Chemistry* **2004**, 385-394.

67. R. S. H. Liu; A. E. Asato, "Tuning the color and excited state properties of the azulenic chromophore: NIR absorbing pigments and materials." *Journal of Photochemistry and Photobiology, C: Photochemistry Reviews* **2003**, 4, (3), 179-194.
68. R. S. H. Liu; R. S. Muthyala; X.-s. Wang; A. E. Asato; P. Wang; C. Ye, "Correlation of substituent effects and energy levels of the two lowest excited states of the azulenic chromophore." *Organic Letters* **2000**, 2, (3), 269-271.
69. S. Ito; A. Nomura; N. Morita; C. Kabuto; H. Kobayashi; S. Maejima; K. Fujimori; M. Yasunami, "Synthesis and two-electron redox behavior of diazulenol[2,1-a,2-c]naphthalenes." *Journal of Organic Chemistry* **2002**, 67, 7295-7302.
70. R. Muthyala; D. Watanabe; A. E. Asato; R. S. H. Liu, "The nature of the delocalized cations in azulenic bacteriorhodopsin analogs." *Photochemistry and Photobiology* **2001**, 76, (6), 837-845.
71. H. Dorothee; S. Siegfried; G. Wolfgang; B. Volker; M. Hans-Dieter, "Merocyanines as extremely bathochromically absorbing chromophores in the halobacterial membrane protein bacteriorhodopsin." *Angewandte Chemie International Edition* **1997**, 36, (15), 1630-1633.
72. L. C. Wang; Y. Li; H. G. Yan, "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* **1997**, 272, 1541-1547.
73. A. W. Norris; L. Cheng; V. Giguere; M. Rosenberger; E. Li, "Measurement of subnanomolar retinoic acid-binding affinities for cellular retinoic acid-binding proteins by fluorometric titration." *Biochimica Biophysica Acta-Protein Structure and Molecular Enzymology* **1994**, 1209, 10-18.
74. M. J. Mintz; C. Walling, "t-Butyl hypochlorite." *Organic Synthesis Collection* **5**, 184-187.

75. F. Lambertin; M. Wende; M. J. Quirin; M. Taran; B. Delmond, "New retinoid analogs from γ -pyronene a natural synthon." *European Journal of Organic Chemistry* **1999**, 1489-1494.
76. D. C. Lee; J. A. Hayward; C. J. Restall; D. Chapman, "2nd-Derivative infrared spectroscopic studies of the secondary structures of bacteriorhodopsin and Ca-2+-ATPase." *Biochemistry* **1985**, 24, (16), 4364-4373.
77. A. C. Sehgal; R. Tompson; J. Cavanagh; R. M. Kelly, "Structural and catalytic response to temperature and cosolvents of carboxylesterase EST1 from the extremely thermoacidophilic archaeon *Sulfolobus solfataricus* P1." *Biotechnology and Bioengineering* **2002**, 80, (7), 784-793.
78. M. P. Horvath; R. A. Copeland; M. W. Makinen, "The second derivative electronic absorption spectrum of cytochrome c oxidase in the Soret region." *Biophysical Journal* **1999**, 77, (3), 1694-1711.

Chapter 7

Attempts to develop an assay to identify rhodopsin surrogates

7.1 Introduction

As discussed in Chapter 4 development of a rhodopsin mimic that formed a protonated Schiff Base with all-*trans*-retinal was accomplished. In order to re-engineer CRABP II into a rhodopsin protein mimic, an active site Lys residue was necessary for the formation of a Schiff base. *In silico* mutagenesis^{1,2} and minimizations based on the published crystal structure of Cellular Retinoic Acid Binding Protein II (CRABP II) with retinoic acid led to position 132 as the ideal choice for the installation of the Lys residue. As discussed in detail by Rachel M. Crist (Michigan State University, Thesis). The use of rational design has provided a number of good protein mimics.

However, the rate at which successful mutants were obtained was time-consuming. For example, in addition to incorporating Lys at position 132, which provided the best rhodopsin surrogates, introduction of a Lys at position 134 seems promising given that



Figure 7-1. Different positions to possibly insert the nucleophilic lysine CRABP II wild type •R132, •Y134 and •T54.

this residue is the closest to the retinal carbonyl carbon. Also Lys at position 134 will approach the carbonyl in a different trajectory than Lys at position 132. Lys 134 approaches in a more head-on trajectory. Thus, it is possible that at least a few different positions could be potentially good to position the nucleophilic lysine. Also, in the progress of this project it was observed that a counter anion was important to stabilize the PSB, and the position Glu121

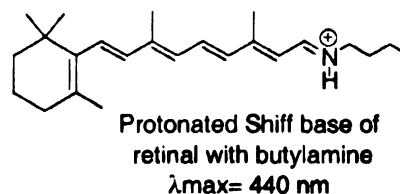
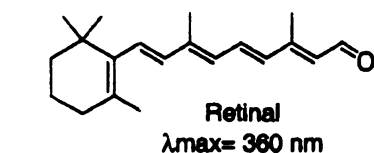


Figure 7-2. Retinal $\lambda_{\text{max}} = 360$ in ethanol, yellow. Protonated Schiff base of retinal with butylamine has a $\lambda_{\text{max}} = 440 \text{ nm}$ in EtOH.

Therefore, we would like to use random mutagenesis³⁻⁵ to increase the number of different mutants. In more detail, error-prone PCR is a random mutagenesis technique for introducing amino acid changes into proteins. Mutations are deliberately introduced during PCR through the use of error-prone DNA polymerases and reaction conditions.⁶ Randomized DNA sequences are cloned into expression vectors and a large mutant library is obtained. When random mutagenesis is used a vast number of different mutants can be obtained. Because of the generation of such large number of different mutants with this technique, the necessity of an assay to efficiently identify the desired mutants (in our case, mutants that form a protonated Schiff base vs. the mutants that do not) is essential. As a result, we attempted to develop an assay for the identification of CRABP II mutants that form a protonated Schiff base with retinal. The assay design exploits

the difference in UV-vis absorptions of retinal and the protonated Schiff base of retinal. The wavelength absorption for retinal in ethanol is 360 nm, and the protonated Schiff base is > 440 nm (Figure 7-2).⁷

7.2 *In vitro* test by using pORANGE, pBCDOX-1 and pCRABPII

As previously discussed in Chapter 2, we have in hand a system that can produce retinal in *E. coli*. This system consists a of XL1-Blue host that contains a plasmid producing β -carotene (pORANGE)^{8,9} and a plasmid that can oxidize β -carotene to retinal (pBCDOX-1)¹⁰ (Figure 7-3). When no IPTG is added the

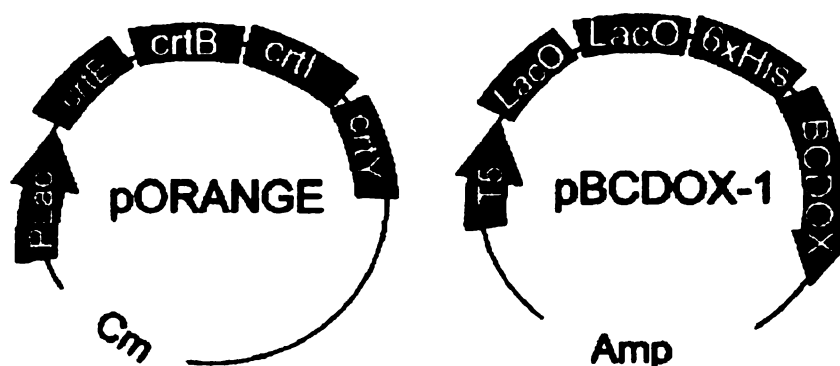


Figure 7-3. pORANGE plasmid contains a gene cluster with crtE, crtB crtI and crtY. These genes produce β -carotene from dimethyl allyl pyrophosphate and isopentyl pyrophosphate. pBCDOX-1 is a vector contains a mice gene for β -carotene-15-15' dioxygenase cloned in a pQE30.

phenotype of this bacteria is an orange color. When the BCDOX is expressed (addition of IPTG, see details in Chapter 2) the phenotype is yellow due to the conversion of β -carotene to retinal, (Figure 7-4). Thus, a host that contains a pORANGE, pBCDOX and pCRABPII could be used as a system to identify good mutants (Figure 7-5). When retinal is produced in the presence of a mutant that forms a protonated Schiff base a shift in the color would be observed, from yellow (380 nm) to red (> 440 nm). Thus, by the difference in colors (yellow vs. red) the mutants that are successful in forming a protonated Schiff base could be identified (Figure 7-6).

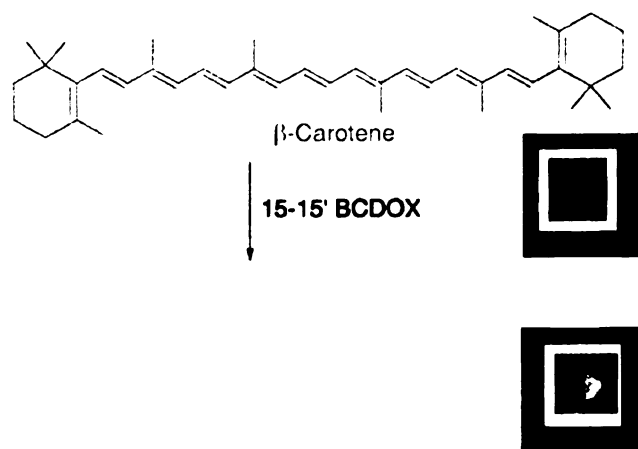
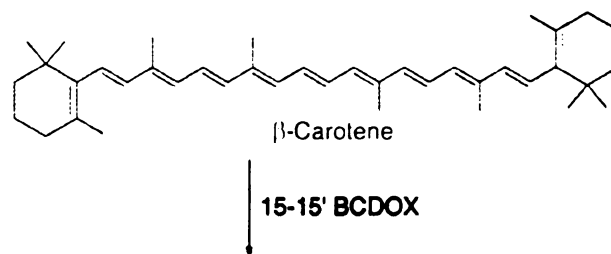
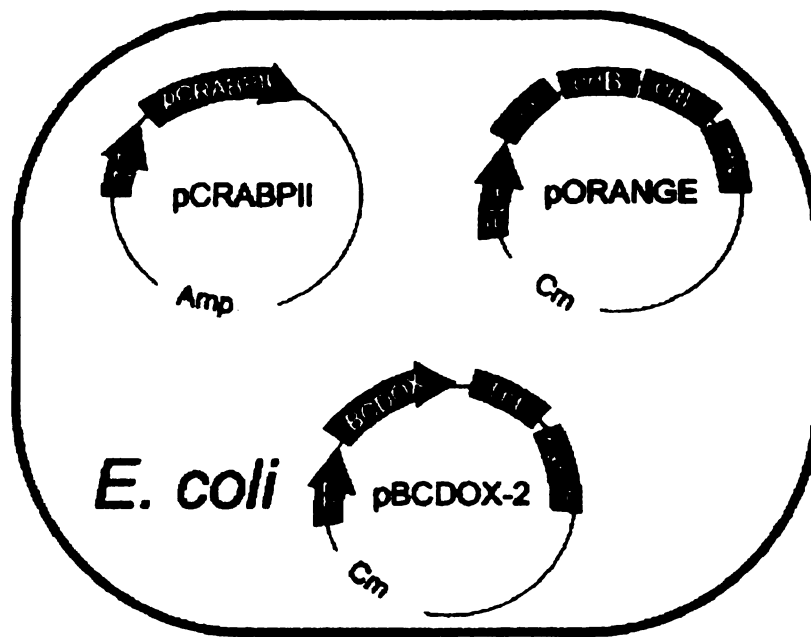


Figure 7-4. Production of retinal *in vivo*. 1. XL1-Blue transformed with pORANGE and pBCDOX-1 with out induction (addition of IPTG). Orange color due to formation of β -carotene 2. XL1-Blue transformed with pORANGE and pBCDOX-1 after addition of IPTG. The yellow coloration is due to oxidation of β -carotene to form retinal.



15-15' BCDOX

Formation of a protonated Schiff base between CRABPII mutants and retinal

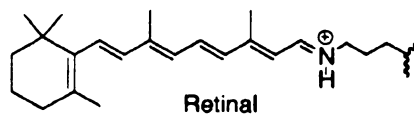


Figure 7-5. Production of retinal *in vivo* could be accomplished by cloning pORANGE and pBCDOX-2 in a host *E. coli*. When the host contains a plasmid that expressed a mutant that forms a Schiff base a change in the color should be observed.

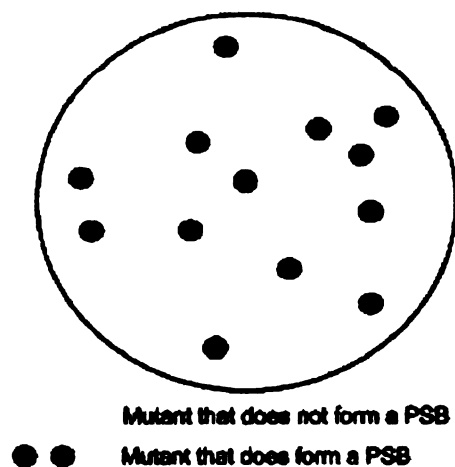


Figure 7-6. The colonies that are yellow would correspond to mutants that do not form a protonated Schiff base. On the other hand the red colonies should contain mutants that form a protonated Schiff base with retinal.

The expression of CRABP_{II} is under the control of a T7 promoter and under the selection of chloramphenicol. The *E. coli* system described in Chapter 2 uses XL1-Blue as the host. In the development of this assay we could not use this host because the CRABP_{II} gene expression is under the control of a T7 promoter. XL1-Blue does not have the capacity of producing the T7 polymerase therefore the CRABP_{II} gene would not be expressed. Thus a different host, ER2566 was used, that contains a copy of the T7 polymerase gene in its genome (Figure 7-7). After the *E. coli* was transformed with pORANGE, it was grown under red safe light. The obtained colonies did not show the expected orange coloration due to the production of β -carotene in the *E. coli*. The reason why the ER2566 *E. coli*

does not produce β -carotene is unclear. As discussed previously when XL1-Blue host was used the coloration of the cells indicated the production of β -carotene, and a comparison of both genotypes (XL1-Blue and ER2566) does not provide an obvious reason of why the production of β -carotene in ER2566 is so low or non-existent.

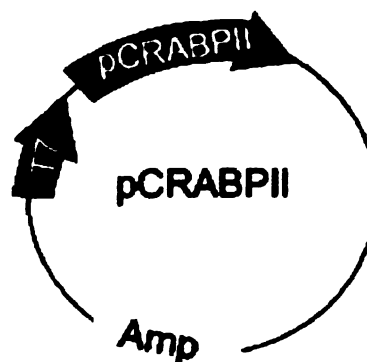


Figure 7-7. CRABP11 is cloned into a pET-17b plasmid under a T7 polymerase control.

This is not the first time that this kind of results are obtained. Wurtzel and colleagues had observed that many unpigmented colonies were obtained among the different hosts.⁸ Also, they noticed that no obvious correlation between genetic markers and carotenoid pigmentation exist. But they did suggest that it is likely that unidentified markers affect pigment accumulation. On the other hand, Sandmann¹¹ observed that when zeaxanthin was produced using a plasmid that was under the control of a T7 promoter in BL21(DE3) *E. coli*, no expression was observed. But when a JM101 strain (that contained a mGP1-2 plasmid that expressed T7 RNA polymerase) was used, production of the carotenoid was observed.

Different groups have expressed carotenoids in a variety of *E. coli* and the production of carotenoids was also variable (Table 7-1).^{8,9,12-18} In general it has been observed that the presence of a high copy plasmid (puck) the production of carotenoids decreases. The deprived production of carotenoid could be due to the

Table 7-1. Expression of carotenoid genes from *E. uredoovora* in different *E. coli*. strain. Lycopene (L), zeaxanthin (Z), β -carotene (BC).

<i>E. coli</i> strain	Expression
TOP10F'	High (L) (Z)
AB2463	High (L) (Z)
XL1-Blue	High (BC)
JM101	High(BC)
JM109	Medium (BC)
HB101	Medium (L) (Z)
SOLR	Medium (L) (Z)
V73	Medium (L), high (Z)
AB1884	Medium (L) (Z)
AB2480	Medium (L) (Z)
AB1886	Medium (L) (Z)
SK2267	Medium/unstable (L) (Z)
V71	Medium/unstable (L) (Z)
KL168	Medium/ very unstable (L) (Z)
SK3451	Medium/ very unstable
Y1088	Low (L) (Z)
SURE	Low (L) (Z)
BL21(DE3)	Low (BC)

fact that a high-copy number plasmid mediates the synthesis of large amounts of β -lactamase, thus it might exhaust the *E. coli* metabolism in a way that negatively affects the synthesis of carotenoids.

7.3 *In vivo* test for protonated PBS formation using pORANGE pRET-CRABPII

It seems evident that to develop this assay, a gene that is not under the control of a T7 promoter would be optimal for the expression of β -carotene. The new strategy consisted of cloning the CRABPII gene next to the BCDOX gene (pBCDOX-1). As discussed in Chapter 2 the pBCDOX plasmid is under the

control of a T5 promoter, and expression β -carotene and BCDOX were performed using XL1-Blue *E. coli*. After the CRABP_{II} mutant was cloned, the protein expression would also be under the control of T5 promoter allowing the use of XL1-Blue *E. coli* as the host. To prove the validity of this assay the CRABP_{II}-tetra mutant, R132K::Y134F::R111L::L121E, (this was the best mutant developed so far, for details refer to Rachel Crist's Thesis) was chosen to be cloned next to the BCDOX. As a control experiment the CRABP_{II} wild type was to be cloned next to BCDOX.

A comparison of the spectroscopic data of the KLFE and WT is shown in Figure 7-8. The UV-vis spectra of retinal incubated with the CRABP_{II}-tetra mutant shows maximal absorbances at 378 nm and 438 nm, a 61 nm red shift as compared to the CRABP_{II} wild type at 377 nm. The peak at 438 nm suggests the formation of the protonated Schiff base. We decided that initially we would prefer to express both proteins (BCDOX and CRABP_{II}) in their native form, and not as fused proteins. The reasoning behind this was that BCDOX is a protein that tends to aggregate (see Chapter 2 for details). Therefore addition of CRABP_{II} as a fusion protein would increase the possibility of the expression of an insoluble protein. When the primers to clone CRABP_{II} were designed, a stop codon after the BCDOX gene was added, and a different ribosome-binding site for CRABP_{II} was introduced. The same RBS that was present in the previous plasmids 8 bp far away from the starting ATG was used. The same RBS was used for the CRABP_{II}-wild type and CRABP_{II}-tetra mutant. The CRABP_{II}-wild type and CRABP_{II}-tetra mutant were PCR out from the original pET-17b vector. The

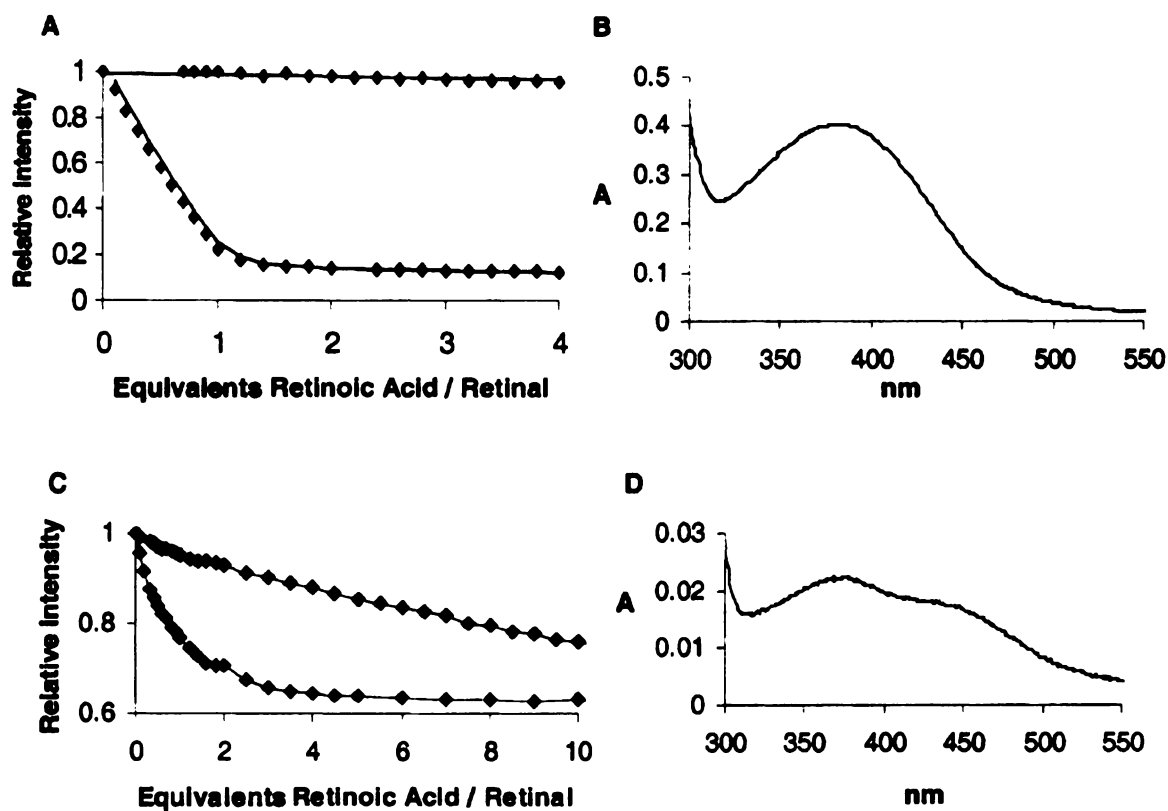


Figure 7-8.. **A.** Fluorescence titration curves for the retinal, $K_d=6600\pm360$ nM.; **B.** UV-vis of retinal incubated with wild-type CRABP II portrays a maximal absorbance at 377 nm. **C** Fluorescence titration curves for the retinoic acid and retinal K_d determination with the tetra mutant CRABP II R132K::Y134F::R111L::L121E. The retinal $K_d=224 \pm 18$ nM. **D.** UV-vis of retinal incubated with the CRABP II protein shows maximal absorbances at 378 nm and 438 nm, an 61 nm red shift as compared to the wild type CRABP II (377 nm), suggesting protonated Schiff base formation.

primers introduced *KpnI* and *Sall* restriction sites (Primer 1 (*KpnI*), 5'-CCCTCTAGAAATAATTTT GGGTACCTTAAGAAGG-3' and Primer 2 (*Sall*) 5'-GTGATGGATATCTGCAGTCGACTCACTCTC-3'). The cloning plan consisted of three steps: PCR, followed by digestion of the genes and plasmid, and finally, ligation of the gene into the desired vector (pBCDOX-1).

The PCR was performed uneventfully and provided the RBS-CRABP_{II}-wt and RBS-CRABP_{II}-tetra mutant gene with the corresponding *KpnI* and *Sall* cut sites. A sequential digestion (with *KpnI* and *Sall*) of the plasmid pBCDOX-1 and the PCR products (RBS-CRABP_{II} and RBS-CRABP_{II}-tetra mutant) were performed. According to the manufacturer a sequential digestion is recommended. The first enzyme used was *Sall* because it requires the highest number of bp next to the cut site. After digestion, purification (using QIAquick gel extraction kit) of the correspondent vector and genes was performed. To avoid re-ligation, the digested vector and genes were treated with Calf Intestine Phosphatase. The DNAs were quantified and the ligations were performed at different molar ratios. The ligation products were transformed in XL1-Blue. A few colonies were obtained, and DNA was isolated to identify the cloned products. After electrophoretic analysis of the plasmid DNA confirmed the existence of a new gene in CRABP_{II}, the isolated plasmids were sent for sequencing. The sequences of both plasmids indeed confirmed the presence of BCDOX and CRABP_{II} genes in the new vector (Figure 7-9). The orange cells were prepared by transformation of *E. coli* XL1-Blue with pORANGE (cells that produce β -carotene), followed by preparation of the competent cells. These

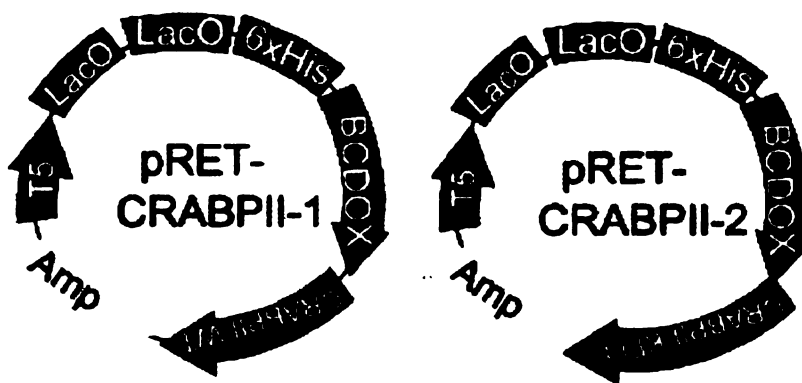


Figure 7-9. The CRABPII genes were cloned into pBCDOX-1. The plasmid pRTE-CRABPII-1 contains the mice BCDOX in frame with the CRABPII wild type. pRTE-CRABPII-2 contains the mice BCDOX in frame with the CRABPII R132K::Y134F::R111L::L121E. Both plasmids are under the control of T5 promoter.

competent cells were transformed with pRET-CRABPII-1 and pRET-CRABPII-2. Digestion of the isolated plasmids was performed to confirm the existence of pRET-CRABPII vectors and pORANGE (Figure 7-10).

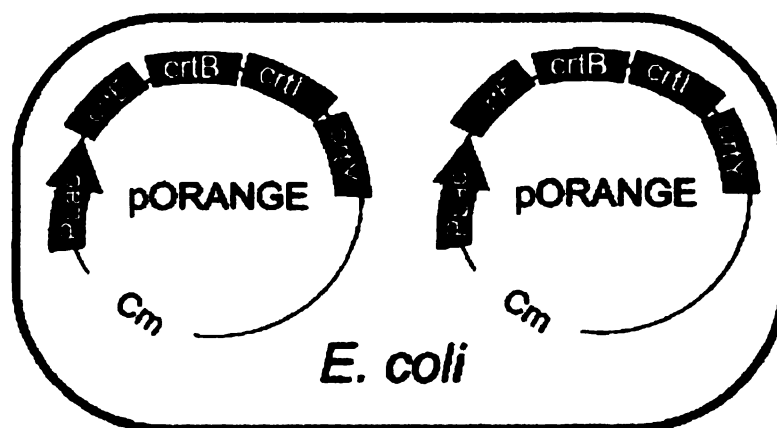


Figure 7-10. Transformation of pRET-CRABP1 or 2 in to the *E. coli* host that produced β -carotene (XL1-Blue-pORANGE).

The cells were grown at 37 °C for about three hours under red safe light, and the production of retinal was induced by addition of IPTG. The results (Table 7.2) showed no obvious difference in the coloration between the cells that contained pBCDOX1 and CRABPII-wt and the cells that contained pBCDOX1-CRABPII-tetra mutant. The cells that contained the pORANGE plasmid did produce the orange coloration due to the presence of β -carotene. When the BCDOX and CRABPII wild type were induced the expected shift in color is

Table 7.2. The developed assay to identify CRABPII mutants that form PSB with retinal did not show the expected results. XL1-Blue was transformed with pORANGE and pRET-CRABPII-1 or pORANGE and pRET-CRABPII-2. The orange coloration was due to the accumulation of β -carotene (pORANGE). The yellow coloration is due to the presence of retinal and but with out the formation of a PBS. The coloration between the wild type and the tetra mutant was the same.

	Expression system	IPTG	Color
1	XL1-Blue/pORANGE	x	orange
2	XL1-Blue/ pORANGE /pRET-CRABPII-1	✓	yellow
3	XL1-Blue/ pORANGE /pRET-CRABPII-1	✓	yellow
4	XL1-Blue/pORANGE	x	orange
5	XL1-Blue/ pORANGE /pRET-CRABPII-2	✓	yellow
6	XL1-Blue/ pORANGE /pRET-CRABP	✓	yellow



observed (a change from orange to yellow). The yellow coloration was due to the formation of retinal not binding as a PSB. However, when the BCDOX and CRABP_{II}-tetra mutant were induced, the expected red shift coloration was not observed. Instead a yellowish color was obtained. Thus no obvious difference in the coloration between the cells that contained pRET-CRABP_{II}-1 and pRET-CRABP_{II}-2 was found.

Since BCDOX is a very sensitive enzyme and could become inactive very easily, extraction of the retinoids was performed. Analysis of the extracted retinoids and carotenoids verified the presence of retinal and β -carotene. The presence of retinal confirms the activity of BCDOX. Due to the fact that the T5 promoter is such a weak promoter, it is possible that the CRABP_{II} protein was not expressed. The strength at which a protein is expressed is proportional to how far the promoter is located from the gene's RBS. In this case the second RBS for CRABP_{II} is about 1.7 KDa bp from the promoter, therefore it is reasonable to believe that probably CRABP was not expressed as expected. To test whether the CRABP_{II} mutants were expressed, a large-scale expression of the XL1-Blue pORANGE-/pRET-CRABP_{II}-1 or pORANGE-/pRET-CRABP_{II}-2 was performed. An SDS gel showed no over expressed protein at about 15 KDa (the CRABP_{II} approximate molecular weight). This could be explained due to the fact that T5 promoter is not as strong as a T7 promoter, thus its expression could be low. Next, an attempt to isolate the expressed CRABP_{II} using Q SepharoseTM was performed. CRABP_{II} has a very high affinity to Q SepharoseTM Fast Flow resin, but no binding to the resin was observed. This was

probably due to the fact that if CRABP_{II} was expressed it was in very low concentrations. To improve the sensitivity of this test, isolation using the FPLC was performed. Conditions to selectively isolate CRABP_{II} mutants using the FPLC have been previously established by Chryssoula. The FPLC elution showed a protein eluting about the same time as CRABP_{II}. However when an acrylamide gel of the eluted protein was performed, only a high molecular protein was identified around 66 kDa. Some times in the purification process of CRABP_{II} mutants some times a protein is eluted at the same time as the expected CRABP_{II}, thus it could be possible that the eluted protein was not CRABP_{II}, instead it might have been just a contaminant. Even though the T5 promoter is a weak promoter, in bacteria, the expression of several genes under the control of a single promoter is very common in the so-called clusters, thus it is a little puzzling why CRABP_{II} is not produced. The lack of difference in color between the CRABP_{II}-native plasmid and the CRABP_{II}-mutant could also be due to the fact that retinoids, as well as carotenoids are stored in the membrane (hydrophobic regions of the cell), and therefore the chromophore does not have the opportunity of binding to CRABP_{II} protein. It is important to keep in mind that as shown in Figure 7-7 there is an equilibrium between the PSB (440 nm) and the non protonated form. Thus, it could be that the amount of PSB is not enough to cause a high number of CRABP_{II} mutants that form a protonated Schiff base, therefore the coloration caused by these mutants would be almost negligible. It is clear, however, that the use of a pQE30 plasmid to express BCDOX and CRABP_{II} does not provide the desired results.

Unfortunately, as previously discussed, the expression of β -carotene in *E. coli* cells that produce the T7 polymerase was impossible. Only in XL1-Blue or JM109 *E. coli* hosts the β -carotene was accumulated in significant amounts. Therefore, we considered inserting the T7 polymerase gene into XL1-Blue genome. But the commercially available systems to infect cells with T7 polymerase cannot be used in XL1-Blue hosts. Bacteriophage CE6 can be used to provide a source of T7 RNA polymerase to susceptible host cells carrying pET recombinant plasmids.¹⁹ But usually after infection with the bacteriophage the cells go to lysogenic growth, thus are destroyed in about 3-4 hours. In order to avoid the lysogenic growth the *E. coli* should express *supF* and the XL1-Blue strains doesn't express *supF*, and neither does JM109. As a result, this experiment cannot be used as an assay for the screening of mutants that form protonated Schiff base with retinal.

7.4 Materials and Methods.

A. Plasmid Purification

Bacteria from a single colony was grown (500 mL, LB, containing appropriate antibiotics) and purified with via Qiagen[®] column purification. Upon completion of the Qiagen[®] purification, the recovered DNA was resuspended in (400 µL) sterile water, and analyzed by UV-visible for concentration determination and purity. The concentration of DNA obtained is µg/µL

Concentration of DNA sample (µg / µL) =

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

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

= 1.8, pure

>1.8, RNA contamination

< 1.8, protein contamination

B. Sample preparation for sequencing DNA

In a sterilized eppendorf tube (0.5 mL), DNA (2 µg) and primer (30 pmol) were mixed. When sequencing is performed, the largest amount of base pairs that are sequenced with accuracy is about 400 bp. Therefore, to sequence the complete gene, five different primers were used.

1. 5'-GATCTCGATCCCGCGAAATTAATACGAC-3'
2. 5'-CCAGACCCTAGAGACCTTGGAGAAGG-3'

3. 5'-CGAGGAGAAGTCCAGGCTGACC-3'
4. 5'-GATCGATCTCGATCCCGCG-3'
5. 5'-GCAGACTGGAATGCAGTGAAGC-3'
6. 5'-GCAGACTGGAATGCAGTGAAGC-3'

C. Melting temperature calculation for primers

Primers should ideally have melting points ≥ 78 °C, and end in at least one, if not more, GC base pair

$$T_m = 81.5\text{ °C} + (0.41)(\% \text{ GC}) - 675 / N - \% \text{ Mismatch}$$

where N = primer length

Optimal PCR conditions. A small scale gradient PCR was performed to optimize the extension temperature.

PCR recipe	
Template DNA	100 ng
Primer 1	1 mM
Primer 2	1 mM
DNTP	200 μ M
Deep vent	2 U
10 x deep vent buffer	5 μ L (1x)
MgSO ₄	5 μ M
H ₂ O	50-Rxn μ L

PCR program		
1 X	94 °C	5 min
30 X	94 °C	1 min
	48 °C	3 min
	72 °C	3 min
1 X	72 °C	10 min
1 X	25 °C	10 min

D. PCR Primers

	Primers Sequence	PCR Template
pRET-CRABPII-1	5'-CCCTCTAGAAATAATTTTGGGTACCTTAAGAAGG-3' 5'-5'-GTGATGGATATCTGCAGTCGACTCACTCTC-3'	pET17-CRABPII-WT
pRET-CRABPII-2	5'-CCCTCTAGAAATAATTTTGGGTACCTTAAGAAGG-3' 5'-GTGATGGATATCTGCAGTCGACTCACTCTC-3'	pET17-CRABPII-KFLE

E. Purification of DNA from agarose gel

GENCLEAN Turbo Qiagen[®] Protocol: The desired band was excised from the agarose gel. The gel was cut into smaller pieces and transferred into an eppendorf tube. Then a GENCLEAN Turbo solution (100 µL per 100 µg of agarose) was added and the mixture was melted at 55 °C for 5 min. The melted solution was transferred into GENCLEAN Turbo cartridge. The filter was spined for 5 seconds. The filter was washed twice by addition of GENCLEAN Turbo Wash (500 µL). The DNA was eluted using GENCLEAN Turbo solution (30 µL). The DNA obtained was used directly in the enzyme digestions.

F. Digestions and ligation reactions

As mentioned previously the double digestion did not afford the desired product therefore, subsequent single digestions were performed. Same protocols were followed for the gene (PCR product) and the pET29b(+). The first digestion was performed using *Sall* followed by purification using GENCLEAN. The second digestion was performed using *KpnI*.

F.1 Digestions

Digestions				
	CRABPII		pQE30-BCDOX	
<i>KpnI</i>	1 μ L	-	1 μ L	-
<i>Sall</i>	-	1 μ L	-	1 μ L
DNA 30 μ L	30 μ L	30 μ L	30 μ L	30 μ L
Buffer 1 or 2 (10X)	1 X	1X	1 X	1X
H ₂ O	6 μ L	6 μ L	6 μ L	6 μ L

- Incubation at 37 °C for 2 hours

F.2 Calf Intestine Phosphatase treatment

CIP treatment		
	Digested pBCDOX1	Digested CRABPII gene
DNA obtained from DNA purification	30 μ L	30 μ L
CIP 1 μ L	1U	1U
Buffer (10X)	1 X	1 X
H ₂ O	4 μ L	4 μ L

- Incubation at 37 °C for 2 hours
- The DNA was purified following the same procedure and quantified before the ligations were performed

F.3 Ligations

Ligations	
	1/10
Gene	4.5x10 ⁻¹³ moles
Vector	4.5x10 ⁻¹⁴ moles
T4 ligase 1 μ L	1U
Buffer (10X)	1 X
H ₂ O	

- Gene 1= CRABPII-WT + *KpnI* and *Sall*
- Vector =pET30-BCDOX + *KpnI* and *Sall*
- Incubation at 16 °C for 24 hours
- JM109 *E. coli* were transformed with 5 μ L of the ligation reactions

G. Typical preparation of competent cells

The *E. coli* strain of interest was grown (37 °C until OD₆₀₀ of 0.4-0.6, the media contained LB, the corresponding antibiotic). After about three h The cells were harvested by centrifugation at 5,000 RPM for 10 min at 4 °C. The cells were re-suspended (100 mL of 0.9% NaCl) and then centrifugated at 5,000 RPM for 10 min at 4 °C. The cells were re-suspended (50 mL of 100 mM CaCl₂) and incubated for 30 min at 0 °C. The cells were centrifuged at 5,000 RPM for 10 min at 4 °C, and the cells were re-suspended (4 mL of 100 mM CaCl₂, 15% glycerol). The cells in suspension were aliquoted (0.1 mL) and were quickly frozen with liquid nitrogen.

H. Typical transformation of competent cells

Sterile conditions were maintained throughout the entire protocol. The competent cells were incubated at 0 °C for 5 min and the plasmid DNA (10 ng/μL) was added. The mixture was incubated for 30 min and the cells were heat shocked at 42 °C for 2 min at 0 °C. LB (antibiotic if necessary) was added and the cells were incubated at 37 °C for 1 h. Then the cells were transfered to a plate containing LB/antibiotic and were grown overnight at 37 °C.

I. Expression of BCDOX and CRABPII *in vivo*

The expression of BCDOX and CRABPII in *E. coli* was monitored *in vivo* by the colorimetric change of the oxidation of β-carotene (orange) to retinal

(yellow) to retinal-CRABP_{II} red. In detail an *E. coli* XL1-Blue was transformed with pORANGE (plasmid that contains the genes necessary for the synthesis of carotene). The transformed cells were grown and they showed an orange phenotype. These cells were converted into competent cells and then transformed with the pRET-CRABP_{II}-1 or pRET-CRABP_{II}-2 plasmid. These transformed cells (Cells that contained both plasmids) were grown under LB/chloramphenicol, carbenicillin and 1% glucose (5 mL), to an OD of 1.00. Then expression of BCDOX and CRABP_{II} was induced by addition of IPTG (1 mM final concentration). This culture was grown for 6 more hours. The cells were spun down at 10000 RPM and rinsed with water. Analysis of the phenotype of the cells showed that no clear difference could be observed between the cells that expressed the BCDOX and CRABP_{II} to the ones that did not.

7.5 References

1. C. Papworth; J. C. Bauer; J. Braman; D. A. Wright, *Strategies* **1996**, *9*, (3), 3-4.
2. Stratagene, "QuikChange Site -Directed Mutagenesis Kit."
3. R. Fujii; M. Kitaoka; K. Hayashi, "One-step random mutagenesis by error prone rolling circle amplification." *Nucleic Acid Research* **2004**, *32*, (19), 145-150.
4. D. W. Leung; E. Chen; D. W. Goeddel, "A method for random mutagenesis of a defined DBA segment using a modified polymerase chain reaction." *Techniques* **1989**, *1*, 11-15.
5. A. Parikh; F. P. Guengerich, "Random mutagenesis by whole-plasmid PCR amplification." *Biotechniques* **1998**, *24*, (3), 428-431.
6. S. Shafkhani; R. A. Siegel; E. Ferrari; V. Schellenberger, "Generation of a large libraris of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization." *BioTechniques* **1997**, *23*, 304-310.
7. Y. Gat; M. Sheves, "A mechanism for controlling the pKa of the retinal protonated Schiff base in retinal proteins. A study with model compounds." *Journal of the American Chemical Society* **1993**, *115*, 3772-3773.
8. E. T. Wurtzel; G. Valdez; P. Matthews, "Variation in expression of carotenoid genes in transformed *Escherichia coli* strains." *Bioresearch Journal* **1997**, *1*, 1-11.
9. N. Misawa; Y. Stomi; K. Kondo; A. Yokoyama; S. Kajiwara; T. Saito; T. Ohtani; W. Miki, "Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway proposed at the gene level." *Journal of Bacteriology* **1995**, *195*, 6575-6584.

10. J. von Lintig; K. Vogt, "Filling the gap in vitamin A research - Molecular identification of an enzyme cleaving beta-carotene to retinal." *Journal of Biological Chemistry* **2000**, 275, (16), 11915-11920.
11. A. Ruther; N. Misawa; P. Böger; G. Sandman, "Production of zeaxanthin in *Escherichia coli* transformed with different carotenogenic plasmids." *Applied Microbiology and Biotechnology* **1997**, 48, 162-167.
12. G. Sandmann, "Carotenoid biosynthesis and biotechnological application." *Archives of Biochemistry and Biophysics* **2001**, 385, 4-12.
13. P. D. Matthews; E. T. Wrtzel, "Metabolic engineering of carotenoid accumulation in *Escherichia coli* by modulation of the isoprenoid precursor pool with expression of deoxyxylulose phosphate synthase." *Applied Microbiology and Biotechnology* **2000**, 53, 396-400.
14. M. Albrecht; N. Misawa; G. Sandmann, "Metabolic engineering of the terpenoid biosynthetic pathway of *Escherichia coli* for production of the carotenoids carotene and zeaxanthin." *Biotechnology Letters* **1999**, 21, 791-795.
15. M. R. Farmer; J. C. Liao, "Precursor balancing for metabolic engineering of lycopene production in *Escherichia coli*." *Biotechnology Progress* **2001**, 17, (1), 57-61.
16. C. Schmidt-Dannet; D. Umeno; F. Arnold, "Molecular breeding of carotenoid biosynthetic pathways." *Nature Biotechnology* **2000**, 18, 750-753.
17. M. Albrecht; S. Takachi; S. Steiger; Z. Wang; G. Sandman, "Novel hydroxycarotenoids with improved antioxidative properties produced by gene combination in *Escherichia coli*." *Nature Biotechnology* **2000**, 18, 843-846.
18. C. Wang; M. Oh; J. Liao, "Directed evolution of metabolically engineered *Escherichia coli* for carotenoid production." *Biotechnology Progress* **2000**, 16, 922-926.

19. F. Miao; S. K. Drake; D. S. Kompala, "Characterization of gene expression in recombinant *Escherichia coli* cells infected with phage lambda." *Biotechnology Progress* **1993**, 9, 153-159.