THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN NATURAL SCIENCE

Conformational Dynamics of Rhodopsins Visualized by Time-resolved Wide Angle X-ray Scattering

ERIK MALMERBERG

UNIVERSITY OF GOTHENBURG

Department of Chemistry – Biochemistry

Göteborg, Sweden

2011

Thesis for the Degree of Doctor of Philosophy in Natural Science

Conformational Dynamics of Rhodopsins Visualized by Time-resolved Wide Angle X-ray Scattering

Erik Malmerberg

Cover: Schematic of the conformational dynamics displayed by bacteriorhodopsin (purple), proteorhodopsin (pink) and visual rhodopsin (red) in solution as determined by time-resolved wide angle x-ray scattering measurements. Light-induced structural rearrangements are depicted in yellow and detergent/lipid molecules in grey.

ISBN 978-91-628-8371-3
Available online at http://hdl.handle.net/2077/27026

Department of Chemistry
Biochemistry and Biophysics
SE-413 90 Göteborg, Sweden

Printed by Aidla Trading AB/Kompendiet
Göteborg, Sweden 2011

Till Johanna

ABSTRACT

Rhodopsins are a family of light-sensitive proteins found in the cellular membranes of a wide range of living organisms. These membrane proteins share a common molecular architecture and are able to use light energy to perform a variety of different biological functions. Mapping the conformational changes required for these proteins to function is important for understanding how light energy is used for energy transduction and sensory perception in biological systems.

In order to visualize these conformational changes over time, the emerging technique of time-resolved wide angle X-ray scattering (TR-WAXS) was employed. Several technical and analytical developments of this solution based method were made during the course of this work, including the development of a new data collection strategy based on a rapid readout X-ray detector.

The light-driven proton pumps bacteriorhodopsin and proteorhodopsin were the first membrane proteins to be characterized using TR-WAXS. The results from these studies indicated that significant α-helical rearrangements precede the primary proton transfer event in bacteriorhodopsin. Comparison with the evolutionary related proteorhodopsin provided important insights into shared conformational dynamics between the two proton-pumps.

Proteorhodopsin was further investigated by probing the conformational changes occurring within its chromophore binding pocket, where the chromophore of proteorhodopsin was substituted with a chemically modified retinal analogue. Comparison between the native and modified form of proteorhodopsin indicated significant chromophore dependant differences in their conformational kinetics. These differences provided new insights into the coupling between retinal isomerisation and protein conformational changes.

The conformational dynamics within visual rhodopsin, the primary light sensor of vertebrate vision, were also investigated using TR-WAXS. By using the rapid readout X-ray detector we were able to follow the activation of this G-protein coupled receptor in real-time. Structural analysis further indicated that dramatic conformational changes are associated with the activation of this receptor.

CONTRIBUTION REPORT

There are several authors on the papers presented here and my contribution to each of them is listed below. The focus of my thesis is on areas where I have made major contributions.

Paper I. I took part in preparing figures and writing of the manuscript.

Paper II. I was involved in planning the project, sample preparation, performing the scattering experiment and was responsible for the spectroscopic characterization. I performed the data reduction, kinetic analysis and participated in the structural analysis. I took a major part in preparing figures and writing of the manuscript.

Paper III. I planned the project, produced and purified the protein, performed the spectroscopic characterization and took part in the scattering experiment. I performed the data reduction, kinetic analysis and structural analysis. I took a major part in interpretation of the results, preparing figures and writing of the manuscript.

Paper IV. I was involved in planning the project, developing the experimental setup, performing the experiment and was responsible for sample preparation. I took part in data reduction, interpretation of the results, preparing figures and writing of the manuscript.

Paper V. I planned the project, prepared the samples and took part in the scattering experiments. I performed the data analysis and modelling of the data. I took a major part in the interpretation of the results, preparing figures and writing of the manuscript.

LIST OF PUBLICATIONS

This thesis is based upon the following papers, which will be referred to in the text by their Roman numerals.

Paper I. Westenhoff, S., Nazarenko, E., **Malmerberg, E**., Davidsson, J., Katona, G. and Neutze, R. (2010) Time-resolved structural studies of protein reaction dynamics: a smorgasbord of x-ray approaches. *Acta Cryst,* A66, 207-219. (*Review*)

Paper II. Andersson, M*., **Malmerberg, E***., Westenhoff, S., Katona, G., Cammarata, M., Wohri, AB., Johansson, LC., Ewald, F., Eklund, M.., Wulff, M., Davidsson, J and Neutze, R. (2009) Structural Dynamics of Light-Driven Proton Pumps. *Structure*, 17, 1265-1275.

Paper III. **Malmerberg, E**., Omran, Z., Hub, JS., Li, X., Katona, G., Westenhoff, S., Johansson, LC., Andersson, M., Cammarata, M., Wulff, M., van der Spoel, D., Davidsson, J., Specht, A. and Neutze, R. (2011) Time-Resolved WAXS Reveals Accelerated Conformational Changes in Iodoretinal-Substituted Proteorhodopsin. *Biophys J,*101, 1345-1353.

Paper IV. Westenhoff, S*., **Malmerberg, E.***, Arnlund, D., Johansson, L., Nazarenko, E., Cammarata, M., Davidsson, J., Chaptal, V., Abramson, J., Katona, G., Menzel, A. and Neutze, R. (2010) Rapid readout detector captures protein time-resolved WAXS. *Nat Methods*, 7, 775-776.

Paper V. **Malmerberg, E**, Katona**,** G., Bovee, P., Westenhoff, S., Johansson, LC., Arnlund, D., Nazarenko, E., Menzel, A., de Grip, WJ. and Neutze, R. (2011). Conformational Activation of Rhodopsin Probed by Time-resolved Wide Angle X-ray Scattering. (*manuscript*)

* Equal contributing authors

Related publications

Paper VI. 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. (2008) Optimized in vitro and in vivo expression of proteorhodopsin: A seven-transmembrane proton pump. *Protein Expr Purif*, 58, 103-113.

Paper VII. Vincent, J., Andersson, M., Eklund, M., Wohri, AB., Odelius, M., **Malmerberg, E**., Kong, QY., Wulff, M., Neutze, R. and Davidsson J. (2009) Solvent dependant structural perturbations of chemical reaction intermediates visualized by time-resolved x-ray diffraction. *J Chem Phys*, 130, 154502.

Paper VIII. Wohri, AB., Wahlgren, WY., **Malmerberg, E.**, Johansson, LC., Neutze, R and Katona G. (2009) Lipidic Sponge Phase Crystal Structure of a Photosynthetic Reaction Center Reveals Lipids on the Protein Surface. *Biochemistry*, 48, 9831-9838.

Paper IX. Wohri, AB., Katona, G., Johansson, LC., Fritz, E., **Malmerberg, E.**, Andersson, M., Vincent, J., Eklund, M., Cammarata, M., Wulff, M., Davidsson, J., Groenhof, G. and Neutze, R. (2010) Light-induced Structural Changes in a Photosynthetic Reaction Center Caught by Laue Diffraction. *Science*, 328, 630-633.

CONTENTS

ABBREVIATIONS AND SYMBOLS

bR	Bacteriorhodopsin
cSAXS	Coherent small angle X-ray scattering
E. coli	*Escherichia coli*
ESRF	European Synchrotron Radiation Facility
fs	Femtosecond (10^{-15} s)
GeV	Gigaelectron volt (10^{9} eV)
GPCR	G-protein coupled receptor
FWHM	Full width half maximum
MD	Molecular dynamics
μs	Microsecond (10^{-6} s)
ms	Millisecond (10^{-3} s)
ns	Nanosecond (10^{-9} s)
OG	n-octyl-β-D-glucopyranoside
ps	Picosecond (10^{-12} s)
pR	Proteorhodopsin
rho	Visual rhodopsin
RMSD	Root mean square deviation
SAXS	Small angle X-ray scattering
SLS	Swiss Light Source
SVD	Singular value decomposition
WAXS	Wide angle X-ray scattering
Å	Ångström (10^{-10} m)

1 INTRODUCTION

Sunlight is a fundamental energy source for life on earth. On the molecular level light energy can be harnessed to perform a range of functions which are vital for the survival of the host organism. This work contributes to the understanding of the molecular changes that occur in one such family of light-sensitive molecules, the rhodopsins. In order to visualize these molecular changes over time, the technique of time-resolved wide angle X-ray scattering was employed. Several technical and analytical developments of this method had to be made to capture these fascinating and complex molecules in action.

1.1 THE BIOLOGICAL MEMBRANE

An essential property defining all organisms is the ability to form biological barriers to separate itself from the surrounding environment. These biological membranes effectively compartmentalize organisms into tissues, tissues into cells and cells into sub-cellular organelles in a hierarchal fashion. By selectively regulating the passage or exclusion of certain molecules to these compartments these membranes establish unique chemical environments where reactions can occur.

Biological membranes, such as the envelope that surrounds the cell, are comprised out of two major biological building blocks; lipids and membrane proteins. The lipid molecules, consisting of polar head groups and nonpolar hydrocarbon tails, form the basic bilayer framework of the membrane. In the bilayer structure (Figure 1) the lipid molecules are arranged in two leaflets; with the hydrocarbon tails of each layer oriented towards the centre of the bilayer and the head groups extending at opposite ends forming the edges of the membrane. The nonpolar core region formed at the centre of the bilayer establishes a permeability barrier between the polar aqueous environment of the inside and the outside of the cell. This permeability barrier prevents ions and other polar solutes to freely traverse the bilayer and thus enables charge and concentration gradients to be maintained across the membrane, a vital prerequisite for the cell metabolism [1].

Membrane proteins, accounting for 20-80 % by dry weight of the membrane[2], constitute a crucial part of the biological membrane. Membrane proteins, as is true for all proteins, are constructed from linear chains of covalently bonded amino acid residues with a sequence encoded by the DNA of its gene. The distinct chemical properties of the amino acids side chains enable the linear chains to fold into a more ordered secondary structure (α-helices and β-sheets), which in turn can pack together to form even more complex tertiary and quaternary structures. The global structure and properties of each membrane protein is as such uniquely determined by its amino acid sequence.

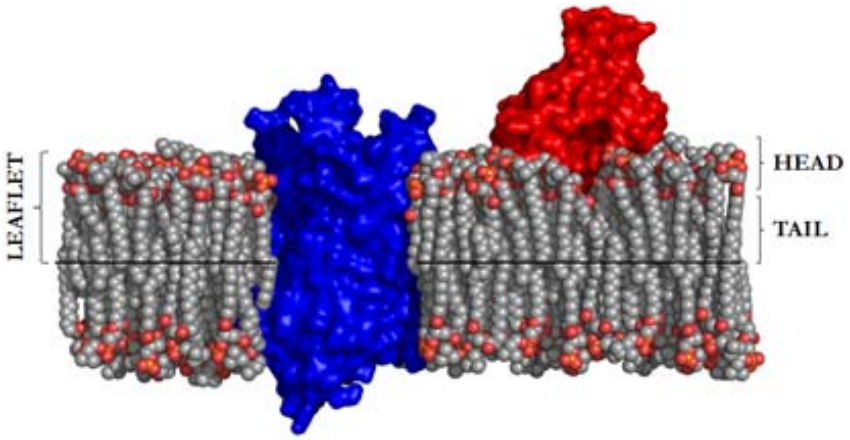


Figure 1. The biological membrane. Lipid molecules are depicted with their polar head groups (red) and nonpolar aliphatic tails (grey). Integral (blue surface) and peripheral (red surface) membrane proteins catalyzes enzymatic reactions and regulates the passage of ions and solutes across the membrane bilayer.

Depending on its surface polarity, membrane proteins can either be partly associated to (peripheral) or embedded into the lipid bilayer (integral). While the lipid bilayer forms the skeleton of the membrane, the membrane proteins infer specificity and functionality to the membrane. The functional role of membrane proteins is highly diverse and range from selectively transporting solutes or transducing signals across the membrane to catalyzing enzymatic reactions or imposing cellular structure[3]. Not surprisingly the vital role for cellular function has made membrane proteins important targets for academic as well as pharmaceutical studies[4].

1.2 RETINYLIDENE PROTEINS

The retinylidene proteins, commonly referred to as rhodopsins, are a diverse superfamily of membrane proteins that uses vitamin-A aldehyde (retinal) as a chromphore for light reception. Over three hundred of these photochemcially reactive proteins of both prokaryotic and eukaryotic origin have been described to date[5]. Between them they form the molecular basis for a variety of light-sensing biological processes ranging from the flagellar movement observed in microorganims to eyesight in animals.

All rhodopsins share a common topology consisting of seven membrane embedded α-helices that form an internal pocket in which the retinal chromophore resides (Figure 2). Despite their shared secondary and tertiary structure, there are considerable differences between rhodopsins in their primary structure. Based on amino acid sequence alignment they can be divided into two distinct families. The type 1 rhodopsins (or microbial rhodopsins) are predominantly found in prokaryotes and algae and function as light-drive ion pumps, channels or receptors for phototaxis. The type 2 rhodopsins (or visual rhodopins) are almost exclusively photosensitive receptor

proteins and are found in higher eukaryotes, including the human rod and cone visual pigments [5].

Crucial for the functionality of all retinylidene proteins is the light sensitive retinal chromophore. In the absence of which, the protein remains in an inactive apoprotein (or opsin) state. The retinal molecule covalently attaches to the ε-amino group of a lysine residue in the seventh helix of the apoprotein and forms a protonated retinylidene Schiff base linkage (Figure 2). Upon binding to the protein the absorption properties of the polyene retinal change due to interactions with surrounding residues in the retinal binding pocket. These interactions are specific for the different rhodopsins and cause the proteins to display characteristic colours. The ability of retinylidene proteins to tune its spectral properties to match the light conditions available for the host organism has been of key importance for the evolution of rhodopsin function[6-7].

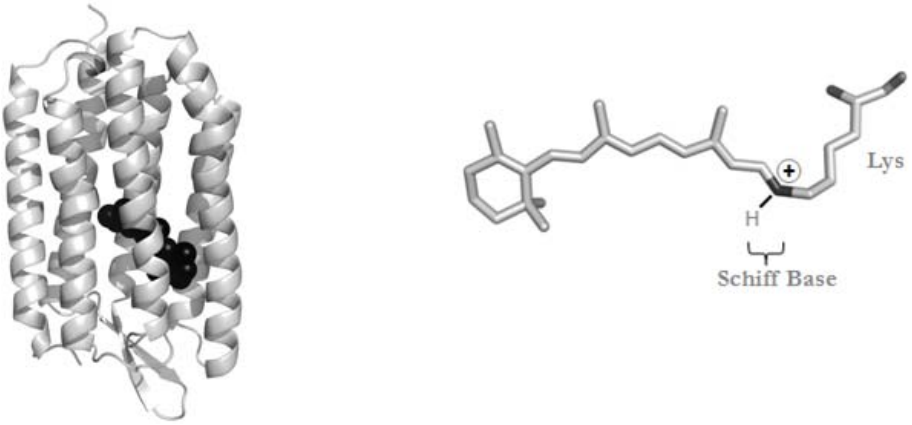


Figure 2. **Left:** Seven transmembrane α-helical topology of rhodopsins with the bound retinal chromophore (black). **Right:** All-*trans* retinal bound to the lysine residue via a protonated Schiff base.

The first functional event in all rhodopsins is the absorption of light by the retinal molecule. The energy absorbed causes the chromophore to isomerize around one of its polyene double bonds, forcing the retinal to adopt a new configuration. In type 1 rhodopsins this transformation causes the retinal to change from an all-*trans* to a 13-*cis* configuration. In type 2 rhodopsins the isomerization predominantly occurs from a 11-*cis* to all-*trans* state[5] (Figure 3). The photoisomerization of the retinal in turn initiates a sequence of events necessary for the protein to perform its specific function. In the subsequent sections we will take a closer look into the structure and function of three different retinylidene proteins; two type 1 rhodopsin (bacteriorhodopsin and proteorhodopsin) and one type 2 rhodopsin (visual rhodopsin).

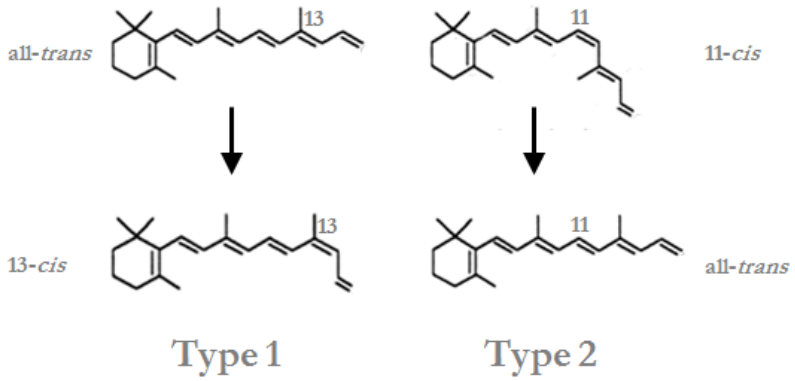


Figure 3. Light-induced isomerization of the retinal chromophore. In type 1 rhodopsins the *trans-cis* isomerization occurs around the C13=C14 bond and in type 2 rhodopsin the *cis-trans* isomerization is localized to the C11=C12 bond.

1.2.1 BACTERIORHODOPSIN

Since its first discovery in the archea *Halobacter salinarium* in 1973[8], bacteriorhodopsin (bR) has become the most well characterized of all the retinylidene proteins. These haloarchea naturally grow in high-temperature salt brines exposed to bright sunlight. Since the abundance of oxygen is sparse in this environment, alternative means to respiration are necessary for organisms in this habitat to generate energy and survive. *H. salinarium* is able to thrive under these conditions by utilizing sun light as it main source of energy[9]. At the heart of this energy transducing mechanism is the light-driven proton pump bacteriorhodopsin.

Bacteriorhodopsin naturally assembles into concentrated patches, called purple membrane, in the cellular membrane of *H. salinarium*[8]. When exposed to sun light the bacteriorhodopsin molecules in the purple membrane use the light energy to transports protons from the negatively charged inside (cytoplasmic side) to the positively charged outside (extracellular side) of the cell. The net translocation of protons establishes an electrochemical proton gradient across the membrane which in turn is harvested by the enzyme ATPase to generate chemical energy in the form of ATP[10].

The different chemical steps (intermediates) involved in the proton translocation pathway of bacteriorhodopsin have been extensively studied using a wide variety of different biochemical and biophysical techniques. Including spectroscopic methods (such as UV/VIS, FTIR and Raman spectroscopy) and methods for three dimensional structure determination (such as electron and x-ray crystallography) and have often been combined with site-directed mutations. Together these studies provide a detailed and sometimes heterogeneous mechanistic view into the inner workings of the bacteriorhodopsins photocycle[11]. From this work the following picture of the key residues and steps involved has emerged:

In the resting state of bacteriorhodopsin (bR) the all-*trans* retinal is covalently linked to Lys216 in helix G forming the protonated Schiff base (the primary donor). A complex

counter ion surrounds the chromophore pocket and effectively stabilizes the high pKa~13.5 of the primary donor and the low pKa~2.2 of the primary proton acceptor, Asp85, thereby preventing proton transfer to occur in the resting state [12-14]. The complex counter ion, comprising a network of water molecules and charged residues, further links the primary acceptor via Arg82 to the proton release group, consisting of Glu194 and Glu204, at the extracellular side of the membrane [15-16]. The photocycle starts when the resting state absorbs a photon and the retinal rapidly isomerizes from an all-*trans* to 13-*cis* configuration. This initiates a series of structural rearrangements in the chromophore environment correlating with spectrally distinguishable intermediates referred to as K, L, M, N and O (Figure 4). First to form, after a few picoseconds, is the K-intermediate [17-18]. Here the isomerized retinal reorients the Schiff base N-H dipole toward the cytoplasmic side[19] causing key interactions with the complex counter ion to break and the accessibility to Asp85 to temporarily decrease[19].

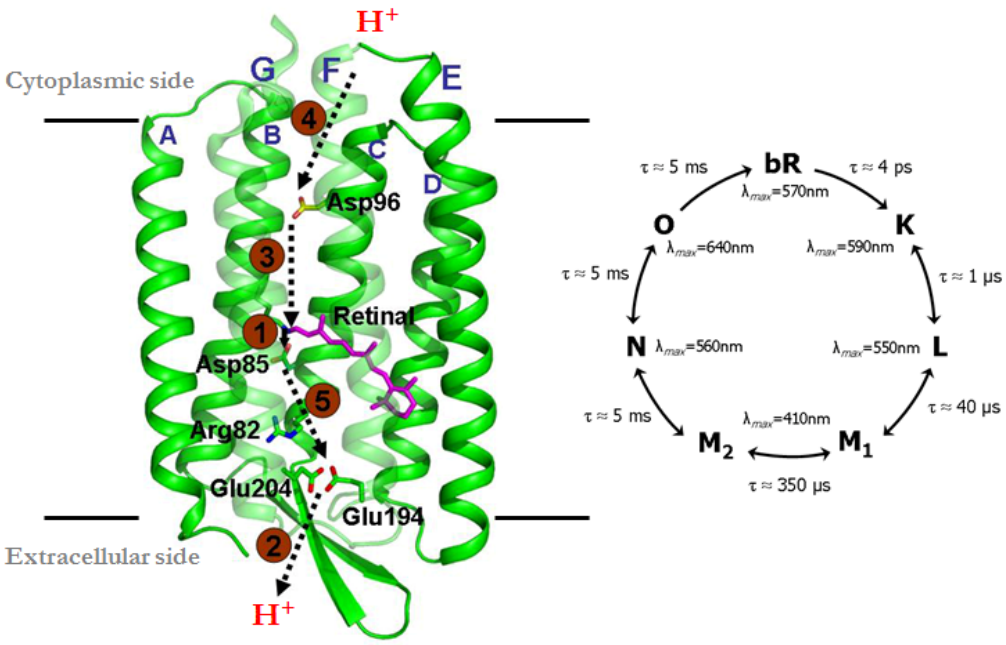


Figure 4. **Left**: Schematic illustration of vectorial transport of protons (H^+) in bacteriorhodopsin. The primary proton transfer (1) occurs between the Schiff base and Asp85. A proton is then released (2) to the extracellular side by the proton release group formed by Glu194 and Glu204. The Schiff base is subsequently reprotonated from Asp96 (3) which is followed by a reprotonation of Asp96 from the cytoplasmic medium (4). The photocycle ends by the shuttling of a proton from Asp85, via Arg82, to the proton release group (5). **Right:** The spectral intermediates of the bacteriorhodopsin photocycle along with their characteristic absorption maxima and life times at room temperature.

The reorientation also lowers the pKa of the Schiff base[20] bringing it closer in pKa to the donor. On the order of a microsecond the L-intermediate forms where the proton donor-acceptor relationship is temporarily restored again as a local inward flex of helix C brings Asp85 closer to the Schiff base[19]. This movement is aided by the reorientation of Arg82 which simultaneously contribute to raising the pKa of Asp85, priming the acceptor for receiving the proton. The primary proton transfer from the Schiff base to Asp85 occurs on the tens of microsecond time-scale (M_1- intermediate) following light absorption. In the same time period the proton release group releases a proton to the

extracellular side[21]. In the submillisecond time-range a spectrally silent transitions occurs where the early M state transforms into a late M (M_2-intermediate)[22]. This transformation has been proposed to be accompanied by large structural rearrangements, involving the cytoplasmic side of helices G and F[23]. Such changes would ensure a switch in accessibility of the retinal towards the cytoplasm, allowing the Schiff base to be reprotonated again[23-24]. This is also supported by the formation of the subsequent N-intermediate which occurs after a few milliseconds and where Asp96, located towards the cytoplasmic side, first reprotonates the Schiff base and then gets reprotonated from the cytoplasm[24]. Although several studies[23,25] supports the notion that the late M and N states displays similar conformational changes at the cytoplasmic end, the nature and magnitude of these changes is inconclusive[11]. The retinal eventually thermally reisomerizes to the all-*trans* state on the millisecond time-scale, marking the transition from the N- to the O-intermediate. As a consequence of the reisomerization, the steric strain between the retinal's C-20 methyl group and nearby residues (such as Trp182) is relaxed, causing the cytoplasmic side to close. Finally the bacteriorhodopsin resting state is recovered when the release group is reprotonated by Asp85, marking the end of the photocycle[24].

1.2.2 PROTEORHODOPSIN

Proteorhodopsin (pR) was first discovered in 2000[26] during gene sequencing of marine bacterioplankton recovered from oceanic seawater. This eubacterial rhodopsin, named after its γ-proteobacteria host, displayed light-driven proton pumping capabilities similar to that of the archeal homologue bacteriorhodopsin[26-27]. Furthermore the proton gradient generated by heterologously expressed proteorhodopsin was demonstrated to be sufficient to generate ATP[28]. The discovery challenged the assumption that chlorophyll *a* was the only light-capturing pigment in oceanic surface waters and raised the prospect that a substantial amount of energy might be harvested by these alternative pigments[26,29]. This notion has since been further developed with the discovery of proteorhodopsin genes in several of the most abundant taxa of marine bacteria (α-, β- and γ-proteobacteria)[30-31], marine archea[32] and even in marine eukaryotes[33] in various oceanic waters. Considering the vast biomass of these marine microorganism, it is conceivable that proteorhodopsin-mediated phototrophism has a significant impact on the biogeochemical carbon cycling and energy fluxes in the worlds oceans[34].

Further characterization was able to demonstrate the existence of two related proteorhodopsin subgroups, absorbing light with absorption maxima centred in the green (λ_{max}=525nm) and blue (λ_{max}=490nm) light region. The natural occurrence and vertical distribution of these proteorhodopsins was further related to oceanic depth, with the 'green-absorbing' pigments predominantly found in surface waters and the 'blue-absorbing' pigments in deeper water, indicating an evolutionary adaptation according to the light available at these strata[34]. The molecular origin of this 'spectral tuning mechanism' was located to the difference in a singular amino acid residue (Leu

or Gln at position 105) in the retinal binding pocket[35-36]. The 'blue-absorbing' proteorhodopsin also displayed a slower photocycle than the 'green-absorbing' counterpart[37]. This prompted the suggestion that the 'blue-absorbing' proteorhodopsin might not have the potential to contribute significantly to solar energy capture and might play a different regulatory role[37-38].

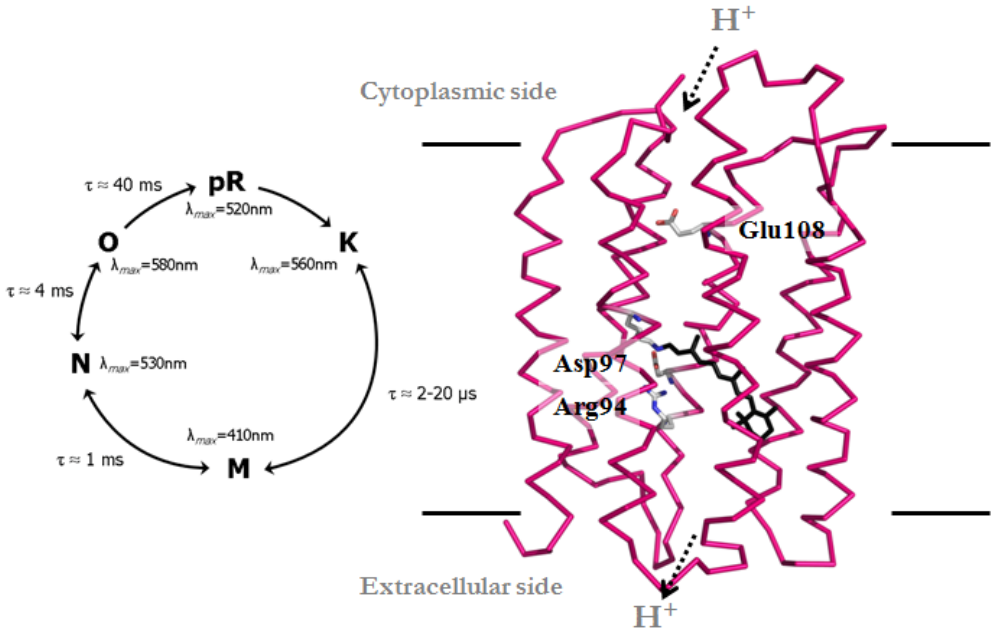


Figure 5. **Left:** The spectral intermediates of the proteorhodopsin photocycle along with their absorption maxima and life times at room temperature. **Right:** Homology model of proteorhodopsin based on the bacteriorhodopsin structure. Key residues are highlighted.

The 'green-absorbing' proteorhodopsin originating from the SAR86 clade of γ-proteobacteria, hereafter only referred to as proteorhodopsin, is the best characterized of the eubacterial proton pumps. Spectroscopic[27,39-43], as well as mutational studies[35,44-45], indicate that the photocycle of proteorhodopsin (Figure 5) shares many similarities with that of bacteriorhodopsin. Baptized with a similar nomenclature of its intermediates (K, M, N and O), the proteorhodopsin photocycle however displays a somewhat slower turnover rate then its archeal homologue[40]. Many of the key residues important for proton pumping appear to be conserved, including the Schiff base formed between Lys231 and the all-*trans* retinal as well as the primary proton acceptor Asp97. In contrast to bacteriorhodopsin, the primary proton acceptor displays an unusual high pKa~7.7[40], indicating that outward directed proton transport can only occur at alkaline pH[27,40]. This adaptation has been attributed to the alkaline nature of seawater[27], having a pH~7.6-8.2[46]. Analogous to bacteriorhodopsin, the retinal interacts with a complex counter ion involving charged residues (Arg94 and Asp227) while Glu108 (instead of Asp96 in bacteriorhodopsin) is responsible for the cytoplasmic reprotonation of the Schiff base[40]. It is worth noting that no clear proton release group, corresponding to the carboxylate Glu194/Glu204 dyad in bacteriorhodopsin, has yet to be identified in proteorhodopsin[45]. Instead, the fast

proton-release observed in proteorhodopsin has been suggested to occur directly from Arg94 (Arg82 in bacteriorhodopsin) without any intermediate transfer events[39,42].

Proteorhodopsin and bacteriorhodopsin share a 30% primary sequence identity in their transmembrane region[47]. Although no high-resolution structure of proteorhodopsin exists to date, an advanced homology model based on the bacteriorhodopsin structure has been reported[47]. In this homology model the sequence alignment was carefully followed up with an energy minimization and a molecular dynamics simulation to account for spectral properties characteristic of the 'green-absorbing' proteorhodopsin.

1.2.3 VISUAL RHODOPSIN

In stark contrast to the two previously described microbial proton-pumps, visual rhodopsin (or visual purple[48]) acts as a photoreceptor in the rod cells in the vertebrate eye. These cells, localized to the retina, are highly light-sensitive and the capture and conversion of light into a chemical signal in these neurons is the first step in the visual process of vertebrate night vision. The rod cell (Figure 6) is comprised of two parts, an inner segment and an outer segment. While the inner segment harbours the nucleus and mitochondria, the outer segment is specialized for photon capture and amplification and contains ~1000-2000 stacked disc membranes enriched with rhodopsin molecules. Rhodopsin occupies ~50% of the disc surface area, the remainder of which is filled with phospholipids and cholesterol, and typically account for >90% of the total protein content[1]. The outer segment also contains Na^+ and Ca^{2+} channels which open in response to cyclic guanosine monophosphate (cGMP). When light is absorbed by the outer segment, the concentration of cGMP decreases in the rod cell, which in turn causes the sodium and calcium channels to close. This results in a dramatic change in electric potential (hyperpolarization) across the cell membrane. The hyperpolarization event prevents the release of neurotransmitters from the rod cell and this lack of stimulus result in visual perception[1].

The main events involved in the hyperpolarization are catalyzed by the visual rhodopsin (a G-protein coupled receptor) and its interaction partner transducin (a G–protein). When rhodopsin is light activated, the soluble G-protein is able to interact with the receptor at the cytoplasmic surface of the membrane. This interaction stimulates the replacement of GTP to GDP in transducin, causing its α-subunit to dissociate from the complex, bind to and activate the enzyme cGMP phosphodiesterase (PDE), which degrades cytosolic cGMP and thereby causes the hyperpolarization[1].

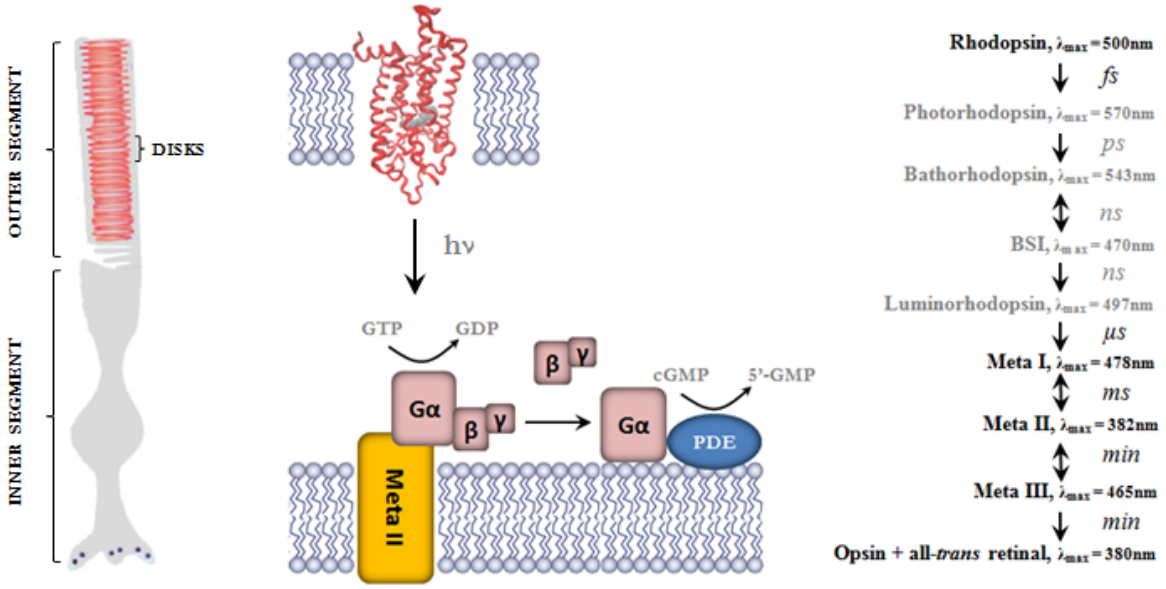


Figure 6. **Left:** Diagram of the rod cell. The outer segment contains the rhodopsin enriched disk membranes. **Center:** Schematic of rhodopsin activation and subsequent binding and activation of the G-protein transducin. **Right:** The spectral intermediates of the rhodopsin photoreaction along with their absorption maxima and approximate time-scales at room temperature.

The activation mechanism of rhodopsin has been characterized using both spectroscopic[49-55] and structural techniques[56-72] and differs in many ways from that of the rapidly reversible proton-pumps. The protonated Schiff base in rhodopsin is formed between Lys296 (in Helix 7) and the 11-*cis* retinal chromophore. Upon absorption of a photon the retinal first isomerizes to form a highly strained and distorted retinal conformation (photorhodopsin) which thermally relaxes through a number of intermediate steps (bathorhodopsin, blue-shifted intermediate (BSI), luminorhodopsin) to form an all-*trans* conformation in the MetaI intermediate[73]. This process occurs on the picoseconds to microsecond time-scale while the Schiff base remains protonated, most likely due to the low pKa of the stabilizing counter ion Glu113[74]. The transition from MetaI to MetaII occurs after a few milliseconds and involves large conformational changes, predominantly on the cytoplasmic side[75-78], and is believed to be coupled to the Schiff base deprotonation[79]. The formed MetaII represents the active receptor state which is capable of interacting with the G-protein transducin. After several minutes MetaII decays, either directly or via the reprotonated MetaIII intermediate, into the apoprotein opsin and free all-*trans* retinal as a result of irreversible hydrolysis of the Schiff base linkage. Under physiological conditions, fresh 11-*cis* retinal is metabolically supplied and slowly taken up by opsin to regenerate new rhodopsin molecules able to absorb light[74]. It is in many ways remarkable to think that activation of a single rhodopsin molecule, by a single photon, yields amplification such that within milliseconds hundreds of ion channels transiently close to hyperpolarize the membrane.

1.3 PROTEIN CONFORMATIONAL DYNAMICS (PAPER I)

Proteins are dynamic molecules and in order to full fill their biological function they often have to undergo structural rearrangements. The transition from one energetically favourable structural state (conformation) to another is commonly referred to as a conformational change. These changes may be induced by several factors such as temperature, pH, light or ligand binding. The nature of the conformational changes can range from subtle relocations of individual amino acid residues to concerted movements of secondary structural elements, or even rearrangements of entire subunits. Usually the sequence of these motions is well coordinated with time in order to optimize the catalysis of the specific reaction. Information about the nature and dynamics of these conformational changes is therefore of fundamental importance to understand how these molecular machines work.

There are several techniques for determining the three dimensional structure of proteins, including Nuclear Magnetic Resonance (NMR) and X-ray crystallography, each with their particular advantages and disadvantages. X-ray crystallography has proven to be one of the most successful techniques by providing high-resolution structural information as well as being applicable to a range of different proteins. There are currently more than 66,000 deposited structures (87 %) in the Protein Data Bank[80] (www.pdb.org) which has been solved using this method. A major limitation of X-ray crystallography however, is that it relies upon the ability of proteins to form well ordered arrangements of identical copies, crystals. Regions in the protein that are disordered therefore tend to be difficult to visualize in the crystal structure. Moreover, proteins that are inherently flexible, such as membrane protein receptors and transporters, might be difficult if not impossible to crystallize. Further complications arise when attempting to structurally determine transient intermediate conformations of the protein. In particular, apart from the technical challenge of triggering the conformational change in the crystalline state, there is the possibility that the conformational changes break the crystal lattice, making high resolution structural determination impossible. Thus, despite the success of X-ray crystallography, most of the deposited structures available are also limited to the resting conformation of proteins. To address the difficulties of visualizing conformational changes in proteins, several crystal and solution based techniques have emerged. They all exploit the use of X-rays to probe the protein structural dynamics and combined they can be used to record the molecular events of protein reactions.

In kinetic crystallography, temperature and chemical substances are commonly used to control the reaction kinetics in such a way that causes a significant population of a desired intermediate to be trapped within the crystal. Prior to collecting crystallographic data, the conditions necessary to yield a highly populated intermediate can be assessed in the crystalline state using complimentary techniques such as optical[81] and Raman spectroscopy[82-84]. If the reaction is sufficiently slow the induced intermediate can be kinetically trapped by rapidly lowering the temperature and flash

freezing the crystal. This low-temperature trapping technique has successfully been demonstrated for a number of systems[85-91] ranging from superoxide reductase to photoactive proteins such as bacteriorhodopsin. Several different strategies have likewise been used to trigger the formation of the trapped intermediates. Such strategies include; light activation of photoactive proteins, X-rays to stimulate redox-sensitive proteins but also by chemically soaking protein crystals in order to produce diffusion triggered pH shifts.

Whereas kinetic crystallography relies on producing long lived intermediates by slowing down the reaction, time-resolved Laue diffraction is able to record the transient intermediate in real time and at ambient temperatures. In this approach the reaction is initiated within the crystal *in situ* and then rapidly probed using a short polychromatic X-ray pulse. The use of a polychromatic rather than monochromatic beam, traditionally used in crystallography, has two principal advantages; (i) it increases the available X-ray flux to the crystal (ii) it enables a large number of full, rather than partial, X-ray diffraction reflections to be collected without the need to rotate the crystal. The latter is essential when studying events occurring in the sub-millisecond time range. Several light activated systems have been successfully characterized in the nano to millisecond regime using time-resolved Laue diffraction including; myoglobin in complex with CO[92-94], photoactive yellow protein[95-98] and photosynthetic reaction centre[99]. Although these studies have provided unique high-resolution structural insights into the reaction dynamics of these proteins, a widespread use of time-resolved Laue diffraction has been limited by the high demands on crystal quality and the experimental limitations of reversibly triggering the reaction within the crystal.

A major shortcoming with both kinetic crystallography and time-resolved Laue diffraction is that the protein is confined within the crystal lattice. The specific crystallisation conditions (pH, high salt concentration etc) could also negatively affect and even inhibit the reaction conditions[100]. This raises the concern that the trapped intermediates or transient Laue structures do not necessarily describe the full extent of conformational changes, but rather reflect the structural displacements permitted within the crystal conditions and confines. A further complication is that large scale conformational changes (such as helical motions) can break the crystal contacts between adjacent molecules and disrupt the three-dimensional crystal lattice, limiting crystal diffraction. X-ray methods applicable to proteins in the solution phase can explicitly avoid many of these limitations and complications of crystal based approaches.

Time-resolved small angle X-ray scattering (SAXS) and wide angle X-ray scattering (WAXS) are two closely related methods that are directly applicable to probe proteins structural dynamics in solution (Figure 7). SAXS takes advantage of the structural information content available at small angles (~0.3-4º) in the diffuse scattering pattern of protein solutions.

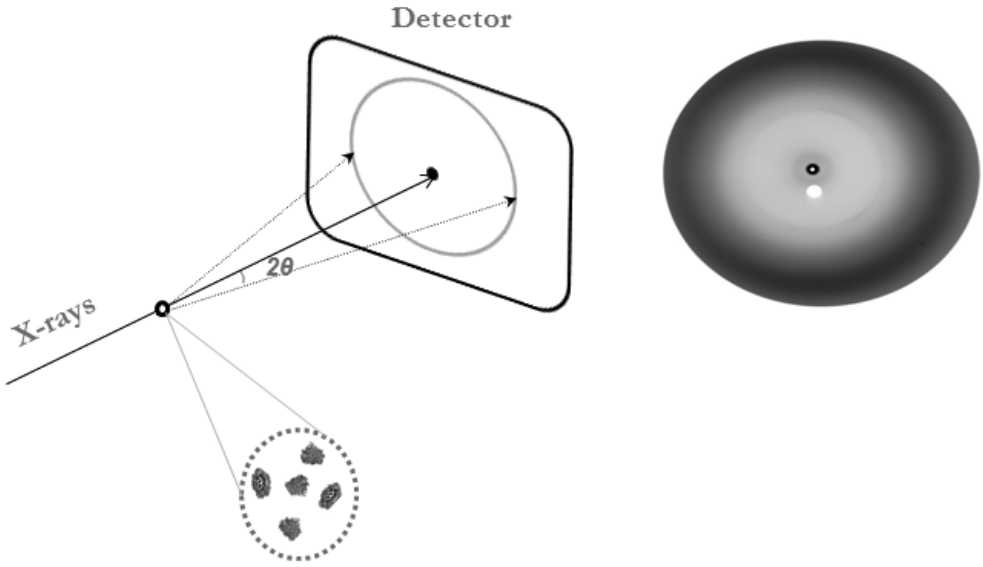


Figure 7. **Left: Schematic of** SAXS/WAXS experiments on protein solutions (dotted circle) where 2θ indicates the measured scattering angle. **Right:** Typical diffuse scattering patterns collected from such experiments.

This region typically provides low-resolution information (~20-250 Å) related to the overall shape and size of the measured macromolecule. In contrast to the crystallographic methods, however, there are fewer demands on the sample, such as the ability of the protein to form well ordered crystals and reproducible SAXS profile can often be collected on limited amounts of protein sample. SAXS is also less hindered by the size of the particle and therefore permits the characterization of large multimeric systems and protein complexes[101]. Advances in sample delivery systems, stopped-flown mixing devices and data collection strategies have also permitted automated sample characterization as well as time-resolved data collection[101]. A major challenge however has been to retrieve the structural information from the one-dimensional SAXS scattering profiles. The last decades have seen considerable advances in this area and to date several publicly available software tools exist for three-dimensional structure reconstruction[102], ranging from *ab initio* shape reconstruction by simulated annealing[103] to crystal structure based rigid-body modelling[104].

While SAXS provides information on the shape and size of the protein, WAXS scattering data emphasises on the wider scattering angles (~3-20º) which contain information related to protein intramolecular structure, such as tertiary and secondary structure fold in proteins[105]. Since the scattering intensity in this regime is typically 1-3 orders of magnitude[106] weaker than at small angles highly concentrated samples are a prerequisite for these measurements. Although a less mature technique then SAXS, the advent of third generation synchrotron sources have provided the X-ray flux necessary to collect high statistic wide angle X-ray data comparable to a distance resolution of ~2.5Å[105].

Time-resolved WAXS was initially applied to study the formation of transient intermediates of small photosensitive molecules in solution[107-110], all of which contained one or several heavy atoms to increase the scattering power of the solute. One of the fundamental challenges in these experiments was to isolate the signal related to the conformational changes in the solute within a background of solvent molecules undergoing structural changes due to solvent heating[111-112]. Recently time-resolved WAXS was also extended to probe the structural changes in the first biological systems[113]. The proof-of-principle experiment was conducted on highly concentrated haemoglobin in complex with carbon monoxide using a pump-probe data collection strategy. In this study the tertiary and quaternary conformational changes of haemoglobin were followed in the nanosecond to millisecond time regime after CO was dissociated by a laser flash. It was further demonstrated that the scattering changes observed were comparable to those predicted from crystallographic structures. The same study was also able to demonstrate the applicability of time-resolved WAXS for two other soluble proteins systems; the photo dissociation of CO from a myoglobin:CO complex and the refolding of cytochrome c after light-induced denaturation. Although limited new structural insight was gained from these proof-of-principle experiments, it demonstrated the feasibility of collecting high statistic time-resolved WAXS data on macromolecules at synchrotrons and showed the sensitive of the technique to protein structural changes.

This provided a starting point for studying the structural dynamics of more complex biological systems, including membrane proteins, and to gain new biological insights by performing comparative studies of different systems. In order for the method to grow in applicability, there is also the need to extend the availability of the technique beyond that of the highly specialized beamline of the original study.

1.4 SCOPE OF THE THESIS

The aim of this thesis has been to develop a solution based X-ray methodology to study time-resolved conformational changes in membrane proteins. The advent of such a technique would be able to complement existing crystal based techniques (Paper I) while simultaneously addressing many of their limitations. The strategy was to extend the technique of time-resolved WAXS, previously applied to small inorganic molecules and soluble proteins, in order to gain new biological insight into membrane protein dynamics. Specifically the technique was applied to study the time-dependant conformational changes in three different members of the rhodopsin family.

The first step was to demonstrate that the emerging technique of time-resolved WAXS was applicable to study conformational dynamics in membrane proteins. As proof-of-principle systems for membrane transport proteins, we studied the light-driven proton pumps bacteriorhodopsin and proteorhodopsin. From this work we were able demonstrate the feasibility of collecting difference WAXS data on membrane proteins

using the pump-probe data collection strategy at the ID09B beamline at the ESRF. In parallel we developed analytical tools to analyse this time-resolved data and to quantify the structural motions associated with the one-dimensional scattering patterns. The results from this analysis shed new light into the structure and kinetics of the conformational changes occurring in these light-driven proton-pumps (Paper II).

To further extend the resolution of the technique we designed an experiment to investigate the conformational changes occurring in the chromophore binding pocket. Here we substituted the chromophore in proteorhodopsin with a chemically modified retinal analogue, 13-desmethyl-13-iodoretinal, and collected time-resolved pump-probe WAXS data at the ESRF. Cross comparison of WAXS and spectroscopic data from the native and modified proteorhodopsin indicated significant chromophore dependant differences in their conformational kinetics. These differences, interpreted in the light of free-energy calculations, provide new insights into the coupling between retinal isomerisation and protein conformational changes (Paper III).

Next, we addressed the fact that a major limitation of the technique is that there are only two highly specialized beamlines in the world with the necessary requirements to collect time-resolved protein WAXS data. To solve this problem we developed a rapid-readout WAXS data collection strategy based on the fast readout capabilities of a newly developed pixel detector. The strategy was evaluated using proteorhodopsin as a test system at the cSAXS beamline at the SLS. (Paper IV)

Finally, we employed the rapid-readout strategy to characterize the irreversible activation mechanism of the G-protein coupled receptor rhodopsin. Using this strategy the full timing history of the active state formation was captured at the cSAXS beamline at the SLS. In order to accurately interpret the structural changes involved in this activation process, efficient strategies for WAXS data modelling had to be developed. The result from the analysis provides new insight into the conformational changes that occur in this receptor in its native environment (Paper V).

2 METHODOLOGY

2.1 X-RAY SCATTERING

2.1.1 SCATTERING BY A SINGLE ATOM

X-rays are electromagnetic radiation with wavelengths in the range of 0.1 to 1 Å (0.01-0.1nm). The short wavelength of X-rays make them ideal to measure objects on the molecular level since the wavelength is comparable to the distance between individual atoms. X-rays are scattered by the atomic electrons. The scattering process can be understood in terms of the oscillating electromagnetic fields of the incident X-rays interacting with the electron cloud of the atom, which causes the electrons to oscillate and in turn emit radiation. The combined scattering from all the electrons thus produces a spherical wave emanating from the atom. Although each electron scatters X-rays with the same amplitude, the spatial extent of the electron cloud causes the amplitude of the X-rays scattered by the atom to decrease as a function of the scattering angle. The atomic scattering amplitude as a function of scattering angle is commonly denoted the form factor (*f*) of an atom, and both the amplitude and the angular dependence is characteristic for each element[114]. Lighter elements (such as H and He) contain fewer electrons and consequently have lower scattering amplitudes and stronger angular dependence then heavier elements (such as I and Hg).

2.1.2 SCATTERING VECTOR AND RESOLUTION

Consider the scattering by a wave in a system consisting of two points A and B, separated by a distance AB (Figure 8). An incident wave ($\mathbf{s_0}$) will have a wave front BC as it reaches point B. At point B the wave scatters at an angle 2θ and the resulting wave ($\mathbf{s_1}$) has a wave front defined by BD.

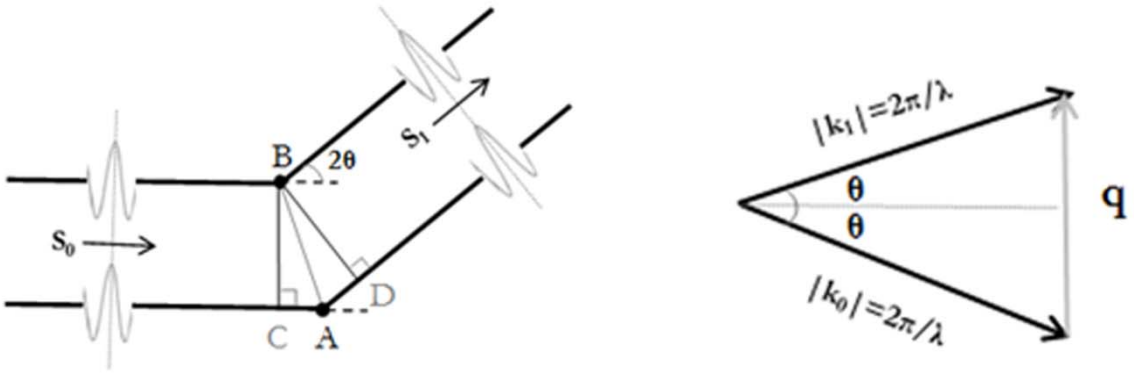


Figure 8. **Left:** X-ray scattering from two points (A and B) in space. **Right:** The scattering vector, **q**, is defined as the difference between the normalized wave vectors $\mathbf{k_0}$ and $\mathbf{k_1}$.

The phase difference, δ (in radians), between $\mathbf{s_0}$ and $\mathbf{s_1}$ can then be expressed as a function of the path difference, Δ, travelled by the two waves such that,

$$\Delta = \text{AD-AC} = \mathbf{r} \cdot (\boldsymbol{s_1} - \boldsymbol{s_0}) \,, \tag{1}$$

$$\delta = 2\pi \frac{\Delta}{\lambda} = 2\pi \frac{\mathbf{r} \cdot (\boldsymbol{s_1} - \boldsymbol{s_0})}{\lambda} \,, \tag{2}$$

where $\mathbf{r}$ defines the vector AB, $\mathbf{s_0}$ and $\mathbf{s_1}$ being unit vectors, and λ is the wavelength of the incident beam. The scattering vector, $\mathbf{q}$, can furthermore be defined as the difference between the wave vectors $\mathbf{k_1}$ and $\mathbf{k_0}$ such that,

$$\boldsymbol{q} = \boldsymbol{k_1} - \boldsymbol{k_0} = \frac{2\pi}{\lambda}\boldsymbol{s_1} - \frac{2\pi}{\lambda}\boldsymbol{s_0} \,, \tag{3}$$

From geometrical arguments (Figure 8) the corresponding magnitude of the scattering vector can in turn be derived,

$$q = \frac{4\pi \sin\theta}{\lambda} \tag{4}$$

and the phase difference can now be expressed as,

$$\delta = \boldsymbol{r} \cdot (\, \boldsymbol{k_1} - \boldsymbol{k_0} \,) = \boldsymbol{r} \cdot \boldsymbol{q} \,, \tag{5}$$

Now imagine that the two points represents two atoms with a form factor *f*. The amplitude (F) of the scattered wave at point B with respect to point A can then be formulated as a function of the scattering vector,

$$F(\boldsymbol{q}) = f e^{i\delta} = f e^{i\boldsymbol{r}\cdot\boldsymbol{q}} \,, \tag{6}$$

Constructively interfere between the waves however only occur when the scattering vector $\mathbf{q}$ is parallel to $\mathbf{r}$ and thus for q equal to 2π/r. This reciprocal relationship between the geometrical space described by q ('reciprocal space') and r ('real space') provides information about the magnification that can be achieved in a scattering experiment. In order to resolve shorter distances between scattering objects we thus have to go to larger q-values, i.e. wider angles. This can be achieved by increasing the observation angle and/or by reducing the wavelength. The resolution, i.e. the minimum distance between points that can be observed separately, achieved in a scattering experiment is therefore approximately $2\pi/q_{max}$, where q_{max} represents the maximum q-value for which scattering intensity is observed[115].

In traditional light- and electron microscopy lenses can be used to attain an image of a particle directly from $F(\mathbf{q})$. There are however no comparable lenses in the Å-range and conversely only the scattering intensity $S=|F(\mathbf{q})|^2$ is recorded in X-ray scattering experiments.

2.1.3 SCATTERING BY PROTEIN SOLUTIONS

The two-atom case explored in Section 2.1.2 can readily be extended to include an assembly of atoms if one considers the scattered wave to be sum of the scattered waves from all atom pairs in the assembly,

$$\mathrm{F}(\boldsymbol{q}) = \sum_j f_j e^{i\boldsymbol{q}\cdot\boldsymbol{r}_j} , \qquad (7)$$

where f_j denotes the q-dependant form factor and $\mathbf{r}_j$ is the position vector (with respect to an arbitrarily defined origin) of an atom *j*. If we now consider the scattering from a protein in solution, as shown schematically in Figure 9, the sum should be over all the atoms in the system, including all the atoms of the solvent and all the atoms in the protein. For this purpose however, where the protein is surrounded by a bulk of infinitely homogenous solvent, it is more useful to describe the form factor (f_j) and scattering vector ($\mathbf{r}_j$) of the protein atoms separately, and to instead define a scattering density (i.e. the scattering amplitude per volume) for the solvent, ρ_s. The system can now be divided into three scattering contributions. The first part (i) comes from the protein, the second part (ii) is the bulk solvent and a third contribution (iii) which represent the scattering from the excluded volume occupied by the protein.

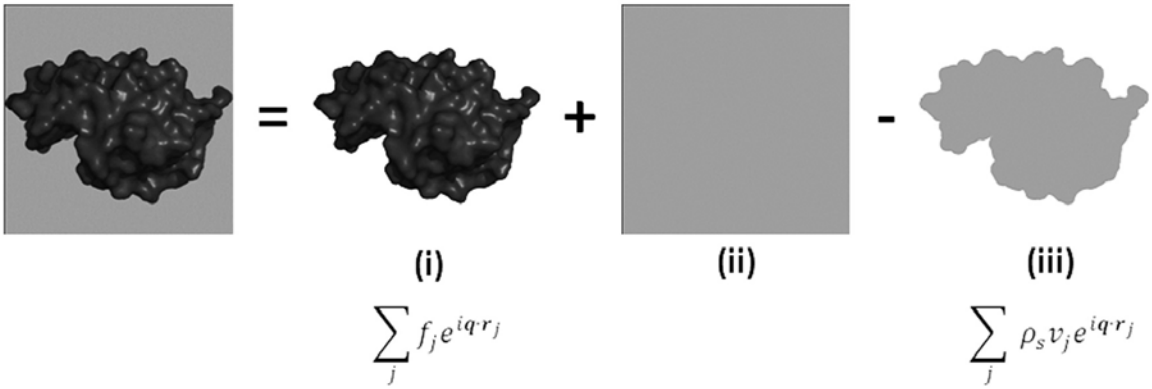


Figure 9. Scattering from a protein in solution. Indicated are the scattering contributions from **(i)** all the atoms in the protein, **(ii)** bulk solvent and **(iii)** the excluded volume.

Under these assumptions the scattering from the bulk solvent is dominant only at small angles, close to q=0, and contributes little to the resulting scattering at low to intermediate angles, so that the sum of the scattering contributions can be described by terms (i)-(iii),

$$\mathrm{F}(\boldsymbol{q}) = \sum_j (f_j - \rho_s v_j) e^{i\boldsymbol{q}\cdot\boldsymbol{r}_j} , \qquad (8)$$

where v_j represents the volume of atom *j*. and where the term $f_j - \rho_s v_j$ represents the contrast amplitude between atom *j* in the protein and the bulk solvent. In practice the bulk solvent does contribute to the scattering and this is traditionally dealt with by experimentally subtracting this contribution.

In a protein solution containing N identical protein molecules (monodisperse system) and where there is no particle interference between the molecules, the intensity of the

scattered waves can be found by summing up the intensities over all the molecules such that,

$$S(\boldsymbol{q}) = \sum_{n=1}^{N} |F_n(\boldsymbol{q})|^2 \,, \qquad (9)$$

but the scattering properties of the molecules differ from each other due to the fact that they take up different orientations. By using the famous Debye formula, which accounts for the rotational average in space $\langle \cdots \rangle$, the scattering intensity can thus be written as,

$$S(q) = N\langle |F(\boldsymbol{q})|^2 \rangle = N \sum_j \sum_k \left(f_j - \rho_s v_j\right)\left(f_k - \rho_s v_k\right) \frac{sinqr_{jk}}{qr_{jk}} \,, \qquad (10)$$

where r_{jk} is the distance between atoms *j* and *k* [115]. Therefore in order to calculate the expected scattering intensity from a particular protein solution we would need information about the number of molecules, *N*, in the solution, the form factors, *f*, and the atomic volume, *v*, of the atoms in the protein, the relative distance between all the atoms, r_{jk}, in the protein and the density of the solvent, ρ_s. In short, we would need information about the three-dimensional structural coordinates of the protein.

2.2 TIME-RESOLVED WIDE ANGLE X-RAY SCATTERING

2.2.1 SYNCHROTRON RADIATION

Synchrotron radiation is the electromagnetic radiation emitted by an accelerated charge travelling at relativistic speed. It is generated for experimental purposes in cyclic particle accelerators (Synchrotrons) where electrons are accelerated, in several stages, to reach velocities close to that of the speed of light in the synchrotron storage ring. The high-energy electron beam, having a final energy in the GeV range, is then directed into auxiliary devices such as insertion devices (undulators or wigglers) or bending magnets where strong magnetic fields are applied perpendicular to the beam. The magnetic fields cause a transverse acceleration of the relativistic electrons, causing them to emit radiation. Synchrotron radiation covers a broad spectrum of electromagnetic light ranging from microwaves to X-rays, and the desired wavelength region can be isolated by use of optical mirrors and by tuning the auxiliary device[116]. Synchrotron radiation also provides excellent characteristics as an X-ray source for solution scattering experiments. It provide an intensity and brightness several orders of magnitudes higher than conventional X-ray tubes and the short-pulsed nature of synchrotron radiation, due to the electrons travelling in bunches in the storage ring, can be used to measure transient events down to ~100 ps[117]. The time-resolved WAXS data in this thesis was collected at two third generation synchrotron facilities; the European Synchrotron Radiation Facility (ESRF) in France and the Swiss Light Source (SLS) in Switzerland.

2.2.2 PUMP-PROBE DATA COLLECTION

When measuring transient phenomena, a system is perturbed from its resting state and the resulting effects are monitored as a function of time. The time-resolution with which one can characterize the change in the system is often limited by how fast the data can be detected and stored, usually defined by the integration and readout time of the detector. An alternative approach which circumvents the temporal limitations in detector hardware is the pump-probe technique. In this technique, commonly used in time-resolved spectroscopy, the system is perturbed using an activating pulse (pump) and then monitored (probed) after a desired delay time (Δt) using a measuring pulse. The procedure is repeated for many such pump-probe cycles while the detector continuously integrates the data. In this case the time-resolution is limited by the length of the measuring pulse. By carefully selecting the time-delays between pump and probe the reaction process can be accurately sampled over several orders of magnitude.

The time-resolved pump-probe WAXS experiments in this thesis were conducted at the dedicated time-resolved beamline ID09B at ESRF. For these experiments a short laser pulse was used to pump the photoactive samples, placed in a capillary, and the resulting protein structural changes were probed after a preset time-delay by a polychromatic ($\Delta E/E=3\%$) X-ray pulse. At the ID09B single bunch X-ray pulses can be isolated by the use of a high-speed rotating chopper (Jülich chopper)[118-119].

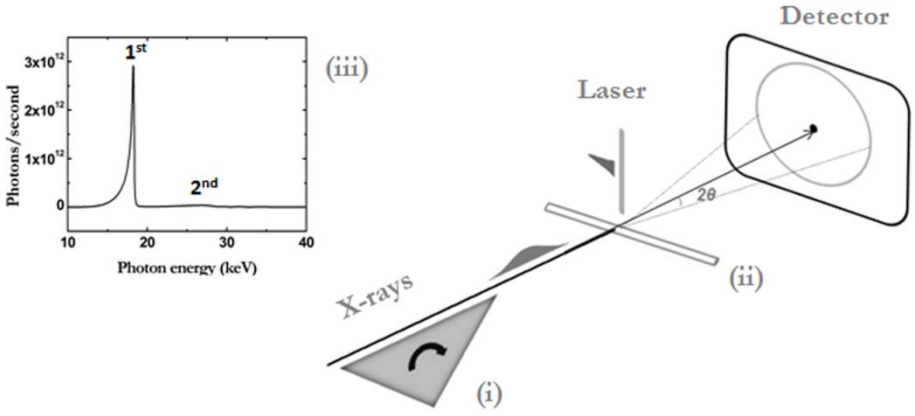


Figure 10. Schematic illustration of the pump-probe experimental setup at ID09B. Illustrating **(i)** the rotating Jülich chopper and (ii) the capillary containing the protein sample. For clarity the He-chamber located between the sample and the detector has been omitted. The polychromatic beam consisted of the first harmonic of the undulator emission spectrum (iii).

The length of the single ~100 ps duration X-ray pulse also establishes the fastest processes that can be studied. For processes occurring on slower time-scale it is preferable to use a multiple bunch X-ray pulse train to increase the flux through the sample. This can be achieved by varying the opening time of the Jülich chopper. The time-delay between the arrival of the laser pump and the X-ray probe is in turn set electronically by synchronizing the laser to the chopper. Between each pump-probe cycle the sample volume is also continuously replaced in order to minimize radiation

damage to the sample. The scattered X-ray are continuously collected onto a 2D CCD-camera which is set to integrate over several pump-probe cycles per image in order to accumulate high photon statistics. To further improve the statistics a helium filled chamber is placed between the sample and the detector. Helium, which contains fewer electrons then oxygen and nitrogen present in air, acts to reduce the background scattering inevitably caused by the primary beam. The two-dimensional concentric scattering patterns are then angularly integrated to yield one-dimensional scattering curves, displaying the scattering intensity as a function of the magnitude of the deflected beam (q). In order to compare the scattering from the light-perturbed system (positive time-delay) with that of the resting state, interleaving dark images are collected at a negative reference time delay. This enables a difference scattering analysis to be performed.

2.2.3 DIFFERENCE SCATTERING ANALYSIS

Since X-rays scatter from all the atoms in the sample, the measured scattering intensity contains the contribution, not only from the solute, but also from the solvent and the solvent-solute cross-terms. To increase the structural sensitivity towards the protein and to extract only the signal related to protein structural changes, difference scattering curves are calculated[120].

The raw scattering curves are first normalized according to the total scattering intensity recorded in an angular region (q_0) were little scattering changes are expected to occur (see Section 2.2.4),

$$S^{norm}(q,\Delta t) = \frac{S(q,\Delta t)}{\sum_{q0} S(q,\Delta t)}, \qquad (11)$$

The difference scattering is then calculated by subtracting the scattering curve from a negative reference time-delay (the X-rays arrive before the laser flash) from those of the positive time-delays,

$$\Delta S(q,\Delta t) = S^{norm}(q,\Delta t) - S^{norm}(q,\Delta t < 0), \quad (12)$$

Due to error propagation when calculating the difference scattering, the error due to measurements increases by approximately a factor √2 in the difference curve, the signal-to-noise ratio is inferior to that of the raw scattering curve. To ameliorate this effect it is practice to use the average of two or more flanking "dark" curves as a reference when calculating the differences. There are, however, many advantages of working with difference curves that more than compensates for the lower S/N[120]. These include (i) that the atomic scattering related to the background (such as capillary and solvent) cancels out in the differences, (ii) systematic errors intrinsic to the experimental setup and hardware are reduced, and perhaps most importantly (iii) that the relative scattering contribution from the transient species is greatly enhanced over those in the resting state. Despite these advantages, some of the scattering cross-terms such as those between the solute-solvent, and in the case of solubilised membrane

proteins between the protein-micelle and micelle-solvent, are not cancelled out using difference scattering analysis. These terms therefore have to be accounted for when doing data interpretation and structural predictions (see Section 2.3.1).

2.2.4 THERMAL RESPONSE OF THE SOLVENT

Photoactivation of proteins in solutions not only triggers the structural transition of the protein but also causes excess energy to be transferred into the surrounding solvent. For neat liquids the deposited energy in the centre of a laser excited volume is converted from a constant volume heating (Δt <10ns) to a constant pressure heating (Δt > 200ns) which causes the solvent to thermally expand[121]. For liquid water the average distance between molecules is ~3 Å and this spacing gives rise to a ~2Å^{-1} peak seen in the scattering curve for water solutions. As the solvent expands the average distance between water molecules increases which causes the scattering around the water peak to change. This signature effectively translates into a transient scattering change in the light-induced difference data (Figure 11). Although the major contribution from this thermal signal arise is in the 1.2-2.5 Å^{-1} q-region, subtle changes will influence the difference curves throughout the scattering region.

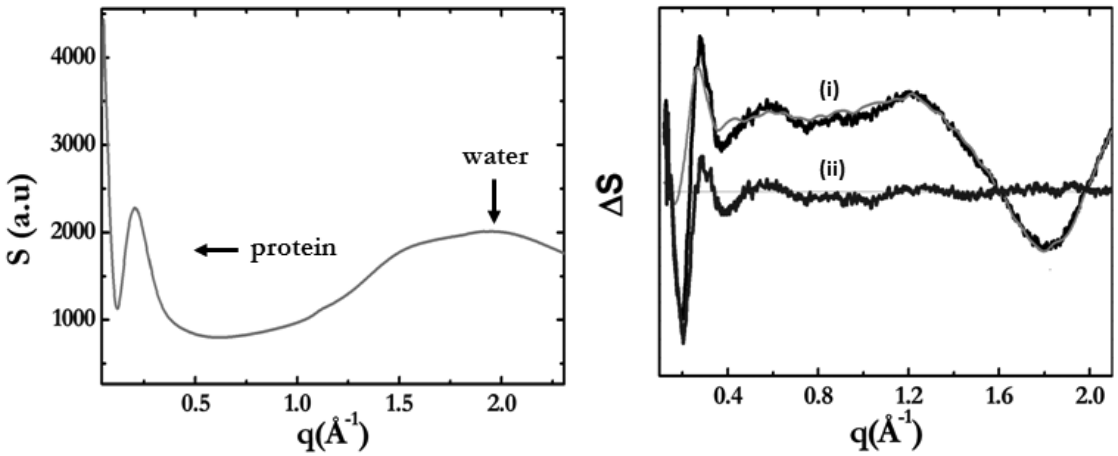


Figure 11. **Left:** Scattering intensity as a function of q for bR in solution. **Right:** (i) Laser-induced difference scattering signal (black) and pure thermal signal (light grey) for bR. (ii) The pure structural signal (dark grey) can be retrieved by scaling and subtracting the thermal signal from each time-delay.

This effectively means that at each time-delay the difference scattering signal is a mixture of a light-induced structural signal and a thermal signal. The challenge lies in separating these two contributions and the most effective way to do this is by isolating and subtracting the pure thermal signal from that of the mixture such that,

$$\Delta S^{light}(q,\Delta t) = \Delta S^{light+heat}(q,\Delta t) - k(\Delta t)\Delta S^{heat}(q)\,, \qquad (13)$$

where $k(\Delta t)$ is the time-delay dependant scaling factor for matching the high q-regions of both curves. The pure thermal signal (ΔS^{heat}) can in turn be identified by (i) performing static temperature measurements on the sample, (ii) geometrically displacing the pump source with respect to the probe, (iii) using an infrared laser to heat the solvent in the sample, or, (iv) for the fortuitous cases where the structural

signal decays on a much faster time-scale then the heating signal, by using time-delays only containing the thermal signal. The effect of the thermal signature can however also be used as an advantage. Regions in the difference scattering curve that undergo little or no change due the solvent expansion (so called isosbestic points) can for instance be used as normalization regions (q_0).

2.2.5 SINGULAR VALUE DECOMPOSITION

The structural signal observed in the difference scattering time-delays (ΔS^{light}) often represents a mixture of several structural species, each with their particular structural fingerprint, evolving at different rates. In order to identify the number of such independent structural finger prints it is necessary to subject the data to some sort of kinetic analysis. For this purpose, the use of singular value decomposition provides a model free starting point.

Singular value decomposition (SVD) is an algorithm commonly used in linear algebra for factorizing matrices. The idea is that for a $m \times n$ rectangular matrix A of real elements ($m \geq$ n) the SVD is defined by,

$$A = USV^T, \qquad (14)$$

where U is an $m \times m$ matrix having the property that $U^TU=I_n$ (where I_n is the $n \times n$ identity matrix), V is an $n \times n$ matrix such that $V^TV=I_n$ and S is a diagonal $n \times n$ matrix of nonnegative elements. The diagonal elements of S ($s_1, .., s_n$) are called the singular values of A, the columns of U and V are called the left and right single vectors of A, respectively. Traditionally the singular values, along with their corresponding columns in U and V, are ordered such that $s_1 \geq s_2 \geq s_n \geq 0$. With this ordering, the largest index r such that $s_r \geq 0$ is defined as the rank of A, and the first r columns of U comprise an orthonormal basis of A. An important property of the SVD is that these first columns of U, together with the corresponding columns in V and elements in S, provide the best least-squares approximation to the matrix A[122].

The concept can be readily applied to analyze time-resolved datasets from WAXS measurements. Here the matrix A is represented by the difference scattering dataset $\Delta S(q,\Delta t)$ containing m number of q-points and n number of time-delays. The corresponding SVD is then defined as,

$$\Delta S(q, \Delta t) = U(q)SV(\Delta t)^T, \qquad (15)$$

where U(q) represents the q-dependant orthonormal basis spectra (related to the structural fingerprints), V(Δt) represents the time-dependant amplitudes of the corresponding basis spectra and S represent the relative significance of each component. Although all the first components having $s \geq 0$ are necessary to minimize the difference between the SVD and the data, experimental data always contains a certain level of noise which is not related to the transient reaction studied. The task is

then to identify the minimum number of components necessary to reconstruct the data without adding additional noise components. This can be done by considering the magnitude of the corresponding singular values as well as by visual inspection of the singular vectors. Another objective approach is to evaluate the signal-to-noise ratio in the columns of U and V by calculating their autocorrelations[123] defined by,

$$C(U_i) = \sum_{j=1}^{m-1} U_{j,i} U_{j+1,i} \,, \qquad (16)$$

$$C(V_i) = \sum_{j=1}^{n-1} V_{j,i} V_{j+1,i} \,, \qquad (17)$$

where $U_{j,i}$ and $V_{j,i}$ represents the jth elements of the ith column of U and V, respectively. Since the column vectors are all normalized to unity, the vectors which display slow variations from row to row ('signal') will have autocorrelation values that are close to, but less, than 1 while rapid variations between rows ('noise') will result in autocorrelation values much less the 1. For column vectors with many elements autocorrelation values less then ~0.8 indicate a signal-to-noise ratio approaching 1.

Although SVD can limit the number of significant signal containing components in the dataset as well as provide information about the corresponding structural fingerprints and time-dependence it does not directly provides information about the rates at which the changes occur. For this purpose kinetic models usually have to be derived and tested.

2.3 STRUCTURAL ANALYSIS

2.3.1 PREDICTING THE X-RAY SCATTERING FROM ATOMIC STRUCTURES

The program CRYSOL[124] evaluates the scattering intensity directly from atomic coordinates of proteins by calculating the average scattering intensity over all particle orientations, using the multipole expansion method to describe scattering amplitudes,

$$S(\boldsymbol{q}) = \langle |F_a(\boldsymbol{q}) - \rho_s F_s(\boldsymbol{q}) + \Delta\rho F_b(\boldsymbol{q})|^2 \rangle \,, \qquad (18)$$

where $F_a(\mathbf{q})$ is the scattering amplitude from the protein *in vacuo* equivalent to Eq. 7, $F_s(\mathbf{q})$ is the scattering amplitude from the excluded volume, ρ_s is the density of the solvent (0.334 e/Å^3 for water), $F_b(\mathbf{q})$ is the scattering amplitude from a border layer of water molecules (hydration layer) and $\Delta\rho$ is the density contrast between the bulk and border layer. For wide-angles, however, the scattering from the border layer is minor[125] and for these purposes the last term can be omitted for this analysis. This reduces Eq. 18 to,

$$S(\boldsymbol{q}) = \langle |F_a(\boldsymbol{q}) - \rho_s F_s(\boldsymbol{q})|^2 \rangle \,, \qquad (19)$$

with the scattering intensity comparable to that of Eq. 10.

Membrane proteins, which are either solubilised in detergents or in lipids, tend to form a tight complex with their detergent/lipids molecules (protein-micelle complex). The scattering intensity from the particle therefore contains the contribution both from the protein and the micelle molecules. For solubilised membrane proteins this protein-micelle effect, which is due to the density contrast between the electron dense head groups of the detergent/lipid and the low density aliphatic tails[126], is commonly observed as a strong peak in the intermediate q-range (Figure 12). To account for this scattering effect in the experimental data, protein-micelles were constructed *in silico* and the scattering intensity was calculated for the entire protein-micelle complex.

Furthermore, to compensate for scattering effects due to random fluctuations of the atomic positions within the protein-micelle, a temperature term was introduced such that,

$$\mathrm{S}(\boldsymbol{q}) = e^{-B\left[\frac{q}{4\pi}\right]^2} \langle |F_a(\boldsymbol{q}) - \rho_s F_s(\boldsymbol{q})|^2 \rangle \,, \qquad (20)$$

where the *B*-factor represents the estimated mean square displacement of the atoms due to thermal disorder and where $F_a(\mathbf{q})$ and $F_s(\mathbf{q})$ represents the atomic scattering and excluded solvent amplitudes for the protein-micelle complex.

The difference scattering, $\Delta S(\mathbf{q})$, between the resting state and a potential active state structures was calculated based on Eq. 20 but with the assumptions that; (i) the change in micelle structure upon activation is negligible compared to the change in protein structure and (ii) the change in excluded volume for the resting and activated state is small. Under these assumptions the following expression for the difference scattering intensity between the active and resting state structures can be derived (see Appendix),

$$\Delta\mathrm{S}(\boldsymbol{q}) = e^{-B\left[\frac{q}{4\pi}\right]^2} g(\boldsymbol{q}) \left[S_a^{active}(\boldsymbol{q}) - S_a^{resting}(\boldsymbol{q})\right] , \qquad (21)$$

Where S_a^{active} and $S_a^{resting}$ is the atomic scattering for the protein in the active and resting state respectively and where,

$$g(\boldsymbol{q}) = \frac{1}{4}\left[1 + \left(\frac{S_a^{pm}(\boldsymbol{q}) - S_a^{m}(\boldsymbol{q})}{S_a^{p}(\boldsymbol{q})}\right)\right]\left[1 - \left(\frac{S_s^{pm}(\boldsymbol{q}) - S_{tot}(\boldsymbol{q})}{S_a^{pm}(\boldsymbol{q})}\right)\right] , \qquad (22)$$

represents scattering contrast terms, where the first term describes the contrast between protein (p) and micelle (m) and the second term describes the contrast between protein-micelle (pm) and excluded solvent (s). S_a^{pm} is the atomic scattering for the resting state protein-micelle complex, S_a^m is the atomic scattering from the resting state micelle, S_a^p is the atomic scattering from the resting state protein, S_s^{pm} is the excluded volume scattering for the resting state protein-micelle complex and S_{tot} is the total scattering intensity for the resting state.

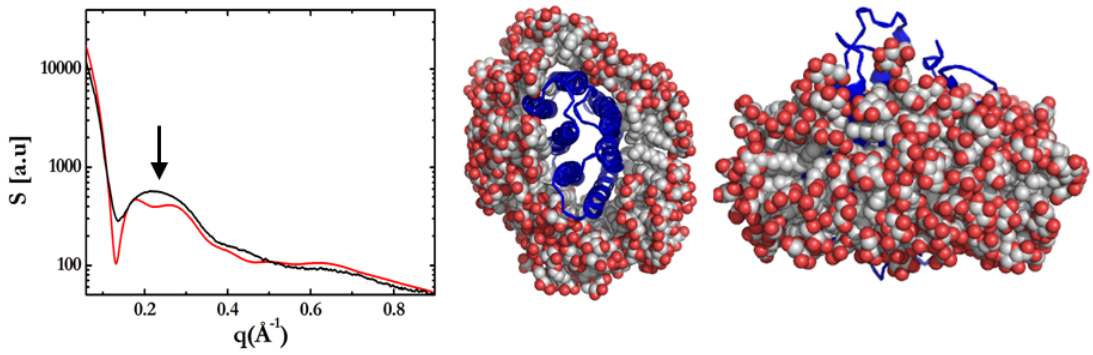


Figure 12. Scattering profile for detergent solubilized pR (black) and theoretical scattering prediction (red) using an atomistic protein-micell description. The arrow indicates the peak associated with interference between head groups and tails in the detergent/lipid protein-micelle.

2.3.2 THE INFORMATION CONTENT OF THE SCATTERING PATTERN

The structural information content in the scattering intensities varies as a function of the scattering vector and different q-regions are associated with different length distributions in the protein (Figure 13). At small angles (I) the electrons in the protein scatter in phase which contributes to a significant scattering intensity in this region. The information content in this region relates to overall size and shape of the particle and molecular properties such as the radius of gyration (R_g) and maximum distance of the particle (D_{max}) can readily be extracted from this region[101].

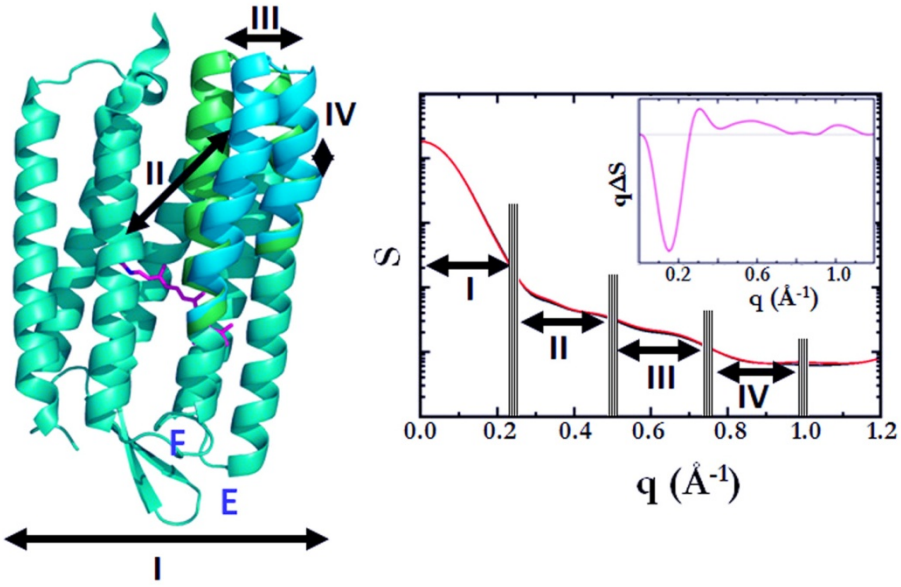


Figure 13. **Left:** Structural model of bR in its resting state conformation (green) and for a putative conformation (cyan) where helices E and F have been displaced. **Right:** The predicted scattering intensity from the same two structures (black and red) together with the approximate regions (I)-(IV) for which the corresponding distances in the protein would be observed. The inset shows the predicted difference scattering intensity (qΔS) from these structures.

Intermediate to wide angles (II-III) are representative of intramolecular structures (α- and β-structural content) and the corresponding distances between individual domains

and secondary structural elements within the protein. At even wider angles (IV) local distances involving individual side chains in turn contributes to the scattering intensity. The resolution range, however, quickly becomes limited at wider angles by the dominant influence of solvent scattering in the region >1.2 Å^{-1} accompanied by the limited scattering intensity at these angles. In difference scattering curves (Figure 13, inset) the q-regions undergoing changes are characterized by oscillating patterns (troughs and peaks in the intensity) related to the formation and subsequent depletion of distances in the protein structure. Coherent changes in secondary structural conformation (as exemplified in Figure 13) produce a characteristic difference scattering patterns with oscillations extending into the wide angle region.

The challenge lies in reversing the process in order to map the three-dimensional conformational changes related to the one-dimensional experimental difference scattering fingerprint. This has also been an important aspect of this thesis.

2.3.3 MOLECULAR DYNAMICS SIMULATIONS

In order to investigate the dynamics associated with protein structural rearrangements it can be useful to employ molecular dynamics simulations. In a molecular dynamics simulation, an atomic (microscopic) *in silico* system is constructed that mimics the conditions of the experimental (macroscopic) system as close as possible. In addition to the protein molecule, environmental components such as solvent molecules, ions and (in the case of membrane proteins) detergents or lipids are added to describe the system. From the atomic positions ($\mathbf{r}_i$) and mass ($\mathbf{m}_i$) and the knowledge about the forces ($\mathbf{F}_i$) governing their interactions, it is possible to simulate the time-evolution of the system by using Newton's law of motions,

$$\boldsymbol{F}_i = \boldsymbol{m}_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2}, \tag{23}$$

$$\boldsymbol{F}_i = -\frac{\partial V}{\partial \boldsymbol{r}_i}, \tag{24}$$

where V is the potential energy of the system. By combining the two equations,

$$-\frac{\partial V}{\partial \boldsymbol{r}_i} = \boldsymbol{m}_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2}, \tag{25}$$

it is possible to relate the derivative of the potential energy to the changes in atomic position as a function of time. Numerical integration of the equations is performed in small time-steps (~fs) and the system is typically followed for several nanoseconds. The coordinates as a function of time represents the trajectory of the system and after initial changes, the system eventually reaches an equilibrium state. By averaging over an equilibrium trajectory, macroscopic properties can be extracted. The latter relates to the ergodic hypothesis which states that the time average of a system equals that of the ensemble average. In this thesis molecular dynamics simulation was used to estimate free energy differences and simulate the protein-detergent system in Paper III.

3 RESULTS AND DISCUSSION

3.1 STRUCTURAL DYNAMICS OF PROTON PUMPS (PAPER II)

Since its first discovery over four decades ago[8], bacteriorhodopsin has developed into a bench-mark model system, not only for understanding proton-pumping in bioenergetics but also for understanding membrane protein transport in general. However, it was not until the discovery of the eubacterial homologue proteorhodopsin in seawater that the widespread role of retinylidene proteins in supplying energy to the marine biosphere was fully appreciated[26,34].

		Helix								R	
	Reference	A	B	C_{EC}	C_{CP}	D	E_{CP}	F_{CP}	G_{CP}	Int.	Late
K	Schobert et al., 2002	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00	0.96
	Edman et al. 1999	0.02	0.02	0.12	0.01	0.01	0.01	0.01	0.12	1.00	0.96
	Matsui et al., 2002	0.12	0.07	0.09	0.06	0.09	0.08	0.07	0.06	0.75	0.92
L	Lanyi and Schobert, 2007	0.09	0.08	0.11	0.09	0.07	0.06	0.10	0.13	1.00	0.93
	Kouyama et al., 2004	0.20	0.23	0.18	0.13	0.16	0.14	0.17	0.23	0.98	0.88
	Lanyi and Schobert, 2003	0.12	0.29	0.30	0.29	0.32	0.18	0.38	0.32	1.00	0.96
	Royant et al., 2000	0.36	0.26	0.54	0.28	0.30	0.23	0.44	0.42	0.76	0.86
	Edman et al., 2004	0.37	0.40	0.59	0.47	0.37	0.48	0.61	0.50	0.68	0.81
WAXS	bR Intermediate state	0.00	0.00	1.00	0.09	0.00	0.31	2.23	0.22	0.34	-
M	Schobert et al. 2003	0.13	0.24	0.25	0.26	0.24	0.26	0.38	0.29	0.91	0.96
	Lanyi and Schobert, 2006	0.23	0.19	0.30	0.25	0.19	0.18	0.38	0.39	1.00	0.96
	Takeda et al., 2004	0.22	0.39	0.39	0.14	0.20	0.21	0.30	0.48	1.00	0.96
	Facciotti et al., 2001	0.26	0.23	0.38	0.28	0.28	0.23	0.42	0.38	1.00	0.96
	Luecke et al., 2000	0.32	0.28	0.36	0.38	0.21	0.34	0.44	0.41	0.94	0.92
	Lanyi and Schobert, 2002	0.44	0.46	0.59	0.42	0.41	0.31	0.66	0.66	0.96	0.96
	Luecke et al., 1999	0.28	0.38	0.38	1.04	0.39	0.38	0.91	0.49	0.75	0.70
	Sass et al., 2000	0.52	0.57	0.77	0.56	0.50	0.40	0.90	0.82	0.98	0.96
	Subramaniam and Henderson, 2000	0.40	0.55	0.32	0.51	0.39	0.61	1.84	1.05	0.69	0.82
WAXS	bR Late state	0.00	0.00	1.54	0.13	0.00	0.45	2.92	0.22	-	0.25
N	Schobert et al., 2003	0.27	0.49	0.56	0.54	0.49	0.40	0.65	0.66	0.77	0.80
O	Rouhani et al., 2001	0.71	0.86	1.00	0.49	1.19	0.56	0.50	0.57	0.68	0.54
	Rouhani et al., 2001	0.54	0.90	1.13	0.55	1.16	0.59	0.50	0.53	0.92	0.66
WAXS	pR Late state	0.00	0.00	0.87	0.00	0.00	0.77	2.62	1.17	-	0.28

Table 1. Overview of the helical movements observed in intermediate trapping and mutational studies of bacteriorhodopsin. The RMSD (Å) as compared to the corresponding resting state structure is indicated along with the R-factors calculated against the intermediate or late state basis spectra from bR. Yellow indicates that the movement is larger than the average listed movement. Orange indicates that the movement is larger than the average listed movement plus one standard deviation. The optimal results from WAXS modeling are highlighted in grey and green.

Bacteriorhodopsins remarkable stability combined with a high natural abundance has made it an excellent candidate for validating new biophysical techniques. A plethora of spectroscopic studies have been able to pinpoint the key important residues (see Section 1.2.1) involved in the vectorial proton transport and electron diffraction

studies[127-128] of bacteriorhodopsin provided early insights into the structure of this membrane protein. Developments in the field of membrane protein X-ray crystallography[129] has since delivered a number of high-resolution crystal structures of the bacteriorhodopsin resting state[130-131]. As the method of producing well diffracting crystals evolved, attention was focused on resolving structural intermediates of bacteriorhodopsin. This was principally pursued using intermediate-trapping strategies (see Section 1.3), where illuminated crystals were rapidly trapped at cryogenic temperatures, or by using site-directed mutagenesis to produce intermediate state analogous[132].

Using these techniques structures pertaining to, three K-intermediate[90,133-134], five L-intermediate[135-139], nine M-intermediate[132,140-147], one N-intermediate[146] and one O-intermediate[148] have been reported (Table 1). The structural picture that emerged from this wealth of crystal structures was however in many respects ambiguous. Only three significant movements, involving the cytoplasmic portions of helices F and G and the extracellular side of helix C, were reproducibly observed, and the magnitudes of these conformational changes differed significantly between different structures[149]. Furthermore, rearrangements in water mediated hydrogen bonding networks were observed to be coupled to these movements[135-136], although the timing between the movements and the primary proton transfer event could not be unambiguously established from these static structures.

In order to simultaneously study the timing and conformational changes occurring in bacteriorhodopsin, alternative experimental approaches are thus essential. In this study time-resolved WAXS was applied to study the conformational dynamics of both bacteriorhodopsin and the eubacterial homologue proteorhodopsin.

3.1.1 TIME-RESOLVED WAXS DATA COLLECTION

Time-resolved pump-probe WAXS data was collected on concentrated solutions of detergent solubilised bacteriorhodopsin and proteorhodopsin at the time-resolved beamline ID09B at the ESRF. Samples were mounted in a sealed quartz capillary and the photoreaction was initiated using a 527 nm laser pulse with 100ns FWHM pulse duration. The experimental time delay (Δt) between the arrival of the laser pulse and the polychromatic X-ray pulse was controlled electronically, and the duration of the probing X-ray pulse (either 360 ns, 2μs or 20μs) was adjusted to the suit the time-delays probed. The WAXS data was recorded onto a CCD camera where each image was exposed for several pump-probe cycles. The capillary was also continuously translated between each pump-probe cycle to prevent radiation damage accumulating in the sample volume. A total of nine time-delays were collected in the range of 360 ns to 100 ms and additional "dark" images were collected as reference points at the start and end of each time-series. The time-window was chosen to include all the major spectroscopic intermediates identified for both samples (see Section 1.2.1 and 1.2.2). Following radial integration and normalization of each WAXS image, the difference

scattering intensities were calculated by subtracting the average of the flanking "dark" curves. The thermal signal due to laser-induced heating was in turn removed by scaling and subtracting an experimentally determined pure heating signal from each time-delay (see Section 2.2.4).

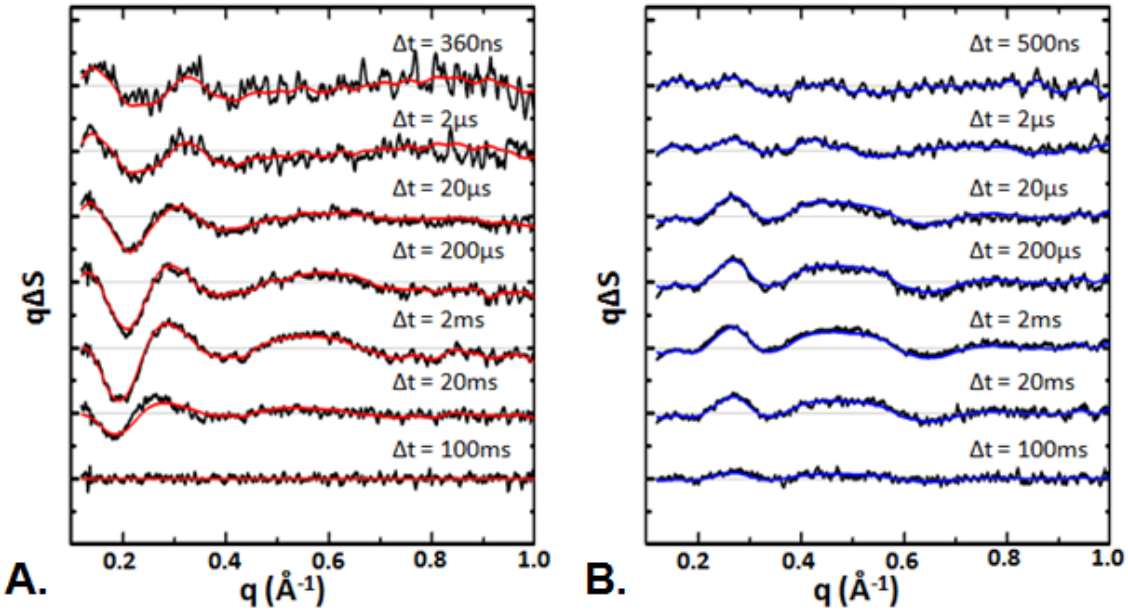


Figure 14. Representative difference WAXS data for (A) bacteriorhodopsin and (B) proteorhodopsin as a function of time-delay after the arrival of the laser pulse. For visualization purposes the difference scattering ΔS has been multiplied with q. Thin grey lines indicate ΔS=0 and red/blue lines corresponds to the predicted time-delays from the kinetic model.

The resulting thermal free difference scattering data ΔS shown in Figure 14 thus represent the scattering changes solely due to light-induced structural changes occurring in the protein sample. The WAXS difference data indicated that major oscillations occur, for both bacteriorhodopsin and proteorhodopsin, in the q-region of $0.4 < q < 0.6$ Å^{-1} where changes in secondary structure would be expected. Although the difference data appeared similar for the two proteins, there were distinct differences in positions and amplitudes of the oscillating features as well as in the time-scales by which they evolved.

3.1.2 KINETIC ANALYSIS

In order to extract time-independent basis spectra U(q) from the time-resolved dataset ΔS(q, Δt), each dataset was subjected to spectral decomposition according to the following sequential model,

$$Early \xrightarrow{k_1} Intermediate \xrightarrow{k_2} Late \xrightarrow{k_3} Resting\ , \qquad (26)$$

such that the optimal rate constants and basis spectra were retrieved by least-squares refinement according to,

$$\min\ [\Delta S(q,\Delta t) - U(q)\cdot C(\Delta t)]^2, \qquad (27)$$

where C(Δt) describes the time-dependant population of the basis spectra (Figure 15A and 16A) and was calculated from the integrated rate equations of Eq. 26. The fact that three components was sufficient to provide a complete description of the time-evolution of both datasets is apparent from the recalculated time-delays shown as solid lines in Figure 14 as well as from the amplitudes of the best-fit linear combinations of the three basis spectra to each time-delay in Figure 15 and 16.

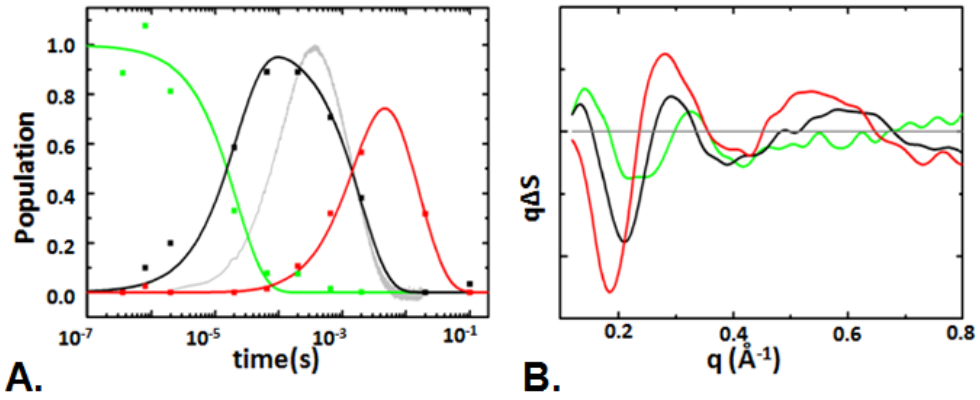


Figure 15. Spectral decomposition of WAXS data from bR. (A) Transient population of the early (green), intermediate (black) and late (red) basis spectra illustrated in (B). Squares indicate the predicted amplitudes for optimal linear combinations of the three basis spectra to each time-delay. The transient change in absorption at 410 nm is shown as a comparison in grey.

The spectral decomposition indicated that for bacteriorhodopsin the early component (green) decays to the intermediate component (black) with a time constant of 22 ± 2 μs, the intermediate component in turn decays to the late component (red) with a lifetime of 1.9 ± 0.4 ms and the late component finally relaxes back to the bacteriorhodopsin resting state (zero line) on the time-scale of 16 ± 1.5 ms.

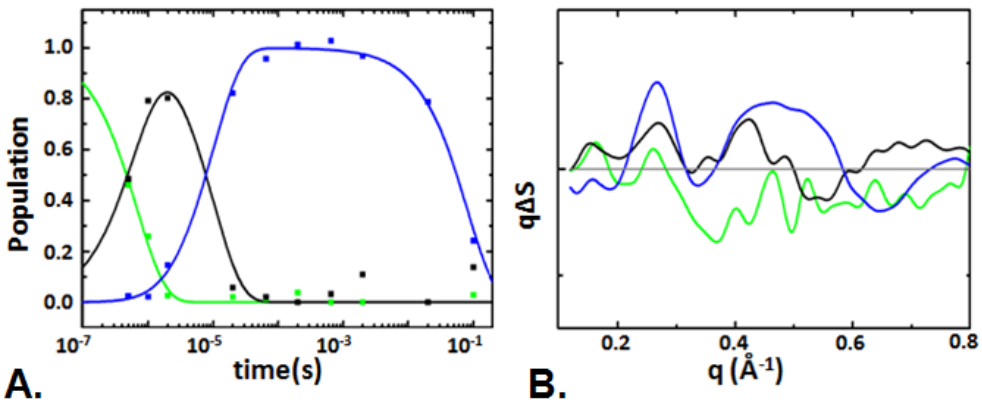


Figure 16. Spectral decomposition of WAXS data from pR. (A) Transient population of the early (green), intermediate (black) and late (blue) basis spectra in (B). Squares indicate the predicted amplitudes for optimal linear combinations of the three basis spectra to each time-delay.

For proteorhodopsin, the species follow a slightly different time-evolution, where the early component (green) more rapidly transform, 670 ± 140 ns, into a transient intermediate component (black) which in turn converts to a long-lived late component (blue) with a time constant of 10 ± 1 μs. The late component persists for 79 ± 4 ms until the proteorhodopsin resting state is regenerated. The relaxation times for both proteins were also in good agreement with those from spectroscopic characterization.

Further mechanistic insight was gained from comparing the WAXS time-scales with those from transient absorption spectroscopy. The 410 nm transient signal (indicative of the bacteriorhodopsin M-intermediate) showed that the early-to-intermediate conformational state transition preceded the Schiff base deprotonation event. Moreover the temporal overlap between the intermediate conformational state and the time-scales of the L- and early M-intermediates (see Section 1.2.1) implied that these spectral intermediates have very similar global conformations. Conversely the intermediate-to-late conformational state transition also occurs on comparable time-scales, such that the late conformation encompasses the late M-, N- and O-intermediates. For proteorhodopsin a similar conceptual framework applies, but where the formation of the M-intermediate occurs on an even faster time-scale[40] and where the population does not build up to the same appreciable levels as for bacteriorhodopsin. The picture that emerged from these kinetic considerations of the WAXS data implied that spectrally distinct photocycle intermediates involve rather similar global conformational changes.

3.1.3 STRUCTURAL ANALYSIS AND REFINEMENT

As a first attempt to interpret the nature of the conformational changes residing in the WAXS basis spectra the theoretical difference scattering was calculated for all deposited intermediate structures of bacteriorhodopsin. The agreement between the experimental basis spectra and theory was then quantified using an R-factor, comparable to that used in crystallography,

$$R = \frac{\sum_q \sqrt{\left(\Delta S^{theory} - \Delta S^{data}\right)^2}}{\sum_q \sqrt{\left(\Delta S^{data}\right)^2}}, \qquad (28)$$

The result indicated that, in contrast to earlier studies of the haemoglobin:CO complex[113], none of the intermediate state structures yielded an entirely satisfactory match (R > 0.50) to either the intermediate or late conformational basis spectra (Table 1). Due to the poor signal-to-noise ratio of the early basis spectra for both proteins these were excluded from the comparison. In order to improve the structural modelling of the data, individual motions were extracted from the set of crystallographic structures and used for iterative refinement against the data. Regions prone to undergo motion were in turn identified based on the heuristic notion that movements in α-helices are restricted almost only to regions where the hydrogen bonding pattern is perturbed by the presence of proline residues. In order to expand

the conformational space investigated beyond the motions observed in the trapped crystallographic intermediates, additional putative motions were also added to the refinement protocol. For bacteriorhodopsin a total of 8 moving regions (Figure 17) were selected based on these considerations. For proteorhodopsin, for which no resting state crystallographic structure exists, a set of 7 similar motions were applied based on the proteorhodopsin homology model and the motions observed for bacteriorhodopsin.

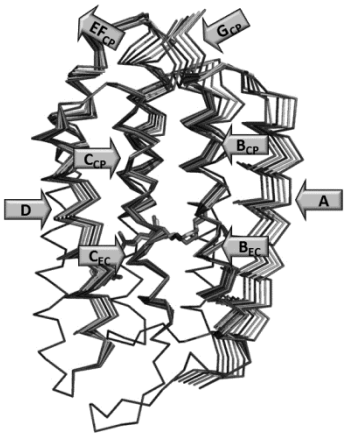


Figure 17. Schematic illustration of the helical movements included in the bR refinement. A similar set of movements was also used for the pR refinement.

Linear movements were then generated for both proteins by calculating the difference vector between each intermediate (I) and corresponding resting state (R) structure, and then adding the difference vector to a common resting state (R') used for refinement. Such that,

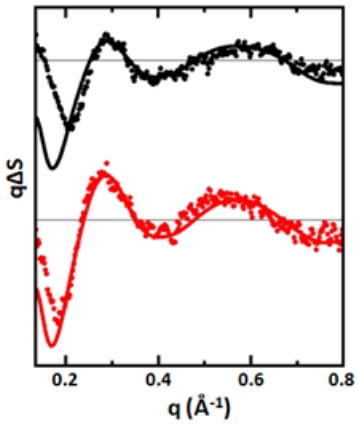
$$I'\begin{bmatrix} \boldsymbol{x_0} + k(\boldsymbol{x_2} - \boldsymbol{x_1}) \\ \boldsymbol{y_0} + k(\boldsymbol{y_2} - \boldsymbol{y_1}) \\ \boldsymbol{z_0} + k(\boldsymbol{z_2} - \boldsymbol{z_1}) \end{bmatrix} = R'\begin{bmatrix} \boldsymbol{x_0} \\ \boldsymbol{y_0} \\ \boldsymbol{z_0} \end{bmatrix} + k\left(I\begin{bmatrix} \boldsymbol{x_2} \\ \boldsymbol{y_2} \\ \boldsymbol{z_2} \end{bmatrix} - R\begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{y_1} \\ \boldsymbol{z_1} \end{bmatrix} \right), \qquad (29)$$

where **x**, **y** and **z** indicates the coordinates of the atoms included in the moving region. By varying the magnitude k of the difference vector, models (I') displaying different degree of movement could be produced. A number of models with varying degrees of motions were produced for each moving region as well as for combinations of movements, so as to produce a conformational space grid.

The theoretical difference scattering curves, calculated between each candidate model and the resting state, were then globally fitted to the intermediate and late basis spectra for each protein. The significance of the movement was subsequently scored based on the concomitant decrease in R-value. To prevent over fitting of the data, the refinement was performed in a hierarchal manner where each step in the refinement only added one additional movement.

The refinement result for bacteriorhodopsin (Figure 18) indicated that a co-varying motion of helices E and F produced the most significant drop in R-value (0.55).

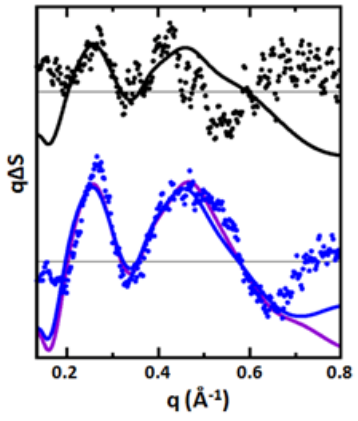
Combining this motion with an inward movement of the extra cellular half of helix C further decreased the R-value (0.32), while adding a third component did not significantly improve the agreement to the experimental data.



Residues		R-factor		
Bacteriorhodopsin		1st component	2nd component plus helix EF_{CP}	3rd component plus helices EF_{CP} & C_{EC}
Helix EF_{CP}	150-183	**0.546**		
Helix C_{EC}	79-91	0.989	**0.317**	
Helix G_{CP}	221-232	0.989	0.492	**0.279**
Helix C_{CP}	89-103	0.986	0.431	0.280
Helix B_{CP}	48-63	0.989	0.545	0.298
Helix D	103-128	0.989	0.585	0.309
Helix B_{EC}	35-52	0.989	0.542	0.319
Helix A	5-39	0.989	0.545	0.334

Figure 18. Refinement results for bR. The best fit predictions (solid lines) for the intermediate (R_{int}=0.34, black) and late (R_{late}=0.25, red) state basis spectra. The table shows the refinement statistics for the global fitting of both basis spectra to the indicated component. R represents the global R-factor

The results from the proteorhodopsin refinement (Figure 19) also indicated that the combined EF motion (0.83), followed by a cytoplasmic motion of helix C (0.66) provided the largest decrease in R-value. A further decrease was achieved when allowing for an inward movement of the cytoplasmic half of helix G (0.51). It was, however, apparent that the agreement for the intermediate conformational state was considerably worse than for the late state. This is likely due to the poorer signal-to-noise of this component as a consequence of lower sampling in this time regime. The same refinement procedure was therefore explored for only the late state basis spectrum. This picked out the same movements but produced a lower R-value (0.28).



Residues		R-factor		
Proteorhodopsin		1st component	2nd component plus helix EF_{CP}	3rd component plus helices EF_{CP} & C_{EC}
Helix EF_{CP}	164-197	**0.831**		
Helix C_{EC}	234-246	0.893	**0.656**	
Helix G_{CP}	93-101	0.834	0.748	**0.506**
Helix D	120-142	0.854	0.822	0.671
Helix B	57-81	0.905	0.884	0.679
Helix C_{CP}	104-114	0.864	0.833	0.706
Helix A	25-51	0.876	0.847	0.707

Figure 19. Refinement results for pR. The best fit predictions (solid lines) for the intermediate (R_{int}=0.81, black) and late (R_{late}=0.33, blue or R_{late}=0.28, violet) state basis spectra. The table shows the refinement statistics for the global fitting of both basis spectra to the indicated component. R represents the global R-facor.

The magnitudes of the refined helical movements are summarized for both proteins in Table 1. For bacteriorhodopsin it is worth noting that the 1.0 Å inward movement of helix C and the 2.2 Å outward tilt of helix F observed in the intermediate state become significantly larger during the transition to the late conformational state, for which the corresponding magnitudes are 1.5 Å and 2.9 Å. For proteorhodopsin it is striking that not only are the principal motions shared with bacteriorhodopsin but also the amplitudes of these motions.

3.1.4 UNIQUENESS OF THE SOLUTION

Questions that naturally arise include; how unique are these solutions? Would it be possible to achieve a similar agreement to the data by a random combination of helical movements? To test this, 1000 random candidate conformations were generated for each protein using the program CONCOORD[150] and the corresponding R-values were calculated.

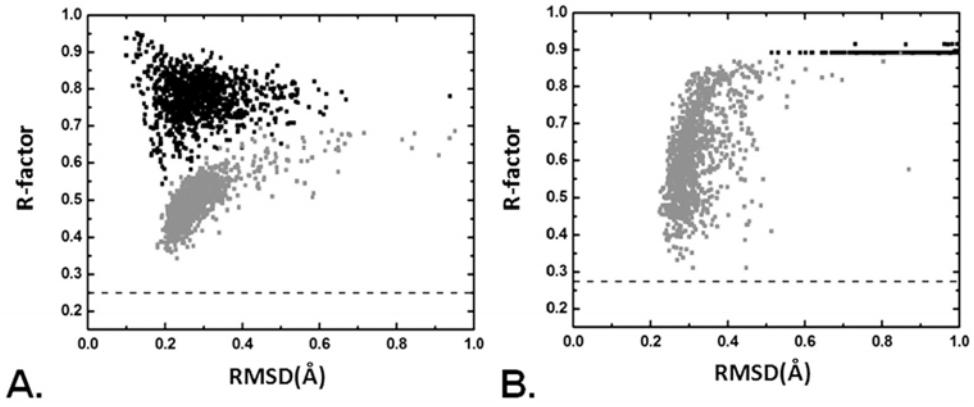


Figure 20. R-values as determined for (A) bR or (B) pR against the late state basis spectrum when 1000 random structures were generated; starting from either the resting state structure (black) or the refined late state structure (grey). The RMSD was calculated for the C_α-atoms relative to the starting structure. The dashed lines indicate the R-value from structural refinement.

The result was plotted as a function of root mean square deviation (RMSD) of the Cα-atoms from the starting model, where the CONCOORD starting model consisted of either the resting state structure or the refined intermediate/late state structures. It was apparent from this analysis that the further one departs from the refined solutions the worse the agreement to the data becomes. More importantly none of the randomly generated structural models were able to improve the R-values, indicating that the solution from structural refinement reflects a global minimum.

3.1.5 SUMMARY

The conformational picture (Figure 21) that emerges from the time-resolved WAXS data in this work differs from that gained using kinetic crystallography. First, the light-induced conformational changes are considerably larger at room temperature and in solution as a consequence of the protein not being constrained by a crystal lattice. For bacteriorhodopsin this was manifested in a 60 % larger movement of helix F and a two and a half times larger movement of helix C in the late conformational. Secondly, and

perhaps most importantly, is that the primary proton transfer event from the Schiff base to Asp85 does not mark a boundary separating two different types of motions. It rather appears that the inward collapse of the extracellular half of the proton transport channel and the subsequent opening of cytoplasmic half are part of the same principal motion, where the deprotonation of the Schiff base marks the threshold governing the magnitudes of these motions.

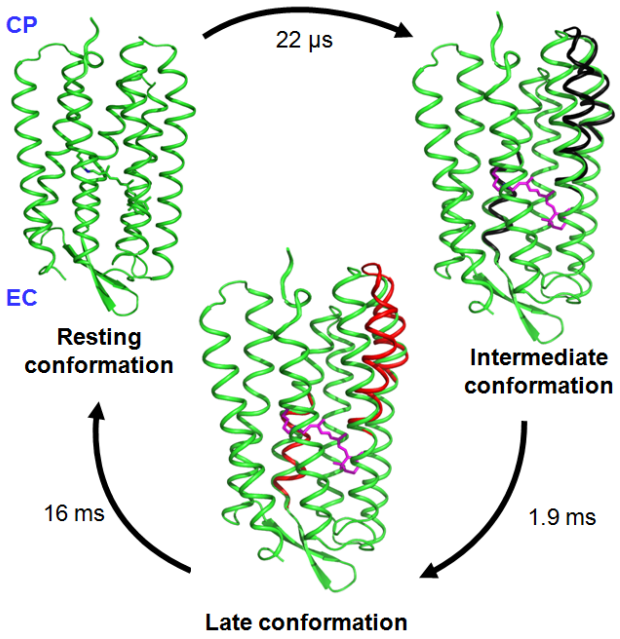


Figure 21. Conformational cycle of bR as determined using time-resolved WAXS. The resting conformation (green), intermediate conformation (black) and late conformation (red) are illustrated along with their respective time-constants. Significant inward movements of helix C and outward movements of the cytoplasmic portions of helices E and F were observed already in the intermediate conformation. These motions grew in amplitude following Schiff base deprotonation before the protein returned to the resting state. A similar picture also emerged for pR. CP: cytoplasmic side, EC: extracellular side.

This structural mechanism simplifies the description of the protein conformational dynamics involved in proton translocation and further supports the notion that all major motions in retinylidene proteins originate in a steric clash with the photoisomerized retinal. Comparison of the major structural findings for bacteriorhodopsin and proteorhodopsin also suggest that the conformational changes that occur during proton pumping in archael bacteriorhodopsin also occur in the eubacterial proteorhodopsin. Considering the vast biomass of bacterioplankton present in the world's oceans[26,34], we conclude that this basic structural mechanism is responsible for a sizeable input of energy into the marine ecosystem.

3.2 CONFORMATIONAL ACCELERATION IN IODORETINAL-SUBSTITUTED PROTEORHODOPSIN (PAPER III)

The first functional event in all rhodopsins is the absorption of light by the retinal molecule. In proteorhodopsin (pR) the retinal molecule is covalently attached via a protonated Schiff base to Lys321. After the absorption of a single photon the *all*-trans retinal rapidly isomerizes to adopt a 13-cis conformation, which induces a sequence of specific structural changes in the vicinity of the retinal that with time propagate throughout the protein. A crucial role in the process of the coupling the isomerisation event to the protein conformational changes has been attributed to steric interactions between the retinal's C-20 methyl group and the surrounding chromophore environment[151-152].

In this study, we explored the influence of this group by chemically modifying the retinal chromophore of proteorhodopsin and using time-resolved WAXS to characterize the resulting perturbations. More specifically, the retinal of native proteorhodopsin was replaced with a halogenated derivative, 13-desmethyl-13-iodoretinal, whereby the retinal's C-20 methyl group (which is attached to C-13 of the retinal) was substituted with iodine to form the bulkier I-20. In order to evaluate the effect of the retinal substitution, comparative measurements were done for both the native and the substituted proteorhodopsin.

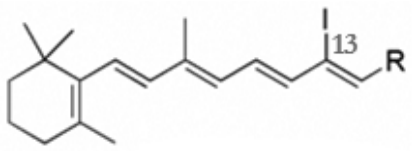


Figure 22. Illustration of 13-desmethyl-13-iodoretinal chromophore where the C-20 methyl group has been substituted with an iodine atom.

Proteo-opsin (the protein lacking the chromophore) was therefore first heterologously expressed in *E.coli*. The retinal was then reconstituted at the stage of cell lysis by adding either all-*trans* retinal (yielding native pR) or 13-desmethyl-13-iodoretinal (yielding 13-I-pR) to the apoprotein membranes. Following subsequent steps of solubilisation in octyl-glucopyranoside (OG) and purification, the two detergent solubilised forms of proteorhodopsin were both functionally and structurally characterized using transient absorption spectroscopy and time-resolved WAXS. The time-resolved WAXS data was collected at the ID09B beamline at the ESRF using an identical pump-probe protocol developed and described for Paper II. The time-delays collected (ranging from 2μs to 100ms) were chosen as to coincide with the major late state structural species previously observed for the native form (see Paper II).

3.2.1 SPECTROSCOPIC CHARACTERIZATION

The absorption spectrum of 13-I-pR indicated that the replacement of the methyl group attached to C-13 with iodine results in a substantial change in the optical

properties of the protein. As a consequence of the iodine substitution the absorption maximum of 13-I-pR was red shifted 23-26 nm compared to native proteorhodopsin. These observations are in agreement with earlier studies of bacteriorhodopsin[153] where a similar substitution was made. Moreover, 13-I-pR also displayed reversible pH dependant absorption characteristics and consequently underwent a ~20nm blue shift when the pH was elevated from 7 to 9. This indicated that the protonation state of the retinal and all key residues in the vicinity of the Schiff base were not significantly affected by the iodine substitution. All time-resolved measurements were subsequently conducted at pH 9 which mimics the alkaline environment of oceanic seawater.

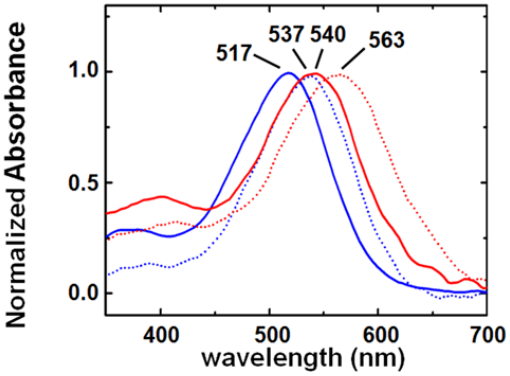


Figure 23. Optical absorption spectra of native pR (blue) and 13-I-pR (red) with their respective absorption maxima at pH 7 (dotted lines) and pH 9 (solid lines).

Time-resolved spectroscopic characterization of 13-I-pR further indicated that the kinetics of the resting state recovery (500 nm) as well as the decay of the final O-intermediate (610 nm) was approximately an order-of-magnitude faster than for the native proteorhodopsin. Furthermore, while the two spectroscopic states for the native pR displayed an almost perfect kinetic correlation, there was a clear separation of the two intermediates when the retinal had been substituted.

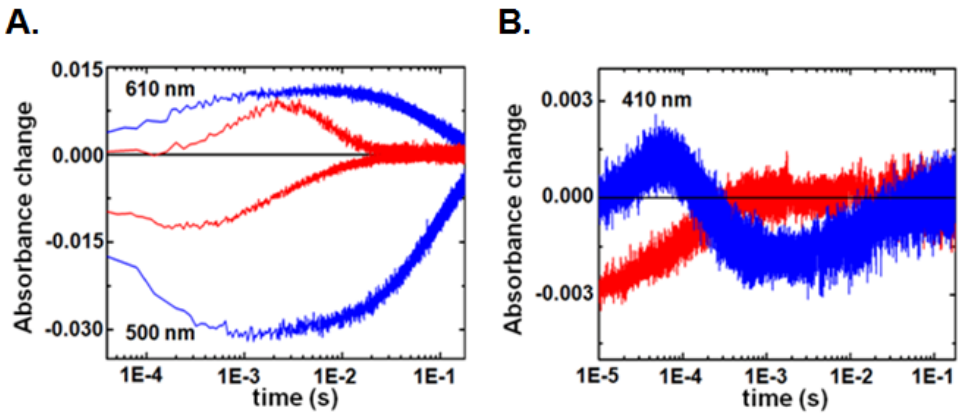


Figure 24. Single wavelength transient absorption changes for native pR (blue) and 13-I-pR (red). The kinetics associated with the formation and relaxation of (A) the resting state (500 nm) and final O-state (610nm) and (B) the M-state (410nm).

Characterization of the M-intermediate (410 nm) also showed that while native pR produced a spectroscopically distinguishable protonated Schiff base intermediate, no significant population was detected for 13-I-pR over this time-window. Thus from the time-resolved absorption changes it was apparent that a more complex photochemistry governs the photocycle of 13-I-pR.

3.2.2 TIME-RESOLVED WAXS CHARACTERIZATION

Representative WAXS time-delays for both native pR and 13-I-pR are shown in Figure 25. From the difference scattering data alone it was already apparent that the amplitude of the signal at $\Delta t = 2\mu s$ was stronger for the modified retinal form compared to the corresponding time-delay in the native form. It was also clear that the two later time-delays (Δt = 20 ms, 100 ms) for 13-I-pR contained no discernable difference signal due to protein conformational changes, whereas significant oscillations persisted for the native form until Δt = 100 ms.

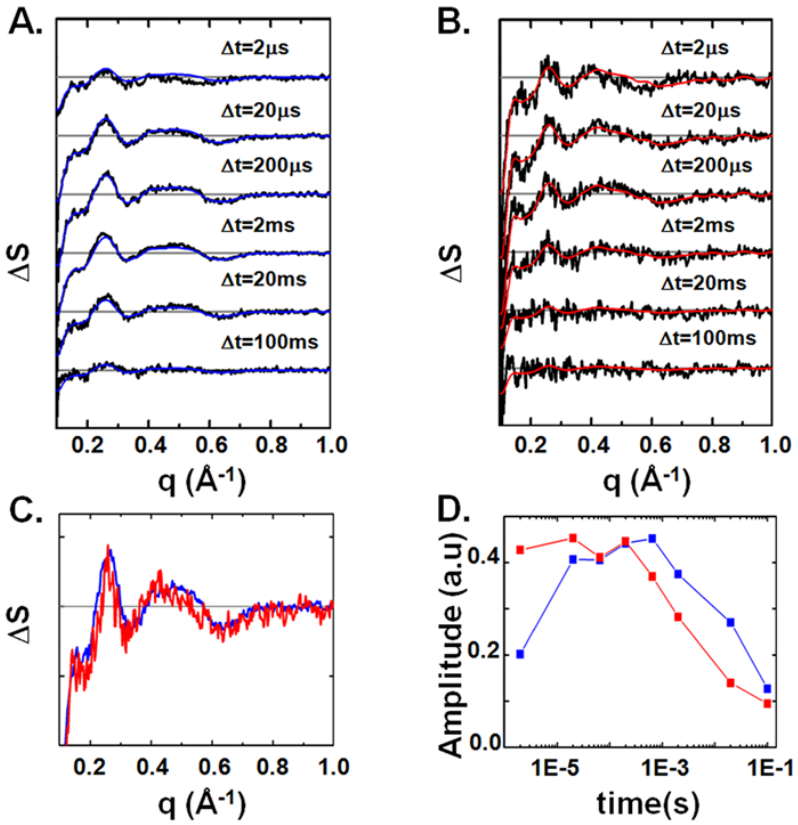


Figure 25. Difference WAXS time-delays for (A) pR and (B) 13-I-pR. Solid lines indicate the recalculated time-delays using the first component from SVD. (C) Basis spectra and (D) corresponding amplitudes from singular value decomposition analysis for pR (blue) and 13-I-pR (red).

In order to quantify the experimental differences as a function of time, singular value decomposition of each dataset was performed. In this context of analyzing data from two similar protein forms, the use of singular value decomposition has the principal advantage that it does not assume any prior kinetic knowledge and therefore avoids the potential problem of model bias. Based on the evaluation of the singular vectors,

singular values and the singular vector autocorrelations (see Section 2.2.5), one-signal containing component was identified for each of the native pR and 13-I-pR datasets.

The time dependence of the light-induced conformational changes is represented by the amplitudes of the singular vectors. Modelling of these amplitudes with single-exponentials rate constants indicated that the transient conformational state of native pR rises and decays with half-times of 2±1 μs and 49±11 ms, respectively. These values were also in agreement with respective time-constants previously observed for the late conformational state of native pR (Paper II). In contrast when the C-20 methyl group was chemically substituted with iodine, these rates were significantly accelerated, such that the transient conformation of 13-I-pR had reached a maximum population already after 2 μs. The population of the transient conformation furthermore decayed significantly faster, with a half-time of 2±0.4 ms. The accelerated structural changes observed for proteorhodopsin in its iodinated form were also consistent with the conclusions drawn from the spectroscopic analysis. Despite the clear differences in the rates of the structural components, it was remarkable that the corresponding basis spectra of the two forms were almost identical (displayed a 91 % correlation) within the observed signal-to-noise. This indicated that the structural changes occurring within the protein was not substantially affected by the retinal substitution.

3.2.3 STRUCTURAL ANALYSIS USING MOLECULAR DYNAMICS

In order to model the structural changes that occur in the transient conformations, the basis spectrum of each protein form was first evaluated against a grid of linear helical movements (see Paper II). The refinement identified the same major motions and amplitudes, involving the cytoplasmic parts of helices E, F and G and the extracellular portion of helix C, previously observed for the late state conformation of proteorhodopsin for both protein forms. Structural mobility in these parts of proteorhodopsin is also consistent with observations made using solid-state NMR[154].

The structures generated by these linear movements, however, reflect tense structural states, where displacements of secondary structural elements have occurred without the corresponding side-chains of the amino acids having the ability to adjust to the changes in protein environment. In order account for these effects and thus more closely mimic the experimental conditions used, we extended the refinement protocol by introducing molecular dynamics simulations to sample the conformational space around both the resting state and the refined active state structure for each protein form. In the molecular dynamic system, each structural model was placed in a pre-equilibrated detergent micelle (consisting of 125 OG molecules) and solvated using water molecules (25000 tip3p) and sodium ions. Following energy minimization, each system was simulated for 3 ns at 300 K, using harmonic restraints on the backbone atoms. The use of harmonic restrains has the further advantage that it allows for a restrained sampling (RMSD <0.5Å) around the optimal secondary structure positions without deviating too far from the starting positions.

For each resting state and transient conformational state trajectory, 10-picosecond spaced snapshots of the protein were then extracted and the theoretical difference scattering for these structures was evaluated pair wise, using an error weighted χ^2-estimator,

$$\chi_i^2 = \sum_q \frac{[\Delta S^i(q) - \Delta S^{\mathrm{data}}(q)]^2}{\sigma_q^2 \, (N-M-1)} \quad , \qquad (30)$$

where $\Delta S^i(q)$ is the theoretical scattering for each structure pair i, $\Delta S^{data}(q)$ is the experimental basis spectrum from singular value decomposition, σ_q is the standard deviation (noise) in each q-point of the data as compared against the same basis spectrum after filtering, N is the number of data points included in the fitting range and M is the number of parameters used for the fitting. The use of a χ^2-estimator, as compared to the previously defined R-factor (PaperII), has the principal advantages that it (i) accounts for differences in S/N over the refined q-region and (ii) normalizes the sum of least squares with respect to the number of q-points included in the refinement.

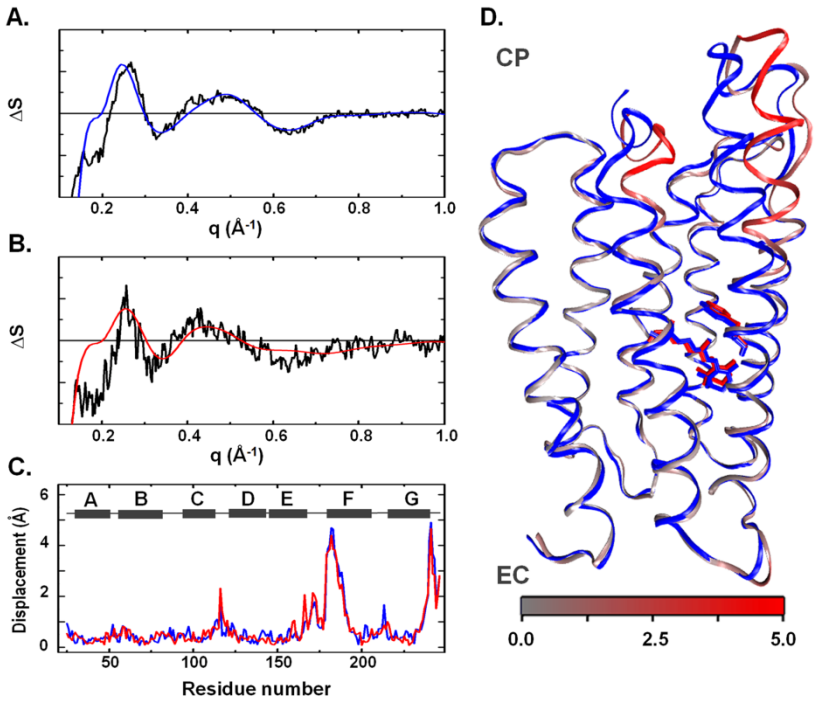


Figure 26. Overlay between the experimental basis spectrum and the optimal solution from an equilibrated molecular dynamics trajectory for (A) pR and (B) 13-I-pR. (C) Displacement of C_α-atoms between the resting state and the transient conformational state as a function of residue number for pR (red) and 13-I-pR (blue). The corresponding helix positions are indicated at the top. (D) Overlay between the resting state homology model (blue) and the transient conformation, colored grey to red according to the amplitude of the C_α displacement, for native pR. CP: cytoplasm, EC: extracellular

The predicted difference from the optimal structure pairs are overlaid against the experimental data in Figure 26. The corresponding displacement of the protein C_α-atoms between the resting state and the transient conformation in turn indicated that the secondary structural differences were almost indistinguishable for the two protein forms. This is consistent with the overall similarities already observed directly from the WAXS basis spectra. It was therefore concluded that the substitution of the C-20 methyl group with an iodine atom did not significantly affect the global conformational changes that occur during the proteorhodopsin photocycle.

3.2.4 FREE ENERGY CONSIDERATIONS

To explore the impact of the iodine substitution on the accelerated conformational transition, free energy differences (ΔG) were estimated using molecular dynamics simulations. This was done by alchemically transforming the C-20 methyl group into I-20, with the retinal in either *all*-trans or 13-*cis* state. For each of the retinal states this transformation was performed with the retinal either inside the protein ($\Delta G^{protein}$) or with the retinal in vacuum ($\Delta G^{protein}$). The difference between these terms,

$$\Delta\Delta G = \Delta G^{protein} - \Delta G^{vacuum} , \qquad (31)$$

quantifies the impact of the iodine atom on the surrounding protein matrix, where a positive/negative value of ΔΔG indicates a destabilization/stabilization of the protein conformation. The results from these calculations suggested that the iodine substitution significantly destabilized the 13-*cis* conformation of the protein by ~10-20 kJ/mol.

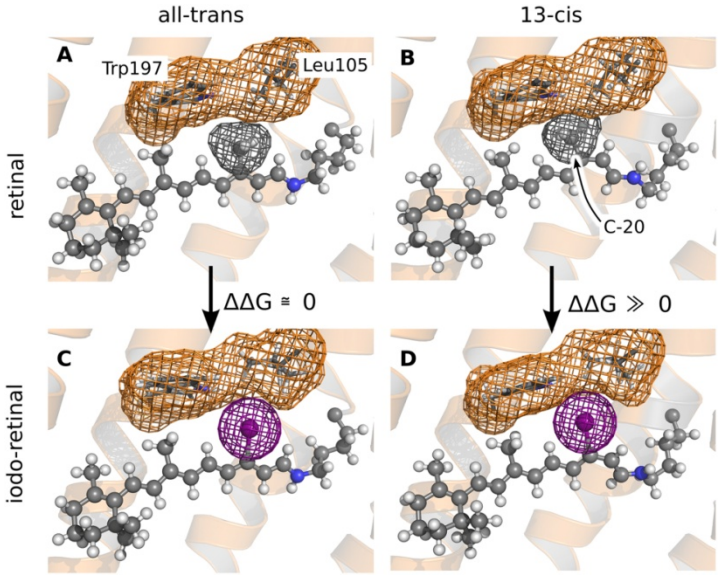


Figure 27. (A, B) Native retinal and (C, D) iodoretinal in either the all-*trans* (A, C) or 13-*cis* (B, D) conformation. The van der Waals surface of the C-20 methyl group, the iodine and of Leu105 and Trp197 are shown in a mesh representation. Transformation of the C-20 methyl group into iodine has little impact on the free energy differences in the all-*trans* conformation. In the 13-*cis* conformation this transformation leads to an unfavorable increase in free-energy due to steric interactions with neighboring Leu105 and Trp197.

This in turn implied that the conformational transitions which followed the isomerization event would also be considerably accelerated by this substitution. Interestingly, the stability of the *all*-trans conformation was not affected by the presence of the iodine atom.

This was understood when considering that the iodine atom is bulkier then the methyl group (Figure 27). In the all-*trans* state, the relaxed retinal prevents a steric clash between the iodine atom and the nearby side-chains of Trp197 and Leu105. In the 13-*cis* state, however, the same relaxation is obstructed by the kink in the retinal, causing an unfavorable steric interaction between the iodine atom and the same side-chains which in turn result in an increased free energy. From this perspective, it was also interesting to note that when the corresponding leucine residue in bacteriorhodopsin (Leu93) was mutated to a smaller alanine, the recovery of the resting state was considerably retarded[155]. In that study it was argued that the steric interaction with the leucine restrict the flexibility of the retinal, causing an acceleration of the retinal reisomerization back to the all-*trans* state. It can similarly be argued that the accelerated relaxation back to the resting state of 13-I-pR observed in this study is due to the even stronger van der Waals contacts of the iodine atom with Leu105 for the isomerized iodoretinal. In both respects, these findings confirm the importance of the C-20 methyl group in coupling the isomerization event to the global rearrangements in the protein.

3.2.5 SUMMARY

The conclusions drawn from this study point to the fact that, although the chemical, absorption and kinetic properties of proteorhodopsin are perturbed by the replacement of the C-20 methyl group with an iodine atom, the primary conformational changes undertaken by the protein are preserved. It thus appears that the underlying structural mechanism of the proteorhodopsin photocycle is tolerant to the increased volume of the iodinated retinal without any considerable structural consequences. This lead us to the conclusion that even though local interactions in the retinal binding pocket steer the onset and offset of the structural changes, it is subtly balanced interactions within the protein scaffold that determine the nature of proteorhodopsin's major transient conformational state.

3.3 TIME-RESOLVED WAXS BY RAPID READOUT DETECTION (PAPER IV)

All previous time-resolved WAXS studies of proteins[94,113] have relied on the pump-probe data collection strategy to study transient biological reactions. A major limitation of this approach is that only two highly specialized polychromatic beamlines in the world exists with the necessary requirements to collect time-resolved protein WAXS data; the ID09B beamline at the European Synchrotron Radiation Facility, previously discussed, and the ID14B beamline at the Advanced Photon Source.

In addition, far from all biological systems have reversible activation pathways. This means that in order to study irreversible systems the sample must continuously be exchanged after each activation event. Since each activation event in the pump-probe scheme equates to a single WAXS time-delay and several repeats have to be collected per time-delay to gain a sufficient signal-to-noise level, the amount of temporal data that can be collected rapidly becomes limited by the amount of sample available.

To address these problems we developed an alternative data acquisition strategy that uses a conventional monochromatic X-ray beam and records time-dependant WAXS data based on the millisecond readout capabilities of a newly developed pixel detector (PILATUS[156]). The strategy was evaluated using proteorhodopsin as a test system at the coherent small angle X-ray scattering (cSAXS) beamline at the Swiss light source.

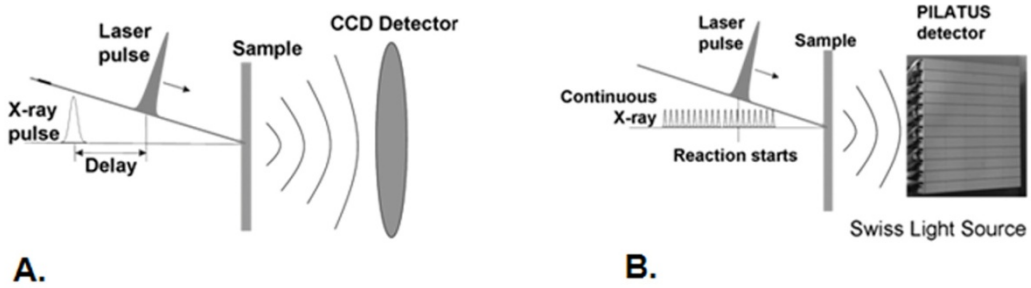


Figure 28. Schematic illustration highlighting the principal differences between the (A) pump-probe and the (B) rapid readout WAXS data collection strategies.

3.3.1 EXPERIMENTAL SETUP AND DATA ACQUISITION

The detergent solubilised proteorhodopsin samples were delivered into the X-ray beam by pumping the solution, through a 1 mm quartz capillary, using a motorized syringe pump. The sample was in turn photoactivated using a continuous wave 532 nm laser diode aligned perpendicular to the X-ray beam. The pump triggering was integrated into the beamline control system, which also controlled the timing of the arrival of the laser pulse and triggered a single module PILATUS 100k X-ray detector[156]. The data collection was then performed in a "toggle mode" where alternating "laser off" and "laser on" acquisition cycles were recorded consecutively.

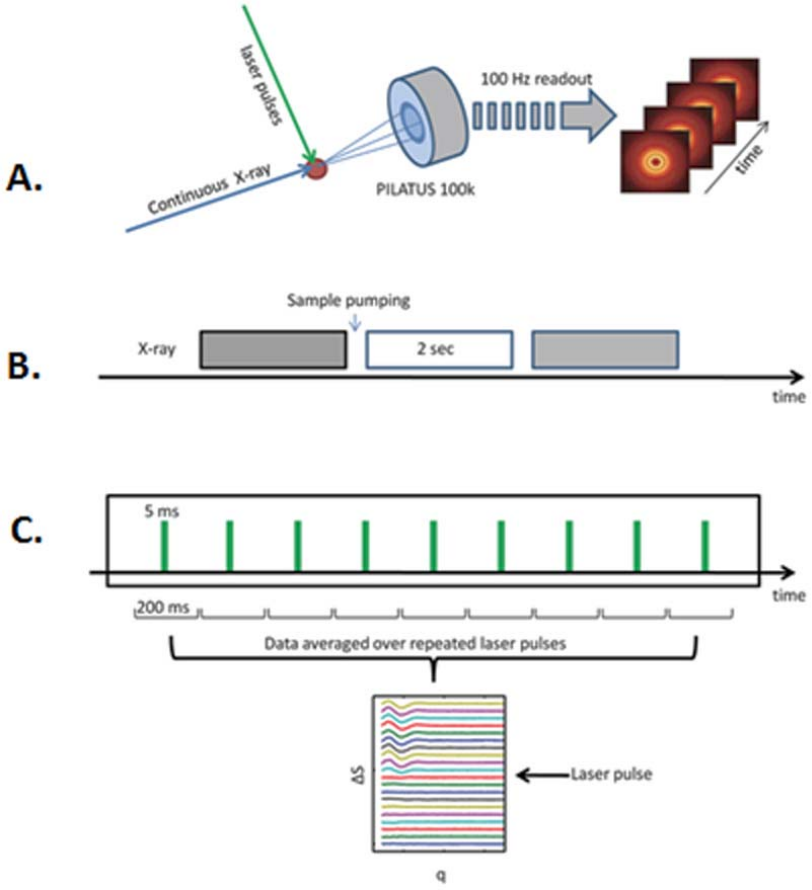


Figure 29. Overview of the data acquisition strategy. (A) The 532 nm laser was directed perpendicular to the X-ray beam and the principal axis of the sample capillary. (B) The data was collected using consecutive "laser off" (grey) and "laser on" (white) X-ray exposures. Between each 2 second X-ray exposure a fresh sample volume was pumped into position. (C) During the 2 s "laser on" exposure a sequence of nine 5 ms duration laser pulses photoactivated the sample.

In the "laser off" cycle a fresh sample volume (3 μl) was pumped into position, exposed to X-rays for 2 seconds while the triggered PILATUS detector continuously recorded the scattered X-rays with a 100 Hz frequency (10 ms per frame). The start of the PILATUS recording defines t=0, and the time-resolution is determined by the combined read out time (3 ms) and integration time (7 ms) of the detector. The "laser on" cycle" followed an identical scheme to the "laser off" cycle with the exception that sample was exposed to a sequence of 9 laser pulses (arriving at t= 200 ms, 400 ms, 600 ms, 800 ms, 1000 ms, 1200 ms, 1400 ms, 1600 ms and 1800 ms) during the X-ray exposure, where each laser pulse had a duration of 5 ms and was isolated from the laser diode using a mechanical shutter.

The thermal signal in this study was in turn measured by geometrically displacing the 532 nm laser with respect to the X-ray focus, allowing heat to be conducted into the region sampled by the X-rays. From the time-dependence of the difference WAXS data, a basis spectrum corresponding to the thermal signal was extracted.

3.3.2 CORRECTION FOR RADIATION DAMAGE AND HEATING

X-rays interacting with the sample deposits heat and induces a certain degree of radiation damage to the sample[157]. This leads to a small time-dependant change in the WAXS data, which evolves linearly with the time throughout the 2s X-ray exposure. In order to compensate for this systematic effect, the difference scattering intensity for each time point *t* was calculated by subtracting the average scattering intensity from the two flanking "laser off" cycles from the scattering intensity in the "laser on" cycle,

$$\Delta S(q,t) = S_{on}^{i}(q,t) - 0.5(S_{off}^{i-1}(q,t) + S_{off}^{i+1}(q,t)), \qquad (32)$$

where *i* represents the cycle number, *q* is the magnitude of the momentum change of the beam and *t* is the time-point recorded by the PILATUS. The thermal signal, due to laser-induced heating, was then removed from each difference scattering time point.

Furthermore, since each laser pulse represents a separate triggering event, the "laser on" cycle probes the same reversible reaction a total of nine times (100<t<300 ms ... 1700<t< 1900 ms). By averaging the difference WAXS data over all nine time-windows, the light-driven reaction of proteorhodopsin was effectively characterized over the temporal window of -100 < Δt < 100 ms, where Δt=0 is defined by the arrival of the 5 ms laser pulse.

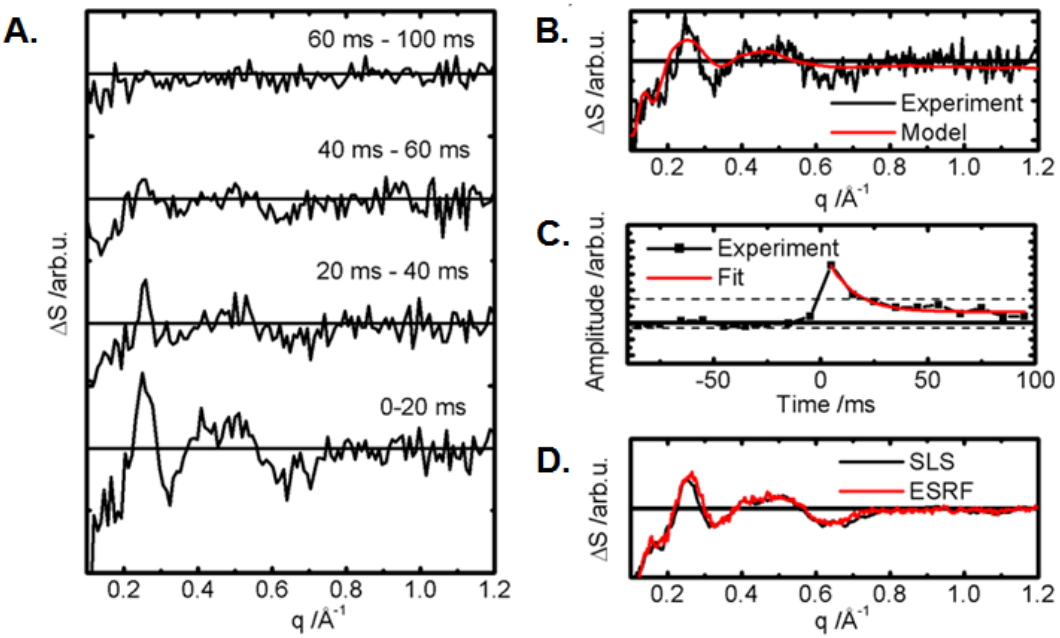


Figure 30. (A) Time-resolved difference WAXS data collected using the rapid readout data collections strategy and a monochromatic beam. (B) The principal basis spectrum retrieved from SVD analysis (black) together with the predicted difference scattering (red) from structural refinement using linear movements (see Paper II). (C) Time-evolution of the principal component (black squares) together with the result from mono-exponential modeling (red). (D) Overlay between the late state basis spectrum retrieved from polychromatic data collection at the ESRF (red) and the basis spectrum from SLS after convolution with the undulator spectrum used in the ESRF experiment (black).

3.3.3 DATA EVALUATION

The time-resolved WAXS data was subjected to singular value decomposition which yielded a single principle component. The time-dependence of the component could be described by a single exponential decay function with a time-constant of 34 ± 6 ms. This rate was slightly faster then what was previously observed (see Paper II and III) for native proteorhodopsin. This discrepancy is however likely associated with the relaxation from a photostationary state reached using a 5 ms pulse, as compared to the flash-induced relaxation previously observed. It was furthermore apparent when overlaying the difference WAXS basis spectrum from this experiment with that previously obtained in the same time regime by pump-probe data collection at the ESRF, that these two basis spectra were almost indistinguishable.

3.3.4 SUMMARY

These results established that the structural evolution in proteorhodopsin can be recorded on the millisecond time-scale by reading out the PILATUS X-ray detector with a frequency of 100 Hz. It further demonstrated that since the PILATUS detector does not add read out noise to the recorded data[156], collecting and averaging over many repeats, can compensate for the lower flux of the monochromatic X-ray beam. The method can therefore be employed at any SAXS/WAXS beamline with limited technical modifications and will thus be vital for increasing the accessibility of time-resolved WAXS. More importantly, the full timing history of the WAXS data is recorded per trigger event. This is an essential requirement when studying irreversible biological systems. Most protein reactions are also optimized to occur in flat potential energy landscapes and therefore take place on millisecond to second timescales. This means that a millisecond time-resolution is sufficient for the study of a large number of biological dynamics. Even so, current developments in pixel-detector technology[158] have already extended the temporal resolution available into the microsecond realm. Initial experiments conducted using these prototypical models at cSAXS have already showed promising sub-millisecond WAXS results.

3.4 CONFORMATIONAL ACTIVATION OF A GPCR (PAPER V)

G-protein coupled receptors (GPCRs) are the largest family of plasma membrane receptors in eukaryotes, and are responsible for mediating the cellular responses to a wide range of external stimuli, including neurotransmitters, odorants, hormones and light. As such, they also represent important targets for pharmaceutical treatments and an estimated 25-40% of all currently available drugs target these membrane proteins[159]. Given their importance, understanding the structural mechanism behind their action has been a long sought after goal.

Visual rhodopsin (or rhodopsin) belongs to the largest subgroup of these seven-transmembrane (TM) helix receptors (class A). Extensive biochemical and biophysical characterization of this light-sensitive GPCR have made rhodopsin the work horse for understanding GPCR activation[160]. Rhodopsin was the first GPCR for which a high-resolution structure of the resting state was reported[62]. Considerable effort has also been made to determine the structural changes associated with the activation of rhodopsin using indirect structural techniques (such as FTIR[161], EPR[162-163] and fluorescence spectroscopy[164-165]) as well as by X-ray crystallography[59,66-69,78,166-167].

From these studies significant structural rearrangements of the cytoplasmic portions of TM5 and TM6 have been reported for an opsin intermediate[67], a constitutively active intermediate[66] and a metarhodopsin II intermediate[78]. However, even though the constitutively active[66] and a metarhodopsin II[78] intermediates both represents activated forms of the protein, based on the observed binding of a C-terminal fragment of the transducin Gα subunit, these intermediates were activated by exogenous addition of all-*trans* retinal to opsin rather than by direct light-activation of the rhodopsin resting state. These studies were furthermore conducted at low temperatures ($\leq 10°C$) and in a detergent environment. These facts combined raise the concern to what extent the movements reported reflects the nature of the light-induced activation that occurs in the native membrane environment and under physiological conditions.

In this study time-resolved WAXS was applied to investigate the light-induced conformational changes that occur in bovine rhodopsin disk membranes at room temperature and in solution. Since the active state (meta II) of rhodopsin forms on the millisecond time-scale and persists for several minutes, the rapid readout data collection strategy developed in Paper IV was employed.

3.4.1 TIME-RESOLVED WAXS MEASUREMENTS

Time-resolved WAXS data were collected on concentrated solutions of bovine rhodopsin disk membranes, at room temperature, at the coherent SAXS beamline at the Swiss Light Source using the PILATUS rapid readout data collection scheme (Paper IV). The sample was photoactivated using a 532 nm laser pulse arriving 200 ms after the start of the PILATUS data recording, with the detector continuously streaming out images at a rate of 100 Hz (10 ms per frame). The data were collected in a "toggle mode" and a total of 300 time points were collected in the temporal domain

of -200 ms < Δt <2800 ms, where Δt=0 defines the arrival of the laser pulse. Following integration, normalization and averaging the laser-induced thermal signal (determined by heating the sample using an IR laser) was removed from each time-delay.

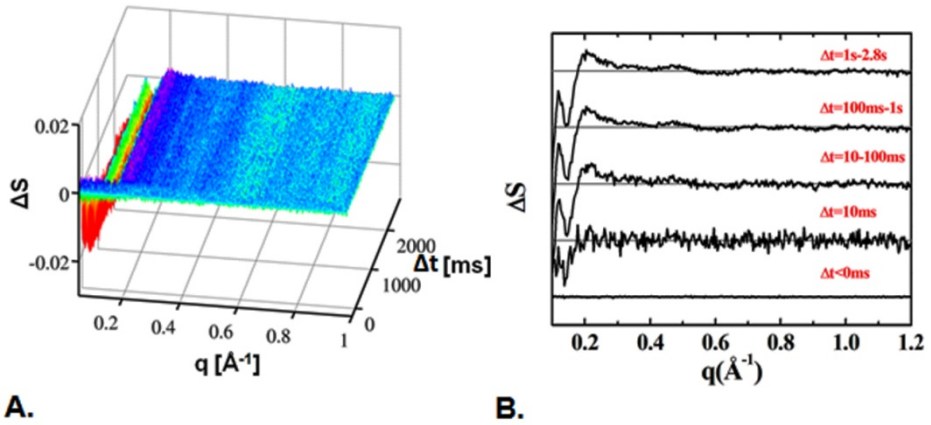


Figure 31. (A) Time-resolved difference WAXS data recorded from bovine rhodopsin disk membranes. (B) The same dataset, where the time-delays have been averaged over five representative time-windows.

From the full timing history of the difference WAXS data it was apparent that immediately following photoactivation (Δt > 0 ms) significant oscillations in the difference scattering pattern occur in the reciprocal space region < 0.70 Å^{-1}, and that this oscillating pattern formed and reached maximum amplitude already in the 10-100 ms domain. This was also confirmed by SVD analysis of all the difference scattering time-delays which yielded one principal component over the temporal window probed (Figure 32).

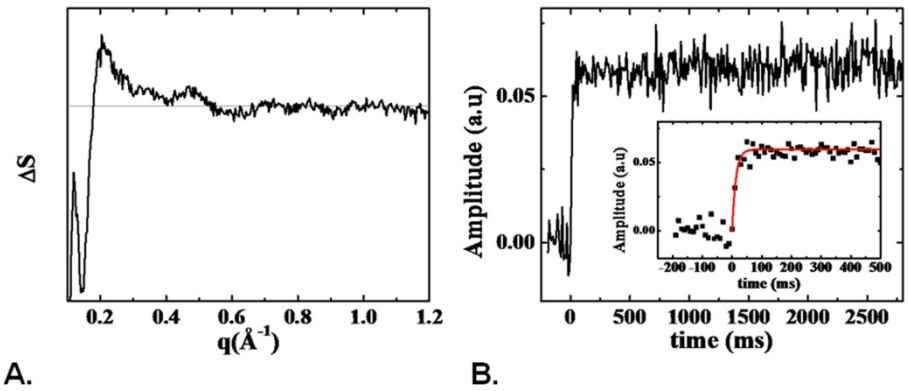


Figure 32. Principle component from SVD analysis. Showing (A) the basis spectrum and (B) the temporal dependence of the principle component. The inset illustrates the single exponential fit (red line) to the rise of the same component.

The time-dependence of the amplitudes of this component indicated an almost instantaneous formation of the structural species upon light-activation. It further showed that the population of this species persist at a constant level throughout the 2.8 s measured. Moreover, a single-exponential approximation of the rise of this structural component yielded a rise-time of 13±5 ms. This time-scale, although bordering the

limit of the achievable time-resolution, is in good agreement with the reported rise of the Meta II photointermediate in disk membranes at room temperature from both absorption[168] and fluorescence[169] spectroscopy.

3.4.2 THEORETICAL PREDICTIONS FROM DEPOSITED INTERMEDIATES

Analogous to the procedure developed in Paper II, the WAXS basis spectrum was first evaluated against the scattering differences predicted from deposited crystal structures. In order to reduce the number of resting and active state structures a structural comparison, based on the estimated coordinate uncertainties[170], was first performed. Comparison of all deposited resting state structures[56-57,59,61-63,171-173] yielded two subpopulations of rhodopsin resting states showing significant structural differences in transmembrane regions. Such secondary structural differences could have a significant impact on the calculated difference scattering. For this reason, one representative high-resolution resting state structure was selected from each population (PDB ID: 1GZM[171] and 1U19[173]) and all subsequent analysis was performed against each of those two resting states. A similar analysis was performed by comparing all active[59,66-69,78,166-167] state structures against the two representative resting states. Of these structures, the opsin (PDB ID: 3CAP[67]), the constitutively active (PDB ID: 2X72[66]) and the metarhodopsin II (PDB ID: 3PXO[78]) structures displayed significant structural displacements in helical regions, compared to both of the two resting states.

Moreover, there is ample evidence, based on atomic force microscopy (AFM) [174-175] and cross-linking data[176], which indicate that rhodopsin forms rows of dimers in the native disk membranes. To investigate the potential effect of a multimeric rhodopsin arrangement, scattering differences was therefore predicted for both monomeric and dimeric models of rhodopsin (Figure 33).

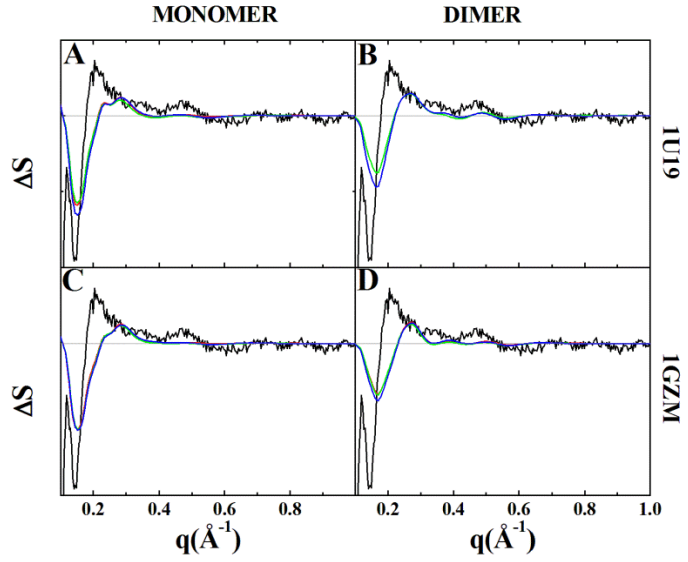


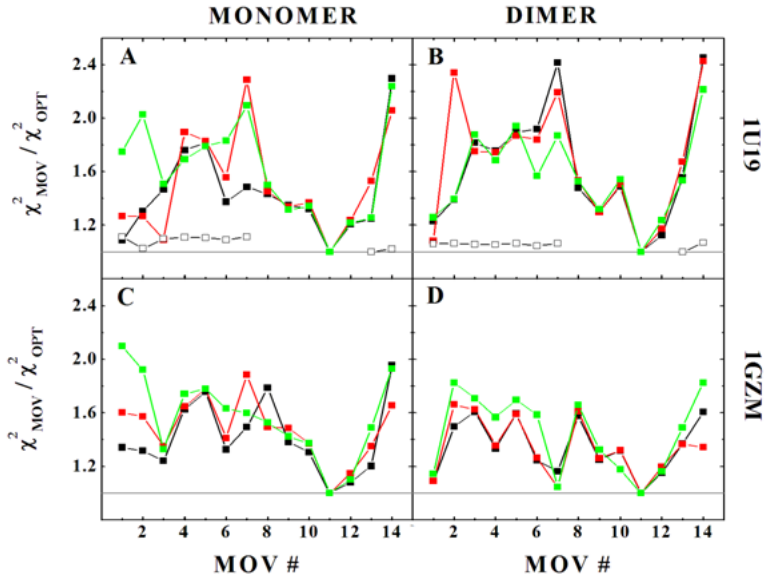
Figure 37. Predicted scattering from deposited crystallographic structures. The theoretical scattering was calculated between each of the active state structure (3CAP red, 3PXO green, 2X72 blue) and each resting state (1U19 or 1GZM) . The prediction was performed using either a monomer (A and C) or a dimer (B and D) description. The experimental basis spectrum is shown in black.

3.4.3 STRUCTURAL REFINEMENT OF THE DIFFERENCE WAXS SIGNAL

Since none of these structure pairs were able to fully capture the experimental difference signal, an iterative structural refinement was explored (see Paper II). From each of the three active state crystal structures fourteen candidate moving regions were identified. The amplitudes of each of these movements were then refined by scoring the theoretical difference WAXS spectrum against the experimental TR-WAXS basis spectrum using a reduced χ^2-estimator (see Paper III). This procedure was repeated for each candidate movement, for each pair of active state and resting state structures, and in either a monomeric or dimeric form. The best solution for each movement (χ^2_{MOV}) was then evaluated against the optimal solution (χ^2_{OPT}) from all candidate movements for each structure pair according to a two sample F-test[177]. From the ratio,

$$F = \frac{\chi^2_{MOV}}{\chi^2_{OPT}}, \qquad (33)$$

and the number of degrees of freedom in the experiment (N-M-1, where N is the number of radial observations and M is the number of refined parameters) a p-value was recovered from the corresponding F-distribution. This p-value provided an estimate of the probability that either movement could equally well describe the data given the statistical variation in the measurements.



E.

Nr	MOV	Residues
1	TM1 EC	35-52
2	TM1 CP	54-70
3	TM2 EC	72-89
4	TM2 CP	91-100
5	TM3 EC	108-120
6	TM3 CP	122-141
7	TM4	150-170
8	TM5 CP	216-238
9	TM 6 CP	240-264
10	TM6	240-284
11	TM56 CP	216-264
12	TM56	216-284
13	TM7	286-302
14	Helix 8	311-326

Figure 38. Structural refinement overview. The F-value (=χ^2_{MOV}/ χ^2_{OPT}) for the pair wise comparison between each movement (χ^2_{MOV}) against the optimal movement (χ^2_{OPT}), for each combination of active state (3CAP black, 3PXO red and 2X72 green) and resting state (1U19 or 1GZM) structures, using either the monomer (A and C) or dimer (B and D) model description. An F-value of 1 (grey line) would indicate that the two solutions produce an equal good agreement with the experimental data. White squares illustrates the two-component refinement, where TM5 and TM6 (i.e. MOV 11) was combined with each of the remaining helical movements. (E) Definition of the moving regions (MOV) used in the refinement. TM: Transmembrane helix, EC: Extracellular part, C: Cytoplasmic part.

A systematic comparison of the refinement results from each structure pair and oligomeric form (Figure 38) indicated that a concerted cytoplasmic movement of TM5 (inward) and TM6 (outward) provided the best agreement with experimental data for all combinations of active and resting state structures (p-values of 0.04 to 0.23). Interestingly, the same concerted movement also provided the best solution in the context of a dimeric model for all structural pairs (p-values from 0.03 to 0.34). Of the models explored, the motion extracted from the opsin[67] structure in a dimeric form provided the best solution to the WAXS signal, although the other two active[66,78] state structures yielded statistically comparable solutions as dimers (Figure 39).

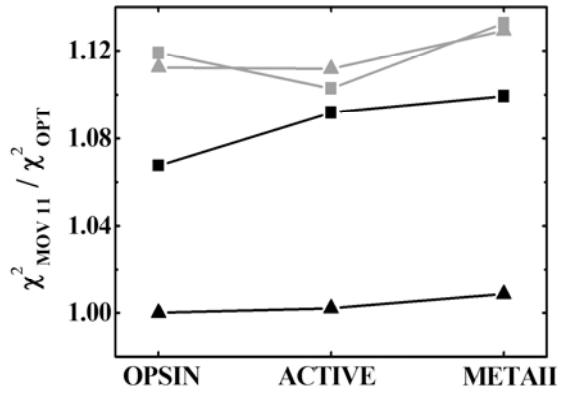


Figure 39. Cross-comparison of the optimal one component movement (i.e. MOV 11) for each combination of structure pairs and oligomeric states. Black: 1U19 resting state structure, Grey: 1GZM resting state structure, Squares: Monomer system and Triangles: Dimer system.

Addition of a second component to the concerted movement of TM5 and TM6 was also assayed. However, the result from this two-component refinement indicated that it was not possible to statistically discriminate between these additional motions.

Following refinement the amplitudes of the motions observed in each monomer of the optimal active state dimer indicated that both TM5 (4.0 ± 0.2) and TM6 (10.7 ± 0.4Å) were considerably displaced as compared to the resting state. Interestingly, a similar degree of motion was also observed for the optimal monomer refinement (4.4 ± 0.4Å for TM5 and 11.8 ± 1.0Å for TM6). The agreement to the WAXS difference signal (Figure 40) also indicated that both models were able to capture the major oscillations in the difference scattering pattern. It was, nevertheless, noticeable that the monomeric case appeared more successful in capturing the data at low angles ($q < 0.2Å^{-1}$) whereas the dimeric form more closely followed the experimental differences at wider angles ($0.25 < q < 0.5\ Å^{-1}$).

3.4.4 SUMMARY

It is striking that the movements observed in native rhodopsin membranes at room temperature are ~50% larger than those observed in the corresponding crystallographic structures[66-67,78]

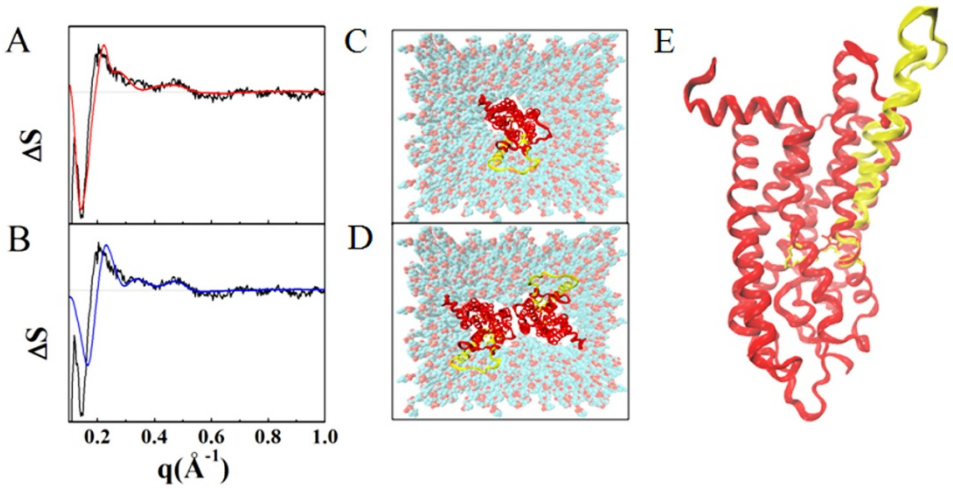


Figure 40. Structural refinement of the difference WAXS signal. Optimal difference scattering solutions retrieved for a (A) monomeric (solid red line) and (B) dimeric (solid blue line) model systems, where helical motions were isolated from the opsin structure (3CAP) and compared against the rhodopsin resting state (1U19). Cytoplasmic view of the corresponding resting (red) and active (yellow) state models in a POPC lipid bilayer for a (A) monomer and (B) dimer system. (E) Side view of the active state monomer model from WAXS refinement (yellow) compared with resting state structure (red).

In this context it is important to note that these crystallographic structures were kinetically restricted by temperature and crystal lattice constraints, and they also represented either chromophore deficient[66-67] or chromophore induced active state structure[78]. It is also noteworthy that the best approximation of the experimental data was observed for the rhodopsin dimer model recovered from atomic force microscopy studies of native membranes[174]. Thus the time-resolved WAXS data in this study of native rhodopsin membranes provide additional support for this proposed oligomeric arrangement.

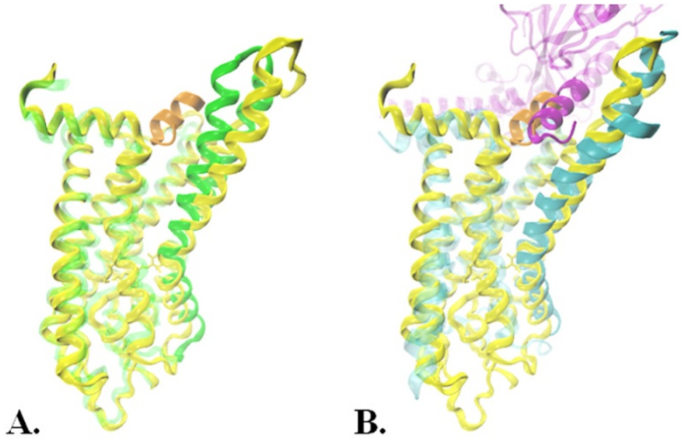


Figure 41. Structural comparison. (A) WAXS model (yellow) and the opsin structure (green) with the bound G-peptide (brown). (B) WAXS model (yellow) and β2 adrenergic receptor (blue) with the full length G-protein (purple). For comparison the transducin G-peptide is shown in brown.

In light of these findings, it is also worth considering that a recent crystal structure of the β2 adrenergic receptor in complex with its G-protein[178] also showed movements 50 % larger than those observed for the active rhodopsin structures. In this study it was suggested that the smaller movements observed for the activated rhodopsin structures was likely the consequence of a partial interaction between the receptor and the G-protein fragment, rather than a significant conformational difference in the G-protein binding pocket between these two GPCRs. The time-resolved WAXS data in this work support this conclusion and imply that photoactivated rhodopsin in its native membrane rapidly achieves the full conformational change required to effectively bind its G-protein transducin. The alternative interpretation, that light-activation of rhodopsin merely achieves two thirds of the helical movements required for rhodopsin activation, and that G-protein binding inputs additional energy to further change the structure of the GPCR, appears less optimized for efficient signal transduction.

4 CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The field of membrane protein biology has rapidly grown over the last decades, in large part driven by the increasing awareness that membrane proteins constitute a majority of all medical drug targets. An enhanced focus on membrane proteins is also reflected by the exponential growth in the number of determined membrane protein structures. As the amount of structural information grows there is an increasing demand to understand how structural changes in these proteins are coupled to its physiological function.

This work has presented time-resolved wide angle X-ray scattering as a method to study membrane protein dynamics in real-time and in solution. The method was applied to study rhodopsins, a superfamily of light-sensitive membrane proteins of diverse origin and function.

The first membrane proteins to be studied using this technique was the light-driven proton pumps bacteriorhodopsin and proteorhodopsin. The experimental data collected in these studies and the subsequent analysis and structural interpretation demonstrated the ability of time-resolved WAXS to characterize the time-evolution of secondary structural changes in complex biological systems. This provided new insights in to the timing, sequence and magnitude of the conformational changes that occur in these two proteins. The method was further extended by introducing a chemically modified retinal chromophore into proteorhodopsin. This time-resolved WAXS study, between a native and modified proteorhodopsin, demonstrated the applicability of the technique for doing comparative structural studies. The results indicated that although local steric interactions in the ligand binding pocket can drastically perturb the rate of conformational changes, it is rather subtly balanced interactions within the protein scaffold that determine the nature of these changes.

In the light of making the technique of time-resolved WAXS accessible for a wider user community, a technically less demanding data collection strategy was also developed. This approach employs the use of fast readout detector to capture conformational events occurring on the millisecond time-scale. Finally, it was demonstrated that time-resolved WAXS using rapid readout detection can be applied to study the irreversible activation of the GPCR rhodopsin. Systematic structural analysis of this data provided new insights into conformational activation of this receptor.

Taken together, the conformational changes observed in these time-resolved WAXS studies suggest that not only do these functionally and genetically diverse proteins share conformational motifs, all displaying significant cytoplasmic movements of transmembrane helices 5 and 6 upon activation, but the magnitudes of these movements are considerable larger than what has been previously suggested. This is most likely a consequence of measuring these conformational events in a more native

like solution-environment where the conformational flexibility of the protein is less restrained by molecular crowding effects such as crystal packing.

Although several technical and analytical developments have been presented here, time-resolved membrane protein WAXS is in many aspects a budding field undergoing exciting developments and challenges. Further developments and implementations of computational tools for data reduction and structural analysis are foreseeable. This is particular true for proteins where there is little *a priori* structural information regarding their conformational flexibility. For these proteins the use of coarse grain modeling[179] or conformational sampling based on geometrical constraints[180] would present an option for structural refinement. A further development in describing the influence of the micelle environment is another interesting aspect which could help guide the structural interpretation of WAXS data. This aspect could, for instance, be explored by performing contrast variation studies[181] or by using uniform lipid environments such as nanodiscs[182].

More importantly, time-resolved WAXS is generic methodology and can be used to explore other important biological processes such as protein complex formation (for instance rhodopsin-transducin interaction), protein folding or co-transport of solutes in membrane protein transporters. The use and development of alternative means of reaction triggering such as pH jumps, caged compounds and stopped-flow substrate mixing also hold great potential for expanding the use of time-resolved WAXS beyond light-sensitive proteins. Such preliminary tests have already indicated promising new directions for the characterization of membrane protein dynamics.

ACKNOWLEDGEMENTS

Time flies when you're busy! It feels like it was only yesterday I started my journey to become a PhD at the Lundberg Lab, and suddenly here I am writing acknowledgments in my thesis. During my time in the Neutze group I've been fortunate to meet a great crowd of people and here is my special- thank you- to you.

First of all I want to thank you **Richard** for inviting me to join your group. It has truly been a privilege and experience to work with you. I really appreciate that you given me the freedom to explore different scientific ideas as well as providing me with the opportunity to travel the world in the pursuit of science. Throughout this period you have been a constant source of inspiration, support and optimism. I'm positive you're going to get "the prize".

Gergos, as my co-supervisor you've been brilliant. Your broad range of knowledge and understanding never seizes to amaze me. Thank you for always having a door open for discussion and for patiently giving me advice whenever it was needed.

The seniors in the group **Kristina, Susanna, Sebastian, Rosie and Urszula.** I really appreciated your experience during my "prepping period" as well as our little luncheons to ~~Lunchen~~ Lyktan. Great thanks to Sebastian for successful scientific collaborations.

The "old crowd" **Magnus, Annemarie, Pontus, Maria and Rob.** I can't remember the countless beer clubs I've been trying to catch up with you guys. Special thanks to Magnus for all long nights in front of the computer hacking PERL trying to get those curves to make sense.

Anna, Fredrik and Gerhard my peers. We started our PhD together and had great times hitting those (ski) slopes from all directions. It's now up to you Anna to preserve the legacy of our generation ☺

"The time-resolved" crowd **Linda and David.** Thank you for making those long synchrotron nights enjoyable. I guess you don't really know a person until you have shared a major "heat load chopper malfunction" / "DM50 laser breakdown" together. Keep up the good work. In doubt, you can always call the head of the beamline☺

"Roomies" **Pia**, **Elena** and **Ida**. Thanks for sharing 4 m^2 of office space with me without going bananas. I have enjoyed your company.

Xuewen, thanks for a great collaboration and for being such a sport during those long synchrotron trips.

Jennie, Mike, Mikael, Rhawnie, Petra and Cecilia, I'm impressed with your organizing skills and social activities, keep up the good work. Rhawnie, thanks for being so supportive during my final hours. Cecilia, welcome to the lab!

Weixiao and **Anette,** keep those xtals diffracting. Thank you Anette for co-hosting G's defense as a champ.

"The CMBs" **Karin L, Karin R** and **Madde**. It is sort of ironic that you guys move away when the Neutze group finally goes CMB. I've appreciated the "outside perspective" and the long beer clubs (Karin R). Don't forget to start a BC in Lund!

Alex and **Philipp,** welcome to the (laser) lab? It's actually quite decent once you've gotten used to it☺

Etienne and **Monika,** it's great to have the international atmosphere back in the lab again.

"The Uppsala crowd"**, Jochen, David and Jan.** Thanks for great collaborations and for teaching me MD and aligning lasers.

"The LA crowd" **Jeff and Vincent**. Thanks guys for showing me how it's done Cali-style ☺ . I really enjoyed our collaboration and hope we'll work together in the future.

Wim and Petra in Nijmegen. Thank you for inviting me to your lab and showing me how to dissect those cow eyes in pitch darkness.

Elias, Bartek, Petter, Lotta och Mattias, tack för att ni har orkat lyssna på mitt labb-tjat och för att ni har dragit med mig ut på äventyr utanför labbets väggar. Jag lovar att jag ska bli mer social när avhandlingen är ur världen.

Tack! **Christina, Bo, David och Elisabeth,** ni har varit ett stort praktiskt, men framför allt moraliskt, stöd under den här perioden.

Min familj, Tack för att ni har stött och uppmuntrat mig under de här åren. Jag ser fram emot en period när jag kan åka och hälsa på er alla snart igen.

Johanna, utan din förbehållslösa uppmuntran och ditt kärleksfulla stöd hade detta arbetet aldrig varit möjligt. Jag ser fram emot att kunna återgälda det inom en snar framtid. Puss!

REFERENCES

1 Luckey, M. (Cambridge University Press, 2008).
2 Jain, M. K. & Wagner, R. C. (Wiley, New York, 1988).
3 Voet, D. & Voet, J. G. *Biochemistry*. 3 edn, (John Wiley & Sons, Inc., 2004).
4 Lundstrom, K. Structural genomics: the ultimate approach for rational drug design. *Molecular Biotechnology* **34**, 205-212 (2006).
5 Spudich, J. L., Yang, C., Jung, K. & Spudich, E. N. RETINYLIDENE PROTEINS: Structures and Functions from Archaea to Humans. *Annual Review of Cell and Developmental Biology* **16**, 365-392 (2000).
6 Kakitani, T., Beppu, Y. & Yamada, A. Color tuning mechanism of human red and green visual pigments. *Photochem Photobiol* **70**, 686-693 (1999).
7 Kochendoerfer, G. G., Lin, S. W., Sakmar, T. P. & Mathies, R. A. How color visual pigments are tuned. *Trends in Biochemical Sciences* **24**, 300-305 (1999).
8 Oesterhelt, D. & Stoeckenius, W. Functions of a new photoreceptor membrane. *Proc Natl Acad Sci U S A* **70**, 2853-2857 (1973).
9 Oesterhelt, D. & Tittor, J. Two pumps, one principle: light-driven ion transport in halobacteria. *Trends Biochem Sci* **14**, 57-61 (1989).
10 Stock, D., Leslie, A. G. & Walker, J. E. Molecular architecture of the rotary motor in ATP synthase. *Science* **286**, 1700-1705, (1999).
11 Hirai, T. & Subramaniam, S. Protein Conformational Changes in the Bacteriorhodopsin Photocycle: Comparison of Findings from Electron and X-Ray Crystallographic Analyses. *Plos One* **4**, (2009).
12 Sheves, M., Albeck, A., Friedman, N. & Ottolenghi, M. Controlling the pKa of the bacteriorhodopsin Schiff base by use of artificial retinal analogues. *Proc Natl Acad Sci U S A* **83**, 3262-3266 (1986).
13 Chang, C. H., Jonas, R., Govindjee, R. & Ebrey, T. G. Regeneration of Blue and Purple Membranes From Deionized Bleached Membranes of Halobacterium halobium. *Photochem Photobiol* **47**, 261-265 (1988).
14 Gat, Y. & Sheves, M. A mechanism for controlling the pKa of the retinal protonated Schiff base in retinal proteins. A study with model compounds. *Journal of American Chemical Society* **115**, 3772-3773 (1993).
15 Brown, L. S. *et al.* Glutamic acid 204 is the terminal proton release group at the extracellular surface of bacteriorhodopsin. *J Biol Chem* **270**, 27122-27126 (1995).
16 Balashov, S. P. *et al.* Glutamate-194 to cysteine mutation inhibits fast light-induced proton release in bacteriorhodopsin. *Biochemistry-Us* **36**, 8671-8676, (1997).
17 Braiman, M. & Mathies, R. Resonance Raman spectra of bacteriorhodopsin's primary photoproduct: evidence for a distorted 13-cis retinal chromophore. *Proc Natl Acad Sci U S A* **79**, 403-407 (1982).
18 Doig, S. J., Reid, P. J. & Mathies, R. A. Picosecond time-resolved resonance Raman spectroscopy of bacteriorhodopsin's J, K and KL intermediates. *Journal of Physical Chemistry* **95**, 6372-6379 (1991).
19 Edman, K. *et al.* Early structural rearrangements in the photocycle of an integral membrane sensory receptor. *Structure* **10**, 473-482, (2002).
20 Marti, T., Rosselet, S. J., Otto, H., Heyn, M. P. & Khorana, H. G. The retinylidene Schiff base counterion in bacteriorhodopsin. *J Biol Chem* **266**, 18674-18683 (1991).
21 Cao, Y. *et al.* Relationship of Proton Release at the Extracellular Surface to Deprotonation of the Schiff-Base in the Bacteriorhodopsin Photocycle. *Biophys J* **68**, 1518-1530 (1995).
22 Hessling, B., Herbst, J., Rammelsberg, R. & Gerwert, K. Fourier transform infrared double-flash experiments resolve bacteriorhodopsin's M-1 to M-2 transition. *Biophys J* **73**, 2071-2080 (1997).
23 Subramaniam, S. *et al.* Protein conformational changes in the bacteriorhodopsin photocycle. *Journal of Molecular Biology* **287**, 145-161 (1999).
24 Neutze, R. *et al.* Bacteriorhodopsin: a high-resolution structural view of vectorial proton transport. *Biochim Biophys Acta* **1565**, 144-167, (2002).
25 Vonck, J. Structure of the bacteriorhodopsin mutant F219L N intermediate revealed by electron crystallography. *Embo Journal* **19**, 2152-2160 (2000).
26 Beja, O. *et al.* Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**, 1902-1906, (2000).
27 Dioumaev, A. K. *et al.* Proton transfers in the photochemical reaction cycle of proteorhodopsin. *Biochemistry-Us* **41**, 5348-5358, (2002).
28 Martinez, A., Bradley, A. S., Waldbauer, J. R., Summons, R. E. & DeLong, E. F. Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host. *Proc Natl Acad Sci U S A* **104**, 5590-5595, (2007).
29 Fuhrman, J. A., Schwalbach, M. S. & Stingl, U. Proteorhodopsins: an array of physiological roles? *Nature Reviews* **6**, 488-494 (2008).

30 DeLong, E. F. *et al.* Proteorhodopsin genes are distributed among divergent marine bacterial taxa. *P Natl Acad Sci USA* **100**, 12830-12835 (2003).

31 DeLong, E. F. & McCarren, J. Proteorhodopsin photosystem gene clusters exhibit co-evolutionary trends and shared ancestry among diverse marine microbial phyla. *Environmental Microbiology* **9**, 846-858 (2007).

32 DeLong, E. F., Frigaard, N. U., Martinez, A. & Mincer, T. J. Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature* **439**, 847-850 (2006).

33 Slamovits, C. H., Okamoto, N., Burri, L., James, E. R. & Keeling, P. J. A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nat Commun* **2**, 183, (2011).

34 Beja, O., Spudich, E. N., Spudich, J. L., Leclerc, M. & DeLong, E. F. Proteorhodopsin phototrophy in the ocean. *Nature* **411**, 786-789, (2001).

35 Man, D. *et al.* Diversification and spectral tuning in marine proteorhodopsins. *EMBO J* **22**, 1725-1731,(2003).

36 Wang, W. W., Sineshchekov, O. A., Spudich, E. N. & Spudich, J. L. Spectroscopic and photochemical characterization of a deep ocean proteorhodopsin. *J Biol Chem* **278**, 33985-33991,(2003).

37 Spudich, J. L. The multitalented microbial sensory rhodopsins. *Trends Microbiol* **14**, 480-487, (2006).

38 Brown, L. S. & Jung, K. H. Bacteriorhodopsin-like proteins of eubacteria and fungi: the extent of conservation of the haloarchaeal proton-pumping mechanism. *Photochem Photobiol Sci* **5**, 538-546, (2006).

39 Krebs, R. A., Alexiev, U., Partha, R., DeVita, A. M. & Braiman, M. S. Detection of fast light-activated H+ release and M intermediate formation from proteorhodopsin. *BMC Physiol* **2**, 5 (2002).

40 Friedrich, T. *et al.* Proteorhodopsin is a light-driven proton pump with variable vectoriality. *J Mol Biol* **321**, 821-838, (2002).

41 Dioumaev, A. K., Wang, J. M., Balint, Z., Varo, G. & Lanyi, J. K. Proton transport by proteorhodopsin requires that the retinal Schiff base counterion Asp-97 be anionic. *Biochemistry-Us* **42**, 6582-6587, (2003).

42 Xiao, Y., Partha, R., Krebs, R. & Braiman, M. Time-resolved FTIR spectroscopy of the photointermediates involved in fast transient H+ release by proteorhodopsin. *J Phys Chem B* **109**, 634-641, (2005).

43 Bergo, V. B. *et al.* His-75 in proteorhodopsin, a novel component in light-driven proton translocation by primary pumps. *J Biol Chem* **284**, 2836-2843, (2009).

44 Yoshitsugu, M., Yamada, J. & Kandori, H. Color-changing mutation in the E-F loop of proteorhodopsin. *Biochemistry-Us* **48**, 4324-4330, (2009).

45 Hempelmann, F. *et al.* His75-Asp97 cluster in green proteorhodopsin. *J Am Chem Soc* **133**, 4645-4654,(2011).

46 Millero, F. J. Thermodynamics of the carbonate system in seawater *Geochimica et Cosmochimica Acta* **43**, 1651-1661 (1979).

47 Rangarajan, R., Galan, J. F., Whited, G. & Birge, R. R. Mechanism of spectral tuning in green-absorbing proteorhodopsin. *Biochemistry-Us* **46**, 12679-12686, (2007).

48 Kühne, W. Beiträge zur Optochemie. *Untersuchungen aus dem physiologischen Institute der Universität Heidelberg* **4**, 169-249 (1882).

49 Birge, R. R. *et al.* Two-photon spectroscopy of locked-11-cis-rhodopsin: evidence for a protonated Schiff base in a neutral protein binding site. *Proc Natl Acad Sci U S A* **82**, 4117-4121 (1985).

50 Fahmy, K., Siebert, F. & Sakmar, T. P. Photoactivated State of Rhodopsin and How It Can Form. *Biophys Chem* **56**, 171-181 (1995).

51 Han, M. & Smith, S. O. Nmr Constraints on the Location of the Retinal Chromophore in Rhodopsin and Barthorhodopsin. *Biochemistry-Us* **34**, 1425-1432 (1995).

52 Kliger, D. S. & Lewis, J. W. Spectral and kinetic characterization of visual pigment photointermediates. *Israel J Chem* **35**, 289-307 (1995).

53 Palings, I. *et al.* Assignment of Fingerprint Vibrations in the Resonance Raman-Spectra of Rhodopsin, Isorhodopsin, and Bathorhodopsin - Implications for Chromophore Structure and Environment. *Biochemistry-Us* **26**, 2544-2556 (1987).

54 Popp, A., Ujj, L. & Atkinson, G. H. Bathorhodopsin structure in the room-temperature rhodopsin photosequence: picosecond time-resolved coherent anti-Stokes Raman scattering. *Proc Natl Acad Sci U S A* **93**, 372-376 (1996).

55 Sakmar, T. P. & Fahmy, K. Properties and photoactivity of rhodopsin mutants. *Israel J Chem* **35**, 325-337 (1995).

56 Teller, D. C., Okada, T., Behnke, C. A., Palczewski, K. & Stenkamp, R. E. Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry-Us* **40**, 7761-7772 (2001).

57 Stenkamp, R. E. Alternative models for two crystal structures of bovine rhodopsin. *Acta Crystallogr D Biol Crystallogr* **D64**, 902-904, (2008).

58 Standfuss, J. *et al.* Crystal structure of a thermally stable rhodopsin mutant. *J Mol Biol* **372**, 1179-1188, doi:S0022-2836(07) (2007).

59 Salom, D. *et al.* Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc Natl Acad Sci U S A* **103**, 16123-16128, (2006).
60 Okada, T. *et al.* The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. *J Mol Biol* **342**, 571-583, (2004).
61 Okada, T. *et al.* Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc Natl Acad Sci U S A* **99**, 5982-5987, (2002).
62 Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739-745,(2000).
63 Makino, C. L., Riley, C. K., Looney, J., Crouch, R. K. & Okada, T. Binding of more than one retinoid to visual opsins. *Biophys J* **99**, 2366-2373, (2010).
64 Li, J., Edwards, P. C., Burghammer, M., Villa, C. & Schertler, G. F. Structure of bovine rhodopsin in a trigonal crystal form. *J Mol Biol* **343**, 1409-1438, (2004).
65 Ruprecht, J. J., Mielke, T., Vogel, R., Villa, C. & Schertler, G. F. Electron crystallography reveals the structure of metarhodopsin I. *EMBO J* **23**, 3609-3620, (2004).
66 Standfuss, J. *et al.* The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* **471**, 656-660, (2011).
67 Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H. W. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454**, 183-187, (2008).
68 Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497-502, (2008).
69 Nakamichi, H., Buss, V. & Okada, T. Photoisomerization mechanism of rhodopsin and 9-cis-rhodopsin revealed by x-ray crystallography. *Biophys J* **92**, L106-108, (2007).
70 Nakamichi, H. & Okada, T. X-ray crystallographic analysis of 9-cis-rhodopsin, a model analogue visual pigment. *Photochem Photobiol* **83**, 232-235, (2007).
71 Yeagle, P. L., Choi, G. & Albert, A. D. Studies on the structure of the G-protein-coupled receptor rhodopsin including the putative G-protein binding site in unactivated and activated forms. *Biochemistry-Us* **40**, 11932-11937, (2001).
72 Choi, G. *et al.* Structural studies of metarhodopsin II, the activated form of the G-protein coupled receptor, rhodopsin. *Biochemistry-Us* **41**, 7318-7324, (2002).
73 Ridge, K. D. & Palczewski, K. Visual rhodopsin sees the light: structure and mechanism of G protein signaling. *J Biol Chem* **282**, 9297-9301, (2007).
74 Ernst, O. P. & Bartl, F. J. Active states of rhodopsin. *Chembiochem* **3**, 968-974, (2002).
75 Dunham, T. D. & Farrens, D. L. Conformational changes in rhodopsin. Movement of helix f detected by site-specific chemical labeling and fluorescence spectroscopy. *J Biol Chem* **274**, 1683-1690 (1999).
76 Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc Natl Acad Sci U S A* **105**, 7439-7444, (2008).
77 Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L. & Khorana, H. G. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **274**, 768-770 (1996).
78 Choe, H. W. *et al.* Crystal structure of metarhodopsin II. *Nature* **471**, 651-655, (2011).
79 Arnis, S. & Hofmann, K. P. Two different forms of metarhodopsin II: Schiff base deprotonation precedes proton uptake and signaling state. *Proc Natl Acad Sci U S A* **90**, 7849-7853 (1993).
80 Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res* **28**, 235-242, (2000).
81 McGeehan, J. *et al.* Colouring cryo-cooled crystals: online microspectrophotometry. *J Synchrotron Radiat* **16**, 163-172, (2009).
82 Katona, G. *et al.* Raman-assisted crystallography reveals end-on peroxide intermediates in a nonheme iron enzyme. *Science* **316**, 449-453, (2007).
83 McGeehan, J. E., Carpentier, P., Royant, A., Bourgeois, D. & Ravelli, R. B. X-ray radiation-induced damage in DNA monitored by online Raman. *J Synchrotron Radiat* **14**, 99-108, (2007).
84 Bourgeois, D., Katona, G., de Rosny, E. & Carpentier, P. Raman-assisted X-ray crystallography for the analysis of biomolecules. *Methods Mol Biol* **544**, 253-267, (2009).
85 Kack, H., Gibson, K. J., Lindqvist, Y. & Schneider, G. Snapshot of a phosphorylated substrate intermediate by kinetic crystallography. *Proc Natl Acad Sci U S A* **95**, 5495-5500 (1998).
86 Wilmouth, R. C. *et al.* X-ray snapshots of serine protease catalysis reveal a tetrahedral intermediate. *Nat Struct Biol* **8**, 689-694, (2001).
87 Berglund, G. I. *et al.* The catalytic pathway of horseradish peroxidase at high resolution. *Nature* **417**, 463-468, (2002).
88 Chu, K. *et al.* Structure of a ligand-binding intermediate in wild-type carbonmonoxy myoglobin. *Nature* **403**, 921-923, (2000).
89 Schlichting, I., Berendzen, J., Phillips, G. N., Jr. & Sweet, R. M. Crystal structure of photolysed carbonmonoxy-myoglobin. *Nature* **371**, 808-812, (1994).
90 Edman, K. *et al.* High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature* **401**, 822-826, (1999).

91 Katona, G. *et al.* Conformational regulation of charge recombination reactions in a photosynthetic bacterial reaction center. *Nat Struct Mol Biol* **12**, 630-631, (2005).
92 Srajer, V. *et al.* Photolysis of the carbon monoxide complex of myoglobin: nanosecond time-resolved crystallography. *Science* **274**, 1726-1729 (1996).
93 Bourgeois, D. *et al.* Complex landscape of protein structural dynamics unveiled by nanosecond Laue crystallography. *Proc Natl Acad Sci U S A* **100**, 8704-8709, (2003).
94 Schotte, F. *et al.* Watching a protein as it functions with 150-ps time-resolved x-ray crystallography. *Science* **300**, 1944-1947, (2003).
95 Genick, U. K. *et al.* Structure of a protein photocycle intermediate by millisecond time-resolved crystallography. *Science* **275**, 1471-1475 (1997).
96 Perman, B. *et al.* Energy transduction on the nanosecond time scale: early structural events in a xanthopsin photocycle. *Science* **279**, 1946-1950 (1998).
97 Ren, Z. *et al.* A molecular movie at 1.8 A resolution displays the photocycle of photoactive yellow protein, a eubacterial blue-light receptor, from nanoseconds to seconds. *Biochemistry-Us* **40**, 13788-13801, (2001).
98 Anderson, S. *et al.* Chromophore conformation and the evolution of tertiary structural changes in photoactive yellow protein. *Structure* **12**, 1039-1045, (2004).
99 Baxter, R. H. *et al.* Time-resolved crystallographic studies of light-induced structural changes in the photosynthetic reaction center. *Proc Natl Acad Sci U S A* **101**, 5982-5987, (2004).
100 Schlichting, I. & Goody, R. S. Triggering methods in crystallographic enzyme kinetics. *Methods Enzymol* **277**, 467-490 (1997).
101 Svergun, D. I. & Koch, M. H. J. Small-angle scattering studies of biological macromolecules in solution. *Rep Prog Phys* **66**, 1735-1782 (2003).
102 Svergun, D. I., Petoukhov, M. V., Konarev, P. V. & Kikhney, A. G. ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. *J Appl Crystallogr* **40**, S223-S228, (2007).
103 Svergun, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* **76**, 2879-2886 (1999).
104 Svergun, D. I., Konarev, P. V. & Petoukhov, M. V. MASSHA - a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data. *J Appl Crystallogr* **34**, 527-532 (2001).
105 Hirai, M., Iwase, H., Hayakawa, T., Miura, K. & Inoue, K. Structural hierarchy of several proteins observed by wide-angle solution scattering. *Journal of Synchrotron Radiation* **9**, 202-205 (2002).
106 Makowski, L. Characterization of proteins with wide-angle X-ray solution scattering (WAXS). *J Struct Funct Genomics* **11**, 9-19, (2010).
107 Neutze, R. *et al.* Visualizing photochemical dynamics in solution through picosecond x-ray scattering. *Phys Rev Lett* **87**, (2001).
108 Plech, A. *et al.* Visualizing chemical reactions in solution by picosecond x-ray diffraction. *Phys Rev Lett* **92**, (2004).
109 Davidsson, J. *et al.* Structural determination of a transient isomer of CH2I2 by picosecond x-ray diffraction. *Phys Rev Lett* **94**, (2005).
110 Ihee, H. *et al.* Ultrafast x-ray diffraction of transient molecular structures in solution. *Science* **309**, 1223-1227, (2005).
111 Ihee, H. *et al.* Impulsive solvent heating probed by picosecond x-ray diffraction. *J Chem Phys* **124**, (2006).
112 Davidsson, J. *et al.* Picosecond calorimetry: Time-resolved x-ray diffraction studies of liquid CH2Cl2. *J Chem Phys* **124**, (2006).
113 Cammarata, M. *et al.* Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering. *Nat Methods* **5**, 881-886, (2008).
114 Serduk, I. N., Zaccai, N. R. & Zaccai, J. *Methods in Molecular Biophysics.* (Cambridge University Press, 2007).
115 Serdyuk, I. N., Zaccai, N. R. & Zaccai, J. *Methods in Molecular Biophysics.* (Cambridge University Press, 2007).
116 Thompson, A. *et al. X-ray data booklet.* (Lawrence Berkeley National Laboratory, 2009).
117 Ihee, H., Wulff, M., Kim, J. & Adachi, S. Ultrafast X-ray scattering: structural dynamics from diatomic to protein molecules. *Int Rev Phys Chem* **29**, 453-520, (2010).
118 Cammarata, M. *et al.* Chopper system for time resolved experiments with synchrotron radiation. *Rev Sci Instrum* **80**, 015101, (2009).
119 Wulff, M. *et al.* The realization of sub-nanosecond pump and probe experiments at the ESRF. *Faraday Discussions* **122**, 13-26, (2003).
120 Lee, J. H. & Ihee, H. Advantages of time-resolved difference X-ray solution scattering curves in anlyzing solute molecular structure. *Structural chemistry* **21**, 37-42 (2009).
121 Cammarata, M. *et al.* Impulsive solvent heating probed by picosecond x-ray diffraction. *J Chem Phys* **124**, (2006).

122 Henry, E. R. & Hofrichter, J. Singular Value Decomposition: Application to Analysis of Experimental Data. *Methods in Enzymologi* **210**, 129-192 (1992).
123 Shrager, R. I. & Hendler, R. W. Titration of Individual Components in a Mixture with Resolution of Difference Spectra, Pks, and Redox Transitions. *Anal Chem* **54**, 1147-1152 (1982).
124 Svergun, D., Barberato, C. & Koch, M. H. J. CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**, 768-773 (1995).
125 Zhang, R. T., Thiyagarajan, P. & Tiede, D. M. Probing protein fine structures by wide angle solution X-ray scattering. *J Appl Crystallogr* **33**, 565-568 (2000).
126 Lipfert, J., Columbus, L., Chu, V. B. & Doniach, S. Analysis of small-angle X-ray scattering data of protein-detergent complexes by singular value decomposition. *Applied Crystallography* **40**, 235-239 (2007).
127 Henderson, R. & Unwin, P. N. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* **257**, 28-32 (1975).
128 Henderson, R. *et al.* Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J Mol Biol* **213**, 899-929, (1990).
129 Landau, E. M. & Rosenbusch, J. P. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci U S A* **93**, 14532-14535 (1996).
130 Belrhali, H. *et al.* Protein, lipid and water organization in bacteriorhodopsin crystals: a molecular view of the purple membrane at 1.9 A resolution. *Structure* **7**, 909-917, (1999).
131 Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P. & Lanyi, J. K. Structure of bacteriorhodopsin at 1.55 A resolution. *J Mol Biol* **291**, 899-911, (1999).
132 Subramaniam, S. & Henderson, R. Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature* **406**, 653-657, (2000).
133 Matsui, Y. *et al.* Specific damage induced by X-ray radiation and structural changes in the primary photoreaction of bacteriorhodopsin. *J Mol Biol* **324**, 469-481, (2002).
134 Schobert, B., Cupp-Vickery, J., Hornak, V., Smith, S. & Lanyi, J. Crystallographic structure of the K intermediate of bacteriorhodopsin: conservation of free energy after photoisomerization of the retinal. *J Mol Biol* **321**, 715-726, (2002).
135 Edman, K. *et al.* Deformation of helix C in the low temperature L-intermediate of bacteriorhodopsin. *J Biol Chem* **279**, 2147-2158, (2004).
136 Kouyama, T., Nishikawa, T., Tokuhisa, T. & Okumura, H. Crystal structure of the L intermediate of bacteriorhodopsin: evidence for vertical translocation of a water molecule during the proton pumping cycle. *J Mol Biol* **335**, 531-546, (2004).
137 Lanyi, J. K. & Schobert, B. Mechanism of proton transport in bacteriorhodopsin from crystallographic structures of the K, L, M1, M2, and M2' intermediates of the photocycle. *J Mol Biol* **328**, 439-450,(2003).
138 Lanyi, J. K. & Schobert, B. Structural changes in the L photointermediate of bacteriorhodopsin. *J Mol Biol* **365**, 1379-1392, (2007).
139 Royant, A. *et al.* Helix deformation is coupled to vectorial proton transport in the photocycle of bacteriorhodopsin. *Nature* **406**, 645-648, (2000).
140 Facciotti, M. T. *et al.* Structure of an early intermediate in the M-state phase of the bacteriorhodopsin photocycle. *Biophys J* **81**, 3442-3455, (2001).
141 Lanyi, J. K. & Schobert, B. Crystallographic structure of the retinal and the protein after deprotonation of the Schiff base: The switch in the bacteriorhodopsin photocycle. *Journal of Molecular Biology* **321**, 727-737, (2002).
142 Lanyi, J. K. & Schobert, B. Propagating structural perturbation inside bacteriorhodopsin: Crystal structures of the M state and the D96A and T46V mutants. *Biochemistry-Us* **45**, 12003-12010, (2006).
143 Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P. & Lanyi, J. K. Structural changes in bacteriorhodopsin during ion transport at 2 Angstrom resolution. *Science* **286**, 255-260 (1999).
144 Luecke, H. *et al.* Coupling photoisomerization of retinal to directional transport in bacteriorhodopsin. *Journal of Molecular Biology* **300**, 1237-1255 (2000).
145 Sass, H. J. *et al.* Structural alterations for proton translocation in the M state of wild-type bacteriorhodopsin. *Nature* **406**, 649-653 (2000).
146 Schobert, B., Brown, L. S. & Lanyi, J. K. Crystallographic intermediates of structures of the M and N bacteriorhodopsin: Assembly of a hydrogen-bonded chain of water molecules between Asp-96 and the retinal Schiff base. *Journal of Molecular Biology* **330**, 553-570, (2003).
147 Takeda, K. *et al.* Crystal structure of the M intermediate of bacteriorhodopsin: Allosteric structural changes mediated by sliding movement of a transmembrane helix. *Journal of Molecular Biology* **341**, 1023-1037, (2004).
148 Rouhani, S. *et al.* Crystal structure of the D85S mutant of bacteriorhodopsin: Model of an O-like photocycle intermediate. *Journal of Molecular Biology* **313**, 615-628 (2001).
149 Hirai, T., Subramaniam, S. & Lanyi, J. K. Structural snapshots of conformational changes in a seven-helix membrane protein: lessons from bacteriorhodopsin. *Curr Opin Struc Biol* **19**, 433-439, (2009).
150 de Groot, B. L. *et al.* Prediction of protein conformational freedom from distance constraints. *Proteins* **29**, 240-251, (1997).

151 Subramaniam, S., Gerstein, M., Oesterhelt, D. & Henderson, R. Electron diffraction analysis of structural changes in the photocycle of bacteriorhodopsin. *EMBO J* **12**, 1-8 (1993).

152 Kralj, J. M., Spudich, E. N., Spudich, J. L. & Rothschild, K. J. Raman spectroscopy reveals direct chromophore interactions in the Leu/Gln105 spectral tuning switch of proteorhodopsins. *J Phys Chem B* **112**, 11770-11776, (2008).

153 Hiraki, K. *et al.* Bacteriorhodopsin analog regenerated with 13-desmethyl-13-iodoretinal. *Biophys J* **83**, 3460-3469, (2002).

154 Yang, J., Aslimovska, L. & Glaubitz, C. Molecular Dynamics of Proteorhodopsin in Lipid Bilayers by Solid-State NMR. *J Am Chem Soc* **133**, 4874-4881, (2011).

155 Delaney, J. K. & Subramaniam, S. The residues Leu 93 and Asp 96 act independently in the bacteriorhodopsin photocycle: Studies with the Leu 93->Ala, Asp 96->Asn double mutant. *Biophys J* **70**, 2366-2372 (1996).

156 Heinrich, B. *et al.* PILATUS: A single photon counting pixel detector for X-ray applications. *Nuclear Instruments & Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **607**, 247-249 (2009).

157 Fischetti, R. F. *et al.* High-resolution wide-angle X-ray scattering of protein solutions: effect of beam dose on protein integrity. *J Synchrotron Radiat* **10**, 398-404, (2003).

158 Dinapoli, R. *et al.* A new family of pixel detectors for high frame rate X-ray applications. *Nucl Instrum Meth A* **617**, 384-386, (2010).

159 Weis, W. I. & Kobilka, B. K. Structural insights into G-protein-coupled receptor activation. *Curr Opin Struc Biol* **18**, 734-740, (2008).

160 Smith, S. O. Structure and activation of the visual pigment rhodopsin. *Annu Rev Biophys* **39**, 309-328, (2010).

161 Vogel, R. *et al.* Tracking G-protein-coupled receptor activation using genetically encoded infrared probes. *Nature* **464**, 1386-U1314, (2010).

162 Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L. & Khorana, H. G. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **274**, 768-770 (1996).

163 Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *P Natl Acad Sci USA* **105**, 7439-7444, (2008).

164 Farrens, D. L. & Dunham, T. D. Conformational changes in rhodopsin - Movement of helix F detected by site-specific chemical labeling and fluorescence spectroscopy. *J Biol Chem* **274**, 1683-1690 (1999).

165 Terakita, A., Tsukamoto, H., Farrens, D. L. & Koyanagi, M. The Magnitude of the Light-induced Conformational Change in Different Rhodopsins Correlates with Their Ability to Activate G Proteins. *J Biol Chem* **284**, 20676-20683, (2009).

166 Nakamichi, H. & Okada, T. Crystallographic analysis of primary visual photochemistry. *Angew Chem Int Ed Engl* **45**, 4270-4273, (2006).

167 Nakamichi, H. & Okada, T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc Natl Acad Sci U S A* **103**, 12729-12734, (2006).

168 Kuwata, O., Yuan, C., Misra, S., Govindjee, R. & Ebrey, T. G. Kinetics and pH dependence of light-induced deprotonation of the Schiff base of rhodopsin: Possible coupling to proton uptake and formation of the active form of meta II. *Biochemistry-Moscow+* **66**, 1283-1299 (2001).

169 Heyn, M. P., Hoersch, D., Otto, H. & Wallat, I. Monitoring the Conformational Changes of Photoactivated Rhodopsin from Microseconds to Seconds by Transient Fluorescence Spectroscopy. *Biochemistry* **47**, 11518-11527, (2008).

170 Schneider, T. R. A genetic algorithm for the identification of conformationally invariant regions in protein molecules. *Acta Crystallogr D Biol Crystallogr* **58**, 195-208, (2002).

171 Li, J., Edwards, P. C., Burghammer, M., Villa, C. & Schertler, G. F. Structure of bovine rhodopsin in a trigonal crystal form. *J Mol Biol* **343**, 1409-1438, (2004).

172 Standfuss, J. *et al.* Crystal structure of a thermally stable rhodopsin mutant. *J Mol Biol* **372**, 1179-1188, (2007).

173 Okada, T. *et al.* The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. *J Mol Biol* **342**, 571-583, (2004).

174 Fotiadis, D. *et al.* Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **421**, 127-128, (2003).

175 Liang, Y. *et al.* Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. *J Biol Chem* **278**, 21655-21662, (2003).

176 Suda, K., Filipek, S., Palczewski, K., Engel, A. & Fotiadis, D. The supramolecular structure of the GPCR rhodopsin in solution and native disc membranes. *Mol Membr Biol* **21**, 435-446, (2004).

177 Montgomery, D. C. (John Wiley & Sons, Inc., 2005).

178 Rasmussen, S. G. *et al.* Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549-555, (2011).

179 Feig, M., Stumpff-Kane, A. W., Maksimiak, K. & Lee, M. S. Sampling of near-native protein conformations during protein structure refinement using a coarse-grained model, normal modes, and molecular dynamics simulations. *Proteins* **70**, 1345-1356, (2008).

180 Seeliger, D. & De Groot, B. L. tCONCOORD-GUI: visually supported conformational sampling of bioactive molecules. *J Comput Chem* **30**, 1160-1166, (2009).

181 Engelman, D. M. & Bu, Z. M. A method for determining transmembrane helix association and orientation in detergent micelles using small angle x-ray scattering. *Biophys J* **77**, 1064-1073 (1999).

182 Sunahara, R. K. *et al.* Efficient coupling of transducin to monomeric rhodopsin in a phospholipid bilayer. *Journal of Biological Chemistry* **283**, 4387-4394, (2008).

Appendix

DERIVATION OF DIFFERENCE SCATTERING EXPRESSION

To derive the difference scattering expression in Eq. 21 and 22, we start by expanding Eq. 19,

$$S= \langle |F_a - \rho_s F_s|^2 \rangle = \langle |F_a|^2 - \rho_s F_a F_s^* - \rho_s F_s F_a^* + \rho_s^2 |F_s|^2 \rangle =$$

$$\langle |F_a|^2 - \rho_s |F_a||F_s| e^{i(\varphi_a - \varphi_s)} - \rho_s |F_a||F_s| e^{i(\varphi_s - \varphi_a)} + \rho_s^2 |F_s|^2 \rangle =$$

$$\langle |F_a|^2 - \rho_s |F_a||F_s| (e^{i\Delta\varphi_{as}} + e^{-i\Delta\varphi_{as}}) + \rho_s^2 |F_s|^2 \rangle =$$

$$\langle |F_a|^2 - 2\rho_s |F_a||F_s| \cos\Delta\varphi_{as} + \rho_s^2 |F_s|^2 \rangle \,, \qquad \text{(A.1)}$$

where $|F_a|$ and $|F_s|$ represents the magnitudes of the atomic (a) and excluded volume (s) scattering amplitudes of the protein-micelle complex, $\Delta\varphi_{as}$ is the phase difference between the two terms and ρ_s is the density of the solvent. By rearranging Eq. A.1 it is possible to formulate the term describing the phase difference as,

$$\cos\Delta\varphi_{as} = \frac{|F_a|^2 + \rho_s^2 |F_s|^2 - S}{2\rho_s \langle |F_a| \rangle \langle |F_s| \rangle} \,, \qquad \text{(A.2)}$$

I. The total difference scattering, ΔS

We now consider the total difference scattering between an activated state (') and a resting state,

$$\Delta S = S' - S =$$

$$\langle |F_a'|^2 - 2\rho_s |F_a'||F_s'| \cos\Delta\varphi_{as}' + \rho_s^2 |F_s'|^2 \rangle - \langle |F_a|^2 - 2\rho_s |F_a||F_s| \cos\Delta\varphi_{as} + \rho_s^2 |F_s|^2 \rangle \,,$$

we now make the assumption that the change in excluded volume scattering is relatively small so that (i) $|F_s'| \approx |F_s|$ and (ii) $\cos\Delta\varphi_{as}' \approx \cos\Delta\varphi_{as}$ so that,

$$\langle |F_a'|^2 - 2\rho_s |F_a'||F_s| \cos\Delta\varphi_{as} + \rho_s^2 |F_s|^2 \rangle - \langle |F_a|^2 - 2\rho_s |F_a||F_s| \cos\Delta\varphi_{as} + \rho_s^2 |F_s|^2 \rangle =$$

$$\langle |F_a'|^2 - 2\rho_s |F_s| (|F_a'| - |F_a|) \cos\Delta\varphi_{as} - |F_a|^2 \rangle \,,$$

by recognizing that $\Delta S_a = |F_a'|^2 - |F_a|^2 = (|F_a'| - |F_a|)(|F_a'| + |F_a|)$ we can then write,

$$\Delta S = \Delta S_a (1 - 2\rho_s \frac{\langle |F_s| \rangle}{\langle |F_a'| + |F_a| \rangle} \cos\Delta\varphi_{as}) \approx \Delta S_a (1 - \rho_s \frac{\langle |F_s| \rangle}{\langle |F_a| \rangle} \cos\Delta\varphi_{as}) \qquad \text{(A.3)}$$

By combining Eq. A.1 and Eq. A.3 we can now express the difference scattering between the two systems as,

$$\Delta S = \Delta S_a (1 - \rho_s \frac{\langle |F_s| \rangle}{\langle |F_a| \rangle} \frac{(|F_a|^2 + \rho_s^2 |F_s|^2 - S)}{2\rho_s \langle |F_a| \rangle \langle |F_s| \rangle}) = \Delta S_a (1 - \frac{(|F_a|^2 + \rho_s^2 |F_s|^2 - S)}{2|F_a|^2}) \qquad \text{(A.4)}$$

We can analogously express the atomic scattering for the protein-micelle system (a) as a function of the individual atomic scattering contributions from the protein (p) and the micelle (m) such that,

$$S_a = \langle |F_a|^2 \rangle = \langle \left|F_p + F_m\right|^2 \rangle = \langle \left|F_p\right|^2 + F_p F_m^* + F_m F_p^* + |F_m|^2 \rangle =$$

$$\langle \left|F_p\right|^2 + \left|F_p\right||F_m| e^{i(\varphi_p - \varphi_m)} + \left|F_p\right||F_m| e^{i(\varphi_p - \varphi_m)} + |F_m|^2 \rangle =$$

$$\langle \left|F_p\right|^2 + \left|F_p\right||F_m| (e^{i\Delta\varphi_{pm}} + e^{-i\Delta\varphi_{pm}}) + |F_m|^2 \rangle =$$

$$\langle \left|F_p\right|^2 + 2\left|F_p\right||F_m| \cos\Delta\varphi_{pm} + |F_m|^2 \rangle \qquad \text{(A.5)}$$

where, $\Delta\varphi_{pm}$ is the phase difference between the protein and the micelle and can be written,

$$\cos\Delta\varphi_{pm} = \frac{|F_a|^2 - \left|F_p\right|^2 - |F_m|^2}{2\langle |F_p| \rangle \langle |F_m| \rangle}, \qquad \text{(A.6)}$$

II. The atomic difference scattering, ΔS_a

We now consider the atomic difference scattering between the activated state (') and the resting state protein-micelle ,

$$\Delta S_a = S_a{}' - S_a =$$

$$\langle \left|F_p'\right|^2 + 2\left|F_p'\right||F_m'| \cos\Delta\varphi_{pm}' + |F_m'|^2 \rangle - \langle \left|F_p\right|^2 + 2\left|F_p\right||F_m| \cos\Delta\varphi_{pm} + |F_m|^2 \rangle ,$$

we make the assumption that the change in micelle scattering is small so that (i) $|F_m'| \approx |F_m|$ and (ii) $\cos\Delta\varphi_{pm}' \approx \cos\Delta\varphi_{pm}$ so that,

$$\langle \left|F_p'\right|^2 + 2\left|F_p'\right||F_m| \cos\Delta\varphi_{pm} + |F_m|^2 \rangle - \langle \left|F_p\right|^2 + 2\left|F_p\right||F_m| \cos\Delta\varphi_{pm} + |F_m|^2 \rangle =$$

$$\langle \left|F_p'\right|^2 + 2|F_m|(\left|F_p'\right| - \left|F_p\right|)\cos\Delta\varphi_{pm} - \left|F_p\right|^2 \rangle ,$$

again by recognizing that $\Delta S_p = \left|F_p'\right|^2 - \left|F_p\right|^2 = (\left|F_p'\right| - \left|F_p\right|)(\left|F_p'\right| + \left|F_p\right|)$ we can then write,

$$\Delta S_a = \Delta S_p (1 + 2\frac{\langle |F_m| \rangle}{\langle |F_p'| + |F_p| \rangle} \cos\Delta\varphi_{pm}) \approx \Delta S_p (1 + \frac{\langle |F_m| \rangle}{\langle |F_p| \rangle} \cos\Delta\varphi_{pm}) \qquad \text{(A.7)}$$

By combining Eq. A.6 and Eq. A.7 we can now express the atomic difference scattering between the two protein-micelles as,

$$\Delta S_a = \Delta S_p (1 + \frac{\langle |F_m| \rangle}{\langle |F_p| \rangle} \frac{\left(|F_a|^2 - \left|F_p\right|^2 - |F_m|^2\right)}{2\langle |F_p| \rangle \langle |F_m| \rangle}) = \Delta S_p (1 + \frac{\left(|F_a|^2 - \left|F_p\right|^2 - |F_m|^2\right)}{2\left|F_p\right|^2}) \qquad \text{(A.8)}$$

III. The resulting difference scattering expression

By inserting Eq. A.8 in Eq. A.4 the total difference can be expressed scattering as,

$$\Delta S = \Delta S_p \left[1 + \frac{\left(|F_a|^2 - \left|F_p\right|^2 - |F_m|^2\right)}{2\left|F_p\right|^2}\right] \left[1 - \frac{(|F_a|^2 + \rho_s^2 |F_s|^2 - S)}{2|F_a|^2}\right] =$$

$$\frac{\Delta S_p}{4}\left[1+\left(\frac{|F_a|^2-|F_m|^2}{|F_p|^2}\right)\right]\left[1-\left(\frac{\rho_s^2|F_s|^2-S}{|F_a|^2}\right)\right],$$

By recognizing that $|F_a|^2$ is the atomic scattering intensity for the protein-micelle, $|F_m|^2$ is the atomic scattering intensity for the micelle, $|F_p|^2$ is the atom scattering intensity for the protein and $\rho_s^2|F_s|^2$ is the excluded volume scattering for the protein-micelle in the resting state conformation we can formulate the total difference scattering as a function of intensities,

$$\Delta S = \Delta S_p \ \frac{1}{4}\left[1+\left(\frac{S_a^{pm}-S_a^m}{S_a^p}\right)\right]\left[1-\left(\frac{S_s^{pm}-S}{S_a^{pm}}\right)\right], \qquad \text{(A.9)}$$

where $\Delta S_p = S_p' - S_p$, is the atomic difference scattering between active state and resting state of the protein. The program CRYSOL can in turn be used to calculate the individual scattering intensities directly from the atomic coordinates.

It should be emphasised that the above expression is an approximation which assumes that the average conformation of the detergent/lipid micelle surrounding the protein does not significantly change upon conformational changes in the protein. For a complete description of the activated micelle a more rigourous model would have to be explored.