

Syracuse University

SURFACE

Chemistry - Dissertations

College of Arts and Sciences

12-2011

Citrate Binding to the Membrane Protein Proteorhodopsin

Farhana F. Syed
Syracuse University

Follow this and additional works at: https://surface.syr.edu/che_etd



Part of the [Chemistry Commons](#)

Recommended Citation

Syed, Farhana F., "Citrate Binding to the Membrane Protein Proteorhodopsin" (2011). *Chemistry - Dissertations*. 182.

https://surface.syr.edu/che_etd/182

This Dissertation is brought to you for free and open access by the College of Arts and Sciences at SURFACE. It has been accepted for inclusion in Chemistry - Dissertations by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

ABSTRACT

Proteorhodopsin (pR) is an intrinsic membrane protein with an important role in solar-energy storage of the biosphere. Earlier work in our lab has shown that polyhistidine-tagged pR can be purified by means of selective precipitation with citrate under specific conditions, as can a number of mutants based on this His₆-tagged pR. Purification of a heterologously-expressed trans-membrane protein by a simple salt such as citrate is novel. However, such a phenomenon leads to several questions: How does citrate cause pR precipitation? Does the polyhistidine-tag assist in such a precipitation? Is this precipitation pH-specific? Does citrate affect the function of pR? Does citrate-induced pR precipitation have any biological significance? Are there other ions that could cause pR precipitation?

This dissertation focuses on understanding the nature of the interaction of pR with citrate and other anions, and in particular on trying to take advantage of this interaction in order to develop a novel membrane protein purification method. The end goal that branches out of these two aims is to utilize the compact citrate interaction site identified in pR, by incorporating it into other membrane proteins and using it to permit their purification by similar simple procedures.

In Chapter 1, I briefly provide some background information on the wide variety of structurally-similar proteins as rhodopsins that include pR. I also describe the general

importance of developing purification methods for 7-helix membrane proteins, including pR.

Chapter 2 focuses on the investigation of the nature of citrate-binding site of pR. To address the main question of how citrate aids in pR purification, site-directed mutagenesis technique was applied to generate several single, double, triple or quadruple mutants of pR in a histidine-tag free background, which were then tested for their reactivity to citrate. Several different anions were tested to examine if precipitation of pR was specific to citrate or whether the precipitation is susceptible to other negatively charged salts. Photocycle of pR progresses through several intermediates, each with a distinct absorption maximum (described in subsection 1.3.1). M-intermediate is detected at $\text{pH} \geq 8$ with $\lambda_{\text{max}} = 410 \text{ nm}$. Flash spectroscopy involves excitation of pR at a particular wavelength that leads to transient absorption, thus, signaling the formation of the corresponding intermediate. Flash-induced transient visible absorption measurements were used to assay the effect of exposing pR to citrate on its physiological function.

Chapter 3 describes the development of a method of purification of pR using simple salts, citrate and phosphate. Chapter 4 begins an exploration of a future direction. The ultimate objective is to apply the above techniques for the general purpose of 7-helix membrane protein purification, especially for the important class of pharmacological receptors known as GPCRs. An attempt at heterologous expression in *E. coli*, and purification, of a mammalian GPCR, is described therein. Such a method would be desirable for obtaining proteins for structural, functional and pharmacological studies.

CITRATE BINDING TO THE MEMBRANE PROTEIN PROTEORHODOPSIN

By

Farhana F. B. Syed

M.Sc., University of Pune, 1996

B.Sc, University of Pune, 1994

Dissertation

Submitted in partial fulfillment of the requirements for the degree of Doctor of
Philosophy in Chemistry in the Graduate School of Syracuse University

December 2011

Copyright© Farhana F. B. Syed 2011

All rights reserved

Table of Contents

Content	Page
1. INTRODUCTION	01
1.1 General background	01
1.1.1 Rhodopsins	01
1.1.2 Microbial rhodopsins	03
1.2. Proteorhodopsin	09
1.2.1 Photocycle of pR	12
1.2.2 Energy transduction and proton pumping or pR	15
1.3 Visual rhodopsins	19
1.4 Biological significance of pR	21
1.4.1 Evolutionary basis	21
1.4.2 pR based phototrophy	21
1.4.3 Energy balance	23
1.5 Variants of proteorhodopsin	27
1.6 Applications of pR	29
1.6.1 Holography and information storage	29
1.6.2 Cosmetic ingredient	32
1.6.3 Bioenergy production	33
1.7 Purification of protein	35
1.7.1 General process of purification of protein	36
1.7.2 Membrane protein purification	37
1.7.3. Salting out: the Hofmeister series	39
1.8 Widely used method of pR purification	43
1.9 Our method of pR purification	43
1.10 Citrate	44
1.11 Citrate does not degrade pR	45
1.12 My Research	46
2. Citrate binding site in proteorhodopsin involves two lysines in the first cytoplasmic loop	47
2.1 Abstract	47
2.2 INTRODUCTION	48
2.3 MATERIALS AND METHODS	50
2.3.1 Protein Expression	50
2.3.2 Protein Purification	51
2.3.3 UV-visible absorbance measurements	53
2.3.4 Site Directed Mutagenesis	53
2.3.5 Test for aggregation with diverse anions	54
2.3.6 Reconstitution of pR into DOPC liposomes	54
2.3.7 Flash spectroscopy and kinetic analysis	55
2.3.8 Homology modeling and molecular graphics imaging	56

2.3.9	Distance Tree of proteorhodopsin related proteins	56
2.4	RESULTS	57
2.4.1	Purification of proteorhodopsin without His-tag	57
2.4.2	Lysines in the first cytoplasmic loop on pR mediate citrate interactions	58
2.4.3	Proteorhodopsin interacts with several anions	63
2.4.4	Model for pR-citrate interaction	68
2.4.5	Citrate precipitates pR at physiological pH	69
2.4.6.	Effect of citrate on pR physiological activity	71
2.4.7	Conservation of the citrate binding site	74
2.5	DISCUSSION	77
2.5.1	Purification of pR TCM without His-Tag	77
2.5.2	pR–citrate interaction and aggregation	78
2.5.3	pR interaction with diverse anions	80
2.5.4	Citrate has no discernible effect on pR reconstituted in a membrane	88
2.6	CONCLUSIONS	89
2.7	Acknowledgements	90
3.	Column-free preparation of proteorhodopsin, an integral membrane protein, by selective precipitation with simple salts	91
3.1	INTRODUCTION	91
3.1.1	Membrane proteins—heterologous expression and purification	91
3.1.2	Use of citrate in proteorhodopsin purification	94
3.1.3	Overview of the development of using phosphate in conjunction with citrate for pR preparation	95
3.2	MATERIALS AND METHODS	96
3.2.1	Protein Expression and Purification using citrate and hydroxylapatite column chromatography	96
3.2.2	Partial purification of pR	97
3.2.3	Preparation of pR using phosphate-citrate buffers in competition	98
3.2.4	Polyacrylamide Gel Electrophoresis	98
3.2.5	Absorption Spectroscopy	98
3.3	RESULTS	99
3.3.1	Phosphate precipitates pR between 50-400 mM phosphate concentration and 0.3 -0.4% OG concentration	99
3.3.2	Determining the minimum phosphate concentration that induces pR precipitation	100
3.3.3	Phosphate precipitation by itself does not eliminate impurities left behind by citrate-induced precipitation	101
3.3.4	Combination of phosphate inhibition with citrate-induced precipitation promotes the selectivity of pR precipitation (Done in collaboration with Jonathan Kim)	103
3.3.5	Phosphate likely binds to pR at the same site as citrate	105
3.3.6	Optimization of purification conditions, using phosphate in combination with citrate	106

3.3.7	Scaled-up test of a simplified citrate/phosphate protocol for preparation of pR	113
3.4	DISCUSSION	118
3.4.1	Phosphate-induced pR precipitation	118
3.4.2	Preparation of pR using citrate-phosphate buffers	121
3.5	Conclusions	121
3.6	Acknowledgements	123
4.	Heterologous expression of bovine rhodopsin, a mammalian rhodopsin, in <i>E. coli</i>	124
4.1.	INTRODUCTION	124
4.2	METHODS AND RATIONALE	129
4.2.1	Construction of chimera plasmid coding for bovine rhodopsin fused to N-terminal leader sequence from pR	129
4.2.2	Transfection of Super-competent <i>E. coli</i> with the chimera Plasmid	134
4.2.3	Transfection of <i>E. coli</i> UT5600 with the chimera plasmid	134
4.3	RESULTS	138
4.3.1	Analysis of the plasmid construct of pR-Bovine rhodopsin chimera	138
4.3.2	Attempt at heterologous expression of chimera protein in <i>E. coli</i>	145
4.4	Follow-up Experiments	146
4.4.1	To detect the expression of chimera protein	146
4.4.2	To detect the location of chimera protein	146
4.4.3	Purification of chimera protein with citrate	148
4.5	Future Experiments	149
4.5.1	Modification of expression experiments	149
4.5.2	Modification of Cell wall lysis	149
4.5.3	Alteration in the size of the leader sequence	149
4.5.4	Stabilization of membrane proteins by Genetic Engineering	149
4.6	Conclusions	150
4.7	Acknowledgement	151
5.	Appendix	152
2.9	UV-visible spectra of pR	152
2.10	Western Blot of pR expressed in <i>E. coli</i>	154
2.11	Flash Photolysis data	155
2.12	pR-citrate aggregation	156
2.13	Interaction of pR with diverse ions	158
2.13.1	Precipitation of pR by Lithium salts	158
2.13.2	Preparation of buffers of Lithium salts	159
2.13.3	Precipitation of pR by salts by potassium salts	161

2.13.4	Precipitation with bromide as the anion	161
3.8	Purity of pR can be predicted by the $A_{280/520}$ ratio	162
3.9	N-terminal sequencing results	164
References		190
1.	References for Chapter 1	190
2.	References for Chapter 2	204
3.	References for Chapter 3	211
4.	References for Chapter 4	214
5.	References for Appendix	216
Vita		217

ACKNOWLEDGEMENT

I thank my advisor, Prof. Mark Braiman, for taking me under his wings when my future was bleak. His confidence in me and constant support were inspiring and encouraging. He let me think independently and at the same time guided me to stay on track, whether it was designing experiments or analyzing data. He was always readily available for discussions or feedback. Besides being a mentor, he has been extremely considerate and compassionate about my limited time-schedule. That motivated me to do even better as a scientist. I am extremely grateful to him for making me a better thinker.

I thank my committee members who were patient and accessible despite their busy schedules. I am thankful to Prof. Robert Doyle for taking time to be a critical reader of my dissertation. I express many thanks to Prof. Michael Cosgrove, the chairman of my committee, for encouraging me through my lows and providing both scientific and moral support. I am grateful to him for letting me use his lab equipment without any hesitation. I would like to thank Prof. Jerry Goodisman, Prof. Yan Yeung Luk and Prof. Philip Borer for their compassion and understanding at various times.

I am also thankful to my previous advisor, Prof. Richard Hallberg. Working in his lab provided me an opportunity to learn about cell biology and yeast genetics. I am grateful and privileged to work with Dr. Matthew Gentry who trained me with all the yeast-related techniques, shared information and also provided a healthy atmosphere for learning and doing science. Thanks to Dr. Yin Yin with whom I share a lot of memories.

I have had an opportunity to work with many bright undergraduate researchers who contributed to my research. Thank you to Jonathan Kim, Alicia Barnes, Farina Mahmud, Chauncy Brown, Soyika Richardson, Euginia Im, Danny Heng and Tiffany Dunston, who have helped me at various times with different projects.

I thank Michael Brandt who designed software so I could convert data into readable form from the flash spectrophotometer. I am also thankful to Prof. Joseph Chaiken who let me borrow his oscilloscope so I could visualize my flash spectroscopy data.

I thank the Department of Chemistry which financially supported me with a TAship through all my graduate student years. I also want to thank the Department of Biology for supporting me with a TAship for one year. Thanks to David Aldinger (Biology) who let me use any lab equipment I needed in the Biology Department.

I am thankful to Valarie Vought and Dr. Anamika Patel who helped me use their lab equipment. Collin Fischer, besides our discussions about technical difficulties, along with Valarie and Anamika, also shared the understanding of mothers trying to balance research. I am also highly appreciative of the Chemistry and Biology office members – Cathy Voorhees, Joyce Lagoe, Deb Maley, Jodi Randall and Deb Herholtz. They took me into their “working mothers club” and expressed great understanding. I will always be grateful to Cathy for thinking about me for any TA- or job-opportunity.

I am fortunate to be born into a family that values education. I am so glad that they encouraged me to take my GRE and then let me fly to a foreign land to get the best education. I will never be able to comprehend the pain my mother must have felt to see her daughter go halfway across the world. My father's unique way of celebrating achievements – short and quick, and then move on to the next goal – has been my lighthouse even when the going got tough. Thanks to all my siblings who constantly showered me with their encouragement and love, and always put me up on a pedestal. Thanks to the prayers my family always sends my way.

Without the love, faith, encouragement and support of my loving husband David, this journey would have been an excruciating ordeal. Our graduate school days have been full of challenges but his unwavering support through the tough times so far is a testament of his loyalty. I am blessed with Sophia (6 years) and Elias (2 years), who have been extremely understanding especially about the weekends when I had to leave them at home. Sophia has extended her help with my lab work – filling tip boxes and taking pictures of my results. I am also thankful to the little baby who co-operated well during my experiments, writing and defense. Zakaria was born 12 days after the defense. I also thank all my in-laws who have accepted me very warmly from the first day I met them. They have helped me in various ways, especially my parents-in-laws who have reduced my pain of being so far away from home.

I also thank the Almighty who gave me strength to endure the hardships and challenges and all the opportunities and reasons to achieve my goal.

Chapter 1

1. INTRODUCTION

1.1 General Background

1.1.1 Rhodopsins

Proteorhodopsin belongs to a large group of proteins, called rhodopsins, found in all domains of life. Rhodopsins are also known as retinylidene proteins, because they consist of the apoprotein opsin and chromophore retinal. They are hepta-helical trans-membrane proteins that form an interior retinal binding pocket in the hydrophobic core of the protein. The apoprotein is reversibly and covalently bound to its cofactor chromophore retinal at a lysine residue via a protonated Schiff base (PSB) linkage as evidenced by the shift in the λ_{\max} . Rhodopsins, although they all share this common architecture, include members of two distinct superfamilies of proteins, with an as-yet uncertain phylogenetic relationship.

The first rhodopsin superfamily is often referred to as type 1 rhodopsins or, alternatively, as the microbial rhodopsins. Some of these trans-membrane proteins have a purely sensory function, apparently including some types of pR [Hillebrecht J. R., *et al.*, 200; Kunz W., *et al.*, 2004; Spudich J. L., 2006]. However, most pR proteins are included among the energy-transducing microbial rhodopsins, which act as light-driven ion pumps. That is, they translocate ions, generally protons or chloride ions, across the

membrane generating an ion gradient. This electrochemical gradient is then coupled to ATP synthase that synthesizes the energy-rich molecule ATP. Examples include bacteriorhodopsin, a proton pump in Archea; halorhodopsin, a chloride pump in Archea; and proteorhodopsin (pR), a proton pump in Bacteria. The microbial rhodopsins are distributed broadly among the three kingdoms of life. This widespread distribution raises the question of whether they were present in the last common ancestor to both Archea and Bacteria, or even the last universal common ancestor (LUCA) [Lane N. *et al.*, 2010]. There is yet little direct evidence for autotrophy based on microbial rhodopsins [Gómez-Consarnau L. *et al.*, 2007; Gómez-Consarnau L. *et al.*, 2010] but that does not rule out the likelihood that, before the advent of chlorophyll-based oxygenic-photosynthesis 2.4-3.5 billion years ago, these ion pumps served a functional purpose in the process of solar energy transformation [Hohmann-Marriott M. F. and Blankenship R. E., 2011].

Multicellular animals, including human beings, possess membrane proteins in the other major family of rhodopsins, often referred to as type 2 rhodopsins. These are the visual-pigment rhodopsins. The visual rhodopsins are in the G-protein coupled receptor (GPCR) superfamily, but are also photoreceptor proteins that contain retinal, a vitamin A derivative, as their photo-pigment. Retinal has an absorption maximum at $\lambda_{\text{max}} = 380 \text{ nm}$ when measured as the free aldehyde in ethanol, but exhibits different absorption maxima when it forms a protonated Schiff base chromophore linked to different opsins. This absorption maxima shift is also responsible for color vision. Rhodopsins in animals (including both vertebrates and multicellular invertebrates) are responsible only for the

function of vision, unlike the microbial rhodopsins that are also involved in photo-transformation of energy as discussed above.

Figure 1.1 shows the different regions at which a variety of microbial and visual rhodopsins, as well as other pigments, exhibit maximum absorption.

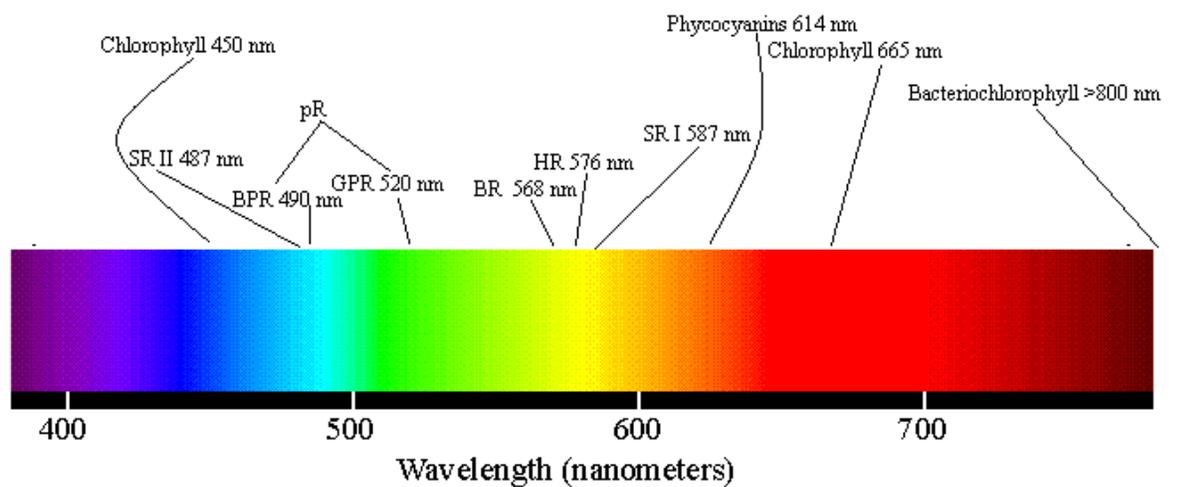


Fig. 1.1: Absorption maxima of various rhodopsins and other photopigments: SR refers to sensory rhodopsin, pR is proteorhodopsin which is divided into Green-absorbing pR (GPR) or Blue-absorbing pR (BPR), and HR is halorhodopsin.

1.1.2 Microbial rhodopsins

Only recently were homologs of archaeal rhodopsins discovered in the domains Bacteria and Eukarya. Proteorhodopsin (pR), the first example of which was found in oceanic proteobacteria in 2000 [Béjã O., *et al.*, 2000], is a light-driven proton pump that functions like its homolog bacteriorhodopsin (BR) (see below). However, its absorption maxima are dependent on the aquatic environment in which it is found. The variants of

pR are either green-absorbing pR with $\lambda_{\max} = 525$ nm or blue-absorbing pR with $\lambda_{\max} = 490$ nm.

There is an extensive older literature on the archaeal homologs of pR, because the first four microbial rhodopsins were discovered in the domain Archea, specifically in *Halobacterium salinarum* and a number of closely related halophilic species. Two of them are light driven pumps viz., bacteriorhodopsin (BR) and halorhodopsin (HR), and the other two are phototaxis receptors, namely sensory rhodopsin (SR) I, and SR II.

Archeal rhodopsins show 80% identity in the 22 residues that form the retinal binding pockets in BR, HR and SR II (Fig.1.2). However, outside the pocket, there is limited identity between the amino acids. A homolog found in *Natronobacteria pharaonis* (NpSR II) shares only 27% identity to BR in the overall primary sequence, and only ~40% identity with other halobacterial sensory rhodopsins [Spudich J. L. and Jung K. H., 2005].

The archaeal homolog of pR, bacteriorhodopsin (BR), is the prototype of a microbial rhodopsin, having been the first one discovered in the 1970s. BR is a light-driven proton pump whereas HR is a light-driven chloride pump. BR pumps proton to the outside of the cell whereas HR pumps chloride inside the cell. Both these pumps hyperpolarize the cell by generating a positive charge outside the cell, thus creating an inwardly directed proton motive force (pmf). Their absorption maxima are in the green-

orange region of the spectrum [Váró G., 2000] with BR absorbing at $\lambda_{\max} = 568$ nm [Albeck A., *et al.*, 1989] and HR at $\lambda_{\max} = 576$ nm [Stenkamp R. E., *et al.*, 2002].

Sensory rhodopsins sense the changes in the intensity or the color of light and then produce signals that affect the flagellar rotor, i.e. the archaeal cell's swimming behavior. SR I attracts halobacteria towards orange light and away from harmful UV-light [Spudich J. L. and Bogomolni R. A., 1984]. SR II is expressed only during the aerobic growth when the cells are not dependent on light for their energy. It serves only a repellent function [Takahashi T., *et al.*, 1990]. In *H. salinarum*, sensory rhodopsin SRI absorbs maximally at $\lambda_{\max} = 587$ nm [Spudich J. L. and Bogomolni R. A., 1984] whereas the SR II absorbs at $\lambda_{\max} = 487$ nm.

The crystal structure of BR has been elucidated [Subramaniam S. and Henderson R., 2000], and its photocycle has been understood quite well [Heberle J., 2000; Lanyi J. K., 2006]. In BR and pR, the retinylidene chromophore alternates between *all-trans* and *13-cis* isomers (Fig. 1.4). Crystal structures of HR [Kolbe M., *et al.*, 2000] and SRII [Gordeliy V. I., *et al.*, 2002] are also known (Fig. 1.3).

srIINatronomonas	-----MVGLTTLFWLGAIGMLVGT---LAFAWAGR-----	28
srIIHaloarcula	-----MATITWFVLGGLGELLGT---AVLAYGYTL-----	28
srIIFalobacterium	-----MALTTFWVWGAVMGLAGT---VLPIRDCTIR-----	27
Bacteriorhodopsin	-----QAQITGRPEWIWLAGTALMGLTLYFLVKGMGV-----	34
Halorhodopsin	-----AVRENALLSSSLWVVA---LAGIAILVVFVYMGRTI-----	33
proteorhodopsin	MKLLLIILGSVIALPTFAAGGGDLASDYTVGSFWLVTAALLAST---VFFVVERDR	53
Bovine	--MNGTEGPNFYVPSNKTGVVRSFPEAPQYLLAEPWFQSMALAAVYMLLIMLGFINFLT	58
srAnabaena	-----	
srIINatronomonas	--AGSGERRYVTVLGVISGIAAVAYVVMALGVGWVPV-----AERTVFAPRYID-	75
srIIHaloarcula	--VPEETRKRYLLLIAPGIAIVAYALMALGFGSIQS-----EGHAVYVRYVD-	75
srIIFalobacterium	--HPS-HRRYDLVLAGITGLAAIAYTTMGLGITATTV-----GDRTVYLARYID-	73
Bacteriorhodopsin	--SDPDAKFFYAITTLVPAIAFTMYLSMLLGYGLTMVFPFG-----EQNPYIWARAD-	85
Halorhodopsin	--RPGRPRIWGATLMIPLVSISSYLGLLSGLTVGMIEMPAGHALAGEMVRSQWGRYLT-	90
proteorhodopsin	--VSARKWTSLSVSGLVGTIAFWHYMYMR-GVWIETG-----DSPTVFRYID-	97
Bovine	LYVTVQHKLRTPLNLYILLNLAVADLFMVFGGFTTTLTSLHGYVFGPTGCNLEGGFAT	118
srAnabaena	---MHHHHHSLSIGRTCWAIAEGYIPPYGNGPEPQ-----FIS-	36
srIINatronomonas	-----WILTTPLIVYFLGLLAGLD---SREFGIVITLNTV-VMLAGFAGAMVPG----	120
srIIHaloarcula	-----WLLTTPLNWVFLALLAGAS---REDTVKLVVLQAL-TIVFGFAGAVTPS----	120
srIIFalobacterium	-----WLVTTPLIVLYLAMLARPG---HRTSAWLLAADVF-VIAAGIAAALTGT-----	118
Bacteriorhodopsin	-----WLFPTPLLLLDLALVDAD---QGTILALVGDGI-MIGTGLVGLTKVY----	131
Halorhodopsin	-----WALSTPMILLALGLLADVD---LGSFTVIAADIG-MCVTGLAAAMTTSAL---	137
proteorhodopsin	-----WLLTVPLLICEFYLLIAAA---TNVAGSLFKKLLVGLVMLVFGYMGEGAI---	145
Bovine	LGGEIALWLSLVLAIERVYVVCCKPMSNFRFGENHAIMGVAFPTVMALACAAAPPLVWGSRY	178
srAnabaena	-----HETVCILNAGD---EDAHVEITIIYSDKEPVGYPYRLTVPAR----	74
srIINatronomonas	-----IERYALFGMGAVAFGLVYVYLVGP-----MTESASQR	152
srIIHaloarcula	-----PVSYALFAVGGALFGGVYIYLLYRN-----IAVAAKST	152
srIIFalobacterium	-----VQRWLFVAVGAAGYAALLYGLLGT-----LPRALGDD	150
Bacteriorhodopsin	-----SYRFVWVAISTAAMLYIYLVVFFG-----FTSKAESM	163
Halorhodopsin	-----LFRWAFYAIACAFFVVLVSLVTD-----WAASASSA	169
proteorhodopsin	-----MAAWPAFIIGCLAWVYMIYELWAG-----EGKSACNT	177
Bovine	IPEGMQCSGIDYYTPHEETNNEFVYIMFVVHFIIPLIVIFFCYGQLVFTVKEAAAQQ	238
srAnabaena	-----	
srIINatronomonas	SSGIK-SLYVRLRNLTVILWAIYPIWLLGP-----PGVALLT-	189
srIIHaloarcula	LSDIEVSLYRTRLNRFVVLWLVYVWVLLGA-----AGVGLMD-	190
srIIFalobacterium	-PRVR-SLFVTLRNIIVVLTLYPVVWLLSP-----AGIGILQ-	186
Bacteriorhodopsin	RPEVA-STFKVLRNVTVVLSAYPVVWLLIGS-----EGAGIVP-	200
Halorhodopsin	GTAEI---FDTLRVLTVVWLVWLYPVVAVGV-----EGLALVQS	205
proteorhodopsin	ASPAVQSAYNTMMYIIIFGWAIYVGYFTG-----YLMGDGGS	215
Bovine	ESATTQKAEKVTVMVIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFMTIPAFFAKTS	298
srAnabaena	--RTKHVRFNDLNDPAPIPHDITDFASVIQSN-----	103
srIINatronomonas	PTVDVALIVYLDLVTKVGFGFIALDAAATLRAEHGESLAGVDTDAPAVAD-	239
srIIHaloarcula	VETATLVVVYLDVVTKVGFGVIALLAMIDLG----SAGETAEEPTAVAGD	236
srIIFalobacterium	TEMYTIVVVYLDVSKVAFVAVLGDADAVSRLVAADAAAPATAEPTPDGD	237
Bacteriorhodopsin	LNIEITLFLMVLDSAKVGFGLILLRSRAIFGEAEPEPSAGDGAATS---	248
Halorhodopsin	VGATSWAYSVLDFVAKYVFAFILLRWVANNERTVAVAGQTLGTMSDD---	253
proteorhodopsin	ALNLIYVNLADFVNKILFGLIWNVAVKESNA-----	249
Bovine	AVYNPVIYIMMKNQFRNCMVTLCCGKNPLGDDEASTTVSKTETSQVAPA-	348
srAnabaena	---VPIVQHTRLDSRQAENALLSTIAYANT-----	131

Fig. 1.2: Multiple sequence alignment of rhodopsins: The conserved residues are shown in red. The following rhodopsins are used for comparison: Sensory rhodopsins II of *Natronomonas pharaonis*, *Haloarcula* and *Halobacteria*, ion pumps of bacteriorhodopsin (proton pump), halorhodopsin (chloride pump), proteorhodopsin (proton pump), sensory rhodopsin of *Anabaena* and eukaryotic GPCR, bovine rhodopsin.

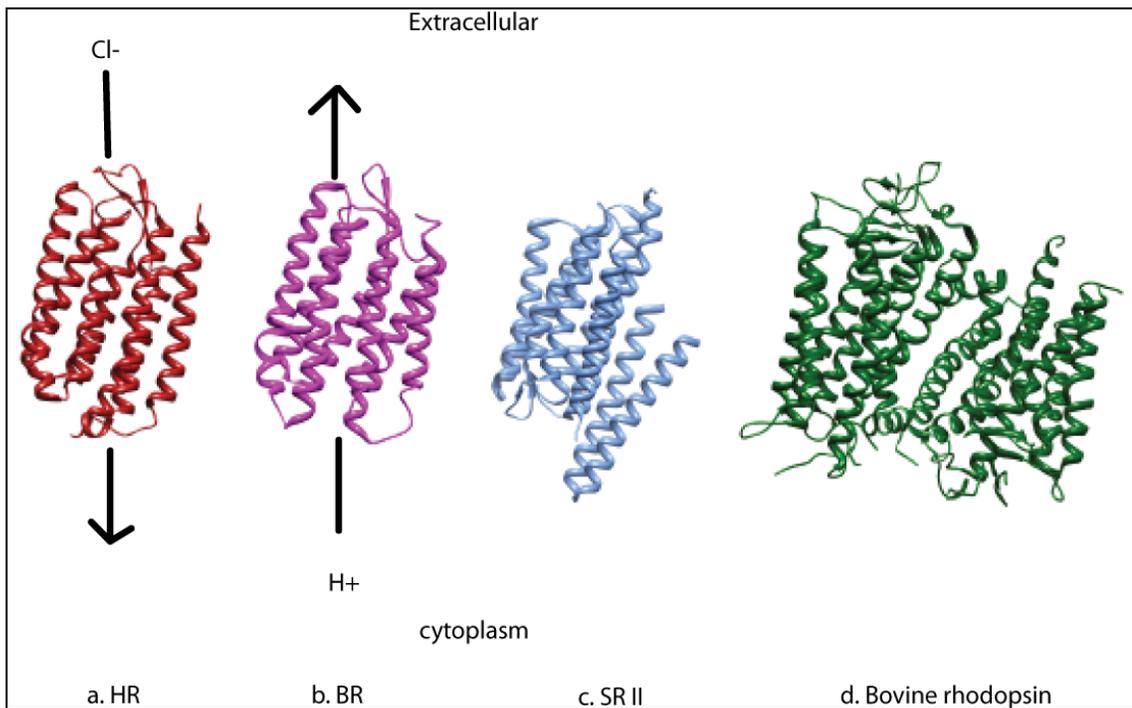
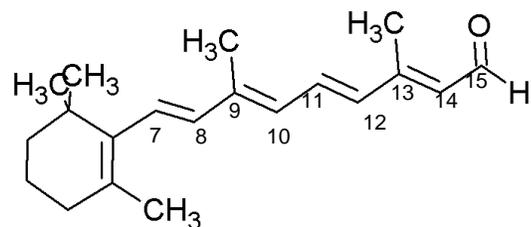
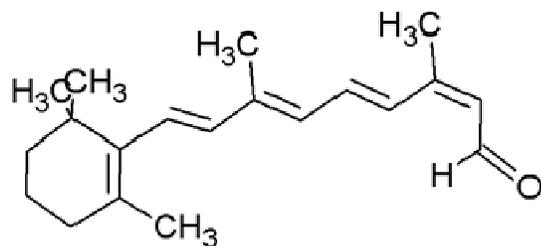


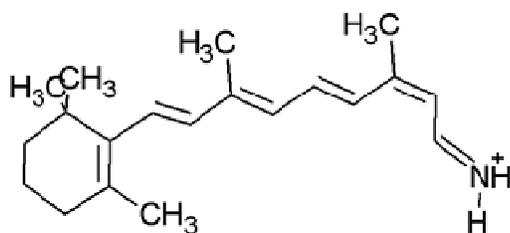
Fig. 1.3: Crystal structures of rhodopsins with Protein Data Bank (PDB) numbers: These rhodopsins serve different functions but share the same overall topology of heptahelical trans-membrane structure. (a) HR PDB: 1E12 (*Halobacterium salinarum*) transports chloride into the cytoplasm from extracellular medium, (b) BR PDB: 1FBB (*Halobacterium salinarum*) pumps proton from cytoplasmic side to the outside of the cell, (c) SR II PDB: 1H2S (*Natronomonas pharaonis*) senses harmful light and repels the cells away to a safer environment, (d) Bovine rhodopsin PDB: 1F88 is a G-Protein Coupled Receptor (GPCR). The PDB files show a dimer of bovine rhodopsin crystal. The above PDB files were manipulated with UCSF Chimera package (described in chapter 2).



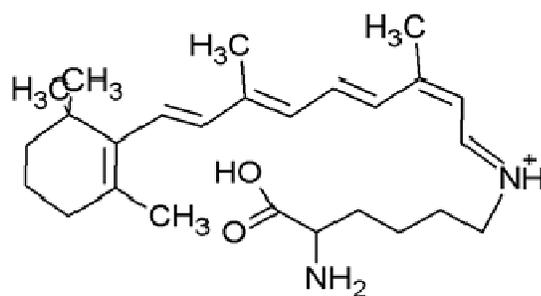
a. All-*trans* retinal



b. 13-*cis* retinal



c. Protonated Schiff Base (PSB)



d. 13-*cis* retinal with PSB linkage to lysine

Fig. 1.4: Structure of free retinal and retinal covalently bonded to lysine: In microbial rhodopsins, All-*trans* retinal (a) is converted to 13-*cis* retinal (b) upon photon absorption. c. Structure of 13-*cis* retinal showing a protonated Schiff base linkage (PSB) d. Structure of 13-*cis* retinal bonded to the Lysine via a protonated Schiff base linkage. These structures were created in ChemSketch [ACD Labs].

1.2 Proteorhodopsin

Within the past decade it has been shown that among the type 1 rhodopsins, the most abundant and widely distributed is proteorhodopsin. Proteorhodopsin was the first-discovered eubacterial homolog of archaeal rhodopsins [Béjà O., *et al.*, 2000]. A bacterial artificial chromosome (BAC) library of naturally occurring marine bacterioplankton from Monterey Bay was constructed and sequenced. Sequence analysis of a 130-kb genomic fragment that encodes ribosomal RNA (rRNA) operon from an as-yet uncultivated member of marine bacteria revealed a gene encoding a putative rhodopsin [Béjà O., *et al.*, 2000]. The rRNA sequence on the same fragment identified the organism as γ -proteobacteria belonging to SAR86 group. Hence, the protein was named proteorhodopsin.

The native pR consists of 249 amino acids [Béjà O., *et al.*, 2000]. The amino acid composition as shown in Table 1.1 and the secondary structure is shown in Fig. 1.5A. The active protein is covalently bound to its chromophore all-*trans* retinal at the conserved lysine, Lys²³¹, via a Schiff base linkage (Fig. 1.4). Proteorhodopsin has been heterologously expressed in *E. coli*, and addition of retinal to the pR expressing *E. coli* bacteria causes a red-coloration of cells (Fig. 1.5B) with an absorption maximum at 520nm or 540nm depending on the pH. The residue that is responsible for the two spectral forms is Asp97, the homolog of Asp85 in BR [Dioumaev A. K. *et al.*, 2002]. The Asp97 has to be deprotonated in order for pR to perform the photocycle function. Asp97 is deprotonated at pH 7 or above where pR exhibits $\lambda_{\text{max}} = 520$ nm. Thus, it is the 520 nm form of pR that is physiologically important.

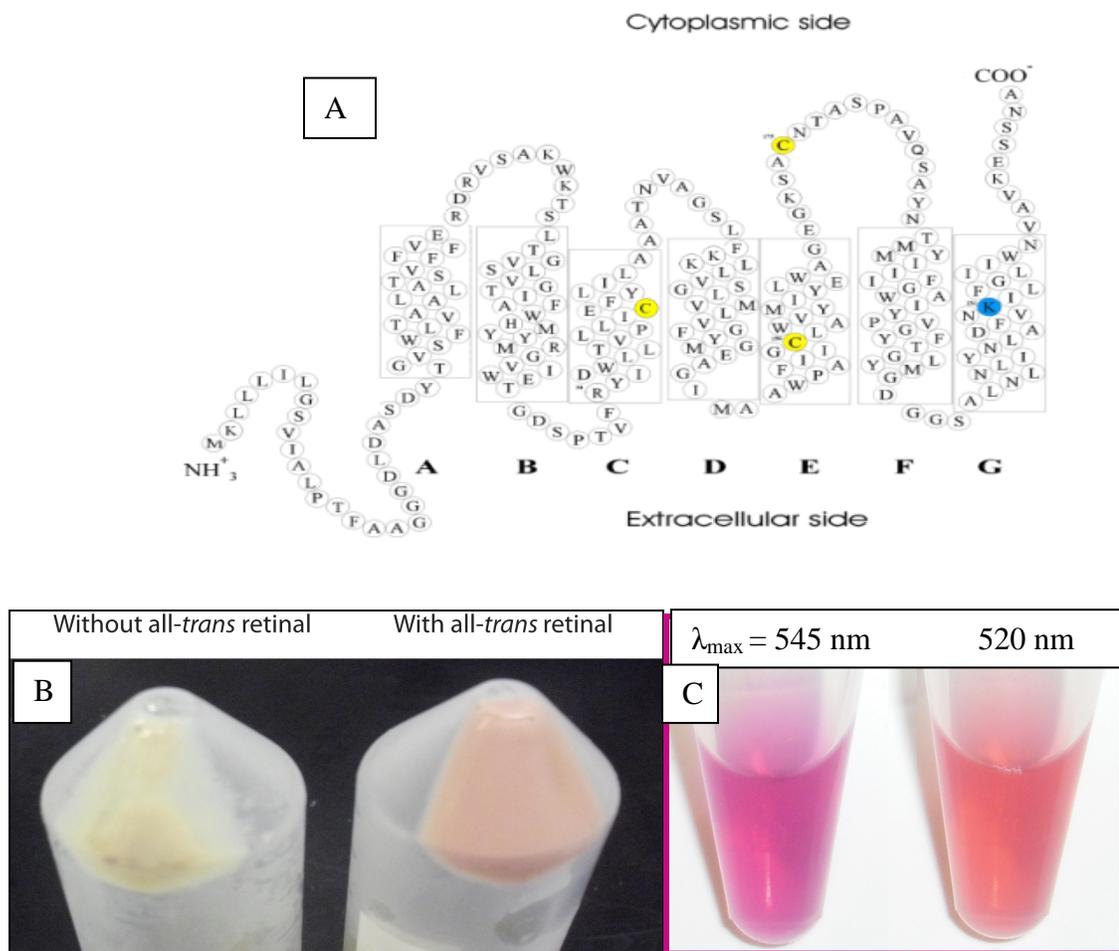
Table 1.1: Amino acid composition of pR (ProtParam)

Amino acid	Number of amino acids present in pR	Percent composition
Ala (A)	28	11.2
Arg (R)	4	1.6
Asn (N)	9	3.6
Asp (D)	8	3.2
Cys (C)	3	1.2
Gln (Q)	1	0.4
Glu (E)	7	2.8
Gly (G)	24	9.6
His (H)	1	0.4
Ile (I)	19	7.6
Leu (L)	31	12.4
Lys (K)	8	3.2
Met (M)	10	4.0
Phe (F)	15	6.0
Pro (P)	6	2.4
Ser (S)	16	6.4
Thr (T)	14	5.6
Trp (W)	10	4.0
Tyr (Y)	14	5.6
Val (V)	21	8.4
Total	249	100

Total number of negatively charged residues (Asp + Glu): 15

Total number of positively charged residues (Arg + Lys): 12

Fig. 1.5: Secondary structure and color of pR: **A) Secondary structure of pR:** Proteorhodopsin is a hepta-helical trans-membrane protein. The 7-helices are labeled from A-G. The N-terminus is on the extracellular side whereas the C-terminus faces the cytoplasmic side of the cell. Lysine²³¹ residue that forms a protonated Schiff base linkage with retinal is depicted in blue color. Three cysteines depicted in yellow are mutated to serine, pR is referred to as pR triple cysteine mutant (TCM). **B) UT5600 cells expressing pR:** The left tube shows cells that were induced with L-arabinose, the right tube shows same cells after the addition of *all-trans* retinal. **C) pR TCM at pH 6 and 9:** pR TCM absorbs at different λ_{max} depending on pH. This is due to the titration of D97 at higher pH. High pH form is the physiologically active form that initiates the photocycle of pR.



1.2.1 Photocycle of pR

Due to a lack of a three-dimensional crystal structure of pR, the photocycle has been studied using various biophysical techniques such as time-resolved and static Fourier-transform infrared (FTIR) or visible spectroscopy, and flash-induced laser spectroscopy. In summary, the photocycle begins upon photon absorption, the retinal chromophore undergoes isomerization along C₁₃-C₁₄ bond which leads to a conformational change in the retinal form from all-*trans* to 13-*cis* [Friedrich T. *et al.*, 2002; Krebs R. A., *et al.*, 2003]. The protein responds to change in the retinal conformation by cycling through a series of intermediate stages named K, L, M, N and O [Béjà O., *et al.*, 2000; Dioumaev A. K. *et al.*, 2002; Friedrich T. *et al.*, 2002; Váró G., *et al.*, 2003]. At the end of the cycle, the protein regains its original state and the retinal chromophore thermally re-isomerizes to an all-*trans* state. The physical significance of this photocycle is that a proton is pumped from the cytoplasm to the extracellular side of the membrane. Fig. 1.6 illustrates the photocycle of pR.

The photocycle of pR is considered to be similar to its archeal homolog bacteriorhodopsin (BR) [Béjà O., *et al.*, 2000] given the conservation of several functionally important residues (Fig. 1.2). Lys231 (Lys216 in BR) present in the helix G (Fig. 1.5A) forms a protonated Schiff base with retinal (Fig. 1.4). The pR with photo-isomerized retinal is referred to as the K-state [Bergo V., *et al.*, 2004; Dioumaev A. K. *et al.*, 2002]. There exists an L-state in BR but it has not been isolated kinetically or cold-trapped in pR yet, although its presence was clearly identifiable in time-resolved FTIR spectral features that decay during the temporal rise of characteristic M spectral features

[Xiao Y., *et al.*, 2004]. In addition, the multi-exponential pattern of K to M transition implies the presence of an L-intermediate [Dioumaev A. K. *et al.*, 2002]. In the L-to-M switch, the proton from the protonated Schiff base is transferred to a proton acceptor Asp97, the homolog of Asp85 in BR [Béjà O., *et al.*, 2000; Dioumaev A. K. *et al.*, 2003; Dioumaev A. K. *et al.*, 2002; Friedrich T. *et al.*, 2002]. Asp227 (Asp212 in BR) probably in concert with another conserved residue Tyr200 (Tyr85 in BR) aids in providing the environment for the proton channel from the Schiff base to the proton acceptor [Béjà O. *et al.*, 2000; Heberle J., 2000]. Simultaneously, another proton is released to the extracellular side during the formation of the M intermediate. Arg94 (Arg82 in BR) appears to be involved in the proton release mechanism at the extracellular side before the proton uptake from the cytoplasm [Govindjee R., *et al.*, 1996; Partha R., *et al.*, 2005]. In the M-to-N transition, the Schiff base is reprotonated by the proton donor Glu108 (Asp96 in BR), located toward halfway between the Schiff base group and the cytoplasmic side of the membrane [Dioumaev A. K. *et al.*, 2002]. The transition from N to O [Béjà O., *et al.*, 2000] involves protonation of Glu108, 13-*cis* retinal returns to all-*trans* state [Dioumaev A. K. *et al.*, 2002] and the proton acceptor D97 becomes deprotonated thus completing one photocycle. The proton pathway is shown in Fig. 1.7.

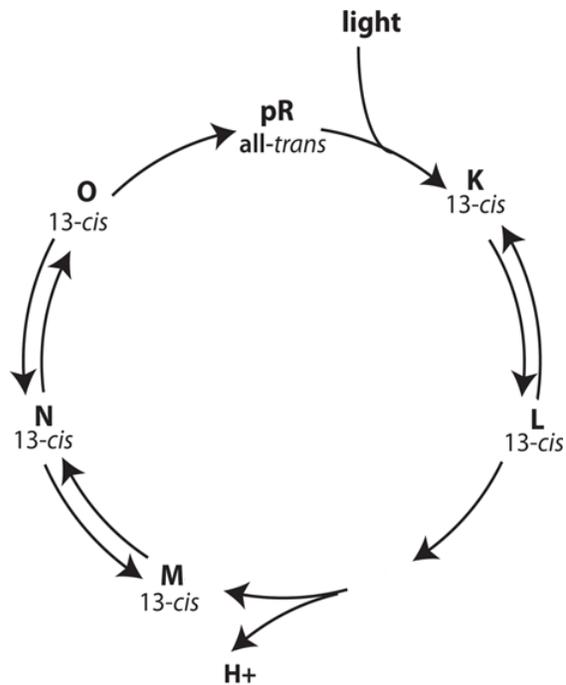


Fig. 1.6: Photocycle of pR: The pR photocycle was elucidated by using biophysical techniques such as time-resolved and static FT-IR or flash-induced laser spectroscopy. The existence of some of the intermediates, e.g. L is based most strongly on evidence from the bR photocycle, and has not been directly observable in pR. There is a possibility of existence of an intermediate between L and M, therefore, a space has been left. The same is true of the reversibility of some of the steps shown. The time constant for K to M state is in tens of milliseconds. Decay of half of M intermediate to N state is approximately 2 ms. The recovery of the native pR takes the longest time, approximately corresponding to 200 ms.

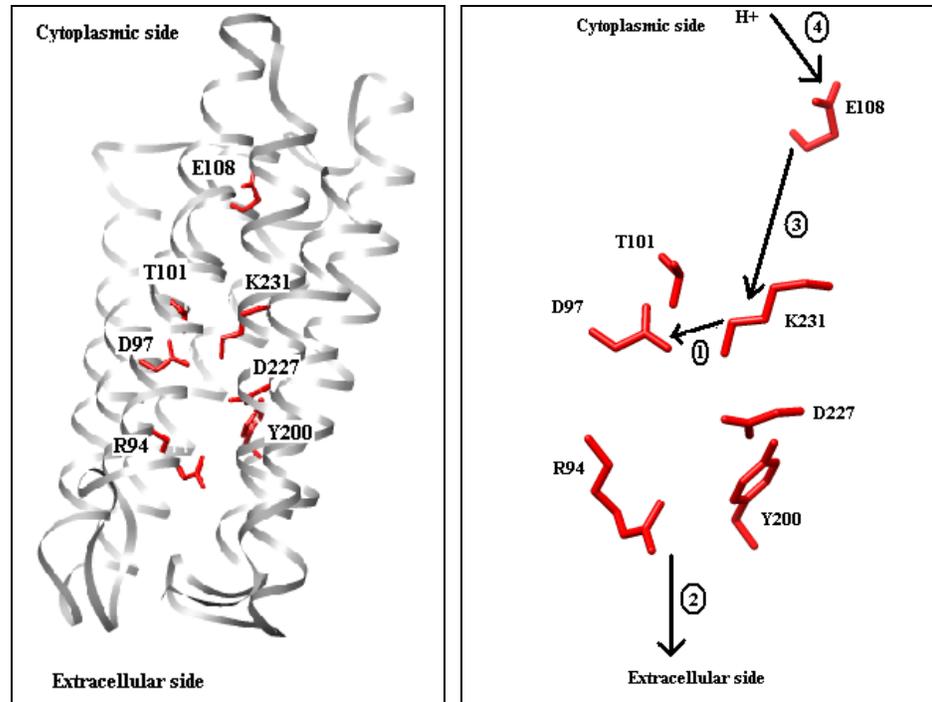


Fig. 1.7: Proton pathway of pR: **Left:** Homology model of pR based on BR is shown with the important residues involved in proton pump colored in red. The Homology model was made and manipulated as described in Chapter 2. **Right:** Only the residues are magnified here to show the proton pathway. The numbers refer to the different intermediates. K^{231} is covalently bonded via protonated Schiff base linkage to retinal chromophore that undergoes all-*trans* to 11-*cis* isomerization upon photon absorption (not shown). This is referred to as K state. (1) In L to M transition, the Schiff base proton is transferred to D^{97} , the proton acceptor and another proton is simultaneously released at the extracellular side (2), while D^{97} is still protonated. (3) In M to N transition, Schiff base is reprotonated from cytoplasmic side by the proton donor E^{108} . (4) In N to O transition, E^{108} is reprotonated from the cytoplasm. At the end, the retinal returns to its all-*trans* state.

1.2.2 Energy transduction and proton pumping of pR

Several lines of evidence prove that pR functions as a proton pump and not as a sensory rhodopsin. One characteristic difference between the sensory rhodopsins and proton-pumps is that the genes for sensory rhodopsins are generally located physically near to, and co-transcribed with, genes that encode additional transducer proteins needed

to convert the signals detected by the sensory rhodopsins into physiological responses. *Htr* is an example of such a co-transcribed regulator/transducer gene [Seidel R., *et al.*, 1995; Yao V. J. and Spudich J. L., 1992]. Sensory rhodopsins generally exhibit light-mediated proton pumping activity, but to a lesser degree, and only when the transducer is removed [Bogomolni R. A. *et al.*, 1994]. The proteorhodopsin gene in γ -proteobacteria does not sit near any *Htr*-like gene. Additionally, pR-expressing *E. coli* cells clearly show a net outward translocation of protons in the presence of retinal and light [Béjà O., *et al.*, 2000; Béjà O., *et al.*, 2001]. The transport of protons creates a proton gradient and proton motive force (pmf) which generates membrane electrical potential, -90 to -100 mV, in physiologically relevant range [Béjà O., *et al.*, 2000]. The pmf is utilized by the ATP synthase to synthesize ATP that can be utilized by the cell for its energy needs. Fig. 1.8 illustrates the mechanism of proton pump and energy production.

Furthermore, flash-induced absorption changes determined in heterologously expressed *E. coli* produced evidence of the existence of two intermediates, M and O [Béjà O., *et al.*, 2000]. Flash spectroscopy of native membranes of the bacterioplankton samples from Monterey Bay also confirmed the photocycle characteristics [Béjà O., *et al.*, 2001]. This also indicates that marine bacteria have pR as well as endogenous retinal synthesizing genes for the pR to be able to function as a proton pump. Flash spectroscopy also showed that pR is present at high density in these membranes to about 2.4×10^4 pR molecules per cell [Béjà O., *et al.*, 2001]. These facts show that pR is functional as a proton pump in its native environment.

The ion pumps differ from sensory rhodopsins with respect to the photocycle kinetics. The ion pumps such as BR and HR have faster photocycling rates. They show photocycle with a half-time typically ~20 ms whereas the sensory rhodopsins exhibit photocycle half-time >300 ms [Hoff W. D., *et al.*, 1997]. This large kinetic difference is significant because the proton-pumping is efficient at a faster photocycling rate whereas the slow rate is important for the light-sensing function of the sensory rhodopsins. The photochemical reaction cycle of pR has been estimated to be 20-45 ms [Béjà O., *et al.*, 2000], yet another fact that confirms the light-dependent proton-pumping function of pR.

BR and pR contain a proton donor Asp96 (BR) or Glu108 (pR) [Dioumaev A. K. *et al.*, 2002] whereas the haloarcheal sensory rhodopsins lack a corresponding proton donor carboxylate group. They contain Tyr or Phe in that position. The significance of proton donor in the pumps lies in the fact that an immediate proton transfer to the unprotonated Schiff base intermediate, M, improves the pumping efficiency by speeding up the decay of M. In the sensory rhodopsins, the M intermediate is a signaling state as shown for SRI [Spudich J. L. and Jung K. H., 2005]. The longer lifetime of the M intermediate is advantageous for the signaling efficiency and therefore, the lack of a proton donor is beneficial.

It should however be noted that although the photocycling rates are considered a criterion for distinction between the transport proteins and sensory proteins in Archea, it is not generally true for the proteorhodopsin. The pR proton-pump found on surface waters of Monterey Bay shows a faster rate of ~20 ms whereas the pR proton-pump

found at a depth of 75 m from Hawaiian Ocean-Time station exhibits ~10-fold difference in its rate, showing a slower rate [Wang W. W., *et al.*, 2003].

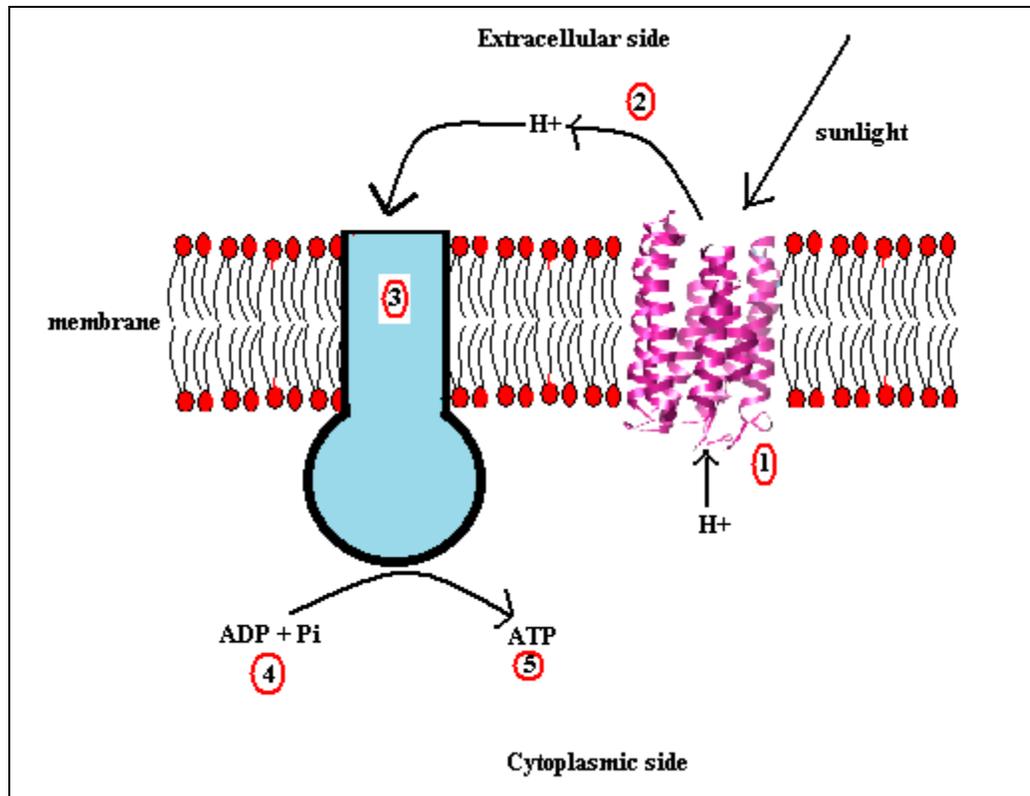


Fig. 1.8: Proton pump coupled to ATP synthesis: Label 1 shows pR that absorbs sunlight and functions as a proton pump. Label 2 shows the expelled protons that generate a proton gradient that creates a proton motive force (pmf). Label 3 represents ATP synthase that utilizes this pmf to synthesize high energy ATP (labeled as 5) from a lower energy metabolic intermediate ADP (labeled as 4).

1.3 Visual rhodopsins

Mammalian rhodopsins are found in the photoreceptor cells, rods and cones, of the retina and are responsible for vision in vertebrates. They are seven-trans-membrane α -helical proteins that bind to 11-*cis* retinal. Photon absorption by the chromophore 11-*cis* retinal isomerizes it to all-*trans* form which is the first step of the photochemical reaction cycle, also known as photocycle. The isomerization of retinal causes conformational changes in the structure of rhodopsin, eventually leading to the formation of an intermediate metarhodopsin I which becomes deprotonated to give metarhodopsin II, the photoactivated rhodopsin which initiates an enzymatic cascade. While the cascade is in motion, the metarhodopsin II is hydrolyzed into opsin and all-*trans* retinal which diffuses away as it cannot fit into the binding site of 11- *cis* retinal. All-*trans* retinal is then converted to 11-*cis* retinal by the action of several enzymes leading to the regeneration of rhodopsin. [Stryer L]

Each photoactivated rhodopsin catalyzes the activation of multiple copies of the G-protein known as transducin, which initiates a further amplifying cascade of additional enzymes, eventually leading to the closure of ion channels in the membrane, leading to the hyperpolarization of the rod cell, a neuronal signal that can be propagated to the optic nerve. Fig. 1.9 shows the visual cycle of eukaryotes [Fain G. L. *et al.*, 2010].

Mammalian rhodopsins absorb in the green-blue region of the spectrum with absorption maximum at $\lambda_{\text{max}} = 500 \text{ nm}$ in the rods. Cones absorb at three distinct

wavelengths, $\lambda_{\max} = 420 \text{ nm}$ (blue cones), $\lambda_{\max} = 530 \text{ nm}$ (green cones) and $\lambda_{\max} = 560 \text{ nm}$ (red cones).

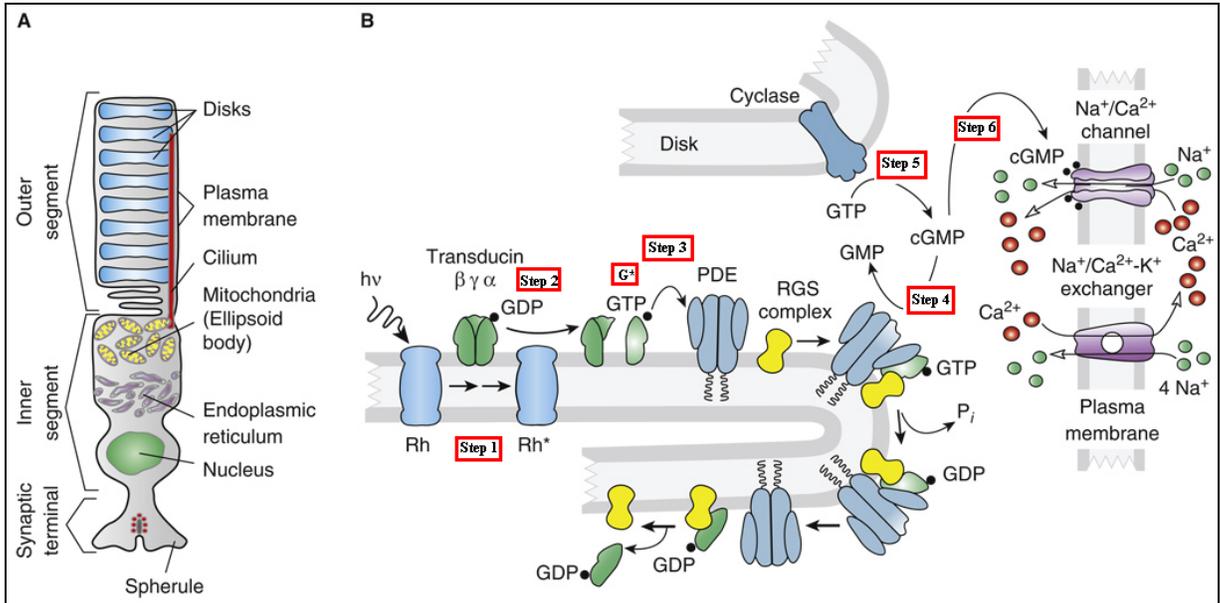


Fig. 1.9: Visual phototransduction in Rod cells: (A) Photoreceptors are present in the disks in the outer segment in a rod cell. (B) In step 1, light ($h\nu$) is absorbed by the rhodopsin (R) molecule which is converted into an active form (R^*). In step 2, R^* activates the downstream G-protein, transducin, to G^* . Transducin is a heterotrimer made up of α , β , and γ subunits, and is bound to a GDP molecule in its inactive form. After activation, it exchanges a GTP molecule at its α subunit and dissociates from the β , and γ subunits. In step 3, G^* binds to the α -subunit of phosphodiesterase (PDE) thereby activating its β , and γ subunits. In step 4 the activated PDE hydrolyzes cGMP to 5'-GMP. In step 5, guanylyl cyclase (GC) synthesizes cGMP. In step 6, the lower concentration of cGMP closes the cGMP-gated Na^+ channel preventing Na^+ and Ca^{++} from entering the cells. This leads to hyperpolarization of the cell causing the voltage-gated Ca^{++} channels to close. This causes inhibition of release of the neurotransmitter, glutamate which is an inhibitor of the bipolar cells. The rod cells are connected via bipolar cells to the ganglion. Since the bipolar cells are released from the inhibition, they begin to release their own transmitter at the synapse of the bipolar cells and ganglion leading to the excitation of the synapse. The axons of the ganglion cells form the optic nerve, which relays the impulse to the brain. This leads to interpretation of the image. This figure is a modified version of a figure taken with permission from Fain G. *et al.* [Fain G. L. *et al.*, 2010].

1.4 Biological significance of pR

1.4.1 Evolutionary basis

The widespread distribution of proteorhodopsins and other microbial rhodopsins among widely divergent classes of unicellular organisms has evolutionary significance. The type 1 rhodopsins i.e. the microbial rhodopsins, are found in all three domains of life, namely Archaea, Bacteria and Eukarya. Therefore, it can be assumed that the progenitors of these proteins may have existed in the common ancestor before these domains underwent divergence. If this is the case, then the pR and other light-driven pumps may have been the primary means of solar energy capture and its conversion into cellular energy. The first chlorophyll-based photosynthetic organisms were the blue-green algae which date back somewhere between 3.5 and 2.4 billion years ago [Hohmann-Marriott M. F. and Blankenship R. E., 2011]. Proteorhodopsin-like homologs could have played an important role in the generation of energy between 3.8 and 3.5 billion years ago before the advent of chlorophyll-based photosynthesis. With the evolution of multicellular and complex organisms with more efficient energy conversion systems, pR-like proteins remained restricted to the simpler unicellular organisms.

1.4.2 pR based phototrophy

The original discovery of pR was from γ -proteobacterium which belongs to SAR86 clade [Béjà O., *et al.*, 2000]. However, since these bacteria have not yet been cultured, all the functional properties of pR were originally studied by heterologous expression in *E. coli*. Proteorhodopsin was also found in *Candidatus pelagibacter* [Govindjee R., *et al.*, 1996], the first cultivated bacteria belonging to SAR11 clade which

is the most abundant clade of oceanic bacteria [Morris R.M., *et al.*, 2002; Rappé M. S., *et al.*, 2002]. Studies were conducted on *Pelagibacter unique* obtained from Oregon coasts, to examine the effect of pR on cell growth rate or cell yield in the light and the dark [Giovannoni S. J., *et al.*, 2005]. These results were not conclusive due to the fact that the composition of the sea water used for the experiments is not known completely. Therefore, the limiting factors for *Pelagibacter* growth could not be identified [Govindjee R., *et al.*, 1996]. In 2007, marine bacteria belonging to the class *Flavobacteria* were isolated from the surface water of the Mediterranean Sea and were cultured [Gómez-Consarnau L. *et al.*, 2007]. Their genome-sequence analysis uncovered the presence of pR genes and consequent experiments on the growth or cell yield in light and dark showed the importance of pR *in vivo*. The bacteria showed increase in number when exposed to light. An inverse relationship between abundant organic matter and a functional pR was suggested by both the groups [Giovannoni S. J., *et al.*, 2005; Gómez-Consarnau L. *et al.*, 2007]. That is, pR-mediated phototrophy is functional only during the periods when organic food is scarce.

Another group working with heterologously expressed pR in *E. coli* also showed the importance of pR *in vivo* [Walter J. M., *et al.*, 2007]. They demonstrated that pR can augment cell mobility and increase cell survival when the cells are subjected to energy- and oxygen-depletion by azide. This observation confirms the notion that pR is most functional during lean times. Another lab working with pR in *E. coli* showed that pR catalyzes light-dependent proton translocation across the membrane which generates chemiosmotic potential [Martinez A., *et al.*, 2007]. The proton translocation thus results

in photophosphorylation in *E. coli*, i.e., synthesis of ATP in the presence of light [Martinez A., *et al.*, 2007].

Recently a study of the marine bacteria *Vibrio sp.* showed that pR enables the cells to survive starvation much better in the presence of light than in dark [Gómez-Consarnau L., *et al.*, 2010]. They also showed that the survival ability was abolished upon deletion of the pR gene, and again restored when pR was supplied in *trans* (i.e. on an exogenous plasmid introduced into the cells). This study definitively determined a physiological role for pR in the native marine bacterium.

The most recent study on marine *Flavobacterium* concluded increase in cell yields and growth rates in the presence of light and low carbon growth conditions. This study showed a significant up-regulation in the pR and retinal biosynthetic genes in light. [Kimura H., *et al.*, 2011].

These studies show that pR supplements the energy needs of the bacteria in the presence of light and retinal only when the nutrition supplies are limited. Thus, it enables the cells to survive in the nutritionally unpredictable environment.

1.4.3 Energy balance

Solar energy is trapped in the biosphere mainly by the chlorophyll-based protein clusters known as photosystems I and II, found mainly in cyanobacteria and the closely-related eukaryotic symbiont organelles known as chloroplasts. However, since the

discovery of pR from surface waters in the year 2000, numerous pR-directed surveys have identified a wide variety of similar genes in picoplankton from various geographic ocean environments including Antarctic, Central North Pacific, Mediterranean Sea, Red Sea, and the Atlantic Ocean as well as fresh waters [Atamna-Ismaeel N., *et al.*, 2010; Bèjà O., *et al.*, 2001; de la Torre J. R. *et al.*, 2003; Man D., *et al.*, 2003; Man-Aharonovich D., *et al.*, 2004; Sabehi G., *et al.*, 2003]. Over 4000 variants of pR have been isolated from both surface and deep-water samples, as well as from both open-seas and coastal areas. Sea water samples from Sargasso Sea alone identified 782 different partial sequences homologous to pR. Another global shipping expedition uncovered 2,674 putative pR genes [Rusch D. B, *et al.*, 2007; Venter J. C., *et al.*, 2004].

Given the ubiquity and abundance of pR with an estimated average of at least 10^{28} pR-producing bacteria in the photic zone [Morris R.M., *et al.*, 2002] and a more recent estimate of 80% oceanic bacteria harboring pR [DeLong E. F. and Bèjà O., 2010] it is evident that pR can affect the solar energy balance of the Earth.

Fig. 1.10 shows the Earth's solar energy balance [Walter J. M., *et al.*, 2010]. About 50% is reflected by the Earth or clouds, leaving only 50% of the total solar irradiance to reach the Earth. According to this article [Walter J. M., *et al.*, 2010], when average solar illumination of North American latitudes is considered to be 200 watts per square meter (W/m^2) [Programme UND, 2000], and land area is approximated at $10^{13} m^2$, solar illumination of US is estimated to be 2000 trillion Watts (TW). Even if 10% of the land is illuminated with 2% efficiency, it would produce 4 TW of energy [Walter J. M., *et*

al., 2010]. A comparison is made to four different energy resources as described in the Fig. 1.10 caption.

Conservative calculations, based on the amount of pR present in a liter of ocean water (using 0.35 $\mu\text{g/L}$ pR, 30,000 Da, 0.1 V electro-potential gradient), show that oceans can capture at least 1 TW of annual solar energy, compared to 3.5 TW of annual non-biological energy consumption by the USA [Brenner M. P. *et al.*, 2006]. Proton pumps such as bacteriorhodopsin and proteorhodopsin have therefore attracted attention for the purpose of solar energy capture and harvest [Rusch D. B, *et al.*, 2007; Venter J. C., *et al.*, 2004; Walter J. M., 2010].

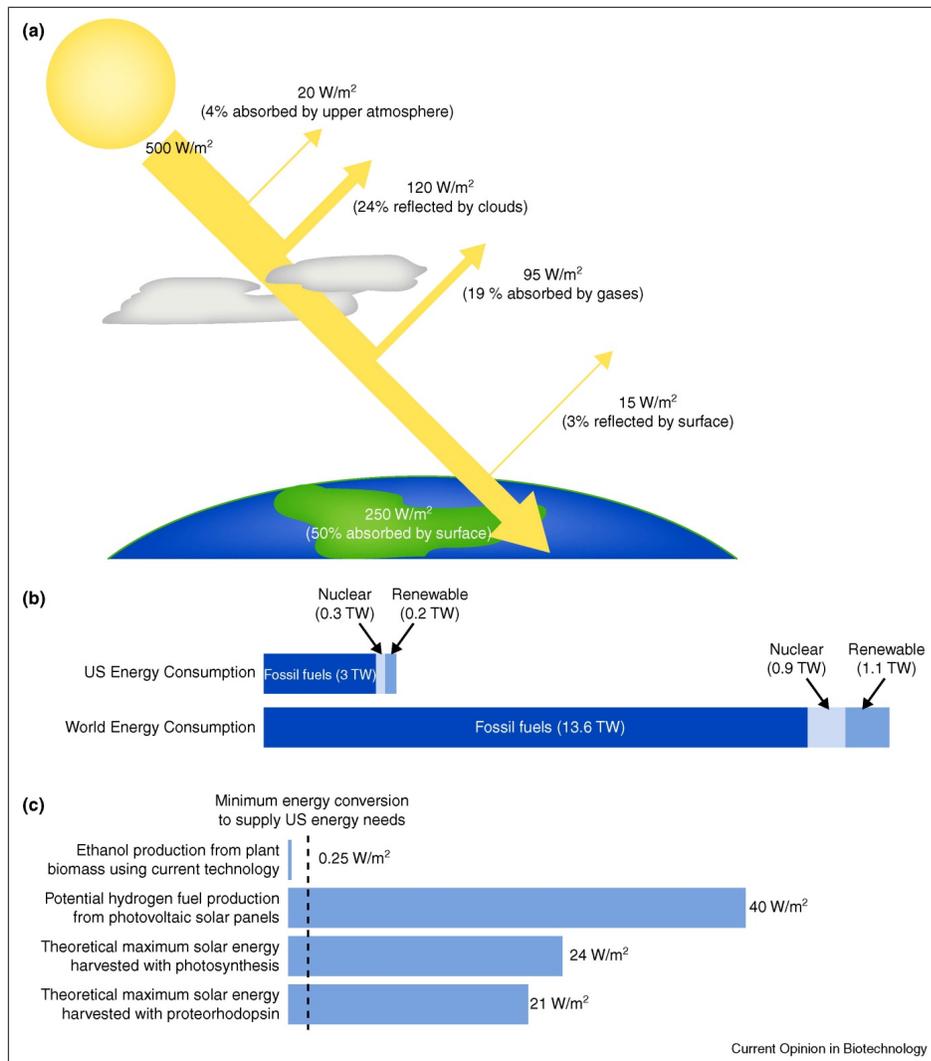


Fig. 1.10: Earth's solar energy balance: This figure is taken with permission [Walter J. M *et al.*, 2010] (A) Solar energy that reaches our planet is only 50% of the original irradiance of the Sun and then 50% of that is absorbed by the Earth's surface. The remaining energy could provide enough power to sustain our current needs. (B) Energy consumption of the USA and world relies mainly on the fossil fuels. However, interest in renewable sources is on the rise. (C) shows a comparison of four energy resources. The dashed line represents the minimum energy need of US (4 TW). Ethanol production from biomass is low efficiency [Brenner M. P. *et al.*, 2006]. Solar panels are efficient at harvesting sunlight and creating hydrogen fuel but are not cost-effective and lack hydrogen transportation infrastructure. Photosynthesis has higher energy conversion efficiency (33%) [Brenner M. P. *et al.*, 2006]. Proteorhodopsin exhibits an energy conversion efficiency of 21% - calculated using the energy of 525 nm photons and dividing it by the energy stored in a proton pump across a 0.2 V membrane potential. However, pR can absorb photons over a wide spectrum, thus capturing higher energy (50%) compared to only 37% by photosynthesis.

1.5 Variants of proteorhodopsin

The pR-expressing variants can be generally classified into two major groups based on their absorption maxima. The green absorbing pR is termed as GPR (Green absorbing pR) with a $\lambda_{\max} = 525$ nm, and the blue absorbing pR are BPR (Blue absorbing pR) with a $\lambda_{\max} = 490$ nm. These absorption maxima correspond to the depth of the habitat of bacteria; the bacteria that live on the surface of the water can absorb the longer wavelengths whereas the ones that live in deeper water obtain only the shorter wavelengths, in accordance with the spectral quality at the depths [Béjà O., *et al.*, 2000; Man D., *et al.*, 2003]. The difference in the GPR and BPR comes from the ability of the chromophore, retinal, to undergo “spectral tuning” according to its microenvironment [Birge R. R., 1990; Ottolenghi M. and Sheves M., 1989]. “Spectral tuning” refers to the large variation of the absorption spectrum of the chromophore as dictated by its interaction with the apoprotein or its environment. For example, in methanol, the protonated retinylidene Schiff base shows a $\lambda_{\max} = 440$ nm but in the protein environment, it exhibits an “opsin shift” [Nakanishi K., *et al.*, 1979] to a longer wavelength such as 525 nm in GPR or 490 nm in BPR.

The various mechanisms attributed to the phenomenon of “opsin shift” as studied in BR are: (1) Coplanarized conformation of retinal structure induced by the protein environment from a twisted β -ionone ring/polyene chain in solution [Nakanishi K., *et al.*, 1979], (2) A weakened PSB/counteranion association resulting either from an increase in the interionic distance from a solvation of counterion may lead to further shift in the absorption maximum to longer wavelengths [Blatz P. E. *et al.*, 1972; Blatz P.E. and

Mohler J. H. 1972], (3) Presence of negative charges i.e., polarizable groups and permanent dipoles along the polyene chain may lead to stabilization of the excited state and cause a red shift [Hubbard R. and Kropf A., 1958; Yan B., *et al.*, 1995], (4) The aromatic retinal binding pocket of the protein [Briman M. S. *et al.*, 1988; Rothschild K. J., *et al.*, 1989], composed of Trp residues, stabilizes the excited state of retinal leading to red shift.

The contribution of chromophore/protein interactions in the process of spectral tuning and opsin shift has been actively investigated in pR. Based on structural modeling and site directed mutagenesis studies, the single amino acid that is responsible for the difference between GPR and BPR was found to be located at the position 105; GPR contains Leu¹⁰⁵ and BPR has Gln¹⁰⁵. In BPRs, Gln105 is directly hydrogen-bonded to the Schiff base (SB) whereas in GPRs, a water molecule bonds to the SB [Kelemen B. R., *et al.*, 2003; Kralj J. M., *et al.*, 2008; Man D., *et al.*, 2003; Man-Aharonovich D., *et al.*, 2004; Wang W. W., *et al.*, 2003].

GPRs and BPRs have a similar primary structure but possess some other differences besides the absorption maximum. GPRs have a faster photocycle (half-time 20-45 ms) whereas BPRs show a slower photocycle (half-time 200 ms) [Béjà O., *et al.*, 2001; Man D., *et al.*, 2003; Wang W. W., *et al.*, 2003]. Another point of difference was found in the protonation state of a glutamate residue at the position 142. In GPRs Glu142 is normally protonated but in BPRs, it was found to be deprotonated at pH 7 or above [Kralj J. M., *et al.*, 2008].

It has been suggested that spectral tuning could also result from a combined effect of amino acids that form the retinal binding pocket and those that can impact changes in the conformation of the binding pocket [Bielawski J. P., *et al.*, 2004]. A recent study using random PCR mutagenesis to screen for color-tuning mutations found about 20 single residue pR mutants that showed a red- or blue-shift compared to the wild type [Kim S. Y., *et al.*, 2008]. Quite confirmatory to the above prediction, most of these mutations are located far from the retinal binding pocket. Other effects that can influence spectral tuning include factors that influence protein stability [Bielawski J. P., *et al.*, 2004] or interaction between helices [Shimono K., *et al.*, 2003].

1.6 Applications of proteorhodopsin

1.6.1 Holography and information storage

Information storage is mainly done via electronic memory, magnetic tape or optical discs. With chip sizes, capacities and speeds approaching the limits of technology, attention is being directed towards biomolecules with photochromic properties for the purpose of developing various electronic and computational devices: information storage, both two- and three-dimensional, and optical processing. The basic criterion for a molecule to qualify for such applications is the photochromic property, i.e., the ability to change color upon excitation by light, with two or more intermediates displaying distinct spectral properties. It should possess photo-cyclicity, meaning that it can undergo repeated transitions between these various photochromic states. Additionally, the molecule should also be small in size, as well as thermally and chemically stable.

Holography refers to use of light for the process of storing data. Photons, as they travel at the speed of light, enhance the speed at which data can be stored and accessed. Three-dimensional (3-D) information storage is ideal for high density memories. Proteins qualify for such an application as they exist as relatively small molecules. Another factor that is important for holographic application is the refractive index and change in the light-induced refractive-index [Hampp N. A., 2000]. The refractive index of bacteriorhodopsin is 1.47-1.55 [Zeisel D. and Hampp N. A., 1992; Zhang C., *et al.*, 1994], compared to 1.33 for water.

Bacteriorhodopsin (BR) is found in halobacteria that survive in extreme conditions of salt, such as 25% (5 M) NaCl [Wise K. J., *et al.*, 2002]. BR is found densely packed in the cell membranes of these cells arranged as uniformly oriented hexagonal crystalline arrays occurring as a two-dimensional crystal. BR is a small photochromic molecule that undergoes photocycle through several intermediates, each with a distinct spectral property. For example, BR (B state) absorbs at 570 nm whereas its most long-lived intermediate with deprotonated Schiff base (M state) exhibits an almost separated absorption maximum at 409 nm. These different wavelength maxima are the source of difference between written and blank spots. Despite the fact that M intermediate is the most stable intermediate for cellular proton pump function, it is not necessarily the optimal stability for the storage applications. A mutant, D96N, was found to increase the lifetime of this intermediate [Hampp N. A., 2000]. Therefore, the D96N mutant of BR has received more attention as a holographic memory material [Hampp N. A., 2000]. The optical properties of B and M states confer on BR the ability of being applied as

photochromic optical recording, long term information storage and holographic biomaterial.

A 3-D memory system was built by placing BR inside polyacrylamide gel in a cuvette [Birge R. R., 1995]. A combination of two one-photon excitations is crucial for applications. One wavelength can be used for writing information into BR and the other can be used for erasing it. The system was read by two lasers, one to activate the protein in a section of the cuvette and the other to read/write information in that section. The information is stored similarly to binary code where the initially excited state is assigned a binary value of 0 and the long-lived intermediates are designated 1. The storage capacity is enhanced because multiple sections can be activated and multiple data locations can be written simultaneously. [Birge R. R., 1995]

Proteorhodopsin is a homolog of BR with added advantages. Unlike BR, pR has been expressed heterologously in *E. coli* which divides every 20-30 minutes. Thus, pR can be made in larger quantities in relatively less time than BR. Detergent-solubilized pR has been found to be as stable as BR in the purple membrane. Proteorhodopsin exhibits a photocycle with similar intermediates and therefore, pR is also being tested for application in 3-D information storage and holography [Ranaghan M. J., *et al.*, 2010; Xi B., *et al.*, 2008].

1.6.2 Cosmetic ingredient

Color changing technology is the future of personal care and cosmetics. Some examples of such items have already been around for a while. Examples are “mood lipstick” that contains weak acid pigments and a conjugate base that absorbs at a different wavelength than the acid pigments; luster pigments that are refractive and reflective; thermochromic pigments that change color when exposed to heat found in the nail-polish; paint or kids’ garments; interactive cosmetics that change color to complement the complexion of the skin, interactive toothpastes with magnetic cleaning action [Narang R., 2008]; and photochromic pigments, such as the AgCl used in the Transitions™ lenses from Bausch and Lomb.

Photochromic materials such as metal oxides, for example, molybdenum and tungsten oxides, have been suggested for use in cosmetics [Chaiken J. and Birge R. R., 2002]. The idea is that upon light absorption, these photochromic materials would maintain or evolve a particular coloration. These materials are light-dependent; therefore, they would change the color depending upon where the person is. BR was also included in this patent as a cosmetic application. BR exhibits a wide absorption in the visible region, 350-630 nm. By selecting certain mutations, the coloration effect can be manipulated to last for a longer duration. These inorganic metal oxides or protein-based formulations are relatively risk-free compared to organic dyes such as azo-dyes or quinolines. Additionally, due to the fact that BR contains 8 tryptophan, 11 tyrosine, and 13 phenylalanine residues, it absorbs strongly in the UV region, 180-330 nm. This has an added advantage of making BR also effective as a UV blocking agent.

Proteorhodopsin exhibits similar photochromic properties as BR. Therefore, it can also be used for such an application. It is a brightly purple colored protein with an intermediate (M) that has a blue-shifted λ_{\max} . The M-intermediate can be made to live longer by creating a mutation in the proton donor group (E108Q). The corresponding lifetime increases by 10-20 times that of the wild type pR [Dioumaev A. K. *et al.*, 2002]. Additionally, proteorhodopsin contains 10 tryptophan, 14 tyrosine, and 15 phenylalanine residues (Table 1.1) which renders it suitable as a sun-blocking agent, similar to BR.

1.6.3 Bioenergy production

The increasing demand for energy is a global issue. Concerns of security and longevity of traditional energy sources has directed research towards alternative fuels and energy production. The annual energy consumption of the US is close to 3.5 Terawatts (TW), and the world energy consumption is approximated at 16 TW. More than 80% of the world energy supply is dependent upon fossil fuels and the current level of CO₂ emissions is 380 ppm compared to the pre-industrial values of 280 ppm. If the current levels are maintained, the CO₂ levels have been projected to rise to 550 ppm by the end of the century [Brenner M. P. *et al.*, 2006]. In the year 2010, the nation's energy supply was dominated by petroleum, followed by natural gas, coal and nuclear power in that order (Fig. 1.11). Compared to these traditional energy resources, the renewable energy consumption was relatively low. However, its use has grown dramatically over years. Renewable alternatives, such as wind, solar, thermal power, and biofuels have attracted much attention in the recent times, both to help achieve fuel-independence and to limit CO₂ emissions.

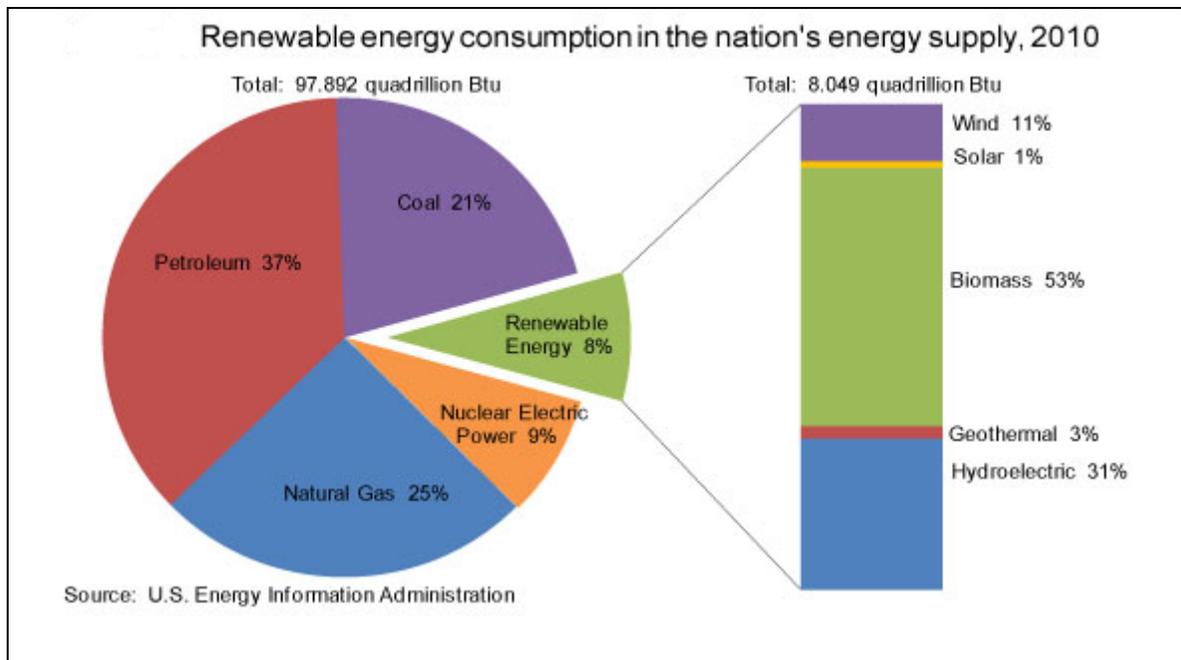


Fig. 1.11: Energy chart: Left: Various energy sources available in the US, Right: Types of renewable energy sources. The consumption values are based on the year 2010. Figure is taken from U. S. Energy Information Administration website (www.eia.gov).

In recent years, research on the use of microorganisms for energy production has been on the rise. For example, ethanol is produced as a metabolic by-product and is being used as an alternative fuel. The aim is to manipulate these microorganisms for the production of liquid hydrocarbons. Since the biological systems are regulated by numerous feedbacks, there are technical engineering hurdles to be overcome. Meanwhile, we can focus on the microorganisms or systems with relatively simpler modes of energy production.

Bacteria and Archea use relatively simple light driven proton-pumps to harvest solar energy. The proton pumps generate proton motive force (pmf) which can then be used to synthesize energy in the form of ATP. The fact that pR can transform solar

energy into useful cellular energy and that it can be functionally expressed in a heterologous system make pR a potential candidate for use in energy production [Walter J. M., 2010]. There is an issue of coupling pR to another protein or system that would synthesize a useful form of energy instead of ATP. Another issue is that pR functions only under circumstances where the cells are limited for ATP or pmf. Furthermore, pR functions to increase the pmf only when the pmf falls below the threshold pmf [Walter J. M., 2010; Walter J. M., *et al.*, 2007]. Metabolic engineering of microbes has had some success, for example, introduction of a glucose transporter gene into an obligate photoautotroph converted it into a heterotroph. It was shown that this microalga survived on glucose in the dark instead of light [Zaslavskaja L. A., *et al.*, 2001]. Proteorhodopsin therefore has great potential for being used in conjunction with another enzyme(s) to produce energy.

1.7 Purification of protein

Why is purification of pR important? The closest relative of pR is the archeal BR which has been a popular protein for applications such as the development of optoelectronic, 2-D and 3-D optical storage, cosmetics and potentially for bioenergy production. However, growth of halobacteria is quite slow with doubling time typically at least 10-12 h [Deshpande A. and Sonar S., 1999; Seehra J. S. and Khorana H. G., 1984]. At the same time, heterologous expression and purification of the native BR has been a considerable challenge due to the inadequate accessibility to the genetic systems for optimization of BR production [Wise K. J., *et al.*, 2002]. On the other hand, pR has been successfully expressed in *E. coli*, possesses similar properties as BR and therefore, has

attracted much attention for similar applications [Ranaghan M. J., *et al.*, 2010; Xi B., *et al.*, 2008].

In addition, the three-dimensional structure of pR has not been solved yet. In order to study the 3D structure, one needs a good quality crystal that can diffract light. To utilize the protein for either its application purposes or carry out crystallography, one needs a significantly high amount of functionally active and pure protein, typically 10-15 mg/ml to at least >90% homogeneity. Since pR is readily expressed in *E. coli*, it is easy to grow these cells in larger volumes. The next step is the isolation and purification of the protein.

1.7.1 General Process of Purification of Protein

Protein purification is the process of isolation and separation of a particular protein from the cellular extract that contains thousands of proteins. Each protein has its own physical and chemical properties and therefore, they can be separated from each other on the basis of the properties such as solubility, size, charge, isoelectric point, hydrophobicity, binding-affinity, post-translational modification, etc. Different methods of purification are used, depending on the property of the protein.

Generally, for purifying intrinsic membrane proteins, the cells that express the proteins are lysed and the membranes are subjected to solubilization in a detergent, either before or after a centrifugation step to separate the membranes from water-soluble cytosolic components. The crude detergent-solubilized protein is then subjected to a

series of steps that gradually isolate the protein from the mixture and eventually yield the pure protein. Traditional protein purification methods include the use of salts or solvents that aid in fractionating the protein of interest from the extract. It is followed by a combination of different column chromatography methods, dialysis and filtration.

Solubility of proteins is one of the properties that is utilized in the initial steps of protein purification. The differing solubilities of proteins depend on pH of the solution, ionic strength of the salts, polarity of the solvent, nature of the ions, temperature, and isoelectric point. Proteins are generally least soluble at their isoelectric point, since there is less Coulombic repulsion between identically-charged molecules of the protein.

The aim is to either get the best yield of the protein of interest or to have the most pure form of protein. In order to keep a track of the protein of interest, one needs an activity assay to measure the protein's concentration and activity. At each step, the different fractions are assayed for protein. Then, that particular fraction that contains the protein is subjected to chromatography technique(s) to separate it from the other contaminant proteins.

1.7.2 Membrane protein purification

Membrane proteins are the proteins that are found associated with the membranes of a cell or its organelles. They are generally categorized as trans-membrane proteins, also known as integral proteins, and peripheral membrane proteins. Since the membrane proteins are surrounded by lipids, their isolation is considerably challenging as compared

to the soluble cytoplasmic proteins. The membrane proteins are hydrophobic and therefore, when they are separated from the membranes, they tend to aggregate and precipitate from an aqueous solution. In order to keep the membrane proteins solubilized, one has to employ detergents. Detergents have affinity for water as well as the hydrophobic amino acid residues. They can break down the phospholipid bilayer of the membrane and solubilize the lipids and proteins. In this process, they form micelles and therefore prevent the protein from forming aggregates.

Integral and peripheral membrane proteins are involved in vital roles in cellular functions, such as cell signaling, cell adhesion, transport of molecules, and maintenance of cell or organelle structure, to name a few. About 30% of the mammalian genome Open Reading Frames (ORFs) has been found to encode integral membrane proteins [Stevens T.J. and Arkin I.T., 2000]. Despite their importance in the key functions, only 290 membrane proteins have had their three-dimensional (3D) structure elucidated to date [Stephen W., membrane protein structure database] compared to the available 3D structures of >43000 soluble proteins [NCBI database].

About 2-4% of the human genome has been predicted to code for G-protein coupled receptors (GPCRs), which are trans-membrane proteins. They are involved in a variety of functions such as visual sensing, smell sensing, regulation of the immune system etc. Many diseases involve the malfunctioning of GPCRs, some of which are blindness, allergies, diabetes, cardio-vascular defects. Approximately 50% of currently marketed drugs target GPCRs [Lundstrom K., 2006]. Yet, the 3D structures of only the

following six GPCRs are known so far: Bovine rhodopsin [Palczewski K., *et al.*, 2000], β 2 Adrenergic Receptor [Cherezov V. *et al.*, 2007; Rasmussen S. G., *et al.*, 2007], β 1 Adrenergic Receptor [Warne T., *et al.*, 2008], human A_{2a} Adenosine receptor [Jaakola V. P., *et al.*, 2008], CxCR4 chemokine receptor [Wu B., *et al.*, 2010] and, Dopamine D3 receptor [Chien, E.Y. *et al.*, 2010].

The major obstacles in obtaining structures of GPCRs, or membrane proteins in general, involve protein expression, large scale production and purification, protein stability, and homogeneity. Selecting the appropriate detergent is one of the challenges that faces the process of membrane protein purification. The chosen detergent should keep the protein solubilized and stable, but also not interfere in the crystal formation. For example, the detergent, nonyl- β -glucoside, that was used for rhodopsin crystallization [Stenkamp R. E., *et al.*, 2002] is not suitable for the crystal formation of another GPCR, β 2 adrenoceptor (β 2AR). Similarly, the detergent, dodecyl maltoside, used for β 2AR was not found to be suitable for rhodopsin crystallization.

1.7.3 Salting out: the Hofmeister series

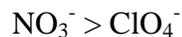
When an impure protein solution is subjected to an increasing amount of salt it begins to precipitate out of the solution. This phenomenon is known as “salting-out”. The salt concentration at which the precipitation begins is different for different types of protein. Thus, different proteins are separated from the solution at different concentrations of the salt and are thus fractionated.

Ammonium sulfate is often used for the salting-out of proteins because it is highly water soluble, and does not denature the proteins. It also utilizes two ions, ammonium and sulfate, that are high in the Hofmeister series (see below), which gauges the relative effectiveness of different types of ions at causing salting-out of proteins. At a fixed salt concentration, ions higher in the Hofmeister series are more likely to cause salting out.

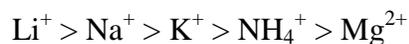
Specific ion effects were long ago found to display recurring trends in the salting-out of proteins [Hofmeister F., 1888; Kunz W., *et al.*, 2004]. Such trends are more pronounced for anions than for cations.

Hofmeister measured the concentration of various salts needed to precipitate the proteins from the whole egg white. He arranged the salts as a series of anions and cations referred to as the Hofmeister series. More salts have been examined and added to the series, a partial list is shown below.

For the same cation with different anions, the order is:



For the same anion, the order for cations is:



Ions in the Hofmeister series were originally thought to cause their effects indirectly by influencing the structure of water, specifically by changing its hydrogen-

bonding properties [Collins K. D. and Washabaugh M. W., 1985]. Therefore, the Hofmeister series is arranged in the order of these ions' ability to make or break hydrogen-bonded water structures. The species to the left of Cl⁻ are known as kosmotropes, while those to its right are called chaotropes. Kosmotropes were believed to promote hydrogen-bonded water structure. They are strongly hydrated, stabilize the macromolecules and exhibit salting-out effects on proteins. Chaotropes are the 'water structure breakers'; they destabilize and cause salting-in of the macromolecules. Hofmeister [Hofmeister F., 1888; Kunz W., *et al.*, 2004] concluded that the better the capacity of a salt to order the water i.e., to be solvated (hydrated), the more effective it is in precipitating the protein.

Recently, it has been shown by various research groups that ions do not affect the bulk water properties. It was shown using femtosecond mid-infrared pump-probe spectroscopy that anions have no influence on the hydrogen bonding network outside the direct vicinity of the anion [Omta A.W. and Bakker H.J., 2003]. It was concluded that, at least at the relatively low concentrations that can lead to salting-out phenomena, neither kosmotropes nor chaotropes are involved in long-range structure making or structure-breaking effects. Another group [Batchelor J.D., *et al.*, 2004] conducted thermodynamic studies by using different solutes known for their stabilizing or destabilizing effects on proteins. They collected data as a function of temperature and calculated the coefficient of thermal expansion. Their results showed no obvious correlation between kosmotropes/chaotropes and their sign (+/-) of the coefficient value. Yet one more challenge to Hofmeister's conclusion came from a study of the effect of Hofmeister

anions on phase transition of a surfactant monolayer of octadecylamine and its adjacent water structure. Gurau *et al.* [Gurau M. C., *et al.*, 2004] showed that anions followed the Hofmeister series in the stabilization of the monolayer, but showed a significant deviation when it came to the structuring of water. The ordering of the monolayer by the anions is suggested to be due to their ability to penetrate the headgroup region and cause disruption in the hydrocarbon packing [Aroti A., *et al.*, 2004].

It has been suggested that direct ion-macromolecule interactions are responsible for the phenomenon of salting-in or salting-out behavior [Zhang Y. and Cremer P. S., 2006]. Anions show two trends in the way they effect protein precipitation, depending upon the pH of the solution [Bostrom M. *et al.*, 2005]. At a pH above the isoelectric point (pI) of the protein, the protein possesses a net negative charge and a direct Hofmeister series is observed. “Direct Hofmeister series” refers to the phenomenon where the kosmotropes stabilize the folded state of the protein and lead to the salting-out effect, whereas the chaotropes lead to destabilization and salting-in of proteins. At a pH below the pI, the net charge is positive. In this range, an inverse Hofmeister effect is observed, where Hofmeister’s “kosmotropes” lead to salting-in and “chaotropes” cause the protein to salt-out of the solution.

The above mentioned effects of salts are observed mostly on soluble proteins. The effects of the Hofmeister series on membrane proteins have only recently become a matter of interest. Conformational equilibrium studies were performed on the artificial visual pigment 9-demethyl-rhodopsin (9dm-Rho), where the retinal was replaced by 9-

demethyl retinal [Vogel R. and Siebert F., 2001]. It was found that the well known Hofmeister cosmotropes were unsuccessful in stabilizing 9dm-Rho because it was mainly stabilized by the phospholipid bilayer. All the salts also caused an equilibrium shift from the MI intermediate to the less compact MII intermediate. Another study on bacteriorhodopsin led to similar conclusion [De´r, A. and Ramsden J. J., 1998]. The denaturant, NaSCN, accelerated the decay of the M intermediate of bR whereas the stabilizer, NaF, did not show any kinetic effect. Thus, the effect of Hofmeister salts on the membrane protein is not as strong as the soluble proteins.

1.8 Widely used method of pR purification

Proteorhodopsin was originally cloned into the plasmid pBAD TOPO TA which adds a polyHis-tag at the C-terminus of pR. The purification of proteorhodopsin has normally depended on the use of the polyHis-tag that binds to the Nickel column or Nickel-nitroloacetic acid (NTA) agarose resin beads [Bergo V., *et al.*, 2004; Friedrich T. *et al.*, 2002; Gourdon P. *et al.*, 2008; Jung J. Y., *et al.*, 2008; Kelemen B. R., *et al.*, 2003], where Nickel is bound to agarose via NTA. In short, Histidine-binding to Ni²⁺ is a type of affinity binding that has been exploited as a column chromatography for purification purposes. The His-tagged protein binds to the Ni²⁺ on the resin, the other proteins are washed away with a very low concentration of imidazole, and thus removed. The bound protein is then eluted with a high concentration of imidazole that competes with the poly-histidine group on the target protein. This type of purification yields a fairly pure protein. However, there can be some contamination from the rare proteins that contain poly-histidine residues or from non-specific binding to Ni²⁺.

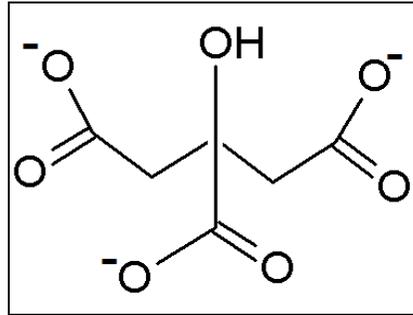
1.9 Our method of pR purification

Our lab developed an inexpensive method of pR purification using citrate as the precipitant [Partha R. and Braiman M.S., 2003; Partha R., *et al.*, 2005]. The utility of citrate was discovered serendipitously. Citrate was initially employed as one of the buffers during the pH titrations in order to determine the pK_a of the proton acceptor group Asp⁹⁷ [Partha R., *et al.*, 2005]. It was observed that upon incubation with citrate, pR begins to precipitate out of the solution. This observation led to the hypothesis that citrate could be used as a selective precipitant of pR and thus, aid in pR purification. It was found that a detergent concentration of 0.3% at pH 5.5 with a 50 mM citrate concentration yielded the most intense purple pellet of pR.

His-tagged pR and several select site-directed pR mutants have been purified with citrate to 50% purity and 48% yield. The advantages of using citrate include its ability to purify the protein in a few steps, eliminating the need for high-speed centrifugation, which is typical in membrane protein purification procedures. Thus, overall, the citrate method of protein purification is inexpensive at various levels, easy and fast. Further purification was carried out using Nickel-column to yield 95% purity and 41% yield.

1.10 Citrate

Citric acid is a tricarboxylic acid. Its conjugate base is known as citrate. Citric acid is a weak organic acid and is found in citrus fruits. The IUPAC name of citric acid is 2-hydroxypropane-1, 2, 3-tricarboxylic acid. The structure of citrate and its pK_a values are shown below:



$pK_1 = 3.13$
 $pK_2 = 4.76$
 $pK_3 = 6.40$

Citrate appears as an intermediate in the citric acid cycle, also known as tricarboxylic acid (TCA) cycle or Krebs cycle. After the food has been broken down into glucose, glucose enters a cycle known as glycolysis. Glycolysis occurs in the cytosol of the cell, it precedes the TCA cycle which is followed by the electron transport chain. The end product of glycolysis is pyruvate. Pyruvate is oxidatively decarboxylated to form acetyl CoA which enters TCA cycle. A four-carbon compound, oxaloacetate, condenses with a two-carbon acetyl unit, acetyl CoA, to generate a six-carbon tricarboxylic acid compound known as citrate. Citrate undergoes a series of reactions to generate energy rich molecules, NADH, FADH₂ and GTP. NADH and FADH₂ are oxidized in the electron transport chain to generate ATP. Besides, the intermediates of TCA cycle also serve as a source of building blocks for biosynthesis. For example, α -ketoglutarate and oxaloacetate, the intermediates of TCA, are the source of many amino acids. [Stryer L]

1.11 Citrate does not degrade pR

The purification process requires that the reagents do not destabilize and inactivate the protein. Proteorhodopsin is a retinal protein and absorbs at $\lambda_{max} = 520-545$ nm depending on the pH. Citrate precipitated pR is found to remain stable, evident from the 545 nm absorption maximum at pH 7.0. This indicates that pR is properly folded and

is bound to its chromophore retinal. A similar absorption maximum and an identical pK_a of 8.1 ± 0.2 was obtained for the protonated Schiff base counterion Asp97 from partially purified citrate extracts as from a freshly-prepared pR purified by using other methods [Partha R., *et al.*, 2005].

1.12 My Research

My research has focused on two main issues. One was on understanding the nature of the interaction of pR with citrate and other anions. The other was to extend this study to develop a simple method for the purification of membrane proteins. The final goal of my study is to apply these methods first to other similar membrane proteins and then to extend it to all the membrane proteins.

Study of the nature of pR-citrate interaction revealed the importance of four positively charged amino acid residues found in the first intracellular loop of pR. Study of interaction of several other anions led to the development of a purification protocol of a membrane protein pR exclusively using simple salts.

The future prospect of this study lies in the application of these two findings for the purification of other membrane proteins of general and pharmaceutical importance. An attempt was made first to express mammalian trans-membrane protein in *E. coli* and then to purify it using our established citrate method of protein purification. There are several hurdles that remain to be overcome before the protocol can be standardized.

Chapter 2

CITRATE BINDING SITE IN PROTEORHODOPSIN INVOLVES TWO LYSINES IN THE FIRST CYTOPLASMIC LOOP

2.1 ABSTRACT

Proteorhodopsin (pR) has one or more binding sites for citrate, as evidenced by aggregation of native purified pR from octylglucoside detergent solutions, at citrate concentrations as low as ~10 mM (pH 5.5-7.5). The aggregation is reversible by dilution of citrate and addition of detergent, at pH>9. Other anions, e.g. glutarate, glutamate, phosphate, and chloride, also cause aggregation, but only at higher concentrations. Citrate binding requires lysines 57 and 59, two of the four cationic residues in the first intracellular loop. That is, citrate no longer causes precipitation of pR when these lysines are mutated. The loss of citrate sensitivity is nearly complete for the quadruple mutant R51Q/R53Q/K57Q/K59Q, the triple mutant R53Q/K57Q/K59Q, and the double mutant K57Q/K59Q, even with 1 M citrate. In contrast, the single-site mutant R51Q shows only a ~2-fold increase in the citrate concentration required to induce pR aggregation (to ~30 mM), while R53Q appears not to fold stably. Thus, the 2 arginines play a less certain role in the citrate binding site. Lysines 57 and 59 are conserved among disparate γ -proteobacteria. Aggregation of pR in detergent micelles requires citrate or glutamate concentrations similar to endogenous levels, ~10 mM and ~100 mM respectively, that

have been measured inside thriving bacteria. However, we have obtained no evidence for citrate binding for pR in a reconstituted lipid bilayer environment. In particular, flash photolysis of pR liposomes shows no effect of citrate on the yield or kinetics of the M intermediate.

2.2 INTRODUCTION

Proteorhodopsin (pR), originally discovered in an as-yet uncultivated species of the SAR86 group of γ -proteobacteria, was the first-discovered eubacterial homolog of bacteriorhodopsin [1, 2]. It is estimated that the world's oceans contain at least 10^{28} pR-producing bacteria in the photic zone [3]. Since its first discovery, numerous pR-directed surveys of marine DNA samples have been conducted, and subjected to metagenomic analysis. Variants of pR genes have been found widespread in diverse bacterioplankton groups from widely-dispersed geographic locations [4-9]. A global ocean sampling expedition uncovered more pR related genes from the surface marine environment [9]. More recently, new pR genes have been discovered in estuarine and inland bodies of water, leading to the conclusion that pR genes are widespread, i.e., not just confined to marine environments [10].

When heterologously expressed in *E. coli*, pR functions as a light-driven proton pump, with a net outward transport of protons, but only in the presence of exogenously added retinal [1]. The ubiquity of pR in the Earth's oceans and fresh water bodies, and its ability to pump protons resulting in photophosphorylation, i.e. ATP formation [1, 11-13],

together raise the possibility that pR photochemistry contributes a considerable fraction of solar energy utilization by the biosphere.

Recently, our lab discovered that citrate, at concentrations of 20 mM, i.e., only a bit above those previously measured intracellularly [14], selectively precipitates pR from a detergent extract of pR-expressing *E. coli* cells, making this a useful step in the purification of a membrane protein [16, 17].

The current work was aimed primarily at identifying the structural basis for pR aggregation in the presence of citrate, specifically, the nature of interaction of pR with the three carboxylate groups of citrate. The degree of specificity of such an interaction is studied by altering the primary structure of the pR protein through site-directed mutagenesis. It was found that a positively-charged cluster of amino acid residues in the first loop, between the 1st and 2nd helices of pR, is involved in mediating the interaction between citrate and pR.

It was also found that besides citrate, pR can also be precipitated by other polycarboxylates, phosphate and chloride, albeit at concentrations that are higher, especially relative to their physiological levels.

Aggregation of pR from detergent solutions is initiated over the physiological concentration range of citrate in bacteria. This is of particular interest in light of the conclusion [11] that marine bacteria might also employ pR as a solar energy transducer

selectively under conditions associated with low respiratory rates, as a result of such conditions triggering pR expression or activity [11]. Citrate is a key intermediate in the tricarboxylic acid (TCA) cycle and is an allosteric regulator of many enzymes in bioenergetic pathways. Therefore, it was hypothesized that the utility of citrate as an *in vitro* precipitant for pR might in fact reflect an *in vivo* interaction between citrate and pR, i.e., citrate might mediate an “on/off” switch for pR activity as a function of respiratory state of the bacteria.

However, experiments on pR in reconstituted liposomes failed to support this hypothesis. In fact, flash photolysis of the reconstituted pR shows no effect of citrate on the millisecond kinetics or amplitude of the photocycle. This kinetics result casts doubt on the hypothesis that citrate could be regulating pR activity in the bacterial cell membrane.

2.3 MATERIALS AND METHODS

2.3.1 Protein Expression

Proteorhodopsin triple cysteine mutant (pR-TCM), i.e. the protein corresponding to accession number AAG10475 but with all 3 cysteines mutated to serine, was expressed from a pBAD-TOPO plasmid in *E. coli*, with minor modifications to protocols described previously [18, 19]. The pR-TCM was expressed in the native form i.e. without the C-terminal polyhistidine tag. Cultures (50-500 mL) were grown in an enriched LB medium (0.5% yeast extract, 1% peptone, 50 mM KH₂PO₄, 50 mM Na₂HPO₄, 1% glycerol, pH-adjusted to 7.0).

For induction of pR, the temperature of the shaking incubator was reduced to 15°C, L-arabinose was added to a final concentration of 0.2%, and all-*trans*-retinal was added to a final concentration of 5 μ M [19]. The culture was left shaking at 50 rpm in the dark at 15°C overnight [19]. Cells were harvested by centrifugation in 500-mL bottles (6000 rpm, 4°C, 10 min).

2.3.2 Protein Purification

A crude detergent extract of pR was prepared by gently resuspending the cell pellet in 10 volumes of 10 mM HEPES buffer, pH 7.1, containing 3% *n*-octyl- β -D-glucoside (OG), 0.08% lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF) protease inhibitor, and 0.01 mg/mL DNase, and incubating at room temperature for 3 h or at 4°C overnight. Cellular debris was removed by centrifuging the lysate at 2500 \times *g*.

To the purple-colored supernatant, a 1/3-volume of 100 mM citrate buffer (pH 5.5) was added with gentle mixing, to give a citrate concentration of 33 mM. This sample was incubated on ice for 15 min, and then centrifuged at 3000 \times *g*, 4°C to precipitate out whitish impurities. Addition of a same-sized second volume of the same 100 mM citrate buffer (to 40 mM) was followed by incubation and centrifugation to remove more of the uncolored contaminants, while losing at most a small amount of the pR as precipitate. A third addition of 100 mM citrate, to a final citrate concentration of 50 mM and final OG concentration of 1.5%, followed by centrifugation, yielded a bright purple pellet containing most of the pR, with almost no color left in the supernatant.

The volume of the purple pellet was estimated, then it was gently re-solubilized in a 10-fold larger volume of 50 mM Tris-Cl buffer, pH 9.1, containing 3% OG. This sample was then centrifuged to remove grayish insoluble material. If any significant purplish color remained in the pellet, it was re-extracted with the same pH 9.1 buffer containing 3% OG. The (combined) clear purple supernatant was again subjected to selective precipitation by addition of aliquots of 100 mM citrate, pH 5.5, as described in the preceding paragraph. However, with increasingly pure pR, the total citrate concentration needed to precipitate the pR decreases, so smaller relative volumes were used.

A total of 3-4 cycles, each consisting of selective precipitation of pR by citrate followed by resolubilization in Tris buffer/detergent, yielded pR with A_{280}/A_{520} ratio of 4-10. This sample was then applied to a freshly-poured 10 cm×1 cm hydroxylapatite, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$, (BioGel HTP, BioRad) column, poured in 500 mM KCl, 100 mM acetate (pH 6). The bound protein was eluted with a 0–600 mM phosphate gradient (pH 6) containing a constant 0.4% OG. The uncolored eluate was discarded and 2 mL fractions were collected as soon as the purple color began to elute.

The eluted purple fractions were pooled, concentrated with Amicon filters (MW cut-off 10,000 kDa), and washed in the filters with 10 mM HEPES, 0.4% OG. The concentrated sample was further dialyzed in a Tube-O-dialyzer (Gene Mate, MW cut-off 4,000 kDa) against 10 mM HEPES 0.4% OG, with several changes, to remove the phosphate completely.

For the quadruple mutant R51Q/R53Q/K57Q/K59Q, the triple mutant R53Q/K57Q/K59Q, and the double mutant K57Q/K59Q, even at 1 M citrate it was not possible to get the pR to precipitate from the crude cell extract. For these mutants, ~1 mL of this crude detergent extract of the cell pellet was applied directly to the hydroxylapatite column, and eluted as described above. This yielded only partial purification, i.e. an A_{280}/A_{520} ratio of ~10.

2.3.3 UV-visible absorbance measurements

A Shimadzu UV-265 spectrophotometer was used to measure quantities of pR remaining in the supernatant of pR samples purified as above, both before and after addition of various amounts of 100 mM citrate buffer (pH 5.5), followed by centrifugation to remove aggregated pR. The data collected were analyzed using MIDAC-GRAMS.RTM. (Galactic Industries) software. The concentration of pR in the sample was determined by the value of A_{520} , after correction of the baseline by subtraction of the absorbance reading at 700 nm.

2.3.4 Site Directed Mutagenesis

Proteorhodopsin mutants, based on the pR TCM template DNA, were generated using a QuikChange Site Directed Mutagenesis kit (Stratagene), with primers synthesized by Integrated DNA Technology. The mutated plasmid DNA was recovered from *E. coli* cultures and was sequenced at the Sequencing Facility of SUNY Upstate Medical University.

2.3.5 Test for aggregation with diverse anions

Proteorhodopsin was purified with several rounds of citrate buffer as mentioned above. It was resolubilized in Tris 50 mM, pH 9 and also containing 3% OG. It was tested for precipitation with anions of various monocarboxylates (acetate), dicarboxylates (succinate, dodacanedioate, aspartate, glutamate and glutarate) and tricarboxylates (isocitrate and cis-aconitate); as well as with monovalent and polyvalent ions such as phosphates, sulfates (ammonium sulfate, sodium sulfate, sodium thiosulfate), chlorides (sodium chloride, potassium chloride), sodium chlorite, larger polyvalent ions such as glucose-6-phosphate (G6P), adenosine triphosphate (ATP), and adenosine diphosphate (ADP). All of the above salts and acids were buffered to pH 7.5 with Tris, 10 mM. The final concentration of OG was 1%. Proteorhodopsin was incubated at 4°C with above mentioned buffers ranging from 5 mM to 200 mM. Then the samples were centrifuged at 2000 rpm, 20°C, 5 minutes.

2.3.6 Reconstitution of pR into DOPC liposomes

Proteorhodopsin was purified with citrate (as above) to an $A_{280/520}$ ratio of about 3. The resolubilized pR was diluted to a concentration of 1 mg/mL in 50 mM Tris-Cl, 3% OG, pH 9. Reconstituted membrane vesicles of pR were prepared with some modifications of protocol described previously [20]. OG (800 mg) was added to 10 mL of 20 mg/mL dioleoylphosphatidylcholine (DOPC) in chloroform, and the mixture was dried using SpeedVacTM. The dry powder of the lipid-detergent mixture was resuspended in 40 mL of 50 mM MES and 300 mM NaCl, pH adjusted to 6, to achieve a final lipid concentration of 5 mg/mL. It was then sonicated until a clear solution was obtained. This

lipid-detergent solution (60 μ L) was added to the partially pure pR (1 mL, 1mg/mL). The lipid-protein mixture was incubated at room temperature for 1h, then transferred to dialysis tubing (4,000 kDa cut-off) and dialyzed for 1 week at room temperature against 50 mM Tris-Cl, 100 mM NaCl, 10 mM MgCl₂, 3 mM NaN₃, and 5 mM DTT, pH 8.5. Unlike in reference 20, Methyl 2-4, Pentanediol (MPD) was not included in the dialysis buffer. After a week of dialysis, the reconstituted sample was washed several times with 50 mM Tris buffer containing 3 mM NaN₃, pH 8.5.

Citrate-containing pR vesicles were made by including 20 mM citrate in the vesicle-forming dialysis step.

2.3.7 Flash spectroscopy and kinetic analysis

Flash-induced kinetics of the M intermediate of the reconstituted pR vesicles were measured in a custom-built system using a 3-mL quartz cuvette. The excitation laser was a frequency-doubled Nd⁺-YAG laser flash (532 nm, 8 ns, \sim 10 mJ cm⁻²). The measuring light was a battery-powered 400-nm light-emitting diode, using a series of lenses to focus its output first through the illuminated region of the cuvette, and then onto an AC-coupled silicon photodiode detector. A 532-nm notch filter was used to reduce the size of the laser flash artifact. Phototransients in the millisecond time range were acquired by using a LeCroy Model 9400A Dual 250 MHz Digital Oscilloscope. The total number of points measured was 25000 at a spacing of 4 μ s, for a total duration of 100 ms. Each sample's transient intensity-change signal was obtained by averaging 5000 acquisition sweeps, then converted to a percent transmission change by dividing by the

intensity drop produced with a mechanical chopper wheel blocking the measuring beam intermittently at ~30 Hz. After removal of the first 0.1 ms of data following the flash, to remove the flash artifact and the rise of the M intermediate, each averaged absorption transient was fitted to a first-order kinetic equation for M decay, $A_{400}(t) = B + A_0 e^{-kt}$ by using the solver function in Excel, optimizing for values of the baseline absorbance B , the transient amplitude A_0 , and the first-order decay constant k .

2.3.8 Homology modeling and molecular graphics imaging

Homology modeling of pR was based on the crystal structure of bacteriorhodopsin (bR) from *Halobacterium salinarum* (PDB: 1FBB). It was produced using Geno3D, an automated protein-modeling Web server [21]. Molecular graphics images of pR-citrate interactions were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco [22].

2.3.9 Distance Tree of proteorhodopsin related proteins

A distance tree of bR/pR homologs was generated using BLAST pairwise alignment [23]. The protein sequence of pR was aligned against the protein database using the BLASTp algorithm, and the distance tree was produced using fast minimum evolution tree method. The resulting file (newick format) was manipulated with TreeViewX [24].

2.4 RESULTS

2.4.1 Purification of proteorhodopsin without His-tag

Proteorhodopsin (pR) with the polyHis-tag has previously been purified by selective precipitation with sodium citrate, from a level of 0.1% of total protein in the crude cell lysate to 20-25% [16, 17]. Those earlier results were confirmed now also with pR TCM lacking the non-native C-terminal His₆-tag, thereby showing that citrate is not likely interacting with these histidines. By using this rapid citrate purification method, followed by a single column step using hydroxylapatite (BioGelHTP), pR-TCM (without any non-native His tag) can be routinely purified to give an A_{280}/A_{520} ratio of <3.0. When extra care is taken, this can be improved to a ratio of 2.0, indicative of >95% purity (Fig. 2.1). With an elution buffer containing 0.4% OG, most of the pR elutes from the hydroxylapatite at 570-600 mM phosphate (pH 6.0).

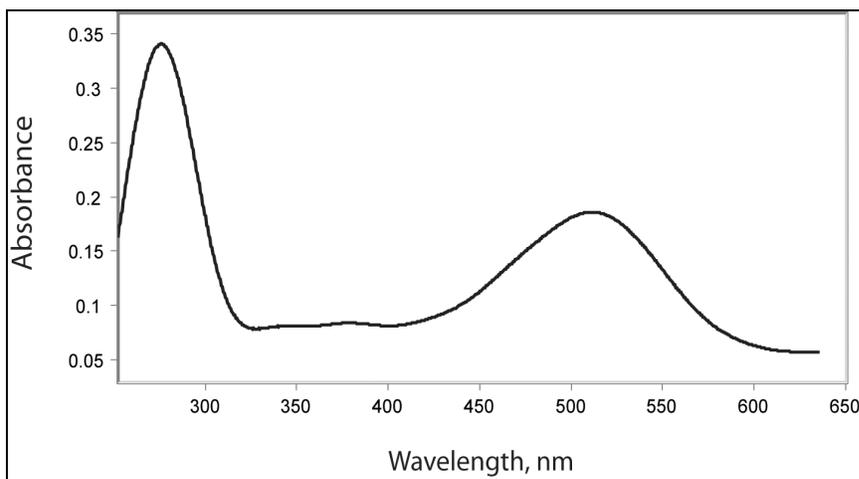


Fig. 2.1: Absorption spectrum of pR triple cysteine mutant without the His₆-tag: The protein was purified by 3 repeated cycles of citrate aggregation followed by resolubilization, then loaded on hydroxylapatite column, eluted with phosphate gradient containing 0.4% OG, and dialysed against 10mM HEPES, 0.4% OG. The A_{280}/A_{520} ratio for this sample is 2.1, indicating >90% purity.

It was also of interest to confirm that this column purification did not strip the pR of any co-factors that might be required for precipitating it with citrate. To establish that, the samples of pR purified by the BioGel HTP column were concentrated, phosphate was removed by extensive dialysis against 10 mM HEPES containing 0.4% OG, then 1/3 volume of 100 mM citrate buffer (pH 5.5) also containing 0.4% OG was added, i.e. to a final citrate concentration of 33 mM. After incubation on ice for 10 min followed by centrifugation, the purified pR TCM sample precipitated out of the solution quantitatively. This indicates that pure pR specifically interacts with citrate, with no dependence on co-precipitating protein(s) or other co-factors. Thus, pR itself has binding site(s) for citrate.

It is of interest that failure to remove phosphate after the hydroxylapatite column results in insensitivity to citrate. That is, even at fairly low (50 mM) concentrations, phosphate competes for the citrate binding site on pR, completely inhibiting the ability of similar (50 mM) citrate concentrations to precipitate pR. This observation establishes definitively that the citrate-induced aggregation of pR is not merely a salting-out of the protein. A salting-out process by 50 mM citrate should be enhanced, rather than inhibited, by the additional presence of 50 mM phosphate.

2.4.2 Lysines in the first cytoplasmic loop on pR mediate citrate interactions

It was hypothesized that citrate, a tricarboxylate with average charge of -2.5 in aqueous solution at pH 5.5, would interact with positively charged amino acid(s) on pR. There are 12 positively charged lysine and arginine residues in the entire primary

sequence, of which 8 are on the cytoplasmic surface, and 4 are grouped in the first intracellular loop (IC1): Arg⁵¹, Arg⁵³, Lys⁵⁷ and Lys⁵⁹ (Fig. 2.2). A reasonable hypothesis to explain the citrate-induced aggregation of pR is that it may be mediated by some of these 4 cationic residues on IC1. There was also a possibility of other lysines or arginines being involved in such an interaction, but the ones present in the membrane-spanning helices were ruled out easily. That is, direct interaction of citrate with two of the arginines (Arg⁸⁰ and Arg⁹⁴) and one lysine (Lys²³¹) were thought to be very unlikely, on the grounds that these residues are buried within the membrane protein, and Lys²³¹ is the site of Schiff base formation with the retinal chromophore. Additionally, our lab had previously made the R94C mutant [16] and shown it to be susceptible to citrate-induced precipitation.

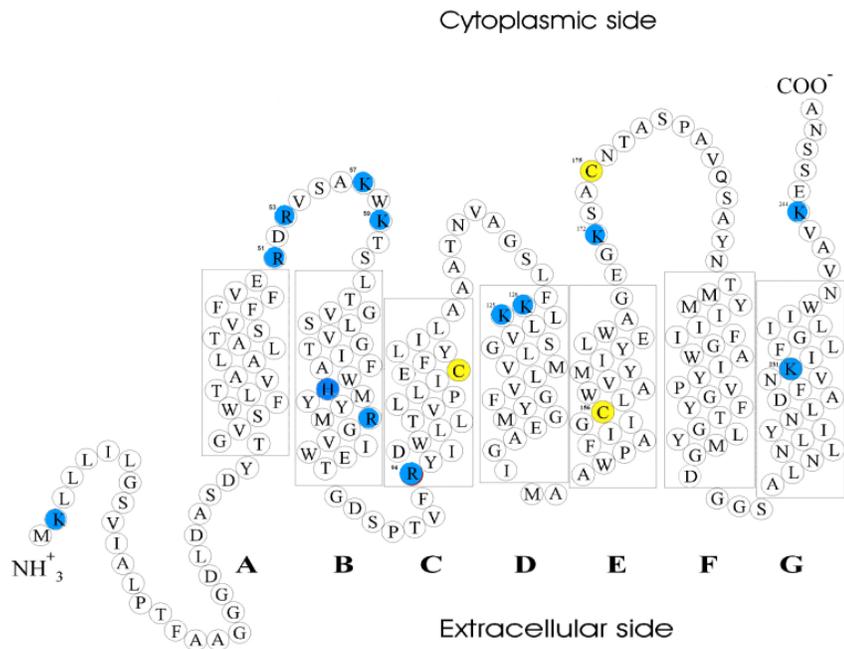


Fig. 2.2: Locations of cationic residues in pR: The native sequence of SAR86 pR (accession #AAG10475) is shown, with approximate division into helices A-G, resulting in 3 interhelical loops on each side of the membrane. The 3 cysteines (yellow) were all mutated to serine in Triple Cysteine Mutant proteorhodopsin (pR TCM), which served as a baseline for all of the further pR mutations studied in the current work. Positively charged residues (lysine, arginine, and histidine) are colored blue.

To test the role of the four amino acids in loop IC1 in binding to the citrate molecule, pR mutations were generated at these residues by using site-directed mutagenesis. Then these mutants were examined for the amount of pR remaining in the supernatant, after standardized additions of increasing amounts of citrate. The results are summarized in Fig. 2.3.

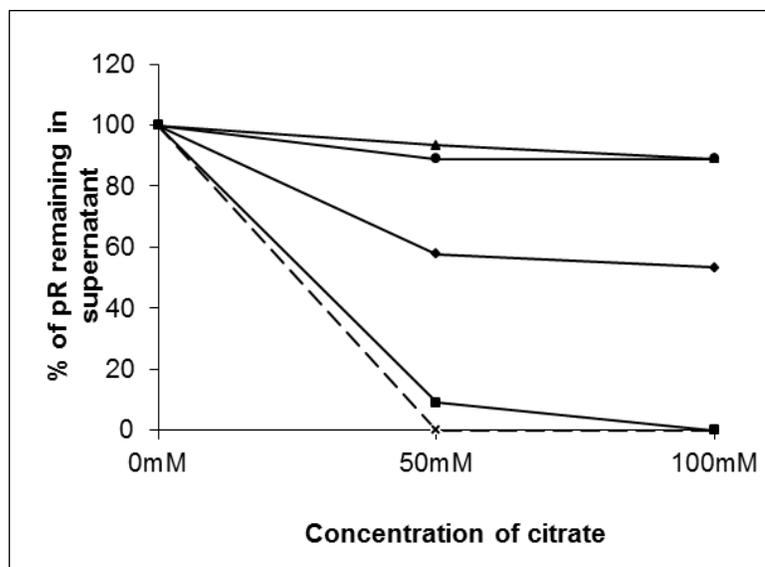


Fig. 2.3: Citrate precipitation of wild type pR and mutants: “Wild type” pR (TCM without His₆ tag) and mutants were purified by means of citrate precipitation and subsequent hydroxylapatite column. The pR content of an aliquot of the sample was determined by measuring the 520-nm absorbance. Each sample was then subjected to increasing concentrations of citrate, and with centrifugation at 4°C (6000 × g, 10 min) prior to measurement of the visible absorption spectrum. The portion of the initial pR remaining in the supernatant, as measured by 520-nm absorbance, is plotted as a function of the citrate concentration. Dashed line represents wild type pR; ■ mutants R51Q, K125Q/K126Q, K172Q and K244Q, which superimpose; ♦ triple mutant R53/K57/K59; ● double mutant K57Q/K59Q; and ▲ quadruple mutant R51Q/R53Q/K57Q/K59Q.

The initial mutation was carried out to change all four residues simultaneously, allowing to rapidly screen for any effect on citrate precipitation. Specifically, the quadruple mutant pR R51Q/R53Q/K57Q/K59Q (on a background pR sequence containing no His tag, but with the 3 native cysteines all mutated to serine) was tested for its reaction to citrate. It exhibited a remarkable difference compared to the control, unmutated protein (pR TCM). Unlike the control (pR TCM lacking His₆), which completely precipitated from 0.4% OG solution at 50 mM citrate, pH 5.5, the quadruple mutant did not precipitate even at citrate concentrations of 100 mM or higher (Fig. 2.3).

This result with quadruple mutant indicates strongly that one or more of these four amino acid residues in IC1 play a role in mediating the interaction between the pR and citrate.

In order to identify the amino acids most directly responsible for citrate interaction with pR, single mutants R51Q, R53Q, K57Q and K59Q, the double mutant K57Q/K59Q and the triple mutant of pR R53Q/K57Q/K59Q were made. These mutant proteins were purified as above, and then precipitated with varying concentrations of citrate buffer. With the triple mutant R53Q/K57Q/K59Q, even though a small fraction of it precipitated at 50 mM citrate, the pR never completely precipitated, even at higher (~100 mM) concentrations of citrate. The double mutant K57Q/K59Q also showed a similar effect as the triple and quadruple mutants. Only a minute fraction of the total pR precipitated with 50 mM citrate, 0.4% OG; and the precipitation did not increase with increasing concentrations of citrate. This suggests that the two lysines are most likely involved in the citrate interaction for the process of pR aggregation.

To assay the individual roles of Arg⁵¹ and Arg⁵³ in the pR binding site for citrate, the single-site arginine mutants, R51Q and R53Q, were tested for sensitivity to citrate-induced precipitation. The R51Q mutant did not show significant alteration with respect to its citrate sensitivity. That is, most of this mutant protein precipitated with about the same concentration of citrate as required for precipitating pR TCM. Thus, Arg⁵¹ is not required for citrate interaction. On the other hand, cells expressed with the single mutant R53Q did not exhibit a normal purplish color. This appears to be due to an inability of this single-site pR mutant to express and/or fold properly.

Since the double lysine mutant showed a dramatic decrease in the precipitation, it was interesting to know whether a single lysine residue mutation would also show a similar decrease and help pinpoint the citrate interaction to just one residue. It was hoped that a single residue might bring about a significant change in citrate sensitivity. The single lysine mutants, K57Q and K59Q were generated and the mutant proteins were expressed. However, these mutants were also not as strongly colored as the wild type and most other mutants, either in cell pellets or at any point during purification. That is, they appeared not to be as stably folded; and the purified proteins lost their faint brown-purple color completely when citrate was added.

The remaining lysines present on the cytoplasmic side of the pR, such as Lys¹²⁵/Lys¹²⁶ near the second loop, Lys¹⁷² on the third intracellular loop and Lys²⁴⁴ near the C-terminal end of pR, were also mutated to glutamine individually or in pairs, in order to examine their role in citrate-mediated pR aggregation. Lys¹²⁵/Lys¹²⁶, Lys¹⁷² and Lys²⁴⁴ pR mutants were expressed as colored proteins and precipitated out of the detergent containing solution at 100 mM citrate concentration, similar to the Arg⁵¹ mutant as well as the unmutated control (pR TCM). These results indicate that lysines 125, 126, 172, and 244 do not help mediate citrate-induced aggregation of pR.

2.4.3 Proteorhodopsin interacts with several anions

Interaction of pR with citrate appears to involve a specific region of the protein, namely the positively charged IC1 loop. In order to investigate the degree of specificity for citrate as the anion, several different anions were tested under similar conditions. As an initial investigation mono-, di-, and tri-carboxylic acids anions were tested in the range

of 10-100 mM, always buffered with 10 mM Tris-Cl at pH 7.5 and in the presence of 1% OG. The control experiment--adding 10 mM Tris-Cl, pH 7.5 while maintaining the detergent concentration at 1%--caused no precipitation. Table 2.1 shows the results of interaction of all anions tested with pR. The possible physiological significance of the values is represented as a physiologically indexed concentration required to initiate precipitation. A value near physiological index means that the concentration of that particular anion required to induce precipitation of pR is close to the physiological concentration of that anion inside healthy growing bacteria.

Acetate, a mono-carboxylate with a negative charge of 1 at pH 7.5, did not cause precipitation, even at the highest concentration measured, 200 mM (see Table 2.1). This was the case even at pH of 5.2, closer to its pK_a (data not shown.). Likewise, several dicarboxylic acids (succinate and dodecanedioate) also failed to precipitate pR under similar conditions, even at a concentration as high as 200 mM.

However glutarate, which has a similar spacing between its carboxylate groups as citrate, did cause pR precipitation, at a slightly higher concentration than citrate. Glutamate, which is just α -amino glutarate, showed a similar pattern of precipitation as glutarate. Interestingly, aspartate, which is α -amino succinate, was more similar to glutamate and glutarate than to succinate. However, both dicarboxylic amino acids (glutamate and aspartate) were unable to cause complete precipitation of pR; even at the most effective concentrations, they always left a substantial fraction of the pR in the supernatant. These results suggest that the spacing between a pair of carboxylate groups

is the major, but perhaps not the only, structural determinant of anion-induced precipitation of pR

Given that several divalent anions precipitated pR, albeit not as well as citrate itself, it was important to examine whether other simple trivalent anions, like isocitrate and aconitate, might interact with the citrate binding site, and thereby cause pR precipitation. In addition to citrate itself, two other tricarboxylic acids showed evidence of binding. That is, precipitation of pR was observed for *cis*-aconitate and isocitrate at 50 mM (but not at 10 mM, and notably not at the higher concentration of 100 mM).

Phosphates are multiply-charged anions pervasive in biological systems. Some evidence of a possible interaction between pR and phosphate was already seen while trying to precipitate pR eluted from a hydroxylapatite column. Therefore, pR samples were treated with increasing concentrations of phosphate, buffered at pH 7.5 and at constant 1% OG concentration. It was found that pR precipitation was not observable until phosphate reached 100 mM phosphate. In order to determine if phosphate was interacting at the same site on pR as the citrate, partially purified quadruple (R51Q/R53Q/K57Q/K59Q) and double (K57Q/K59Q) mutants were also tested with increasing concentrations of phosphate at pH 7.5. It was found that, as with citrate, these IC1 mutants failed to precipitate. This remained true at even higher (200 mM) concentrations of phosphate, strongly suggesting that phosphates may bind to the same site on pR as the citrate.

Larger phosphate-containing molecules were also tested that are found intracellularly, such as glucose-6-phosphate (G6P), ADP, and ATP. G6P has a single phosphate group, ADP has two, and ATP has three. G6P showed precipitation of pR at 50 mM, but not at concentrations above 100 mM. ADP showed precipitation up to 40 mM while ATP exhibited no precipitation at any of the concentrations tested (20-150 mM). Additionally, concentrations of ADP or ATP above 50 mM led to denaturation of the pR with white precipitation. That is, while no pR remained in the supernatant, the precipitate showed a yellow color instead of the characteristic red-purple. It would be interesting to investigate any further details of this intriguing behavior.

Inorganic anions such as sulfate, chloride, and chlorite were investigated for their ability to cause pR to aggregate. Sulfate (sodium or ammonium), sodium thiosulfate and sodium chlorite did not cause any precipitation even at 100 mM. Chloride showed initiation of pR precipitation at 40-50 mM, depending on whether the counter-ion used was potassium or sodium. These results suggest that chloride also may have a specific effect on the IC1 anion-binding site, even though it is monovalent.

Table 2.1: Precipitation of pR with diverse anions: Anions were buffered to pH 7.5, also contained 10 mM Tris-Cl at pH 7.5, and a sodium counterion. The concentration of OG was maintained at a constant 1% for all tests.

Anion tested	Min. concentration, mM (pR precipitation)	Physiologically-indexed concentration	Physiological concentration (mM) ^{14, 15, 25, 30}
Citrate	++ (10)	0.76, 5	13 ¹⁴ , 2 ¹⁵
Cis-aconitate	+ (50)	3125	0.016 ^{14, 15}
Isocitrate	+ (50)	2.5	20 ¹⁴
Succinate	-	N/A	0.57 ¹⁵
Dodecanedioate	-	N/A	N/A
Acetate	-	N/A	N/A
Aspartate	+ (50)	12	4.2 ^{14, 15}
Glutamate	+ (50)	0.52	96 ^{14, 15}
Glutarate	+ (50)	N/A	N/A
Sulfate	-	N/A	N/A
Thiosulfate	-	N/A	N/A
Phosphate	++ (100)	13.3	7.5 ³⁰
ATP	-*	4.2	9.6 ^{14, 15}
ADP	+ (20)*	71.4	0.56 ^{14, 15}
Glucose-6-phosphate	++ (50)	62.5, 41.6	0.8 ¹⁴ , 1.2 ¹⁵
Chloride	++ (40)	0.07	535 ²⁵
Chlorite	-	N/A	N/A

- ++ indicates complete precipitation of pR, + indicates partial precipitation leaving significant color in the supernatant, - indicates lack of precipitation with all the color in the supernatant.
- The references for concentrations for physiologically important anions available are labeled as superscripts. Concentration of phosphate is from *Streptococcus bovis* [30] and that of chloride is from marine bacteria *Pseudomonad* [25]. All other anion concentrations are from *E. coli* [14, 15].
- N/A refers to data Not Available
- * indicates anomalous behavior. ATP showed white precipitate with faintly colored supernatant up to 40 mM. It started to show degradation of pR beginning at 50 mM with yellow precipitate and a colorless supernatant. ADP showed similar behavior starting at 60 mM

2.4.4 Model for pR-citrate interaction

A homology model of pR was created using Geno3D molecular modeling web server [21] based on the crystal structure of bR (PDB: 1fbb). Based on the results of the double lysine mutants, it can be speculated that two of the negatively charged carboxylate groups of citrate bind to the positively charged ϵ -amino groups of the two lysines, Lys⁵⁷ and Lys⁵⁹ through H-bonded salt bridges. Using the homology model of bR, a structural model for pR-citrate interaction was built by using the UCSF Chimera package [22] (Fig. 2.4). It can be seen that Lys⁵⁷ and Lys⁵⁹ are likely close enough for such an interaction to occur. The distance between the farthest carboxyl groups of citrate is about 6 Å, while the distance between the ϵ -amino groups of the two lysines, Lys⁵⁷ and Lys⁵⁹, is approximately 8 Å in the homology model. At physiological pH, citrate is completely dissociated and has a charge of -3, while the lysine side chains each carry a +1 charge. Assuming an H-bonding distance between the proton donor nitrogen of lysine and proton acceptor oxygen of citrate to be 2.5-3Å, the two lysines can easily form salt bridges to a single citrate molecule.

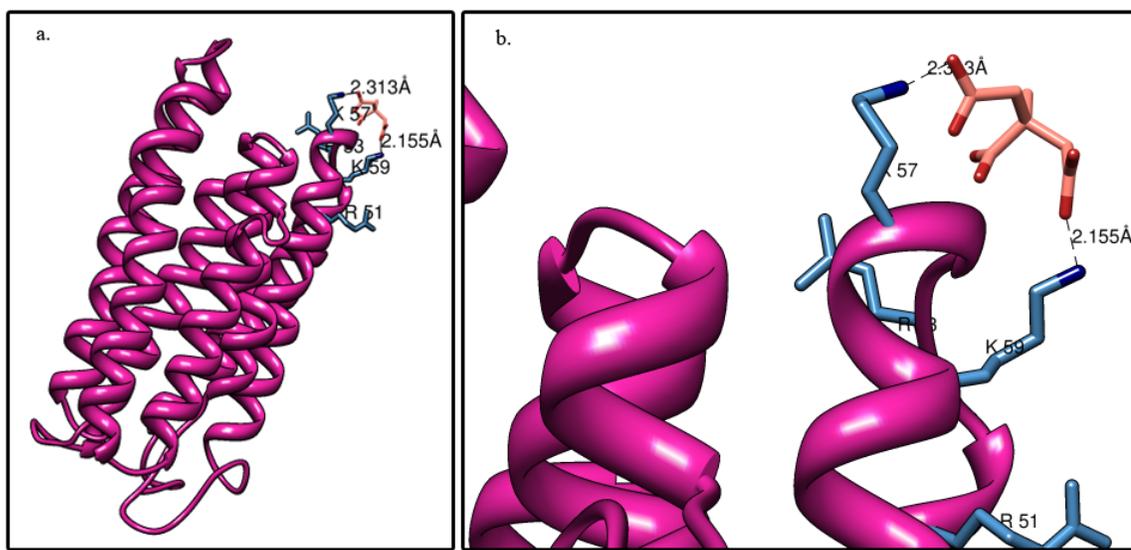


Fig. 2.4: Model of pR-citrate interaction: **a.** The homology model of pR, based on the crystal structure of native conformation of bR from *Halobacterium salinarum* (PDB: 1FBB), was produced using Geno3D. Molecular graphics images of pR-citrate interactions were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco. Arg⁵¹, Arg⁵³, Lys⁵⁷, and Lys⁵⁹ side chains are depicted in blue. The lysine side chain bond torsion angles were manually adjusted to permit reasonable H-bonding interactions with two of the carboxylates of a citrate molecule, shown in pink and red. **b.** Close-up image of the citrate-lysine interaction.

2.4.5 Citrate precipitates pR at physiological pH

Proteorhodopsin shows precipitation at pH 5.5 from a 3% detergent solubilized extract or Tris-Cl buffered solution also containing the detergent. To assess if pR precipitates at a pH higher than 5.5, especially near the physiological pH, a detergent extract of pR containing cells was made and subjected to several rounds of citrate precipitation. The precipitated pellet was resolubilized in 3% OG buffered at pH 9. The

detergent extract and the resolubilized pellet were then treated with equal volume of 100 mM citrate buffered at pH 5.5, 6.0, 7.0, and 7.5. The protein completely precipitated out of the solution at all the pH values indicating that pR precipitation likely occurs throughout the physiological pH range.

Another aim was to find out the exact concentration of citrate that can cause pR precipitation out of the detergent solubilized solution at the likely physiological pH (7.5). In order to avoid possible effects of dilution of the OG detergent, the same 1% (34 mM) concentration of OG was used in all the precipitant solutions as in the starting pR sample. This concentration is just above the critical micellar concentration (CMC) of OG (23-25 mM). The pR in detergent solution was subjected to an increasing concentration of citrate, buffered at pH 7.5 (with Tris), while maintaining 1% OG. It was found that pR precipitates within minutes at a citrate concentration of 10 mM, and indeed nearly quantitatively at all 5 concentrations measured between 10 and 200 mM. It was also observed that upon incubation at 4°C for several h, even 5 mM citrate proved sufficient to cause pR precipitation.

Another significant question regarding the role of pH is whether the lysine and arginine mutations of IC1 would react similarly to citrate precipitation at pH 7.5 as they do at pH 5.5. Partially purified samples of quadruple and double lysine mutants were tested with an increasing citrate concentration at pH 7.5 at 1% OG concentration. It was found that both the quadruple and the double mutants failed to precipitate, even at citrate

concentrations as high as 100 mM. This supports the hypothesis that Arg⁵¹, Arg⁵³, Lys⁵⁷ and Lys⁵⁹ are part of a citrate-binding site in pR.

2.4.6 Effect of citrate on pR physiological activity

Based on site-directed mutagenesis, the citrate binding site on pR appears to be located on the intracellular side of the membrane. This site appears to interact with citrate at a concentration close to that of the endogenous citrate concentration inside bacteria [14, 15]. These findings led to the consideration of the possibility that citrate could potentially have a physiological significance in the regulation of pR photocycle activity.

To assess this possibility, pR was purified (by using the citrate purification method described), and reconstituted in citrate-free DOPC vesicles. The pR-containing liposomes were then subjected to flash photolysis. An observed transient 410-nm absorbance signaled the formation of the M intermediate. A representative plot of transient absorbance against time is shown in Fig. 2.5A. A curve fit to the single exponential is also shown along with the residual plot. A log plot of absorbance vs. time (Fig. 2.5B) yielded a straight line, consistent with a first-order exponential decay. The exact value of the 410-nm absorbance decay constant was calculated from a nonlinear least-squares curve fit as 45 ms^{-1} , which corresponds to a half-life of 22 ms. This is very similar to that reported previously [26].

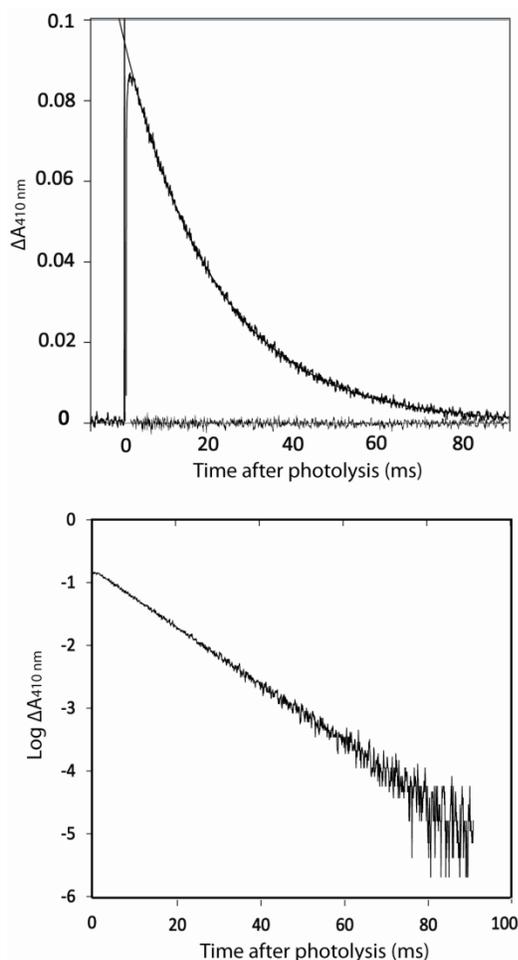


Fig. 2.5: Flash photolysis: Flash-induced transient absorbance changes in pR-containing DOPC liposomes (buffered at pH 8.5) show the decay of the M intermediate. These data were obtained from a sample containing no citrate. In part **A**, the measured absorbance data were fit to a first-order kinetic equation, $A_{400}(t) = B + A_0 e^{-kt}$, by using the solver function in Excel, optimizing the values of B , A_0 and k . The fitted exponential curve is superimposed on the measured transient absorbance at 410 nm. The residual trace (lightly-dotted line) demonstrates the high quality of the fit. Part **B** shows a log plot of absorbance against time. This plot yielded a straight line signifying a first-order exponential decay of the M intermediate. The decay constant, k , calculated from the slope of this curve matches that obtained from the non-linear least-squares optimization. (In **B**, and for the exponential fitting procedure in **A**, all digitized points up to the commencement of the M decay phase, corresponding to a total of ~8 ms of time, were omitted in order to allow the approximation of a single-exponential process.)

These results by themselves indicate that the use of citrate-induced aggregation as a step during pR purification does not significantly alter the photocycle, i.e. the resultant pR retains its native functionality. As a further control, we added increasing amounts of citrate to the external solution surrounding the pR-containing liposomes and subjected them to flash photolysis. No effect of citrate addition on the kinetics (Fig. 2.6) or amplitude (not shown) of the 410-nm (M) absorbance decay was observed. This was as expected, since pR has been found to insert right-side-out in vesicles [27, 28], and citrate added after vesicle formation would therefore not be expected to be able to reach the intravesicular binding site on IC1.

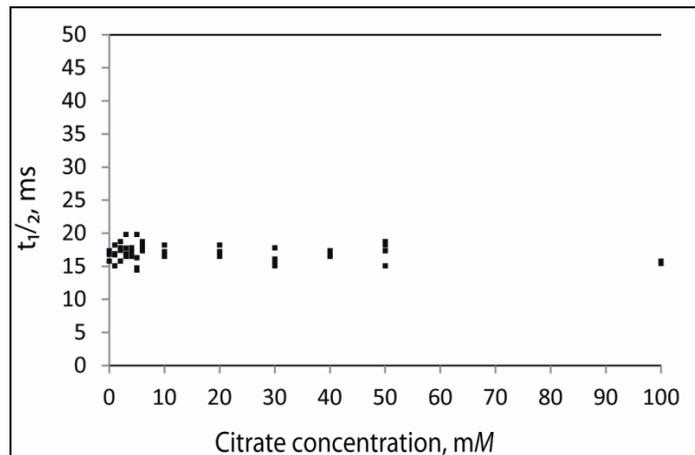


Fig. 2.6: Dependence of the fitted half-life for M (410-nm absorbance) decay on external citrate concentration: The 410-nm absorbance transient for each citrate concentration was fitted as in Fig. 2.5.

Finally, to test the effect of intra-vesicular citrate, the experiment was repeated but with citrate buffer included during the vesicle-forming dialysis step. If intracellular

citrate binding had a physiological role in regulating pR activity, then intravesicular citrate might be expected to affect either the amplitude or kinetics of the M intermediate. However, this was not observed (data not shown). That is, citrate inclusion in the reconstituted liposomes did not significantly alter the M-intermediate amplitude or its decay rate. Overall, the flash spectroscopy experiments with pR liposomes provide no evidence for a physiological role for citrate.

2.4.7 Conservation of the citrate binding site

Based on the protein sequence alignment of pR to the members of the microbial rhodopsin superfamily and the resulting distance tree of proteins (Fig. 2.7), it is evident that the pR representatives belonging to the proteobacteria group share conserved amino acid residues at the citrate binding site we have identified on loop IC1. That is, while there is no evidence for conservation of these residues in the archaeal branch of this tree (corresponding roughly to the top half of this figure), both the lysines, K57 and K59, are highly conserved among the members of bacteria belonging to the γ -proteobacteria. In less closely-related α -proteobacteria, the conservation is not as strong, e.g. *Candidatus pelagibacter* lacks both the lysines, and possesses just one arginine instead.

At least one of the two lysines is conserved in other members of phylum Bacteroidetes, such as Flavobacteria, Cellulophaga and Polaribacter. The other lysine is generally substituted by arginine. This is the most conservative possible substitution, since arginine conserves both the positive charge of lysine and its ability to donate

protons to an H-bond. Thereby, pR in Bacteroidetes would be predicted also to interact with citrate. This large phylum Bacteroidetes, previously known as the CFB (Cytophaga-Flexibacter-Bacteroides) group of bacteria, diverged from a common ancestor about 2.5 Ga [29], and from the γ -proteobacteria even more distantly in the past. The significant degree of conservation of the cluster of two charged cationic residues in loop IC1 among such distantly-related bacteria is worth noting.

Fig. 2.7: Conservation of citrate-binding lysine pair in pR: pR and bR sequences from multiple organisms were aligned with BLASTp algorithm, and a distance tree was produced using fast minimum evolution tree method. The following color code is used to indicate the degree of conservation of the two lysine residues in loop IC1, associated with the putative citrate binding site of the proteins. ■ K.K ■ K.R ■ ..R ■ R.R ■ ..K.



2.5 DISCUSSION

2.5.1 Purification of pR TCM without His-Tag

It was shown previously that selective precipitation of pR with citrate works even better and more reproducibly when the 3 native cysteines are mutated to serine [16, 17]. Furthermore, the resulting protein is somewhat more stable. The reasons for this are still unclear, but TCM was used as the background “wild type” for the studies here of the citrate-induced aggregation phenomenon.

Nearly all previous studies on purified pR were carried out on protein expressed with six non-native histidine residues added to its C-terminus. This method of purification of pR using citrate as a selective precipitant is equally effective with pR in its native form, i.e., without the His₆-tag. This finding is significant because it has the potential of being applied for purification of pR from natural source organisms, which of course lack the His-tagged version of the protein. Additionally, this method also eliminates the use of expensive Ni²⁺-derivatized resin, instead using a relatively inexpensive hydroxylapatite column. Besides, it has been observed that the non-native His₆-tag is unstable over long periods of storage at 4°C (our unpublished data). It also shows that pR can be purified to A_{280/540} ratio of less than 3 (>50% purity) with only citrate and hydroxylapatite column. The pR purified by using citrate aggregation steps shows normal M decay kinetics, and is therefore likely to represent the same native structure as for pR prepared by other methods.

2.5.2 pR–citrate interaction and aggregation

Citrate-induced aggregation of pR occurs at concentrations around the endogenous free citrate concentrations that range between 2 mM and 13 mM, that have been measured inside healthy bacterial cells [14, 15]. The apparent citrate concentration required for aggregation of pR TCM from crude cell extract is ~50 mM, but the required free citrate concentration is considerably less than this, because a significant portion is likely consumed by other proteins present in the crude extract that bind citrate, either specifically or randomly. Thus, with purified pR less citrate is needed to aggregate the protein, indeed approaching the physiological levels of citrate inside the cell. Specifically, it has been observed that nearly-quantitative precipitation of pure pR in 0.4% OG can be achieved with as low 10 mM citrate at pH 5.5 (data not shown). The experiments at pH 7.5 and in the presence of 1% OG also showed that precipitation of pR can be achieved rapidly at 10 mM citrate. Prolonged incubation of pR under above mentioned conditions has also shown a precipitation at 5 mM citrate. The intracellular concentration of citrate inside a healthy growing bacteria ranges from 2 mM to 13 mM. An intracellular concentration of 2 mM was obtained from cells growing on filters which made them grow slightly slower than those in liquid media [15]. This could partially explain why Bennett B.D. *et al.*, [15] obtained a lower concentration of citrate compared to the previous value of 13 mM by another group [14]. It should also be noted that native pR is found in γ -proteobacteria, and in other bacteria from oceanic or fresh water environment. The difference in the environment as well as the genotype might also probably affect the metabolism and metabolite concentrations in cells yielding different numbers than those estimated in the *E. coli* cells under standard lab conditions.

In this study, two lysine residues of pR have been specifically identified that help mediate its interaction with citrate. The two lysine residues appear to function together to form a binding site for citrate. However, the binding that was observed was a highly cooperative process, rather than simple equilibrium involving a 1:1 stoichiometry. Therefore it was not easy to define the binding reaction so as to permit fitting to a simple binding isotherm.

The quadruple mutant of R51Q/R53Q/K57Q/K59Q shows the clearest reduction in citrate-induced aggregation, followed by the triple mutant R53Q/K57Q/K59Q. The R51Q single mutation does not show any effect on citrate-induced precipitation. But when only the two lysines, Lys⁵⁷ and Lys⁵⁹, are mutated, then the protein behaves like the quadruple mutant, i.e., citrate does not cause aggregation. This strongly suggests that the two lysines are critical for the citrate mediated aggregation, and thus form the essential locus of initial citrate binding to pR.

Interestingly, single mutations at either of the 2 lysines in IC1 (K57Q or K59Q) or at one of the arginines (R53Q) did not express cells with their normal strong pink color, but at most a much more subdued brownish-orange. Furthermore, it was much more difficult, or impossible, to purify pR from any of these 3 single-site mutants than for the triple- or quadruple-mutants, and the resulting purified pR proteins lost all pinkish color to give an absorption conforming to free retinal, with $\lambda_{\max}=380$ nm, upon addition of as little as ~30 mM citrate. This suggests that pR is unable to fold stably when certain of these mutations are present individually. It is not clear why these mutations, especially

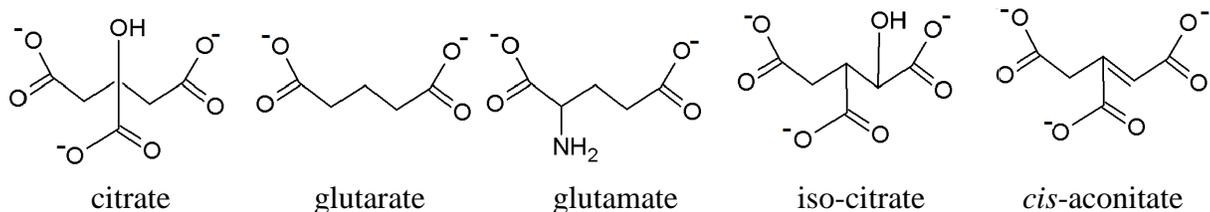
that of Arg⁵³, should cause such protein instability only when present individually, and not in combination. However, as a result of such protein instability, the degree of involvement of R53 in citrate-mediated pR-aggregation cannot be clearly defined.

Phylogenetics comparisons show that these 4 cationic residues in IC1 are quite strongly conserved, even among pR sequences from organisms that likely diverged nearly 2.5 billion years ago. Such a degree of conservation is usually a sign of physiological function. It seems unlikely that these residues are a requirement merely for structural stability, since they are not conserved in the somewhat more-distantly-related retinal-binding proteins found in archaea (bacteriorhodopsin, halorhodopsin, and sensory rhodopsin), nor in their fungal homologs. Our site-directed mutagenesis experiments and the phylogenetic comparisons together hint at a possible physiological role for citrate binding at IC1 in pR.

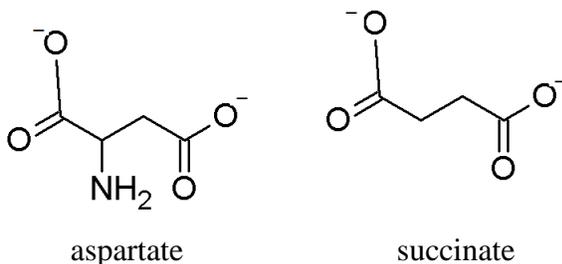
2.5.3 pR interaction with diverse anions

Citrate induces pR aggregation at a pH ranging from 5.5 to 7.5, in the presence of 1% OG. Since citrate causes pR aggregation at a physiological pH 7.5, other organic and inorganic anions were tested at the same pH, and in the presence of detergent up to a concentration of 1% OG. Monocarboxylates such as acetate, and several dicarboxylates such as dodacanedioate and succinate, could not induce pR precipitation. Other dicarboxylates such as glutarate, aspartate, and glutamate, and tricarboxylates (iso-citrate, *cis*-aconitate) appear to induce considerable precipitation of pR at particular concentrations. The concentration of glutamate and isocitrate required to induce

between these two tricarboxylates is in the placement of the OH group. This suggests that pR-citrate interaction is very specific. The spacing between the carboxylates might be only a part of the steric requisite to produce precipitation of pR.



The results that most contradicted a simple model of carboxylate-distance as determinant of binding were those from aspartate, a dicarboxylate that resembles succinic acid except for the additional presence of an amino group. Succinate lacks the ability to induce pR precipitation, whereas aspartate precipitated pR, albeit only partially.



Nevertheless glutamate does interact considerably strongly with pR than aspartate, although not as strongly as citrate. The concentration required to induce pR precipitation by glutamate is only half of its physiological concentration. In *E. coli* cultures fed on glucose, acetate or glycerol, glutamate was found to be the most abundant intracellular metabolite, constituting over 40% of the total metabolome [15]. In comparisons of glucose-fed cultures, glutamate was followed by glutathione, fructose-1,6-bisphosphate (FBP) and ATP. The abundance of glutamate is justified on two

grounds. First, it is the major nitrogen donor in the cell, involved in ~88% distribution of nitrogen that is donated via transamination reactions [31]. Second, it acts as the major intracellular counterion to potassium [32]. It can be speculated that a high concentration of glutamate could potentially aid in the regulation of pR activity intracellularly. In the nutrient-limited conditions, glutamate could potentially support the bacteria up to a certain period of time while also keeping the pR inactive. Only under conditions where glutamate concentration falls below the levels that signals starving conditions, then the pR pumping activity takes control of the energy provision.

Other anions such as phosphate and chloride completely precipitated pR, at higher concentrations compared to citrate, albeit within their physiological ranges. It appears that pR-anion interaction is quite specific, in the sense that only a few anions can cause the pR to aggregate and precipitate.

The aggregation phenomenon observed is unlikely to be merely a salting-out of the protein. Salts typically used for salting-out (e.g. 10-100 mM ammonium sulfate) do not cause aggregation of pR, whereas a mere 10 mM concentration of citrate suffices, measured under a specific set of similar conditions, namely at pH 7.5 and in 1% OG. It should be noted, however, that at lower pH and either higher or lower detergent concentrations, there is less distinction in the behavior of pR with sodium citrate and other salts. Salting-out certainly appears likely to be occurring under such conditions. It is also possible that the precipitation of pR by higher concentrations of other salts, e.g. 40-

50 mM chloride or 100 mM phosphate at pH 7.5 and 1% OG, could be a salting-out phenomenon.

Furthermore, NaCl caused pR aggregation at a lesser concentration than KCl, indicating some possibility of cation-specific effect/interaction. However, this trend too is opposite of what is expected from the Hofmeister series, if these salts are merely causing salting-out. This might suggest a cation-dependent protein interaction. Likewise, potassium phosphate precipitated pR at a lesser concentration than sodium phosphate. However, there is not a monotonic dependence of the pR precipitation phenomenon on KCl concentration: KCl precipitates pR at 50 mM but not at 100 mM. However, neither potassium nor sodium acetate causes any precipitation, so the cation appears to have much smaller effect than the anion.

Hofmeister [33, 34] showed that different ions show a consistent trend in effecting precipitation of wide variety of proteins. For the same cation with different anions, the order is $\text{SO}_4^{2-} > \text{HPO}_4^{2-} > \text{acetate} > \text{citrate} > \text{tartarate} > \text{bicarbonate} > \text{chromate} > \text{Cl}^- > \text{NO}_3^- > \text{ClO}_4^-$. For the same anion, the order for cations is $\text{Li}^+ > \text{Na}^+ > \text{K}^+ > \text{NH}_4^+ > \text{Mg}^{2+}$. Additional salts have been studied by subsequent workers. It has also been shown that the order of both, the anions and cations, is reversed depending upon the pH. Proteins show a “direct order” i.e. the order as shown by Hofmeister, at a pH above their isoelectric point (pI) whereas the order is reversed at a pH below the pI [35, 36]. This reversal is theoretically explained by a combination of Coulombic screening effect and ion-surface interactions [37]. The order of cations may reverse with different cations,

due to ion pairing [38, 39]. The order of anions may also reverse depending on the hydrophobicity and hydrophilicity of the surface of the peptide. Ions that are repelled and depleted from the hydrophobic surface (such as F^-) raise the interfacial tension strongly and enhance the precipitation of a protein. Such a repulsion is directly proportional to the ionic-surface charge density [40, 41]. For a surface that possesses both hydrophobic and hydrophilic patches or groups, a smaller anion, Cl^- binds strongly to the cationic surface compared to a larger anion, I^- [42].

Proteorhodopsin precipitation with diverse anions has been carried out at a pH higher than the pI of pR, estimated to be 4.6 (Braiman M. S., unpublished data) indicating that the anions should have exhibited a direct order of Hofmeister series. The reversal of trend suggests something more than Hofmeister effect as a cause of citrate-induced precipitation.

Of the anions beside chloride and polycarboxylates, the one that appears to interact most strongly with pR is phosphate, which causes precipitation at concentrations somewhat higher than citrate, i.e. a few times greater than the physiological range. The physiological concentration of phosphate varies between a wider range depending on the metabolic status of the cell. Although the precipitation of pR by phosphate occurs at a concentration higher than that of citrate, it is still within the physiological range of the phosphate. Intracellular concentration of phosphate is affected by the metabolic state of the cell; when the glucose consumption is high, the inorganic phosphate concentration decreases from 45.5 mM to 7.5 mM [30]. At higher concentration of phosphate, i.e., 45.5

mM, the cells are probably at metabolic rest or at a level of lower metabolic activity. That is possibly the concentration at which phosphate shows its inhibitory effect on pR by precipitating it. It should be noted that these values of inorganic phosphate concentrations are derived from *Streptococcus bovis* as those of *E. coli* are not available. It is quite likely that they might be different from *E. coli* and, from oceanic bacteria which harbor pR.

The two mutants that most clearly interfere with citrate binding (K57Q/K59Q and R51Q/R53Q/ K57Q/K59Q) also interfered with phosphate-induced precipitation. This indicates that phosphate and citrate likely interact with pR through the same site. Interestingly, phosphate at a higher concentration (>400 mM) not only fails to precipitate wild type pR, but can also suppress its interaction with citrate, even when the detergent concentration is lower (~0.4%), a condition that usually promotes the citrate-induced aggregation. This was observed previously when pR eluted from a hydroxylapatite column with phosphate gradient failed to undergo citrate-induced precipitation, and required dialysis to remove the phosphate prior to use of the citrate. Competition of phosphate and citrate for the same site could help to explain why citrate loses the ability to precipitate pR when the phosphate concentration is high.

Phosphate esters and anhydrides that are found intracellularly gave surprising results. G6P showed no precipitating effect at higher concentrations, whereas ADP and ATP led to pR denaturation.

Anion binding has been studied in bacteriorhodopsin (bR) [43], halorhodopsin (hR) [44] and mammalian rhodopsin [45]. A recent study on anion binding and its effect on the pK_a of proteorhodopsin [46] included sulfate, phosphate and chloride as well as other anions. Based on the effect on pK_a of pR, it was proposed that chloride and sulfate have a binding site on pR. However, these workers [46] did not report any precipitation effect on pR by any of these salts, even at a 300 mM anion concentration. This likely arises from the use of dodecyl maltoside for their measurements on pR, at 5-6 times its critical micelle concentration (CMC). Citrate-induced precipitation of pR solubilized in such concentrations of dodecyl maltoside, or even with OG concentrations at 5 times its CMC fails to cause pR precipitation. Our lab's initial experiments with pR and citrate were standardized with OG near its CMC, and therefore the studies were carried out with these conditions.

However, there is clearly a dependence of the aggregation phenomenon on the detergent as well as the anion. At even lower concentrations of OG, e.g. 0.5%, pR precipitation from OG solutions has been observed by using tricarboxylates (iso-citrate, *cis*-aconitate) as well as numerous other salts such as sodium sulfate, sodium chlorite, potassium chloride, sodium thiosulfate, G6P, ATP and ADP. When OG concentration was maintained at 0.5%, most of the salts and carboxylates showed at least a scanty precipitation of pR, even at quite low anion concentration. This interplay between salt and detergent concentration is somewhat surprising. Increasing salt concentration normally has the effect of increasing micellar size and reducing CMC of non-ionic detergents such as OG. Thus addition of salt to pR in 0.5-1% OG might be expected to

lead to formation of more stable mixed micelles containing pR and detergent, rather than leading to aggregation.

Citrate has already been employed in the process of purification of pR and the possibility of using some of these other salts as alternative purifying agents is being explored. The salts that are most attractive are NaCl and phosphate. They are inexpensive and easily available in abundance and based on our preliminary experiments, do not show any effect on the characteristic absorption of pR at 520 nm.

2.5.4 Citrate has no discernible effect on pR reconstituted in a membrane

Normal kinetics following pulsed-laser photolysis of the reconstituted pR vesicles prepared in the presence of citrate, suggest that citrate fails to exert an aggregation effect on proteorhodopsin in DOPC liposome environment, just as it apparently does not have an effect when non-ionic detergents are used well above their CMC. It may be that the geometry adopted by the lipid and/or detergent molecules interferes with the contact of the some of the pR surfaces involved in aggregation.

The absence of a measurable effect of citrate on the pR photocycle in reconstituted phospholipid vesicles makes it very unlikely that citrate binding is involved in any simple binary regulation scheme. However, these experiments in purified reconstituted systems from heterologously-expressed pR may have excluded some necessary component, e.g. another bacterial protein found in the oceanic γ -proteobacterium, that is involved in a ternary interaction involving also pR and citrate.

Thus, the possibility cannot be excluded, that the binding and self-aggregation of pR in detergent micelles in the presence of citrate (and phosphate) is a non-physiological effect that is nevertheless reflective of anion binding involved in a more complicated physiological regulatory process.

2.6 CONCLUSIONS

Citrate in a physiological concentration range can cause aggregation and precipitation of proteorhodopsin from detergent solution, without denaturation or loss of function. In this study, the structural basis for the interaction between pR and citrate was investigated. Site-directed mutants of a number of positively charged amino acids in pR shows that those found in the first intracellular loop (IC1), R51/R53/K57/K59, are clearly involved in binding to citrate, with the strongest effect from the lysines. Phylogenetic comparisons indicate that these four amino acid residues are strongly conserved among pR sequences belonging to organisms that diverged more than 2.5 billion years ago. In the physiological pH range of oceanic bacteria, the concentration of citrate required for selective pR precipitation from OG detergent solution, 10 mM, is also in the intracellular physiological range. However, flash spectroscopy experiments with pR liposomes showed no effect of citrate on photocycle yield or kinetic, and thus fail to support the hypothesis of a physiological role for citrate in pR regulation.

Besides citrate, pR also interacts with other anions such as phosphate, glutamate and chloride, all of which can also lead to selective precipitation under certain conditions. However, at high concentrations phosphate not only fails to cause precipitation, but also

blocks citrate-induced precipitation, suggesting that both may bind at the same site in IC1. Thus, in addition to citrate, manipulation of phosphate and chloride concentrations is potentially useful in the purification of pR by selective precipitation.

2.7 Acknowledgements

This work was supported by Syracuse University. I express thanks to Prof. Joseph Chaiken for lending me his oscilloscope that enabled me to visualize flash spectroscopy measurements. I thank Mr. Michael Brandt for his help with Flash spectroscopy instrumental expertise.

Chapter 3

COLUMN-FREE PREPARATION OF PROTEORHODOPSIN, AN INTEGRAL MEMBRANE PROTEIN, BY SELECTIVE PRECIPITATION WITH SIMPLE SALTS

3.1 INTRODUCTION

3.1.1 Membrane proteins—heterologous expression and purification

Membrane proteins are involved in essential roles in cellular functions, such as cell signaling, cell adhesion, transport of molecules, and maintenance of cell or organelle structure. About 30% of the mammalian genome open reading frames (ORFs) have been found to encode integral membrane proteins [1].

Despite their key functions, only 290 membrane proteins have had their three-dimensional (3D) structure elucidated to date [2] compared to the available 3D structures of >42000 soluble proteins. One major difficulty lies in the inability to achieve successful expression and purification of membrane proteins. These goals are generally more challenging than for soluble proteins.

Selective precipitation of proteins with various salts has been traditionally used as one of the earliest steps in the process of purification of soluble proteins. The process of

salting out was originally proposed and studied by Hofmeister [3, 4]. Ions are thought to cause their effect indirectly by influencing the structure of water [5]. Therefore, the ions are divided as kosmotropes that promote hydrogen-bonded water structure leading to salting-out of proteins, and chaotropes that break the water structure consequently causing salting-in of proteins. Solubility of proteins also depends on the salt concentration in the solution. At lower concentrations, the salts stabilize the charged groups of the proteins thus assisting in protein solubility or salting-in. However, at higher concentrations, they bind to the water molecules apparently causing a decrease in the water activity that leads to protein precipitation or salting-out of proteins.

Hofmeister established a series of anions and cations according to their efficacy in inducing protein precipitation from aqueous solutions [3, 4]. Despite continued differences in opinion about the mechanism of salting out, it is nevertheless a valuable method for protein biochemists. Ammonium sulfate is a prime example of a salt used for salting out, mostly for soluble proteins. Other commonly used salts include sodium chloride, sulfate and phosphate [6].

The effect of Hofmeister ions on membrane proteins has only recently gained some attention. Studies performed on membrane proteins such as rhodopsin [7] and bacteriorhodopsin [8] showed that membrane proteins tend to show deviations from the Hofmeister series. It was suggested that these deviations could be due to the presence of a membrane environment that prevents the ions from exerting their influence on local water structure as strongly as in bulk aqueous phase. Salt-induced precipitation of proteins

requires that they be capable of existing in a soluble form. For membrane proteins, a suitable detergent is required to keep the membrane proteins stable as well as soluble, in particular to prevent them from aggregating prior to addition of a salt.

There are general guidelines that can be followed when choosing a detergent to solubilize membrane proteins during their purification. However, these cannot be used strictly [9]. The general guidelines use a classification of detergents as mild or harsh. Mild detergents have a good chance of maintaining a physiologically active protein structure. Harsh detergents exhibit a high probability of denaturing membrane proteins. Two structural factors that are most important in determining mildness or harshness of a detergent are the head group and the hydrophobic tail. Harshness increases as the polar head group is altered from large to small and from polar to zwitterionic to charged. It also increases as the hydrophobic tail decreases in length. However, there are additional factors to consider before selecting a particular detergent. There is also often a specific “chemical effect” between the detergent and the protein; for example, thermostabilized β_1 -adrenergic receptor is very stable in a relatively harsh detergent decylmaltoside (DM) [10] but is extremely unstable in a mild polyoxyethylene detergent like C12E8. In contrast, another protein, the sarcoplasmic reticulum Ca^{2+} -ATPase was found to be especially stable in C12E8 and indeed has been successfully crystallized in C12E8 [11].

At the same time as it maintains a significant degree of solubility by mimicking a membrane environment, the detergent ideally should not interfere with the protein crystallization process. However, a detergent that works for crystallizing one membrane

protein does not necessarily perform equally well with another protein. For example, the detergent, nonyl- β -glucoside, that was used for rhodopsin crystallization [12, 13] is not suitable for the crystal formation of another membrane protein, β 2 adrenoceptor (β 2AR). Similarly, the detergent, dodecyl maltoside, used successfully for crystallizing β 2AR was not found to be suitable for rhodopsin crystallization.

3.1.2 Use of citrate in proteorhodopsin purification

It was discovered serendipitously several years ago in the Braiman lab that citrate can selectively induce the precipitation of the transmembrane protein, proteorhodopsin (pR) under certain conditions [14, 15]. Beginning with pure solutions of pR solubilized in octylglucoside (OG) detergent, the precipitation occurred maximally, with minimal protein denaturation, at a detergent concentration of 0.3% OG, a pH of 5.5, and a citrate concentration exceeding 50 mM.

In chapter 2 of this thesis, the mechanism of citrate-induced pR precipitation was discussed. It was shown that citrate can induce pR precipitation at a broad range of pH as well as OG concentrations; and also that pR does not depend on the non-native His-tag for its precipitation by citrate. In the course of working with His-tag-free pR, a purification method was developed which included the use of citrate followed by a hydroxylapatite column chromatography. This method was originally developed due to a need to purify pR mutants that had to be prepared without the His-tag. However, producing pR without a His₆ tag is advantageous for other reasons: Firstly, the His-tag tends to be unstable over periods of days as evidenced by lower molecular weight bands

gradually appearing on SDS-PAGE gels (data not shown) with apparent self-cleavage occurring; whereas non-His-tagged pR is stable over many months of storage. Secondly, this method should be applicable for the eventual purification of homologously-expressed pR, which is produced at a low concentration in γ -proteobacteria.

3.1.3 Overview of the development of using phosphate in conjunction with citrate for pR preparation

For the original method of purification of His-tag-free pR, a phosphate gradient was used in order to elute pR from a hydroxylapatite column [15] (see also preceding chapter). The pR was eluted at 540 mM phosphate when using an OG concentration of 0.4%. While attempting to ascertain that the eluted pR was still susceptible to citrate induced precipitation, it was found that no amount of added citrate could precipitate pR out of the 540 mM phosphate solution. This led to a hypothesis that the phosphate has a binding site on pR, and that this binding site overlaps with the citrate binding site, resulting in competitive inhibition. As a result, phosphate can keep pR solubilized, and the presence of phosphate inhibits or prevents pR from being precipitated by citrate. This hypothesis was tested, and confirmed as described herein.

The ultimate aim has been to take advantage of this set of properties, by employing phosphate in combination with citrate in the process of pR purification. In addressing this aim, it is first important to know whether phosphate, by itself, can induce pR precipitation. This was answered, in the affirmative, in the preceding chapter. At 1% OG concentration, phosphate precipitates pR over a concentration range of 100-250 mM.

Several questions remain: (1) Does phosphate-induced precipitation depend on the detergent (OG) concentration? (2) Does phosphate-induced precipitation lead to further purification of pR i.e. beyond what can be achieved with citrate alone? After answering these questions, I proceeded to find a combination of citrate and phosphate that would yield a substantially pure pR, without requiring columns. A short protocol was established for preparation of *E. coli*-expressed pR using only selective precipitation with inexpensive salts of phosphate and citrate.

Preparations of pR using citrate and phosphate are likely to be directly usable for many commercial purposes, such as for holography and cosmetic applications. Combination of citrate and phosphates and perhaps other salts, may conceivably result in pR crystallization. However, use of these salts will need to be fine-tuned before a purer form of pR can be obtained that would be appropriate for structural studies.

3.2 MATERIALS AND METHODS

3.2.1 Protein Expression and Purification using citrate and hydroxylapatite column chromatography

Protocols as described in chapter 2 were followed for protein expression and purification of pR. Plasmid pBAD TOPO TA containing pR triple cysteine mutant (TCM) without the C-terminal His-tag was used for all the experiments. Mutants were prepared using this background sequence. Mutants were made using site directed mutagenesis as described in chapter 2.

3.2.2 Partial purification of pR

Initial partial purification of pR TCM was carried out using citrate buffer only. The detergent extract of pR-expressing *E. coli* was made using lysis buffer described in chapter 1. A 1/10 volume of 100 mM citrate buffer, pH 5.5 was added to the extract. After incubation on ice for 30 min, the citrate-treated extract was centrifuged at 5000 rpm, 4°C. The whitish impurities were discarded and another 1/10 volume of citrate buffer was added to the pink supernatant. More white impurities sedimented upon centrifugation and were discarded. To the pink supernatant, an equal volume of citrate was added and it was incubated overnight at 4°C. The pink precipitate was then resolubilized in 50 mM Tris buffer pH 9 also containing 3% OG. The solution was centrifuged at 5000 rpm, 4°C. Any impurities were discarded. A 1/10 volume of citrate buffer was again added to the reddish-pink supernatant. Impurities were discarded with each addition of citrate buffer. A bright pink pellet was finally obtained which was again resolubilized in Tris buffer pH 9 containing 3% OG. Addition of a 1/10 volume of citrate buffer was repeated one last time. This last addition yielded an intense purple pellet. If needed, this pellet was subjected to another round of Tris buffer resolubilization and consequent precipitation with citrate buffer. Two or three rounds of such purification by citrate-precipitation produced a pellet, which when resolubilized in 3% OG gave an $A_{280/520}$ ratio of 4-6.

Partial purification of pR TCM with mutations at IC1 cationic residues was done as described in chapter 2.

3.2.3 Preparation of pR using phosphate-citrate buffers in competition

Proteorhodopsin was partially purified as described in the preceding subsection, to $A_{280/520}$ ratio of 6. This pellet was resolubilized in 100 μL of 50 mM Tris-Cl buffer, pH 9 containing 3% OG. To this solution, 70 μL of a buffer containing 600 mM potassium phosphate and 30mM citrate, pH 7.0, also containing 0.4% OG, was added and it was incubated on ice for 1 h. Another addition of 70 μL of the above solution yielded a white pellet that was discarded. To the pink colored supernatant, 240 μL of 100 mM citrate, pH 7.0 also containing 0.4% OG was added that yielded a purple pellet.

3.2.4 Polyacrylamide Gel Electrophoresis

A 13% and an 18% SDS-polyacrylamide gel were used to analyze the purification of the protein. The purity level was estimated using the Kodak gel documentation system or BioRad Molecular Imager[®] ChemiDoc[™] XRS+ Imaging System.

3.2.5 Absorption Spectroscopy

UV-vis spectra measurements were recorded on Shimadzu UV-265 spectrophotometer and the data were analyzed using MIDAC GRAMS (Galactic Industries) software.

3.3 RESULTS

3.3.1 Phosphate precipitates pR between 50-400 mM phosphate concentration and 0.3-0.5% OG concentration

An initial aim was to confirm the conditions, described briefly in Chapter 2, for using phosphate by itself to cause the precipitation of pR. It was also of interest to see whether this phenomenon could be used to aid in the purification of pR, or if phosphate is acting simply as a non-specific protein precipitant. Another aim was the determination of the lowest detergent concentration that allows pR to be precipitated by phosphate without undergoing denaturation.

From prior work in which pR was eluted from hydroxylapatite columns at 540 mM phosphate 0.4% OG, and pH 7.0, it was apparent that pR is not susceptible to precipitation at this concentration of phosphate—at least at this pH and in this range of OG concentration. Nevertheless, it was found that pR in 0.3-0.5% OG was precipitated by a lower phosphate concentrations, specifically in a range of 50-400 mM phosphate buffered at pH 7.0 (See table 3.1). Once again, no precipitation was seen at 500 mM or 600 mM phosphate. On the other hand, increasing the OG concentration to 0.6% or above led to a much scantier precipitation when using 50-400 mM phosphate. Table 3.1 shows the effect of increasing concentration of OG on the precipitation of pR by phosphate buffer.

To summarize, phosphate by itself has the capacity to induce precipitation of pR throughout the tested concentration range of 50 - 400 mM, when using 0.3% to 0.5% OG.

However, higher concentrations of phosphate clearly inhibit precipitation of pR. As noted previously in section 3.1.2, this inhibition persists even when a different precipitant such as citrate is added at a concentration normally sufficient to cause precipitation. This suggests some kind of competitive inhibition is occurring, at least in this detergent concentration range.

Table 3.1: Determination of an optimum OG concentration under increasing concentration of phosphate at pH 7.0

Phosphate concentration → OG concentration ↓	50 mM	100 mM	200 mM	300 mM	400 mM	500 mM
0.3%	++	++	++	++	++	-
0.4%	++	++	++	++	++	-
0.5%	++	++	++	++	++	-
0.6%	+/-	+/-	+/-	+/-	+/-	-
0.7%	-	-	-	-	-	-
0.8%	-	-	-	-	-	-
0.9%	-	-	-	-	-	-

++ signifies essentially complete precipitation; +/- indicates scanty precipitation with much pR remaining in the supernatant; - refers to absence of precipitation

3.3.2 Determining the minimum phosphate concentration that induces pR precipitation

In order to pinpoint the minimum concentration of phosphate that induces precipitation of pR at pH 7 and at 0.4% OG, phosphate concentrations between 10 mM and 50 mM were tested for their ability to induce precipitation of pR, while holding the OG concentration constant at 0.4% (in both the pR sample and in the added phosphate buffer). It was found that 40 mM potassium phosphate is sufficient to cause complete pR precipitation at this detergent concentration and at pH 7 (Table 3.2).

Table 3.2: Determination of minimum concentration of phosphate that induces precipitation of pR at 0.4% OG

Phosphate concentration	10 mM	15 mM	20 mM	25 mM	30 mM	35 mM	40 mM	50 mM
Precipitation detected	-	-	-	-	-	-	++	++

++ indicates complete precipitation with a colorless supernatant; + indicates partial precipitation, leaving significant pink color in the supernatant, - indicates no precipitation.

3.3.3 Phosphate precipitation by itself does not eliminate impurities left behind by citrate-induced precipitation

The simplest possibility of using citrate and phosphate in a complementary fashion to purify pR would be to alternate these two precipitants. In order for the strategy to work however, it would be necessary for phosphate alone to selectively increase the purity of pR, beyond what had been achieved with citrate alone.

That is, it was necessary to determine whether precipitation with phosphate alone leads to further purification of pR, or simply co-precipitates pR along with the same contaminant proteins left after selective precipitation with citrate. This was measured by analyzing the purity level after successive precipitations with citrate alone, then phosphate alone, at 0.4% OG concentration. After each precipitation, pR was re-solubilized in Tris-Cl, pH 9, also containing 3% OG; and its UV-vis spectrum was obtained after a 10-fold dilution with water.

The $A_{280/520}$ ratio provides an estimate of purity (Fig. 3.1). There was not a significant difference in the $A_{280/520}$ ratio between the two samples, suggesting that additional precipitation with phosphate, by itself at 0.4% OG concentration, does not lead to any further purification of pR beyond what can be achieved with citrate precipitation alone.

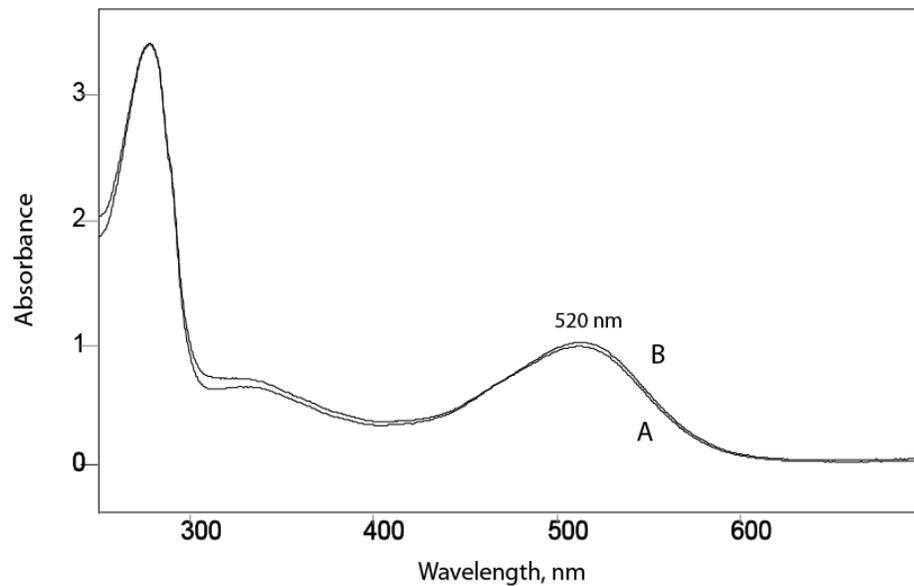


Fig. 3.1: UV-vis spectra of pR before and after phosphate precipitation: Spectrum (A) represents citrate-precipitated pR solubilized in 3% OG buffered with 50 mM Tris-Cl, pH 9. Spectrum (B) represents the this same pR sample, subsequently re-precipitated with 50 mM potassium phosphate buffer pH 7, also containing 0.4% OG, then resolubilized in 3% OG buffered with 50 mM Tris-Cl, pH 9.

3.3.4 Combination of phosphate inhibition with citrate-induced precipitation promotes the selectivity of pR precipitation (Done in collaboration with Jonathan Kim)

As discussed above, it was observed that pR precipitation is inhibited by phosphate at concentrations higher than 500 mM or at OG concentrations 0.6% or higher. Therefore, while the precipitation by phosphate alone does not lead to any further purification, it was hypothesized that the inhibition might be utilized for purification purposes.

That is, the next question to be addressed was whether the inhibitory effect of phosphate on citrate-induced precipitation was selective for pR, and in particular whether it was more selective than the citrate-induced precipitation itself. If so, it could be exploited in a purification step by combining phosphate and citrate treatment. In order to test this possibility, the phosphate concentration was maintained at 200 mM, high enough in order to inhibit the citrate-induced precipitation. Meanwhile, the concentration of OG was maintained more than 0.7% to prevent phosphate-induced pR precipitation. Higher concentrations of phosphate, such as 500 mM, were not used because addition of even a very high citrate concentration fails to precipitate pR under these conditions.

Partially purified pR samples, with an initial $A_{280/520}$ ratio of 5, were treated with three citrate concentrations - 15 mM, 30 mM, and 65 mM - while maintaining a constant concentration of 0.4% OG and 200 mM phosphate buffered at pH 7.0. All the three samples produced pR precipitates. These were then resolubilized in 50 mM Tris buffer,

pH 9.0 containing 3% OG, and UV-visible spectra were obtained. Citrate concentrations of 15 mM and 65 mM yielded $A_{280/520}$ ratio of 4 whereas 30 mM citrate sample produced a ratio of 3 (Fig. 3.2). These results show that phosphate at 200 mM aids most in the purification process of pR in combination with an optimum concentration of citrate. A concentration of 30 mM citrate is required to achieve the greatest degree of purification: lower or higher citrate concentrations than 30 mM of citrate, as tested above, do not yield as good results.

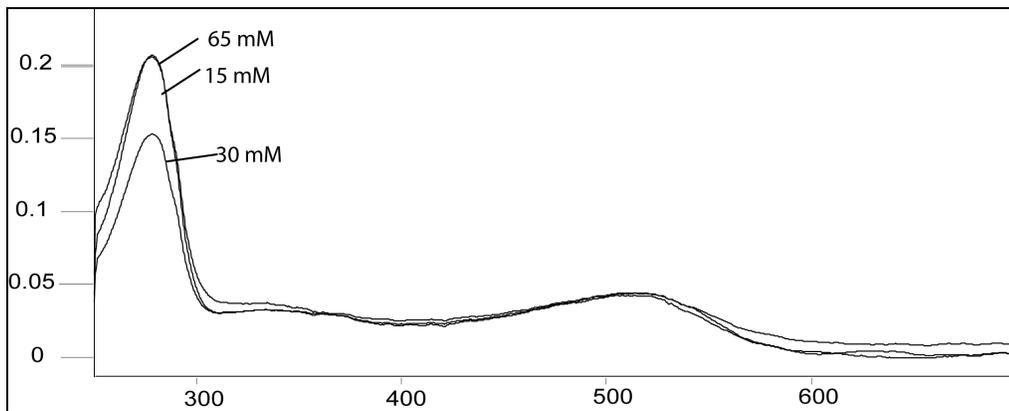


Fig. 3.2: Combination of phosphate inhibition with citrate-induced pR precipitation promotes pR purification: Citrate-purified pR was resolubilized in 50 mM Tris also containing 3% OG, pH 9. Equal volume of phosphate (400 mM, pH 6 also containing 0.4% OG), was added and it was divided into three parts. Citrate, buffered at pH 5.5, and also containing 0.4% OG was added to a final concentration of 15, 30 or 65 mM; then the three samples were centrifuged. The three pellets were separately resolubilized in 50 mM Tris buffer, pH 9, containing 3% OG, and UV-vis spectra were obtained as described in the text.

3.3.5 Phosphate likely binds to pR at the same site as citrate

Citrate binds to pR at the positively charged residues found at the first intracellular loop (see Chapter 2). Since phosphate appears to compete with the citrate, the citrate-binding sites were examined for phosphate-pR interaction. Partially purified double (K57Q/K59Q) and quadruple (R51Q/R53Q/K57Q/K59Q) mutants as well as a control wild type (pR TCM) were treated with increasing concentrations of phosphate ranging from 40 mM and 200 mM, buffered at pH 7 while maintaining 0.4% concentration of OG. Proteorhodopsin TCM was confirmed to precipitate at phosphate concentration of 40 mM, as shown in Tables 3.2 and 3.3. However, the mutants failed to precipitate with phosphate (Table 3.3). This observation strongly suggests that phosphate binds to the same site on pR as the citrate, since the same mutations eliminate binding effects. Other mutants will need to be examined to determine all the sites that are involved in phosphate binding.

Table 3.3: Determination of phosphate-pR interaction site using two citrate-binding site mutants

Phosphate concentration	40 mM	75 mM	100 mM	200 mM
Wild Type, pR TCM	++	++	++	++
Double mutant(K57Q/K59Q)	-	-	-	-
Quadruple mutant (R51Q/R53Q/K57Q/K59Q)	-	-	-	-

++ means complete precipitation; - refers to absence of precipitation

3.3.6 Optimization of purification conditions, using phosphate in combination with citrate

To explore the possibility of using phosphate in combination with citrate for the purpose of pR purification, initially a high phosphate concentration was maintained to keep the other contaminant proteins in solution while incrementally increasing the concentration of citrate to selectively precipitate the pR. The concentration of OG was also maintained above 0.7% to avoid any phosphate-dependent pR precipitation.

For an optimized purification protocol, solutions were simplified so as to contain fewer ingredients. This also resulted in higher yields of pR in a shorter time, as it was possible to simultaneously increase citrate and decrease phosphate, with the addition of a single concentrated citrate solution acting as precipitant.

Proteorhodopsin was initially purified to an $A_{280/520}$ ratio of 5-6 using two cycles of citrate precipitation followed by resolubilization in 10 mM Tris pH 9 containing 3% OG. Then to the Tris-OG-resolubilized pR, an equal volume of 500 mM phosphate also containing 0.4% OG was added, and the mixture was incubated at 4°C for 30-45 min. Centrifugation yielded a white pellet and a pink supernatant.

To the supernatant initially containing 250 mM phosphate and 0 mM citrate, a 100 mM citrate buffer, pH 5.5 was incrementally added. An initial addition of 1/10 volume produced a white pellet, corresponding to contaminants, and pink supernatant corresponding to pR. Repeated addition of similar volumes of citrate buffer caused

additional white impurities to precipitate. Fig. 3.3 shows the incremental purification of pR with citrate, in the presence of phosphate. The top panel summarizes the experiment: as the concentration of citrate increases, the concentrations of phosphate and OG decrease. When the concentration of citrate reached 34 mM and that of phosphate decreased to 165 mM, a light pink pellet was obtained. Upon further additions of the citrate buffer, each followed by centrifugation, it was noticed that the pellet improved both in size (yield) and intensity of color (purity). It was also observed that a point was reached where no amount of citrate could precipitate any additional pR from the faintly pink supernatant.

It can be seen that precipitation of whitish impurities begins with the first addition of about 9 mM citrate and 228 mM (Fig. 3.3 lane 2). However, most of the pR remains solubilized. The purity level of pR is 5.3% at this stage, as judged by SDS-PAGE analysis, meaning that 5.3% of the total precipitate is pR. This level of purity and yield clearly demonstrate that phosphate competes with citrate, i.e. phosphate blocks citrate-induced precipitation.

However, with each further addition of citrate and concomitant dilution of phosphate, more impurities are precipitated out, while leaving increasingly pure pR in the supernatant. The cause of the precipitation of these impurities is not clear; they could be binding to citrate either non-specifically or by specific protein-citrate interaction. Addition of citrate also causes precipitation of pR leading to loss of pR.

It can be seen that with the progress of citrate addition and simultaneous dilution of phosphate, more pR is precipitated while also being purified (Fig 3.3). The majority of impurities are removed at the earlier phase of purification leaving few residual contaminants. It is intriguing that three contaminants co-precipitate with pR at apparent molecular weights of ~31, 45, 66 and about 97 kDa (See appendix 3.9 for the identity of contaminants). It should also be noted that the amount of sample loaded in lane 7 is half of the amount loaded in lanes 3-6. This observation means that the pR yield may be under-estimated.

Fig 3.3 (lane 7) also demonstrates that pR TCM can be substantially purified with 124 mM phosphate and 50 mM citrate. Fig. 3.4 shows the UV-vis spectra and an intense purple pellet of the precipitate obtained with 124 mM phosphate and 50 mM citrate. Phosphate aids in the purification of pR; part of the reason for this is that phosphate raises the concentration of citrate required to precipitate pR by 5-fold, but does not increase the citrate concentration needed to precipitate impurities. Note that complete precipitation of pR is obtained with 10 mM citrate in the absence of phosphate (chapter 2). The purification thus arises from the competing nature of phosphate for citrate, and the selectivity of this competition process for pR.

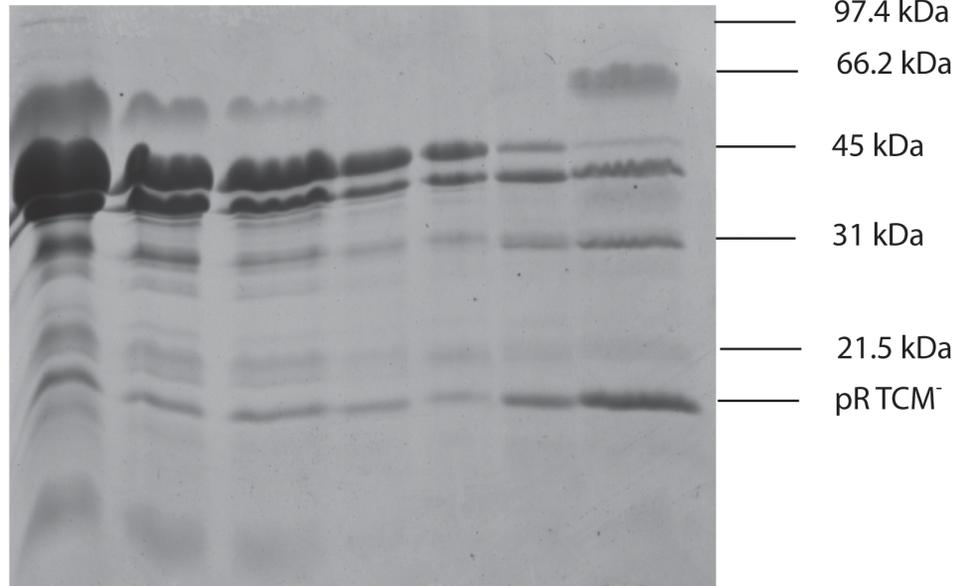
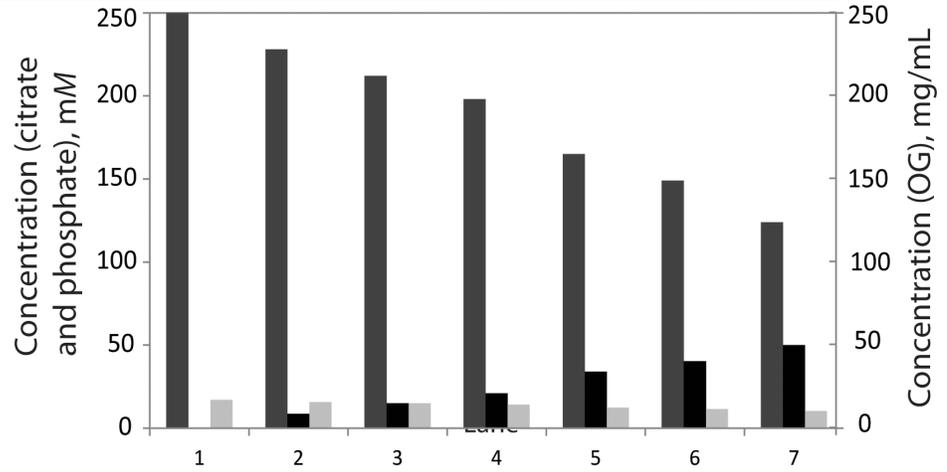
It should be pointed that even 50 mM citrate is unable to completely precipitate pR in the presence of phosphate, yet more evidence of phosphate-citrate competition. The precipitation reaches a saturation point leaving a faint pink pR in the supernatant adding

to the loss of recovery. Conditions will need to be explored to minimize the overall losses of pR recovery.

Changing the concentrations of phosphate and citrate simultaneously provides a sharp solubility transition for pR. With the above conditions, this occurs at a concentration of 149 mM phosphate and 40 mM citrate (Fig. 3.3 Lane 5).

The yield and purity results at each step of citrate-phosphate purification are summarized in Table 3.4. Calculation of yield and purity of the supernatant was done using the spectra (data not shown) obtained from the supernatants. The corresponding precipitate obtained after removing the supernatants was analyzed on SDS-PAGE. Purity of the precipitates was estimated from the gel. Fig. 3.5 shows the increase in purity of pR in supernatant and precipitate. Initially, the supernatant contains more pR that becomes purer as more impurities are removed. The precipitate contains more impurities with a small fraction of pR. As purification progresses, the precipitate contains most pR, achieving 40% purity.

Lane	1	2	3	4	5	6	7
Phosphate, mM	250	228	212	198	165	149	124
Citrate, mM	0	8.75	15	20.8	33.9	40.4	50
OG, mg/mL	17	15.8	15	14.3	12.4	11.5	10.3



Lane 1 2 3 4 5 6 7

Fig. 3.3: SDS-PAGE analysis of citrate-phosphate purification of pR:

Each lane (in the third panel) contains the proteins in the pellet obtained by sequential precipitations of a single sample of partially-purified pR. The pR was initially in a buffer containing 50 mM Tris-Cl, 250 mM sodium phosphate, pH 7 containing 1.7% OG, and was subjected to sequential additions of a 1/10 volume of 100 mM citrate buffer, pH 5.5, to attain the final concentrations as shown in the top panel. The wavy shape of the bands in the first 3 lanes is due to the presence of high concentration of OG. Detergent tends to cause spreading and anomalous behavior of protein bands.

The second panel shows the variations in the concentrations of citrate, phosphate and OG as more citrate was added. Column colored ■ represents concentration of phosphate, ■ depicts citrate concentration and □ shows the concentration of OG.

The last panel shows the color of the pellets obtained at various stages with citrate-phosphate purification.

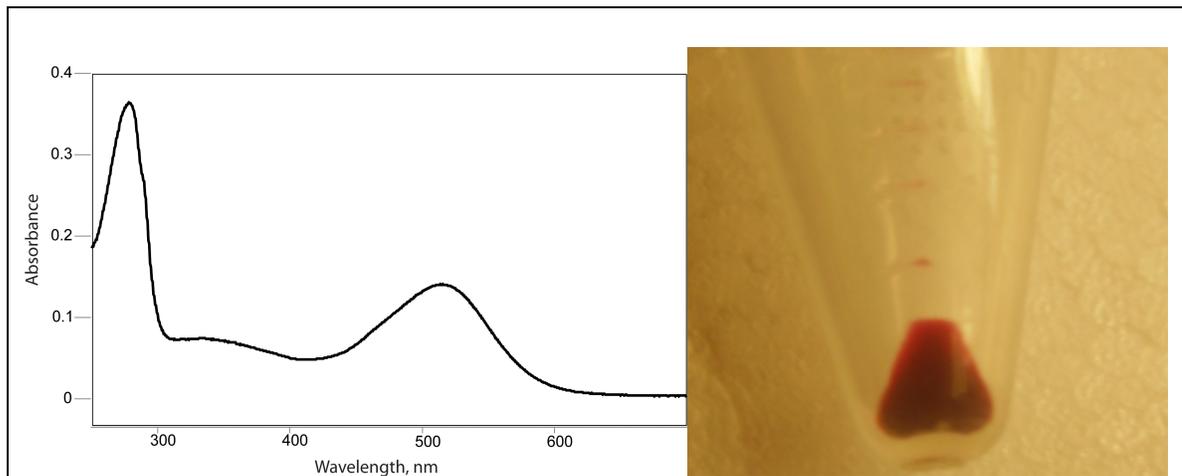


Fig. 3.4: Citrate-phosphate purified pR TCM: Proteorhodopsin TCM was purified using the protocol described in the text. The spectrum (at left) and the purified protein pellet (at right) correspond to lane 7 of Fig. 3.3. This is an uncorrected spectrum showing $A_{280/520}$ ratio of 2.57. The yield of the pellet, relative to the total pR present in the original cell culture, is 34.68%; and the purity is 40-48% as judged by SDS-PAGE (with Coomassie Blue staining and optical scanning) and UV-vis spectra.

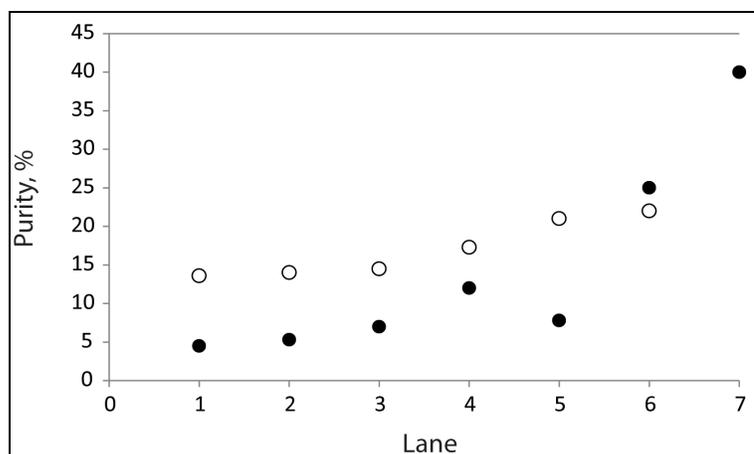


Fig. 3.5: Purity of supernatant and precipitate: Progress of purification of pR using citrate-phosphate buffers. ○ depicts progress of purity in supernatant, ● depicts progress of purity in precipitate.

Table 3.4: Yield and purity of pR at various stages of purification with citrate and phosphate

Lanes	1	2	3	4	5	6	7
Yield of pR in supernatant, % (from spectra)	53	49	46	44	42	38	--
Purity of pR in supernatant, % (from spectra)	13.6	14	14.5	17.3	21	22	--
Purity of pR in precipitate, % (from gel)	<4.5	5.3	7	12	7.8	25	40

The yield is calculated by considering the total pR (mg) of the lysate as 100. Purity refers to the fraction of pR present in the total protein, expressed as percent.

3.3.7 Scaled-up test of a simplified citrate/phosphate protocol for preparation of pR

A 1200 mL *E. coli* culture of pR-expressing cells was subjected to purification using citrate and phosphate buffers. A summary of this process is as follows: a buffer, containing both a high concentration of phosphate and a low concentration of citrate, was added initially to a partially purified pR, in order to selectively precipitate the impurities. High initial concentration of phosphate and effective OG ensured pR was soluble at the start. At this stage, citrate was kept to a minimum concentration in order to prevent any citrate-induced pR precipitation early in the process. However, whitish impurities precipitated allowing their removal. Subsequently, phosphate-free citrate buffer was added to compete with phosphate, and thereby selectively precipitate pR.

Detailed protocol

As described in the preceding subsection, proteorhodopsin was partially purified with two cycles of citrate precipitation, each followed by 3% OG/Tris buffer resolubilization, to an $A_{280/520}$ ratio of 6. The pellet was again resolubilized in minimal volume of 3% OG buffered with 50 mM Tris-Cl, pH 9.0. A buffer of 600 mM phosphate and 30 mM citrate, pH 7.0, containing 0.4% OG was added gradually with gentle mixing, followed by centrifugation. When the concentrations of phosphate and citrate reached 350 mM and 17.5 mM, respectively, whitish impurities settled out while pR remained in the solution. Complete precipitation of pR was observed by adding to this solution an equal volume of phosphate-free citrate buffer, 100 mM, pH 7.0. This yielded final concentrations of 175 mM phosphate and 58.75 mM citrate. The final concentration of OG was close to 1%. Results above, showing that a similar concentration of phosphate

by itself in the presence of 1% OG fails to cause precipitation, show that it is citrate, and not phosphate, that led to precipitation at this point. But the competition by phosphate is what leads to purification of pR.

In support of these conclusions, increasing the concentration of citrate while proportionately increasing the phosphate in the same ratio, does not lead to precipitation of pR. Furthermore, a phosphate:citrate concentration ratio of 20:1 is adequate to provide partial, but not complete inhibition of the citrate induced precipitation.

Fig. 3.6 shows the spectra of pR at various stages of purification, and Fig. 3.7 shows an SDS-PAGE analysis. Proteorhodopsin is easily purified to $A_{280/520}$ ratio of 6 (Fig. 3.7, B) with two cycles of citrate only to a purity level of 35% (Fig. 3.7, lane B). This ratio was improved to 3.3 (Fig 3.7, C) and to a purity of 67% as measured by SDS-PAGE analysis (Fig. 3.7, lane C) after hydroxylapatite column chromatography. Fig. 3.7 also demonstrates that pR with $A_{280/520}$ ratio of 6 can be enhanced to an even better $A_{280/520}$ ratio of 2.6 (Fig. 3.7, D), corresponding to a purity level on SDS gel of 50% (Fig 3.7, lane D), by simply using a combination of phosphate and citrate. These results suggest that phosphate can be used in the method of pR preparation, mainly as a selective competitor of the precipitating citrate.

Table 3.5 shows the purification yield and purity of pR obtained using citrate method followed by either hydroxylapatite column chromatography or a combination of citrate and phosphate buffers. The amounts of pR, total protein and yields were calculated

using the A_{520} and A_{280} values from the spectra (Fig. 3.6) as described in the footnote of Table 3.5. Comparison of purity levels obtained from the spectra and those from the SDS-PAGE shows a discrepancy between the values. This could be attributed to the fact that the specific A_{280} value at the earlier steps of purification reflect that most of the protein content was average soluble proteins, for which $A_{280} = 1$ for a 1 mg/mL solution, with a smaller contribution from pR, for which $A_{280} = 1$ for a 0.4 mg/mL solution. However, as the purity of pR increases, the A_{280} value due to contaminants is smaller compared to that by pR. For the purity calculations, since the amount of pR is added to the total protein content, and then the amount of pR purity is obtained by dividing the entire protein content, the values appear smaller.

Besides the above explanation, it is also possible that the protein undergoes denaturation and loses its chromophore retinal. Therefore, the absorption at 520 nm is relatively low giving it a low ratio of $A_{280/520}$. However, the protein still shows up on the SDS-PAGE even though it has lost its chromophore retinal, giving an impression of higher yield.

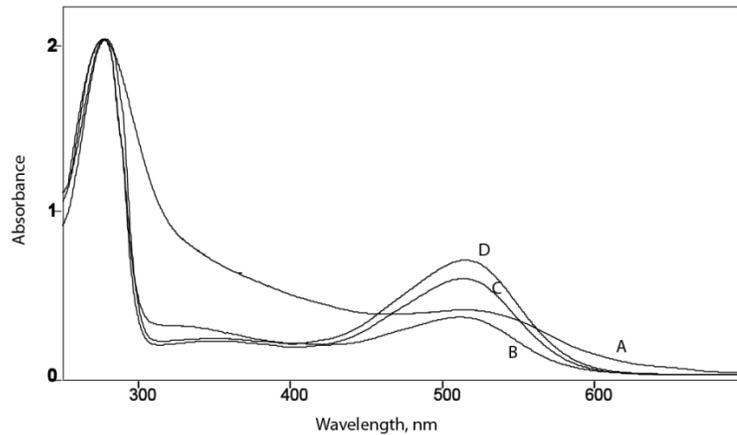


Fig. 3.6: UV-visible absorption of pR at various stages of purification using different methods: All the spectra were measured in the presence of 3% OG at pH 9 buffered by 50 mM Tris-Cl. Spectra A, B, C and D correspond to pR samples in Lanes A, B,C and D of Fig. 3.2.

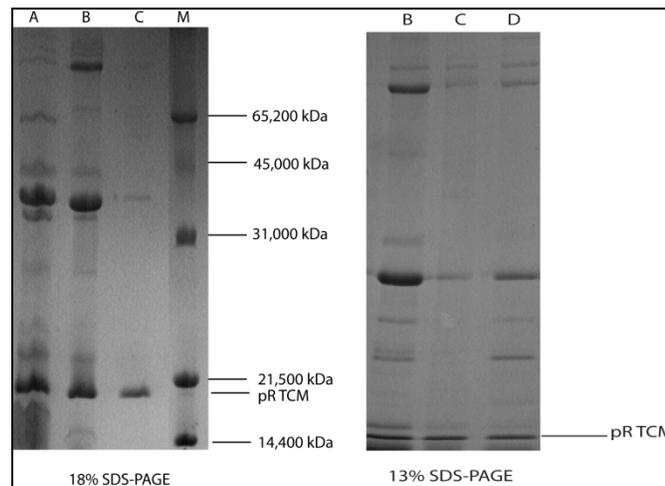


Fig. 3.7: SDS-PAGE analysis of pR purification: **Left, 18% SDS-PAGE:** Lane A contains pR purified by one cycle of citrate precipitation and resolubilization, Lane B shows pR purified by two cycles of citrate precipitation/resolubilization, Lane C is the hydroxylapatite-column-purified pR, Lane M contains broad-range SDS-PAGE molecular weight markers, BioRad. **Right, 13% SDS-PAGE:** Lane B and Lane C are the same samples as correspondingly labeled in the 18% gel. Lane D depicts the pR sample purified exclusively through differential precipitations with different combinations of citrate and phosphate buffers, as described in the text.

Table 3.5: Proteorhodopsin purification using citrate followed by hydroxylapatite column chromatography or combination of citrate and phosphate buffers

Purification stage	$A_{280/520}$ Ratio	Total protein (mg)	Amount of pR obtained (mg)	Purity (total pR/total protein) (%)	Enrichment Factor	Yield (%)
1. Whole cell extract*	44.4	406	6.0	1.47	1	100
2. After 1 citrate precipitation cycle (spectrum A)	28	251	5.67	2.3	1.5	94.5
3. After 2 citrate precipitation cycles (spectrum B)	6.36	41	4.95	12	8.2	82.5
†4a. After hydroxylapatite column (spectrum C)	3.33	5.4	1.66	31	21	28
†4b. After citrate-phosphate purification (spectrum D)	2.6	4.34	2.08	48	32.4	35

Proteorhodopsin TCM was expressed in *E. coli* UT5600. Purification was carried out as described in the text. The amount of pR obtained at each step was calculated based on A_{520} , the absorbance due to retinylidene Schiff base at pH 9, of an appropriately diluted aliquot. Molar extinction coefficient at A_{520} considered was $50,000 \text{ M}^{-1} \text{ cm}^{-1}$, molecular weight of pR was considered to be 30,000 Da. The non-pR contaminants of the same sample were estimated based on the amount of A_{280} that could not be contributed to pR, estimating 1 mg/mL for $A_{280} = 1.0$, this was added to the concentration of pR to determine the total protein content. Purity was calculated by dividing the amount of pR by the total protein.

† The sample from step 3 was divided into two parts: one part was subjected to hydroxylapatite column purification while the other was purified with a combination of citrate and phosphate buffers. Total yield is a combination of both 4a and 4b.

*The A_{520} of whole cell extract was corrected by adjusting the baseline to negate the effect of light scattering.

3.4 DISCUSSION

3.4.1 Phosphate-induced pR precipitation

Precipitation of pR by phosphate suggests that either there is a binding site on pR for phosphate, like citrate, or that phosphate has a salting-out effect [3, 4]. At 0.3-0.5% OG phosphate effects precipitation of pR only within a range of 40 – 400 mM. With only these two species present, i.e., phosphate and OG, at any concentration above the indicated ranges, pR precipitation is inhibited. This behavior, taken by itself, could be consistent with a salting-out effect (and salting-in at the higher phosphate concentrations).

However, it is more likely that a specific binding site for phosphate is present on pR, because of an observed competition between phosphate and citrate. Specifically, it was found that 250 mM phosphate in 1.7% OG causes precipitation only of impurities, but not pR. This concentration of phosphate also inhibits citrate from binding and leading to pR precipitation, but in a competitive fashion. That is, an increase in the concentration of citrate can overcome the inhibitory effect of phosphate. Under At 0.3-0.5% OG, and in the absence of phosphate, pR can be completely precipitated with 10 mM citrate. However, in the presence of phosphate, the citrate concentration required for pR precipitation is increased to about 35 mM. This competitive effect of elevated citrate would not be expected if the high level of phosphate were simply causing a “salting-in” effect.

An even more important piece of evidence favoring a specific competitive binding is that the two IC1-loop mutants that we examined for effects on phosphate binding - both the double (K57Q/K59Q) and quadruple (R51Q/R53Q/K57Q/K59Q) charge neutralization - eliminated both citrate and phosphate induced precipitation at all concentrations (at pH 7 and 0.4% OG).

This observation strongly suggests that phosphate binds to pR at the IC1 loop, specifically at the positively charged residues. We conclude further that phosphate at higher concentrations likely binds at the same site (and possibly others); and this form of binding by phosphate can inhibit precipitation of pR; also, that a 3-4 fold higher-than-usual concentration of citrate can out-compete the phosphate binding on the IC1 loop, and thereby restore the citrate-induced precipitation. Thus the phosphate and citrate binding sites overlap, and phosphate binds to that site preventing 10 mM citrate from binding to it. Therefore, pR cannot be precipitated by citrate, except at significantly elevated concentrations, in the presence of the inhibitory phosphate.

It has previously been shown that 300 mM phosphate binds pR and lowers the pK_a of Asp, the counter-ion of Schiff base of pR [16]. Other anions such as chloride, sulfate and halogeno-acetates were also shown to influence the pK_a of pR at even lower concentrations and were suggested to have a binding site on pR. These anions are assumed to bind away from the active site, Asp97, based on the fact that none of them affects the absorption maxima of pR. They are speculated to bind either at the extracellular side or possibly cytoplasmic side of pR. The results of my experiments with

the pR mutants is consistent with these earlier conclusions, and in particular indicates a binding site at the cytoplasmic side on the intracellular loop, IC1.

In a previous study with bovine rhodopsin, the presence of solute anions, including sulfate, chloride and nitrate, was shown to increase the pK_a of Schiff base group of the chromophore [17]. These anions were also shown to influence the conformational equilibrium of Meta I/Meta II intermediate states, suggesting a well-defined ion binding site at or within the protein. It has been observed that several anions, including sulfate and chloride, cause the precipitation of pR under select counterion conditions, as mentioned in chapter 2.

Phosphate-induced pR precipitation is dependent on the concentration of OG. In the presence of 0.7% OG, which is the critical micelle concentration (CMC), or higher, phosphate loses its ability to precipitate pR. Phosphate groups perhaps bind to or interact with certain amino acid groups that become buried inside the pR-detergent micelle as more detergent is added.

The presence of phosphate does not affect the absorption maximum of pR, as evidenced by the presence of a peak at 520 nm at pH 9. The effect of phosphate on the pR photocycle remains to be examined. Sharaabi et al. [16] showed that pR binding of other anions such as chloride, sulfate and halogeno-acetates affects the photocycle by altering the decay rate of the M intermediate.

3.4.2 Preparation of pR using citrate-phosphate buffers

The combination of citrate and phosphate for the purpose of pR preparation has many of the same advantages as the citrate method by itself, including low cost and rapidity. The combination of citrate and phosphate gives a higher purity than citrate alone, and can effectively eliminate the requirement of any chromatography column in order to obtain pR samples suitable for some research purposes, e.g. for simple spectroscopic measurements screening for alterations in absorption maximum.

A pR band is present in all the lanes in Fig. 3.3, showing that pR is partially lost along with the bulk of the impurities, before a relatively pure pR band can be precipitated with a higher concentration of citrate. Conditions need to be explored where the impurities or contaminant proteins can be either precipitated before pR, or that they can be kept solubilized in the phosphate solution while pR selectively and completely precipitates out.

3.5 CONCLUSIONS

Phosphate by itself precipitates pR selectively but does not purify pR beyond what is achieved with citrate alone. However, it binds to the impurities allowing citrate to selectively precipitate pR. Phosphate shares at least one of its binding sites of pR with citrate, i.e. the lysines and arginines located at the IC1. There might be other binding sites of phosphate but this particular overlap of the interaction sites explains why citrate cannot precipitate pR in the presence of phosphate.

While proteins have been studied for their binding to various salts, there is no report yet on the use of salts in the purification of membrane proteins. This study underscores the application of inexpensive salts for the purification of membrane protein such as pR, more importantly, without the use of any chromatographic columns. The entire process of pR production is rendered extremely inexpensive, beginning with the protein expression because pR is expressed in *E. coli* making the protein production fast and less labor-intensive.

Proteorhodopsin purified by citrate-phosphate method lacks the polyHis-tag which makes the process applicable to the purification of pR from natural source. The three-dimensional structure of pR has not been solved yet. Because pR is a proton pump and the His₆ tag would be expected to affect binding of protons to various groups with the protein, it is desirable to produce pure protein without the tags for structural and functional studies [18]

Proteorhodopsin shares properties with BR that make them both attractive biomaterials for certain applications, e.g. as a holographic recording medium [19, 20]. Unlike BR, pR can be expressed rapidly in *E. coli* in a fully functional form, which makes it more attractive for such uses. Due to its photochromic properties, pR is also being pursued as a cosmetic ingredient as was suggested in a patent issued for similar use of BR [21]. These applications demand significant amount of protein. The above method of pR preparation has the potential to produce such quantities in short time period.

This method of purification might also potentially be applied toward the purification of pharmaceutically important membrane proteins known as GPCRs. This somewhat-speculative potential application will be explored further in the next chapter of this thesis.

3.6 Acknowledgements

Jonathan Kim, a hard-working and devoted undergraduate student, worked with me on the role of phosphate in purification of pR. I am very thankful for his contribution to this chapter. I also thank another undergraduate, Halli Benson, who worked with me on fine-tuning the concentration of phosphate.

Chapter 4

HETEROLOGOUS EXPRESSION OF BOVINE RHODOPSIN, A MAMMALIAN RHODOPSIN, IN *E. coli*

4.1 INTRODUCTION

Based on the identification of a short compact motif responsible for pR-citrate interaction, it was hypothesized that it might be possible to exploit this motif for the purification of other membrane proteins. This chapter describes a strategy that was developed to test this hypothesis, and preliminary attempts to express a mammalian membrane protein, bovine rhodopsin, in *E. coli* with the aim of adapting the citrate purification method to G-protein-coupled receptors (GPCRs). Similar to pR, GPCRs generally contain a cluster of positively charged amino acid residues in their first intracellular loop (Fig. 4.1 and 4.2). If their spacing in the native primary sequence is inadequate, site directed mutation(s) can be carried out on the amino acids present in the loop to generate positively charged amino acid (s) with spacing matching that is present in pR.

```

MNGTEGPNFYVPFSNKTGVVRSPEAPQYYLAEPWQFSMLAAYMFLIMLGFPINFLTLYVTVQH
KKLRTPLNYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGFFATLGGEIALWSLVVL
AIERYVVVCKPMSNFRFGENHAIMGVAFTWVMALACAAPPLVGWSRYIPEGMQCSCGIDYYTPH
EETNNESFVIYMFVVHFIPLIVIFFCYGQLVFTVKEAAAQQQESATTQKAEKEVTRMVIIMVIAFLI
CWLPYAGVAFYIFTHQGSDFGPIFMTIPAFFAKTSAVYNPVIYIMMNKQFRNCMVTTLCCGKNPLG
DDEASTTVSKTETSQVAPA

```

Fig. 4.1: Primary structure of bovine rhodopsin showing positively charged residues: The underlined segments represent the loops, beginning with the extracellular loop and ending with the cytoplasmic loop. The first intracellular loop is shown in bold and the positively charged residues are highlighted.

To use citrate for the purpose of GPCR purification, the first step was to express a GPCR protein in bacteria. *E. coli* was chosen as the expression tool because it multiplies fast, with a doubling time of ~30 minutes, consequently producing the protein in large quantities in a shorter time period. Also, with *E. coli*, the behaviors of lipids and co-contaminants that are co-purified at each step when using the citrate purification method for pR are recognizable. Switching to a different expression organism would be expected to introduce complications resulting from a different mix of contaminant proteins and lipids.

However, membrane proteins, especially integral membrane proteins, are difficult to express in *E. coli*. High-level expression of such proteins is often toxic to the host cell. To circumvent such difficulties, it is often possible to take advantage of the technique of

gene fusion. Another membrane protein, native to *E. coli*, is genetically fused to the protein of interest at either the C- or N-terminus. These fusion proteins are also known as tags which are either short fragments or the entire length of protein. Bacteriorhodopsin (BR), an archaeal homolog of pR, has been expressed in *E. coli* using tags such as maltose-binding protein (MBP) [1] or β -galactosidase [2]. These larger fusion proteins allow efficient expression of the protein of interest. The drawback of tag-fusion is that it adds another step in the process of purification; the tags typically need to be cleaved in order to produce a pure protein in the native form for structural studies.

Because it is possible to obtain relatively high levels of pR expression in *E. coli* - up to 10 mg per liter of culture, an uncommonly high level of expression for an heptahelical membrane protein - fusion with pR or with pR segment(s) appears as an attractive strategy for expression of GPCRs.

As a first candidate for heterologous expression of a GPCR in *E. coli*, an attempt was made to express mammalian rhodopsin, specifically bovine rhodopsin. GPCRs are of great pharmacological importance, with about 30-50% of drugs being targeted at GPCRs [3]. Some of the drugs that target GPCRs are shown in Table 4.1 [4]. Table 4.2 shows some of the hereditary diseases associated with malfunctioning GPCRs [4].

The binding sites of GPCR ligands are buried in between the seven transmembrane helices. Structure-guided design of drugs that target GPCRs has been limited due to the lack of three-dimensional structures of these binding sites. This is

largely due to the challenge in purification of these proteins in sufficient quantities to carry out crystallization studies. Bovine rhodopsin was the first GPCR to be crystallized mainly due to its abundance and high concentration in an easily-obtained tissue, cow retinas [5]. Methods for purification of it were first developed in the 1940s, allowing widespread research attempts on its crystallization. However, it still took a span of around 50 years to get the protein pure enough that is suitable for three-dimensional (3D) structural studies.

Homologous expression of most other mammalian GPCRs in large quantities is difficult, since they are generally expressed in very low quantities in the cell. Considerable effort is being applied into heterologous expression of functional GPCRs, but production of even modest quantities of membrane proteins is a challenge. Rat neurotensin receptor is a GPCR that has been expressed in *E. coli* as a dual fusion protein with MBP-tag at the N-terminus and thioredoxin (Trx)-tag at the C-terminus. This yielded 10 mg of the fusion protein from 200 L of culture [6]. (It should be noted that only half of the 10 mg was attributable to the native rhodopsin sequence; the other half was MBP itself). Another example of *E. coli* expression of a GPCR is that of chemokine receptors, CR, CCR3, CCR5, and CX3CR1 [7]. The final yields of purified receptors obtained were between 0.1 to 3 mg/L of culture.

Our lab and others have been successful in expressing 5-10 mg/L of pR in *E. coli*. [8]. It should be noted that while expression of pR is by definition considered heterologous in *E. coli*, the native pR was originally found in *E. coli* related bacteria, γ -

proteobacteria. This might explain the ease and the high success rate of expression of pR in *E. coli*. If a membrane protein such as a GPCR could be expressed in *E. coli* and purified with the citrate method, the consequences of such results could be immense, as the protein could be obtained in larger quantities and purified with a simple, inexpensive and labor-easy method.

Table 4.1: GPCR-targeted drugs

Trademark	Generic name of drug	Disease	Target receptor
Allegra	Fexofenadine	Allergies	H1 antagonist
Atrovent	Ipratropium	Asthma	Mixed muscarinic antagonist
Axid	Nizatidine	Renal dysfunction	H ₂ antagonist
Betaloc	Metoprolol	Asthma/COPD	Beta1 antagonist
BuSpar	Buspirone	Depression	5-HT ₁ agonist
Cardura	Doxazosin	hypertension	Alpha ₁ antagonist
Claritin	Loratidine	Allergies	Antihistamine H1 antagonist
Cozaar	Losartan	Hypertension	AT ₁ antagonist
Diovan	Valsartan	Hypertension	AT1 antagonist
Duragesic	Fentanyl	Pain	Opioid agonist
Gaster	Famotidine	Ulcers	H ₂ antagonist
Hytrin	Terazosin	Hypertension	Alpha ₁ antagonist
Imigran	Sumatriptan	Migraine	5-HT ₁ agonist
Lupron	Leuprolide	Cancer	LH-RH agonist
Neurotonine	Gabapentin	Neurogenic pain	GABA B agonist
Pepcidine	Pamotidine	Ulcers	H ₂ antagonist
Plavix	Clapidogrel	Stroke	P2Y ₁₂ antagonist
Prepulsid	Cisapride	Digestion	5-HT ₄ ligand
Prostap SR	Leuprorelin	Prostate cancer	LH-RH agonist
Risperdal	Risperidone	Psychosis	Mixed D ₂ /5-HT ₂ antagonist
Seloken	Metoprolol	Hypertension	Beta1 antagonist
Serevent	Salmeterol	Asthma	B2 agonist
Singulair	Montelukast	Asthma	LTD ₄ antagonist
Tagamet	Cimetidine	Ulcers	H ₂ antagonist
Tenormin	Atenolol	Hypertension	Beta2 antagonist
Ventolin	Salbutamol	Asthma/breathing problems	Beta2 agonist
Zantac	Ranitidine	Ulcers	H ₂ antagonist
Zofran	Ondansetron	Antiemetic	5-HT ₃ antagonist
Zoladex	Goserelin	Cancer	LH-RH agonist
Zyprexa	Olanzapine	Schizophrenia	Mixed D ₂ /D ₁ /5-HT ₂ antagonist
Zyrtec	Cetirizine	Allergies	Antihistamine H1 antagonist

Source: The table is a redrawn version of the table taken with permission from Flower D. R [4]. More drugs and the diseases have been added to this table.

Table 4.2: Diseases associated with malfunctioning GPCRs

Disease	Receptor	Type/function of mutation
Color blindness	Red and green opsins	X chromosome rearrangements
Stationary night blindness	Rhodopsin	Missense mutation
Retinitis pigmentosa	Rhodopsin	Apoptosis of rod cells
Retinitis pigmentosa	Rhodopsin	Null mutations
Nephrogenic DI	V2- receptor	Loss of function
Isolated glucocorticoid deficiency	ACTH receptor	Loss of function
Hyperfunctioning thyroid adenomas	TSH receptor	Missense
Familial precocious puberty	LH receptor	Missense
Familial hypocalciuric hypercalcaemia	Ca ²⁺ -sensing receptor	Missense
Neonatal severe hyperparathyroidism	Ca ²⁺ -sensing receptor	Missense

Source: The table is a redrawn version of the table, taken with permission from Flower D. R. [4].

4.2 Methods and Rationale

4.2.1 Construction of chimera plasmid coding for bovine rhodopsin fused to N-terminal leader sequence from pR

Initially, the unmodified bovine rhodopsin gene (originally obtained from Jeremy Nathans) was cloned into the expression site of commercially available plasmid pBAD-TOPO-TA from Invitrogen. The resulting recombinant plasmid pBAD TOPO-TA had its sequence verified in order to confirm the correct direction of insertion, then was

transformed in *E. coli* Strain UT5600, the same that has been used for pR expression. In this construct, bovine rhodopsin is under the control of arabinose promoter. However, even after addition of 0.2% arabinose and 11-cis-retinal, the protein failed to express, fold, and/or form a holoprotein, as evidenced by lack of a reddish chromophore.

This was a generally-expected result, as it has not been possible in general to express GPCRs in *E. coli*. A widely-accepted reason for this is that *E. coli* is unable to insert the N-terminal (N-T) region of bovine rhodopsin into its membrane and therefore is incapable of expressing it without an N-terminal fusion to a native *E. coli* protein.

Therefore, we adopted the strategy of fusing bovine rhodopsin to an N-terminal leader sequence from pR. Since pR is successfully expressed by *E. coli*, it is apparent that the N-terminal portion of pR is able to translocate through the membrane and consequently lead the entire protein to span the membrane seven times. It was hypothesized that if the N-T region of bovine rhodopsin that forms the extracellular segment was replaced by the N-terminal portion of pR corresponding to the extracellular segment of pR, then bovine rhodopsin might also express and properly fold in *E. coli*.

Fig 4.2 and 4.3 show the secondary structures of bovine rhodopsin and pR respectively. Construction of the chimeric protein was accomplished by performing a site-directed mutagenesis using pBAD TOPO TA containing bovine rhodopsin as the cloning vector and replacing the N-terminal extracellular loop of bovine rhodopsin with that of pR. To be specific, the first 33 amino acids at the N-terminus in bovine rhodopsin

were replaced by the first 27 amino acids of pR. The gene and protein sequences of pR and bovine rhodopsin are shown in Fig. 4.4 and 4.5.

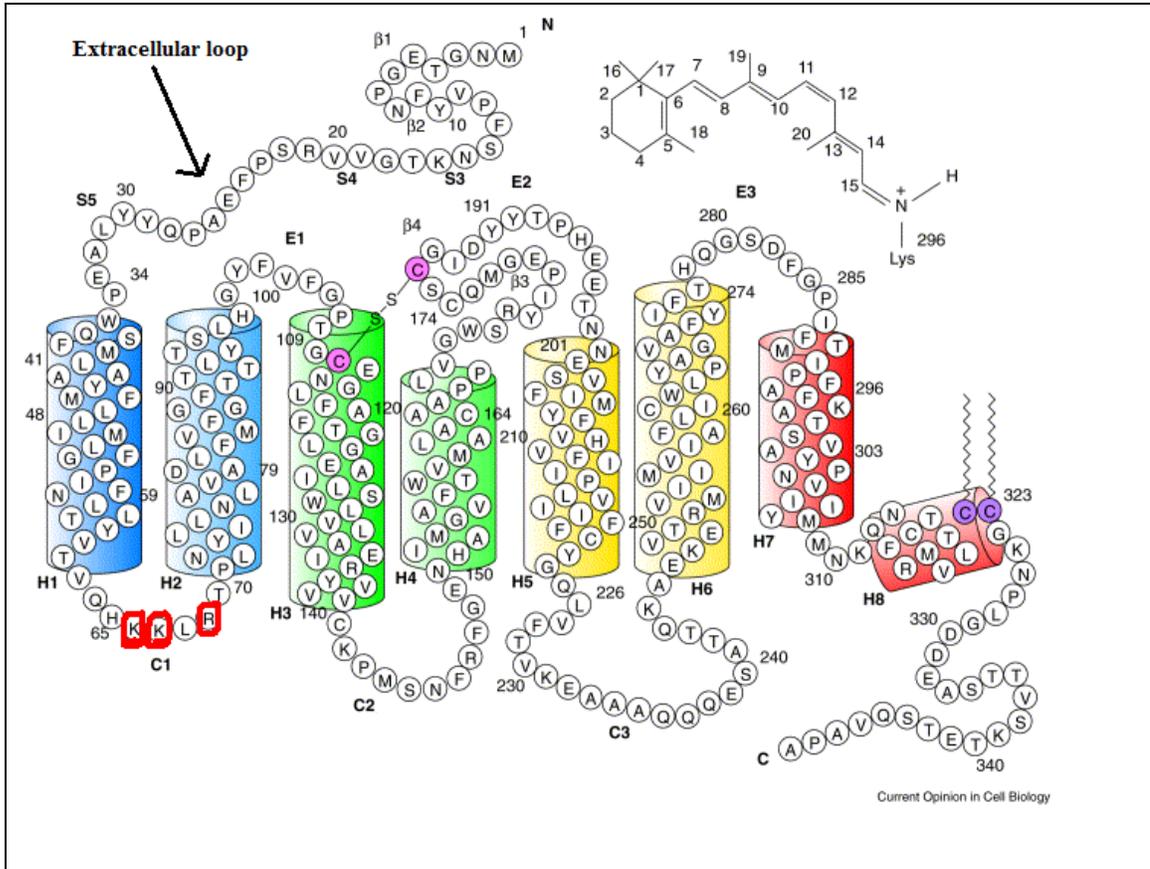


Fig. 4.2 Secondary structure of bovine rhodopsin: This figure is based on Sakmar T. P. [10], taken with permission. The 33 amino acids at N-terminal region on the extracellular side, labeled as extracellular loop, were replaced by N-terminal 27 amino acids of pR. The positively charged amino acids in the intracellular loop 1 (C1) are marked in red.

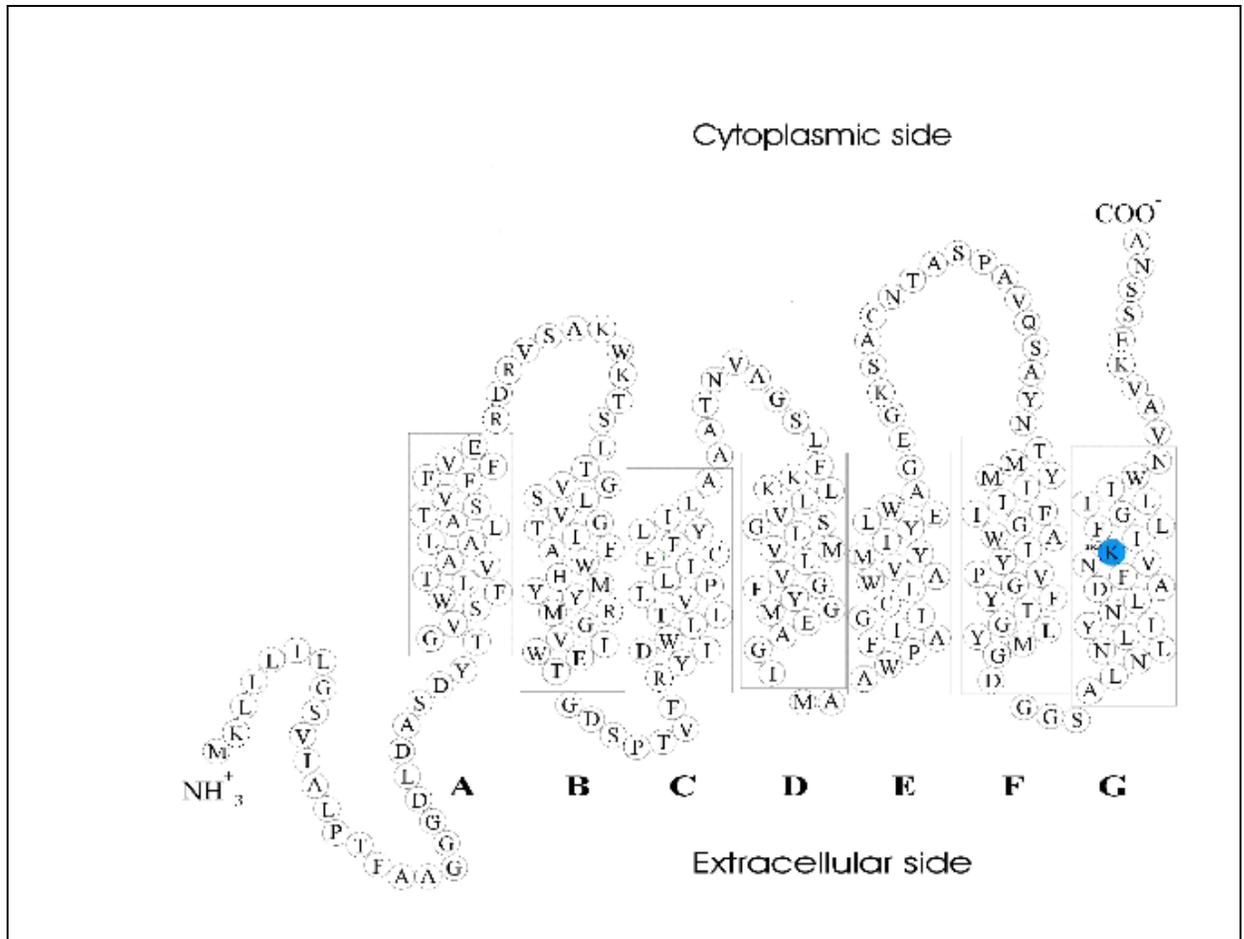


Fig. 4.3: Secondary structure of pR: The first 27 amino acids at the N-terminus of pR on the extracellular side were used to replace the first 33 amino acids of N-terminal extracellular loop of bovine rhodopsin.

The mutagenesis was accomplished by using a polymerase chain reaction (PCR)-mediated method [9]. This method utilizes overlap extension methodology. The overall reaction is divided into two steps: the first step involves synthesis of a double-stranded linear DNA insert containing the desired new sequence, with sequences at both ends that overlap the existing target gene; the second step uses the double-stranded insert as a pair of primers for thermal-cycling DNA synthesis (as in the polymerase chain reaction). The overlapping sequences at the ends of the synthetic DNA are identical to portions of the template for the DNA synthesis in the second step (Fig. 4.6). Thus, the insert becomes

incorporated into the cloning vector in the second step. The PCR-mediated cloning method uses a polymerase known as Phusion®, a high-fidelity DNA polymerase that can potentially insert up to 6.7 kb of insert [9].

A set of primers were designed for the purpose of generating a chimera of bovine rhodopsin with the N-terminus of pR. The sequences of the primers are shown in Fig. 4.7. Schematic diagram of the strategy of primer design is shown in Fig. 4.8. Both the primers included a portion of plasmid pBAD TOPO TA on left side of N-T pR and, a portion of bovine rhodopsin gene on the right side of N-T pR. An extra fragment, known as overhang, was also included in the primers to ease the process of annealing in the second step; the overhang on the 5' side of pR matched a portion of the pBAD-TOPO commercial plasmid sequence, and the overhang on the 3' side corresponded to a portion of bovine rhodopsin sequence.

The overall methodology is summarized in Fig. 4.6. For the first step, the N-terminal portion of pR with overhangs on both its sides was synthesized commercially (Integrated DNA Technology). An approximate portion of pR that was synthesized is shown in Fig. 4.6a. The synthesized linear fragment is shown in Fig. 4.6b. It has the aforementioned overlapping regions on both sides of pR. This linear fragment served as the insert that primed the second step of PCR (Table 4.3 and Fig. 4.6c). Plasmid pBAD-TOPO with bovine rhodopsin (cloning vector) served as the template; thus the overhang and the overlapping sequences of the insert are complementary to segments of this template. The pR sequence that is a part of the primer automatically becomes inserted.

Since the N-terminal portion of bovine rhodopsin is not a part of the insert, it is replaced by pR. The final product obtained is the recombinant plasmid pBAD TOPO TA with N.T. pR-bovine rhodopsin chimera (Fig. 4.6d).

4.2.2 Transfection of supercompetent *E. coli* with the chimera plasmid

pBAD TOPO TA plasmid containing the chimera was digested with DpnI to eliminate the template plasmid. It was then transfected into supercompetent *E. coli* cells (Agilent technologies). Plasmid DNA was isolated for analysis on 0.8% agarose gel as well as DNA sequencing.

4.2.3 Transfection of *E. coli* UT5600 with the chimera plasmid

Plasmid containing chimera DNA was transfected into UT5600 cells for the purpose of protein expression.

Fig. 4.4a: Sequence of Bovine rhodopsin gene (Nucleotide accession number: AH001149): The red region is the sequence included in the primers. The region, 99 bp, 5' side to the red region (*italics*) is the sequence on the N-terminus that forms the first extracellular loop, which was replaced by N-terminus of pR.

```
ATGAACGGGACCGAGGGCCCAAACCTTCTACGTGCCTTTCTCCAACAAGACGGGCGTGGTGCGCAGCCCCT
TCGAGGCCCCGCAGTACTACCTGGCGGAGCCATGGCAGTTCTCCATGCTGGCCGCCTACATGTTCTCTGCT
GATCATGCTTGGCTTCCCATCAACTTCCTCACGCTGTACGTACAGTCCAGCACAGAAGCTGCGCACA
CCCCCTCAACTACATCCTGCTCAACCTGGCCGTGGCCGACCTCTTCATGGTCTTCGGGGGCTTCACCACCA
CCCTCTACACCTCTCTGCACGGGTACTTCGTCTTTGGGCCCACGGGCTGCAACCTGGAGGGCTTCTTTGC
CACCTTGGGCGGTGAAATTGCACTGTGGTCTTGGTGGTCCCTGGCCATCGAGCGGTACGTGGTGGTGTGC
AAGCCCATGAGCAACTTCCGCTTCGGGGAGAACCACGCCATCATGGGCGTCGCCTTCACCTGGGTTCATGG
CTCTGGCCTGTGCCGCGCCCCCTCGTCCGCTGGTCCAGGTACATCCCGGAGGGCATGCAGTGTCTCGTG
CGGGATTGACTACTACACGCCCCACGAGGAGACCAACAATGAGTCGTTTCGTTCATCTACATGTTTCGTGGTC
CACTTCATCATCCCCCTGATTGTTCATATTTCTTCTGCTACGGGCAGCTGGTGTTCACCGTCAAGGAGGCGG
CTGCCCAGCAGAGGTGCGCCACCACTCAGAAGGCCGAGAAGGAGGTACCCGCATGGTGATCATCAT
GGTCATCGCTTTCTAATCTGCTGGCTGCCCTACGCTGGGGTGGCGTTCTACATCTTCACCCATCAGGGC
TCTGACTTTGGCCCCATCTTCATGACCATCCCGGCTTTCTTTGCCAAGACTTCTGCCGTCTACAACCCCG
TCATCTACATCATGATGAACAAGCAGTTCGGAACTGCATGGTCAACACTCTCTGCTGTGGCAAGAACCC
GCTGGGTGACGACGAGGCCTCCACCACCGTCTCCAAGACAGAGACCAGCCAAGTGGCGCCTGCCTAA
```

Fig. 4.4b: Sequence of pR (Nucleotide accession number: AF279106): Note: The sequence provided with this accession number is the reverse complement of the following sequence. Here, the underlined sequence represents the sequence used for the primers. It is the portion of pR, the N-terminus pR, that replaced the N-terminus of bovine rhodopsin.

ATGAAATTATTACTGATATTAGGTAGTGTATTGCACTTCCTACATTTGCTGCAGGTGGTGGTGACCTTGA
TGCTAGTGATTACACTGGTGTTCCTTTTTGGTTAGTTACTGCTGCTTTATTAGCATCTACTGTATTTTTCT
 TTGTTGAAAGAGATAGAGTTTCTGCAAAATGGAAAACATCATTAAGTATCTGGTCTTGTTACTGGTATT
 GCTTTCTGGCATTACATGTACATGAGAGGGGTATGGATTGAACTGGTGATTTCGCCAACTGTATTTAGATA
 CATTGATTGGTTACTAACAGTTCCTCTATTAATATGTGAATTCTACTTAATTCTTGCTGCTGCAACTAATG
 TTGCTGGATCATTATTTAAGAAATTACTAGTTGGTTCCTTGTATGCTTGTGTTTTGGTTACATGGGTGAA
 GCAGGAATCATGGCTGCATGGCCTGCATTATTGGGTGTTTAGCTTGGGTATACATGATTTATGAATT
 ATGGGCTGGAGAAGGAAAATCTGCATGTAATACTGCAAGTCTGCTGTGCAATCAGCTTACAACACAATGA
 TGTATATTATCATCTTTGGTTGGGCGATTTATCCTGTAGGTTATTTACAGGTTACCTGATGGGTGACGGT
 GGATCAGCTCTTAACTTAAACCTTATCTATAACCTTGCTGACTTTGTTAACAAGATTCTATTTGGTTTAAAT
 TATATGGAATGTTGCTGTTAAAGAATCTTCTAATGCTTAA

Fig. 4.5a: Sequence of bovine rhodopsin protein (Protein accession number: AAA30674): The N-terminal 33 amino acids of bovine represented in bold italics form the first extracellular loop. They replaced by N-terminus of pR in the chimera plasmid construction.

MNGTEGPNFYVPFSNKTGVVRSFPFEAPQYYLAEPWQFSMLAAYMFLIMLGFPINFLTLYVTVQHKKLRT
 PLNYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGGFFATLGGEIALWLVVLAIERVYVVC
 KPMSNFRFGENHAIMGVAFTWVMALACAAPPLVGWSRYIPEGMQCSCGIDYYTPHEETNNE SFVIYMFVV
 HFIIPLIVIFFCYGQLVFTVKEAAAQQESATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQG
 SDFGPIFMTIPAFFAKTSAVYNPVIYIMMNKQFRNCMVTTLCCGKNPLGDDEASTTVSKTETSQVAPA

Fig. 4.5b) Sequence of pR protein (Protein accession number: AAG10475): The N-terminal 27 amino acids of pR represented in bold italics form the extracellular loop. They replaced the N-terminus of bovine rhodopsin in the chimera plasmid construction.

MKLLLLILGSVIALPTFAAGGDLDASDYGVSFWLVTAALLASTVFFFVERDRVS AKWKTS LTVSGLVTG
 IAFWHYMYMRGVWIE TGDSP TVFRYIDWLLTVPLLI CEFYLILAAATNVAGSLFKLLVGLVMLVFGYM
 GEAGIMAAWPAFIIIGCLAWVYMIYELWAGEGKSACNTASPAVQSA YNTMMYIIIFGWAIYPVGYFTGYLM
 GDGGSALNLIYNLAD FVNKILFGLI IWNVAVKESSNA

Fig. 4.6: PCR-cloning method: First step is shown in (a) which involves synthesizing a linear insert. The primers used for this step include a sequence of pR to be synthesized flanked by overlapping sequences as described in the text. (b) is the product of the first step that serves as an insert for the second step. NOTE: The first step was omitted, and the linear fragment was directly obtained as a commercially synthesized product (Integrated DNA Technology). (c) Second step uses cloning vector as the template into which the insert (linear product) becomes inserted. (d) shows the final product with N-T pR-bovine rhodopsin chimera.

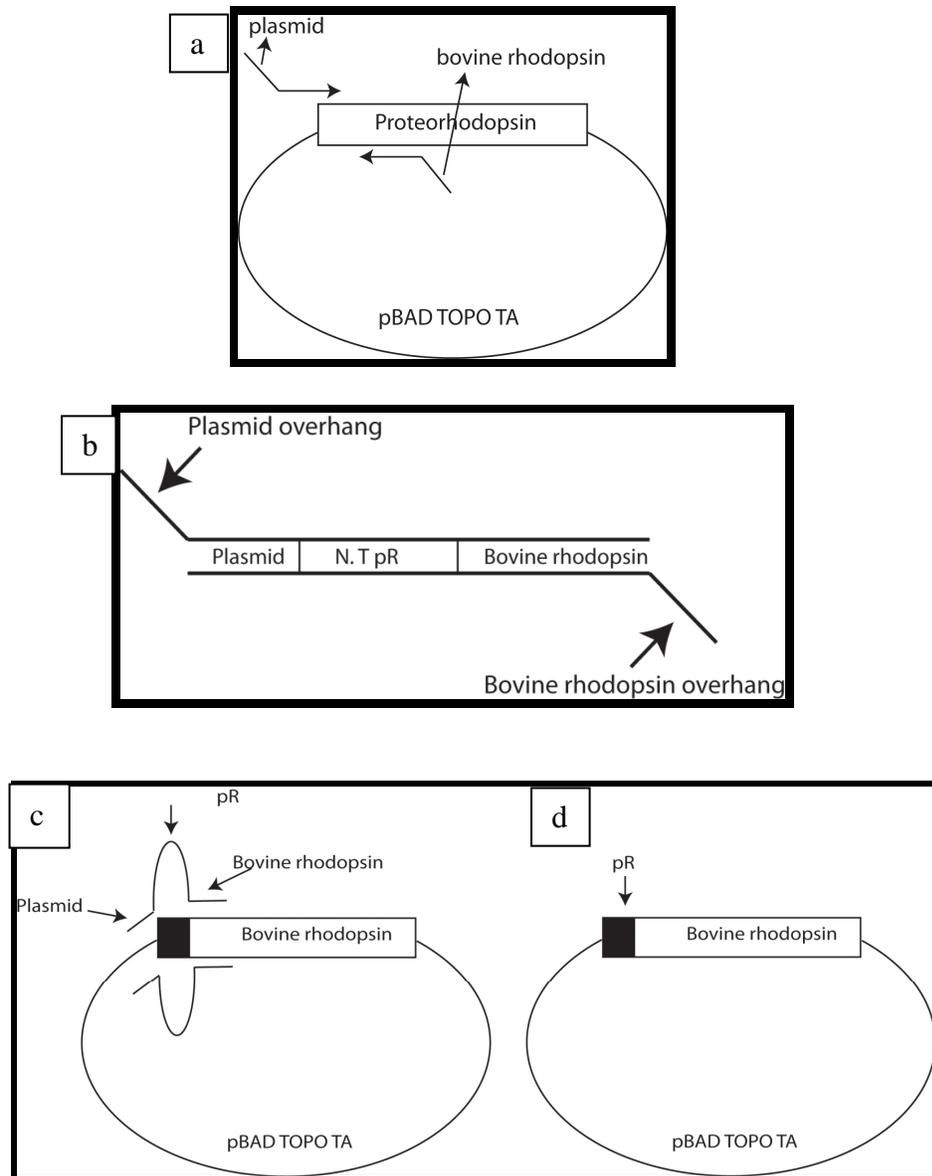


Fig. 4.7: Sequences of primers, obtained commercially, used for construction of chimera plasmid N-T pR-bovine rhodopsin

a) Forward primer: Fwd/pBAD_overhang-pR-BovR: Blue portion represents the sequence of pBAD TOPO TA plasmid, black underlined sequence is the N-T of pR, red portion represents that of bovine rhodopsin sequence beginning at amino acid # 34. The first 19 bases on the 5' side (blue bold part) is the sequence of pBAD TOPO TA that serves as an overhang, i.e., it does not have a complementary sequence in the reverse primer.

5'**TCCATACCCGTTTTTGGGCTAGAAATAATTTTGTTTAACTTTAAGAGGAGA**
TATACATACCCATGAAATTACTGATATTAGGTAGTGTTATTGCACTTCCT
ACATTTGCTGCAGGTGGTGGTGACCTTGATGCTAGTGAT**CCATGGCAGTTCTC**
CATGCTGGCCGCCTACATGTTCTCTGCTGATC-3'

b) Reverse primer: Rev/BovR_overhang-pR-pBAD: The color coding is same as the Forward primer. Similar to the Forward primer, there is an overhang on the 5' side of reverse primer. However, this 19 bases overhang is the sequence of bovine rhodopsin that is in continuation with the sequence used in the Fwd primer. It does not have a complementary sequence in the Fwd primer.

5'**TGATGGGGAAGCCAAGCATGATCAGCAGGAACATGTAGGCGGCCAGCAT**
GGAGAACTGCCATGGATCACTAGCATCAAGGTCACCACCACCTGCAGCAAAT
GTAGGAAGTGCAATAACACTACCTAATATCAGTAATAATTTTCATGGGTATGT
ATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAG-3'

c) The following is a reverse sequence of the Reverse primer. The 3' to 5' sequence of the above Rev primer to show where it complements the Fwd primer. Note: the bold portions do not have a complementary sequence.

3'GATCTTTATTAAAACAAATTGAAATTCTTCCTCTATATGTATGGGTACTTTA
ATAATGACTATAATCCATCACAATAACGTGAAGGATGTAAACGACGTCCACC
ACCACTGGAAGTACGATCACTAG**GGTACCGTCAAGAGGTACGACCGGCGGATG**
TACAAGGACGACTAGTACGAACCGAAGGGGTAGT-5'

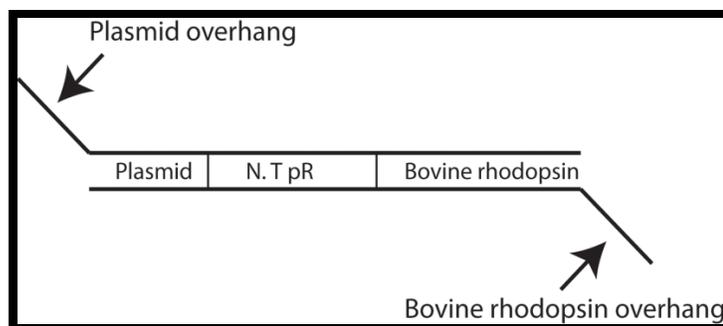


Fig. 4.8: Schematic diagram for primer design: This represents the strategy of primer design used for the construction of chimera plasmid, as described in the text. Total number of bases in each primer is 190.

Table 4.3: PCR parameters

Thermo-cycling reaction for Step 2

Step	Temperature, °C	Time
1	98	30 sec
2	98	10 min
3	72	2.5 min
4	35 × 4	
5	72	10 min
6	4	hold
	End	

4.3 Results

4.3.1 Analysis of the plasmid construct of pR-Bovine rhodopsin chimera

A chimera of N-T pR-Bovine rhodopsin was cloned into the plasmid pBAD TOPO TA. The expected size of the chimera plasmid is 5128 bp. This plasmid DNA isolated from *E. coli* was electrophoresed on 0.8% agarose gel (Fig. 4.11). Chimera plasmid shows at least two forms, one at an apparent size of 3 kb (supercoiled) and

another at an apparent size of about 7 kb (circular). A control plasmid pBAD TOPO TA containing bovine rhodopsin was also run on the gel. The size of this control plasmid is 5202 bp, which is very close in size to the chimera plasmid. Therefore, to ascertain the identity of chimera plasmid, chimera was analyzed by two methods.

The first one was a restriction digestion analysis. Two restriction enzyme Nde1 sites are present in the original control plasmid pBAD TOPO TA-bovine rhodopsin; one of them is located at the N-terminal sequence of bovine rhodopsin. This site is lost in the chimera plasmid, since that sequence is replaced by N-terminal sequence of pR. When the control and the chimera plasmids were subjected to digestion with Nde1, two bands for the control plasmid and only one band for the chimera plasmid were expected. Figs. 4.9 and 4.10 describe the position of Nde1 cuts and the expected band sizes. Fig. 4.11 shows the Nde1 restriction digestion analysis. The expected sizes of bands were obtained; two bands of sizes 3544 bp and 1658 bp for bovine rhodopsin and one band corresponding to size 5128 bp for chimera.

The second, more definitive method of analysis was DNA sequencing. The sequence result of the pR-bovine opsin chimera is shown in Fig. 4.12. These results show that the N-T pR was cloned and replaced the N-T of bovine rhodopsin gene in the chimera plasmid.

Fig. 4.9a: pBAD-TOPO TA-bovine rhodopsin showing Nde1 restriction sites:

The following sequence shows plasmid pBAD TOPO TA containing Bovine rhodopsin gene (red). The purple color represents Nde1 restriction recognition sites (CATATG). There are 2 sites in this sequence. The total size of this plasmid is 5202 bp. Nde1 cuts at positions 399 and 3943. Therefore two bands are obtained upon Nde1 digestion corresponding to 3544 bp and 1658 bp.

```
AAGAAACCAATTGTCATATTGCATCAGACATTGCCGTCACTGCGTCTTTACTGGCTCTTCTCGCTAACCAAACCGGTA
ACCCCGCTTATTAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGTCTATAATCA
CGGCAGAAAAGTCCACATGATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTATCCATAAGATTAGCGG
ATCCTACCTGACCTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTGGGCTAGAAAATAATTTGTTTAA
CTTTAAGAAGGAGATATACATACCCATGGGCTCGGATCCGGTATGACGATGACAAGCTCGCCCTtagaggatcccat
ATGAACGGGACCGAGGGCCAACTTCTACGTGCTTTCTCCAACAAGACGGGCGTGGTGCGCAGCCCTTCGAGGCC
GCAGTACTACCTGGCGGAGCCATGGCAGTTCTCCATGCTGGCCGCTACATGTTCTGCTGATCATGCTTGGCTTCCCA
TCAACTTCTCACGCTGTACGTACAGTCCAGTCCAGCACAAGAAGCTGCGCACACCCCTCAACTACATCTCCTCAAC
GTGGCCGACCTCTTCATGGTCTTCGGGGCTTACCACCACCTCTACACCTCTGACGGGTACTTCTGCTTTGGCC
CAGCGGGTGCAACTGGAGGGCTTCTTGGCACCTTGGCGGTGAAATGCACTGTGGTCTTGGTGGTCTGGCCATCG
AGCGGTACGTGGTGGTGTGCAAGCCATGAGCAACTCCGCTTCCGGGAGAACCACGCCATCATGGGCGTCCGCTTACC
TGGGTCATGGCTTGGCCGTGTCGGCGCCCTTCCGCTGGTCCAGGTACATCCCGAGGGCATGCAAGTCTCGTG
CGGGATTGACTACTACGCCCCAGGAGACCAACAATGAGTCTGTTCTGTCATCTACATGTTCTGGTCCACTTCA
TCCCCCTGATTGTCATATTTCTGCTACGGGCGAGTGGTGTCCACCGTCAAGGAGGGCGGCTGCCAGCAGAGGAGTCC
GCCACCACTCAGAAGGCCGAGAAGGAGTCAACCCGATGGTATCATATGGTTCATCGCTTCTTAATCTGCTGGTGGC
CTACGCTGGGGTGGCGTTTCACTCTTCCACCATCAGGCTCTGACTTGGCCCCATCTTCATGACCACTCCCGGCTT
TTGCCAAGACTTCTGCGCTTACAACCCCGTCTACATCATGATGAACAAGCAGTTCGGGAAGTGCATGGTCAACCT
CTCTGCTGGCAAGAACCCGCTGGGTGACGACGAGGGCTCCACCACCGTCTCCAAGACAGAGACCAGCCAAGTGGCGCC
TGCCATAcgataagcttaattAAGGGCGAGCTTGAAGGTAAGCCTATCCCTAACCTCTCCTCGGTCTCGATTCTACGC
GTACCCGTCATCATCACCATCACCAATTGAGTTTAAACGGTCTCCAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTCA
GCCTGATACAGATTAATCAGAACGCAGAAGCGGTCTGATAAAAAGAAATTTGCTGGCGGAGTAGCGGGTGGTCCCA
CCTGACCCCATGCGCAACTCAGAAGTGAACGCGGTAGCGCCGATGGTAGTGTGGGGTCTCCCATGCGAGAGTAGGGAA
CTGCCAGGCATCAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCCTTTCGTTTTATCTGTTGTTGTGCGGTGAACGCT
CTCCTGATAGGACAAAATCCGCGGGAGCGGATTTGAACGTTGCGAAGCAACGCGCCCGGAGGGTGGCGGGCAGGACGCC
GCCATAAACTGCCAGGCATCAAAATTAAGCAGAAGGCCATCTGACGGATGGCCTTTTTGCGTTTTTACAAAACCTTTTTTG
TTTTTTTTCTAAATACATTTCAATATGATATCCGCTCATGAGACAATAACCCGTGATAAATGCTTCAATAATATGAAAA
GGAAGAGTATGATATTAACATTTCCGTGTCGCCCTTATCCCTTTTTGCGGCATTTTGCCTTCTGTTTTTGTCTCAC
CCAGAAACGCTGTTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCAACGAGTGGTTACATCGAATCGAATCAGAG
CGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAAGCTTTTCCAATGATGAGCACTTTTAAAGTCTGCTATGTGGCGGG
TATATCCCGTGTGACGCCGGCAAGAGCACTCGGTGCGCCGATACACTATCTCAGAATGACTTGGTGGTACTTCA
CCAGTACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATATGCACTGCTGCCATAACCATGAGTGATAACAC
TGCGGCCAATTTACTCTGACAAAGATCGGAGGACCGAAGAGCTAACCCGTTTTTTGACACAACATGGGGGATCATGTAA
CTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCAGATGCCTGTAGCAATG
GCAACAACGTTGCGCAAACTATTAACCTGGCGAACTACTTACTTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGG
GGATAAAGTTGCAAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGT
GTGGGCTCTGCGGTTCTATTGACACTTGGGGCAGATGGTAACCCCTCCCGTATCGTAGTTATCTACACGACCGGGAGT
CAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTGACACCA
AGTTTACTCATATATACTTTAGATTGATTTAAAACCTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATA
ATCTCATGACAAAATCCCTTAACTGAGTTTTCTGTTCCACTGAGCGTCAACCCCGTAGAAAAGATCAAAGGATCTTCT
TGAGACTTTTTTCTGCGCGTAAATCTGCTGTTGCAAAACAAAAAACCCGCTACCAGCGGTGGTTTTGTTGCTGGT
TCAAGAGTACCAACTCTTTTTCCGAAGGTAAGTGGCTTACAGAGAGCGCAGATACCAAACTGTCTTCTAGTGTAGC
CGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGTAATCCTGTTACCAGTGGCTGCT
GCCAGTGGCGATAAGTCTGTCTTACCAGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGGTGAAC
GGGGGTTCTGTCACACAGCCAGCTTGGAGCGAAGCAGCTACACCGAACTGAGATACCTACAGCTGAGGATGAGAAA
GCGCCACGCTTCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCCGAACAGGAGAGCGCACGAGGGAG
CTTCCAGGGGAAACGCTGATCTTTATAGTCTGTGGGTTTCCGCCACTCTGACTTGGAGCGTCAATTTTGTGATG
CTCGTCAAGGGGGCGGACCTATGGA AAAACGCGCAGCAACGCGGCCCTTTTTACGGTTCTTGGCCTTTTTGCTGGCCTTTTG
CTCACATGTTCTTTCGCTTATCCCTGATTTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGC
CGCAGCCGAAACGACGCGCAGCTCAGTGTGCGGAGGAAGCGGAAGAGCGCCTGATGCGGATTTTTCTCCTACGCA
TCTGTGCGGTATTTACACCGCATATGTCAGTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACAC
TCCGCTATCGCTACGTGACTGGTCAATGGCTGCGCCCCGACACCCGCCAACACCCCGTGCAGCGCCCTGCAGGGCTTGT
TGCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCGGGAGCTGCATGTGTCAGAGTTTTTACCAGTATACCCG
AAACGCGGAGGACGAGATCAATTCGCGCGAAGGCGAAGCGGCATGCATAATGTGCTTGTCAAATGGAAGCAAGGAG
GATTCTGCAAACCTATGCTACTCCGTCAGCCGTCATTTGTCTGATTCTGTTACCAATATGACAACCTGACGGCTACAT
CATTCACTTTTTCTTCAACCGGCACGGAACCTGCTCGGGTGGCCCCGGTGCATTTTTTAAATACCCGCGAAGAAATAG
AGTTGATCTGCAAAACCAATTGCGACCGAGGTCGATAGGCATCCGGTGGTGTCTCAAAGCAGCTTCCGCTGGCT
GATACGTTGGCTTCCGCGCAGCTTAAGACGCTAATCCCTAAGTCTGTCGGCGAAAAGATGTGACAGACGCGCAGCACA
AGCAAACATGCTGTGCGACGCTGGCGATATCAAATTTGCTGTCTGCCAGGTGATCGCTGATGACTGACAAGCCTCGCGT
ACCCGATATCCATCGGTGGATGGAGCGACTCGTTAATCGCTTCCATGCGCCGAGTAACAATGCTCAAGCAGATTTAT
CGCCAGCAGCTCCGAATAGCGCCCTTCCCTTGGCCGGCGTTAATGATTTGCCAAACAGGTCTGAAATGCGGCTGGT
```

GCGCTTCATCCGGGCGAAAGAACCCCGTATTGGCAAATATTGACGGCCAGTTAAGCCATTCATGCCAGTAGGCGCGCGGA
CGAAAGTAAACCCACTGGTGATACCAATTCGCGGAGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGGCGGGAACAG
CAAAATATCACCCGGTCGGCAAACAAATTCGTCCCTGATTTTTCACCACCCCTGACCGCGAATGGTGAGATTGAGAA
TATAACCTTTCATTCACGCGGTCGGTCGATAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTTAAACCCGCCACC
AGATGGGCATTAACGAGTATCCGGCAGCAGGGGATCATTTTTCGCTTCAGCCATACTTTTCATACTCCCGCCATTCAG
AG

Fig. 4.9b: Schematic representation of plasmid pBAD TOPO TA – bovine rhodopsin digestion with NdeI

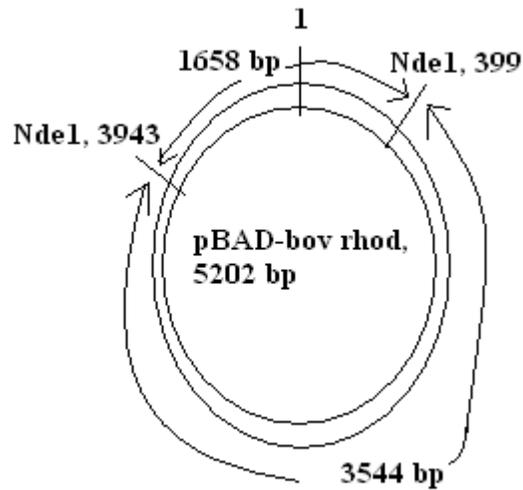


Fig. 4.10a: Chimera plasmid showing Nde1 restriction sites: The following is a manually generated sequence of N-T pR-bovine rhodopsin chimera in order to show the Nde1 restriction site. The same color coding is followed. The purple color represents Nde1 restriction recognition sites (CATATG). There is only one site in this sequence at position 3869. The total size of the chimera plasmid is 5128 bp. Upon Nde1 digestion, only one band corresponding to 5128 bp is obtained.

AAGAAACCAATTGTCATATTGCATCAGACATTGCCGTCACTGCGTCTTTACTGGCTCTTCTCGCTAACCAAACCGGTA
 ACCCCGCTTATTTAAAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGTCTATAATCA
 CGGCAGAAAAGTCCACATGATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGG
 ATCCTACTGACCGCTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTGGGCTAGAAAATAATTTTGTAAAC
 TTTAAGAAGGAGATATACATACCCATGAAATTATTACTGATATTAGGTAGTGTATTGCACCTCCTACATTTGCTGCAGG
 TGGTGGTGACCTTGATGCTAGTGATCCATGGCAGTTCTCCATGCTGGCCGCTACATGTTCTCTGCTGATCATGCTTGGCT
 TCCCCATCACTTCTCAGCTGTACGTACAGTCCAGCACAAAGAGTGCACACCCCTCAACTACATCTGCTCAAC
 CTGGCCGTGGCCGACCTCTTCATGGTCTTCGGGGGCTTACCACCACCTCTACACCTCTCTGCACGGGTACTTCGTCTT
 TGGGCCACCGGCTGCAACCTGGAGGGCTTCTTTGCCACTTGGGCGGTGAAATTGCACCTGTGGTCTTGGTGGTCTGG
 CCATCGAGCGGTACGTGGTGGTGTGCAAGCCATGAGCAACTTCCGCTTCGGGGAGAACACGCCATCATGGGCGTCCGC
 TTCACCTGGGTGATGGCTCTGGCCTGTGCCGCCCCCCCCCTCGTGGTGGTCCAGGTACATCCGGAGGGCATGCAGTG
 CTGGTCGGGATGACACTACACGCCACGAGAGACCAACAATGAGTCGTTCGTCTACATGTTCTGGTCCACT
 TCATCATCCCCCTGATTGTCATATTCTTCTGCTACGGGCGAGTGGTGTTCACCGTCAAGGAGGCGGCTGCCAGCAGCAG
 GAGTCGGCCACCACTCAGAAGGCCGAGAGGAGGTCAACCCGATGGTGTATCATGATGATGATGAAACAAGCAGTTTCCGA
 GCTGCCCTACGCTGGGGTGGGCTTCTACATCTTCCACCATCAGGGCTCTGACTTTGGCCCCATCTTCCATGACCATCCCGG
 CTTTCTTTTGGCAAGACTTCTGCCGTCTACAACCCCGTCTCATCATGATGATGAAACAAGCAGTTTCCGAAGTCTGCTG
 ACCACTCTCTGCTGTGGCAAGAACCCGCTGGGTGACGACGAGGCTCCACCACCGTCTCAAGACAGAGACCAGCCAAGT
 GGGCCTGCCTAATcgataagcttaattAAGGGCGAGCTTGAAGGTAAAGCTTATCCCTAACCTCTCCTCGGTCTCGATT
 CTACGCGTACCGGTATCATCACCATCACCATTGAGTTTAAACGGTCTCCAGCTTGGTGTGTTTTGGCGGATGAGAGAAGA
 TTTTACGGGTGATACAGATTAATCAGAACCGAGAGAGGCTTGATAAAAACAGAATTTGCCTGGCGGAGTAGCGCGGTG
 GTCCACCTGACCCCATGCCAAGTCAAAGTCAAAGCAGGCTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGT
 AGGGAACCTGCACGGCATCAAATAAACGAAAGGCTCAGTCAAAGACTGGGCCCTTCGTTTTATCTGTTGTTGTCCGGTG
 AACGCTCTCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGG
 ACGCCCGCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCTGACGGATGGCCTTTTTGCGTTCCTACAAACTC
 TTTTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCTGATAAATGCTTCAATAATATT
 GAAAAAGGAAGATATGAGTATCAACATTTCCGCTGTCCGCTTATTCCCTTTTTTGGCGCATTTTGCTTCTCCTGTTTTT
 GCTCACCCAGAAACGCTGGTGAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCT
 CAACACGGTGAATCCTTTGAGAGTTTTTCGCCCCGAAGGAGCTTTTCCAAATGATGAGCACTTTTAAAGTCTGCTGATG
 GCGCGGTATTATCCCGTGTGACGCGGGCAAGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGTTGAG
 TACTACCACTGACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGTGCCATAACCATGAGTGA
 TAACACTGCGGCCAAGTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATC
 ATGTAACCTGCACTTACGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAACACCATGATGCTGTC
 GCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCAACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGAT
 GGAGGCGGATAAAGTTGACAGGACCACTTCTGCGCTCGGCCCTCCGGCTGGCTGGTTTATTGTGATAAATCTGGAGCCG
 GTGAGCGTGGGTCTCGCGGATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACG
 GGGATCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTGCTCACTGATTAAGCATTTGGTAACTGTC
 AGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTCATTTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTT
 TTGATAATCTCATGACCAAAAATCCCTTAAACGTGAGTTTTCTGTTCCACTGAGCGTCAAGCCGTTAGAAAAGATCAAAGGA
 TCTTCTTGAGATCCTTTTTTTCTGCGGTAATCTGCTGCTTGCAAAACAAAAAACCCGCTACCAGCGGTGGTGTGTTTT
 GCCGATCAAGACACTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGTCGGAACAGGAGACGACAG
 AGGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGTCTGTCGGGTTTTCGCCACCTCTGACTTGAGCGTGCATTTTT
 GTGATGCTCGTACGGGGGCGGAGCCTATGGAAAACGCCAGCAACGCGGCTTTTTACGGTTCCTGGCCTTTTGTCTGGC
 CTTTTGCTCACAATGTTCTTTCCTGCGTTATCCCTGATTTCTGTGGATAACCGTATTACCGCTTTGAGTGAGCTGATAAC
 CTCGCCGACGCAACGACCGAGCGAGCAGTCACTGAGGAGGAAAGCGGAAAGCGCCTGATCGGTTATTTCTCCT
 TACGCATCTGTGCGGTATTTACACCGCATATGCTACTCTCAGTACAATCTGCTCTGATGCGCATAGTTAAGCCAGT
 ATACACTCCGCTATCGTACGTACTGGGTCACTGGCTGCGCCCGACACCCGCAACACCCGCTGACGCGCCTGACGGG
 CTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGTTTTACCGTCA
 TCACCGAAACGCGGAGGCGAGATCAATTCGCGCGGAGGCGAAGCGGATGCATAATGTGCTGTCAAATGTGACAGCGACG
 AGGAGGATTTCTCAAAACCTTATGCTACTCCGTCGAAGCCGTAATTTGCTGATTTACCAATTAAGCAACTTGACGCG
 CTACATCATTCACTTTTTCTTCAACACCGGCACGGAACCTCGCTCGGGTGGCCCGGTGCATTTTTTAATACCCGCGAG
 AAATAGAGTTGATCGTCAAAACCAACATGCGACCGAGCGGTGGCGATAGGCATCCGGGTGGTCTCAAAGAGCTTCGC
 CTGGCTGATACGTTGGTCTCGCGCCAGCTTAAAGACGCTAATCCCTAAGTGGCTGGCGGAAAAGATGTGACAGACGCGACG
 GCGACAAGCAAACTGTGCGACGCTGGCGGATATCAAAAATTGCTGTCTGCGAGGTGATCGTGTGACAGCAGCCG
 TCGGCTACCCGATTATCCATCGGTGGATGGAGGACTCGTTAATCGCTTCCATGCGCCGAGTAACAATTGCTCAAGCAG
 ATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGGCCGGCGTTAATGATTTGCCAAACAGGTCGCTGAAATGCG
 GCTGGTGGCTTCATCCGGGCAAGAACCCCGTATTGGCAAAATTTAGCGCCAGTTAAGCCATTTCATGCCAGTAGGGC

```
CGCGGACGAAAGTAAACCCACTGGTGATACCATTCGCGAGCCTCCGGATGACGACCGTAGTGATGAATCTCTCTGGCGG
GAACAGCAAAATATCACCCGGTCGGCAAACAATTCTCGTCCCTGATTTTTCCACCACCCCTGACCGCGAATGGTGAGAT
TGAGAATATAACCTTTTCATTCCAGCGGTTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTTAAACCC
GCCACCAGATGGGCATTAAACGAGTATCCCGGCAGCAGGGGATCATTTTTCGCGCTTCAGCCATACTTTTCATACTCCCGCC
ATTCAGAG
```

Fig. 4.10b: Schematic representation of plasmid pBAD TOPO TA N-T pR-bovine rhodopsin chimera digestion with NdeI restriction enzyme

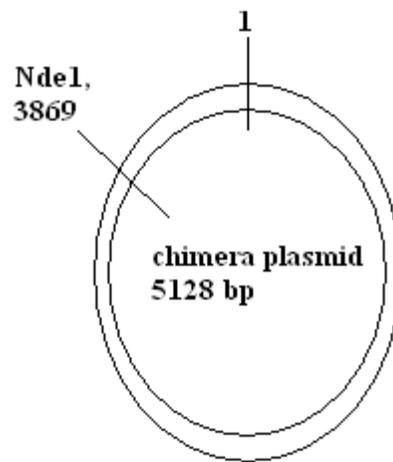


Fig. 4.11: Agarose gel analysis of Nde1 digested plasmids: plasmid pBAD TOPO TA-bovine rhodopsin and pBAD TOPO TA N-T pR-bovine rhodopsin were digested with restriction enzyme Nde1 at 37°C for 1 hour. The digested product was loaded on 0.8% agarose gel and run against undigested control plasmids. Lane 1 contains undigested bovine rhodopsin plasmid; lane 2 contains digested bovine rhodopsin plasmid, lane 3 has digested chimera plasmid and lane 4 shows undigested chimera plasmid. Lane 5 was loaded with 1 kb ladder (NEB).

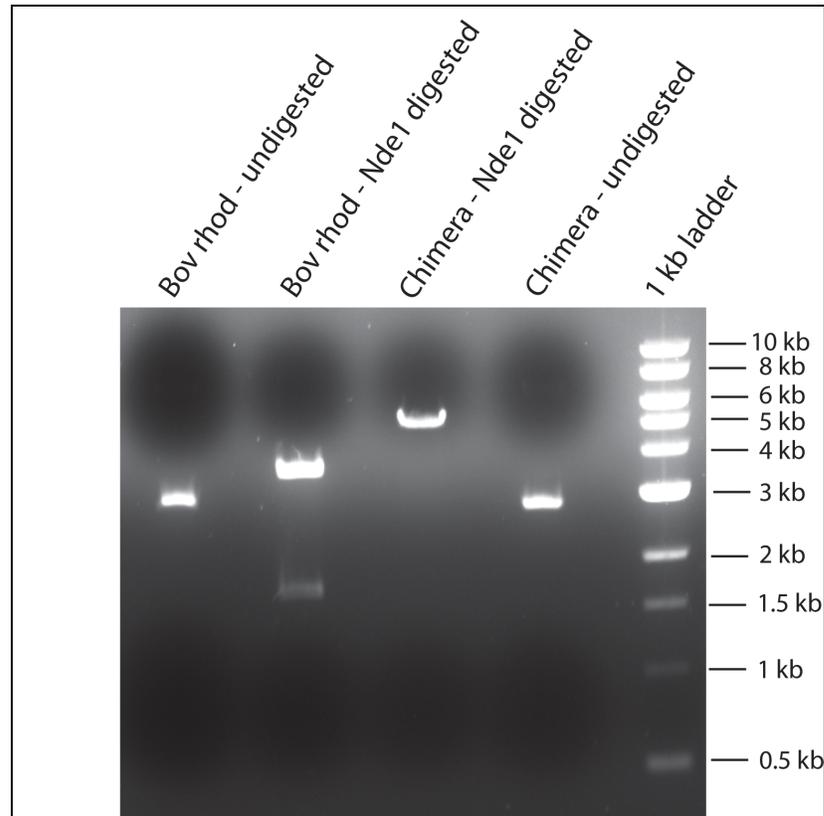


Fig. 4.12: Sequence result of chimera plasmid pBAD TOPO TA N-T pR-Bovine rhodopsin chimera: a) The following is a result obtained from sequencing of the chimera plasmid (SUNY Upstate DNA Sequencing Facility). The underlined portion is the sequence of pR included in the primers, the bold red region corresponds to the sequence of bovine rhodopsin present in the primers.

```
GACGCTTTNATCGAAACTCTCTACTGTTTCTCCATACCCGTTTTTTGGGCTAGAAATAATTTGTTTAACTTTAAGAAG
GAGATATACATACCCATGAAATTATTACTGATATTAGGTAGTGTATTATGCACTTCCTACATTTGCTGCAGGTGGTGGTGA
CCTTGATGCTAGTGATCCATGGCAGTTCTCCATGCTGGCCGCCTACATGTTCCCTGCTGATCATGCTTGGCTTCCCCATCA
ACTTCCTCAGCGTGTACGTACAGTCCAGCACAAGAAGCTGCGCACACCCCTCAACTACATCCTGCTCAACCTGGCCGTG
GCCGACCTCTTCATGGTCTTCGGGGGCTTCACCACCACCCCTCTACACCTCTCTGCACGGATACTTCGTCTTTGGGCCAC
GGGCTGCAACCTGGAGGGCTTCTTTGCCACCCCTGGGCGGTGAAATTGCACTGTGGTCTTGGTGGTCTGGCCATCGAGC
GGTACGTGGTGGTGTGCAAGCCATGAGCAACTTCCGCTTCGGGGAGAACCACGCCATCATGGCGTGCCTTCACCTGG
GTCATGGCTCTGGCCTGTGCCGCGCCCCCTCGTCGGCTGGTCCAGGTACATCCCGAGGGCATGCAGTGCCTGTGCCG
GATTGACTACTANACGCCCCACGAGGAGACCAACAATGAGTCGTTTCGTCATCTACATGTTTCGTGGTCCACTTCATCATCC
CCCTGATTGTCATATTCTTCTGCTACGGGCAGCTGGTGTTCACCGTCAAGGAGGGCGGTGCCAGCAGCAGGAGTCCGGCC
ACCACCTCAGAAGGCCGAGAAGGAGGTCACCCGATGGTGATCATCATGGTCAATCGCTTCCCTAATCTGCTGGCTGCCCT
ACGCTGGGGTGGNGGTTCTACATCTTACCCATCAGGGGCTCTGAANTTTGGNCCCATCTTCATGACCATCCCCGGCTTT
CTTTGCCAAGANTTCTGCCGCTTACAACCCCGTCATCTANATCATGATGAAAAAGCAGTTCNGGAACTGCATGGTACCA
CTCTCTGCTGNGGNAAAAACCCNCTGGGGGANAACAAGNCCCTCCCCACCGTNTCCAAAANAAAAACAGCCAAGNGGG
CCTGCCCTANNGAAAAGCTAAANAAGGNAGNNTGAAGGAANGCTATCCCTNACCNNCTCTCGGTCCCNATTNACCGAAC
GGTANNATCNCATCCNNTTGNNGTTNANNGGCNACNTGTCNTGTTTGGGANAAAANAATTTCCCCGAANNAATAANNNA
CNNAAGGGNTGNAANAATTTGCTGGGGGANANGNGGGGCCCCNTACCCTGCCAANCAAAAAACCCNANCCAAGGGGG
```

4.3.2 Attempt at heterologous expression of chimera protein in *E. coli*

The plasmid pBAD TOPO TA containing pR-Bovine opsin chimera was transfected into *E. coli* UT5600 cells. These cells are protease-deficient which allows for efficient expression of foreign protein. Three transfected colonies were picked, and grown in LB medium supplemented with ampicillin (50 µg/ml). The culture was grown to OD₆₀₀ of about 1. It was then induced with L-arabinose (0.2%) and retinal (a mix of photoisomers of all-trans and 11-cis retinal) was added in order to produce a fully functional protein covalently bound to its chromophore.

The 11-cis retinylidene Schiff base chromophore of bovine rhodopsin absorbs at a λ_{\max} = 500 nm and therefore, cells expressing properly-folded rhodopsin should appear orangish-red. A slightly brown coloration of the cells was seen. An attempt was made to isolate the protein by lysing the cells. A UV-vis spectrum of the extract revealed a strong

absorbance at 280 nm. The characteristic absorbance of retinal bound bovine rhodopsin at 500 nm was not clearly visible.

4.4 Follow-up Experiments

4.4.1 To detect the expression of chimera protein

Expression of chimera protein can be assessed by Western blot using just a fragment of the proteorhodopsin.

4.4.2 To detect the location of chimera protein

Western blot only indicates that the expression of a mammalian rhodopsin, bovine rhodopsin. It does not necessarily demonstrate that the protein is translocated and inserted into the membrane. Several techniques can be utilized to investigate whether the chimera protein is integrated into the membrane.

Differential centrifugation: It is first of all important to know that the protein fractionates with the bacterial inner membrane fraction. To that end, the crude cellular extract will be subjected to a technique known as differential centrifugation that includes use of high speed centrifugal force. For this purpose, the cells expressing chimera protein and bovine rhodopsin (control) will first need to be disrupted. This can be done by various methods such as liquid shear pressure (French press), ultrasonication, osmotic shock, glass bead milling, Freeze-thaw, and enzymatic lysis (lysozyme). Cell disruption yields a suspension of plasma membrane/vesicles/microsomes, cell debris, intact cells, nuclei, soluble proteins and cell organelles. This suspension will be subjected to

differential centrifugation where it is first centrifuged at a lower speed ($10,000 \times g$) to remove the cell debris, nuclei and intact cells. The supernatant will be centrifuged at a higher speed ($20,000 \times g$) to remove other contaminants. The supernatant fraction resulting from this speed will finally be subjected to ultracentrifuge, that is, at very high speeds such as $150,000 \times g$. The pellet obtained would contain the membrane fraction including microsomes (endoplasmic reticulum vesicles). Various fractions of supernatants and pellets will be loaded on SDS-PAGE and the chimera protein (or bovine rhodopsin) will be detected by Western blot. The protein bands would show up in the lower speed supernatant (possibly pellet too due to the unbroken cells), and in the supernatant of the medium speed ($20,000 \times g$). The most intense band would be expected to show up in the high-speed (ultracentrifuge) pellet. In order to ascertain the identity of the membrane fraction, a marker enzyme that is plasma-membrane specific can be detected.

Sucrose gradient: An alternative or additional method to differential centrifugation is density gradient centrifugation. The medium speed pellet from above will be layered on top of a sucrose gradient and centrifuged at high speed. The membrane pellet will settle in the gradient at a point where its density matches that of the sucrose density. The detection of the protein will be done similarly as mentioned above.

GFP- or immunolocalization: In addition to membrane fractionation, cellular localization of the pR can be determined with some direct microscopy methods. For example, green fluorescent protein (GFP) can be tagged at C-terminal end and the cells

can be observed under a fluorescence microscope. Another similar technique known as immunofluorescence can also be employed which utilizes an antibody designed against the protein (or the polyHis tag) that is chemically linked to a fluorophore. The fluorophore can be visualized under a fluorescence microscope.

4.4.3 Purification of chimera protein with citrate

Despite the absence of a significant similarity in the sequence of pR and bovine rhodopsin, a close examination of the first intracellular loop of both the proteins shows that they are somewhat similar. They both contain positively charged amino acid residues, lysine and arginine. Citrate interacts with pR at the positively charged residues found in the first intracellular loop (chapter 2). Therefore, citrate can be used for the purification of chimera protein. For the first trial, same protocol as used for pR purification will be attempted. The concentration and pH can be altered if needed.

Combination of phosphate and citrate will also be used to purify the chimera. The purification process will be monitored by UV-vis spectra as well as SDS-PAGE. The percent yield and purity will be calculated from the λ_{max} absorption.

4.5 Future Experiments

If the follow-up experiments fail to show expression of the chimera protein, some of the following modifications can be attempted, not necessarily in that order.

4.5.1 Modification of expression experiments

Preliminary screening: Several different strains of *E. coli* will be screened to transfect the chimera protein. After growing the best *E. coli* cells containing chimera protein N-T pR-bovine rhodopsin to $OD_{600}=1$, several different temperatures such as 15°C, 22°C, and 37°C, will be tried for inducing the protein. Freshly made solution of 11-*cis* retinal (in acetone) will be added to aid the formation of a holoprotein.

4.5.2 Modification of cell wall lysis

The extraction buffer used for cell lysis will be buffered at several different pH units to keep the protein stable. Several different detergents can also be screened to improve the lysis of the cells and also to solubilize the chimera protein into the detergent.

4.5.3 Alteration in the size of the leader sequence

If the protein is not inserted into the membrane, then the length of N-T pR sequence will be increased to assure the membrane incorporation of the chimera.

4.5.4 Stabilization of membrane proteins by Genetic Engineering

It has been shown that some point mutations, individual or in combination, have a dramatic effect on the thermostability of the membrane proteins. For example, a random mutagenesis on *E. coli* diacylglycerol kinase (DGK) yielded 12 thermostabilising

mutations [11]. Later the same group showed that similar stability was conferred upon bacteriorhodopsin (BR) by using mutations [12]. This method was also applied to the GPCRs such as turkey β_1 -adrenergic receptor (β_1 AR) [13], adenosine A_{2a} receptor (A_{2a} R) [14], and the neurotensin receptor (NTR) [15]. Thermostability was also found to improve the solubility of proteins in the detergents that are otherwise not suitable for the protein. Random mutagenesis will be performed on the chimera protein and a stable temperature will be found. The stability of the protein can be assayed using UV-vis spectrophotometry as bovine rhodopsin shows maximum absorption at 500 nm.

4.6 Conclusions

This study will need further examination to make a definitive conclusion. A higher level of expression, proper native folding as judged by its characteristic absorption maximum, and confirmed cellular localization of the chimera protein will be the primary prerequisites. The purification of the chimera can then be attempted using the well-established protocol using citrate method of purification. Assuming the expression, isolation and purification of chimera follow the steps of pR purification, this study can be extended toward the heterologous expression of pharmaceutically important membrane proteins in *E. coli*. If the N-terminal portion of pR can significantly improve the expression of GPCRs in bacteria such as *E. coli*, then it will not only facilitate in increasing the amount of protein in a shorter time but also help in the expression of proteins that are difficult to express. This will eventually aid in providing large amounts of proteins required for structural studies and thereby, understanding the structure of the

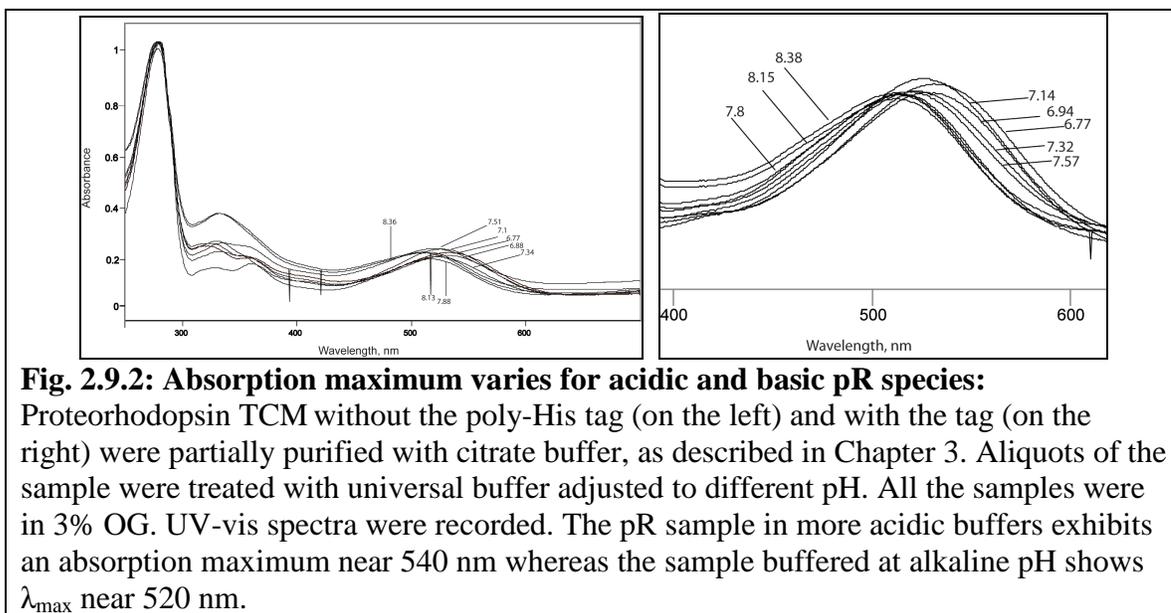
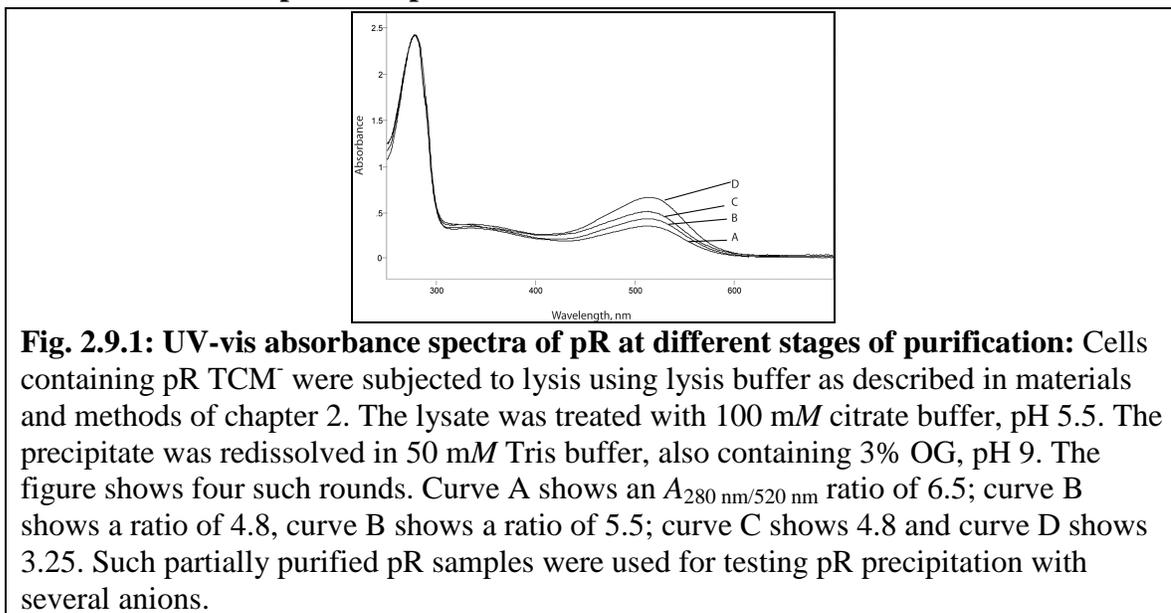
GPCRs and their ligand binding sites. A well-resolved ligand-binding site of a receptor will benefit in designing ligand-site specific drugs.

4.7 Acknowledgements

I thank all the undergraduate researchers who have been involved with this project over the years: Plasmid pBAD TOPO TA containing bovine rhodopsin was constructed in collaboration with Soyika Richardson. Construction of chimera plasmid was done with the help of Chauncy Brown, Alicia Barnes, Farina Mahmud, Euginia Im, Danny Heng, and Tiffany Dunston.

Appendix

2.9 UV-visible spectra of pR



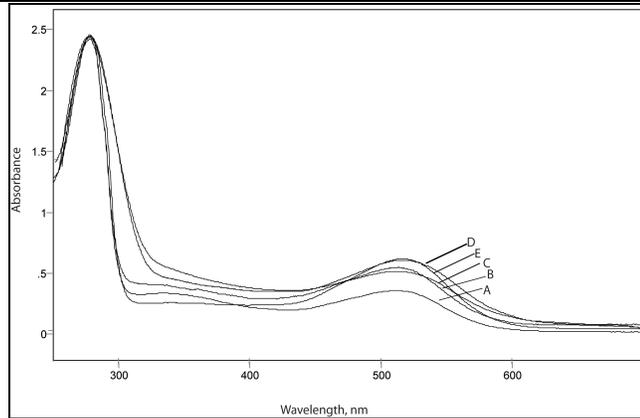


Fig. 2.9.3: UV-vis spectra of pR TCM mutants: Quadruple mutant (R51Q/R53Q/K57Q/K59Q) was purified by loading the cell lysate on the hydroxyapatite column while the other mutants were partially purified using citrate as described in text. Curve A represents double mutant K125Q/K126Q, curve B represents single mutant K172, curve C is that of quadruple mutant, curve D shows spectrum of single mutant R51Q, and curve E shows the spectrum of single mutant K244Q.

2.10 Western Blot of pR expressed in *E. coli*

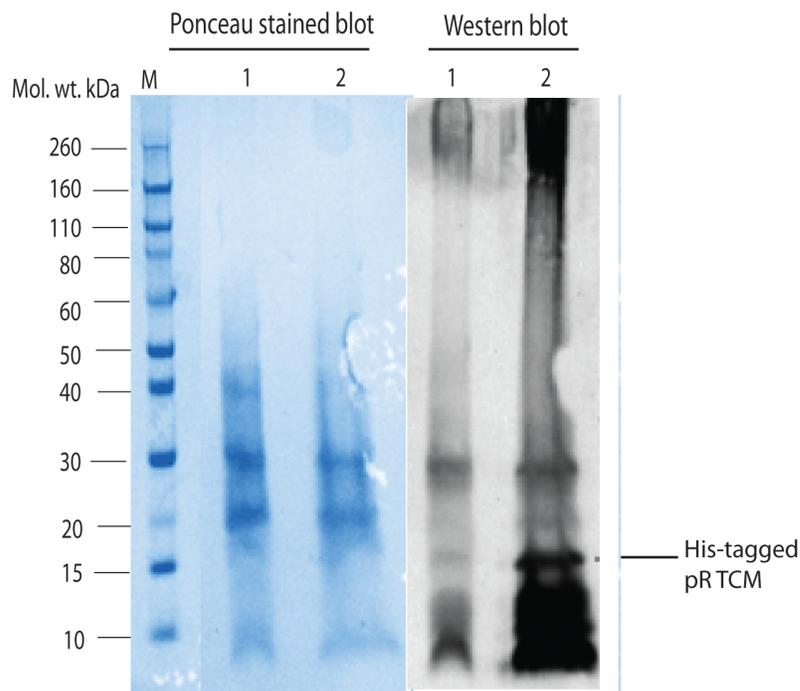
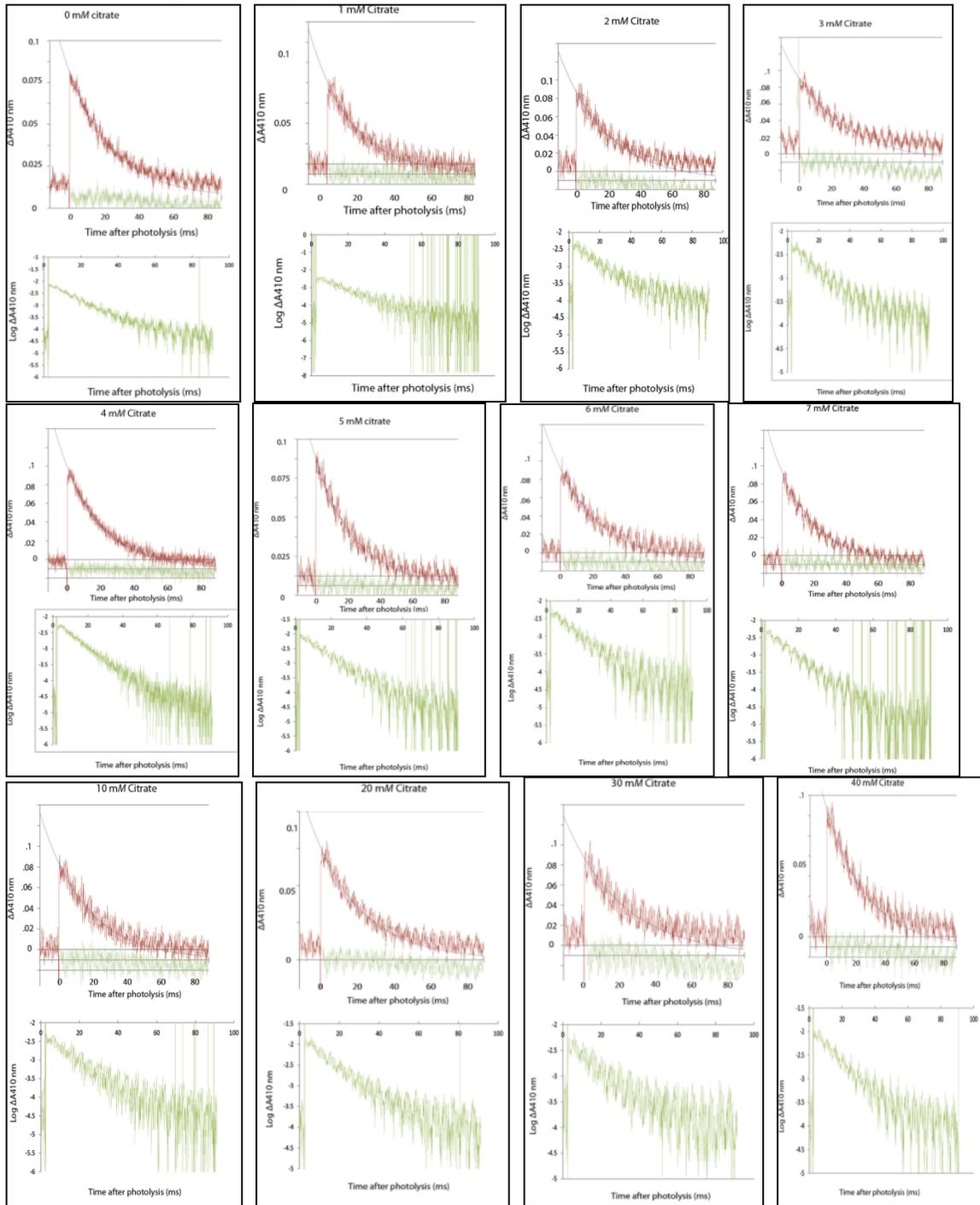


Fig. 2.10: Western Blot of pR expressed in *E. coli*: UT5600 *E. coli* cells transfected with plasmid pBAD TOPO TA containing pR TCM with, and without the polyHis-tag were induced with 0.2% L-arabinose and 5 μ M retinal was added. The Cultures were left at 16°C overnight under shaking conditions, 200 rpm. Cells were harvested and lysed using the lysis buffer described in chapter 2. Whole cell extract was precipitated with equal volume of 50% TCA. The resulting waxy pellet was washed with water and then treated with 50:50 mixture of chloroform:methanol. After drying it overnight in Speedvac^R, SDS-PAGE loading buffer was added and the samples were loaded on 4-12% gradient gel. The proteins were Western blotted on the nitrocellulose membrane, blocked with 5% milk prepared in Tris-buffered saline with Tween (TBST) and incubated with anti-His Antibody conjugated to horseradish peroxidase (HRP) overnight at 4°C. The blot was washed 8 times with 10-15 mL TBST for 15 minutes/wash. It was then treated with Chemiluminescent reagents (Pierce) and the signal was detected with BioRad gel documentation system. To visualize the molecular weight markers, the blot was incubated with Ponceau stain and detected similarly. Molecular weight marker, lane M, is Novex® Sharp Protein Standard, Invitrogen. Lane 3 is pR TCM without His-tag, and lane 2 is pR TCM with His₆-tag.

Acknowledgements for Fig 2.9: Anti-His antibody conjugated to HRP was kindly provided by Dr. Robert Doyle. SDS-PAGE gradient gel, Western blotting unit, molecular weight marker, chemiluminescent reagents and the BioRAD gel documentation system were generously provided by Dr. Michael Cosgrove. I thank both of them along with Brian Huta (helped me use Dr. Doyle's Western blotting unit, unfortunately the results were of poor quality), Valerie Vought, Dr. Anamika Patel and Melody Sanders (for their help with the above blot).

2.11 Flash Spectroscopy data



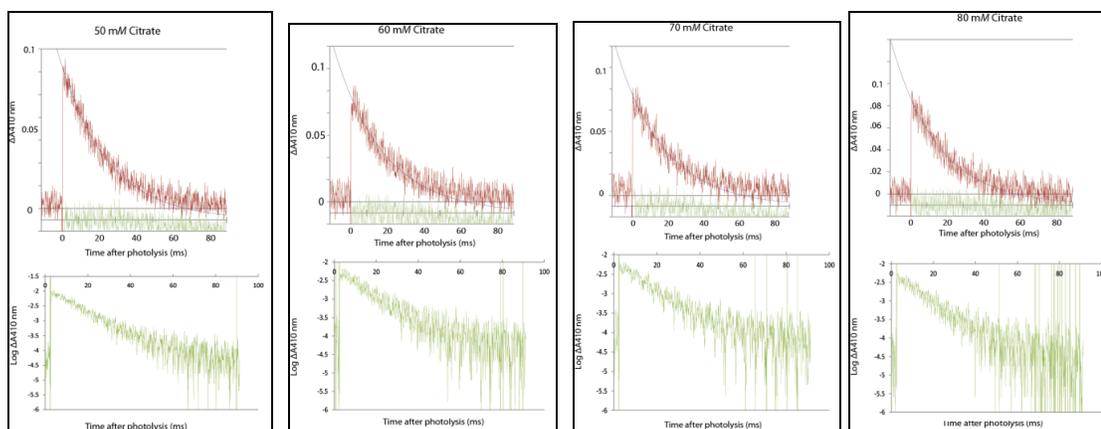


Fig. 2.11: Flash photolysis: Flash-induced transient absorbance changes in pR-containing DOPC liposomes (buffered at pH 8.5) show the decay of the M intermediate. These data were obtained from samples containing increasing amounts of citrate, as labeled. In part **A**, the measured absorbance data were fit to a first-order kinetic equation, $A_{400}(t) = B + A_0 e^{-kt}$, by using the solver function in Excel, optimizing the values of B , A_0 and k . The fitted exponential curve is superimposed on the measured transient absorbance at 420 nm. The residual trace (lightly-dotted line at $x = 0$) demonstrates the quality of the fit. Part **B** shows a log plot of absorbance against time. This plot yielded a straight line signifying a first-order exponential decay of the M intermediate. The decay constant, k , calculated from the slope of this curve matches that obtained from the non-linear least-squares optimization. (In **B**, and for the exponential fitting procedure in **A**, all digitized points up to the commencement of the M decay phase, corresponding to a total of ~8 ms of time, were omitted in order to allow the approximation of a single-exponential process.)

2.12 pR-citrate aggregation

Based on the lack of interaction with R51 (as measured by results with the single site mutant R51Q), and the possible involvement of R53 (as measured by the insensitivity to citrate of multi-site mutations involving this residue, as well as the inability of the single site mutant R53Q to fold properly), the phenomenon of citrate-induced pR aggregation is thus best explained by considering a homology model of three molecules of pR and two molecules of citrate (Fig. 2.12). The three pR molecules can be bridged by two citrate molecules, via salt bridges to both lysines and one arginine of the first intracellular loop of pR (Fig. 2.12). For example, an arginine (R53) of one pR (purple) can H-bond to one

carboxylate group of a citrate, while the two other carboxylate groups of the same citrate can be H-bonded to the lysines (K57 and K59) of a neighboring pR (green). This model indicates how multivalent citrate anions might bind to multiple pR molecules simultaneously, and thus lead to citrate-induced aggregation of pR.

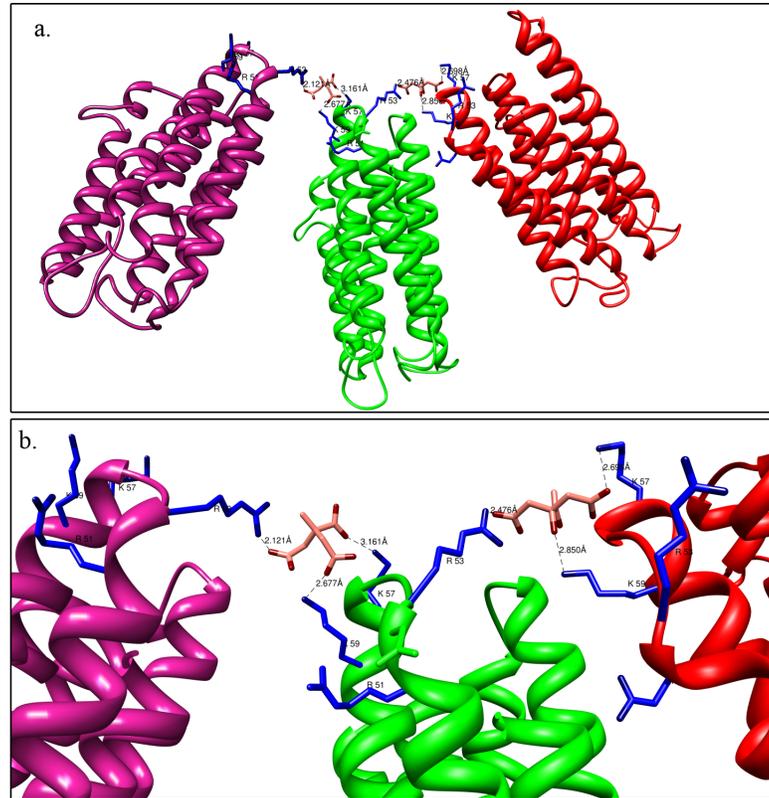


Fig. 2.12: Model of pR-citrate aggregation: **a.** Three molecules of pR, shown in purple, green and red, are bridged by citrate molecules, shown in pink and red, interacting via salt bridges to lysines and arginines of loop IC1 on each pR (blue). For example, Arg⁵³ of one pR (purple) is H-bonded to one carboxylate group of a citrate. The two other carboxylate groups of the same citrate are H-bonded to the lysines (Lys⁵⁷ and Lys⁵⁹) of a neighboring pR (green). This model indicates how multivalent citrate anions might bind to multiple pR molecules simultaneously, and thus to lead to citrate-induced aggregation of pR, **b.** Magnified image of the pR-citrate interactions.

The references to the construction of homology model and its manipulation are described in Chapter 2.

2.13 Interaction of pR with diverse ions

2.13.1 Precipitation of pR by Lithium salts

Citrate buffer and buffers of other diverse anions were prepared with lithium as the counter-ion, and the pR precipitation was examined.

Preparation of 250 mM citrate-LiOH buffer, with 10 mM Tris-Cl, pH 7.5

250 mM citric acid was dissolved in minimum amount of dH₂O. Tris-Cl buffered at pH 7.5 was added to a final concentration of 10 mM. This solution was then titrated with 1 M LiOH solution to attain pH 7.5.

Table 2.13.1: Determination of OG concentration at which pR exhibits the broadest range of precipitation with citrate buffer with lithium as the counterion, buffered at pH 7.5, also containing 10 mM Tris-Cl, pH 7.5

Conc. Of OG →	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
Conc. Of citrate buffer ↓						
5 mM	-	-	-	-	-	-
10 mM	-	++	++	+	-	-
50 mM	++	++	++	+	-	-
100 mM	++	++	++	-	-	-
150 mM	++	++	+	-	-	-

- ++ depicts complete precipitation of pR with no color in the supernatant, + depicts about 50% precipitation, and – shows no precipitation with all the color in the supernatant.

Conclusion: 0.7% OG was chosen for the examination of precipitation of pR by other Lithium salts.

2.13.2 Preparation of buffers of Lithium salts

The following buffers of Lithium were prepared to test the effect of Lithium on the precipitation of pR:

1. LiCl in 10 mM Tris-Cl, pH 7.5
2. Li₂SO₄ in 10 mM Tris-Cl, pH 7.5
3. Lithium acetate in 10 mM Tris-Cl, pH 7.5
4. Phosphate buffer with 10 mM Tris-Cl, titrated with 1M LiOH to pH 7.5
(prepared with phosphoric acid) – Higher concentration could not be prepared as Li-phosphate precipitated, only 100 mM was made.
5. Glutamate with 10 mM Tris-Cl, titrated with 1M LiOH to pH 7.5
(prepared with glutamic acid)
6. Aspartate with 10 mM Tris-Cl, titrated with 1M LiOH to pH 7.5 (prepared with aspartic acid)
7. Succinate with 10 mM Tris-Cl, titrated with 1M LiOH to pH 7.5 (prepared with succinic acid)
8. Cis-aconitate with 10 mM Tris-Cl, titrated with 1M LiOH to pH 7.5
(prepared with cis-aconitic acid)

Note: Other lithium salts with anions such as isocitrate, ATP, ADP, and G6P could not be prepared in similar fashion because these are available only as salts of sodium, rather than as free acids.

Table 2.13.2: Determination of concentration of Lithium salts required to induce pR precipitation at 0.7% OG

Buffer with 10 mM Tris-Cl, pH 7.5	10 mM	20 mM	30 mM	40 mM	50 mM	100 mM	150 mM
LiCl	-	-	+ (<50%)	+ (90%)	++	++	++
Li ₂ SO ₄	-	++	++	++	++	++	+
Lithium acetate	-	ND	-	- (<5%)	+ (50%)	+ (90%)	++
Lithium phosphate	-	++	++	++	++	++ (85 mM)	ND
Lithium glutamate	-	ND	+ (60%)	++	+ (90%)	+ (90%)	+ (90%)
Lithium aspartate	-	ND	+ (60%)	++	++	++	+ (50%)
Lithium succinate	-	ND	++	++	++	++	+ (90%)
Lithium cis-aconitate	+ (50%)	ND	ND	ND	++	++	-

- ++ indicates complete precipitation of pR with no color in the supernatant, + indicates approximately about 50%-90% precipitation, and – indicates no precipitation retaining all the color in the supernatant
- ND means “not determined”

Conclusion: Diverse anions with lithium as the counter-ion tend to precipitate pR.

However, they do so at a lower concentration of OG than that shown by the respective sodium salts.

- **2.13.3 Precipitation of pR by potassium salts**

Table 2.13.3: Determination of pR precipitation by salts of Potassium: Anions were buffered to pH 7.5, also contained 10 mM Tris-Cl at pH 7.5, and a sodium counterion. The concentration of OG is maintained at a constant 1%.

Anion tested	pR precipitation (Min. concentration, mM)	Physiologically - indexed concentration	Physiological concentration (mM) ^{14, 15, 25, 30}
Acetate (potassium)	ND	N/A	N/A
Phosphate (potassium)	40 (++)	7.530	5.3
Chloride (potassium)	50 (++) ND at 100 mM	555.5	0.09

- ++ indicates complete precipitation of pR, + indicates partial precipitation leaving significant color in the supernatant, - indicates lack of precipitation with all the color in the supernatant.
- The references for concentrations for physiologically important anions available are labeled as superscripts. Concentration of phosphate is from *Streptococcus bovis* [30] and concentration of chloride is from marine bacteria *Pseudomonad* [25]. All others are from *E. coli*. (References refer to Chapter 2)
- N/A refers to data Not Available.

- **2.13.4 Precipitation with bromide as the anion**

The aim was to determine if the effect of chloride in the process of pR precipitation could be minimized. Since bromide is placed towards the right side of chloride in the Hofmeister series of anions, it is expected to show less salting-out effect. A buffer of 10 mM Tris was pH adjusted with HBr to pH 7.5, to test the precipitation of pR. Additionally, since all the anions are the sodium salts, it was necessary to examine if sodium bromide by itself or buffered with Tris-Br would induce pR precipitation. Tris-Br buffer or NaBr buffer/solution were prepared as follows:

Tris-Br buffer, pH 7.5: 220 mM Tris was dissolved in minimum dH₂O. It was titrated with HBr (conc.) to attain a pH of 7.5

NaBr solution: 200 mM NaBr was prepared with dH₂O

NaBr buffered with Tris-Br: Tris-Br buffer was added to the reaction tubes to a final concentration of 10 mM followed by addition of NaBr to the required concentration. This ensured a constant concentration of the Tris-Br buffer.

Table 2.13.4: Determination of pR precipitation with Bromide at pH 7.5 and at 1% OG

Precipitant tested: Buffer/solution	10 mM	50 mM	100 mM	150 mM
Tris-Br	-	+	+	+
NaBr solution	-	-	++	++
NaBr buffered with Tris-Br	-	+	+	+

- ++ indicates complete precipitation of pR, + indicates partial precipitation leaving significant color in the supernatant, - indicates lack of precipitation with all the color in the supernatant.

3.8 Purity of pR can be predicted by the $A_{280/520}$ ratio

There is an inverse relationship between $A_{280/520}$ ratio and the purity of proteorhodopsin. As the purity of pR increases, the ratio decreases: the sample contains less impurities and more pR leading to higher A_{520} absorption for the same A_{280} value. The contribution of pR toward A_{280} also increases with purification. Fig. 3.8 shows a plot of purity (%) of pR against the $A_{280/520}$ ratio. The calculation of purity was performed

using equation 1 (Table 3.5) and the values of ratio were ranged from 2-32. Ratio of 2 corresponds to the maximum theoretical purity of pR obtained in our lab [1].

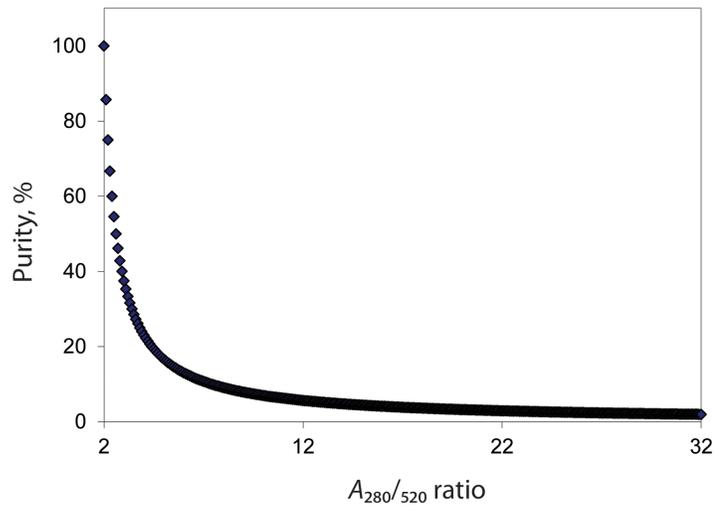


Fig. 3.8: Dependence of purity on $A_{280/520}$ ratio: Purity (%) is plotted against $A_{280/520}$ ratio. The values of ratio are obtained from Table 3.4, the purity (%) values are obtained from equation 1. For comparison purposes, $A_{280/520}$ ratio of 2, a theoretical ratio that has been obtained for pure pR in our lab [1] is also included. It corresponds to a 100% purity.

Table 3.8: Derivation of equation

Plot of % purity against $A_{280/520}$ ratio		
Assume A_{280}	R	
Assume A_{520}	1	
Theoretical ratio $A_{280/520}$ for pure pR	2	
A_{280} due to contaminant proteins	= R-2	
mg/ml contaminant proteins ($A_{280} = 1 = 1$ mg/ml)	= R-2	
mg/ml of pR ($A_{520}/50000*30000$)***	0.6	
total protein contaminants + pR)	(R-2) + 0.6	
therefore, total protein (contaminants + pR)	=R-1.4	
%purity of pR in total protein	$(0.6/R-1.4)*100$	
	= 60/R-1.4 where R denotes Ratio $A_{280/520}$	Equation 1
	R [Ratio($A_{280/520}$)]	% purity
Theoretical value for pure pR	2	100
D	2.8	42.85714
C	3.3333	31.03502
B	6.3636	12.088
A	26.66	2.375297

*** Molar extinction coefficient = $50,000 \text{ M}^{-1} \text{ cm}^{-1}$
Molecular weight of pR = 30,000 mg/mL (or g/L).

3.9 N-terminal sequencing results

Purification of pR with citrate or with a combination of citrate and phosphate yields some contaminants. A band corresponding to 30,000 Da from an SDS-PAGE of pR purification was excised and N-terminal sequencing was carried out using Edman degradation (Iowa State University Protein Facility). The result showed the following sequence:

V/A, D/N, F/L, H/M, Q, Y, A, R, S?/Q?, I?

The detailed N-terminal peptide sequence report is attached at the end of this subsection 3.9.4. The most likely sequence was suggested to be VDFLHQYARSI (See detailed results). This sequence was subjected to BLASTp [2] to identify the protein from

the database. The sequence alignment yielded numerous hits, mainly including maltoporins. Since pR is expressed in *E. coli* (Strain UT5600), interest was concentrated on *E. coli* hits (Table 3.9.1). It was found that the sequence matched (90%) to the LamB protein of *E. coli* as shown below.

```
VDFHQYARS ---N-terminal sequence result
VDFH YARS ---showing the match
VDFHGYARS ---E. coli maltoporin sequence from BLAST
```

The protein sequence of lamB is shown in Fig. 3.9.1., and the details corresponding to this protein, such as, the protein accession number, the authors and publication of this sequence, are provided in Fig. 3.9.2. LamB is expressed with an N-terminal leader sequence that is cleaved off during the translocation into the membrane. The sequence of the entire protein containing the leader sequence is shown in Fig. 3.9.3.

LamB protein is a specific type of porin called maltoporin. Porins are a class of proteins that belong to β -barrel proteins. Maltoporins are therefore β -barrel proteins that are found in the outer membrane of gram negative bacteria including *E. coli*. They are homo-trimers that form water-filled channels that allow passive diffusion of small molecules (<600 Da) [3]. Specifically, LamB is part of the maltose regulon that encodes proteins required for uptake and metabolism of maltose and other linear maltodextrins [3, 4].

The relative molecular weight (Mr) is 142 kDa. However, it appears as a band at ~95 kDa on SDS-PAGE that has been suggested to possess a compact or folded form.

Similarly, the compact monomer has been found to run at ~35-37 kDa, depending upon the concentration of acrylamide in the SDS-PAGE [4]. About 80% folded trimer appears at ~66 kDa, and denatured monomer runs at 45 kDa [4].

Based on above information, it appears that the majority of the contaminants in pR preparation arise from the presence of LamB maltoporin. Since maltoporins are residents of outer membrane of *E. coli*, it can be speculated that removal of the outer membrane during pR purification would eliminate maltoporin contamination. This would yield a fairly pure protein with a purity of >90% using a combination of citrate and phosphate as described in the text.

In order to remove the outer membrane of *E. coli*, penicillin or lysozyme can be used in the presence of EDTA while also maintaining iso-osmotic conditions by using sucrose. Isolation of resulting spheroplasts (cells lacking outer membrane) can be done using centrifugation at higher speeds.

Table 3.9.1: BLAST hits from alignment of N-terminal sequence result against reference sequence database: The result obtained from N-terminal sequencing was BLASTed against the protein database of reference sequences using BLASTp algorithm. The results containing *E. coli* hits were selected

Accession	Description	Max score	Total score	Query coverage	E value	Links
ZP_07185557.1	LamB porin [Escherichia coli MS 69-1] >ref[ZP_08356844.1] maltoporin (Maltose-inducible porin) (Lambda receptorprotein) [Escherichia coli M718]	26.9	26.9	90%	17	
ZP_06660163.1	maltoporin [Escherichia coli B185]	26.9	26.9	90%	17	
YP_002405410.1	maltoporin [Escherichia coli 55989]	26.9	26.9	90%	17	G
YP_002295602.1	maltoporin [Escherichia coli SE11]	26.9	26.9	90%	17	G
ZP_02903949.1	maltose-inducible porin [Escherichia albertii TW07627]	26.9	26.9	90%	17	
NP_290670.1	maltoporin [Escherichia coli O157:H7 EDL933] >ref[NP_313046.1] maltoporin [Escherichia coli O157:H7 str. Sakai] >ref[ZP_02775854.1] maltose-inducible porin [Escherichia coli O157:H7 str. EC4113] >ref[ZP_02780195.1] maltose-inducible porin [Escherichia coli O157:H7 str. EC4401] >ref[ZP_02799382.2] maltose-inducible porin [Escherichia coli O157:H7 str. EC4196] >ref[ZP_02805437.2] maltose-inducible porin [Escherichia coli O157:H7 str. EC4076] >ref[ZP_02791285.2] maltose-inducible porin [Escherichia coli O157:H7 str. EC4486] >ref[ZP_02785890.2] maltose-inducible porin [Escherichia coli O157:H7 str. EC4501] >ref[ZP_02809640.2] maltose-inducible porin [Escherichia coli O157:H7 str. EC869] >ref[ZP_02823327.2] maltose-inducible porin [Escherichia coli O157:H7 str. EC508] >ref[ZP_03081185.1] maltoporin [Escherichia coli O157:H7 str. EC4024] >ref[ZP_03250829.1] maltose-inducible porin [Escherichia coli O157:H7 str. EC4206] >ref[ZP_03255079.1] maltose-inducible porin [Escherichia coli O157:H7 str. EC4045] >ref[ZP_03261262.1] maltose-inducible porin [Escherichia coli O157:H7 str. EC4042] >ref[YP_002273557.1] maltose-inducible porin [Escherichia coli O157:H7 str. EC4115] >ref[ZP_03441540.1] maltose-inducible porin [Escherichia coli O157:H7 str. TW14588] >ref[YP_003080875.1] maltoporin [Escherichia coli O157:H7 str. TW14359] >ref[YP_003502271.1] maltoporin precursor [Escherichia coli O55:H7 str. CB9615]	26.9	26.9	90%	17	G

NP_756858.1	maltoporin [Escherichia coli CFT073] >ref YP_543545.1 maltoporin [Escherichia coli UTI89] >ref YP_859628.1 maltoporin [Escherichia coli APEC O1] >ref YP_002331805.1 maltoporin [Escherichia coli O127:H6 str. E2348/69] >ref YP_002394020.1 maltoporin [Escherichia coli S88] >ref YP_002400534.1 maltoporin [Escherichia coli ED1a] >ref ZP_04004763.1 maltoporin [Escherichia coli 83972] >ref ZP_04534075.1 maltoporin [Escherichia sp. 3_2_53FAA] >ref ZP_07180104.1 LamB porin [Escherichia coli MS 45-1] >ref ZP_07194092.1 LamB porin [Escherichia coli MS 185-1] >ref ZP_07449995.1 maltoporin [Escherichia coli NC101] >ref ZP_08350973.1 maltoporin (Maltose-inducible porin) [Escherichia coli M605] >ref ZP_08361536.1 maltoporin (Maltose-inducible porin) [Escherichia coli TA206] >ref ZP_08386364.1 maltoporin (Maltose-inducible porin) [Escherichia coli H299]	26.9	26.9	90%	17	G
ZP_03000669.1	maltose-inducible porin [Escherichia coli 53638] >ref ZP_07133992.1 LamB porin [Escherichia coli MS 115-1]	26.9	26.9	90%	17	
YP_001746426.1	maltoporin [Escherichia coli SMS-3-5] >ref ZP_03048831.1 maltose-inducible porin [Escherichia coli E110019] >ref YP_002410330.1 maltoporin [Escherichia coli IAI39] >ref ZP_06651609.1 conserved hypothetical protein [Escherichia coli B354] >ref ZP_07152941.1 LamB porin [Escherichia coli MS 21-1] >ref ZP_08366600.1 maltoporin (Maltose-inducible porin) (Lambda receptorprotein) [Escherichia coli TA143]	26.9	26.9	90%	17	G
YP_672105.1	maltoporin [Escherichia coli 536] >ref ZP_03033493.1 maltose-inducible porin [Escherichia coli F11] >ref ZP_07175041.1 LamB porin [Escherichia coli MS 200-1]	26.9	26.9	90%	17	G
ZP_03044172.1	maltose-inducible porin [Escherichia coli E22] >ref ZP_03059296.1 maltose-inducible porin [Escherichia coli B171] >ref YP_003224609.1 maltose outer membrane porin [Escherichia coli O103:H2 str. 12009]	26.9	26.9	90%	17	
YP_312948.1	maltoporin [Shigella sonnei Ss046] >ref YP_691468.1 maltoporin [Shigella	26.9	26.9	90%	17	G

	<p>flexneri 5 str. 8401] >ref YP_001465537.1 maltoporin [Escherichia coli E24377A]</p> <p>>ref YP_001460823.1 maltoporin [Escherichia coli HS]</p> <p>>ref YP_001726919.1 maltoporin [Escherichia coli ATCC 8739]</p> <p>>ref ZP_03029272.1 maltose-inducible porin [Escherichia coli B7A]</p> <p>>ref YP_002389506.1 maltoporin [Escherichia coli IAI1]</p> <p>>ref YP_002415177.1 maltoporin [Escherichia coli UMN026]</p> <p>>ref YP_003232038.1 maltose outer membrane porin [Escherichia coli O26:H11 str. 11368] >ref YP_003237152.1 maltose outer membrane porin [Escherichia coli O111:H- str. 11128] >ref ZP_06646772.1 lamB [Escherichia coli FVEC1412]</p> <p>>ref ZP_06664749.1 maltoporin [Escherichia coli B088]</p> <p>>ref ZP_06988088.1 maltoporin [Escherichia coli FVEC1302]</p> <p>>ref ZP_07098165.1 LamB porin [Escherichia coli MS 107-1]</p> <p>>ref ZP_07103702.1 LamB porin [Escherichia coli MS 119-7]</p> <p>>ref ZP_07118062.1 LamB porin [Escherichia coli MS 198-1]</p> <p>>ref ZP_07124042.1 LamB porin [Escherichia coli MS 84-1]</p> <p>>ref ZP_07140275.1 LamB porin [Escherichia coli MS 182-1]</p> <p>>ref ZP_07208840.1 LamB porin [Escherichia coli MS 124-1]</p> <p>>ref ZP_07223201.1 LamB porin [Escherichia coli MS 78-1]</p> <p>>ref ZP_07591760.1 porin LamB type [Escherichia coli W] >ref ZP_07690255.1 LamB porin [Escherichia coli MS 145-7]</p> <p>>ref ZP_07788346.1 maltoporin [Escherichia coli 1827-70]</p> <p>>ref ZP_08371730.1 maltoporin (Maltose-inducible porin) [Escherichia coli TA271]</p> <p>>ref ZP_08380822.1 maltoporin (Maltose-inducible porin) [Escherichia coli H591]</p> <p>>ref ZP_08393167.1 phage lambda receptor protein [Shigella sp. D9]</p>					
NP_418460.1	<p>maltose outer membrane porin (maltoporin) [Escherichia coli str. K-12 substr. MG1655] >ref YP_001732813.1 maltoporin [Escherichia coli str. K-12 substr. DH10B] >ref ZP_03069347.1 maltose-inducible porin [Escherichia coli 101-1] >ref YP_002928942.1 maltose outer membrane porin (maltoporin)</p>	26.9	26.9	90%	17	

	[Escherichia coli BW2952] >ref YP_003038169.1 maltoporin [Escherichia coli 'BL21-Gold(DE3)pLysS AG'] >ref ZP_04871107.1 maltose- inducible porin [Escherichia sp. 1_1_43] >ref YP_003047080.1 maltoporin [Escherichia coli B str. REL606] >ref ZP_07145223.1 LamB porin [Escherichia coli MS 187-1] >ref ZP_07160863.1 LamB porin [Escherichia coli MS 116-1] >ref ZP_07169764.1 LamB porin [Escherichia coli MS 175-1] >ref ZP_07183347.1 LamB porin [Escherichia coli MS 196-1] >ref ZP_07244983.1 LamB porin [Escherichia coli MS 146-1] >ref ZP_08345891.1 maltoporin (Maltose- inducible porin) (Lambda receptorprotein) [Escherichia coli H736]					
ZP_07779537.1	maltoporin [Escherichia coli 2362-75]	26.9	26.9	90%	17	
ZP_05937728.1	maltose outer membrane porin (maltoporin) [Escherichia coli O157:H7 str. FRIK2000] >ref ZP_05949515.1 maltose outer membrane porin (maltoporin) [Escherichia coli O157:H7 str. FRIK966]	26.9	26.9	90%	17	

Fig. 3.9.1: Sequence of lamB protein from *E. coli*

GenBank: CAA02071.1

```
>gi|4529934|emb|CAA02071.1| lamB [Escherichia coli]
VDFHGYARSGIGWTGSGGEQQCFQTTGAQSKYRLGNECETYAELKLGQEVWKEGDKSFYFDTNVAYSVAQ
QNDWEATDPAFREANVQGNLIEWLPGSTIWAGKRFYQRHDVHMIDFYWDISGPGAGLENIDVGFGLKS
LAATRSSEAGSSSFASNNIYDYNETANDVDFVRLAQMEINPGGTLELGVDYGRANLRDNYRLVDGASK
DGWLF TAEHTQSVLKGFNKFVVQYATDSMTSQGKLSQSGVAFDNEKFAYNINNNGHMLRILDHGAI SM
GDNWDMMYVGM YQDINWDNDNGTKWWTVGIRPMYKWTPI MSTVMEIGYDNVESQRTGDKNNQYKITLAQQ
WQAGDSIWSRPAIRVFATYAKWDEK WGYDYTG NADNNANFGKAVPADFN GGSFGRGDSDEWTFGAQMEIW
W
```

Fig. 3.9.2: Accession number, author, and publication details of lamB [*Escherichia coli*]

GenBank: CAA02071.1

```
LOCUS          CAA02071          421 aa          linear    PAT 20-JUN-1996
DEFINITION    lamB [Escherichia coli].
ACCESSION    CAA02071
VERSION      CAA02071.1  GI:4529934
DBSOURCE     embl accession A31943.1
KEYWORDS     .
SOURCE       Escherichia coli
  ORGANISM   Escherichia coli
             Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
             Enterobacteriaceae; Escherichia.
REFERENCE    1 (residues 1 to 421)
  AUTHORS    Hofnung,M., Bouges-Bocquet,B. and Guesdon,J.L.
  TITLE      Vector modified by at least a part if not the whole of gene Lam B
             and microorganisms transformed with this vector and made capable of
             synthesizing a determined external membrane protein, encoded by an
             insert also included in said vector
  JOURNAL    Patent: EP 0146416-A1 26-JUN-1985;
             INSTITUT PASTEUR; INSERM; CENTRE NATIONAL DE LA RECHERCHE
             SCIENTIFIQUE (CNRS)
FEATURES     Location/Qualifiers
  source     1..421
             /organism="Escherichia coli"
             /db_xref="taxon:562"
  Protein   <1..>421
             /name="lamB"
  Region   1..421
             /region_name="Maltoporin-like"
             /note="The Maltoporin-like channels (LamB porin) form a
             trimeric structure which facilitate the diffusion of
             maltodextrins and other sugars across the outer membrane of
             Gram-negative bacteria. The membrane channel is formed by an
             18-strand antiparallel beta-...; cd01346"
  Site     /db_xref="CDD:30073"
             order(3,9,11..13,40..42,58,60,64..66,68,81..82,84..86,88,
             100,102..104,106,123..127,151..155,168,170,319,351..352,
             361,363,415,417,419,421)
             /site_type="other"
             /note="trimer interface"
             /db_xref="CDD:30073"
  Site     order(8,41,43,82,106,109,116,118,121)
             /site_type="other"
             /note="sugar binding site [chemical binding]"
             /db_xref="CDD:30073"
  CDS     1..421
             /gene="lamB"
             /coded_by="A31943.1:<1..>1263"
             /transl_table=11
ORIGIN
1  vdfhgyarsg igwtgsggeq qcfqttgaqs kyrlgnecet yaelklgqev wkegdksfyf
61  dtnvaysvaq qndweatdpa freanvqgkn liewlpgsti wagkrfyqrh dvhmidfyyw
121 dispggagle nidvgfgkls laatrsseag gsssfasnni ydytnetand vfdvrlaqme
181 inpggtlelg vdygranlrd nyrlvdgask dgwlftaht qsvlkgfnkf vvqyatdsmt
241 sggkglsggs gvafdnkfa yninnnghml rildhgaism gdnwdmmyvg myqdinwdnd
301 ngtkwvtvgi rpmykwtpim stvmeigydn vesqrtgdkn nqykitlaqq wqagdsiwsr
361 pairvfatya kwdekwgydy tgnadnnanf gkavpadfng gsfgrgdsde wtfgaqmeiw
421 w
```

Fig. 3.9.3: Sequence of LamB porin showing the N-terminal leader sequence from LamB porin [Escherichia coli M863]: The leader sequence is shown in bold italicized letters.

GenBank: EGB61149.1

```
>gi|323965697|gb|EGB61149.1| LamB porin [Escherichia coli M863]
MMITLRKLPLAVAVAAGVMSAQAMAVDFHGYARSGIGWTGSGGEQQCFQTTGAQSKYRLGNECETYAELK
LGQEVWKEGDKSFYFDTNVAYSVAQQNDWEATDPAFREANVQGKNLIEWLPGSTIWAGKRFYQRHVDHMI
DFYYWDISGPGAGLENIDVGFGKLSLAATRSSEAGGSSSFASNNIYDYTNETANDVFDVRLAQMEINPGG
TLELGVDYGRANLRDNYRLVDGASKDGWLFTAEHTQSVLKGFNKFVVQYATDSMTSQGKGLSQGSGVAFD
NEKFAYNINNNHMLRILDHGAI SMGDNWDMMYVGMYQDINWDNDNGTKWWTVGIRPMYKWTPIMSTVME
IGYDNVESQRTGDKNNQYKITLAQQWQAGDSIWSRPAIRVFATYAKWDEKWGYDYNGSSSTNPYYGKAVP
ADFNGGSFGRGDSDEWTFGAQMEIWW
```

Fig. 3.9.4: N-terminal peptide sequence report: The excised band corresponding to ~35 kDa was sequenced at the Iowa State University Protein Facility. The details of the report are described therein.

Iowa State University
Protein Facility

PROTEIN/PEPTIDE SEQUENCE
REPORT

Date: July 30, 2007

To: Farhana Syed

Sample Number: 3570

Sample Name: ECO CIT 35k #3

Sample Preparation: The sample was washed 6 times with DI water and loaded onto a filter for sequence analysis.

Instruments: 494 Procise Protein Sequencer/140C Analyzer from

Applied Biosystems, Inc.

Sequencing Method: Edman Degradation

<u>Cycle Number</u>	<u>Amino Acid</u>
1	V, A
2	D, N
3	F, L
4	H, M
5	Q
6	Y
7	A
8	R
9	S?, Q?
10	I?

Note

515-294-3267,

The major amino acid is listed first for each cycle. If you have any questions, feel free to call me.

Prepared by: Joel Tel:

Note

: 515-294-3267, e

e-mail:protein@iastate.edu

Protein/Peptide Sequence Analysis Submission Form

Sequence Number 3570 Date July 20th, 2007
Your Sample ID # Eco CII 35K #3
Name Farhana Syed Supervisor Dr. Mark S. Brauman
Department/Company Syracuse Univ. Phone # 315-443-4691 Fax # 315-443-4070
Mailing Address 1-014 CST, Dept. of Chemistry, Syracuse Univ, Syracuse - NY-13241
E-Mail Address mbrauman@Syr.edu fsyed@Syr.edu

Account or P. O. # _____
For on campus orders please submit an intramural with your order. For off campus orders please submit a purchase order.

How many residues do you need? 10

Sample Information

Sample amount _____ moles; or ~10 micrograms M.W. 35 kDa

For samples in solution: What solvent is the sample in? _____

For samples electroblotted to PVDF: What membrane was used?

Immobilon-P (.45 micron) (Millipore) Problot (.1 micron) (ABI)
 Westran (.45 micron) (Schleicher & Schuell) Trans-Blot (.1 micron) (Biorad)
 Immobilon-PSQ (.1 micron) (Millipore) Fluorotrans (.1 micron) (Pall Corp)

N-Terminal blocked: No _____ Do not know Yes _____

Protein/Peptide Modified: Yes, at _____ with _____

Cysteine modified: Yes _____ If yes, what derivative? _____ No

Enzyme treatment: Yes _____ What enzyme? _____ Cleavage sites _____

Radioactivity: Yes _____ No

Protein sequence known: Yes _____ No (Please attach)

DNA sequence known: Yes _____ No (Please attach)

Amino acid analysis performed: Yes _____ No (Please attach)

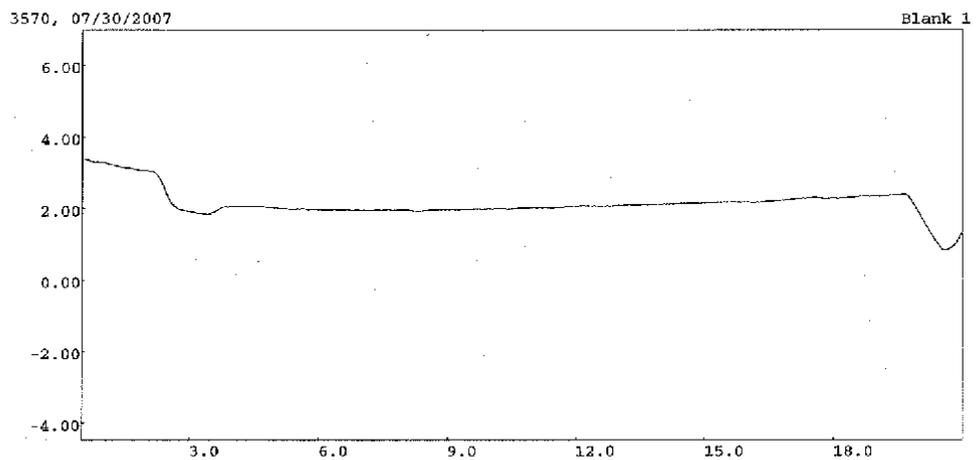
Describe purification steps in detail, especially possible contaminants such as buffer, salts, and SDS:
12.5% SDS PAGE followed by Western blotting on the PVDF membrane. Stained with Coomassie blue, destained, the band (2 of the same sample) cut out & reblotted w/ deionized water.

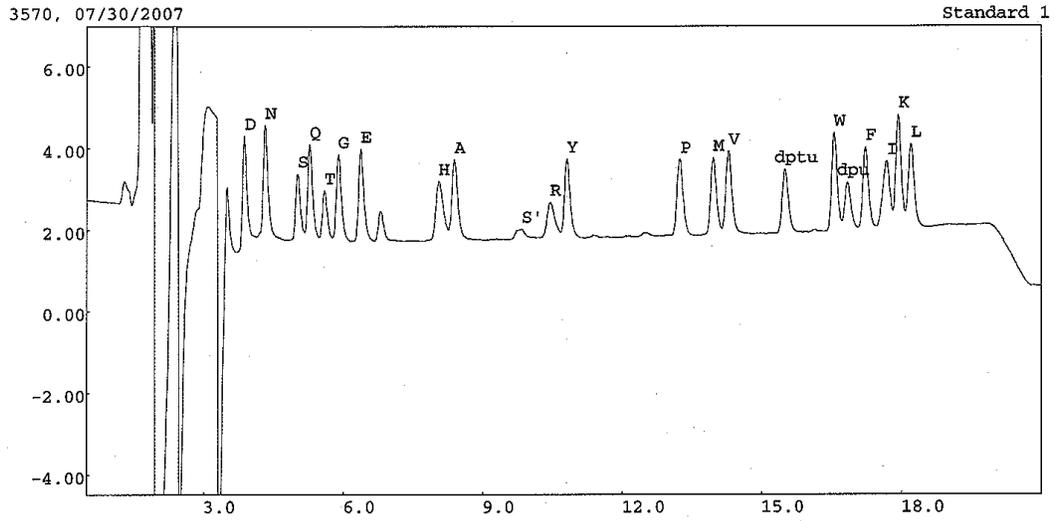
If your sample was collected on an HPLC, please attach the chromatogram with AUFS, gradient, solvents, column and wavelength.

7/30/07 4:41 PM

3570, 07/30/2007

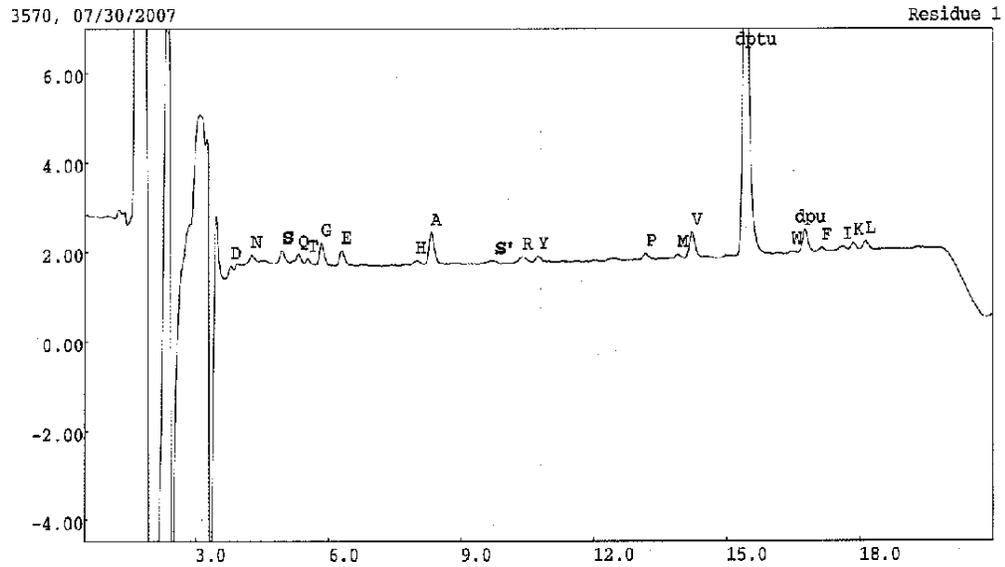
1





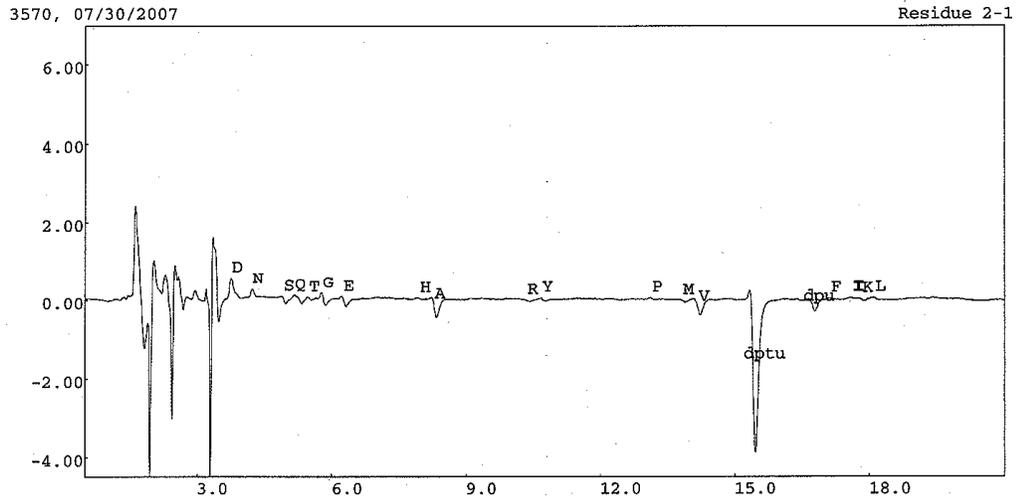
Standard 1, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol Ht
2	3.87	D	c	36251	10.00
3	4.32	N	c	30594	10.00
4	5.01	S	c	13267	10.00
5	5.27	Q	c	19250	10.00
6	5.59	T	c	10102	10.00
7	5.89	G	c	17420	10.00
8	6.37	E	c	18552	10.00
11	8.05	H	c	11985	10.00
12	8.39	A	c	16111	10.00
17	9.83	S'	c	1920	10.00
19	10.45	R	c	7152	10.00
20	10.80	Y	c	15884	10.00
28	13.24	P	c	15442	10.00
30	13.97	M	c	15575	10.00
31	14.29	V	c	16825	10.00
34	15.51	dptu	c	12778	10.00
37	16.56	W	c	19913	10.00
38	16.85	dpu	c	9780	10.00
39	17.23	F	c	16679	10.00
40	17.68	I	c	13636	10.00
41	17.94	K	c	22878	10.00
42	18.21	L	c	16896	10.00



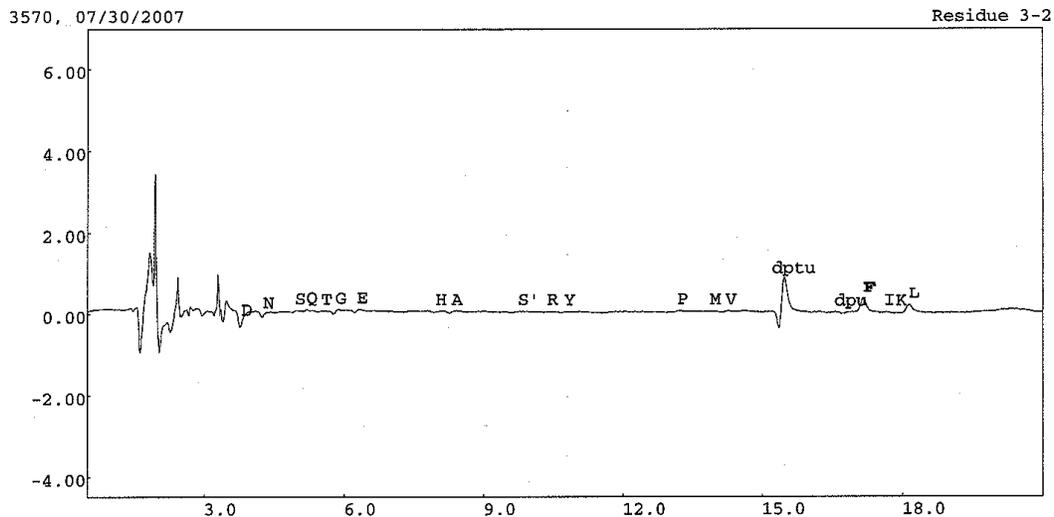
Residue 1, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol Ht
2	3.79	D	c	8422	2.32
4	4.26	N	c	8737	2.86
7	4.95	S	c	6968	5.25
9	5.31	Q	c	4898	2.54
10	5.52	T	c	3165	3.13
11	5.83	G	c	4952	2.84
12	6.29	E	c	2757	1.49
18	7.98	H	c	774	0.65
19	8.32	A	c	5864	3.64
24	9.76	S'	c	480	2.50
26	10.38	R	c	1090	1.52
27	10.74	Y	c	1021	0.64
37	13.18	P	c	1096	0.71
41	13.90	M	c	829	0.53
42	14.22	V	c	4884	2.90
46	15.44	dptu	c	73115	57.22
48	16.48	W	c	352	0.18
50	16.78	dpu	c	4234	4.33
51	17.16	F	c	735	0.44
52	17.62	I	c	765	0.56
53	17.87	K	c	1330	0.58
54	18.14	L	c	1543	0.91



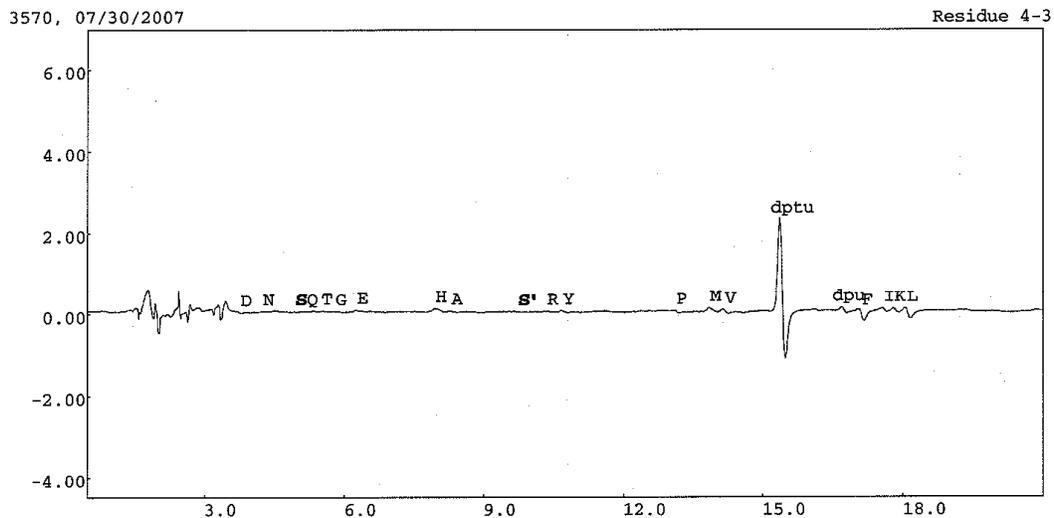
Residue 2, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol Ht
2	3.76	D	c	5249	1.45
4	4.24	N	c	3164	1.03
6	4.92	S	c	1652	1.25
7	5.18	Q	c	1021	0.53
9	5.50	T	c	817	0.81
10	5.80	G	c	3451	1.98
11	6.25	E	c	1626	0.88
17	7.96	H	c	801	0.67
18	8.29	A	c	2341	1.45
26	10.37	R	c	648	0.91
27	10.70	Y	c	757	0.48
34	13.15	P	c	1142	0.74
35	13.86	M	c	153	0.10
37	14.20	V	c	1508	0.90
40	15.43	dptu		45054	35.26
43	16.76	dpu		1900	1.94
44	17.15	F	c	962	0.58
45	17.62	I	c	1138	0.83
46	17.85	K	c	1137	0.50
47	18.12	L	c	1844	1.09



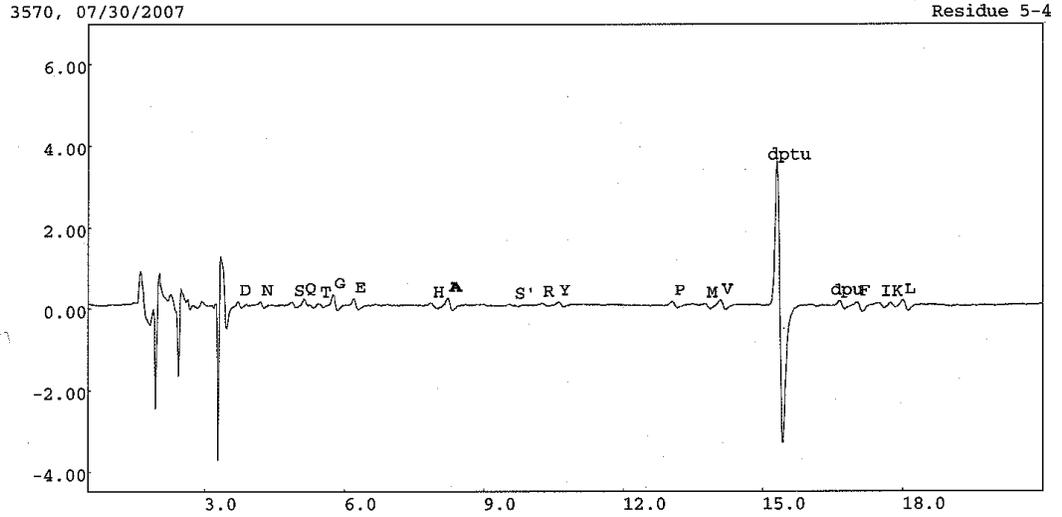
Residue 3, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol Ht
2	3.78	D	c	2324	0.64
4	4.25	N	c	2197	0.72
6	4.93	S	c	1563	1.18
7	5.18	Q	c	1086	0.56
9	5.51	T	c	700	0.69
10	5.82	G	c	3277	1.88
11	6.27	E	c	1656	0.89
18	7.97	H	c	682	0.57
19	8.30	A	c	2239	1.39
24	9.73	S'	c	369	1.92
28	10.36	R	c	610	0.85
29	10.73	Y	c	768	0.48
39	13.16	P	c	1261	0.82
41	13.87	M	c	250	0.16
42	14.21	V	c	1584	0.94
45	15.43	dptu	c	49071	38.40
48	16.77	dpu	c	1675	1.71
49	17.15	F	c	3745	2.25
50	17.60	I	c	1218	0.89
51	17.86	K	c	1059	0.46
52	18.13	L	c	3381	2.00



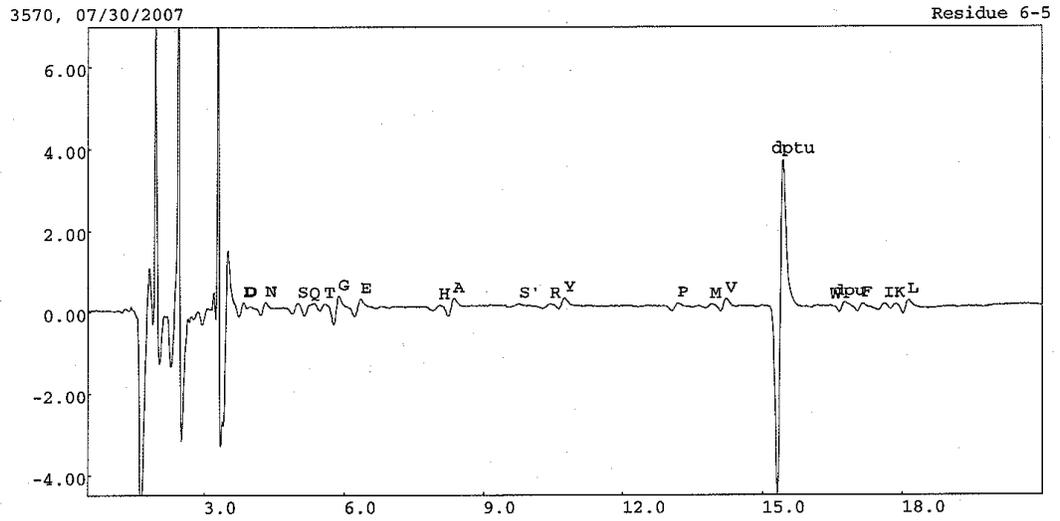
Residue 4, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol Ht
2	3.77	D	c	1989	0.55
4	4.25	N	c	3526	1.15
6	4.93	S	c	2861	2.16
7	5.18	Q	c	2397	1.25
9	5.51	T	c	1609	1.59
10	5.82	G	c	3824	2.20
11	6.27	E	c	2076	1.12
14	7.97	H	c	1326	1.11
15	8.30	A	c	2258	1.40
20	9.73	S'	c	360	1.88
21	10.36	R	c	708	0.99
22	10.70	Y	c	807	0.51
28	13.13	P	c	1224	0.79
30	13.86	M	c	975	0.63
31	14.18	V	c	1630	0.97
37	15.40	dptu		53880	42.17
39	16.74	dpu		1701	1.74
40	17.12	F	c	2121	1.27
41	17.59	I	c	1477	1.08
42	17.83	K	c	1184	0.52
43	18.10	L	c	2528	1.50



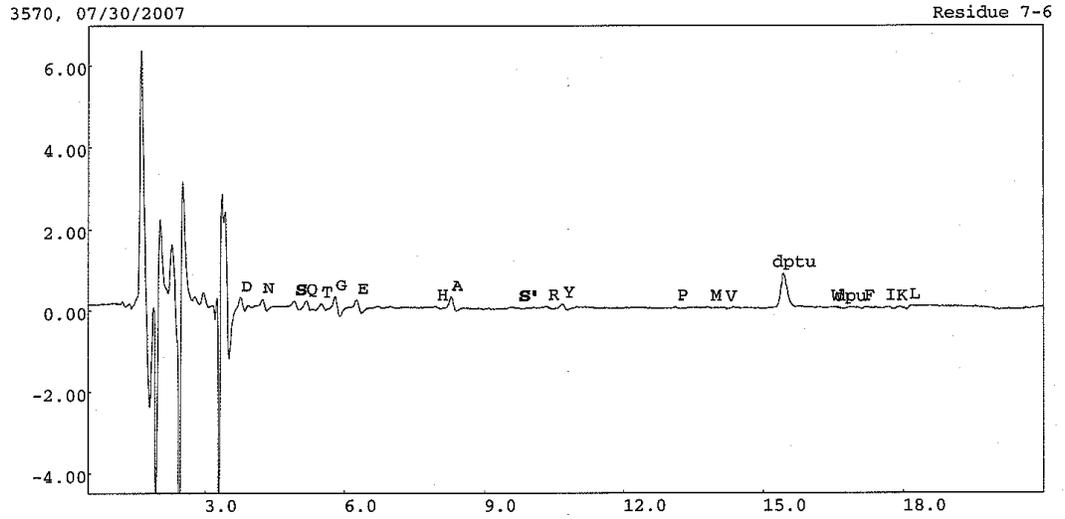
Residue 5, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol	Ht
2	3.75	D	c	1395		0.38
4	4.22	N	c	1778		0.58
6	4.90	S	c	1486		1.12
7	5.14	Q	c	1915		0.99
8	5.47	T	c	706		0.70
9	5.78	G	c	3661		2.10
10	6.22	E	c	1868		1.01
14	7.90	H	c	843		0.70
15	8.25	A	c	2153		1.34
19	9.66	S'	c	125		0.65
23	10.28	R	c	584		0.82
24	10.63	Y	c	755		0.48
32	13.08	P	c	1026		0.66
34	13.79	M	c	444		0.29
35	14.12	V	c	1451		0.86
39	15.35	dptu		49037		38.38
42	16.69	dpu		1521		1.56
43	17.07	F	c	1185		0.71
44	17.53	I	c	1179		0.86
45	17.77	K	c	1009		0.44
46	18.05	L	c	2084		1.23



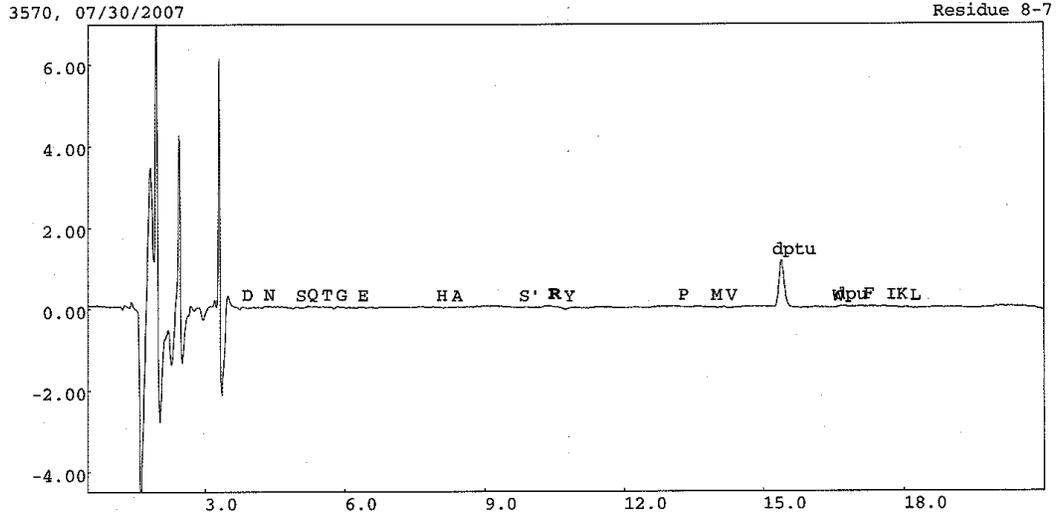
Residue 6, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol Ht
2	3.85	D	c	12945	3.57
4	4.31	N	c	5627	1.84
6	5.00	S	c	1447	1.09
7	5.25	Q	c	1684	0.87
8	5.57	T	c	831	0.82
9	5.88	G	c	3138	1.80
10	6.34	E	c	1845	0.99
16	8.04	H	c	464	0.39
17	8.35	A	c	2114	1.31
22	9.77	S'	c	411	2.14
24	10.43	R	c	695	0.97
25	10.74	Y	c	1880	1.18
29	13.17	P	c	1042	0.67
31	13.89	M	c	474	0.30
32	14.21	V	c	2058	1.22
35	15.43	dptu	c	48649	38.07
38	16.45	W	c	87	0.04
39	16.75	dpu	c	1581	1.62
40	17.14	F	c	1055	0.63
41	17.60	I	c	1291	0.95
42	17.83	K	c	1105	0.48
43	18.12	L	c	2348	1.39



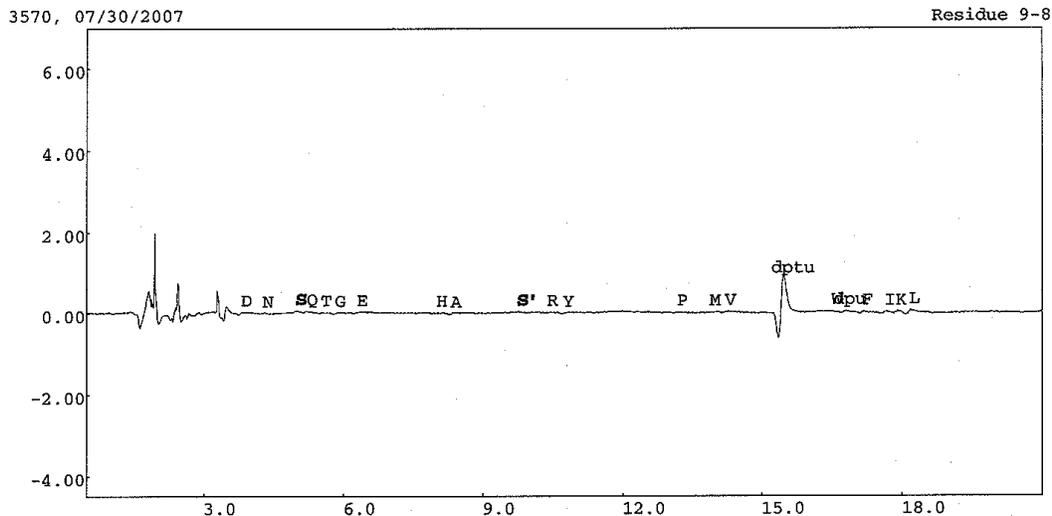
Residue 7, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol	Ht
2	3.77	D	c	2016		0.56
4	4.25	N	c	3781		1.24
6	4.93	S	c	2971		2.24
7	5.18	Q	c	2711		1.41
8	5.50	T	c	1650		1.63
9	5.81	G	c	3342		1.92
10	6.27	E	c	2032		1.10
15	7.98	H	c	187		0.16
16	8.30	A	c	2964		1.84
23	9.73	S'	c	478		2.49
25	10.38	R	c	646		0.90
26	10.70	Y	c	1866		1.17
34	13.15	P	c	991		0.64
35	13.89	M	c	324		0.21
36	14.20	V	c	1746		1.04
39	15.43	dptu		55576		43.49
41	16.45	W	c	89		0.04
42	16.76	dpu		1452		1.48
43	17.15	F	c	1031		0.62
44	17.61	I	c	1203		0.88
45	17.86	K	c	811		0.35
46	18.13	L	c	2127		1.26



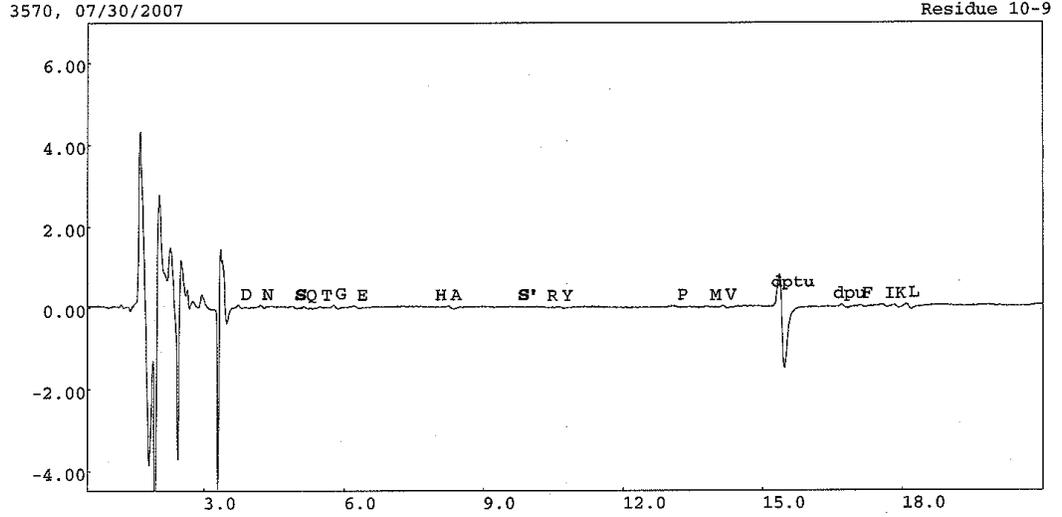
Residue 8, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol Ht
2	3.78	D	c	7012	1.93
4	4.25	N	c	4635	1.52
6	4.93	S	c	1486	1.12
7	5.19	Q	c	1723	0.90
8	5.50	T	c	858	0.85
9	5.82	G	c	2751	1.58
11	6.27	E	c	1744	0.94
15	7.96	H	c	368	0.31
16	8.30	A	c	2953	1.83
22	9.73	S'	c	236	1.23
23	10.36	R	c	954	1.33
24	10.70	Y	c	1465	0.92
32	13.16	P	c	975	0.63
35	13.88	M	c	382	0.25
36	14.20	V	c	1753	1.04
41	15.42	dptu	c	63156	49.43
43	16.48	W	c	58	0.03
44	16.76	dpu	c	1723	1.76
45	17.14	F	c	1169	0.70
46	17.61	I	c	1383	1.01
47	17.85	K	c	972	0.42
48	18.13	L	c	2176	1.29



Residue 9, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol Ht
2	3.79	D	c	10495	2.90
4	4.26	N	c	10770	3.52
6	4.94	S	c	7324	5.52
7	5.19	Q	c	6346	3.30
8	5.51	T	c	3620	3.58
9	5.82	G	c	3838	2.20
10	6.28	E	c	1718	0.93
15	8.00	H	c	192	0.16
16	8.30	A	c	2606	1.62
19	9.72	S'	c	365	1.90
21	10.37	R	c	934	1.31
22	10.71	Y	c	1301	0.82
30	13.16	P	c	986	0.64
32	13.88	M	c	393	0.25
33	14.21	V	c	1736	1.03
37	15.43	dptu	c	66548	52.08
39	16.49	W	c	73	0.04
40	16.77	dpu	c	1709	1.75
41	17.16	F	c	1099	0.66
42	17.63	I	c	1258	0.92
43	17.87	K	c	945	0.41
44	18.15	L	c	2180	1.29



Residue 10, Interpolated baseline

Peak No	RT	Peak ID	Type	Height	Pmol	Ht
2	3.78	D	c	2669	0.74	
4	4.25	N	c	2557	0.84	
6	4.93	S	c	1464	1.10	
7	5.18	Q	c	1564	0.81	
8	5.51	T	c	788	0.78	
9	5.81	G	c	2589	1.49	
10	6.27	E	c	1598	0.86	
15	7.96	H	c	147	0.12	
17	8.29	A	c	2478	1.54	
21	9.73	S'	c	525	2.73	
22	10.38	R	c	771	1.08	
23	10.70	Y	c	995	0.63	
28	13.15	P	c	989	0.64	
30	13.87	M	c	200	0.13	
31	14.20	V	c	1567	0.93	
34	15.42	dptu		60859	47.63	
37	16.75	dpu		1649	1.69	
38	17.14	F	c	1102	0.66	
39	17.61	I	c	1338	0.98	
40	17.85	K	c	922	0.40	
41	18.12	L	c	2092	1.24	

Uncorr

	A	D	E	F	G	H	I	K	L
1	3.64	2.32	1.49	0.44	2.84	0.65	0.56	0.58	0.91
2	1.45	1.45	0.88	0.58	1.98	0.67	0.83	0.50	1.09
3	1.39	0.64	0.89	2.25	1.88	0.57	0.89	0.46	2.00
4	1.40	0.55	1.12	1.27	2.20	1.11	1.08	0.52	1.50
5	1.34	0.38	1.01	0.71	2.10	0.70	0.86	0.44	1.23
6	1.31	3.57	0.99	0.63	1.80	0.39	0.95	0.48	1.39
7	1.84	0.56	1.10	0.62	1.92	0.16	0.88	0.35	1.26
8	1.83	1.93	0.94	0.70	1.58	0.31	1.01	0.42	1.29
9	1.62	2.90	0.93	0.66	2.20	0.16	0.92	0.41	1.29
10	1.54	0.74	0.86	0.66	1.49	0.12	0.98	0.40	1.24

	M	N	P	Q	R	S	S'	T	V
1	0.53	2.86	0.71	2.54	1.52	5.25	2.50	3.13	2.90
2	0.10	1.03	0.74	0.53	0.91	1.25	0.00	0.81	0.90
3	0.16	0.72	0.82	0.56	0.85	1.18	1.92	0.69	0.94
4	0.63	1.15	0.79	1.25	0.99	2.16	1.88	1.59	0.97
5	0.29	0.58	0.66	0.99	0.82	1.12	0.65	0.70	0.86
6	0.30	1.84	0.67	0.87	0.97	1.09	2.14	0.82	1.22
7	0.21	1.24	0.64	1.41	0.90	2.24	2.49	1.63	1.04
8	0.25	1.52	0.63	0.90	1.33	1.12	1.23	0.85	1.04
9	0.25	3.52	0.64	3.30	1.31	5.52	1.90	3.58	1.03
10	0.13	0.84	0.64	0.81	1.08	1.10	2.73	0.78	0.93

	W	Y
1	0.18	0.64
2	0.00	0.48
3	0.00	0.48
4	0.00	0.51
5	0.00	0.48
6	0.04	1.18
7	0.04	1.17
8	0.03	0.92
9	0.04	0.82
10	0.00	0.63

Bkgnd

	A	D	E	F	G	H	I	K	L
1	2.35	2.08	0.38	0.00	0.91	0.00	0.00	0.06	0.00
2	0.12	0.95	0.00	0.00	0.05	0.03	0.01	0.00	0.00
3	0.02	0.00	0.00	1.62	0.00	0.00	0.05	0.00	0.78
4	0.00	0.00	0.08	0.64	0.27	0.59	0.22	0.04	0.26
5	0.00	0.00	0.00	0.07	0.18	0.26	0.00	0.00	0.00
6	0.00	2.05	0.00	0.00	0.00	0.00	0.04	0.03	0.14
7	0.31	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00
8	0.26	0.00	0.00	0.05	0.00	0.05	0.06	0.00	0.01
9	0.01	0.61	0.01	0.00	0.29	0.00	0.00	0.00	0.01
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	M	N	P	Q	R	S	S'	T	V
1	0.19	2.39	0.00	1.81	0.67	3.76	0.82	2.22	1.98
2	0.00	0.41	0.02	0.00	0.02	0.00	0.00	0.00	0.00
3	0.00	0.00	0.11	0.00	0.00	0.00	0.06	0.00	0.00
4	0.33	0.21	0.10	0.33	0.05	0.67	0.00	0.59	0.00
5	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00
6	0.03	0.59	0.00	0.00	0.00	0.00	0.02	0.00	0.23
7	0.00	0.00	0.00	0.32	0.00	0.76	0.28	0.54	0.03
8	0.00	0.00	0.00	0.00	0.28	0.00	0.00	0.00	0.02
9	0.02	1.80	0.00	2.10	0.23	4.04	0.00	2.43	0.00
10	0.00	0.00	0.01	0.00	0.00	0.00	0.26	0.00	0.00

	W	Y
1	0.12	0.21
2	0.00	0.00
3	0.00	0.00
4	0.00	0.00
5	0.00	0.00
6	0.01	0.51
7	0.02	0.46
8	0.01	0.16
9	0.03	0.00
10	0.00	0.00

Lag	A	D	E	F	G	H	I	K	L
1	2.35	2.08	0.38	0.00	0.91	0.00	0.00	0.06	0.00
2	0.12	0.95	0.00	0.00	0.05	0.03	0.01	0.00	0.00
3	0.02	0.00	0.00	1.27	0.00	0.00	0.05	0.00	0.78
4	0.00	0.00	0.08	0.64	0.27	0.59	0.22	0.04	0.26
5	0.00	0.00	0.00	0.07	0.18	0.26	0.00	0.00	0.00
6	0.00	2.05	0.00	0.00	0.00	0.00	0.04	0.03	0.14
7	0.31	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00
8	0.26	0.00	0.00	0.05	0.00	0.05	0.06	0.00	0.01
9	0.01	0.61	0.01	0.00	0.29	0.00	0.00	0.00	0.01
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	M	N	P	Q	R	S	S'	T	V
1	0.19	2.39	0.00	1.81	0.67	3.76	0.82	2.22	1.98
2	0.00	0.41	0.02	0.00	0.02	0.00	0.00	0.00	0.00
3	0.00	0.00	0.11	0.00	0.00	0.00	0.06	0.00	0.00
4	0.33	0.21	0.10	0.33	0.05	0.67	0.00	0.59	0.00
5	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00
6	0.03	0.59	0.00	0.00	0.00	0.00	0.02	0.00	0.23
7	0.00	0.00	0.00	0.32	0.00	0.76	0.28	0.54	0.03
8	0.00	0.00	0.00	0.00	0.22	0.00	0.00	0.00	0.02
9	0.02	1.80	0.00	2.10	0.23	4.04	0.00	2.43	0.00
10	0.00	0.00	0.01	0.00	0.00	0.00	0.26	0.00	0.00

	W	Y
1	0.12	0.21
2	0.00	0.00
3	0.00	0.00
4	0.00	0.00
5	0.00	0.00
6	0.01	0.51
7	0.02	0.46
8	0.01	0.16
9	0.03	0.00
10	0.00	0.00

REFERENCES

1. References for Chapter 1

1. Albeck A., Friedman N., Sheves M. and Ottolenghi M., Factors affecting the absorption maxima of acidic forms of bacteriorhodopsin. A study with artificial pigments, *Biophys J.* **56** (1989) 1259-65.
2. Aroti A., Leontidis E., Maltseva E. and Brezesinski G., Effects of Hofmeister anions on DPPC Langmuir monolayers at the air-water interface, *J Phys Chem B.* **108** (2004) 15238-15245.
3. Atamna-Ismaeel N., Sabehi G., Sharon I., Witzel K. P., Labrenz M., Jürgens K., Barkay T., Stomp M., Huisman J. and Beja O., Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems, *ISME J.* **2** (2008) **656-62**. Erratum in: *ISME J.* **4** (2010) 462.
4. Batchelor J. D., Olteanu A., Tripathy A. and Pielak G. J., Impact of Protein Denaturants and Stabilizers on Water Structure, *J. Am. Chem. Soc.*, **126** (2004) 1958–1961.
5. Béjà O., Aravind L., Koonin E. V., Suzuki M. T., Hadd A., Nguyen L. P., Jovanovich S. B., Gates C. M., Feldman R. A., Spudich J. L., Spudich E. N. and DeLong E. F., Bacterial rhodopsin: evidence for a new type of phototrophy in the sea, *Science* **289** (2000) 1902–1906.
6. Béjà O., Spudich E. N., Spudich J. L., Leclerc M. and DeLong E. F., Proteorhodopsin phototrophy in the ocean, *Nature* **411** (2001) 786–789.

7. Bergo V., Amsden J. J., Spudich E. N., Spudich J. L. and Rothschild K. J., Structural changes in the photoactive site of proteorhodopsin during the primary photoreaction, *Biochemistry*. **43** (2004) 9075-83.
8. Bielawski J. P., Dunn K. A., Sabehi G. and Béjà O., Darwinian adaptation of proteorhodopsin to different light intensities in the marine environment, *Proc Natl Acad Sci U S A*. **101** (2004) 14824-9.
9. Birge R. R., Nature of the primary photochemical events in rhodopsin and bacteriorhodopsin, *Biochim Biophys Acta*. **1016** (1990) 293-327.
10. Birge R. R., Protein-based computers, *Scientific American* **272** (1995) 90-95.
11. Blatz P. E., Mohler J. H. and Navangul H. V., Anion-induced wavelength regulation of absorption maxima of Schiff bases of retinal, *Biochemistry* **11** (1972) 848-55.
12. Blatz P. E. and Mohler J. H., Effect of selected hydrogen-bonding solvents on the absorption maxima of N-retinylidene-n-butylammonium salts, *Biochemistry* **11** (1972) 3240-3.
13. Bogomolni R. A., Stoeckenius W., Szundi I., Perozo E., Olson K. D. and Spudich J. L., Removal of transducer HtrI allows electrogenic proton translocation by sensory rhodopsin I, *Proc Natl Acad Sci U S A*. **91** (1994) 10188-92.
14. Bostrom M., Tavares F. W., Finet S., Skouri-Panet F., Tardieu A. and Ninham B. W., Why forces between proteins follow different Hofmeister series for pH above and below pI, *Biophysical Chemistry* **117** (2005) 217 – 224.
15. Braiman M. S., Mogi T., Marti T., Stern L. J., Khorana H. G. and Rothschild K. J., Vibrational spectroscopy of bacteriorhodopsin mutants: light-driven proton

- transport involves protonation changes of aspartic acid residues 85, 96, and 212, *Biochemistry* **27** (1988) 8516-20.
16. Brenner M. P., Bildsten L., Dyson F., Fortson N., Garwin R., Grober R., Hemley R., Hwa T., Joyce G., Katz J., *et al.* Engineering Microorganisms for Energy Production. Report JSR-05-300. Washington, DC: JASON, Department of Energy (2006).
 17. Chaiken J. and Birge R. R., Photochromic Materials Suitable For Cosmetic and Sunblocking Effects - Patent 6461594 (2002).
 18. Cherezov V., Rosenbaum D. M., Hanson M. A., Rasmussen S. G., Thian F. S., Kobilka T. S., Choi H. J., Kuhn P., Weis W. I., Kobilka B. K. and Stevens R. C., High-resolution crystal structure of an engineered human β 2-adrenergic G protein-coupled receptor, *Science* **318** (2007) 1258-1265.
 19. Chien, E. Y., Liu, W., Zhao, Q., Katritch, V., Han, G. W., Hanson, M. A., Shi, L., Newman, A. H., Javitch, J. A., Cherezov, V. and Stevens, R. C., Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist, *Science* **330** (2010) 1091-1095.
 20. Collins K. D. and Washabaugh M. W., The Hofmeister effect and the behaviour of water at interfaces, *Quarterly Reviews of Biophysics* **18** (1985) 323-422.
 21. de la Torre J. R., Christianson L. M., B ej a O., Suzuki M. T., Karl D. M., Heidelberg J. and DeLong E. F., Proteorhodopsin genes are distributed among divergent marine bacterial taxa, *Proc Natl Acad Sci U S A.* **100** (2003) 12830-5.
 22. De r, A. and Ramsden J. J., Evidence for loosening of a protein mechanism. *Naturwissenschaften.* **85** (1998) 353–355.

23. DeLong E. F. and Béjà O., The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times, *PLoS Biol.* **8** (2010) e1000359.
24. Deshpande A. and Sonar S., Bacterioopsin-triggered retinal biosynthesis is inhibited by bacteriorhodopsin formation in *Halobacterium salinarium*, *J Biol Chem.* **274** (1999) 23535-40.
25. Dioumaev A. K., Wang J. M., Bálint Z., Váró G. and Lanyi J. K., Proton transport by proteorhodopsin requires that the retinal Schiff base counterion Asp-97 be anionic, *Biochemistry* **42** (2003) 6582-7.
26. Dioumaev A. K., Brown L. S., Shih J., Spudich E. N., Spudich J. L. and Lanyi J. K., Proton transfers in the photochemical reaction cycle of proteorhodopsin, *Biochemistry* **41** (2002) 5348-58.
27. Fain G. L., Hardie R. and Laughlin S. B., Phototransduction and the evolution of photoreceptors, *Curr Biol.* **20** (2010) 114-24.
28. Friedrich T., Geibel S., Kalmbach R., Chizhov I., Ataka K., Heberle J., Engelhard M., Bamberg E., Proteorhodopsin is a light-driven proton pump with variable vectoriality, *J Mol Biol.* **321** (2002) 821-38.
29. Giovannoni S. J., Bibbs L., Cho J. C., Stapels M. D., Desiderio R., Vergin K. L., Rappé M. S., Laney S., Wilhelm L. J., Tripp H. J., Mathur E. J. and Barofsky D. F., Proteorhodopsin in the ubiquitous marine bacterium SAR11, *Nature* **438** (2005) 82-5.
30. Gómez-Consarnau L., Akram N., Lindell K., Pedersen A., Neutze R., Milton D. L., González J. M. and Pinhassi J., Proteorhodopsin phototrophy promotes survival of marine bacteria during starvation, *PLoS Biol.* **8** (2010) e1000358.

31. Gómez-Consarnau L., González J. M., Coll-Lladó M., Gourdon P., Pascher T., Neutze R., Pedrós-Alió C. and Pinhassi J., Light stimulates growth of proteorhodopsin-containing marine Flavobacteria, *Nature* **445** (2007) 210-3.
32. Gordeliy V. I., Labahn J., Moukhametzianov R., Efremov R., Granzin J., Schlesinger R., Büldt G., Savopol T., Scheidig A. J., Klare J. P. and Engelhard M., Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex, *Nature* **419** (2002) 484-7.
33. Gourdon P., Alfredsson A., Pedersen A., Malmerberg E., Nyblom M., Widell M., Berntsson R., Pinhassi J., Braiman M., Hansson O., Bonander N., Karlsson G. and Neutze R., Optimized in vitro and in vivo expression of proteorhodopsin: a seven-transmembrane proton pump, *Protein Expr Purif.* **58** (2008) 103-13.
34. Govindjee R., Misra S., Balashov S. P., Ebrey T. G., Crouch R. K. and Menick D. R., Arginine-82 regulates the pKa of the group responsible for the light-driven proton release in bacteriorhodopsin, *Biophys J.* **71** (1996) 1011-23.
35. Gurau M. C., Lim Soon-Mi, Ecasetllana E. T., Albertorio F., Kataoka S. and Cremer P. S., On the Mechanism of the Hofmeister Effect, *J. Am. Chem. Soc.* **126** (2004) 10522-10523.
36. Hampp N. A., Bacteriorhodopsin as a Photochromic Retinal Protein for Optical Memories, *Chem. Rev.* **100** (2000), 1755–1776.
37. Hampp N. A., Bacteriorhodopsin: mutating a biomaterial into an optoelectronic material, *Appl Microbiol Biotechnol.* **53** (2000) 633-9.
38. Heberle J., Proton transfer reactions across bacteriorhodopsin and along the membrane, *Biochim Biophys Acta.* **1458** (2000) 135-47.

39. Hillebrecht J. R., Galan J., Rangarajan R., Ramos L., McCleary K., Ward D. E., Stuart J. A. and Birge R. R., Structure, function, and wavelength selection in blue-absorbing proteorhodopsin, *Biochemistry*. **45** (2006) 1579-90.
40. Hoff W. D., Jung K. H. and Spudich J. L., Molecular mechanism of photosignaling by archaeal sensory rhodopsins, *Annu Rev Biophys Biomol Struct*. **26** (1997) 223-58.
41. Hofmeister F., Zur Lehre von der Wirkung der Salze, *Arch Exp Pathol Pharmacol* **24** (1888), 247-260. [Title translation: About the science of the effect of salts.]
42. Hohmann-Marriott M. F. and Blankenship R. E., Evolution of photosynthesis, *Annu Rev Plant Biol*. **62** (2011) 515-48.
43. Hubbard R. and Kropf A., The action of light on rhodopsin, *Proc Natl Acad Sci U S A*. **44** (1958) 130-9.
44. Jaakola V. P., Griffith M. T., Hanson M. A., Cherezov V., Chien E. Y., Lane J. R., Ijzerman A. P. and Stevens R. C., The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist, *Science* **322** (2008) 1211-7.
45. Jung J. Y., Choi A. R., Lee Y. K., Lee H. K. and Jung K. H., Spectroscopic and photochemical analysis of proteorhodopsin variants from the surface of the Arctic Ocean, *FEBS Lett*. **582** (2008) 1679-84.
46. Kelemen B. R., Du M. and Jensen R. B., Proteorhodopsin in living color: diversity of spectral properties within living bacterial cells, *Biochim Biophys Acta*. **1618** (2003) 25-32..

47. Kim S. Y., Waschuk S. A., Brown L. S. and Jung K. H., Screening and characterization of proteorhodopsin color-tuning mutations in *Escherichia coli* with endogenous retinal synthesis, *Biochim Biophys Acta.* **1777** (2008) 504-13.
48. Kimura H., Young C. R., Martinez A. and Delong E. F., Light-induced transcriptional responses associated with proteorhodopsin-enhanced growth in a marine flavobacterium, *ISME J.* 2011 Apr 7. [Epub ahead of print]
49. Kolbe M., Besir H., Essen L. O. and Oesterhelt D., Structure of the light-driven chloride pump halorhodopsin at 1.8 Å resolution, *Science* **288** (2000) 1390-6.
50. Kralj J. M., Bergo V. B., Amsden J. J., Spudich E. N., Spudich J. L. and Rothschild K. J., Protonation state of Glu142 differs in the green- and blue-absorbing variants of proteorhodopsin., *Biochemistry* **47** (2008) 3447-53.
51. Kralj J. M., Spudich E. N., Spudich J. L. and Rothschild K. J., Raman spectroscopy reveals direct chromophore interactions in the Leu/Gln105 spectral tuning switch of proteorhodopsins, *J Phys Chem B.* **112** (2008)11770-6.
52. Krebs R. A., Dunmire D. A., Partha R. and Braiman M. S., Resonance Raman characterization of proteorhodopsin's chromophore environment, *J. Phys. Chem., B* **107** (2003) 7877–7883.
53. Kunz W., Henle J. and Ninham B.W., ‘Zur Lehre von der Wirkung der Salze’ (about the science of the effect of salts): Franz Hofmeister’s historical papers *Current Opinion in Colloid and Interface Science* **9** (2004) 19-37.
54. Lane N., Allen J. F. and Martin W., How did LUCA make a living? Chemiosmosis in the origin of life, *Bioessays.* **32** (2010) 271-80.

55. Lanyi J. K., Proton transfers in the bacteriorhodopsin photocycle, *Biochim Biophys Acta*. **1757** (2006) 1012-8.
56. Lundstrom K., Structural genomics for membrane proteins, *Cell. Mol. Life Sci.* **63** (2006) 2597–2607.
57. Man D., Wang W., Sabehi G., Aravind L., Post A. F., Massana R., Spudich E. N., Spudich J. L. and Bèjà O., Diversification and spectral tuning in marine proteorhodopsins, *EMBO J.* **22** (2003) 1725-31.
58. Man-Aharonovich D., Sabehi G., Sineshchekov O. A., Spudich E. N., Spudich J. L. and Bèjà O., Characterization of RS29, a blue-green proteorhodopsin variant from the Red Sea., *Photochem Photobiol Sci.* **3** (2004) 459-62.
59. Martinez A., Bradley A. S., Waldbauer J. R., Summons R. E. and DeLong E. F., Proteorhodopsin photosystem gene expression enable photophosphorylation in a heterologous host, *Proc. Natl. Acad. Sci. U. S. A.* **104** (2007) 5590–5595.
60. Morris R. M., Rappe M. S., Connon S. A., Vergin K. L., Sjobold W. A., Carlson C. A., and Giovannoni S. J., SAR11 clade dominates ocean surface bacterioplankton communities, *Nature* **420** (2002) 806-810.
61. Nakanishi K., Balogh-Nair V., Gawinowicz M. A., Arnaboldi M., Motto M. and Honig B., Double point charge model for visual pigments; evidence from dihydorhodopsins, *Photochem Photobiol.* **29** (1979) 657-60.
62. Narang R., Color change technology and the emergence of interactive cosmetics, 2008. www.slideshare.net/rajatnarang
63. NCBI database: <http://www.ncbi.nlm.nih.gov>

64. Omta A.W. and Bakker H.J., Negligible effect of ions on the hydrogen-bond structure in liquid water, *Science* **301** (2003) 347-349.
65. Ottolenghi M. and Sheves M., Synthetic retinals as probes for the binding site and photoreactions in rhodopsins, *J Membr Biol.* **112** (1989) 193-212.
66. Palczewski K., Kumasaka T., Hori T., Behnke C. A., Motoshima H., Fox B. A., Le Trong I., Teller D. C., Okada T. and Stenkamp R. E., Yamamoto M. and Miyano M., Crystal structure of rhodopsin: A G protein-coupled receptor, *Science* **289** (2000) 739-45.
67. Partha R. and Braiman M. S., A rapid and inexpensive method for the purification of proteorhodopsin, *U.S. Provisional Patent Application 60/485*, **272** (2003).
68. Partha R., Krebs R., Caterino T. L. and Braiman M. S., Weakened coupling of conserved arginine to the proteorhodopsin chromophore and its counterion implies structural differences from bacteriorhodopsin, *Biochim Biophys Acta* **1708** (2005) 6-12.
69. Programme UND, Affairs UNDoEaS, Council WE: *World Energy Assessment: Energy and the Challenge of Sustainability* (2000).
70. Ranaghan M. J., Shima S., Ramos L., Poulin D. S., Whited G., Rajasekaran S., Stuart J. A., Albert A. D. and Birge R. R., Photochemical and thermal stability of green and blue proteorhodopsins: implications for protein-based bioelectronic devices, *J Phys Chem B.* **114** (2010) 14064-70.
71. Rappé M. S., Connon S. A., Vergin K. L. and Giovannoni S. J., Cultivation of the ubiquitous SAR11 marine bacterioplankton clade, *Nature* **418** (2002) 630-3.

72. Rasmussen S. G., Choi H. J., Rosenbaum D. M., Kobilka T. S., Thian F. S., Edwards P. C., Burghammer M., Ratnala V. R., Sanishvili R., Fischetti R. F., Schertler G. F., Weis W. I. and Kobilka B. K., Crystal structure of the human β 2 adrenergic G-protein-coupled receptor, *Nature* **450** (2007) 383-387.
73. Rothschild K. J., He Y. W., Gray D., Roepe P. D., Pelletier S. L., Brown R. S. and Herzfeld J., Fourier transform infrared evidence for proline structural changes during the bacteriorhodopsin photocycle, *Proc Natl Acad Sci U S A.* **24** (1989) 9832-5.
74. Rusch D. B., Halpern A. L., Sutton G., Heidelberg K. B., Williamson S., Yooseph S., Wu D., Eisen J. A., Hoffman J. M., Remington K., Beeson K., Tran B., Smith H., Baden-Tillson H., Stewart C., Thorpe J., Freeman J., Andrews-Pfannkoch C., Venter J. E., Li K., Kravitz S., Heidelberg J. F., Utterback T., Rogers Y. H., Falcón L. I., Souza V., Bonilla-Rosso G., Eguiarte L. E., Karl D. M., Sathyendranath S., Platt T., Bermingham E., Gallardo V., Tamayo-Castillo G., Ferrari M. R., Strausberg R. L., Neilson K., Friedman R., Frazier M, and Venter J. C., The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific, *PLoS Biol.* **5** (2007) 398-431 (e77).
75. Sabei G., Massana R., Bielawski J. P., Rosenberg M., Delong E. F. and Béjà O., Novel Proteorhodopsin variants from the Mediterranean and Red Seas, *Environ Microbiol.* **5** (2003) 842-9.
76. Seehra J. S. and Khorana H. G., Bacteriorhodopsin precursor. Characterization and its integration into the purple membrane, *J Biol Chem.* **259** (1984) 4187-93.

77. Seidel R., Scharf B., Gautel M., Kleine K., Oesterhelt D. and Engelhard M., The primary structure of sensory rhodopsin II: a member of an additional retinal protein subgroup is coexpressed with its transducer, the halobacterial transducer of rhodopsin II, *Proc Natl Acad Sci U S A.* **92** (1995) 3036-40.
78. Shimono K., Hayashi T., Ikeura Y., Sudo Y., Iwamoto M. and Kamo N., Importance of the broad regional interaction for spectral tuning in *Natronobacterium pharaonis* phoborhodopsin (sensory rhodopsin II), *J Biol Chem.* **278** (2003) 23882-9.
79. Spudich J. L. and Bogomolni R. A., Mechanism of colour discrimination by a bacterial sensory rhodopsin, *Nature* **312** (1984) 509-13.
80. Spudich J. L. and Bogomolni R. A., Spectroscopic discrimination of the three rhodopsinlike pigments in Halobacterium halobium membranes, *Biophys J.* **43** (1983) 243-6.
81. Spudich J. L. and Jung K. H., Microbial rhodopsins: phylogenetic and functional diversity, Handbook of Photosensory Receptors, Edited by W. R. Briggs, J. L. Spudich, Wiley-VCH Verlag GmbH and Co. KGaA, Weinheim, (2005) 1-23.
82. Spudich J. L., The multitalented microbial sensory rhodopsins, *Trends in Microbiology* **14** (2006) 480-487.
83. Stenkamp R. E., Teller D. C. and Palczewski K, Crystal Structure of Rhodopsin: A G-Protein-Coupled Receptor, *Chembiochem.* **3** (2002) 963-7.
84. Stephen W., Membrane protein structure database,
http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html

85. Stevens T. J. and Arkin I. T., Do more complex organisms have a greater proportion of membrane proteins in their genomes?, *Proteins* **39** (2000), 417–420.
- 86. Stryer L, Biochemistry, Fourth Edition, W. H. Freeman and Company, New York.**
87. Subramaniam S. and Henderson R., Molecular mechanism of vectorial proton translocation by bacteriorhodopsin, *Nature* **406** (2000) 653-7.
88. Takahashi T., Yan B., Mazur P., Derguini F., Nakanishi K. and Spudich J. L., Color regulation in the archaeobacterial phototaxis receptor phoborhodopsin (sensory rhodopsin II), *Biochemistry* **29** (1990) 8467-74.
89. Váró G., Analogies between halorhodopsin and bacteriorhodopsin, *Biochim Biophys Acta.* **1460** (2000) 220-9.
90. Váró G., Brown L. S., Lakatos M. and Lanyi J. K., Characterization of the photochemical reaction cycle of proteorhodopsin, *Biophys J.* **84** (2003) 1202-7.
91. Venter J. C., Remington K., Heidelberg J. F., Halpern A. L., Rusch D., Eisen J. A., Wu D., Paulsen I., Nelson K. E., Nelson W., Fouts D. E., Levy S., Knap A. H., Lomas M. W., Nealson K., White O., Peterson J., Hoffman J., Parsons R., Baden-Tillson H., Pfannkoch C., Rogers Y. H. and Smith H. O., Environmental genome shotgun sequencing of the Sargasso Sea, *Science* **304** (2004) 66-74.
92. Vogel R. and Siebert F., Conformations of the Active and Inactive States of Opsin, *J. Biol. Chem.* **276** (2001), 38487–38493.
93. Walter J. M., Greenfield D. and Liphardt J., Potential of light-harvesting proton pumps for bioenergy applications, *Curr Opin Biotechnol.* **21** (2010) 265-70.

94. Walter J. M., Greenfield D., Bustamante C. and Liphardt J., Light-powering *Escherichia coli* with proteorhodopsin, *Proc Natl Acad Sci U S A.* **104** (2007) 2408-12.
95. Wang W. W., Sineshchekov O. A., Spudich E. N. and Spudich J. L., Spectroscopic and photochemical characterization of a deep ocean proteorhodopsin, *J Biol Chem.* **278** (2003) 33985-91.
96. Warne T., Serrano-Vega M. J., Baker J. G., Moukhametzianov R., Edwards P. C., Henderson R., Leslie A. G., Tate C.G. and Schertler G. F., Structure of a beta1-adrenergic G-protein-coupled receptor, *Nature* **454** (2008), 486-491.
97. Wise K. J., Gillespie N. B., Stuart J. A., Krebs M. P. and Birge R. R., Optimization of bacteriorhodopsin for bioelectronic devices, *Trends Biotechnol.* **20** (2002) 387-94.
98. Wu B., Chien E. Y., Mol C. D., Fenalti G., Liu W., Katritch V., Abagyan R., Brooun A., Wells P., Bi F. C., Hamel D. J., Kuhn P., Handel T. M., Cherezov V. and Stevens R.C., Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists, *Science* **330** (2010), 1066-71.
99. Xi B., Tetley W. C., Marcy D. L., Zhong C., Whited G. Birge R. R. and Stuart J. A., Evaluation of blue and green absorbing proteorhodopsins as holographic materials, *J Phys Chem B.* **112** (2008) 2524-32.
100. Xiao Y., Hutson M. S., Belenky M., Herzfeld J., and Braiman M. S., Role of arginine-82 in fast proton release during the bacteriorhodopsin photocycle: a time-resolved FT-IR study of purple membranes containing ¹⁵N-labeled arginine, *Biochemistry* **43** (2004) 12809-18.

101. Yan B., Spudich J. L., Mazur P., Vunnam S., Derguini F. and Nakanishi K., Spectral tuning in bacteriorhodopsin in the absence of counterion and coplanarization effects, *J Biol Chem.* **270** (1995) 29668-70.
102. Yao V. J. and Spudich J. L., Primary structure of an archaebacterial transducer, a methyl-accepting protein associated with sensory rhodopsin I, *Proc Natl Acad Sci U S A.* **891** (1992) 11915-9.
103. Zaslavskaja L. A., Lippmeier J. C., Shih C., Ehrhardt D., Grossman A. R. and Apt K. E., Trophic conversion of an obligate photoautotrophic organism through metabolic engineering, *Science* **292** (2001) 2073-5.
104. Zeisel D. and Hampp N. A., Spectral relationship of light-induced refractive index and absorption changes in bacteriorhodopsin films containing wildtype BRWT and the variant BRD96N, *J. Phys. Chem.*, **96** (1992), (19) 7788–7792.
105. Zhang C., Ku, C. Y., Song, Q. W., Gross, R. B. and Birge, R. R. *Opt. Lett.* **19** (1994) 1409–1411.
106. Zhang Y. and Cremer P. S., Interactions between macromolecules and ions: the Hofmeister series, *Current Opinion in Chemical Biology* **10** (2006), 658–663.

2. References for chapter 2

1. O. Bèjà, L. Aravind, E.V. Koonin, M.T. Suzuki, A. Hadd, L.P. Nguyen, S.B. Jovanovich, C.M. Gates, R.A. Feldman, J.L. Spudich, E.N. Spudich and E.F. DeLong, Bacterial rhodopsin: evidence for a new type of phototrophy in the sea, *Science* **289** (2000), pp. 1902–1906.
2. O. Bèjà, E.N. Spudich, J.L. Spudich, M. Leclerc and E.F. DeLong, Proteorhodopsin phototrophy in the ocean, *Nature* **411** (2001), pp. 786–789.
3. R.M. Morris, M.S. Rappe, S.A. Connon, K.L. Vergin, W.A. Sjobold, C.A. Carlson, and S.J. Giovannoni, SAR11 clade dominates ocean surface bacterioplankton communities, *Nature* **420** (2002), pp. 806-810.
4. G. Sabehi, R. Massana, J.P. Bielawski, M. Rosenberg, E.F. DeLong and O. Bèjà, Novel proteorhodopsin variants from the Mediterranean and Red seas, *Environ. Microbiol.* **5** (2003), pp. 842–849.
5. G. Sabehi, O. Bèjà, M.T. Suzuki, C.M. Preston and E.F. DeLong, Different SAR86 subgroups harbour divergent proteorhodopsins. *Environ. Microbiol.* **6** (2004), pp. 903-910.
6. J.Y. Jung, A.R. Choi, Y.K. Lee, H.K. Lee and K.H. Jung, Spectroscopic and photochemical analysis of proteorhodopsin variants from the surface of the Arctic Ocean, *FEBS Lett.* **582** (2008), pp. 1679-1684.
7. J.R. de la Torre, L.M. Christianson, O. Bèjà, M.T. Suzuki, D.M. Karl, J. Heidelberg and E.F. DeLong, Proteorhodopsin genes are distributed among

- divergent marine bacterial taxa, *Proc Natl Acad Sci U S A.* **100** (2003) pp. 12830-12835.
8. J.C. Venter, K. Remington, J.F. Heidelberg, A.L. Halpern, D. Rusch, J.A. Eisen, D. Wu, I. Paulsen, K.E. Nelson, W. Nelson, D.E. Fouts, S. Levy, A.H. Knap, M.W. Lomas, K. Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y.H. Rogers and H.O. Smith, Environmental genome shotgun sequencing of the Sargasso Sea, *Science* **304** (2004), pp. 66–74.
 9. D.B. Rusch, A.L. Halpern, G. Sutton, K.B. Heidelberg, S. Williamson, S. Yooseph, D. Wu, J.A. Eisen, J.M. Hoffman, K. Remington, K. Beeson, B. Tran, H. Smith, H.B. Tillson, C. Stewart, J. Thorpe, J. Freeman, C.A. Pfannkoch, J.E. Venter, K. Li, S. Kravits, J.F. Heidelberg, T. Utterback, Y.H. Rogers, L.I. Falcon, V. Souza, G.B. Rosso, L.E. Eguarte, D.M. Karl, S. Sathyendranath, T. Platt, E. Bermingham, V. Gallardo, G.T. Castillo, M.R. Ferrari, R.L. Strausberg, K. Nealson, R. Friedman, M. Frazier and J.C. Venter, The sorcerer II global ocean sampling expedition: northwest Atlantic through eastern tropical Pacific, *PLoS Biol.* **5** (3) (2007), p. e77.
 10. N. Atamna-Ismaeel, G. Sabehi, I. Sharon, K.P. Witzel, M. Labrenz, K. Jürgens, T. Barkay, M. Stomp, J. Huisman and O. Beja, Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems, *ISME J.* **2** (2008) pp. 656-662.
 11. J.M. Walter, D. Greenfield, C. Bustamante and J. Liphardt, Light-powering *E. coli* with proteorhodopsin, *Proc. Natl. Acad. Sci. USA* **104** (2007), pp. 2408–2412.

12. L. Gomez-Consarnau, J.M. González, M.C. Lladó, P. Gourdon, T. Pascher, R. Neutze, C.P. Alió and J. Pinhassi, Light stimulates growth of proteorhodopsin-containing marine flavobacteria, *Nature* **445** (2007), pp. 210–213.
13. Martinez A., A.S. Bradley, J.R. Waldbauer, R.E. Summons and E.F. Delong, Proteorhodopsin photosystem gene expression enable photophosphorylation in a heterologous host, *Proc. Natl. Acad. Sci. U. S. A.* **104** (2007), pp. 5590–5595.
14. K.R. Albe, M.H. Butler and B.E. Wright, Cellular concentrations of enzymes and their substrates, *J.Theor. Biol.* **143** (1990), pp 163-195.
15. B.D. Bennett, E.H. Kimball, M. Gao, R. Osterhout, S.J.V. Dien and J.D. Rabinowitz, Absolute Metabolite Concentrations and Implied Enzyme Active Site Occupancy in *Escherichia coli*, *Nature Chemical Biology* **5** (2009), 593-599.
16. R. Partha, R. Krebs, T.L. Caterino and M.S. Braiman, Weakened coupling of conserved arginine to the proteorhodopsin chromophore and its counterion implies structural differences from bacteriorhodopsin, *Biochim Biophys Acta* **1708** (2005) pp. 6-12.
17. R. Partha and M.S. Braiman, Rapid and inexpensive method for the purification of proteorhodopsin, US Patent # 7517968 (April 14, 2009).
18. R.A. Krebs, U.Alexiev, R. Partha, A.M. DeVita and M.S. Braiman, Detection of fast light-activated H⁺ release and M intermediate formation from proteorhodopsin, *BMC Physiol.* **2** (2002) pp. 2:5.
19. P. Gourdon, A. Alfredsson, A. Pedersen, E. Malmerberg, M. Nyblom, M. Widell, R. Berntsson, J. Pinhassi, M. Braiman, O. Hansson, N. Bonander, G. Karlsson

- and R. Neutze, Optimized in vitro and in vivo expression of proteorhodopsin: a seven-transmembrane proton pump, *Protein Expr. Purif.* **58** (2008) pp. 103-113.
20. U.A. Hellmich, N. Pflieger and C. Glaubit, ¹⁹F-MAS NMR on Proteorhodopsin: Enhanced Protocol for Site-Specific Labeling for General Application to Membrane Proteins, *Photochemistry Photobiology* **85** (2009) pp 535-9.
21. C. Combet, M. Jambon, G. Deléage and C. Geourjon, Geno3D an automated protein modelling Web server, *Bioinformatics* **18** (2002) pp. 213-214.
22. E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng and T.E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis, *J Comput Chem.* **25** (2004) pp. 1605-1612 (supported by NIH P41 RR-01081).
23. S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller and Lipman DJ, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* **25** (1997) pp. 3389-3402.
24. R.D. Page, TREEVIEW: An application to display phylogenetic trees on personal computers, *Computer Applications in the Biosciences* **12** (1996) pp. 357-358.
25. F. P. Takacs, T. I. Matula and R. A. Macleod, Nutrition and Metabolism of marine bacteria, XIII. Intracellular concentrations of sodium and potassium ions in a marine pseudomonad, *J. Bacteriol.* **87** (1964) pp 510-18.
26. S. Y. Kim, S. A. Waschuk, L.S. Brown and K.H. Jung, Screening and characterization of proteorhodopsin color-tuning mutations in *Escherichia coli* with endogenous retinal synthesis, *Biochim. Biophys. Acta*, **1777** (2008), pp 504-13.

27. T. Friedrich, S. Geibel, R. Kalmbach, I. Chizhoy, K. Ataka, J. Herbie, M. Engelbard and E. Bamberg, Proteorhodopsin is a light-driven proton pump with variable vectoriality, *Journal of Molecular Biology* **321** (2002) pp 821-38.
28. Dioumaev A. K., J.M. Wang, Z.Balint, G.Varo and J.K.Lanyi, Proton transport by proteorhodopsin requires that the retinal Schiff base counterion Asp-97 be anionic, *Biochemistry* **42** (2003) pp 6582-7.
29. P.P. Sheridan, K.H. Freeman and J.E. Brenchley, Estimated Minimal Divergence Times of the Major Bacterial and Archaeal Phyla, *Geomicrobiology Journal* **20** (2003) pp. 1–14.
30. D.R. Bond and J.B. Russell, Relationship between intracellular phosphate, proton motive force, and rate of nongrowth energy dissipation (Energy Spilling) in *Streptococcus bovis* JB1, *Appl. Environ. Microbiol.*, **64** (1998) pp 976-981.
31. T. P. Ikeda, A. E. Shauger and S. Kustu, Salmonella typhimurium apparently perceives external nitrogen limitation as internal glutamine limitation, *J. Mol. Biol.* **259** (1996) pp 589-607.
32. D. McLaggan, J. Naprstek, E. T. Buurmann and W. Epstein, Interdependence of K^+ and glutamate accumulation during osmotic adaptation of *Escherichia coli*, *J. Biol. Chem.* **269** (1994) pp 1911-17.
33. Hofmeister F., Zur Lehre von der Wirkung der Salze, *Arch Exp Pathol Pharmacol* **24** (1888), 247-260. [Title translation: About the science of the effect of salts.]
34. W. Kunz, J. Henle and B.W. Ninham, ‘Zur Lehre von der Wirkung der Salze’ (about the science of the effect of salts): Franz Hofmeister’s historical papers *Current Opinion in Colloid and Interface Science* **9** (2004) 19-37.

35. Ries-Kautt M. M. and Ducruix A. F., Relative effectiveness of various ions on the solubility and crystal growth of lysozyme, *J. Biol. Chem.* **264** (1989) 745-8.
36. Schwierz N., Horinek D., and Netz R. R., Reversed anionic Hofmeister series: the interplay of surface charge and surface polarity, *Langmuir.* **26** (2010) 7370-9.
37. Boström M., Tavares F. W., Bratko D., and Ninham B. W., Specific ion effects in solutions of globular proteins: comparison between analytical models and simulation, *J Phys Chem B.* **109** (2005) 24489-94.
38. Lo Nostro, P., Fratoni, L., Ninham, B. W., and Baglioni, P., Water absorbency by wool fibers: Hofmeister effect, *Biomacromolecules* **3** (2002) 1217-1224.
39. Ganguly P., Schravendijk P., Hess B., and van der Vegt N. F., Ion pairing in aqueous electrolyte solutions with biologically relevant anions, *J Phys Chem B.* **115** (2011) 3734-9.
40. Zangi R., Hagen M., and Berne B. J., Effect of ions on the hydrophobic interaction between two plates, *J Am Chem Soc.* **129** (2007) 4678-86.
41. Pal S., and Müller-Plathe F., Molecular dynamics simulation of aqueous NaF and NaI solutions near a hydrophobic surface, *J Phys Chem B.* **109** (2005) 6405-15.
42. Lund M., Vacha R., and Jungwirth P., Specific ion binding to macromolecules: effects of hydrophobicity and ion pairing, *Langmuir.* **24** (2008) 3387-91.
43. L. Kelemen, P. Galajda, S. Száraz and P. Ormos, Chloride ion binding to bacteriorhodopsin at low pH: an infrared spectroscopic study, *Biophys J.* **76** (1999) pp. 1951-8.

44. K. Mevorat-Kaplan, V. Brumfeld, M. Engelhard and M. Sheves, Effect of anions on the photocycle of halorhodopsin. Substitution of chloride with formate anion, *Biochemistry* **44** (2005) pp. 14231-7.
45. R. Vogel, G.B. Fan, F. Siebert and M. Sheves, Anions stabilize a Metarhodopsin II-like photoproduct with a protonated Schiff base. *Biochemistry* **40** (2001), pp. 13342–13352.
46. Y. Sharaabi, V. Brumfeld and M. Sheves, Binding of anions to proteorhodopsin affects the Asp97 pK(a). *Biochemistry*, **49** (2010) pp. 4457-65.

3. References for Chapter 3

1. Stevens T.J. and Arkin I.T., Do more complex organisms have a greater proportion of membrane proteins in their genomes? *Proteins* **39** (2000) 417–420.
2. Stephen W., Membrane protein structure database,
http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html
3. Hofmeister F., Zur Lehre von der Wirkung der Salze, *Arch Exp Pathol Pharmacol* **24** (1888), 247-260. [Title translation: About the science of the effect of salts.]
4. W. Kunz, J. Henle and B.W. Ninham, ‘Zur Lehre von der Wirkung der Salze’ (about the science of the effect of salts): Franz Hofmeister’s historical papers *Current Opinion in Colloid and Interface Science* **9** (2004) 19-37.
5. Collins K. D. and Washabaugh M. W., The Hofmeister effect and the behaviour of water at interfaces, *Q Rev Biophys.* **18** (1985) 323-422.
6. Shih Y. C., Prausnitz J. M. and Blanch H. W., Some characteristics of protein precipitation by salts, *Biotechnol Bioeng.* **40** (1992) 1155-64.
7. Vogel R. and Siebert F., Conformations of the Active and Inactive States of Opsin, *J. Biol. Chem.* **276** (2001) 38487–38493.
8. De´r, A. and J. J. Ramsden. Evidence for loosening of a protein mechanism. *Naturwissenschaften.* **85** (1998) 353–355.
9. Tate C. G., Practical considerations of membrane protein instability during purification and crystallization, Heterologous expression of membrane proteins, *Methods and Protocols*, Edited by Isabelle Mus-Veteau, Humana Press, (2010) 187-203.

10. Serrano-Vega M. J., Magnani F., Shibata Y, and Tate C. G. Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form, *Proc Natl Acad Sci U S A* **105** (2008) 877-82.
11. Toyoshima C., Nakasako M., Nomura H. and Ogawa H., Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution, *Nature* **405** (2000) 647-55.
12. Palczewski K., Kumasaka T., Hori T., Behnke C.A., Motoshima H., Fox B.A., Le Trong I., Teller D.C., Okada T., Stenkamp R.E., Yamamoto M. and Miyano M., Crystal structure of rhodopsin: A G protein-coupled receptor, *Science* **289** (2000) 739-45.
13. Stenkamp R. E., Teller D. C. and Palczewski K, Crystal Structure of Rhodopsin: A G-Protein-Coupled Receptor, *ChemBiochem.* **3** (2002) 963-7.
14. Partha R. and Braiman M.S., A rapid and inexpensive method for the purification of proteorhodopsin, U.S. Provisional Patent Application 60/485, **272** (2003).
15. Partha R., Krebs R., Caterino T.L. and Braiman M.S., Weakened coupling of conserved arginine to the proteorhodopsin chromophore and its counterion implies structural differences from bacteriorhodopsin, *Biochim Biophys Acta* **1708** (2005) 6-12.
16. Sharaabi Y., Brumfeld V. and Sheves M., Binding of anions to proteorhodopsin affects the Asp97 pK(a), *Biochemistry* **49** (2010) 4457-4465.
17. Vogel R., Fan G. B., Siebert F. and Sheves M., Anions stabilize a metarhodopsin II-like photoproduct with a protonated Schiff base, *Biochemistry* **40** (2001) 13342-52.

18. Hempelmann F., Hölper S., Verhoefen M. K., Woerner A. C., Köhler T., Fiedler S. A., Pflieger N., Wachtveitl J. and Glaubitz C., His75-Asp97 cluster in green proteorhodopsin, *J Am Chem Soc.* **133** (2011) 4645-54.
19. Ranaghan M. J., Shima S., Ramos L., Poulin D. S., Whited G., Rajasekaran S., Stuart J. A., Albert A. D. and Birge R. R., Photochemical and thermal stability of green and blue proteorhodopsins: implications for protein-based bioelectronic devices, *J Phys Chem B.* **114** (2010) 14064-70.
20. Xi B., Tetley W. C., Marcy D. L., Zhong C., Whited G. Birge R. R. and Stuart J. A., Evaluation of blue and green absorbing proteorhodopsins as holographic materials, *J Phys Chem B.* **112** (2008) 2524-32.
21. Chaiken J. and Birge R. R., Photochromic Materials Suitable For Cosmetic and Sunblocking Effects - Patent 6461594 (2002).

4. References for chapter 4

1. Chen Guo-Qiang and Gouaux J. Eric, Overexpression of bacterio-opsin in *Escherichia coli* as a water-soluble fusion to maltose binding protein: efficient regeneration of the fusion protein and selective cleavage with trypsin, *Protein Science* **5** (1996) 456–467.
2. Pompejus Markus, Friedrich Karlheinz, Teufel Michael, and Fritz Hans-Joachim, High-yield production of bacteriorhodopsin via expression of a synthetic gene in *Escherichia coli*, *European Journal of Biochemistry*, **211** (1993) 27–35.
3. Lundstrom K., Structural genomics: the ultimate approach for rational drug design, *Mol Biotechnol.* **34** (2006) 205-12.
4. Flower D. R., Modelling G-protein-coupled receptors for drug design, *Biochimica et Biophysica Acta* - Reviews on Biomembranes **1422** (1999) 207-234.
5. Palczewski K., Kumasaka T., Hori T., Behnke C. A., Motoshima H., Fox B. A., Le Trong I., Teller D. C., Okada T., Stenkamp R. E., Yamamoto M., Miyano M., Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289** (2000) 739-45.
6. White J. F., Trinh L. B., Shiloach J., and Grisshammer R., Automated large-scale purification of a G protein-coupled receptor for neurotensin, *FEBS Lett.* **564** (2004) 289-93.
7. Ren H., Yu D., Ge B., Cook B., Xu Z., and Zhang S., High-Level Production, Solubilization and Purification of Synthetic Human GPCR Chemokine Receptors CCR5, CCR3, CXCR4 and CX3CR1. *PLoS One* **4** (2009) doi:10.1371/journal.pone.0004509.

8. Gourdon P., Alfredsson A., Pedersen A., Malmerberg E., Nyblom M., Widell M., Berntsson R., Pinhassi J., Braiman M., Hansson O., Bonander N., Karlsson G. and Neutze R., Optimized in vitro and in vivo expression of proteorhodopsin: a seven-transmembrane proton pump, *Protein Expr Purif.* **58** (2008) 103-13.
9. Bryksin A.V. and Matsumura I., Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids, *Biotechniques.* **48** (2010) 463-5.
10. Sakmar T. P., Structure of rhodopsin and the superfamily of seven-helical receptors: the same and not the same, *Current Opinion in Cell Biology* **14** (2002) 189–195
11. Lau F.W., Nauli S., Zhou Y. and Bowie J.U., Changing single side-chains can greatly enhance the resistance of a membrane protein to irreversible inactivation, *J Mol Biol.* **290** (1999) 559-64.
12. Faham S., Yang D., Bare E., Yohannan S., Whitelegge J. P. and Bowie J. U., Side-chain contributions to membrane protein structure and stability, *J Mol Biol.* **335** (2004) 297-305.
13. Serrano-Vega M. J., Magnani F., Shibata Y., and Tate C. G., Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc Natl Acad Sci U S A.* **105** (2008) 877-82.
14. Magnani F., Shibata Y., Serrano-Vega M. J., and Tate C. G.. Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor, *Proc Natl Acad Sci U S A.* **105** (2008) 10744-9.
15. Shibata Y., White J. F., Serrano-Vega M. J., Magnani F., Aloia A. L., Grishammer R., and Tate C. G., Thermostabilization of the neurotensin receptor NTS1, *J Mol Biol.* **390** (2009) 262-77.

5. References for Appendix

1. Krebs R. A., Alexiev U., Partha R., DeVita A. M. and Braiman M. S., Detection of fast light-activated H⁺ release and M intermediate formation from proteorhodopsin, *BMC Physiol.* **5** (2002) 5-12.
2. <http://blast.ncbi.nlm.nih.gov/Blast.cgi>
3. Schirmer T., Keller T. A., Wang Y. F and Rosenbusch J. P., Structural basis for sugar translocation through maltoporin channels at 3.1 Å resolution, *Science* **267** (1995) 512-4.
4. Baldwin V., Bhatia M., and Luckey M., Folding studies of purified LamB protein, the maltoporin from the *Escherichia coli* outer membrane: trimer dissociation can be separated from unfolding, *Biochim Biophys Acta.* **1808** (2011) 2206-13.

VITA

NAME: Farhana Fahmeeda Bano Syed

PLACE OF BIRTH: Poona (Pune), India

DATE OF BIRTH: October 2nd, 1973

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED

University of Pune, Pune, India

Syracuse University, Syracuse, NY, USA

DEGREES AWARDED

M.Sc. in Biochemistry, 1996

B.Sc. in Chemistry, 1994

PROFESSIONAL EXPERIENCE

Teaching Assistant, Department of Chemistry, Syracuse University, Syracuse, NY (2001-2005, 2007-2010)

Teaching Assistant, Department of Biology, Syracuse University, Syracuse, NY (2006-2007)

Research Assistant, Department of Biology, Syracuse University, Syracuse, NY (2005-2006)

Project Assistant, National Chemical Laboratory, Pune, India (1996-2000)

AWARDS

Outstanding Teaching Assistant – William D. Johnson Award 2009, Chemistry Department, Syracuse University

PUBLICATIONS

1. Syed F.F. and Braiman M.S., Citrate binding site in proteorhodopsin involves two lysines in the first cytoplasmic loop (under revision).
2. Syed F.F and Braiman M.S., Column-free preparation of proteorhodopsin, an integral membrane protein, by selective precipitation with simple salts (manuscript in preparation).
3. Gentry M.S., Li Y., Wei H., Syed F.F., Patel S.H., Hallberg R.L. and Pallas D.C., A novel assay for protein phosphatase 2A (PP2A) complexes in vivo reveals differential effects of covalent modifications on different *Saccharomyces cerevisiae* PP2A heterotrimers. *Eukaryot Cell*. 2005 (4): 1029-40.
4. Syed F.B., Joshi B.N., SivaRaman H., Khire J.M. and Khan M.I. Purification and characterization of a cell-surface lectin (Lectin II) from *Agrobacterium radiobacter* NCIM 2443. *Biochem. Mol. Biol. Int.* 1999 (47): 361-7.