Studying membrane protein structure and function using nanodiscs

Pie Huda

Structural Biophysics
X-ray and Neutron Science
Niels Bohr Institute
University of Copenhagen

Ph.D. thesis
June 2015

Front cover photo: The European Synchrotron Facility & Institute Laue-Langevin, Grenoble, France (credit: ESRF/A. Petricola).

Preface

The present thesis is submitted for the degree of Doctor of Philosophy at the University of Copenhagen. The research described in the thesis was conducted under the supervision of Professor Lise Arleth and carried out at the Niels Bohr Institute and the Department of Plant and Environmental Science at Faculty of Science, University of Copenhagen, Denmark, during the period April 2012 to June 2015. This includes a 6- month's research stay abroad in Professor Jenny Martin's group, at Institute for Molecular Bioscience, The University of Queensland, Australia.

Copenhagen, January 2014

In addition to the presented work I have been involved in other scientific projects which have resulted in three publications:

Huda et al. (2015) PET/CT based *in vivo* evaluation of ^{64}Cu labelled nanodiscs in tumor bearing mice. *PLOS ONE*, *Accepted*

Järvå et al. (2015) SAXS and SANS investigation of SoPIP2;1 Aquaporin-tetramers in POPC-nanodiscs. Bench-mark of the nanodisc approach to extract structural information about membrane proteins, *Submitted*

Stevenson et al. (2015) Investigating the function of Fc-specific binding of IgM to *Plasmodium falciparum* erythrocyte membrane protein 1 mediating erythrocyte resetting. *Cellular Microbiology* **17:** 819-891

Acknowledgements

Abstract

The structure and dynamic of membrane proteins can provide valuable information about general functions, diseases and effects of various drugs. Studying membrane proteins are a challenge as an amphiphilic environment is necessary to stabilise the protein in a functionally and structurally relevant form. This is most typically achieved through the use of detergent based reconstitution systems. However, time and again such systems fail to provide a suitable environment causing aggregation and inactivation.

Nanodiscs are self-assembled lipoproteins containing two membrane scaffold proteins and a lipid bilayer in defined nanometer size, which can act as a stabiliser for membrane proteins. This enables both functional and structural investigation of membrane proteins in a detergent free environment which is closer to the native situation. Understanding the self-assembly of nanodiscs is important for understanding the key mechanisms during reconstitution of membrane proteins in these lipoproteins. In this project the self-assembly of nanodiscs has been structurally characterized with small angle X-ray scattering (SAXS) in a time resolved fashion. This brought knowledge about the structural development as detergent is removed from the solution. This also provided valuable information useful for optimal reconstitution of membrane proteins in nanodiscs. The knowledge was utilized in the reconstitution of proteorhodopsin in nanodiscs where buffer compositions and reconstitution detergents were varied. The different reconstitutions of proteorhodopsin were subsequently analysed by small angle X-ray scattering to evaluate the structural impact of these factors. This shed light on influences that are important to consider in the reconstitution process. In regards to the structure analysis of membrane proteins in nanodiscs it is desirable to acquire structural information supplementary to that obtained through SAXS. Such information can in theory be obtained using small angle neutron scattering (SANS) by using a D_2O based buffer system. Unfortunately D_2O induced aggregation is a commonly observed problem for biomolecules which also hampered SANS studies in this project. For this reason nanodiscs were systematically analysed at different D_2O buffer compositions varying salt, PH etc. to clarify this problem. Unfortunately, a solid solution was not found but an influencing factor identified.

The experiences from the proteorhodopsin experiments provided the basis for obtaining a successful reconstitution of magnesium transporter CorA and of disulphide bond forming protein DsbB.

Reconstitution of CorA in nanodiscs allowed for structural investigation of the membrane protein by SAXS. It was possible to obtain high quality structural information of two conformations of CorA, providing interesting new insights. A conformation of CorA which has so far only been structurally characterized by molecular dynamics simulations was explored. Surprisingly it was also found that the nanodisc is able to assemble from four MSPs to accommodate membrane protein needs. The outcome of the structural analysis clearly exposed the potentials and future perspectives of this method.

Reconstitution of DsbB in nanodiscs allowed for improved activity and inhibition assays of the disulphide bond forming machinery. Additionally membrane protein activity could be kept for substantially longer in nanodiscs compared to the usual conditions. This enables new types of assays where protein stability is not a time limiting factor. The experiment evidently demonstrated the superiority of using nanodiscs in membrane protein assays.

To enable structural dynamics studies of membrane proteins new methods need to be developed. Low sample consumption is crucial when working with this type of proteins and has to be taken into consideration when developing new techniques. To accommodate this, a microfluidic chip for time resolved SAXS of proteins was developed and investigated. It was possible to collect information of kinetic processes with sub-millisecond time resolution of simple protein systems with minimal sample consumption. The obtained results show very promising prospects for this method and for future applications in studying the dynamics of membrane proteins, using nanodiscs as stabilising agent.

Resumé

At have kendskab til struktur og dynamik af membranproteiner kan bidrage med vigtig information omkring deres generelle egenskaber, funktioner i sygdomme og forskellige lægemidlers virkning. At studere membranproteiner er udfordrene, da de er afhængige af et amfifilt miljø for at opretholde deres native struktur og funktion. Dette opnået oftest ved at opløse proteinet med detergenter. Gentagende gange observeres det dog, at detergenter ikke kan reproducere det nødvendige miljø, hvilket resulterer i udfældning og inaktivering af proteinet.

Nanodiske er selvansamlende lipoproteiner bestående af et lipid dobbelt lag, omringet af to membrandefinerende proteiner i nanometer størrelse, som kan fungere som stabilisator for membranproteiner. Membranproteiner kan herved undersøges både strukturelt og funktionelt i et detergentfrit miljø, som bidrager med et mere nativt miljø for membranproteiner. Det er vigtigt at få forståelse for selvansamlingen af nanodiske, da dette bidrager til viden om de essentielle mekanismer der foregår, når membranproteiner indbygges i nanodiske. I dette projekt er selvansamlingen af nanodiske blevet tidsopløst strukturelt karekteriseret med småvinkel røngten spredning (SAXS). Denne karakterisering bidrog til viden om den strukturelle udvikling af nanodiske, hvilket foregår ved fjernelse af detergent. Desuden gav dette brugbar information, der er nyttig for optimering af indbyggelsen af membranproteiner i nanodiske. Denne indsigt blev udnyttet til at indbygge proteorhodopsin i nanodiske, hvor også buffersammensætningen og detergenten brugt til selvansamlingen, blev varieret. De forskellige proteorhodopsinprøver blev derefter undersøgt med SAXS, for at finde den strukturelle inflydelse af disse variationer. Dette belyste hvilke faktorer der er vigtige i selvansamlingsprocessen. Med hensyn til strukturel analyse af membranproteiner i nanodiske, er det attraktivt at få supplerende strukturel information udover SAXS data. Denne inforation kan teoretisk set tilegnes med småvinkel neutron spredning (SANS) ved at anvende en D_2O baseret buffer. Desværre er det ofte observeret at D_2O foresager udfældning af biologiske molekyler, hvilket desuden har ført til foringelser af SANS data i dette projekt. Derfor blev nanodiske systematisk analyseret i forskellige D_2O buffersammensætninger med varierende salt, pH osv., med henblik på at løse dette problem. Desværre var dette ikke muligt, men en vigtig faktor for denne aggregering blev fundet.

Erfaringerne fra proteorhodopsinforsøgene gav grundlaget til at indbygge magnesium transportør CorA og disulfidbinding protein DsbB, i nanodiske med gode resultater.

Indbygning af CorA i nanodiske muliggjorde strukturel undersøgelse af membranproteinet ved hjælp af SAXS. Det var mulig at få strukturel information af høj kvalitet af to forskellige konformationer af CorA, hvilket resulterede i ny spændende viden. En konformation der ind til nu kun har været strukturelt karekteriseret med molecular dynamics simulering, blev undersøgt. Derudover blev det forbavsende observeret, at nanodiske er i stand til at selvansamle med fire membrandefinerende proteiner for at tilfredsstille membranproteiners behov. Den strukturelle analyse viste tydeligt potentiale og fremtidsperspektiv for denne metode.

Indbygning af DsbB i nanodiske muliggjorde forbedrede aktivitets og inhibitor analyser af disulfidbindingsmekanismen. Derudover var det mulig at bibeholde aktiviteten af membranproteinet i væsentlig længere tid i nanodiske når man sammenligner med normalt benyttede betingelser. Dette muliggør nye typer af analyser, hvor proteinaktivitet ikke er en tidsbegrændsende faktor. Eksperimenterne tydeliggjorde hvor suverene nanodiske er i brugen til membranproteinaktivitetsanalyser.

For at studere strukturdynamik af membranproteiner er der behov for at udvikle nye metoder. Et lavt prøveforbrug er afgørende når man arbejder med denne type proteiner, som er vigtigt at tage hensyn til, når der udvikles nye metoder. For at imødekomme dette blev der udviklet en microfluid chip til tidsopløst SAXS, som blev karakteriseret. Det var muligt at indsamle information af kinetiske processer af simple proteiner med en tidsopløsning lavere end 1 millisekund og med et minimalt prøveforbrug. Resultaterne viser lovende usigter for denne metode og fremtidig brug til undersøgelsen af membranproteindynamik, hvor nanodiske kan benyttes som stabilisator.

Contents

Introduction

Membrane proteins contribute to about 30% of the cell's proteome (Stevens & Arkin, 2000). These proteins serve a variety of important roles in the body, such as regulation of blood pressure, fertilization and nerve excitation. Besides from the fact that many diseases are known to be caused by mutations in these proteins (Ashcroft, 2006), membrane proteins are central for pharmaceutical applications. 59% of the United States Food and Drug Administration (FDA) approved drug target proteins are predicted membrane proteins (Uhlen et al, 2015).

Unfortunately, very little is known about the structure-function relationships of membrane proteins. Only 2.5% of the total number of protein structures found in the Protein Data Bank (PDB) (Bernstein et al, 1977) is classified as membrane proteins (PDB, 2015b). The reason for this low fraction is presumably attributable to the general accepted fact that membrane proteins are very hard to handle and crystalize. For this reason, there is a great demand for alternative methods applicable to study the relationship between structure and function of membrane proteins.

During the last three years, I have been involved in different projects related to the application of nanodiscs for structural studies of membrane proteins, using small angle scattering. The overall aim of my PhD project was to combine this knowledge with microfluidics and time resolved small angle scattering to investigate the structure, action and dynamics of membrane proteins under physiological conditions. This has been a collaborative project where my focus mainly has been on optimising sample preparation for these studies. The project has shown to be very ambitious and still needs much work to reach the intended goal. The thesis will cover the developments that I have been involved in during this project.

The thesis is divided into two parts. First part will give a scientific introduction to the background and methods applied in the work presented in thesis, while results of the work will be presented in part two.

Part one – Scientific background and methods, is divided into four chapters

Chapter 1 - Protein structure and dynamics: A general introduction to protein dynamics and methods for studying protein structures and dynamics. The main focus will be on the use of small angle scattering in this context.

Chapter 2 - The cell membrane: An introduction to lipids and proteins in the biological membrane and the arrangement and packing of lipids.

Chapter 3 - Methods in protein isolation: Methods for bacterial protein expression and purification is introduced as well as methods for solubilising membrane proteins including the nanodisc.

Chapter 4 - Small angle scattering of biological systems: A short description of the theory behind small angle scattering, instrumentation and simple data analysis.

Part two – Results, is divided into four chapters covering six projects and a final remarks chapter covering discussion, perspectives and a final conclusion.

Chapter 5 - Understanding the reconstitution process of nanodiscs
        A. Optimising reconstitution of membrane proteins in nanodiscs
        B. The self-assembly of nanodiscs – time resolved
        C. Investigating aggregation of biomolecules in D_2O

Chapter 6 - Reconstitution of membrane proteins in nanodiscs
        D. Solution structure of magnesium transporter CorA in nanodiscs
        E. Activity and inhibition of disulphide bond protein DsbB in nanodiscs

Chapter 7 - Probing protein dynamics
        F. Developing a microfluidic chip for time resolved SAXS

Chapter 8 – Final Remarks

All projects will be presented in individual sections including results, discussion and conclusion. The essence of the different projects will be wrapped up in a final remarks section. **Chapter 5** covers different aspects of understanding the reconstitution of nanodiscs with and without membrane proteins. This includes optimising sample preparation and understanding which important factors are involved, and studying structural features in the self-assembly process of nanodiscs. Project A is a time resolved study of the self-assembly process of nanodiscs. This study elucidates the structural development of nanodiscs during detergent removal, to a better understanding of the self-assembly process. Project B examines important factors in the reconstitution of membrane proteins in nanodiscs aimed for optimisation of sample preparation for small angle scattering. Project C investigates aggregation of biomolecules when kept in D_2O buffer, based on aggregation of nanodiscs. This study was necessitated as exchanging to a D_2O buffer is an important tool for small angle neutron scattering and aggregation is a frequently observed problem diminishing data quality.

The three projects presented in Chapter 5 are presented as individual research reports.

Chapter 6 concerns two case studies of reconstituting membrane proteins in nanodiscs using the tools and knowledge obtained from the projects in chapter 5. Project D is a sophisticated structural study of a large membrane protein undergoing a subtle structural transition during ligand binding. The study proves the power of combining nanodiscs and small angle scattering for elucidating unknown structural conformations. Project E explores a membrane protein important in folding bacterial virulence proteins and clearly demonstrates the superiority of nanodiscs in membrane protein assays.

The two projects in Chapter 6 will be presented in the form of paper drafts.

Chapter 7 deals with research using a microfluidic chip for time resolved structural studies of proteins. Project F is an excellent example towards developing a sophisticated method for

studying protein dynamics with small angle scattering. The aim is to be able to study a system like the presented in Project D and using this setup to obtain time resolved structural information of the action of membrane proteins. This project will be presented as a research report.

Chapter 8 gives a final discussion and conclusions based on the presented projects.

In addition to the six presented projects, a project concerning optimising membrane protein expression in *E. coli* can be found in Appendix 3. The aim with this project was firstly to improve my skills in molecular biology and secondly to identify and optimise expression of possible membrane protein candidates for other projects. However, it was not possible to obtain a good candidate within the three months project. The report found in Appendix 3 summarizes the work carried out during this period.

Part one – Scientific background and methods will now follow. This will eventually lead to presentations of the before mentioned projects and a final discussion in Part two - Results.

Part one - Scientific background and methods

Chapter 1 - Protein structure and dynamics

Proteins are responsible for most functions in the cell. The three dimensional structure of a protein is crucial for its functionality which makes the obtained protein structures fundamental for understanding cellular function and diseases. Knowing the structure of a protein can help us understand basic mechanisms taking place in the cell and in advanced drug design (Bunney et al, 2014; Kolli et al, 2015; Mabanglo et al, 2014; McCusker et al, 2012).

Methods for determination of protein structure

For decades, X-ray crystallography has been and still is the method of choice for obtaining high quality structural information about proteins, recently marked by celebrating 2014 as international year of crystallography (IYCr2014) (Su et al, 2015). This technique provides detailed visualization with a resolution down to atomic level revealing the position of most atoms within the protein. In order to obtain these data, one first needs a crystal. This can be a bottleneck as some proteins require multiple screenings for optimising their crystallization process, and some proteins are natively too disordered and unstable to be crystalized (Tickle et al, 2004). For one type of proteins, namely membrane proteins, crystallization has shown to be a major challenge. First of all, obtaining the quantities needed for crystallization can be a challenge itself because of low expression and next, more sophisticated crystallization screening conditions need to be applied. It is a long and tedious process which in many cases leads to failure, and other complementary techniques may be considered (Kang et al, 2013).

Predicting the function of a specific protein based on its amino acid sequence alone is close to impossible as the function, as before mentioned, is determined by their 3D fold. Even though most proteins with similar sequences hold similar structures, it is shown that proteins holding similar structures and functions may have very low sequence homology (Onuchic et al, 1997). In order to overcome this problem, a number of computational methods have been developed to predict the structure of proteins from their amino acid sequences, each using different strategies (Schmidt et al, 2014; Shenoy & Jayaram, 2010). Of course, the result of such prediction is not a certain answer, but it may be used as a guideline for categorizing proteins. A successful example is the Rosetta project that is based on structure prediction by identifying the structure holding the lowest free-energetic state, based on the protein's amino acid sequence. This method has created several blindfolded protein structures with accuracy down to less than 2 Å (Das & Baker, 2008).

Other experimental "in solution" structural techniques can be applied in order to obtain the needed information and eliminate the need of a crystal. This also allows studying more flexible proteins which cannot undergo crystallization. Nuclear magnetic resonance (NMR) is such a technique, which within the last decade has made impressive advances (Kay, 2005; McDermott, 2009). This technique can provide high resolution structural information comparable with crystal structures, but is generally most applicable for smaller proteins <20 kDa (Yee et al, 2014). The reason for this is that with increasing protein size more crowded correlation spectra are obtained, creating a higher chance of overlaps between peaks and thereby a difficulty to distinguish protons from each other.

7

Low resolution techniques can be applied as well. Within recent years, cryo-EM has advanced significantly and a 3.2 Å resolution structure has now been solved using this strategy (Bartesaghi et al, 2014). The method is best suited for symmetrical larger proteins >100 kDa (Glaeser & Hall, 2011; Kuehlbrandt, 2014); however, it is a rapidly developing field that will be interesting to follow within the coming years.

Also small angle scattering is a low resolution method where a resolution of about 10 Å can be obtained. The technique has the advantage of being non-invasive and proteins can therefore be studied in a close to physiological environment. Also structural dynamics can be followed as a response to environmental changes (Bernado et al, 2007; Levantino et al, 2015). Small angle scattering has been the central structural technique for this thesis and will be explained in more details in Chapter 4 - Small angle scattering from biological systems.

Common for the mentioned experimental techniques is that a pure homogeneous protein sample is crucial. This may pose a major problem for many membrane proteins because of handling difficulties. Some of the challenges regarding membrane proteins will be addressed in Chapter 2 – The cell membrane and Chapter 3 – Methods in protein isolation.

Proteins are dynamic

Proteins are not static molecules, but can be built from multiple flexible domains able to move in relation to each other. The action of proteins is linked to certain changes in the local environment, such as calcium selective channels activated by voltage change over the membrane, caused by local calcium increase (Neely & Hidalgo, 2014). Activation of GPCR β_2-adrenergic receptor is associated with a structural change when binding to catecholamines, initiating a signal cascade (Zhang et al, 2010). The response or activation from this impact is in the form of structural changes, observed at time scales from picoseconds to milliseconds (van den Bedem & Fraser, 2015). Protein dynamics influence all protein function, among other things, protein complex formation, cell signaling and regulation, and the catalysis of different processes (Bu & Callaway, 2011). Put simply, studying the dynamics of a protein can provide valuable information about general function, disease or mutation, but also the effect of different drugs or environmental changes.

Studying protein dynamics on the structural level is a developing field and still many challenges apply, not least to derive different structural states from the obtained data. Circular dichroism (CD) is a method widely used in this field due to its high sensitivity to minor structural changes in a molecule (Whitmore & Wallace, 2008). Another powerful tool is NMR spectroscopy which can provide detailed information about dynamic regions in a molecule (Markwick et al, 2008). However, these techniques cannot provide specific information about the geometric motions alone. Combining biophysical, experimental and computational techniques can be necessary to obtain sufficient structural information (Ward et al, 2013). This does not only set high demands for sample control and quality but also require sophisticated computational models and interdisciplinary collaborations.

Probing structural dynamics of proteins

To trigger a conformational change in a protein, a change in the local environment must occur. This can be by introducing a ligand which interacts with the protein, a change in pH, jump in temperature or for photo sensitive proteins, by irradiating the protein with light of a specific energy. Probing protein structure changes can be achieved statically by trapping a protein in its different conformations and obtaining structures of each conformation, or time resolved in the sense of following the whole dynamic process of the conformational changes in real time. A great example is from the Nissen group who have been able to capture and crystalize Ca^{2+}-ATPase, active in calcium transport, in three different stages of the process (Olesen et al, 2007). Another example is Montoya and coworkers who have specialized in a method to capture catalytic intermediates of enzymes in a crystal, by slowing down the reaction rate. This allows them to obtain detailed structural information on the mechanism with high resolution (Molina et al, 2015).

Obviously, real time studies are more challenging because of the requirement of advanced instrumentation able to follow movements within an extremely short time frame. Another obstacle is to introduce the agent that induces the structural response, faster than the time scale of the kinetic process of interest.

Time resolved small angle scattering

Small angle scattering can be used for probing time resolved structural dynamics of proteins, though some challenges are associated to this.

- The experimental setup
- Sample environment for triggering a structural response
- Sample homogeneity and a detectable structural response

The experimental setup for these kinds of studies puts on high demands for beam line scientists and instrumental technicians. Fast detectors are essential in these experiments and a narrow beam with high flux is preferred. However, a few beam lines have met the challenges and optimised their instruments to allow obtaining time resolved data, such as ID02 (Angelov et al, 2015) and ID09 (Wulff et al, 1997) at ESRF, Grenoble and BioCAT at APS, Chicago (Graceffa et al, 2013).

Instantaneous triggering a structural change of a protein sample is where great challenges still remain. Various approaches are taken: For photo sensitive proteins, it is in principle easily done as photons will reach all proteins simultaneously when irradiating the sample. Using small angle X-ray scattering (SAXS) it has been possible to obtain nanosecond time resolution, probing the structural dynamics of human hemoglobin and inducing a conformational change by laser (Cammarata et al, 2008). This method has also been applied on more challenging proteins, namely the seven helical transmembrane protein bacteriorhodopsin and its closely related proteorhodopsin probing the structural shifts during proton pumping (Malmerberg et al, 2011). Coordinating a photon flash with rapid measurements and also achieve high time resolution is a challenge which has been solved with a sophisticated experimental setup. Finding good light sensitive protein candidates is now the real challenge in these kind of studies as the number of light sensitive proteins is limited (Reddington et al, 2013).

Studying ligand triggered structural changes is significantly more challenging than photo induced triggered, as fast mixing of ligand and protein is crucial for obtaining a simultaneous response. Operating in the narrow time frame of structural changes means that the ligand must be delivered homogeneously to the entire exposed sample within this window. One efficient method is to take advantage of the principle of photo sensitive reactions using photochemical cages. Here the ligand of interest is trapped inside the cage and released by flash photolysis. The caged ligand can be mixed into the protein solution, radiated at energies were the photochemical cage is sensitive, and the ligand is released in the solution within milliseconds, thereby ensuring rapid mixing and a simultaneous release (Giovannardi et al, 1998).

For SAXS, active mixing techniques have generally been used for studying protein dynamics, using stopped flow and continuous flow methods. In a stopped flow experiment, two solutions are rapidly pushed through a mixing chamber and into an observation chamber (see Figure 1). Here the flow is stopped or flowed at low speed, and measurements are recorded as a function of time. The great advantage of this method for SAXS is that the whole observation chamber is filled with mixed solution and therefore a relatively large beam can be applied and a high signal can be obtained. Meanwhile, stopped flow measurements are limited by the dead time of the instrument detector (Grillo, 2009). If a time resolution below this dead time is desired the mixing procedure must be repeated several times, resulting in a high sample consumption of about 150 μl protein solution of typically 10 mg/ml per mixing (Giladi et al, 2012). However, this is a well-established technique and somehow straight forward to apply (Angelov et al, 2015). Continuous flow mixing relies on mixing by passive diffusion, therefore requiring small diffusion distances for a fast mixing process, which can be achieved using a microfluidic device (Park et al, 2008). Solutions of protein and ligand are injected into a channel where they form a laminar flow with a thin sample sheet of protein solution surrounded by sheets of ligand solution. The sample is mixed as ligand diffuses into the sample sheet, where it induces a structural change that can be followed by collecting scattering data at different positions downstream in the channel (see Figure 1). This is a much less applied technique compared to stopped flow as a more sophisticated sample environment is required and much development is still needed. The major advantage of this method is that very little sample is required, with a sample consumption down to 0.55 μl/min protein solution of typically 10 mg/ml (Brennich et al, 2011). The drawback is that it can be quite challenging to achieve adequate measurement signal from the narrow sample sheet and requires an X-ray beam of comparably small dimensions. Further, only small ligands such as ions can be used to obtain rapid mixing, as their diffusion is fast enough. Figure 1 explains the principle of the two mentioned active mixing techniques.

Fast mixing by turbulence can be applied as well, but high flow rates are needed, which again require large sample volumes of 10-20 ml/min, and at least 2-3 mg protein per data point (Graceffa et al, 2013).

Large sample consumption is a general problem in time resolved SAXS, which is rather unfortunate when working with fragile biological samples. To limit sample consumption microfluidic devises can be applied and short mixing times obtained by simple diffusion (Pollack et al, 1999).
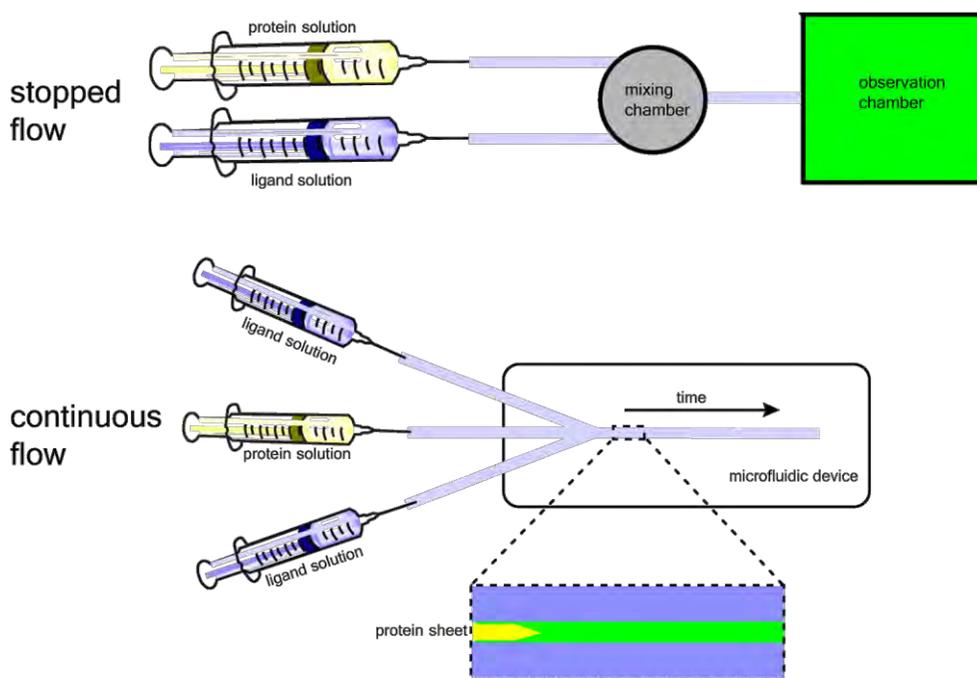
Figure 1. Active mixing techniques. Stopped flow is based on a mixing chamber into which a protein solution and ligand solution are injected and mixed before transferred to an observation chamber. For continuous flow, the solutions are injected into a narrow channel, thereby creating a thin protein sheet surrounded by a ligand solutionthat will diffuse into the protein sheet. Time resolved structural changes can thereby be recorded downstream the channel.

In the present thesis, different methods for studying time resolved protein dynamics with SAXS have been applied, including the following:

- Self-assembly of lipoprotein nanodiscs, found in Project A.
- Structural investigation of magnesium transporter CorA, trapped in its ligand-bound and ligand-free state, respectively, in Project D.
- Probing the structure of soluble proteins in real time using a microfluidic chip, presented in Project F.

Especially the last mentioned project has been a somewhat more technical challenge, than first expected, however, it has now started to show promising results, and it will be very exciting to discover the full potential of this method.

Chapter 2 – The cell membrane

The barrier protecting the cell has been studied for more than 100 years. This barrier consists of a lipid bilayer which was first discovered in 1925 by Gorter and Grendal, who at the same time proposed that the membrane lipids were directed with their polar group to the inside of the cell (Gorter & Grendel, 1925). In 1972, Singer and Nicolson proposed the famous "fluid mosaic model" which has been an inspiration for most of the knowledge we have about the biomembrane architecture today. Their hypothesis was that the membrane consists of a lipid matrix containing many different types of integral and peripheral proteins in copies of more than 10,000 per cell. They also illustrated the membrane as highly protein crowded (Singer & Nicolson, 1972). The presented model was laid out as a homogeneous fluid with proteins diffusing in two dimensions.

The solving of the first structures of membrane proteins led to the proposition of the "mattress model". This refined model considered the hydrophobic and hydrophilic protein-lipid interaction in more details. The interaction was believed to be energy associated with matching amphipathic features in the lipids and protein, meaning mismatch in hydrophobic length of the protein is compensated by local deformation of the lipid bilayer (Mouritsen & Bloom, 1984). Later, evidence was found for the inhomogeneity on lipid composition within the membrane leaflets, so-called lipid rafts, caused by specific lipid-lipid and lipid-protein interaction (Harder et al, 1998). Raft domains are formed in the lipid membrane from clusters of specific membrane components creating a local environment with optimised biophysical properties.

Lipids of the membrane

The nature of biological membranes is very complex, and the membranes are composed by numerous kinds of different lipids, membrane proteins and a small amount of carbohydrates. The distribution of lipid species varies between cell types and also within the different compartments of a cell in order to obtain the unique functionality needed for the membrane and the proteins embedded. Furthermore, the lipid composition in a membrane often has an asymmetric distribution on the two leaflets on the same membrane (Heimburg, 2007a).

Lipids are, in this thesis, defined as amphiphilic molecules with a hydrophilic head group and hydrophobic hydrocarbon chains or tail group with a significant variation in head and tail structure, as described by Heimburg (Heimburg, 2007b). The lipid composition of biomembranes is dominated by different types of phospholipids but does also contain a high number of cholesterols and glycolipids. This membrane composition has a substantial physiological effect on the membrane in regards to environmental changes such as temperature, pH and pressure (Lindblom & Rilfors, 1992).

As mentioned above, the predominant type of lipids in biomembranes is phospholipids, that are often classified by their head group composition, bound to two hydrocarbon chains varying in length and saturation level. The phospholipids found in *E. coli* membranes are mainly phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) in a molar ratio of 73:27 (Sweetman et al, 1996). In mammalian cells, three main types of phospholipids are found in the membranes, namely phosphatidylcholine (PC), PE and phosphatidylserine (PS). The

phospholipid composition of the human erythrocyte membrane, the most studied type of mammalian cells, consists of 77% PC, 20% PE and less than 3% PS; 65-75% of the PC is found in the outer leaflet while the inner contains 80-85% of the PE and more than 96% of the PS (Zachowski, 1993).

Self-organisation of lipids

Lipids do not only form into bilayers in aqueous solutions, they aggregate into various phases to shield their hydrophobic tail groups from water, so only the hydrophilic head groups are exposed. Different types of lipids have different aggregation preferences depending on their specific packing parameter, mentioned later in this chapter. The lipid aggregation is strongly concentration dependent, and the formation of larger aggregates is more favourable at high lipid concentrations.
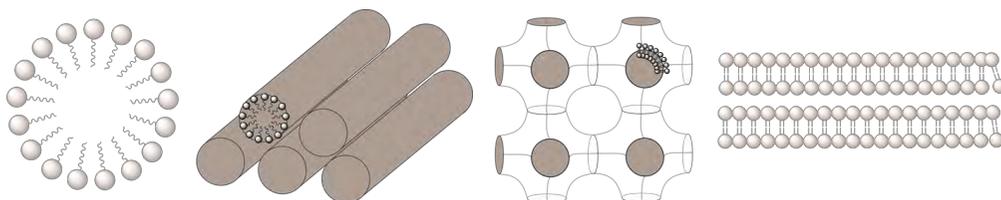
Figure 2. Selection of various lipid phases. From the left, spherical lipid micelle, hexagonal phase consisting of cylindrical micelles, lipid cubic phase, consisting of flexible bilayers aggregating into a cubic structure, and lamellar phase lipid bilayers.

At low concentrations lipids exists as isolated monomers, but above a certain concentration micelles will form. This is called the critical micelle concentration (CMC). The formation of simple spherical micelles may be approximated to simple equilibrium law $K = \frac{[L_n]}{[L]^n}$ where $[L_n]$ is the concentration of lipids consisting of n lipids and $[L]$ the concentration of monomeric lipids (Heimburg, 2007b). Though this is much more complex in reality due to the constant competition between hydrophobic attraction and repulsion within the molecules, and several models to describe micelle aggregation have been proposed (Wennerstrom & Lindman, 1979).

At higher concentration lipids typically grow into more extended lamellar phases, like the earlier mentioned bilayers found in biological membranes. However, some types of lipids prefer to aggregate into non-lamellar phases such as lipid cubic phases or hexagonal phases. These types of lipids are all present in the cell membranes in lower concentrations, which is of great importance for structure and function (Lindblom & Rilfors, 1992). The mentioned lipid phases are illustrated in Figure 2.

Lipid packing parameter

The packing parameter describes the molecule shape and structures that a certain type of lipid will form above CMC. This parameter is defined as $v/a_0 l_c$, where v is the volume of the hydrocarbon chain or chains, a_0 the optimal surface area and l_c the critical chain length that sets a limit for how long the chains can extend. Different forms of packing parameters are illustrated

in Figure 3. Molecules will pack into structures consistent with these geometric constrains. With this said, higher entropy will always favour the smallest aggregation number. However, smaller structures where the surface area is forced above a_0 will be energetically unfavoured (Israelachvili, 2007b; Israelachvili et al, 1976).
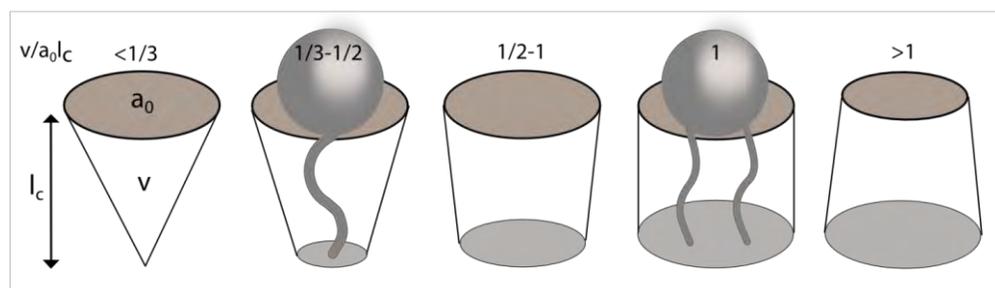
Figure 3. Different lipid shapes and how this relates to the lipid packing parameter v/a_0l_c, v is the volume of the hydrocarbon chain or chains, a_0 the optimal surface area and l_c the critical chain length . The lipid shapes illustrated are from left: cone, forming spherical micelles; truncated cone, forming cylindrical micelles; less truncated cone, forming flexible bilayers; cylinder, forming planar bilayers; and inverted truncated cone, forming inverted micelles.

Membrane proteins

Membrane proteins have been given much attention the last decade especially in relation to pharmaceutical aspects. The reason is their potential as drug targets mainly because of their central roles in cell signalling and transport of nutrients into the cell. Within recent years the importance of this research area has been given much attention as Robert Lefkowitz and Brian Kobilka were awarded the Nobel Prize in chemistry in 2012 for their studies in G-protein coupled receptors (GPCRs) (Lefkowitz & Kobilka, 2012). However, this is not the first time structural membrane protein research has been granted a Nobel price. As early as in 1988 Johan Deisenhofer, Robert Huber and Hartmut Michel received the Nobel Prize in Chemistry "for the determination of the three-dimensional structure of a photosynthetic reaction centre" (Deisenhofer et al, 1988) and several studies have since followed (Agre & MacKinnon, 2003; Axel & Buck, 2004; Carlsson et al, 2000; Gilman & Rodbell, 1994; Neher & Sakmann, 1991; Skou, 1997).

Ten to fourteen percent of the proteins encoded in the human genome are predicted drug targets, meaning they have structural features favouring interaction with drug-like chemical compounds. Half of these drug targets can be categorized as membrane proteins (Hopkins & Groom, 2002; Russ & Lampel, 2005). About 30% of a cell's proteome is membrane proteins with a very small variance in ratio from organism to organism, regardless of proteome size (Stevens & Arkin, 2000). However, considering the large fraction, very little is known about this class of proteins, because of the difficulties in the work with membrane proteins. In the Protein Data Bank (Bernstein et al, 1977), only 2.5% of the total number of protein structures can be classified as membrane proteins. Furthermore, according to the Structural Biology Knowledgebase only 525 unique structures are available (PDB, 2015b; White & Laboratory, 2015).

Membrane proteins are extremely dependent on the lipid membrane, not only because it offers the amphiphilic environment needed to keep the protein in its native fold, but in many cases membrane proteins are also highly dependent on the specific lipids and lipid composition in order to maintain their functionality and activity.

As earlier mentioned, the membrane is highly protein crowded, but differ in the content for different organelles, though the plasma membrane typically contains 50% protein of the total cell mass. Membrane proteins serve an endless number of functions; channels, pores and transporters maintaining the influx and outflow of nutrition, ions and metabolites across the membrane, receptors detecting chemical signals, enzymes catalysing specific reactions and many more.

Proteins can be associated with the membrane in different ways and are in general classified as integral or peripheral membrane proteins. Integral membrane proteins can be transmembrane, extending throughout the membrane with parts of their mass on both sides of the membrane, or membrane associated, located only in the cytosol integrated in the inner leaflet of the membrane. Peripheral membrane proteins are bound to the membrane by polar interactions, present only on one side of the membrane. The binding can be as lipid linked, anchored to lipids by covalent bonds, or protein bound to surfaces of integral proteins (Alberts, 2004; Berg et al, 2012). Different types of membrane proteins are illustrated in Figure 4.
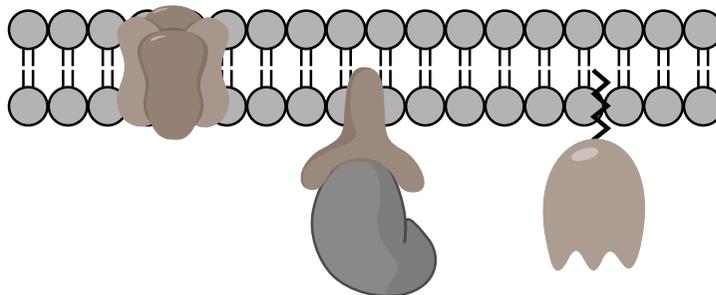
Figure 4. Different types of membrane proteins. From left, a transmembrane integral membrane protein, a membrane associated integral membrane protein with a peripheral protein bound and a peripheral lipid linked membrane protein.

Membrane protein structure

Different classes of integral membrane proteins possess different overall structures; however, there are some common structural features that they all share. The membrane exposed residues are on average more hydrophobic than their interior residues that are opposite to water soluble proteins. The residues exposed to the membrane are mainly Isoleucine, Leucine, Valine and Phenylalanine, which all contain hydrophobic side chains. Transmembrane segments are usually ordered in the form of helixes or beta sheets, as part of beta barrels (Voet & Voet, 2004).

Studying and handling membrane proteins represent major challenges because of their special requirements. The main reason for these difficulties is the amphipathic nature of membrane

proteins. Hydrophobic regions must be scavenged from water for the protein to keep its 3D fold and some proteins have higher restrictions to the type of scavenging molecules than others. Some proteins require interaction with specific lipid types to remain active while others have yet unknown needs to be discovered. This all explains the great challenges associated with working with and studying membrane proteins and is the reason for the many unsolved structures.

Chapter 3 – Methods in protein isolation

Overexpression of proteins

In general, studying proteins in any kind of research requires large quantities of the isolated and purified protein of interest. In some cases, these large quantities of usually 1-1000 mg protein can be obtained from the original host, usually stable, fully characterized proteins. However, in most cases, it is an impossible task to obtain these amounts and it is therefore necessary to produce the protein recombinant. For this purpose *Escherichia coli* (*E. coli*) is by far the most popular and efficient host. 76,125 protein structures out of 98,443 in the Protein Data Bank derive from proteins expressed in *E. coli*, compared to 1,634 from human sources (PDB, 2015a). However, *E. coli* cannot be the go-to-organism in all cases as some proteins need more sophisticated expression routes, such as many eukaryote proteins that need a eukaryotic expression host. Examples of such a host are yeast, insect cells or even mammalian cells. These more advanced hosts are needed when for instance expressing glycoproteins as glycosylation takes place in organelles that are nonexistent in prokaryotes, namely endoplasmic reticulum and golgi (Paulson & Colley, 1989). These higher expression hosts are, unfortunately, much more time consuming, and the yields are lower compared to bacterial expression. For this reason they are also more expensive, and they are only used if it is necessary.

As much of the bacterial proteome remains unexplored in relation to the projects of this thesis, *E. coli* has been the applied expression host. This section will therefore solely focus on using *E. coli* as the expression host.
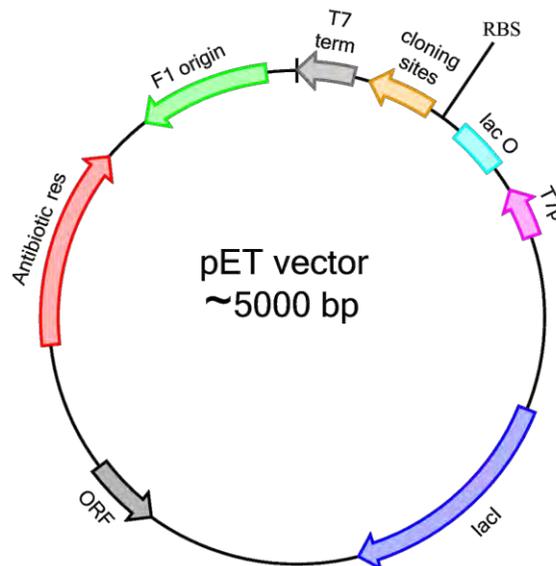
Figure 5. Overall design of a pET vector. The T7 promoter (pink) is followed by the lac operon (turquoise), then a multiple cloning site (orange) for insertion of the expression gene and the T7 terminator (green). The ribosomal binding site (RBS) is found between the lac operon and the expression gene. DNA coding for the lacI protein (blue) and antibiotic resistance (red) is also present as well as an origin of replication (black) and the F1 origin of replication (purple).

In order to utilize *E. coli* as expression host, a plasmid containing the gene for expression is needed. The plasmid is taken up by the host organism, supplementing its existing genome and is copied with every cell division. Most common are the pET vectors; the basic design of such a plasmid is illustrated in Figure 5. An expression vector contains everything needed for producing the protein: an appropriate promoter, a sequence terminating the transcription and a ribosomal binding site. A strong promoter is needed for high level expression, as it has to be able to accumulate protein in amounts up to 10-30% of the total cellular protein (Makrides, 1996).

Promoters are controlled by a regulatory gene present in the plasmid itself or in the host chromosome. pET vectors use a T7 expression system, originating from the T7 bacteriophage, and it is widely used because of its high specificity. In this expression system a T7 promoter controls the transcription of the expression gene. Driving the transcription is the T7 RNA polymerase, expressed from the T7 gene encoded in the host genome. Because of a highly conserved nucleotide sequence, the binding of the T7 RNA polymerase to the T7 promoter is highly selective. The polymerase also has the ability to direct most of the cells resources towards transcription of DNA underlying its promoters (Studier & Moffatt, 1986). However, this high expression activity can create a problem as limited resources will be available for the production of the cell's own proteins, and therefore the T7 polymerase is, in most applied cells, regulated by a lac promoter. This means, only when lactose or a similar molecule is present, the polymerase will be transcribed. However, regulating the gene by the promoter alone has shown to be insufficient leading to a "leaky" expression, which can be very harmful for the cell if toxic proteins are expressed (Mertens et al, 1995). To solve this problem, a lac operon sequence is inserted between the promoter and the gene for the polymerase in the cell genome as well as the expression gene in the expression plasmid. The operon is regulated by a lac repressor protein (Dubendorff & Studier, 1991) encoded in both host and vector DNA for sufficient repression (Glascock & Weickert, 1998). The lac repressor gene also recognizes lactose and in the absence of lactose it will bind the operon blocking access to the promoter. When lactose binds to the lac repressor it falls off the operon leaving access to the promoter. Isopropyl β-D-1-thiogalatopyranoside (IPTG) is most commonly used for mimicking lactose as this molecule cannot be broken down or metabolized by the cell (Hansen et al, 1998).

Other plasmids than pET vectors can be used, such as pBAD vectors. In pBAD vectors, the DNA transcription is tightly controlled by a promoter regulated by arabinose. The advantage of using this vector is very low levels of uninduced expression, which is rather favorable for expression of toxic proteins. Also the level of expression can be highly controlled by the concentration of arabinose added to the medium. However, expression yields cannot match those obtained by using pET vectors (Guzman et al, 1995).

Expression strains

Quite a few *E. coli* strains have been optimised for recombinant protein overexpression, many derived from the K12 strain; which genome was the first *E. coli* strain to be fully sequenced (Blattner et al, 1997). Expression strains should be deficient in harmful proteases and maintain the expression plasmid stable. BL21 and derivatives hereof are the most common strains

because of their robustness and their ability to grow vigorously in minimal media; and they have shown a wide range of applications in standard recombinant expression. BL21(DE3) is the most popular variant for most protein expression. It carries a bacteriophage DE3 DNA fragment, coding for the before mentioned *LacI* gene (Studier et al, 1990).

Despite successful overexpression of thousands of different proteins in BL21(DE3), significant overproduction cannot be achieved for some certain proteins, due to toxicity of the expression protein that consequently kills the host cell. This is a problem often related to overexpression of membrane proteins. Strains optimised for overexpression of membrane proteins have been developed, including the so-called "Walker strains". The strains C41(DE3) and C43(DE3), both derived from BL21(DE3), have been optimised for overexpression of membrane proteins by random mutation selection (Miroux & Walker, 1996). These strains have shown resistance to toxicity associated with overexpression of several different membrane proteins. Also instability of certain plasmids observed in BL21(DE3) can in many cases be overcome or minimized by C41(DE3) and C43(DE3) (Dumon-Seignovert et al, 2004).

A different approach to decrease plasmid instability is to introduce a plasmid coding for T7 lysozyme, referred to as pLysS. T7 lysozyme inhibits the T7 RNA polymerase and the small amount of pLysS provides a reduction of the basal activity of the polymerase, ensuring that expression of the target protein will not occur prior to induction (Studier et al, 1990).

Protein Purification

As mentioned earlier, in order to study proteins, structurally or functionally, a large quantity of purified protein is usually needed. After successful overexpression of a protein in cells, the protein needs to be isolated from the cell material with a reasonable high purity, depending on the experiment.

Affinity tags

For most proteins, introducing an affinity tag is an efficient and easy strategy, as this will allow purification using generalized protocols. Many fusion tags are available. Among popular tags are His, FLAG, Strep II and GST. The choice of tag highly depends on priorities among quantity, purity and cost (Lichty et al, 2005). Polyhistidine (His) tag is frequently used because of its low cost and high yields. His tags consist of 4-10 histidine repeats fused to the N- or C-terminal of the protein, which allows purification by one simple chromatographic step. Including this histidine repeat to the protein of interest, will allow formation of a strong metal complex with Ni. These may in turn be immobilized, once solid supported, such as a metal chelate absorbent charged with nickel ions (Ni^{+2}-NTA) (Hochuli et al, 1988), and protein can be purified by so called metal ion affinity chromatography (IMAC), as shown in Figure 6. Protein can be released from Ni^{+2}-NTA either at denaturing conditions by lowering pH or at native conditions by adding imidazole, a molecule having the same structure as the histidine side chain. At higher concentrations, imidazole efficiently competes with the binding of Ni-NTA and elutes the bound fusion protein (Janknecht et al, 1991).
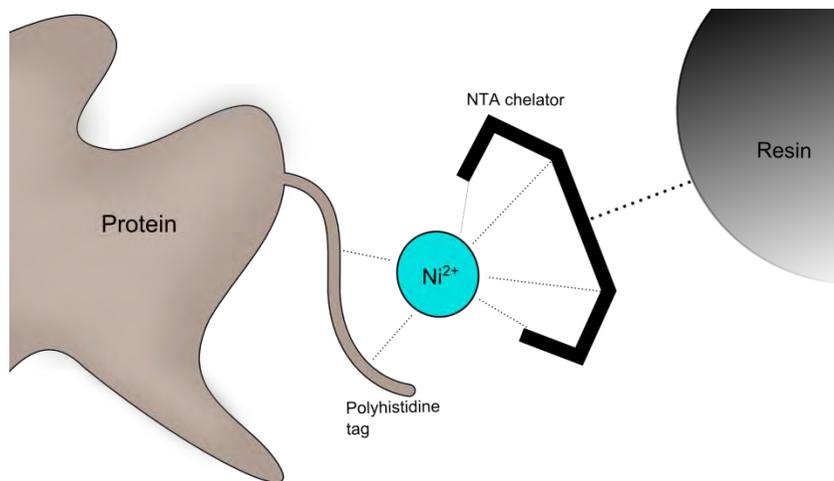
Figure 6. Metal ion affinity chromatography (IMAC) by Ni^{2+}-NTA. Proteins containing a histidine tag can be bound to Ni^{2+} ions. NTA, bound to resin, complex Ni and keeps it coordinated in such a way that it strongly binds histidine and thereby makes it possible to hold back a target protein for purification.

Usually the affinity tag has no effect on the protein but in some cases the tag needs to be removed from the protein. This may be done by enzymatic cleavage using site specific proteinases which recognize a specific amino acid sequence; that is introduced between the protein sequence and the affinity tag. Trypsin, Trombin and Tobacco etch virus (TEV) protease are commonly used proteases. TEV is particular popular because of its high specificity, preventing cleavage at cryptic sites within the target protein (Polayes et al, 1998). As the name states, the 49 kDa TEV proteinase originates from the tobacco etch virus, a plant potyvirus; this proteinase naturally catalyzes the cleavage of TEV polyproteins. It recognizes the amino acid sequence Glu-Xaa-Xaa-Tyr-Xaa-Gln-Ser/Gly and cleaves between Gln and Ser/Gly (Carrington & Dougherty, 1988; Dougherty et al, 1989). TEV protease can be produced recombinant in *E. coli* with high activity containing a histidine tag and is thereby easily removed from the target protein by the use of Ni^{+2}-NTA after cleavage (Parks et al, 1994).

In many cases, a single Ni^{+2}-NTA purification step is not sufficient to achieve the required protein purity. Here a second Ni^{+2}-NTA purification step may be introduced or obtained by introducing a different chromatographic method, such as size exclusion chromatography (SEC).

Size exclusion chromatography

SEC or gel filtration is a commonly used method in protein purification and characterization. The chromatographic method is applied for separating water soluble molecules, typically proteins or larger polymers, according to their hydrodynamic radius. The stationary phase, i.e. the column, consists of spherical porous particles with a controlled pore size. Through this stationary phase molecules diffuse based on their size difference, using an aqueous buffer as mobile phase. In general, the molecules do not react with the spherical particles in the column, but diffuse through the pores of the particles. Smaller particles diffuse into more pores than larger particles and will for that reason have a longer retention time than larger particles. As

most of the size separation occurs within the particle pores, the highest resolution is obtained at large pore volumes (Bouvier & Koza, 2014; Fekete et al, 2014).

Various column matrixes are available to obtain a high separation within the relevant protein sizes.

Solubilisation of membrane proteins

There are many difficulties in studying membrane proteins due to their very special amphipathic nature. Most techniques used for studying proteins are developed for soluble proteins and must be further modified if applied for membrane proteins. Despite great efforts, the handling of these proteins still presents a major problem to researchers. For scientists to apply the methods available for soluble protein characterization and analysis, the membrane proteins ideally have to be kept in solution. Solubilisation of membrane proteins is generally achieved with detergents; nevertheless, detergents offer a somewhat harsher environment than lipids. Evidently, this may have an unfortunate impact on the structure and function of membrane proteins. Low stability, inactivation and aggregation are often observed over time for membrane proteins kept in detergent solution (more about detergents in Project C) (Bowie, 2001). More ideal conditions for working with membrane proteins are met by the application of a lipid bilayer, and here the so-called nanodisc can offer close to natural conditions. Nanodiscs are disc shaped lipoprotein particles based on high density lipoprotein (HDL) particles, holding a lipid bilayer with a diameter around 10 nm. Engineered for the purpose of stabilising membrane proteins, the nanodisc provides a native like membrane environment with much higher particle homogeneity than other bilayer particles, such as liposomes and micelles (Bayburt et al, 2002). In addition to nanodiscs, also bicelles and peptide discs are somewhat monodisperse particles suitable for solubilisation of membrane proteins, as illustrated in Figure 7.
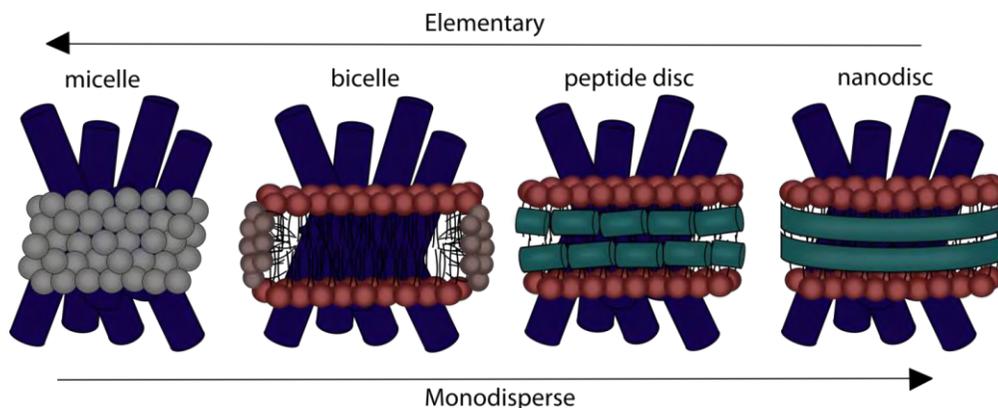
Figure 7. Solubilisation of membrane proteins. Monodispersity and stability of particles increases from left to right, where micelles are more polydisperse than bicelles etc. In opposite direction, from right to left, elementary, where micelles are more straightforward to use than bicelles etc.

These particles are more straightforward to work with than nanodiscs, but do also hold a more polydisperse structure. The forthcoming section will describe the nanodisc in detail following a brief comparison to these other particle types in Figure 7.

From high density lipoprotein to nanodisc

HDL is a complex particle consisting of different proteins and lipids and is responsible for the so-called reverse cholesterol transport in the human body (Glomset, 1968).

Apolipoprotein A-1 (ApoA-1) is the major protein player in HDL and in association with cholesterol and phospholipids pre-β HDL is formed (Kawano et al, 1993). ApoA-1 is usually present in two to four copies per HDL particle, depending on the particle state. ApoA-1 forms various stable structures with lipids, where pre-β HDL is an early state, discoidal, lipid poor particle. This particle is able to interact with plasma lecithin cholesterol (LCAT) to form spherical and more lipid rich HDL which then can be transported to the liver (Jonas et al, 1990).

Human ApoA-1 is a 30 kDa large protein consisting of 267 amino acids produced and excreted from the liver. In its lipid free form, it structurally consists of two helical domains forming a half circle with a height of ~17 Å and a diameter of ~110 Å; its crystal structure is shown in Figure 8. The N-terminal domain of ApoA-1 forms an α-helix followed by an extended strand connected to a second short helix. This short helix connects to a third long helix and a fourth helix formed at the C-terminal. The N-terminal helixes are crucial for maintaining its lipid free structure as they scavenge hydrophobic amino acids in the molecule (Mei & Atkinson, 2011).

Figure 8. Crystal structure of human Apolipoprotein A-1. The structure consists of two helical domains forming a half circle with a height of ~17 Å and a diameter of ~110 Å [pdb entry 3R2P].

The ability of ApoA-1 to form discoidal particles in the presence of lipids caught the interest of Sligar and coworkers, who started investigating the structure of reconstituted HDL (rHDL) from purified ApoA-1 and synthetic phospholipids, by imaging (Carlson et al, 1997). They suggested that these rHDLs could be used as a self-assembling phospholipid bilayer support for trapping molecules of interest. Furthermore, they managed to incorporate the integral membrane protein P450 reductase (Bayburt et al, 1998).

In order to optimise the particle they produced a truncated version of ApoA-1 that lacked the globular N-terminal domain. This modified protein, also called membrane scaffold protein (MSP), is also able to self-assemble into discoidal monodisperse particles in the presence of synthetic phospholipids (Bayburt et al, 2002).

200 amino acid long amphipathic MSPs form together with phospholipids into 10 nm diameter nanodiscs. Each nanodisc consists thereby of two MSPs surrounding a phospholipid bilayer. A molecular dynamics simulation of such a nanodisc is shown in Figure 9.

Moreover, MSP has been optimised for expression in *E.coli;* the protein is purified by simple affinity chromatography and can easily be obtained in large quantities.
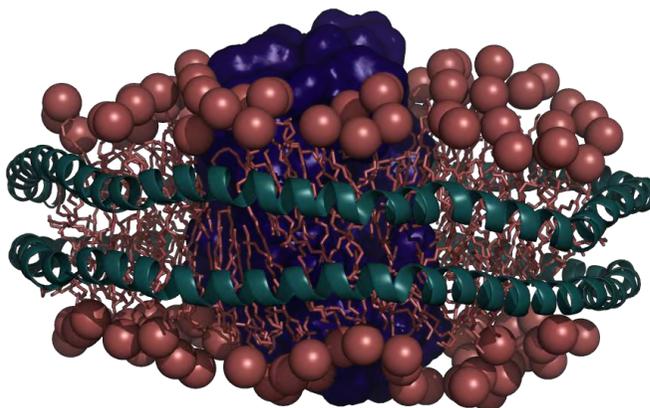
Figure 9. **Simulated nanodisc structure. Two MSPs (green) surround a lipid bilayer, forming the nanodisc in which a membrane protein is embedded (blue). The illustrated structure is a combined molecular dynamics simulation of a nanodisc (Shih et al, 2005) and an NMR structure of proteorhodopsin [pdb ID 2L6X].**

Self-assembly of nanodiscs can be obtained in a few simple steps: phospholipids self-assemble into mixed micelles when dissolved in detergent solution and mixed with a proper ratio of MSP. The detergent is removed by incubation with non-polar polystyrene beads. Subsequently, homogeneous nanodiscs are purified from the reconstitution mixture by gel filtration (Ritchie et al, 2009). Reconstitution of membrane proteins into nanodiscs basically follows the same procedure; this is illustrated schematically in Figure 10. Various lengths of MSPs have been developed in order to form different sizes of nanodiscs, and lipid composition may be varied to provide the optimal conditions for the relevant membrane protein (Denisov et al, 2004).
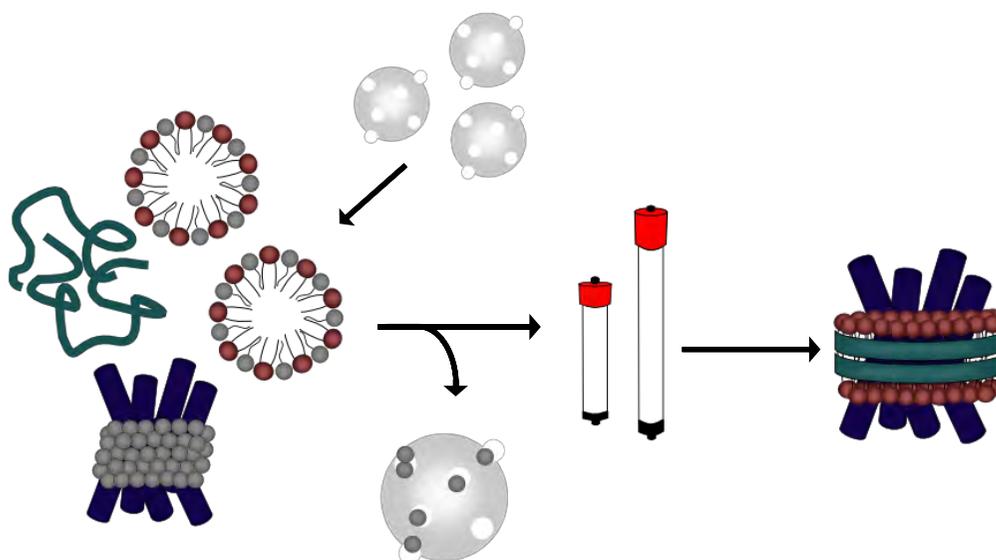
Figure 10. **Reconstitution of membrane proteins into nanodiscs. Phospholipids solubilised in a detergent solution are mixed in appropriate ratios with MSP and membrane proteins, ideally containing an affinity tag, in detergent solution. Detergents are removed by polystyrene beads, and nanodiscs are formed. The sample is then passed through an affinity column to separate membrane protein loaded nanodiscs from discs without membrane protein. The loaded discs are purified from any kind of aggregate by gel filtration.**

In several studies, the nanodisc has been applied as a stabilising sample holder for membrane proteins, primarily in activity and functional studies, leading to a higher or longer activity of the protein comparted to detergent micelles (Bayburt & Sligar, 2010). A broad selection of these can be found on Sligar Labs webpage (Sligar, 2015). On the other hand, structural studies have shown to be far more challenging than expected and this issue will strongly influence the present thesis. However, by the application of small angle scattering it is now possible to determine low resolution structures of membrane proteins employing nanodiscs as sample holders (Kynde et al, 2014). More details on this issue are given in the following section on small angle scattering.

Bicelles and peptide discs

Bicelles are discoidal lipid aggregates formed from long chain phospholipids and detergent or short chain phospholipids. The long chain phospholipids form a lipid bilayer while the short chain lipids or detergents form a rim shielding the long chain lipid tails, as shown in Figure 7. The size of bicelles is controlled by the short chain to long chain phospholipid ratio and the total lipid concentration (Whiles et al, 2002). Bicelles have been successfully used for structure NMR studies of membrane proteins, for review see (Duerr et al, 2012), and have also proven to be useful in crystallization (Rasmussen et al, 2007). Though, nanodiscs have shown to provide a more stable environment for membrane proteins compared to bicelles (Lyukmanova et al, 2012).

Peptide discs are closely related to nanodiscs, containing a lipid bilayer, but are held together by peptides with the affinity to line up or self-assemble into disc shaped particles, as seen in Figure

7. A type of peptides called 18A is synthesized from the sequence of a single α-helix from ApoA-1. Using these peptides overcome the reservations that must be taken when working with proteins. These discs show promising features for stabilising membrane proteins but does also appear to change their stoichiometry over time (Midtgaard et al, 2014). This is an unfavorable feature for small angle scattering studies of membrane proteins as homogeneous samples are desired

Chapter 4 - Small angle scattering from biological systems

The following chapter describes basic theory and instrumentation of small angle scattering and is based on material written by Svergun (Koch et al, 2003; Svergun & Koch, 2003), Als-Nielsen & McMorrow (Als-Nielsen & McMorrow, 2011) and Glatter & Kratky (Glatter & Kratky, 1982).

Small angle scattering (SAS) is a method that may be used for determining particle size and shape of structures at nm length scales. This gives the possibility to study the structure of biological systems in solution and thereby close to their native-like state at physiological conditions. It also makes it possible to study structural changes in response to environmental variations in e.g. temperature, pH or salinity. Two types of radiation sources are used in SAS, namely neutrons (SANS) and X-rays (SAXS).

X-rays are electromagnetic waves. The wavelength λ typically applied in SAXS is around 1 Å setting the lower limit for the resolution that can be obtained through scattering-based techniques like SAXS. Neutrons are uncharged particles released from the atomic nuclei, travelling with velocities corresponding wavelengths in the same region as X-rays and with a half-life of about 10 minutes (Christensen et al, 1967).

The interaction with matter is very different for photons and neutrons. While X-ray photons will interact with the electrons of the atoms, neutrons interact with the nuclei of the atoms. When this interaction occurs, a fraction of the photons or neutrons will be scattered. The ability of an atom to scatter X-rays and neutrons is expressed by their scattering length. Because the scattering length of all electrons is the same, the atomic scattering length for X-rays is given by $f_x = N_e r_0$, where N_e is the number of electrons in the atom and r_0 is the scattering length of an electron, given by the Thomson radius (2.28×10^{-13} cm). This implies that the X-ray scattering length of an atom increases linearly with the number of electrons, and therefore with the atomic number. For neutrons it is far more complicated as the scattering length for the different isotopes varies more and is not easily predicted. For most biological relevant atoms, such as carbon with a scattering length of 6.651×10^{-13} cm, oxygen with 5.803×10^{-13} cm and phosphorous with 5.130×10^{-13} cm, the scattering lengths are similar except from hydrogen with -3.739×10^{-13} cm, that is, a negative scattering length. A commonly used approach is to exchange selected hydrogen atoms in the scattering molecule to its heavier isotope deuterium having a scattering length of 6.671×10^{-13} cm, which interestingly changes the scattering length significantly and thereby makes a huge difference in the visibility of different components in the scattering pattern (NIST, 2015).

Basic SAS theory

SAS is one of many scattering based techniques. The resolution of SAS is typically in the range of 1-100 nm, depending on the scattering angles that data is collected for, typically 0.1° to 10°. In SAS the features at larger scattering angles contain information on the smaller details of the scattering particles, and the smaller angels contain information on the larger structures. It is

hereby a combination of the measured angles and the wavelength which determines the resolution in SAS

It is possible to extend the resolution by decreasing or increasing the measured angle, but both ways will need special setups. Above about 10°, also called wide angle scattering, the distance to the detector must be decreased significantly from a classic setup. Here the features in the scattering pattern are typically very weak. Measuring at small angles (ultra-small angle scattering) requires increased distance to the detector and a very well-collimated beam. However, these two methods have not been applied in this project.

Basic instrumentation

The basic construction of a modern synchrotron based BioSAXS instrument is summarized in Figure 11. It typically holds a tunable double crystal monochromator, for selecting an appropriate wavelength, followed by multilayer mirrors that focus the beam. Multiple slits, made from highly absorbing material, are placed in several positions along the instrument to reduce parasitic scattering around the beam and to narrow the beam size if necessary. The sample cell is typically a capillary made from boron or quartz glass and at most beamlines made as temperature controlled flow cells, keeping the sample at controlled conditions and lowering beam damage. The scattering pattern is typically collected by a solid state pixel detector with a rapid readout, protected from the direct beam by a beam stop and with a photodiode monitoring the direct photon flux.
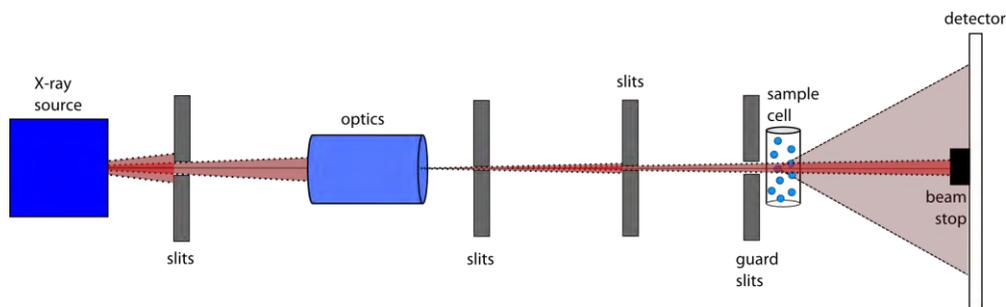
Figure 11. Basic SAXS instrument setup. An X-ray source provides photons for the instrument; the photons are reduced to a monochromatic beam by a monochromator and then focused and sorted for harmonics by a mirror in the instrument optics. Slits are placed along the instrument to reduce beam dispersion and scattering around the beam. The beam passes a sample cell and the scattering of the beam by the sample is detected by a pixel detector with a beam stop protecting the detector from the direct beam. Take notice that the different features are simplified and not drawn to scale.

SANS requires a high flux of neutrons that can only be obtained at large scale facilities with a reactor or spallation source. For SAXS an X-ray source is needed; at large scale facilities, synchrotrons deliver high photon flux to the instruments in the order of 10^{12} photons/s (ESRF, 2015), while smaller "home sources" rely on less powerful sources such as the rotating anodes, which can create a flux up to 10^8 photons/s or the newer liquid metal jet sources delivering up to 10^9 photons/s (Pedersen, 2015). The significant lower flux from "home sources", compared to synchrotrons, means that substantial longer measurement times are necessary.

For biological samples that are often very delicate, a high flux is crucial as low measuring time is needed due to the fact that these kind of samples are unstable over time. High quality data from a weakly scattering sample at a synchrotron beamline can be obtained within a few seconds while it may require hours of measurement time at a home source instrument. The high flux obtained at synchrotron facilities also allows smaller beam size leading to reduction in sample consumption that is often a limiting technicality when working with biological samples.

Scattering geometry

The basic principle behind a SAXS experiment is described in the following. An incoming beam, with a wave vector $\overrightarrow{k_0} = \frac{2\pi}{\lambda}$, that has passed a collimation section illuminates a sample solution containing particles with random orientation. This scatters the beam with an angle of 2θ in relation to the incoming beam. The scattered beam, with the wave vector $\overrightarrow{k_1}$, is then detected by a 2D detector and represented by the scattering vector $\vec{q}$ which is related to the momentum transfer between $\overrightarrow{k_0}$ and $\overrightarrow{k_1}$, as illustrated in Figure 12. The detector records the number of photons received in a given time in such a way that the detector image is built from the scattering intensity I as a function of $\vec{q}$, $I(q)$, containing all scattering information in reciprocal space from the investigated object.
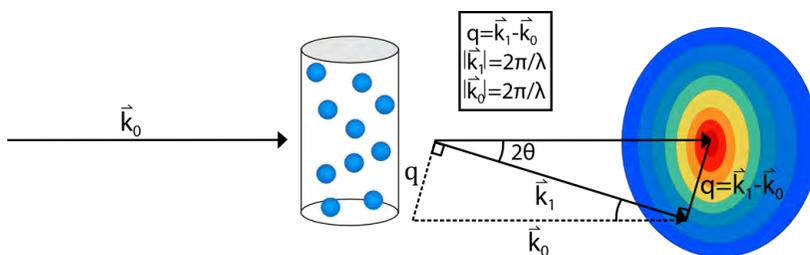
Figure 12. Basic setup of a SAXS experiment. An incoming beam with wave vector $\overrightarrow{k_0}$ illuminates a sample that scatters the beam with an angle 2θ. The scattered beam with wave vector $\overrightarrow{k_1}$ is detected by a 2D detector. The difference between $\overrightarrow{k_1}$ and $\overrightarrow{k_0}$ is described by the scattering vector $\vec{q}$.

When a monochromatic beam illuminates an object the atoms will interact with the radiation giving rise to scattering in spherical waves. Given the relatively low energy of X-rays, used in SAXS, only elastic scattering (Thomson scattering) is considered. This means that no energy is transferred when the incoming beam interacts with the object, therefore the wavelength and frequency remain unchanged for the scattered beam, hence $|\overrightarrow{k_0}| = |\overrightarrow{k_1}|$.

Due to the geometry of a SAXS experimental setup, as shown in Figure 12, the scattering vector is described trigonometrically with the sine relation; $sin\theta = \frac{q/2}{2\pi/\lambda} \Leftrightarrow q = \left|\frac{4\pi sin\theta}{\lambda}\right|$.

In order to be "visible" in SAS, a molecule must have a non-zero scattering contrast compared to the surrounding solvent, given by the excess scattering length density relative to solvent. As X-rays interact with electrons this means that a molecule must hold a different electron density than the surrounding solution in SAXS measurements. In SANS, the atoms of the molecule need to hold different scattering lengths densities than the surrounding solvent.

For SAS a sample of relatively monodisperse particles is preferred as a large sample volume is investigated and an average scattering pattern of all structure parameters present is obtained. Therefore, the presence of a wide variety of particles will result in a scattering pattern where the features of the individual species are not obvious. However, for a monodisperse sample the features are clearly visible and can be associated with particle structure. The detected scattering density is also an average of all orientations of the object, which is associated with observation of isotropic, circular scattering patterns and a high loss in scattering information.

Due to the fact that the scattering intensity does not depend on the direction of the scattering vector $\vec{q}$, but only its length, scattering data are presented as intensity plotted against the length of the scattering vector $\vec{q}$.

Preliminary data analysis - Indirect Fourier transformation

By indirect Fourier transformation of the data, the pair distance distribution function $p(r)$ of the scattering particles can be obtained. It contains, in principle, the same information as $I(q)$, but in real space and is effectively a histogram representing the distances found in the particle; this means that most frequently occurring distances will appear as maxima. It follows that characteristic particle shapes, such as spheres, dumbbells, or elongated particles, can be identified as these will result in very distinct $p(r)$ functions, as illustrated in Figure 13. The distances are represented from 0 to maximum distance, D_{max}, defining the maximum distance within the particle. The distances are weighted by contrast in each of the end-points meaning that $p(r)$ functions do only apply for particles merely consisting of molecules with a contrast significantly different from that of the solvent, such as amino acids. Negative values in the pair distribution function can appear for particles built from multi-contrast molecules. If the contrast in one region is negative, that is with a lower scattering length density than the solvent, such as for X-ray scattering from the hydrocarbon tails of phospholipids, which has an electron density lower than water.
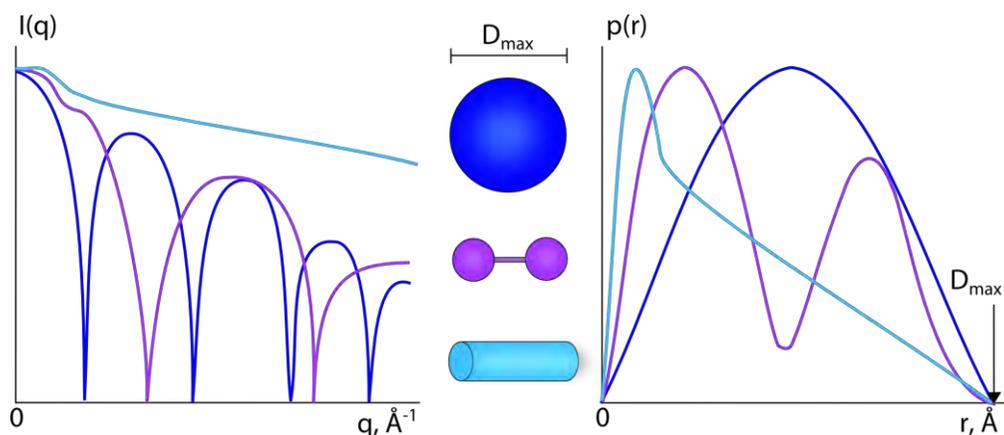
Figure 13. *I(q)* vs *p(r)* in relation to characteristic shapes. Key features for a few characteristic shapes outlined in a scattering intensity plot *I(q)* (left) on a log-lin scale and in its corresponding *p(r)*-function (right) in a lin-lin scale. *D$_{max}$* is the largest distance for all three particles.

29

The pair distance distribution function can be applied as a quick guide for describing sample quality. The reason for this is that aggregation and polydispersity are easily detected, which shows as a larger D_{max} than expected and a change in shape of the characteristic $p(r)$ function, respectively.

Preliminary data analysis - Radius of gyration

The radius of gyration R_g of a particle is the root mean square distance from the end-points in the particle to the center of gravity of the particle, and is determined by using the Guinier approximation, which for low q values applies $I(q) = I(0) \exp\left(-\frac{q^2 R_g^2}{3}\right)$. That is, the radius of gyration can be obtained from monodisperse particles by plotting $\ln I(q)$ as a function of q^2, a so-called Guinier plot, which a low q results in a linear dependency with a slope equal to $-R_g^2/3$ and intercepting in $I(0)$. The linearity of the Guinier plot and the obtained value for R_g may be used as a test for sample quality, in this case sample homogeneity, as well as the pair-distance distribution function. However, as the obtained radius of gyration is also weighted by the contrast of the individual points, this method is less straightforward to interpret for multi contrast particles. Though, it does still contains valuable information about structure, such as particle-particle interaction (Whitten et al, 2007).

Advanced data analysis

A more advanced SAS data analysis approach, frequently used for protein structures is, *ab-initio* methods for shape determination. Using these methods, it is possible to recover three dimensional structures from a one dimensional scattering pattern of a particle solution. One of the most popular models for protein structure determination is DAMMIF, which builds a dummy atom model from the scattering data. First, the model forms a sphere of a radius equal to that of the D_{max} of the particle (obtained from indirect Fourier transformation) with small dummy beads. Subsequently one dummy is moved at the time and a theoretical scattering pattern is calculated for this newly formed structure. All dummies must be interconnected to form a single body. Dummies are randomly selected and moved around until a satisfactory fit between the calculated and measured scattering pattern is obtained (Franke & Svergun, 2009).

As *ab-initio* methods are based on a two phase output, solvent or particle, it is not suitable for multi contrast particles, such as nanodiscs. To obtain structural information from the scattering patterns from these objects, a geometrical model describing the nanodisc has been developed in the Structural Biophysiscs group at the Niels Bohr Institute. Briefly, this model describes the nanodisc as a cylinder with an elliptical cross section built from a combination of form factors fitted to the data. The phospholipid interior of the cylinder is divided into a stack of cylinders with different heights and contrasts, representing head group regions, alkyl chains and methyl end groups. The MSPs surrounding the phospholipids is represented by a hollow cylinder, and when combined with the stack of cylinders representing the phospholipids, the nanodisc model is built. Detailed information about the model can be found in (Skar-Gislinge & Arleth, 2011).

To describe a membrane protein in a nanodisc, the geometrical model for the nanodisc may be combined with a dummy bead model that represents an embedded membrane protein and is

fitted to the data. This so-called hybrid approach has been used to model bacteriorhodopsin (bR) in nanodiscs, built on the nanodisc model with a few modifications, such as introducing curvature in the disc, and a dummy model built from a bR pdb structure, in which each bead corresponds to one residue. For a detailed description of this model, see (Kynde et al, 2014).

Synchrotron facilities

For the projects presented in this thesis, experiments at multiple large-scale X-ray facilities have been conducted to obtain the required high quality data. Facts about the different facilities and beamlines are listed in Table 1; however, the specific settings for each experiment will be stated in the relevant method sections.

Table 1. The different SAXS beamlines where data has been collected. Important specifications for each beamline are listed, obtained from (APS, 2015; Diamond, 2015; EMBL, 2015; ESRF, 2015; MaxLab, 2015; PSI, 2015; Soleil, 2015). Brackets in the energy range column states the applied energies in the experiments described in this thesis.

Facility	Beamline	Energy range [keV]	q-range [Å^{-1}]	beam size [μm]	max flux [photons/s]	Detector	Special feature
ESRF	BM29	0.7-15 (12.5)	0.0025-0.5	700x700	10^{12}	Pilatus M1	BioSAXS
PETRA III	P12	4-12.5 (12.5)	0.003-0.5	200x120	10^{13}	Pilatus M2	BioSAXS
MaxLab	I911-4	13.5	0.01-0.5	300x300	5×10^{10}	Pilatus M1	Location
Soliel	SWING	5-17 (12)	0.006-0.7	450x20	8×10^{12}	AVIEX*	SECSAXS
Diamond	B21	6-23 (12.4)	0.0025-0.4	250x250	10^{11}	Pilatus 2M	BioSAXS
SLS	C-SAXS	4-19	0.005-0.5	20x5	10^{12}	Pilatus 2M	Microbeam
APS	BioCAT	3.5-35 (12)	0.011-0.5	20x5	2×10^{13}	Pilatus 100k	Microbeam

*Gas detector

Pros and cons based on personal experiences

BM29 and I911-4 have been the most used facilities; **BM29** because of its excellent bioSAXS setup specifically engineered to handle small biological samples. With its cooled sample environment, fully automated sample changer and flowed sample capillary, it is possible to obtain high quality data of hundreds of biological samples within a single shift (8 hours) of beamtime. Also the advantage of the neighboring neutron facility (ILL) has made this beamline ideal for combined SAXS/SANS experiments.

I911-4 has been much used for less complicated samples and for prescreening of more complicated samples prior to other beamtime experiments. The facility is a multipurpose setup and not optimised for biological samples, though some add-ons are being developed for this purpose. The photon flux is considerably lower than for the other mentioned beamlines, but because of the favorable location in relation to the Niels Bohr Institute this beamline has been visited many times.

P12 does as well as BM29 have an excellent bioSAXS setup with a fully automated cooled sample environment. A sophisticated setup at this beamline allows measurements at several different detector distances. Unfortunately I have only had few opportunities to visit this beamline.

The **SWING** beamline has been pioneering the SEC-SAXS approach, though this is by now also available on most of the other bioSAXS beamlines. However, these beamlines still lack the experience necessary for optimising the setup as it is at the SWING beamline. Using the online gel filtration, a more monodisperse sample can be obtained compared to a normal setup. The drawbacks of this beamline are its old and easily overloaded detector and unstable sample changer for non SEC samples.

B21 is a recently established bioSAXS beamline, much inspired by P12. Consolidation and few optimisations are still necessary, but the facility will most likely become a high-end bioSAXS beamline in the near future.

CSAXS and **BioCAT** both have the advantage of a micro focus beam that permits more specialized sample environments, such as microfluidic devices. Though, the BioCAT staff is much more experienced and the setup optimised for these specific types of experiments. Both beamlines are less user friendly in terms of software and automated data analysis, compared to the above mentioned beamlines.

All the beamlines visited in this project, regardless of specialization, have delivered great support thanks to enthusiastic beamline scientists. Especially regarding the BioSAXS beamlines, it has been a pleasure for me to follow due to the optimisations that were made for users within the short time frame of my PhD.

Part two – Results

Chapter 5 - Understanding the reconstitution process of nanodiscs

This chapter covers three projects aimed to understand different aspects of the self-assembly of nanodiscs. The following projects will be presented in individual reports:

A. The self-assembly of nanodiscs – time resolved
B. Optimising reconstitution of membrane proteins in nanodiscs
C. Investigating aggregation of biomolecules in D_2O

Project A structurally elucidates the time dependent self-assembly process of nanodisc. This project was a collaboration with Grethe Vestergaard Jensen who performed the data analysis.
Project B concerns reconstitution of nanodiscs with membrane proteins for sample optimisation and identification of important factors in this process.
Project C tests variations in sample preparation of nanodiscs to decrease and ideally remove sample aggregation observed in D_2O based buffers. This project was carried out in collaboration with Søren Roi Midtgaard.
All three projects are based on structural characterization by SAXS.

A. The self-assembly of nanodiscs – time resolved

Abstract

Lipoprotein nanodiscs self-assemble from membrane scaffold protein and phospholipid-sodium cholate micelles, when cholate is removed. The structural development of this assembly has been exploited. Following the assembly over time using small angle X-ray scattering (SAXS) it has been possible to obtain information on the assembly process by fitting a structural model to the data.

Introduction

The nanodisc is a disc-shaped phospholipid bilayer surrounded by two α-helical amphipathic membrane scaffold proteins (MSPs) derived from human Apo lipoprotein A-1. MSPs keep the nanodisc in its defined shape and size. With its lipid bilayer, its diameter of approximately 10 nm, and monodispersity, the nanodisc serves as an excellent tool for stabilising and studying membrane proteins (Bayburt et al, 2002; Bayburt & Sligar, 2010; Denisov et al, 2004).

The self-assembly of nanodiscs is carried out in a few simple steps. Phospholipids dissolved in detergent solution, forming mixed micelles, are added a proper amount of MSP. The detergent can subsequently be removed by incubation with non-polar polystyrene beads leading to formation of nanodiscs (Ritchie et al, 2009).

Macroporous divinyl benzene cross-linked polystyrene beads such as Bio-Beads or Amberlites XAD have high surface area for absorbing organic materials from aqueous solutions. They can be used to remove detergents of both high and low critical micelle concentrations (CMC) (Rigaud et al, 1998).

Removal of the detergent molecules from the MSP-lipid solution results in self-assembly of nanodiscs; a kinetic process which has not before been structurally characterized. Using SAXS we have been able to describe the assembly process to a certain extend by the use of a structural model, optimised to fit the data. The model includes a combination of nanodiscs and micelles consisting of cholate and lipids.

Methods

Sample preparation

MSP1D1 was expressed and purified as earlier described (Ritchie et al, 2009) and measured by SAXS.

1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) (Avanti Polar Lipids) was solubilised in Tris buffer (20 mM Tris/HCl pH 7.5, 0.1 M NaCl) containing 0.1 M sodium cholate, to a concentration of 50 mM. POPC was mixed with MSP1D1 in Tris buffer in a 65:1 molar ratio giving a final concentration of 11.8 mM POPC, 182 μM MSP1D1 and 23.5 mM sodium cholate. Immediately after mixing, a 30 μl sample was taken from the mixture and measured by SAXS. This was repeated 20 and 60 min after mixture.

The mixture was then added Amberlite XAD-2 (SUPELCO) ~0.5 ml/ml sample and 5 min after mixing, a 30 μl sample was taken from the mixture, filtered through a 0.22 μm syringe filter and measured by SAXS. This was repeated after 10, 16, 24, 39, 55, 78, 112 and 137 min.

Additionally, POPC was solubilised in Tris buffer (20 mM Tris/HCl pH 7.5, 0.1 M NaCl) containing 0.1 M sodium cholate. The mixture was diluted with Tris buffer to a final concentration of 11.8 mM POPC, and 23.5 mM sodium cholate. A 30 μl sample was taken from the mixture, and measured by SAXS. This was repeated 20 and 60 min after mixture. Amberlite XAD-2 was added ~0.5 ml/ml sample and left for incubation for 5 min before filtered through a 0.22 μm syringe filter and measured by SAXS. This was repeated after 10, 15, 25, 40, 55, 80, 115 and 140 minutes.

SAXS

The SAXS-data were recorded at the I911-4 beamline at MaxLab (Labrador et al, 2012). The SAXS data were converted to scattering intensity units as a function of the length of the scattering vector $\vec{q}$, $I(q)$, where $q = \left| \frac{4\pi sin\theta}{\lambda} \right|$, 2θ is the scattering angle and λ is the X-ray wavelength. Samples were exposed for 120 s using a nominal X-ray wavelength of 0.91 Å. Experiments were conducted at 10°C and data detected using a PILATUS 1M-detector located 1.98 m from the sample. Data reduction and treatment were performed using the BioXtas raw software (Nielsen et al, 2009).

Data analysis

Data analysis was handled by Grethe Vestergaard Jensen from the structural biophysics group at the Niels Bohr Institute. Data for the MSP alone could be described by a model of aggregated ellipsoidal particles. At $t = 0$, the data were described by the same ellipsoidal particles (now not aggregated) and core-shell particles, introducing a contribution from phospholipid-cholate micelles. Data sets collected after addition of Amberlites ($t > 0$) were fitted by a structural model combined of a nanodisc model based on earlier developements (Skar-Gislinge & Arleth, 2011) and the same core-shell particles. This will be described in more detail in the results section. The POPC-sodium cholate micelles were not modelled.

Results

To follow the structural development of the self-assembly of nanodiscs by detergent removal, scattering data were obtained from the sample mixture at different time points. This was compared to a sample not containing MSP, but treated in the same way and under the same conditions.
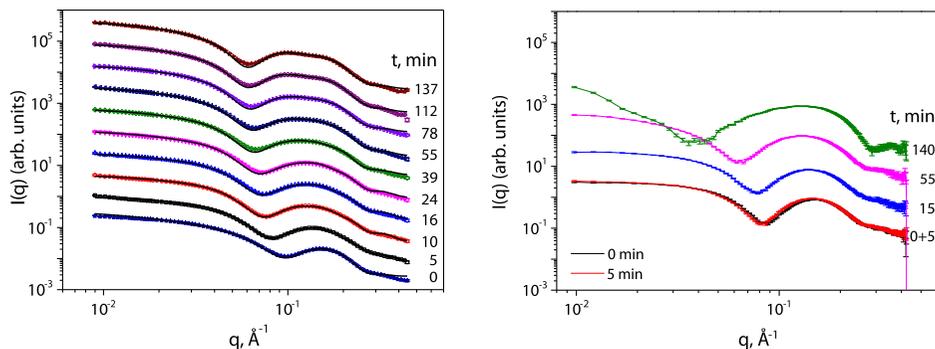
Figure 1. Left: SAXS data at different time points in the self-assembly process of nanodiscs. Time 0 represents the sample before adding Amberlites while the remaining times are minutes after addition of Amberlites. Data were collected at I911-4 in March 2013. Right, SAXS data of mixed cholate-lipid micelles. Time 0 represents before adding Amberlites while remaining time is after adding Amberlites. Data were collected at I911-4 in August 2014.

The measurements obtained at different times before adding Amberlites showed identical scattering patterns (data not shown), and it was concluded that no structural development is taking place during this incubation step. The scattering patterns obtained at different time points after addition of Amberlites together with a dataset before adding Amberlites are presented in Figure 1 and clearly show a structural change of the samples over time. The largest structural change for the nanodisc self-assembly is observed just after adding Amberlites, from $t = 0$ to $t = 5$ min. The structural development of the sample without MSP show much less structure change at the first time points as compared to the sample with MSP. At the latest time points for the nanodisc self-assembly only minor structural change takes place. At 137 minutes the assembly process appears to have finalized, and the data have the characteristic nanodisc features (Skar-Gislinge et al, 2010). For the sample without MSP, it is clearly seen that the later scattering pattern are very different, indicating transition from globular micelles to extended lamellar structures, probably in large vesicles at the late time points. The scattering data of MSP alone (data not shown) was used for determining the scattering contribution from MSP in the nanodisc self-assembly sample at t=0.

Data analysis using a structural model

The obtained scattering data from the nanodisc self-assembly were fitted by a structural model. The following information was implemented in the model: Before the addition of Amberlites, at $t = 0$ min, the solution contains two different particle types; MSP described as ellipsoidal particles as in a solution of MSP alone, and mixed micelles consisting of cholate and POPC in the form of core-shell particles. After the addition of Amberlites, for $t > 0$, the model contains

three different particle types; MSP given by the same ellipsoidal particles, mixed micelles in the same core-shell as obtained for $t = 0$, and nanodiscs. The model applied was therefore based on the assumption that the micelles in the sample do not change their structure or composition over time. The concentration of micelles decreases as cholate is removed and phospholipids are associated with the MSP to form the nanodiscs. In the model, the remaining fraction of micelles is given by the fraction of lipids that are not included in nanodiscs. Thus, the assumption is that the lipids are either present in nanodiscs or in micelles of the original composition and structure. The concentration of nanodiscs is 95% of the total MSP concentration, as determined for the latest data frame. The concentration of nanodiscs was therefore fixed at this value for all data sets, and the concentration of MSP in form of the ellipsoidal particles was correspondingly fixed at 5 % of the total MSP concentration.

The structural parameters of the nanodiscs were introduced as free fitting parameters: The number of POPC included in nanodiscs, N_{lipids}, the height of the hydrophobic section of the lipid bilayer (the traditional nanodisc model (Skar-Gislinge & Arleth, 2011) was simplified slightly so that it did not include a methyl layer with a separate contrast), $H_{hydrophob}$, the height of the curved end caps of the nanodiscs, H_{endcap}, and the height of the MSPs, H_{MSPs}. The height of the MSPs was fixed at 24 Å. These parameters together determine the two radii of the elliptical nanodisc, R_{minor} and R_{major}. An attempt to include cholate molecules into the nanodiscs were made, however leading to insignificant contributions. The slight aggregation observed at low q in the scattering data in Figure 1 is included in the model as the presence of nanodisc dimers with a distance of ~170 Å between the nanodiscs, for 10% of the nanodiscs.

From this model different structural parameters could be obtained, shedding light on the assembly of the nanodiscs.

Nanodisc structure

As described above, fitting the data to a structural model results in values describing developments over time of various structural parameters related to the nanodisc. The obtained parameters are plotted in Figure 2.
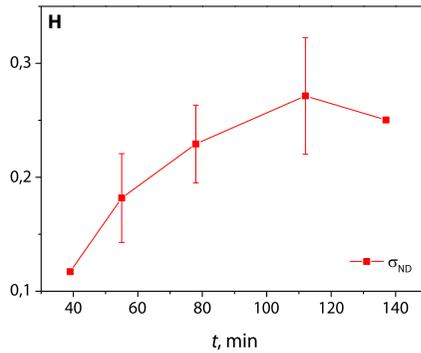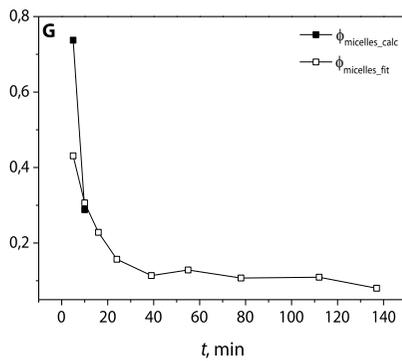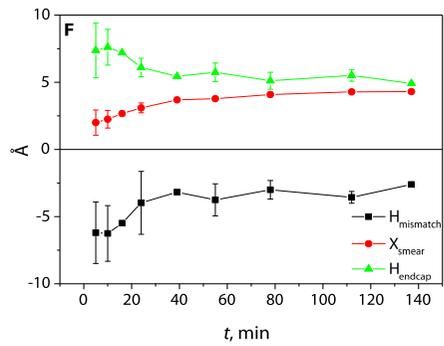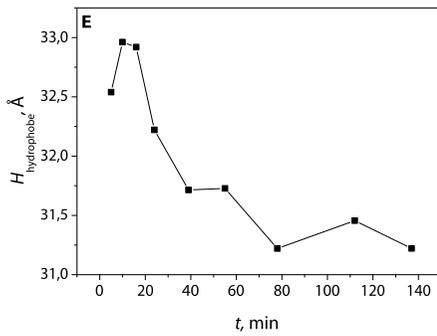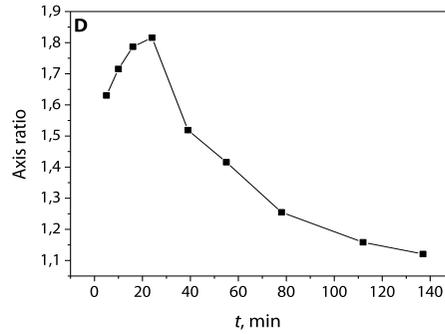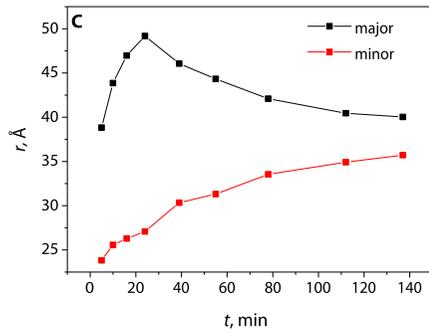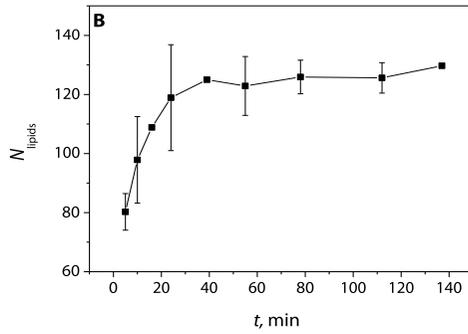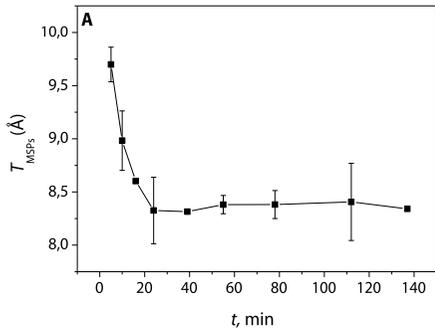
Figure 2. Nanodisc parameters over time. A, changes in thickness of the MSPs (T_{MSPs}). B, changes in number of lipids per disc (N_{lipids}). C, changes in radius of the two disc axis. D, changes in axis ratio of the discs, based on C. E, changes in height of the hydrophobic layer of POPC. F, changes in hydrophobic mismatch between the MSPs and the lipid bilayer (black), in smearing of the nanodisc (red) and in the height of end caps (curvature change) (green) of the nanodisc. Mismatch and endcaps are illustrated in Figure 4. G, micelle concentration of the total sample, based on the number of lipids (N_{lipids}) in nanodiscs shown in B. A mismatch between the model fit ($\phi_{micelles\_fit}$) and the calculated micelle concentration for the first data point was found and is shown as $\phi_{micelles\_calc}$. H, nanodisc polydispersity (σ_{ND}), shown as the relative lipid distribution given by the fraction of nanodiscs in the solution divided by the number of lipids per nanodisc. A, B, E and F are derived fitting parameters while C, D, G and H are calculated parameters, based on fitting parameters.

The thickness of the MSPs decreases over the first 24 minutes (Figure 2A) while the number of lipids increases drastically (Figure 2B). Since the MSPs consist of alpha-helical protein, one might have expected a constant thickness. For the lowest numbers of lipids, this would result in very elongated, asymmetric nanodiscs of high axis ratio, which is not observed. Instead, the MSPs are slightly less stretched, resulting in a larger width (under the applied assumption of a constant MSP height). As the number of lipids increases, the MSPs are stretched. This is associated with an increase in axis ratio as seen in Figure 2C and D. After 24 minutes, the number of lipids increases only moderately (Figure 2B). No further stretching of the MSPs is observed, indicating that they are fully stretched. Instead, the nanodisc becomes more circular to accommodate the higher number of lipids, and the axis ratio decreases and ends close to 1.1. The structural changes of the nanodisc are illustrated in Figure 3, based on the parameters presented in Figure 2.
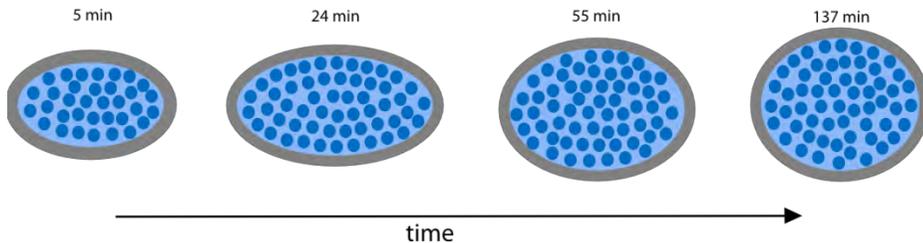
Figure 3. Top view illustration of the self-assembly of nanodiscs. Disc dimensions are based on axis ratio and MSP thickness obtained from the structural model, shown in Figure 2A&C.

The nanodiscs become less curved (less lens-shaped) over time as the height at the rim of the nanodisc increases and the height of the end caps, H_{endcap}, decreases (Figure 2F). The height at the rim of the nanodisc is given by $H_{MSPs} + H_{mismatch}$, where H_{MSPs} is the height of the MSPs (fixed at 24 Å) and $H_{mismatch}$ is the discrepancy between the MSP height and the lipid height. The negative value of $H_{mismatch}$ indicates that the MSPs more than cover the hydrophobic part of the lipids, however less over time. The total height of the hydrophobic part of the bilayer at the highest point of the curved nanodisc is given by $H_{hydrophob} = H_{MSPs} + H_{mismatch} + 2H_{endcap}$, this is illustrated in Figure 4 and plotted in Figure 2E and decreases continuously.

41
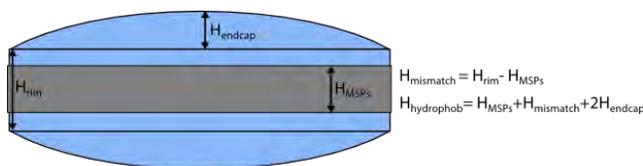
Figure 4. Hydrophobic height of the nanodisc. H_{rim} is the hydrophobic height of the lipids at the rim of the nanodisc. H_{endcap} is the height of the curvature on the hydrophobic height of the lipids at the center of the disc. H_{MSPs} is the height of the MSPs.

Smearing of the nanodisc interfaces was included in the model, resulting in the smearing widths presented in Figure 2F as X_{smear}.

A certain degree of polydispersity might be expected for the final nanodiscs, especially considering that no sample purification steps are taken. Polydispersity in the number of lipids per nanodisc was therefore included in the model as a flat distribution of width $\sigma = \sigma_{ND}N_{lipids}$. It was only necessary to include polydispersity for the later stages of the assembly process.

The relative width of the distribution of number of lipids in nanodiscs, σ_{ND}, is shown in Figure 2H. It appears to grow over time to a value of about 25%, corresponding to a distribution of N_{lipids} of $\pm$ 25% of the values given in Figure 2B

The fraction of micelles compared to the starting concentration, $\phi_{micelles\_calc}$, decreases quickly over time as lipids are built into the nanodisc, down to ~8% of the starting concentration as shown in Figure 2G. $\phi_{micelles\_calc}$ is equal to the fraction of lipids that is not inserted in nanodiscs, and is therefore given by the total concentration of lipids, the concentration of nanodiscs, and the number of lipids per nanodisc. However, for the dataset collected at t = 5 minutes, it was necessary to increase the micelle contribution beyond this value, and a fit parameter, $\phi_{micelles\_fit}$ was introduced. The mismatch between $\phi_{micelles\_calc}$, and $\phi_{micelles\_fit}$, is seen in Figure 2G.

Discussion

The obtained scattering data and the fitted structural model show the slow self-assembly of nanodiscs. The obtained scattering data from nanodiscs show over time very different structures from those of lipid micelles, as seen in Figure 1.

The first dataset of both the nanodisc mixture and the micelles (solution without MSP) look very similar indicating the dominating structure in the nanodisc mixture derives from micelles. However, following data points, the structures deviate strongly for the two solutions as nanodiscs are formed with MSP and a transition to lamellar structures takes place at the late time points without MSP, as seen in Figure 1. The structure change of the lipid micelles just after addition of Amberlites is almost negligible. For the nanodisc self-asembly the largest structure change is found at this point; as soon as Amberlites are added nanodiscs are formed with a relatively high number of lipids. This suggests that the cholate may be more tightly bound in the micelles compared to the nanodisc self-assembly mixture. However, when lipids are built into the nanodisc, cholate may become easily accessible for binding to the Amberlites and removed

from the solution faster. At the later time points very little structural changes are observed for the nanodisc self-assembly while the MSP free solution drastically changes shape. This may be caused to that little or no cholate is left in the solution to influence the lipid micelles resulting in lamellar structures.

The structural parameters in Figure 2 obtained from the model fitted to the nanodisc data reveals how the nanodisc grows over time as more lipids are added to the disc. It is interesting to notice that the nanodisc holds a less elliptical shape in the beginning of the assembly process than later on. As mentioned above, the early state nanodisc appear to arrange its MSPs in a less stretched manner when only few lipids are included in the disc, as illustrated in Figure 3. As lipids are added to the disc the MSPs stretches, resulting in a larger circumference, and the disc turns more elliptical. As the number of lipids increases further, the disc again turn less elliptical to increase its area while maintaining the width of the stretched MSPs, as seen in Figure 2D.

The change in axis ratio show good agreement with a recent study on nanodisc shape and stoichiometry (Skar-Gislinge et al, 2015). Here it was shown that the axis ratio decreases as more lipids are reconstituted into the nanodisc, until a saturation point. However, our model suggests a relatively higher axis ratio at some of the lower lipid stoichiometries. This may be due to the fact that their data were acquired from titration studies. The study presented here concerns structures during a kinetic transition, which do not necessarily correspond to the equilibrium states.

Here we have studied the decrease in cholate concentration in nanodisc self-assembly. The effect of different cholate concentrations on nanodisc structures has earlier been studied. A molecular dynamics (MD) simulation study approaching the inverted scenario of the here presented, adding cholate to the nanodisc, show some different results than the here presented study. The simulation builds cholate into the nanodisc particle, and as cholate concentration increases more globular/sphere-like shapes are formed (Shih et al, 2007). Based on these results, one might also speculate that nanodiscs contain cholate during the process of cholate removal. However, the results presented here indicate that this is not the case. Our model does not include cholate in the nanodisc at any time. The inclusion of cholate in the nanodisc structure was attempted, but only lead to insignificant numbers of cholate per lipid molecule (less than 0.02). However, in the MD simulation, the model is built in such a fashion that cholate can only be introduced in the nanodisc and cannot form detached micelles. It would be interesting to fit a similar model of the one used in the study presented here to the SAXS data included in (Shih et al, 2007).

At the initial states the discs has a high curvature which is associated with a low rim resulting in high mismatch and these parameters are observed to move closer to counterbalance over time. Figure 2E&F shows that the height of the hydrophobic layer and the curvature of the lipid bilayer reduce over time giving a more flat disc. As POPC holds a lipid packing parameter of ~1 it prefers to form flat bilayers. This means that during the assembly process some of the lipids cannot maintain their preferred curvature, creating a stress in the bilayer. As the curvature

decreases the lipids drive towards more favourable packing strains and a more energetically favourable state (Israelachvili, 2007b).

For the later times, increasing polydispersity of the nanodiscs was observed. This is not surprising as it is quite plausible that the discs will contain a distribution of lipids and therefore a variation in disc structure will occur. Exactly this variation in lipid number between discs show to vary up to 25% as seen in Figure 2H. This means that the dispersity in lipid distribution is relatively high between the discs. This is a little unexpected as nanodiscs are assumed relatively homogeneous particles. However, it must be kept in mind that nanodiscs normally, in contrast to this experiment, are purified by gel filtration, decreasing the polydispersity. See Appendix 1 for a standard nanodisc protocol.

The fraction of micelles left in solution at different time points, shown in Figure 2G were calculated on the background of number of lipids per disc obtained from the model presented in Figure 2B. These calculated fractions of micelles decreases over time, as more and more lipids are built into nanodiscs, decreasing the number of lipids left in solution. The calculated fractions determine the contribution of the micelle scattering in the model, resulting in very good agreement with the data, except from the first dataset at 5 minutes. In order to obtain an acceptable fit to the data, about 30 percentage point increase in micelles must be added the model, compared to the calculated fraction. The reason behind this discrepancy is unclear, but might be related to an overestimate of the number of lipids in the nanodiscs, which in turn might be related to deviations of the nanodisc structure or concentration from the model at this point.

For the later times, the assumption of a constant structure of the lipid-cholate micelle might lead to errors in the determined structural parameters: If the micelles in fact change their shape in a way comparable to the transitions (or part of the transitions) observed for the samples without MSP, this would affect the determination of all other parameters in the model. However, the excellent fits to the data obtained, indicate that the applied approach is in fact valid. In order to test this, experiments for determining the cholate content of the micelles in the sample without MSP, as well as in the micelles in the nanodisc samples are planned. This will be done by ^{14}C radiolabelling of cholate which will enable quantification of cholate concentration in the sample at each time point. From the obtained results, the parameters of the assembling of nanodiscs can be determined on a solid ground. Furthermore, knowledge of the cholate content can help to determine the scattering contrasts for the cholate molecule in the micelle (using estimates for the specific volume of cholate together with assumptions of its location in the lipid-cholate micelle did not result in contrasts matching the observed scattering patterns from the micelles) and verify that the cholate content in the nanodiscs is negligible during the entire assembly process.

Conclusion

We have for the first time structurally characterised the self-assembly of nanodiscs in a time resolved manner. This assembly process was found to have a time course of less than 137 minutes. No structural development was observed prior to addition of Amberlites implying this incubation step is unnecessary for the nanodisc formation.

A model based on containing nanodiscs and mixed micelles of lipids and sodium cholate was applied to obtain different structural parameters of the nanodiscs at the different measured time points. It was found that nanodiscs grow from elliptical lipid poor discs to discs with decreasing ellipicity as more lipids are included in the discs and that this causes stretching of the MSPs. According to our model, no cholate is included in the nanodisc at any time point. Quantification of the cholate concentration at each time point would help to determine the cholate contrast and location in the micelles, and thereby to give a final conclusion for the self-assembly process of nanodiscs. Experiments to do this are under planning.

B. Optimising reconstitution of membrane proteins in nanodiscs

Abstract

Reconstitution of membrane proteins in nanodiscs, for structural studies, has proven not to be a trivial task and many parameters may contribute to this. In structural small angle scattering studies, the most common problem observed is the presence of aggregates resulting in scattering data with a D_{max} larger than expected. Varying the buffer or lipid compositions have been explored for solving this problem, in the reconstitution of two types of rhodopsins in nanodiscs. Using POPC lipids, compared to a POPC-POPG mix and adding EDTA to the buffer show positive improvements.

As the choice of reconstitution detergent earlier has shown minor structural influence on nanodiscs without membrane proteins, this was also addressed for membrane protein nanodisc complexes. The reconstitution detergent used for reconstituting membrane proteins into nanodiscs appear to have substantial impact for the structural features and should be carefully considered. Furthermore, the impact of a D_2O based solvent was examined and found to have a significant destabilising effect.

Introduction

Membrane proteins can be reconstituted into nanodiscs and studied in a stabilising native-like environment. While several publications report studies of membrane proteins utilizing the nanodisc as sample holder, the large majority of these studies are on functionality (Bayburt & Sligar, 2010). Obtaining the high quality samples that are needed for structural studies such as small angle scattering has proven not to be a trivial matter. The most common problem observed in scattering data obtained from these complexes is a D_{max} larger than expected, indicating the presence of aggregates. So far, only one publication reports to have obtained scattering data of sufficiently quality for structural modelling of a membrane protein in nanodiscs, namely of bacteriorhodopsin (Kynde et al, 2014).

In order to reduce the aggregation observed in scattering data and analyse the structural effects of various parameters, variations in the reconstitution process has been studied using small angle X-ray scattering (SAXS). To optimise the reconstitution of membrane proteins in nanodiscs, the lipid and buffer composition were varied as well as the reconstitution detergent. Finally, applying a D_2O based solvent, which is widely used in small angle neutron scattering (SANS), was evaluated.

Background to the project

The following background covers brief introductions to detergents and rhodopsins. Information about D_2O and its effect on biomolecules can be found in Project C.

Detergents

In order to study structure or function of membrane proteins, these must first be isolated from the cell membrane and somehow kept stable and functional without its surrounding membrane.

This is where detergent comes into play. Detergents are one of the most important handling tools in the studies of membrane proteins and are in most cases used in the isolation and purification process. They are extensively used for storage, crystallization, NMR and functional studies (Bowie, 2001).

Detergents or surfactants are amphiphilic molecules consisting of a hydrophilic head group and hydrophobic tail. These molecules, just like lipids, spontaneously self-assemble into aggregates in aqueous solutions and share many physical properties with lipids. Detergents can in general be classified into four major categories according to their structure:

- Ionic detergents, containing a cationic or anionic head group and a hydrocarbon or steroidal tail.
- Bile acid salts, which are ionic but with a tail consisting of rigid steroidal groups.
- Non-ionic detergents, which has uncharged head groups of polyoxyethylene or glycosidic groups.
- Zwitterionic detergents, which have combined properties of ionic and nonionic detergents and are generally more deactivating for membrane proteins than non-ionic detergents (le Maire et al, 2000; Seddon et al, 2004).

Examples from the different detergent classes can be found in Figure 1. Non-ionic detergents are in general preferred for working with membrane proteins as they are considered mild and non-denaturing (le Maire et al, 2000; Seddon et al, 2004). Bile salt, non-ionic and zwitterionic detergents have been applied in this project, and their influence on reconstitution of membrane proteins in nanodiscs explored.
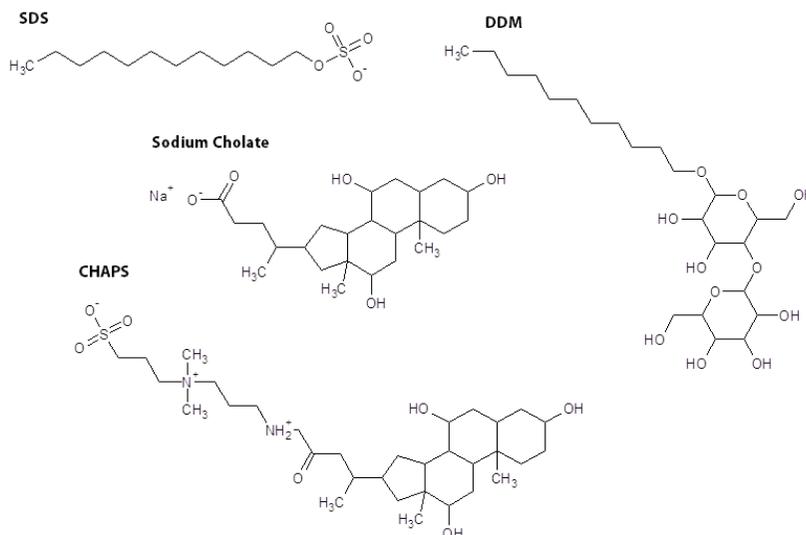
Figure 1. Examples of the four major detergent classes. Top left, the ionic detergent sodium dodecyl sulfate (SDS). Top right, the non-ionic detergent dodecyl maltoside (DDM). Middle, the bile salt sodium cholate. Bottom, the zwitterionic 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate (CHAPS).

Solubilisation of membranes happens as detergent monomers penetrate into the bilayer until saturation and from here mixed micelles of lipids and detergent are formed. How the detergents interact with the bilayer is highly dependent of the hydrophobicity and CMC. The more hydrophobic and the lower CMC, the higher affinity for penetrating lipid bilayers at low concentration (Jones, 1999).

The forces holding amphiphilic molecules together in micelles are weak, compared to conventional colloid particles held together by strong ionic and covalent bonds. Hydrogen and hydrophobic bonds, Van der Walls, and electrostatic interactions hold amphipathic molecules together and make them soft and flexible. Moreover, this also implies that changes in pH or electrolyte concentration in the surrounding solution will affect the interactions of the molecules, modifying the shape or size of the micelles (Israelachvili, 2007a).

Rhodopsins

Rhodopsins are a family of photosensitive membrane proteins containing seven membrane spanning helixes, found in both eukaryotes and prokaryotes. They can be divided into two classes, characterized by their function and presence in different organisms. Class I are light driven ion transporters or sensory transducers found in microbial membranes, such as bacteriorhodopsin, the best understood ion transport protein, found in *Halobacterium salinarium* (Luecke et al, 1999).

Class II rhodopsins function as light sensing receptors through interaction with G proteins in higher eukaryotic organisms, such as rhodopsin found in the mammalian eye, responsible for vision and a member of the so-called G-protein coupled receptor family (GPCRs) (Palczewski et al, 2000).

Rhodopsins are all coupled to a retinal chromophore responsible for the light sensing ability, and when exposed to light a conformational change is induced in the protein (Jung, 2007).

Two class I rhodopsins have been studied in this project, namely proteorhodopsin originating from *Halobacterium salinarium* and sensory rhodopsin II from *Halobacterium halobium*. The three dimensional structures of the two proteins, determined from NMR and crystallization respectively, are shown in Figure 2.

Proteorhodopsin (pR) is a growth stimulating light dependent proton pump found in marine proteobacteria and bacteroidetes. The protein holds a strong red pigmentation easily visible in overexpressed cells and absorbs green light, with a maximum absorption wavelength $\lambda_{max} = 530$ nm (Giovannoni et al, 2005).

Sensory rhodopsin II (sRII) is a phototaxis receptor found in halobacteria, with a high functional similarity to the animal rhodopsin. It has an orange pigmentation and is slightly blue shifted in its absorption maximum compared to proteorhodopsin with a $\lambda_{max} = 487$ nm (Takahashi et al, 1990).
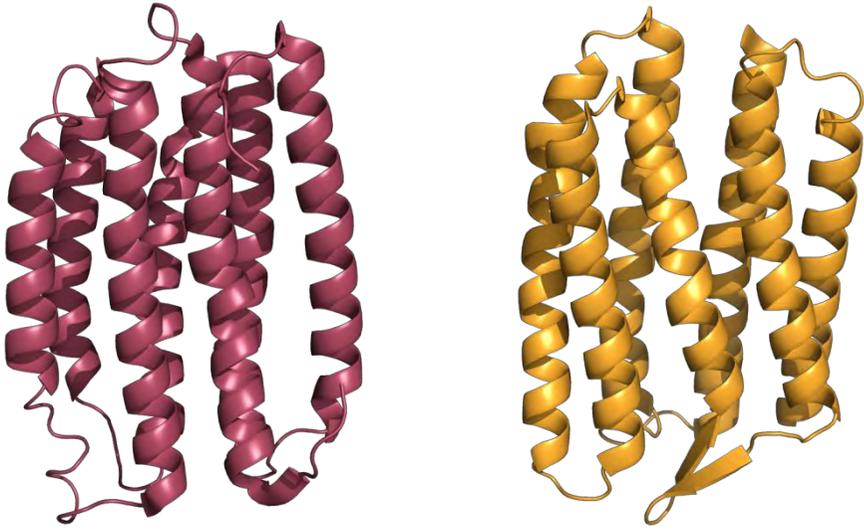
Figure 2. Left, NMR structure of proteorhodopsin from marine γ-proteobacterium [pdb entry 2L6X]. Right, crystal structure of sensory rhodopsin from _Natronomonas pharaonis_ [pdb entry 1H68].

Methods

The structural and stabilising effects of different parameters has been tested in this project, including variation in lipid and buffer composition in the reconstitution of membrane proteins in nanodiscs. The effect of type of reconstitution detergent, centrifugation and D_2O based buffer has as well been studied, using SAXS. The experimental details of the different parameter variations will be described following.

All nanodiscs were prepared using a molar ratio 1:20:1300 for membrane protein:MSP:lipid. Membrane protein and phospholipids were solubilised in detergent prior to mixing. Following mixing of the three components, detergent was removed and membrane protein loaded nanodiscs were purified using affinity chromatography and gel filtration. A detailed protocol for reconstitution of membrane proteins in nanodiscs can be found in Appendix 2. All nanodiscs were prepared in accordance to this protocol and any variations from the protocol are described here.

Proteorhodopsin expression and purification

pR was expressed and purified as earlier described (Gourdon et al, 2008), except from cells were grown in 5 l shaker flasks, without monitoring airflow and pH and at 200 rpm. Cultures were incubated for approximately 4 hours to an $OD_{600\ nm}$ ~1 before induction and stirring speed reduced to 150 rpm post induction.

Lysozyme (0.1 g/g cells), all-trans retinal (10 μg/g cells), 1 mM PSMF and DNase (5 μl/40 g cells) were added harvested cell mass kept on ice and sonicated 4x4 min. at 50% output with a tip sonicator.
Following IMAC purification, the concentrated pooled fractions were loaded onto a column containing 7 ml SP Sepharose high performance beads (GE healthcare).

The purified pR was added glycerol to a final concentration of 10%, flash frozen in liquid nitrogen and stored at -80˚C.

POPC-POPG mix vs POPC

Two samples of pR reconstituted in nanodiscs were prepared, using a 1:2 ratio of POPG:POPC lipid mix for one of the samples. The lipids solubilised in chloroform (Avanti Polar Lipids) were mixed and dried to a lipid film and resuspended in cholate containing buffer. The applied gel filtration buffer was without EDTA. One fraction of the POPG-POPC mix sample was flash frozen using liquid nitrogen immediately after preparation and thawed on ice just before SAXS measuments, while another was kept in solution on ice until measurements. The samples were measured at the I911-4 beamline at MaxLab.

EDTA

pR was reconstituted in nanodiscs and purified with and without 1 mM EDTA in the gel filtration buffer. The samples were measured by SAXS at the I911-4 beamline at MaxLab.

pH and NaCl variation

Four different samples of pR reconstituted in nanodiscs were prepared. At pH 6.5 (100 mM MOPS, 100 mM NaCl), pH 7.5 (20 mM Tris/HCl, 10 mM NaCl), pH 7.5 (20 mM Tris/HCl, 100 mM NaCl) and pH 8.5 (20 mM Tris/HCl, 100 mM NaCl). SAXS data were recorded at the I911-4 beamline at MaxLab.

Reconstitution detergent

Four different samples of pR reconstituted in nanodiscs were prepared using different solubilisation detergents for the lipids for three of the samples. 0.1 M n-dodecyl-β-D-maltoside (DDM), 0.1 M CHAPS or 0.2 M Octyl-β-glycoside (OG) was applied. All samples were measured by SAXS at BM29 at ESRF.

Centrifugation

pR and sRII were reconstituted in nanodiscs and measured by SAXS. The two samples were following centrifuged for 10 minutes at 10000 rpm and the top fraction immediately measured by SAXS. Data were collected at BM29 at ESRF.

H_2O vs D_2O

pR and sRII was reconstituted in nanodiscs and measured by SAXS before solvent exchanged to a D_2O based buffer by spin filtration using a 10 kDa cut off (Amicon). The samples were measured by SANS (data not included in this report) and immediately followed by a SAXS measurement. SAXS data were obtained at BM29 at ESRF.

SAXS

Samples were measured at BM29 or I911-1. In both cases the SAXS data were converted to scattering intensity units as a function of scattering vector $\vec{q}$, $I(q)$, where $q = \left| \frac{4\pi sin\theta}{\lambda} \right|$, where 2θ is the scattering angle and λ is the X-ray wavelength

BM29

SAXS-data were recorded at the BM29 BioSAXS beamline (Pernot et al, 2013) at ESRF, Grenoble. 10 frames of both samples and buffers, each exposed for 1 s, were recorded to ensure that the sample was not damaged by the radiation. The experiments were conducted at 10 °C, and the samples were irradiated using a nominal x-ray wavelength of 1 Å. The instrument uses a PILATUS 1M-detector located 2.85 m from the sample; and the tasks of data reduction and treatment were performed using in-house software at the instrument.

I911-4

Scattering data were recorded at the I911-4 beamline at MaxLab (Labrador et al, 2012). Samples were exposed for 120 s using a nominal X-ray wavelength of 0.91 Å. Experiments were conducted at 10°C and data detected using a PILATUS 1M-detector located 1.98 m from the sample. Data reduction and treatment were performed using the BioXtas raw software (Nielsen et al, 2009).

Results

POPC-POPG mix vs POPC

Small angle scattering data of bacteriorhodopsin (bR) has earlier been obtained with great success using a POPG-POPC lipid mix (Kynde et al, 2014). However, earlier studies of nanodiscs without membrane proteins, has shown that using POPC lipids show different scattering pattern than nanodiscs containing a POPG-POPC mix (Kynde et al, 2014; Skar-Gislinge et al, 2010). From personal experience, POPC has shown greater stability for nanodiscs than other tested lipid types. The difference in using these two lipid compositions where for this reason studied.

pR was reconstituted with a 2:1 POPC:POPG lipid mixture and compared to reconstitution with POPC alone. The sample containing the lipid mix was also compared to a sample fraction kept frozen and thawed just before measurements.
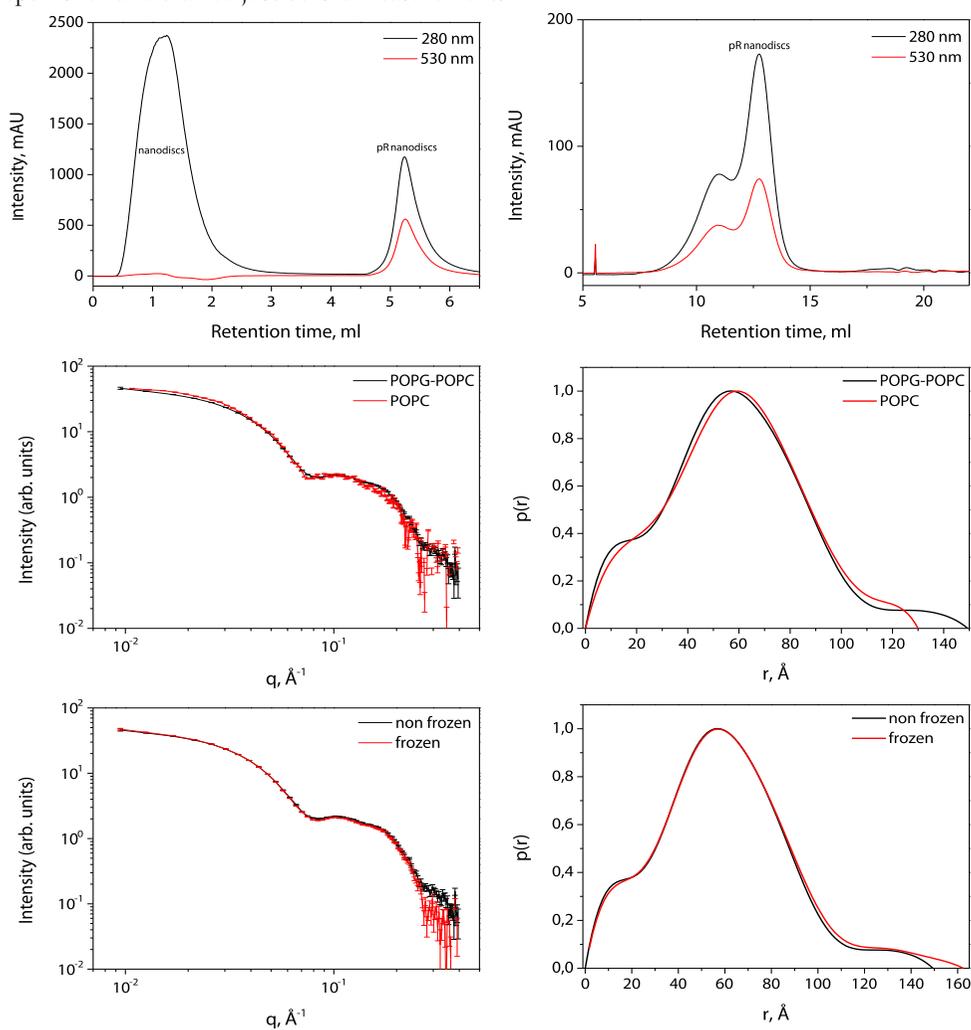
Figure 3. Purification and SAXS data of pR reconstituted in nanodiscs. Top left, IMAC purification using a 1 ml HisTrap column, absorption monitored at 280 (black) and 530 (red) nm. Top right, SEC chromatogram using a Superdex 200 10/300 GL column with absorption monitored as for IMAC. Middle SAXS data (left) of reconstitution with a 2:1 POPC:POPG mix (black) and POPC only (red) with corresponding $p(r)$-functions (right). Bottom, SAXS data (left) of reconstitution with 2:1 POPC:POPG mix which had been frozen (red) compared to a sample maintained in solution on ice (black) with corresponding $p(r)$-functions (right). Scattering data were obtained at I911-4 in May and June 2012.

As seen on the chromatogram from the IMAC purification in Figure 3, only discs containing pR (arbsorbing at 530 nm) is bound to the column while empty discs passes through. From the gel filtration it is seen that some larger aggregates are formed containing pR, aside from the defined nanodisc peak eluting at ~12.7 ml. Some aggregation during concentration prior to gel filtration was observed as the sample turned cloudy, causing loss in sample yield.

The reconstitution with POPC only, showed a chromatogram close to identical with that for the POPC-POPG mix (data not shown). Furthermore, diluting the imidazole concentration immediately after IMAC, eliminated the cloudy aggregation earlier observed.

Using a POPC-POPG lipid mixture gives rise to light aggregation while using POPC lipids decreases the aggregation, judged from the low q-region of the scattering data shown in Figure 3. This is even clearer from the change in D_{max} in the corresponding $p(r)$-functions. The lipid composition also impacts the structure as the "bumps" following the minima in the scattering data have slightly different shapes for the two samples.

Flash freezing the lipid mixture sample does not appear to influence the structure, as the two samples look close identical judged from the scattering data in Figure 3. However, a little aggregation is observed after freezing, seen as an increase in D_{max} in the $p(r)$-function of the frozen sample compared to the sample kept in solution on ice.

EDTA

Divalent ions, represented as impurities in the solution, may interact with lipid bilayer, causing instability. It has earlier been shown that low concentrations of Ca^{2+} can interfere with lipids changing thermodynamic properties, though this can be prevented by adding EDTA to the solution (Riske et al, 2003).

Gel filtration purification of pR nanodiscs with and without EDTA in the elution buffer were compared.
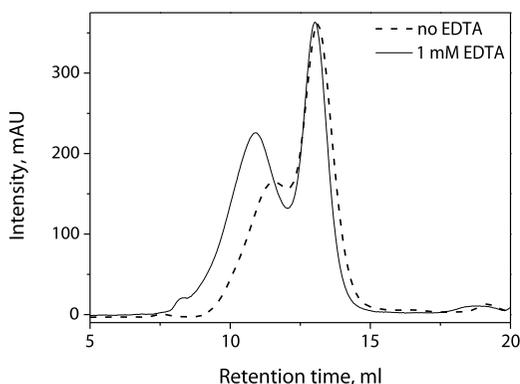
Figure 4. SEC purification chromatograms of pR in naodiscs with (solid) and without (dashed) EDTA, in the elution buffer, using a Superdex 200 10/300 GL column.

As it can be deduced from the chromatogram shown in Figure 4, that adding EDTA to the elution buffer slightly changes the elution time of larger aggregates, and thereby a separation of nanodiscs from aggregates is more easily obtained.

pH and NaCl variation

Varying pH or NaCl concentration can have a protein stabilising effect by changing the ionic interactions in the protein (discussed further in Project C). Though, this did not have any structural effect on nanodiscs without membrane proteins, as shown in Project C.

To address whether this could have a stabilising effect for membrane protein loaded nanodiscs, pH and NaCl concentration were varied in the reconstitution and purification. At pH 6.5 the sample aggregated, hence no reconstitution at this pH was obtained.
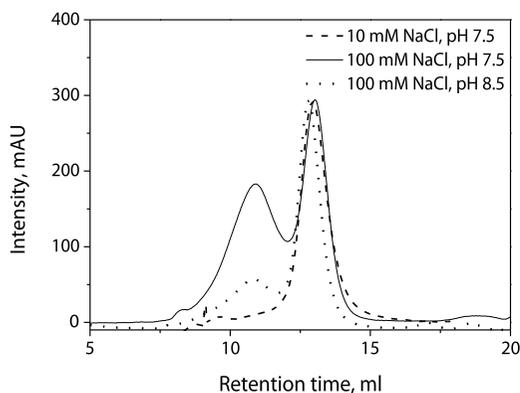
Figure 5. Gel filtration chromatograms of pR containing nanodiscs at 10 mM NaCl, pH 7.5 (dashed) and 100 mM NaCl, pH 8.5 (dotted), compared to the usual 100 mM NaCl, pH 7.5 (solid). A Superdex 200 10/300 GL column was used for the purification and absorbance monitored at 280 nm.

From the chromatogram, shown in Figure 5, of nanodiscs reconstituted in 10 mM NaCl it can be seen that very little aggregates are formed compared to using 100 mM NaCl. The elution peak looks nice and symmetrical. In the chromatogram from nanodiscs reconstituted at pH 8.5,

larger aggregates are formed but in a much lower amount than observed from nanodiscs reconstituted at pH 7.5. Again the peak looks nice and symmetrical.

These samples were measured by SAXS together with a sample reconstituted in 100 mM NaCl, pH 7.5 and scattering data compared in Figure 6.
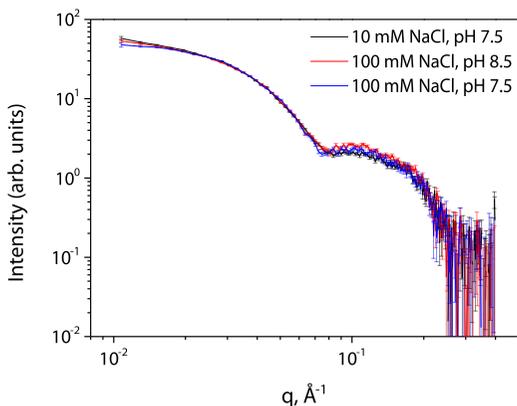
Figure 6. SAXS data of pR containing nanodiscs at different buffer conditions. Black is 10 nM NaCl, pH 7.5, red is 100 mM NaCl, pH 8.5 and blue is 100 mM NaCl, pH 7.5. Scattering data were obtained at I911-4 in June 2012.

From the scattering data, in Figure 6 it can be seen that both lowering the salt concentration and increasing pH result in some aggregation compared to the usual reconstitution conditions. This is seen as a slight upturn in the low-*q* region

Reconstitution detergent

The choice of reconstitution detergent has earlier shown to have a great impact on the nanodisc structure (Skar-Gislinge et al, 2015) and was therefore interesting to examine in what degree this would affect reconstitution of membrane proteins in nanodisc. Also the fact that many membrane proteins are fragile and only tolerate very mild detergents, could dictate the detergent to use for the reconstitution in future experiments.

Reconstitution detergent was varied, using DDM, OG or CHAPS besides from the "usual" sodium cholate. CHAPS appeared to destabilise the sample leading to visible aggregation and turned colourless due to unfolding and loss of the bound retinal of proteorhodopsin, hence no sample was obtained.
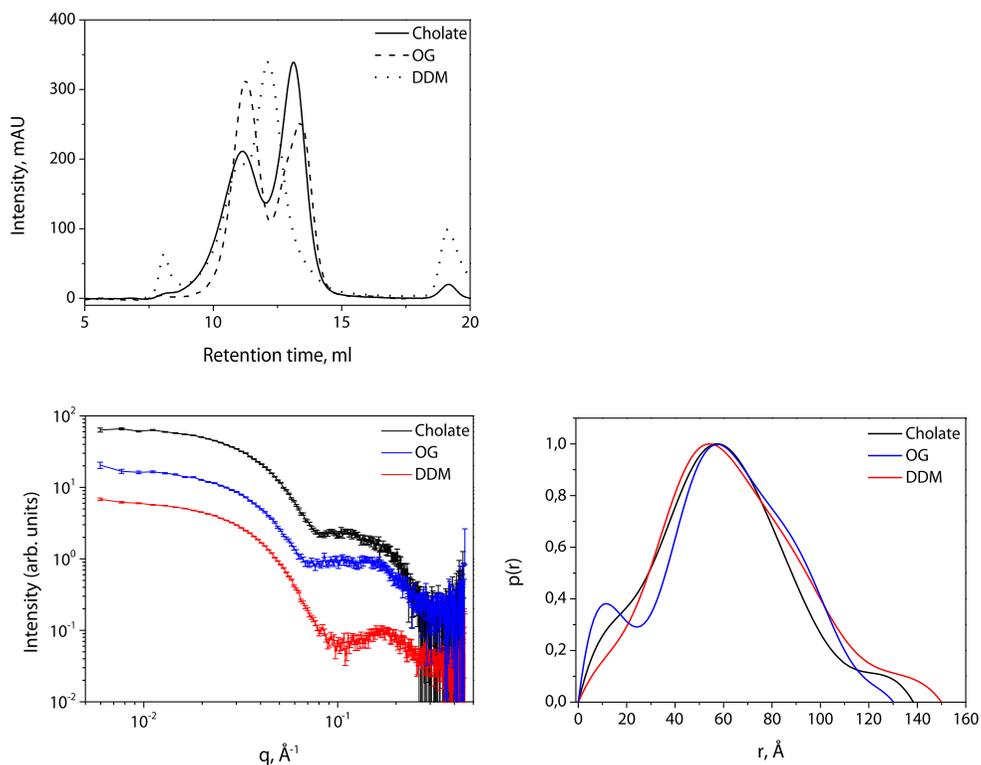
Figure 7. SEC purification chromatograms and SAXS data of pR in nanodiscs using different reconstitution detergents. Top, chromatogram using sodium cholate (solid), DDM (dashed) or OG (dotted) for reconstitution. A Superdex 200 10/300 GL column was used for SEC and absorbance at 280 monitored. Bottom, SAXS data (left) of pR nanodiscs reconstituted using sodium cholate (black), DDM (red) or OG (blue) with corresponding *p(r)*-functions (right). Scattering data were obtained at BM29 in September 2012.

The chromatograms obtained using the three different detergents, found in Figure 7, show very different profiles. Using DDM as reconstitution detergent results in an earlier elution time and thereby a worse separation from larger aggregates. Using OG for reconstitution gives rise to a higher degree of large aggregates and consequently losses in sample yield. Also the elution time is slightly longer than from what observed using sodium cholate and the nanodisc elution peak looks asymmetrical.

From the obtained SAXS data it is clear that the structural features of the three different samples are very different. However, it is possible to obtain samples with all three detergents without any clear signs of aggregation, judged from the scattering data in Figure 7. Hence, it can be deduced from the *p(r)*-functions that D_{max} is largest when using DDM and smallest using OG, as suggested from the gel filtration.

Centrifugation

It is commonly known that macromolecyles can be separated according to mass by centrifugation. It was therefore interesting to analyse if centrifuging the sample prior to SAXS

measurements could reduce aggregation. Reconstituted pR and SRII was measured by SAXS before and after centrifugation (10 min at 10000 rpm).
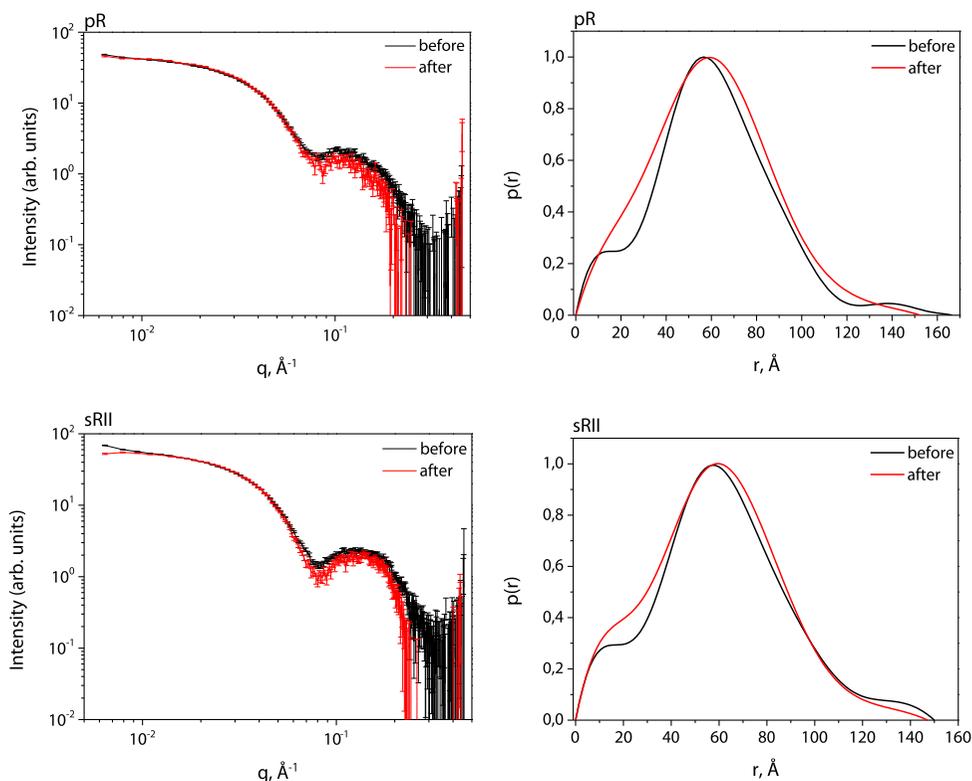
Figure 8. Left, SAXS data of pR (top) and sRII (bottom) before (black) and after (red) centrifugation. Right, corresponding *p(r)*-functions. Scattering data collected at BM29 in June 2012.

From both the scattering data and the corresponding *p(r)*-functions, shown in Figure 8, it is clear that centrifuging the sample before measurement reduces aggregates. The *p(r)*-functions show a smaller D_{max} for both samples as well as change in shape at short distances.

H_2O vs D_2O

Using a D_2O based solvent has favourable properties in SANS measurements, as this alters the scattering contrast significantly compared to H_2O (this is discussed further in project C). But as also observed, D_2O has some destabilising consequences for the nanodisc. To evaluate the consequences of keeping a nanodisc sample with membrane protein in D_2O for SANS measurements, pR and sRII reconstituted in nanodiscs were measured by SAXS before, in H_2O, and after SANS, in D_2O.
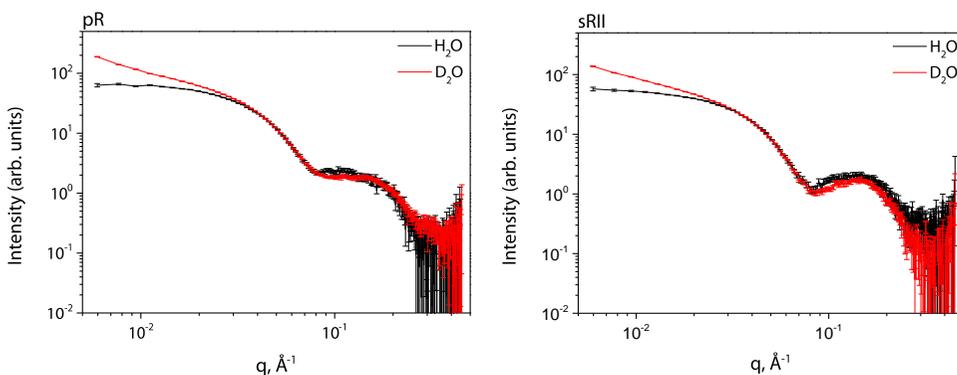
Figure 9. SAXS data of pR (left) and sRII (right) in nanodiscs in H₂O (black) and in D₂O after SANS measurement (red). Data collected at BM29 in September 2012.

From the scattering data in Figure 9 it is clear to see that pronounced aggregation occur over "short" time when using D_2O as solvent compared to H_2O. a very clear upturn is observed at low q for both samples in D_2O.

Discussion

The idea with this project was to evaluate stability and the structural effects of different variables on reconstitution of membrane proteins in nanodiscs. In structural studies, the most common problem observed is the presence of aggregates giving rise to scattering data with a D_{max} larger than expected. Various experiments have been carried out to optimise the reconstitution and to understand which factors impact the structure.

POPC-POPG mix vs POPC

pR was reconstituted in nanodiscs with a 2:1 ratio of POPC:POPG lipids based on a previous reconstitution of bR (Kynde et al, 2014). Using MSP without a His-tag, utilizing the His-tag of pR, only nanodiscs with membrane proteins are retained on the HisTrap column. This was confirmed by the obtained chromatogram from IMAC in Figure 3. Absorption at 530 nm was only detected for the fraction bound to the HisTrap column. The collected fraction from the HisTrap column was concentrated by spin filtration prior to gel filtration. Noticeable aggregation was observed during this process. It was assumed that containing the sample in a high concentration of imidazole might have led to some destabilisation of the protein complex, as this has before been observed for membrane proteins (Pollock et al, 2014). The observed aggregation was later prevented by diluting the imidazole immediately after IMAC. A different approach to decrease the imidazole concentration could be employed, by using a different type of IMAC than Ni-NTA. Using TALON, a cobalt based IMAC, a lower imidazole concentration is needed for protein elution than Ni-NTA. This is due to its lower affinity for polyhistidine, compared to Ni-NTA, which on the other hand can result in a lower yield (Bornhorst & Falke, 2000).

The following purification by gel filtration in Figure 3 showed larger aggregations followed by a dominating symmetrical peak. Judged from the defined peak a monodisperse pR nanodisc

sample could be collected. However, a better separation from larger aggregates would be preferred as well as a decrease in amount of these large aggregates.

Using POPC only for reconstitution did not show any difference for the purification and it was not possible to evaluate if the choice of lipids had any structural influence from the chromatograms.

SAXS data were obtained of the samples to evaluate if using POPC lipids would make any difference from the POPC-POPG mix. Using POPC alone reduced aggregation observed for the lipid mix and suggests that using POPC alone acts more stabilising. This was seen as lower D_{max} was observed from the $p(r)$-functions in Figure 3 of the POPC discs. A minor difference in shape of the two samples with different lipid composition were observed, which is in good agreement with what has earlier been seen for nanodiscs without membrane protein (Kynde et al, 2014; Skar-Gislinge et al, 2010). This means that the choice of lipid should be carefully considered and stability of this evaluated. Also, the structural influence means the type of lipids may contribute to some challenges when analysing the data, using a structural model.

Keeping protein samples at soluble conditions over time is not ideal and experimenting at large scale X-ray or neutron facilities means extensive travelling time and transport of samples. Sample quality will for this reason always be a compromise between favouring sample preparation in an in house lab and the travelling time to the facility. In order to limit the time a sample is kept in solution before a SAS measurement, the sample must be prepared as close as possible to departure. This puts a pressure on experimentalists as there is no time for mistakes. Also the use of lab facilities and instruments among lab users must be coordinated in detail. This unfavourable work situation may be avoided by preparing frozen samples as the sample can be prepared much earlier in advance. Though transporting frozen samples do also open up for other problems as not all airlines allow dry ice on their flights, while using a courier for transportation can be risky in terms of disappearance of samples.

pR in nanodiscs were measured by SAXS, one sample had been kept in solution while another was flash frozen after preparation and thawed at the beamline. No structural differences between the two samples were observed judged from the obtained scattering data in Figure 3. Minor aggregation was observed from the deduced $p(r)$-function, which may be removed simply by centrifugation and definitely by gelfiltration. This means that flash freezing and transporting samples frozen may be a solution, however, this should be investigated more extensively for more membrane proteins before any further conclusion can be made.

EDTA

EDTA was added to the gel filtration elution buffer to remove potential divalent ions present as impurities, which could interact with the lipid bilayer and cause destabilisation. This resulted in a better separation of larger aggregates from the nanodisc fraction as seen in the chromatogram in Figure 4. pR nanodiscs could easier be purified by gel filtration without risking the inclusion of unwanted aggregated in the final sample.

pH and NaCl variation

Varying pH or NaCl concentration can have a protein stabilising effect, but have no structural effect on nanodiscs without membrane proteins, as shown in Project C. pR nanodiscs were prepared at pH 6.5 and 8.5 containing 100 mM NaCl as well as a sample at pH 7.5 and 10 mM NaCl and compared with the usual pH 7.5 and 100 mM NaCl. The sample prepared at pH 6.5 could not be obtained due to aggregation. As it has been possible to prepare nanodiscs at these conditions (Project C), pH 6.5 must have been too acidic for keeping pR stabilised.

Reconstituting pR in nanodiscs at pH 8.5 led to a dramatic decrease in larger aggregates from at the usual condition, judged from the obtained chromatogram from gel filtration in Figure 5. Preparing the sample at 10 nM NaCl did decrease the amount of larger aggregates even more. This was very promising results. Unfortunately, the following SAXS measurements did not deliver an optimal outcome. Minor aggregations could be observed for both the sample at pH 8.5 and at 10 nM NaCl, not present in the sample prepared under usual conditions. Also a minor structural change could be observed for both samples compared to the usual sample, expressed as a less pronounced minima in the obtained scattering data in Figure 6. This means that pH and NaCl concentration does have a minor impact on the scattering data and hence structure of nanodiscs containing membrane proteins. In which level these conditions affect other membrane proteins in nanodiscs is uncertain and should be investigated if relevant. Also a larger variation of pH and NaCl concentration should be investigated in order to find optimal conditions for purification and scattering data, as purification can be optimised.

Reconstitution detergent

The choice of reconstitution detergent has earlier shown to have a great impact on the nanodisc structure (Skar-Gislinge et al, 2015). On the other hand, certain detergents may have unfavorable effects on membrane proteins and dictate which detergent to use. pR was reconstituted in nanodiscs using four different detergents, namely CHAPS, DDM, OG and sodium cholate. Reconstitution in CHAPS could not be achieved despite several attempts and led to visible aggregation. This detergent is clearly too harsh for pR, as reconstitution of nanodiscs without membrane protein has been succeeded with CHAPS (Skar-Gislinge et al, 2015). The reason for this may be that Chaps is a zwitterionic detergent causing the destabilisation of pR (Seddon et al, 2004).

The chromatograms, obtained from gel filtration in Figure 7, using the three other detergents showed very different profiles. For DDM an earlier elution time of the pR nanodiscs were observed and thereby a worse separation from larger aggregates, compared to cholate. This suggests that nanodiscs formed with DDM, as reconstitution detergent, are larger than those formed using sodium cholate. For OG as reconstitution detergent, a higher amount of larger aggregates was present, compared to sodium cholate. This gave losses in sample yield. Also the elution time was slightly longer than from what observed using sodium cholate, suggesting smaller nanodiscs are formed using this detergent. Furthermore, the nanodisc elution peak was asymmetrical indicating an inhomogeneous sample. Judged from the three chromatograms, sodium cholate is the preferred reconstitution detergent giving the best separation and highest yield.

The scattering data collected from the three samples in Figure 7 showed great structural variety. Not much can be concluded about the structures of the different samples other than choice of

reconstitution detergent matters significantly for the structure formed. The scattering data was attempted analysed by a structural model also used for bacteriorhodopsin in nanodiscs (Kynde et al, 2014), but was not fully successful. If a different reconstitution detergent than sodium cholate is applied, new structural characterization and modelling parameters must be implemented before any structural work can be carried out with membrane proteins.

Centrifugation

pR and sRII was reconstituted in nanodiscs and left for a few days before measured by SAXS. This clearly illustrates that aggregation is a typical problem with these kind of samples. In order to remove this, the samples were centrifuged for ten minutes which resulted in a clear reduction of these minor aggregates, as seen in the obtained scattering data and corresponding *p(r)*-functions in Figure 8. It was found that sRII was more prone to aggregation than pR and this is definitely something that will vary from protein to protein. Leaving samples in solution over time clearly has a destabilising effect but minor aggregations can be reduced by centrifugation.

H_2O vs D_2O

To evaluate the consequences of changing the buffer from H_2O based to D_2O based, pR and sRII were reconstituted in nanodiscs and measured by SAXS before and after SANS measurement. Samples were first measured by SAXS and then solvent exchanged to D_2O based buffer before SANS measurement. After ended SANS measurement the samples were again measured by SAXS, now in D_2O. From the scattering data in Figure 9, it is very clear the samples in D_2O contain aggregates and this solvent is destabilising the sample. It must also be kept in mind that this aggregation is a time dependent factor and that the time frame for a SANS measurement is relatively long, because of the low flux and change of detector distance. This may have an influence on the collected data as the sample may change during the actual measurement. It is therefore really important to find a solution for this problem by adding some kind of stabilising agent or somehow modify the protein to become less affected by D_2O. For now, trying to obtain SANS data of membrane proteins in nanodiscs should be carefully considered if worth doing, until a solution is found.

Conclusion

From these studies with rhodopsins in nanodiscs it is clear that many things influence the structure. Any variation from the usual protocol must be carefully evaluated. Changing one variable can lead to unexpected structural changes nanodisc membrane protein complex which might not be possible to model with the tools available. Using POPC lipids and sodium cholate should be employed if possible in order to improve scattering data quality. Adding EDTA to the gel filtration buffer improves purification by better separation of larger aggregates and the nanodisc elution peak. Centrifuging the sample immediately before SAXS measurements improves data quality.

D_2O based buffers have a destabilising effect on proteins compared to H_2O based buffers, clearly detected in small angle scattering experiments. As D_2O is the preferred solvent for SANS measurements, SANS data from membrane proteins in nanodiscs must be critically evaluated if worthwhile obtaining at this point. It is clearly a problem that needs to be solved.

C. Investigating aggregation of biomolecules in D₂O

Abstract

Destabilisation of proteins in D$_2$O is a general problem observed as mild protein aggregation or coagulation. This is among other things observed for nanodiscs in SANS experiments when using a D$_2$O based buffer, leading to poor data collection. Protein aggregation gives rise to larger particles in the sample solution, clearly noticeable at low q-values in SAS data and significantly hampers the data quality due to the inhomogeneity of the measured sample.

Here SAXS has been applied to study different approaches to prevent or remove coagulation of nanodisc samples in D$_2$O based buffer. Adjustable parameters such as varying buffer composition or filtration does not improve sample stability. In contrast, the method for buffer exchange highly impact sample stability and in turn data quality.

Introduction

D$_2$O based solvent for bioSANS is preferred over H$_2$O based solvent because of a reduction in incoherent background scattering (NIST, 2015). For nanodisc studies SANS is very valuable as a supplementary contrast to SAXS, for increasing the number of constrains in the subsequent data analysis. Unfortunately is mild aggregation of nanodiscs in D$_2$O based buffers a general problem in these SANS experiments, leading to poor data quality. Larger particles, caused by destabilisation of proteins, leading to aggregation, can clearly be observed at low q-values of SAS data and makes the scattering data less useful as information, such as the non-aggregated molecule size, is lost. Different approaches to decrease or ideally remove the observed aggregation of nanodisc samples in D$_2$O based buffer, has been investigated in this project using SAXS. SAXS was used rather than SANS as SAXS is more accessible, measurements are faster and the structural information which can be obtained just as useful for this study. Varying the buffer composition in means of pH (pD) and NaCl concentration has been evaluated as well as the addition of methanol or acetonitrile.

Deuterium and proteins

Hydrogen and deuterium behave very similar in molecular compounds, though deuterium holds twice the mass of hydrogen, which is useful for various applications in structural biology. Deuteration of proteins is widely used in neutron diffraction and NMR, for obtaining multiple contrasts and thereby more structural information, and is based on the assumption that deuterium has no influence on protein structures. In 1962 Tomita and coworkers published a paper that challenged this hypothesis by examining the effect of deuteration of the alpha helix. Their findings showed a smaller elongation of the hydrogen bonds of 0.03 Å, and concluded this should not result in any immediate structural change of the overall protein (Tomita et al, 1962). This assumption was indeed supported in a review published in 2008 comparing 23 different proteins in their hydrogenated and deuterated crystal structures, concluding that no significant differences could be found (Fisher & Helliwell, 2008).

Solvent exchange from H_2O to D_2O is a very valuable tool in SANS to increase signal to noise ratio. D_2O is a preferable solvent as background scattering is decreased because of the low incoherent scattering compared to H_2O (NIST, 2015).

D_2O and H_2O are known to hold very similar biophysical properties, due to the similar chemical behaviour of hydrogen and deuterium, and are therefore also expected to possess similar solvent effects. However, it has been observed that biological molecules are affected by this solvent exchange, resulting in change of essential molecular properties. It has been found that the enthalpy and protein denaturation temperature changes in D_2O solvent from H_2O resulting in increased destabilisation of the proteins (Chen et al, 1984; Makhatadze et al, 1995; Meilleur et al, 2004).

The mentioned elongation of hydrogen bonds observed by Tomita can also be observed for molecules when subjected to the so-called isotope effect, caused by solvent exchange. This was discovered and extensively studied in the 1930s by Ubbelohde and co-workers. They concluded that the small elongation of the hydrogen bonds, when hydrogen is exchanged for deuterium, would result in a weakening of the bonds (Robertson & Ubbelohde, 1939). The observed destabilisation of proteins when solvent is exchange to D_2O may be coursed by this isotope effect, altering hydrogen bonds. Interestingly are the hydrogen bonds in D_2O molecules shorter than in H_2O and in addition to that is the bond angle between two heavy water molecules larger (Soper & Benmore, 2008). Furthermore, it is a fact that D_2O holds a higher freezing and boiling point from that of H_2O as well as enthalpy and entropy of vaporization. Also, the viscosity of D_2O is higher than of H_2O and D_2O holds in comparison a more ordered structure (Holtzer & Emerson, 1969). Some other properties of heavy water are a higher neutral pH (pD) (Covingto.Ak et al, 1968) and acids are in general weaker in D_2O than H_2O (Robinson et al, 1969). With these many minor differences it is not surprising that H_2O and D_2O cannot act as identical solvents.

Methods

The different approaches investigated in this project, to decrease nanodisc aggregation in D_2O, were varying ionic strength, adding organic additives, filtration, varying the D_2O concentration and lastly, evaluating the buffer exchange method.

All nanodiscs were prepared by mixing MSP and POPC in the ratio 1:65. Prior to mixing, POPC was solubilised in a cholate containing buffer. After mixing, cholate was removed and nanodiscs purified by gel filtration. A detailed protocol for nanodisc self-assembly can be found in Appendix 1. Any variation from this protocol is described here.

Varying ionic strength

Nanodiscs were prepared using different conditions for gel filtration, varying pH from 6.5 to 8.5 or NaCl concentration from 10 mM to 500 mM. pHs at 6.5 and 7.0 was prepared in 0.1 M $NaPO_4$, while 7.5, 8.0 and 8.5 was in 20 mM Tris/HCl.

Samples were measured by SAXS, then exchanged to D_2O based buffers with same buffer compositions as in H_2O, using a 10 kDa cut off spin filter (Amicon) and again measured by SAXS.

SAXS data were collected at BM29 at ESRF

Addition of organic solvents

Solvent was exchanged to a D$_2$O based buffer, added different concentrations of deuterated methanol or deuterated acetonitrile, using a 10 kDa cut off spin filter (Amicon). Buffers contained methanol or acetonitrile in concentrations 0, 1, 5 or 10%.

SAXS data were following collected at BM29 at ESRF.

Removing aggregates by filtration

Solvent was exchanged on a nanodisc sample to a D$_2$O based buffer at 20 mM Tris/HCl pH 7.5, 100 mM Nacl using a 10 kDa cut off spin filter (Amicon). Three different samples were prepared, where one did not get any further treatment. A second sample was filtered using a centrifugal filter with 50 nm pore size (Nanosep) centrifuged for 10 minutes at 5000 rpm. The third sample was filtered using a 100 kDa cut off spin filter (Amicon) centrifuged for 30 minutes at 5000 rpm. Data were collected at I911-4 at MaxLab.

Varying D$_2$O concentration

Nanodiscs were concentrated to ~ 15 mg/ml using a 10 kDa cut off spin filter (Amicon) and diluted with H$_2$O and D$_2$O based buffer to D$_2$O concentrations from 50-95% to a nanodisc concentration ~ 1 mg/ml. One sample was exchanged to 100% D$_2$O buffer using a 10 kDa cut off spin filter (Amicon).

SAXS data were following collected at I911-4 at MaxLab.

Method for solvent exchange – spin filtration vs dialysis

Solvent was exchanged on a nanodisc sample to a D$_2$O based buffer at 20 mM Tris/HCl pH 7.5, 100 mM NaCl either by using a 10 kDa cut off spin filter (Amicon) or by dialysis using a 8-10 kDa membrane pore size (Spectra/Por). The samples were measured by SAXS and the dialyzed sample was following centrifuged for 10 min at 5000 RPM and again measured by SAXS. Data were collected at I911-4 at MaxLab.

SAXS

Samples were measured at BM29 or I911-1. In both cases, the SAXS data were converted to scattering intensity units as a function of scattering vector $\vec{q}$, $I(q)$, where $q = \left| \frac{4\pi sin\theta}{\lambda} \right|$

BM29

Scattering data were recorded at the BM29 bioSAXS (Pernot et al, 2013) at ESRF, Grenoble. Samples and buffers were exposed for a total of 10 s each during which, 10 frames were recorded to ensure that the sample was not damaged by the radiation. The experiments were conducted at 10 °C, and the samples were irradiated using a nominal x-ray wavelength of 1 Å. The instrument uses a PILATUS 1M-detector located 2.85 m from the sample; and the tasks of data reduction and treatment were performed using in-house software at the instrument.

I911-4

Scattering data were recorded at the I911-4 beamline at MaxLab (Labrador et al, 2012). Samples were exposed for 120 s using a nominal X-ray wavelength of 0.91 Å. Experiments were conducted at 10°C and data detected using a PILATUS 1M-detector located 1.98 m from the sample. Data reduction and treatment were performed using the BioXtas raw software (Nielsen et al, 2009).

Results and discussion

Varying ionic strength

Changing the ionic strength of a protein solution can contribute to stabilisation or destabilisation of the proteins, depending on the different ionic interactions in the molecules. Simply changing the salt concentration or pH in a solution will affect the ionic interactions. Increasing salt concentration in the solution increases screening of electrostatic interactions (Kohn et al, 1997), while changing pH will change the electrostatic free energy and the net charge of a protein. A protein can hereby be stabilised by changing the charge of ionisable groups, moving the net charge in relation to its isoelectric point (Yang & Honig, 1993). Also the fact that heavy water holds a different neutral pH (pD) could suggest the ideal ionic strength for a protein differs from H_2O and D_2O. For this reason, changing the ionic strength of the D_2O buffer, by varying NaCl concentration of pH, was investigated to evaluate if this had any effect on the nanodisc stability.

Nanodiscs were prepared at different NaCl concentrations at pH 7.5 in H_2O and measured by SAXS. Solvent was changed to D_2O based buffers at same buffer conditions by spin filtration.
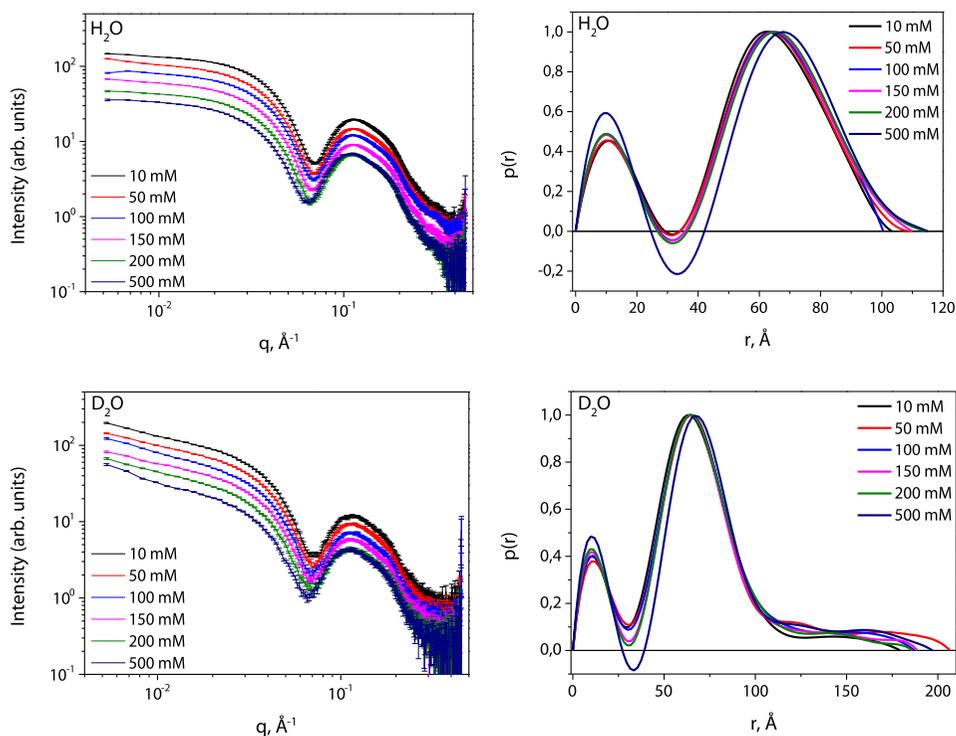
Figure 1. SAXS data and derived *p(r)*-functions of nanodiscs in H₂O (top) and D₂O (bottom) based buffers with varying NaCl concentrations. Left are the obtained scattering data, and corresponding *p(r)*-functions to the right. Data obtained at BM29 ESRF, June 2012.

Judged from the scattering data, seen in Figure 1, varying the NaCl concentration does not affect the aggregation problem in D₂O. All samples in D₂O look aggregated when compared to H₂O samples, which becomes very clear looking at the derived *p(r)*-functions. 500 mM NaCl appears to influence the structure or scattering contrast slightly in both H₂O and D₂O judged from the *p(r)*-functions.

Nanodiscs prepared at different pH with 100 mM NaCl in H₂O was also measured by SAXS. Solvent was changed by spin filtration to D₂O based buffer at the same buffer conditions.
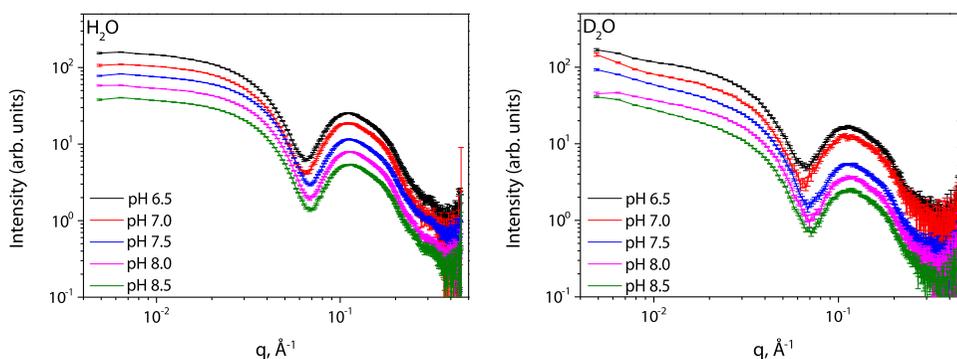
Figure 2. SAXS data of nanodiscs in H₂O (left) and D₂O (right) based buffers with varying pH. Data obtained at BM29 ESRF, June 2012.

Changing pH does not have any effect on the aggregation of the D_2O samples, which still look aggregated compared to the H_2O samples, judged from the scattering data in Figure 2. This is clear from the pronounced upturn in the low-q region. However, using pH 6.5 and 7.0 have some structural effect, as the minima of both the H_2O and D_2O samples are changed slightly toward a lower q. These two samples were prepared in 0.1 M $NaPO_4$ buffer compared to 20 mM Tris for the remaining, which may cause this difference.

The idea with changing the ionic strength was that nanodiscs contained at a certain ionic strength, could hold a higher stability and maybe suffer less from the adverse effects coursed by the surrounding D_2O, though this was unfortunately not observed.

Interestingly it was found, that nanodiscs in a D_2O based buffer subjected to a significant increase in radiation damage as compared to nanodiscs in a H_2O buffer. The X-ray beam had to be attenuated in order to obtain reproducible data frames. Radiation damage in SAXS experiments are observed as irreversible aggregation of protein samples (Kuwamoto et al, 2004). The interaction between X-rays and the solvent water can lead to radiolysis and thereby formation of free radicals which undergo reactions with the proteins at solvent accessible sites (Maleknia et al, 2001). The observed increase in radiation damage may be caused by weaker hydrogen bonds of exchanged deuterium in the sample making the protein more susceptible to reactions with free radicals. Addition of radical scavengers such as dithiothreitol (DTT) to the samples should be considered to decrease this effect.

Addition of organic solvents

Organic solvents are commonly used as additives in protein crystallization and NMR to increase protein stability. In the presence of methanol, hydrogen bond interaction of the protein backbone is strengthened and side chain hydrophobic interactions weakened. This should lead to stabilising the secondary structure and a lower protein-water interaction (Hwang et al, 2011).

Acetonitrile has previously been shown useful as additive in NMR studies for proteins aggregating in water. Acetonitrile should contribute to a destabilisation of quaternary structures while preserving the tertiary protein structure (Bocian et al, 2008).

Adding methanol or acetonitrile in different concentrations were investigated to find if this could have some positive impact on the aggregation problem.

Nanodiscs were exchanged to a D₂O buffer containing different concentrations of acetonitrile or methanol.
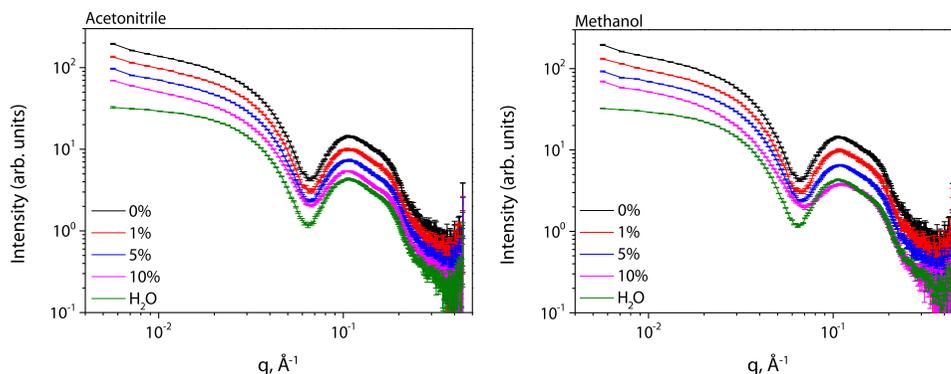
Figure 3. SAXS data of nanodiscs in D₂O added varying concentrations of acetonitrile (left) and methanol (right), compared with a sample in H₂O without additives. Data obtained at BM29 ESRF, July 2012.

No positive effects were observed from the use of any of the two additives, aggregation is still observed in the scattering data shown in Figure 3. Adding methanol was found to affect the structure of the particles, in a concentration dependent manner, as seen on the scattering data in Figure 3. This was observed as a less pronounced minima at higher methanol concentration. This effect is most likely caused by alcohol interacting with the lipid bilayer and not a structural change in the protein. Short chain alcohols have amphiplilic character and interact with lipids. Methanol contains only a single CH_3 group and will interact less with lipids than alcohols with a more CH_2 groups, such as ethanol. Methanol is not able to penetrate the lipid bilayer because of the hydrophobicity of the bilayer but some interaction still occurs. Structural influences such as increase in average lipid area is found for POPC lipids in the presence of methanol (Patra et al, 2006).

Removing aggregates by filtration

Nanodiscs in D₂O buffer was passed through a 50 nm centrifugal filter or filtered through a 100 kDa spin filter in an attempt to move some of the observed aggregates.
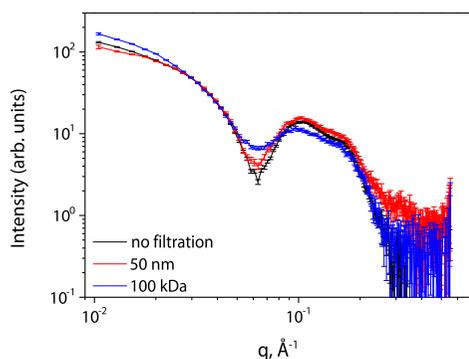
Figure 4. SAXS data of nanodiscs in D₂O filtered through a 50 nm or 100 kDa spin filter, compared to an unfiltered sample. Data obtained at I911-4 MaxLab, October 2012.

Filtering the sample in D_2O through a 50 nm filter makes a slight improvement, as some of the larger particles are removed, judged from the scattering data in Figure 4. However, using a 100 kDa spin filter does not remove any aggregation and does also impact the nanodisc structure. This can be seen as the minima as well as the following "bumps" are less pronounced in the scattering data, compared to the other samples.

As nanodiscs are only ~10 nm in diameter, passing a sample through a 50 nm filter was obviously not sufficient. To remove even more aggregates, the sample was filtered through a 100 kDa spin filter, though this showed to have consequences for the nanodisc structure and did worsen the aggregation as seen in the data in Figure 4. A 100 kDa filter may be too small and a filter with pore size in-between the two applied filters may be attempted.

Varying D₂O concentration

In order to examine how much D_2O needs to be present before this aggregation occurs, a concentration variation series was prepared. Nanodisc samples containing 50-100% D_2O were analysed by SAXS to monitor aggregation levels.

Nanodiscs were concentrated to ~15 mg/ml in H_2O buffer and added H_2O and D_2O buffers to a nanodisc concentration ~1 mg/ml and D_2O concentrations from 50-95%. A sample in 100% D_2O buffer was prepared by spin filtration.
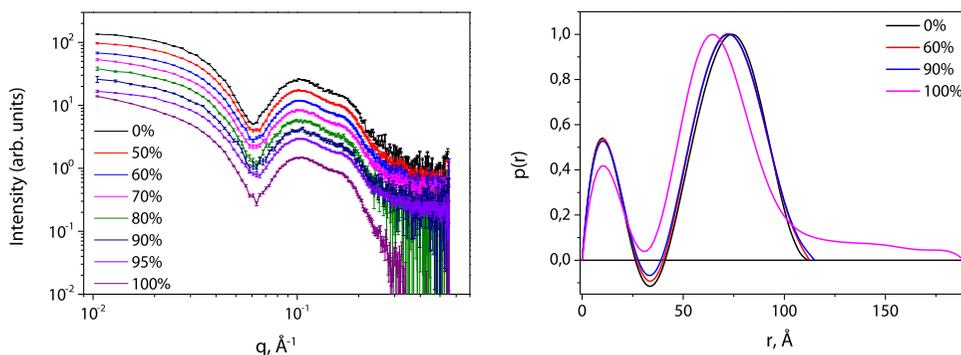
Figure 5. SAXS data of nanodiscs in buffers with varying D₂O concentration. Data obtained at I911-4 MaxLab, October 2012.

Adding up to 60% D_2O does not seem to have any impact on the nanodisc structure, judged from the scattering data in Figure 5. From 70-95% the samples start to look slightly aggregated, judged from the scattering data, while at 100% D_2O, the sample look aggregated as previously observed.

100% D_2O showed significantly more aggregation compared to 95% which may lead to the assumption that proteins only need a small amount of H_2O to keep stabilised. Looking at the

derived *p(r)*-functions in Figure 5, the difference between 100% and the other concentrations are even more striking questioning if this is a consequence of D_2O alone.

The only other difference between the 100% sample and the other concentrations, except for D_2O concentration, was the method used for buffer exchange. While the 50-95% samples were prepared by concentration in H_2O followed by dilution with D_2O to the desired concentration, the 100% sample had been exchanged to D_2O by spin filtration, i.e. concentrated in the presence of D_2O. For this reason the two methods were tested and compared for obtaining a 95% D_2O sample.

Nanodiscs were exchanged to a 95% D_2O buffer by spin filtration or concentrated in H_2O and diluted with D_2O buffer to 95% as in the previous experiment.
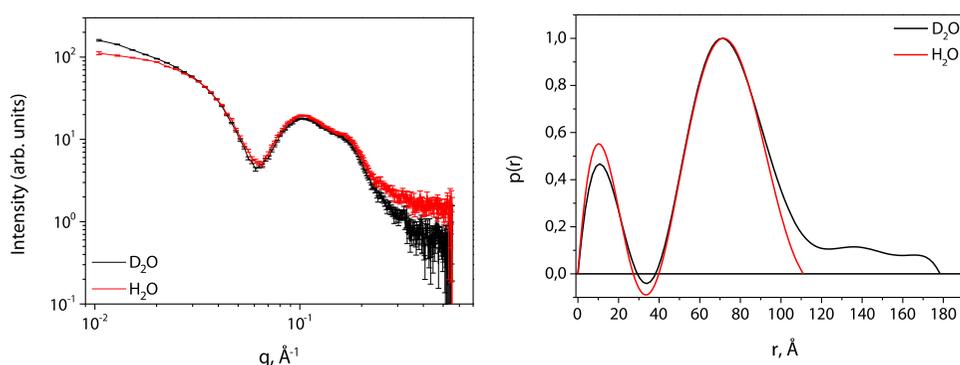
Figure 6. SAXS data of nanodiscs in 95% D_2O. One sample has been prepared by buffer exchange to D_2O using a spin filter (black), while the other was concentrated in H_2O followed by dilution with D_2O (red). Data obtained at I911-4 MaxLab, October 2012.

It appears to have a great impact whether the sample has been concentrated in H_2O or D_2O, which shows on the scattering data and *p(r)*-function in Figure 6. Obtaining a very concentrated sample in H_2O followed by D_2O dilution to 95% lead to considerable less aggregation compared to a sample gone through solvent exchange by spin filtration. A different strategy than spin filtration for solvent exchange to 100% D_2O should for this reason be used.

Method for solvent exchange – spin filtration vs dialysis

Solvent exchange by dialysis was compared to spin filtration as this may act as a milder alternative. Samples in 100% D_2O was prepared by using a 10 kDa cut off spin filter or by dialysis and analysed by SAXS. The dialysed sample was after measurement subjected to centrifuged at 5000 RPM for 10 min to imitate spin filtration conditions and measured again.
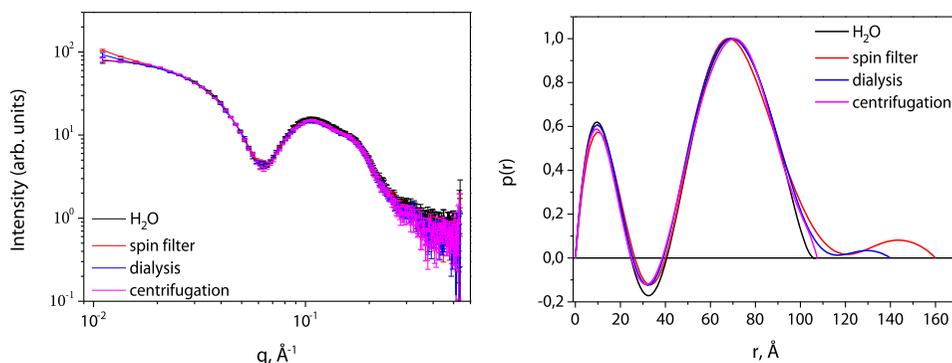
Figure 7. SAXS data (left) and corresponding *p(r)*-functions (right) of nanodiscs in D₂O buffer exchanged by spin filtration (red), dialysis (blue) and dialysis followed by centrifugation (pink), compared to a sample in H₂O (black). Data obtained at I911-4 MaxLab, October 2012.

Exchange of solvent from H_2O to D_2O by dialysis appears to have less impact on the sample compared to using spin filtration, as judged from the scattering data and in Figure 7. Solvent exchange of samples by dialysis caused less aggregation compared to samples prepared by spin filtration. This suggested that the aggregation could somehow be caused by centrifugation of the sample in D_2O.

The dialyzed sample was following centrifuged in an Eppendorf tube at same speed and duration normally used for the spin filters to test if this could cause aggregation. The centrifuged sample showed no further aggregation as can be seen on both scattering data and the derived *p(r)*-functions in Figure 7. The minor aggregation observed after dialysis was even removed by the centrifugation of the sample, showing close to identical scattering pattern of a nanodisc sample in 100% H_2O buffer. However, a less pronounced minimum in the *p(r)*-function of the sample in D_2O compared to that in H_2O suggests some asymmetry in the particles.

Conclusion

Investigating the effect of different factors for decreasing nanodisc aggregation did not lead to any direct solution. Though it could be concluded that spin filters, used for buffer exchange to D_2O based solvent, somehow has a destabilising effect on nanodisc. How this occurs and which interactions are taking place is uncertain but should be examined. Dialysis is clearly a much more gentle method for solvent exchange to D_2O and should be considered.

Protein samples in D_2O solvent are much more subjected to radiation damage from SAXS compared to in H_2O. Caution must be taken by attenuating the beam or considering the addition of radical scavengers to the sample.

From this study it cannot be concluded if varying buffer composition can have a stabilising effect on samples in a D_2O solvent. The effect of changing ionic strength or adding organic solvent should once again be evaluated for dialyzed samples for impact on the long term stability of proteins in D_2O based solvent. Also the D_2O concentration variation should be re-evaluated if adding a small fraction of H_2O can improve long term stability.

Chapter 6 - Reconstitution of membrane proteins in nanodiscs

This chapter will elucidate some sophisticated examples of using nanodiscs for studying membrane proteins. The project is divided into two projects each focusing on a specific membrane protein.

D. Solution structure of magnesium transporter CorA in nanodiscs
E. Activity and inhibition of disulphide bond forming protein DsbB in nanodiscs

Project D presents an investigation of the solution structures of magnesium transporter CorA in its open ligand free state and closed ligand bound state, obtained from SAXS.

Project E concerns the reconstitution of DsbB disulphide bond forming protein in nanodiscs for activity and inhibition assays. This protein is an important player in bacterial toxin folding and therefore an interesting target for inhibition. The project focuses on utilizing the nanodisc for stabilising the protein and retaining protein activity.

Both projects will be presented as paper drafts.

D. Solution structure of magnesium transporter CorA in nanodiscs

Preface

This project has been a collaboration with Martin Cramer Pedersen from Australian National University and Mikaela Rapp from Stockholm University. Martin has been responsible for SAXS data analysis while Mikaela has provided the purified CorA for nanodisc reconstitution.

The following draft has been co-written with Martin Cramer Pedersen and Lise Arleth.

Exploring conformational structures of magnesium transporter CorA embedded in a phospholipid nanodisc using SAXS

Pie Huda[a,1,2], Martin Cramer Pedersen[a,b,1], Hena Sandhu[c], Michaela Rapp[c]
and Lise Arleth[a,2]

[a]Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark

[b]Department of Applied Mathematics, Australian National University, Oliphant Building 60, Mills Road, Canberra ACT 0200, Australia

[c]Department of Biochemistry and Biophysics, Stockholm University, Svante Arrhenius Väg 16C, SE-106 Stockholm, Sweden

[1]These authors contributed equally to this work
[2]Corresponding authors
Pie Huda and Lise Arleth
Universitetsparken 5, DK-2100 Copenhagen, Denmark
Email: huda@nbi.ku.dk, arleth@nbi.ku.dk
Phone: +45 40499247

Abstract

Membrane protein mediated ion transport mechanisms and in particular how cells obtain gradient sensitivity and selectivity towards different cations are not yet well understood. The pentameric CorA is an important and widely studied Mg^{2+} transporter in prokaryotes. The five subunits form a centrally located ion-conducting pore facilitating Mg^{2+} transport. The crystal structure of CorA in the closed, Mg^{2+}-bound conformation became available almost a decade ago and has provided the basis for molecular dynamics based extrapolations to better understand the Mg^{2+}-free open conformation. This recently lead to two rather different hypotheses for the open conformation of CorA: One, where the cytoplasmic part of the pentamer reorganize in a symmetric, iris-opening fashion, and one, where the pentameric subunits are displaced in relation to each other to form an asymmetric bended bell-like shape. Until now, there has not been biophysical data of sufficient structural resolution to critically evaluate these two hypotheses. In the presented work, a recently developed approach combining small-angle scattering and nanodisc was applied to further investigate the structure of CorA under solution conditions. Surprisingly, it was also found that the reconstituted CorA in nanodiscs, have a stoichiometry of four nanodisc belts to one CorA pentamer. This suggests that the CorA acts as a scaffold to form larger discs that what is usually observed. Regarding the solution structure of CorA, it was found that scattering data supported the symmetry-broken, bended bell-conformation of the open, magnesium-free CorA. Furthermore, it could be deduced that magnesium bound CorA in solution holds a conformation different from that of the crystal structure.

Introduction

Magnesium ions are the most abundant divalent cations in living cells and is essential for growth and maintenance of the cell (Lusk et al, 1968). Magnesium serves as regulator for various metabolic enzymes as well as a required cofactor for ATP activity. It is necessary for stabilising structural components such as the ribosomes and the plasma membrane (Moncrief & Maguire, 1999; Moomaw & Maguire, 2008). Magnesium ions are unique compared to other biological relevant cations, as it has the smallest ionic radius, the highest charge density, and the largest hydrated radius, which is 400 times larger than the dehydrated ion. This change is many times greater than for other cations such as Na^+, K^+, and Ca^{2+} and in addition to that, hydration waters are bound 3-4 times more tightly to Mg^{2+} creating an exceptional challenge for Mg^{2+} transporters (Martin, 1990).

The membrane protein CorA is the most abundant Mg^{2+} transporter found in Archaea and Bacteria and is a prokaryotic homologue to the eukaryotic Mrs2-Alr1 Mg^{2+} channel (Moncrief & Maguire, 1999; Payandeh et al, 2013). Three crystal structures from *T. maritima* are available showing CorA is a homopentamer, where each monomer holds two transmembrane helixes oriented in fivefold symmetry about a central pore. This forms a 40 Å long funnel with an interior diameter of 5 Å at the cytosol end increasing to a much wider diameter of 20 Å at the cytoplasmic site. In extension of the funnel, on the cytosol interface, a large intracellular domain is found (Eshaghi et al, 2006; Lunin et al, 2006; Payandeh & Pai, 2006). Given the available crystal structures of CorA are all found in its Mg^{2+}-bound, closed conformation, suggests a more ordered (or less flexible) structure in this conformation than in the ligand-free, open state.

Speculations of the open unbound state conformation have led to different proposals on the structure. One suggests a simple gating mechanism resembling opening and closing of the pore based on a crystal structure of the isolated funnel domain from *Archaeoglobus fulguidus* CorA (Dalmas et al, 2010). MD-simulations of this mechanism gives an iris-like opening when seen from the cytosolic side (Dalmas et al, 2014). Another more surprising proposal for the open structure, also widely based on MD simulations, suggests an asymmetric structure leading to a tilt of individual pore forming helixes influenced by the formation of salt bridges and resulting in shape resembling a bended bell (Pfoh et al, 2012).

Small-angle scattering (SAS) can elucidate the structure of macromolecules in solution with resolution down to about 10 Å. This allows studying proteins in a close-to-natural environment giving users the possibility of obtaining structural information about proteins that are not easily crystallized as well as indicating whether the crystal structures also represent the solution structures or not. Membrane proteins must be reconstituted in a carrier system in order to be compatible with solution studies. It was recently shown that the nanodisc platform (Bayburt et al, 2002; Bayburt & Sligar, 2010) is a very homogenous and optimal carrier system for small-angle X-ray and neutron scattering studies of membrane proteins, and that the methods can be used in combination with advanced modelling to provide detailed structural information about both the nanodiscs and the incorporated membrane proteins (Kynde et al, 2014; Skar-Gislinge et al, 2010)

In the present work, the combined nanodisc and small-angle scattering approach was used to investigate the solution structure of CorA in the presence or absence of Mg^{2+}-ions. In terms of

information content, high quality data were obtained from the system. This allowed for critically investigating the different structural hypotheses for the CorA under close-to-native conditions with respect to the bilayer membrane environment and the solution conditions.

To our surprise, it was found that CorA, when reconstituted with MSP1D1, formed lipid-protein complexes with a stoichiometry of one CorA pentamer to four MSP1D1s. Consequently, the disc diameter was significantly larger than what is normally observed for the discs based on the 8-alphahelical MSP1D1 (Bayburt et al, 2002). This phenomenon has to our knowledge not been previously observed with MSP1D1. However, the MSP1D1 is directly derived from human ApoA1, which is a primary component of HDL/LDL and therefore known to be highly flexible. When CorA was reconstituted with the 11-alphahelical MSP1E3D1 (Denisov et al, 2004), complexes with the expected stoichiometry of one CorA pentamer to two MSP1E3D1 belts were obtained.

Results
CorA can be reconstituted in nanodiscs

Pentameric CorA was reconstituted into POPC nanodiscs and separated from the disc fraction without incorporated CorA through size exclusion chromatography (see Fig. 2 (A)). The heavy fraction of the CorA ND peak, marked in Fig. 2(A), was used for SAXS measurements. Fitting a Gaussian curve to the peak without CorA showed that the fraction used for SAXS measurements should only include insignificant amounts of nanodiscs without CorA.

SDS PAGE indicates that MSP1D1:CorA nanodiscs have a four MSP1D1 to one CorA pentamer stoichiometry

Sample and a known gradient of purified CorA and MSP were run on SDS PAGE and coomasie stained (see Fig. 2 (D)). Quantification of the coomassie stain intensity of each protein band based on the known protein gradients allowed for determinations of the concentrations of MSP and CorA in the sample as well the ratio between these. For the MSP1D1 nanodiscs, the protein ratio of MSP1D1 to CorA pentamer was 3.7 ± 0.6 suggesting that each nanodisc contained one CorA pentamer and four MSPs. This was quite surprising as nanodiscs are normally expected to contain only two MSPs (Bayburt et al, 2002).

A native PAGE would disclose the presence of different disc sizes in the sample and was run on in order to control that the curious finding of four MSP1D1s to one CorA pentamer was not

simply a result of the presence of nanodiscs without CorA. MSP1D1-CorA nanodisc sample was applied on the native gel and compared to a MSP1D1 nanodisc sample without CorA. The gel is shown in Fig. 2 (C) and indicates that there is no visible amount of nanodiscs without CorA in the sample. This supports that the results from protein quantification are purely based on nanodiscs containing CorA. Based on these results, the disc stoichiometry of four MSP1D1 to one CorA pentamer was accepted and applied in the following analysis.

In order to obtain further insight along with an independent control of the experiments, CorA was reconstituted in the slightly larger MSP1E3D1 nanodiscs. Here, a similar SDS PAGE based quantification, showed a protein ratio of 2.0 ± 0.3 in discs without Mg and 2.5 ± 1.1 with Mg. Interestingly, this implied that these CorA containing nanodiscs contain one CorA pentamer and two MSP1E3D1s in agreement with the expected stoichiometry.

POPC contents in the MSP1D1-CorA nanodiscs

The total protein concentration was determined through UV_{280} absorption, the extinction coefficients from MSP1D1 and CorA, and the assumption of a stoichiometry of four MSP1D1s per CorA pentamer (see Table 1). The POPC contents of the samples were determined through a phosphate analysis (Rouser et al, 1966) by exploiting that each POPC contains one phosphate atom. By combining the information about protein concentration and POPC concentration and assuming four MSP1D1's per disc, a CorA loaded MSP1D1-nanodisc was found to contain 198 ± 21 POPC molecules per disc. On the other hand, if the nanodiscs were assumed to hold two MSP1D1s per disc the number would be 174 ± 19 POPC molecules per disc. This is significantly higher than normally found for POPC:MPS1D1 nanodiscs even without membrane protein (Denisov et al, 2004; Skar-Gislinge et al, 2015; Skar-Gislinge et al, 2010). Therefore, it is physically unrealistic that an MSP1D1 nanodisc with two MSPs could contain this amount of lipids as well as the CorA pentamer; again in favour of a structure with four MSPs per disc.

SAXS shows that well-defined MSP1D1-CorA nanodiscs can be prepared

An initial inspection of the SAXS-data presented in Fig. 2 (B) indicated that the sample consisted of well-defined non-aggregated particles with well-defined shape and size; a conclusion, drawn from the presence of Guinier-regions in the low-q part of the data, as well as from the fairly sharp features in the intermediate and high-q part of the data. This claim was further substantiated by the *p(r)*-distributions determined from the SAXS-data using the Indirect

Fourier Transformation (Glatter, 1977; Hansen, 2012) (Fig. 2 (E)). These showed that the CorA-loaded MSP1D1 nanodiscs had sizes in terms of the D_{max} of around 150-170 Å, which is significantly larger that the value of around 120 Å that was previously obtained for bacteriorhodopsin-loaded MSP1D1 nanodiscs (Kynde et al, 2014).

Mg^{2+} introduces a structural transition of CorA visible in SAXS data

A small but significant difference was observed between the SAXS-data from CorA with and without Mg^{2+} (Fig. 2 (B)). This does indeed imply a subtle structural transition the samples in between. We hypothesize that this difference in the recorded scattering patterns is caused by the structural change in CorA induced by the presence of Mg^{2+}-ions. Similar structural transitions could be introduced by introducing the excess $MgCl_2$ (100mM $MgCl_2$) both to the Mg^{2+}-free sample of reconstituted CorA (see SI Fig. 1(A)) and by introducing the 100mM $MgCl_2$ already from the beginning of the reconstitution process (see SI Fig. 1(B)). However, the former strategy led to a weak aggregation of the sample. For this reason, our structural analysis was based on the data obtained using the latter strategy.

Analysis of the SAXS data indicates that nanodiscs with four MSP1D1's per CorA are formed and suggests that CorA without Mg^{2+} is in a bended bell-conformation.

As an investigation of the surprising finding of four MSP1D1's per CorA containing nanodisc, two different models were refined from the SAXS data from the Mg^{2+} free CorA embedded in the nanodisc: one, where the nanodisc was assumed to contain two MSP1D1, and one, wherein the nanodisc contained four MSP1D1. Some controversy can be found in the literature on the topic of the conformation of CorA in absence of Mg^{2+} (Dalmas et al, 2010; Pfoh et al, 2012) and in particular whether or not the CorA is in a symmetry broken state without Mg^{2+}. In line with recent MD predictions (Pfoh et al, 2012), it was assumed that CorA is in the symmetry-broken state in the absence of Mg^{2+}, the so-called bended bell-conformation (see Fig. 1(B)). The fits and the refined models are shown in Fig. 3. The obtained results for the fit-parameters are listed in Table 2. The fit-quality is excellent for both the model with two and four MSP1D1 (see fig. 3). However, physically realistic fit-parameters can only be obtained for the model with four MSP1D1 per disc. In particular, the number of POPC per disc is found to be 262 per disc in the model with two MSP1D1. This should be compared to the value of around 130 POPC per disc in the case of empty POPC-MSP1D1 discs and is clearly not realistic. Based on this in

combination with the above findings, the model with two MSP1D1 per CorA containing disc was finally ruled out.

The data for the open conformation (i.e. Mg^{2+} free) CorA were also attempted fitted with the crystal structure of CorA in its so-called closed Mg^{2+} bound conformation (PDB entry 2BBJ, (Lunin et al, 2006)), see Fig. 1(A). The rationale for this was that recent suggestions for the open conformation CorA, the Coiled-coil mutant [4EED] by Pfoh et al (Pfoh et al, 2012), the CryoEM structure [4CY4] by Cleverly et al (Cleverley et al, 2015) along with the recent iris-opening model (Dalmas et al, 2014) based on EPR and MD all bear very great structural resemblance with this closed conformation structure, in particular within the resolution of SAXS. It was not possible to obtain physically reasonable fits with this symmetric upright model for CorA (data not shown). A large mismatch between the lipid bilayer and the hydrophobic parts of the CorA was obtained, and fits with a low χ^2 could only be achieved if the CorA was allowed to tilt unphysically in relation to the bilayer plane (data not shown).

From this analysis, it was deduced that the bended bell-conformation is in significantly better agreement with the obtained SAXS-data than the above mentioned suggestions for five-fold symmetric Mg^{2+}-free CorA structures.

Investigation of the Mg^{2+} induced structural transition

Similarly, the model for the Mg^{2+}-bound CorA crystal structure was refined against the SAXS data from the Mg^{2+}-bound state. Based on the evidence above, only the case of four MSP1D1s per nanodisc was considered. Interestingly, it was not possible to obtain physically meaningful fits to the SAXS-data based on this structure (see figs 4 (A) and (B)). While the overall fit quality is good, the values for the fit-parameters appear unphysical and in particular it is only possible to fit the data if the pentamer is allowed a significant tilt, hence leaving major parts of the hydrophobic domains exposed to water. Interestingly, a much better but still not perfect model fit could be obtained based on the symmetry-broken bended bell-conformation which should otherwise only be applicable to the Mg^{2+}-free conformation (see figs 4 (C) and (D)). This disagreement between the crystal structure of CorA and the obtained SAXS data, suggests that the crystal structure for Mg^{2+}-bound CorA does not provide a complete description of its solution structure.

Cross-check of CorA reconstitution and sample homogeneity with size-exclusion-chromatography SAXS

As a final check, CorA was reconstituted in MSP1E3D1-POPC discs and analysed by size-exclusion-chromatography SAXS. This provided an additional control of the sample homogeneity and allowed for investigating whether discs based on two MSPs could be prepared with these extended MSP's. Data are plotted in Fig. 5. SEC-profiles obtained in the absence and presence of Mg^{2+} (Fig. 5 (A) and (B)) indicate high sample homogeneities with only a single species. SAXS-data were acquired based on a scan across the SEC-peak (Fig. 5 (C) and (D)). It is clearly seen that the overall structure of the SAXS data changes only very little over the SEC peak. Guinier-regions can be identified at the low-q part of all SAXS-data sets, as well as distinct features in the medium-q values. In combination, the data indicate that the samples are mono-disperse and well-defined.

As for the data obtained from the MSP1D1-based CorA discs (Fig. 2 (B)), a slight difference is observed between the samples with and without magnesium, specifically in the I-range from 0.05 Å^{-1} to 0.1 Å^{-1}. Once again, this suggests that the presence of magnesium induces a structural transition in the sample.

Indirect Fourier Transformation analysis indicate that the size of the particles, in terms of the largest intra-particle distance in the sample, D_{max}, decreases only slightly (Fig. 5 (E) and (F)), as we scan across the SEC-peak. This confirms the previous observation of well-defined particles but show that both samples contain slightly differently sized particles. This structural dispersity is probably related to a smaller intrinsic variation of the lipid-MSP contents within the sample.

Discussion

The crystal structure of CorA in its magnesium-bound conformation has been known for nearly a decade, and several close to identical structures are available (Eshaghi et al, 2006; Lunin et al, 2006; Payandeh & Pai, 2006). However, the structure of the unbound, open conformation of CorA has not yet been solved. The general agreement is that the open conformation is a more flexible structure while the binding of magnesium stabilises the structure, which in turn enables crystallization (Palombo et al, 2012; Pfoh et al, 2012). Hence CorA is an excellent example of the need for developments of complementary solution-based structural techniques to elucidate structures including the general validity of crystallographic conformations in solution.

Combining the nanodisc-platform and SAXS, the structure of CorA was evaluated in its open (Mg^{2+}-free) and closed (Mg^{2+}-bound) conformations. The data from the open conformation were in agreement with a recent MD-based suggestion for a symmetry-broken bended bell-conformation of CorA (Pfoh et al, 2012), but not with a structure resembling the different upright symmetric conformations also suggested for the open state (Lunin et al, 2006). Interestingly, the SAXS-data also suggested that the crystal structure of the closed magnesium-bound CorA did not correspond to the solution conformation. Actually, a better agreement was found between the SAXS-data and the bended bell-conformation suggested for the open conformation. This is highly interesting and suggests that the crystal conformation is not the dominating state in solution. While an upright fivefold symmetry conformation of CorA might be a criteria for crystallization, the data suggest that a larger flexibility or even a breaking of symmetry is present in solution and therefore possibly also when CorA is in its functional state. Hence, while protein crystallography is essential for obtaining the high-resolution structure of a protein, it is also highly relevant to complement crystallography with solution data such as those obtained by SAXS. This consideration of flexibility is probably particularly relevant for membrane proteins with large soluble domains as CorA.

The modelling in this paper was performed based on what we believe is the current state-of-the-art approach to this type of problems. However, several interesting options come to mind, which would improve the overall framework for this type of refinement – i.e. rather than rejecting models based on visual (and thus inevitably subjective) quantization of hydrophobic mismatches or disagreement between the refined parameters and chemical analysis such as the phosphate analysis conducted in this study, one could include this as *a priori* knowledge in a more elaborate Bayesian optimisation framework. This would allow for a more complete and concise refinement taking all the presented evidence and data into account during the refinement.

In this study, the structure from a molecular dynamics simulation was adopted as a likely solution structure of CorA, yet other techniques might be equally feasible, when it comes to the task of producing probable candidates for solution structures of non-crystallisable proteins, in particular CryoEM and NMR are obvious candidates for such studies. However, the lower limit for CryoEM presently lies at around 170 kDa (Lu et al, 2014), while the upper limit for NMR presently lies around 30 kDa. However, with the recent advances made within particularly

CryoEM, it may be possible to obtain high quality data from a larger range of protein sizes with these techniques in the not so distant future.

Currently, the criteria for submissions and acceptance of a structure into the Protein Data Bank (Bernstein et al, 1977) excludes structures proposed from molecular dynamics simulations, small-angle scattering, and chemical determination studies based on e.g. cross-linking and FRET. However, in the phase of developing a hypothesis for an unknown structure, these structures are highly relevant, and we emphasise that a forum for sharing these structures would be highly desired and relevant.

Interestingly, it was found that the nanodiscs with CorA contained four and not the usual two MSP's when using the MSP1D1 variant. This has not been observed earlier and may imply that nanodiscs are more flexible to accommodate the needs of membrane proteins than expected (Bayburt et al, 2002). In this context, it might be more prudent to think of nanodiscs as capable of forming larger supporting structures in order to accommodate larger membrane proteins unable to fit in a nanodisc with 2 MSPs – an effect not unlike the one observed with so-called peptide discs (Midtgaard et al, 2014) yet likely in a more constrained manner. In an recent study of a large tetrameric aquaporin membrane protein, it was observed that it was not possible to carry out the reconstitution with the MSP1D1 variant, whereas using the slightly larger MSP1E3D1 variant was possible (Järvå et al, 2015). This suggests that despite the flexibility demonstrated by the CorA study, there are still size limitations.

As mentioned, the literature holds several proposals for the structure of CorA in the absence of Mg, which has yet to be crystallised (Dalmas et al, 2014; Pfoh et al, 2012). While this article focuses on the differences between the two structures in Fig. 1, the study and applied methodology are easily expanded to include any other relevant structures of CorA. As a final comment, we will emphasise the necessity of state-of-the-art bio-laboratories present at the large-scale facilities necessary to carry out this type of experiment: the larger a part of the sample production and purification that can be done on-site, the better the chances for a successful experiment. When working with fragile biological systems such as membrane proteins, stability is a constant concern, and sample transportation and unnecessary delays should be kept to an absolute minimum, as this enhances the risk of hampered data quality.

Conclusion

CorA was successfully reconstituted in MSP1D1 based phospholipid bilayer nanodiscs and investigated using SAXS and a range of supplementary techniques. It was found that the conformation of a recently proposed MD-simulation for the Mg^{2+}-free open structure of CorA is in agreement with the obtained SAXS-data (Pfoh et al, 2012). Following this result, it was surprisingly found that the well-established crystal structure for CorA was unable to fit the SAXS data from a sample of Mg^{2+}-bound CorA in nanodiscs. This suggested that also Mg^{2+}-bound CorA has a much higher flexibility in solution than suggested by its crystal structure. This flexibility may be the key to better understand the functioning of CorA.

Together with a quantification of the protein-ratio, the analysis also showed that the nanodisc stabilising the CorA-pentamer in solution is composed of four MSP1D1-proteins rather than the two, usually observed for the nanodisc. When MSP1E3D1 was used in the reconstitution process, CorA could be reconstituted with the usual two MSPs per nanodisc as was demonstrated from a SECSAXS analysis.

Materials and methods
Cloning

The *CorA* gene from *Thermotoga maritima* (TmCorA) was cloned with an N-terminal 6xHis affinity tag and TEV protease cleavage site (MHHHHHHENLYFQGM) into pGFPi (a modified version of the pGFPe expression vector (Daley et al, 2005; Rapp et al, 2004). The coding region was under the control of the T7 promoter. Primers were obtained from (MWG, Germany). Cloning was performed in MACH1 cells. The coding region of the his-TEV-TmCorA plasmid was confirmed by sequencing (MWG, Germany).

Protein purification

MSP1D1 and MSP1E3D1 were expressed and purified as described in (Ritchie et al, 2009). CorA protein expression and purification were performed as described by Palombo et al. (Palombo et al, 2012) with some adjustments. The his-TEV-TmCorA plasmid was transformed into the *E. coli* strain BL21(DE3) and grown at 37 °C in Terrific broth (TB). When OD_{600} 0.8 was reached, protein expression was induced by the addition of IPTG (0.4 mM) and grown over night at 20 °C.

In the morning, cells were harvested and resuspended in lysis buffer [20mM Tris-HCl pH 7.9, 0.3 M NaCl with EDTA-free Complete Protease Inhibitor Cocktail (Roche), Benzonase Nuclease (Sigma), and Lysozyme from hen egg white (Boehringer Mannheim)]. All subsequent steps were carried out at 4 °C. Cells were broken by sonication. Unbroken cells were removed by centrifugation at 9300g for 20 minutes, and membranes were isolated by centrifugation at 142000g for 35 minutes. Membranes were washed in 20mM Tris-HCl pH 7.9, 0.3 M NaCl.

Membrane proteins were solubilised with 1% n-Dodecyl-β-D-Maltopyranoside (DDM) (Affymetrix) in 7 ml solubilisation buffer [20 mM Tris-HCl pH 8, 0.3 M NaCl, and 20 mM imidazole] for 1 hour. Insolubilised material was removed by centrifugation at 142000g for 30 minutes.

The supernatant containing the solubilised proteins was applied to 2 ml Ni-NTA Agarose (Invitrogen) and incubated for 30 minutes to allow for His-tag binding to the resin. The column was washed with wash buffer [20 mM Tris-HCl pH 8, 0.3M NaCl, 40 mM imidazole, and 0.03% (w/v) DDM] before elution with 3 ml elution buffer [20 mM Tris-HCl pH 8, 0.1 M NaCl, 0.3 M imidazole, and 0.03% (w/v) DDM]. The elution buffer was changed to gel filtration buffer [20 mM Tris-HCl pH 8, 0.3 M NaCl, and 0.03% (w/v) DDM] using pre-packed PD-10 columns (GE Healthcare). Amicon Ultra Centrifugal Filters 30K (Millipore) were used to concentrate the proteins to approximately 5 mg/ml.

The purity and integrity of the concentrated proteins were verified by denaturing PAGE (NuPAGE 4-12% Bis-Tris Gels (Invitrogen)). The protein was snap frozen in liquid nitrogen and stored at -80 °C.

Reconstitution in nanodiscs

1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine (POPC) (Avanti Polar Lipids) was solubilised in 20 M Tris/HCl pH 7.5, 0.1 M NaCl and 0.1 M sodium cholate to a lipid concentration of 50 mM. CorA was mixed with lipids in a 1:650 molar ratio for MSP1D1 and 1:1300 molar ratio for MSP1E3D1 and left for incubation at 4 °C in approximately 10 minutes.

MSP was added the mixture in a 1:650:4 molar ratio for MSP1D1 and 1:1300:4 for MSP1E3D1 and left for incubation in approximately 30 minutes at 4 °C. This was added approximately 0.5

ml amberlite XAD-2 (SULPECO) to 1 ml sample and left for incubation and continuous stirring for approximately 3 hours at 4 °C. Amberlites were removed by centrifugation and filtration.

The sample was applied to a Superdex 200 10/300 GL (GE) column using 20 mM Tris/HCl pH 7.5, 0.3 M NaCl, and 5 mM EDTA as eluent at a flow rate of 0.5 ml/minute.

Likewise, CorA was incubated in 0.1 M MgCl for 1 hour and mixed with lipids as described above. However, for gel filtration, 20 mM Tris/HCl pH 7.5, 0.3 M NaCl, and 0.1 M MgCl was used as eluent.

Amicon Ultra Centrifugal Filters 30K (Millipore) was used to concentrate the purified nanodiscs and the absorption was measured at 280 nm.

Quantification of MSP:CorA ratio

MSP and CorA were loaded on a 14% SDS PAGE gel in known concentrations of 0.1-1 μg followed by 1 and 2 μl of the nanodisc sample. All wells were added a total volume of 3 μl including 1 μl loading buffer. The gel was made in triplicates and run at 150 V for 90 min.

The gels were stained in a coomassie staining solution and destained with 30% ethanol and 10% acetic acid in deionized water.

Gels were analysed using a Typhoon Trio variable mode imager (GE). Each band was analysed by colour intensity using the standard curve with known concentrations of MSP and CorA for protein quantification of the nanodisc samples.

Native PAGE

Nanodiscs with CorA were loaded on a 4.5% native gel together with empty nanodiscs. Each well was added 2 μl of sample and 2 μl DTT-free loading buffer and run in a DTT-free running buffer at 150 V for 1 hour and 45 minutes. The gels were stained in coomassie staining solution and destained with 30% ethanol and 10% acetic acid in deionized water.

Phosphate analysis

A standard concentration series from Na_2HPO_4 was prepared from 0-150 nM in glass tubes. 400 μl 72% perchloric acid was added to each tube, subsequently cooked for 30 minutes at 180 °C.

After cooling to room temperature, 4 ml of molybdate reagent [2.2 g $(NH_4)_6Mo_7O_{24} \cdot 4 H_2O$ in 14.3 ml conc. H_2SO_4 in 1 l H_2O] were added to each tube and mixed thoroughly. Each tube was added 500 μl ascorbic acid (10% w/v) and cooked for 10 minutes at 80 ˚C. The tubes were cooled in ice water and absolute concentrations of phosphate were obtained through measurements of the absorption at 812 nm.

Likewise, 10 μl of CorA nanodisc sample was treated as the concentration series and the number of lipids was determined on the basis of the absorptions from the concentration series. All solutions were made in duplicates.

Sample preparation – SAXS

Before injection in the SAXS-capillary, the protein concentrations of the samples were assessed using a NanoDrop UV-spectrophotometer with a wavelength of 280 nm. Absorptions of 1.65 and 1.30 were recorded for the samples without and with Mg, respectively. The path length of this instrument is 1 cm, and we used 254320 $M^{-1}cm^{-1}$ and 21430 $M^{-1}cm^{-1}$ as extinction coefficients for the CorA-pentamer and for MSP1D1, respectively.

SAXS

The SAXS-data shown in Fig. 2 were recorded at BM29 (Pernot et al, 2013) at ESRF, Grenoble. Samples and buffers were exposed for 10 s each during which, 10 frames were recorded to ensure that the sample was not damaged by the radiation. The experiments were conducted at 10 °C, and the samples were irradiated using a nominal x-ray wavelength of 1 Å. The instrument uses a PILATUS 1M-detector located 2.87 m from the sample; and the tasks of data reduction and treatment were performed using in-house software at the instrument versus a standard of bovine serum albumin.

The WillItFit framework (Pedersen et al, 2013) and models based on the one introduced in Kynde et al. (Kynde et al, 2014) (however with CorA instead of bacteriorhodopsin) were refined from the presented sets of data, and the resulting structures are shown in Fig. 3 and Fig. 4. As shown in Fig. 3, models were refined models assuming two or four MSP1D1-proteins per disc in order to assess the impact of each of these assumptions.

As described in the literature (Skar-Gislinge & Arleth, 2011; Skar-Gislinge et al, 2010), the nanodisc can be described by a series of form factor amplitudes describing the scattering from

cylinders with elliptic cross-sections (Pedersen, 1997). The scattering from the phospholipid bilayer is described by five continuous slabs describing the headgroups, the CH_2-groups and the CH_3-tails of the molecules. The scattering from the belt of MSP1D1 is described by an elliptical hollow cylinder.

The N-terminal domains of the MSP1D1 were modelled as a statistical average of Gaussian random coils as done e.g. by Skar-Gislinge et al. (Pedersen, 1997; Skar-Gislinge et al, 2010). The values of their radii of gyration were fixed at 12.2 Å in line with the studies of Kohn et al. (Kohn et al, 2004) during the analysis.

The scattering contribution from CorA was calculated by coarse-graining one of the high-resolution structures to amino acid residue level. The form factor amplitudes of these as well of those of the individual parts of the nanodisc were expanded on a basis of spherical harmonics, from which the scattering intensity is easily calculated – see Kynde et al. (Kynde et al, 2014) for details.

The refined parameter values along with estimates of the statistical uncertainty, with which they have been refined, are shown seen in Table 2. We refer to the literature for descriptions of the methods behind these estimates (Pedersen et al, 2014).

The source code for the applied software and the described model can be found online as described by Pedersen et al. (Pedersen et al, 2013).

Sample preparation – SECSAXS

Samples were concentrated to an absorbance of approximately 2.0 using Amicon Ultra Centrifugal Filters 30K (Millipore) as measured by a NanoDrop UV-spectrophotometer with a wavelength of 280 nm.

The sample was then applied on a BioSEC3 (Aligent) column using 20 mM Tris/HCl, 0.3 M NaCl and 20 mM Tris/HCl, 0.3 M NaCl, 0.1 M $MgCl_2$ as eluent for the sample without and with Mg^{2+}, respectively and a flow rate of 0.2 ml/minute.

SECSAXS

In Fig. 5, data recorded at the SECSAXS-beam-line SWING (14) at Synchrotron SOLEIL, Paris, are shown. The eluate from the aforementioned column was irradiated directly using a nominal X-ray wavelength of 1 Å. The instrument uses an AVIEX-detector, which was

positioned 1.79 m or 3.59 m from the sample in these experiments for the sample without and with Mg2+, respectively.

Once again, the data reduction and treatment were done using software readily available at the beam-line. As indicated in Fig. 5 (A) and (B), the capillary is continuously irradiated as the eluted sample flows through. Frames were collected every 2 s during exposure.

Software

The presented data analysis was performed in the WillItFit-framework presented by Pedersen et al. (Pedersen et al, 2013) and the BayesApp-webpage by Hansen (Hansen, 2012). The figures in this manuscript were created using Gnuplot (Williams & Kelley, 2015) and PyMOL (Schrödinger, 2015).

References

Bayburt TH, Grinkova YV, Sligar SG (2002) Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Letters* **2:** 853-856

Bayburt TH, Sligar SG (2010) Membrane protein assembly into Nanodiscs. *Febs Letters* **584:** 1721-1727

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) PROTEIN DATA BANK - COMPUTER-BASED ARCHIVAL FILE FOR MACROMOLECULAR STRUCTURES. *Journal of Molecular Biology* **112:** 535-542

Cleverley RM, Kean J, Shintre CA, Baldock C, Derrick JP, Ford RC, Prince SM (2015) The Cryo-Electron Microscopy Structure of the Cora Channel from Methanocaldococcus Jannaschii in Low Magnesium Conditions. *To be published*

Daley DO, Rapp M, Granseth E, Melen K, Drew D, von Heijne G (2005) Global topology analysis of the Escherichia coli inner membrane proteome. *Science* **308:** 1321-1323

Dalmas O, Cuello LG, Jogini V, Cortes DM, Roux B, Perozo E (2010) Structural Dynamics of the Magnesium-Bound Conformation of CorA in a Lipid Bilayer. *Structure* **18:** 868-878

Dalmas O, Sompornpisut P, Bezanilla F, Perozo E (2014) Molecular mechanism of Mg2+-dependent gating in CorA. *Nature Communications* **5**

Denisov IG, Grinkova YV, Lazarides AA, Sligar SG (2004) Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *Journal of the American Chemical Society* **126:** 3477-3487

Eshaghi S, Niegowski D, Kohl A, Molina DM, Lesley SA, Nordlund P (2006) Crystal structure of a divalent metal ion transporter CorA at 2.9 angstrom resolution. *Science* **313:** 354-357

Glatter O (1977) A new method for the evaluation of small-angle scattering data. *Journal of Applied Crystallography* **10:** 415-421

Hansen S (2000) Bayesian estimation of hyperparameters for indirect Fourier transformation in small-angle scattering. *Journal of Applied Crystallography* **33:** 1415-1421

Hansen S (2012) BayesApp: a web site for indirect transformation of small-angle scattering data. *Journal of Applied Crystallography* **45:** 566-567

Järvå M, Kynde S, Huda P, Törnroth-Horsefield S, Arleth L (2015) SAXS and SANS investigation of SoPIP2;1 Aquaporin-tetramers in POPC-nanodiscs. Bench-mark of the nanodisc approach to extract structural information about membrane proteins *Submitted*

Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiyagarajan P, Sosnick TR, Hasan MZ, Pande VS, Ruczinski I, Doniach S, Plaxco KW (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 12491-12496

Kynde SAR, Skar-Gislinge N, Pedersen MC, Midtgaard SR, Simonsen JB, Schweins R, Mortensen K, Arleth L (2014) Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment. *Acta Crystallographica Section D-Biological Crystallography* **70:** 371-383

Lu P, Bai X-c, Ma D, Xie T, Yan C, Sun L, Yang G, Zhao Y, Zhou R, Scheres SHW, Shi Y (2014) Three-dimensional structure of human gamma-secretase. *Nature* **512:** 166-170

Lunin VV, Dobrovetsky E, Khutoreskaya G, Zhang RG, Joachimiak A, Doyle DA, Bochkarev A, Maguire ME, Edwards AM, Koth CM (2006) Crystal structure of the CorA Mg2+ transporter. *Nature* **440:** 833-837

Lusk JE, Williams RJP, Kennedy EP (1968) MAGNESIUM AND GROWTH OF ESCHERICHIA COLI. *J Biol Chem* **243:** 2618-2624

Martin RB (1990) BIOINORGANIC CHEMISTRY OF MAGNESIUM. *Metal Ions in Biological Systems* **26:** 1-13

Midtgaard SR, Pedersen MC, Kirkensgaard JJK, Sorensen KK, Mortensen K, Jensen KJ, Arleth L (2014) Self-assembling peptides form nanodiscs that stabilize membrane proteins. *Soft Matter* **10:** 738-752

Moncrief MBC, Maguire ME (1999) Magnesium transport in prokaryotes. *Journal of Biological Inorganic Chemistry* **4:** 523-527

Moomaw AS, Maguire ME (2008) The Unique Nature of Mg2+ Channels. *Physiology* **23:** 275-285

Palombo I, Daley DO, Rapp M (2012) The Periplasmic Loop Provides Stability to the Open State of the CorA Magnesium Channel. *J Biol Chem* **287:** 27547-27555

Payandeh J, Pai EF (2006) A structural basis for Mg2+ homeostasis and the CorA translocation cycle. *Embo J* **25:** 3762-3773

Payandeh J, Pfoh R, Pai EF (2013) The structure and regulation of magnesium selective ion channels. *Biochim Biophys Acta-Biomembr* **1828:** 2778-2792

Pedersen JS (1997) Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting. *Advances in Colloid and Interface Science* **70:** 171-210

Pedersen MC, Arleth L, Mortensen K (2013) WillItFit: a framework for fitting of constrained models to small-angle scattering data. *Journal of Applied Crystallography* **46:** 1894-1898

Pedersen MC, Hansen SL, Markussen B, Arleth L, Mortensen K (2014) Quantification of the information in small-angle scattering data. *Journal of Applied Crystallography* **47:** 2000-2010

Pernot P, Round A, Barrett R, Antolinos ADM, Gobbo A, Gordon E, Huet J, Kieffer J, Lentini M, Mattenet M, Morawe C, Mueller-Dieckmann C, Ohlsson S, Schmid W, Surr J, Theveneau P,

Zerrad L, McSweeney S (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *Journal of Synchrotron Radiation* **20:** 660-664

Pfoh R, Li A, Chakrabarti N, Payandeh J, Pomes R, Pai EF (2012) Structural asymmetry in the magnesium channel CorA points to sequential allosteric regulation. *Proceedings of the National Academy of Sciences of the United States of America* **109:** 18809-18814

Rapp M, Drew D, Daley DO, Nilsson J, Carvalho T, Melen K, De Gier JW, Von Heijne G (2004) Experimentally based topology models for E-coli inner membrane proteins. *Protein Science* **13:** 937-945

Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, Sligar SG (2009) *Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs*, SAN DIEGO: ELSEVIER ACADEMIC PRESS INC.

Rouser G, Siakotos AN, Fleische.S (1966) QUANTITATIVE ANALYSIS OF PHOSPHOLIPIDS BY THIN-LAYER CHROMATOGRAPHY AND PHOSPHORUS ANALYSIS OF SPOTS. *Lipids* **1:** 85-&

Schrödinger L. (2015) The PyMOL Molecular Graphics System. pymol.org.

Skar-Gislinge N, Arleth L (2011) Small-angle scattering from phospholipid nanodiscs: derivation and refinement of a molecular constrained analytical model form factor. *Physical Chemistry Chemical Physics* **13:** 3161-3170

Skar-Gislinge N, Høiberg-Nielsen R, Arleth L (2015) What determines the shape and stoichiometry of self-assembled phospholipid nanodiscs: Speed of self-assembly process, initial lipid:MSP stoichiometry or detergent type? *In preparation*

Skar-Gislinge N, Simonsen JB, Mortensen K, Feidenhans'l R, Sligar SG, Moller BL, Bjornholm T, Arleth L (2010) Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. *Journal of the American Chemical Society* **132:** 13713-13722

Williams T, Kelley C. (2015) Gnuplot: an interactive plotting program. gnuplot.sourceforge.net.

Figures
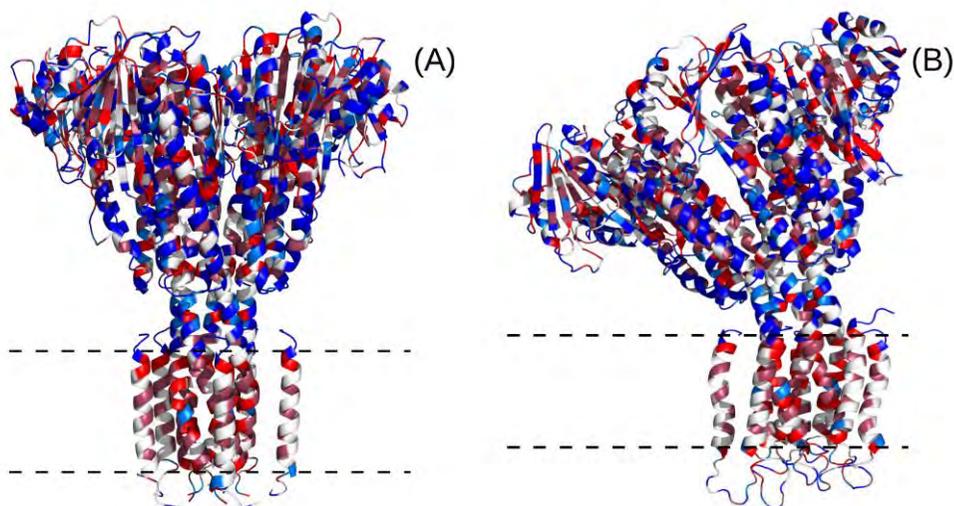
Fig. 1. In (A), the structure of the Mg-bound state of the membrane protein CorA published by Lunin et al. (Lunin et al, 2006) is shown. The structure was refined from x-ray protein crystallography data and can be found in the Protein Data Bank (Bernstein et al, 1977) as structure 2BBJ. In (B), the structure of the same protein in the absence of Mg is shown as published by Pfoh et al. (Pfoh et al, 2012). This structure was produced using molecular dynamics simulations. The structures are coloured by hydrophobicity: red corresponds to a hydrophobic residue, whereas blue corresponds to a hydrophilic residue. Rough estimates of the membrane-spanning regions of the structures are shown with dashed lines.
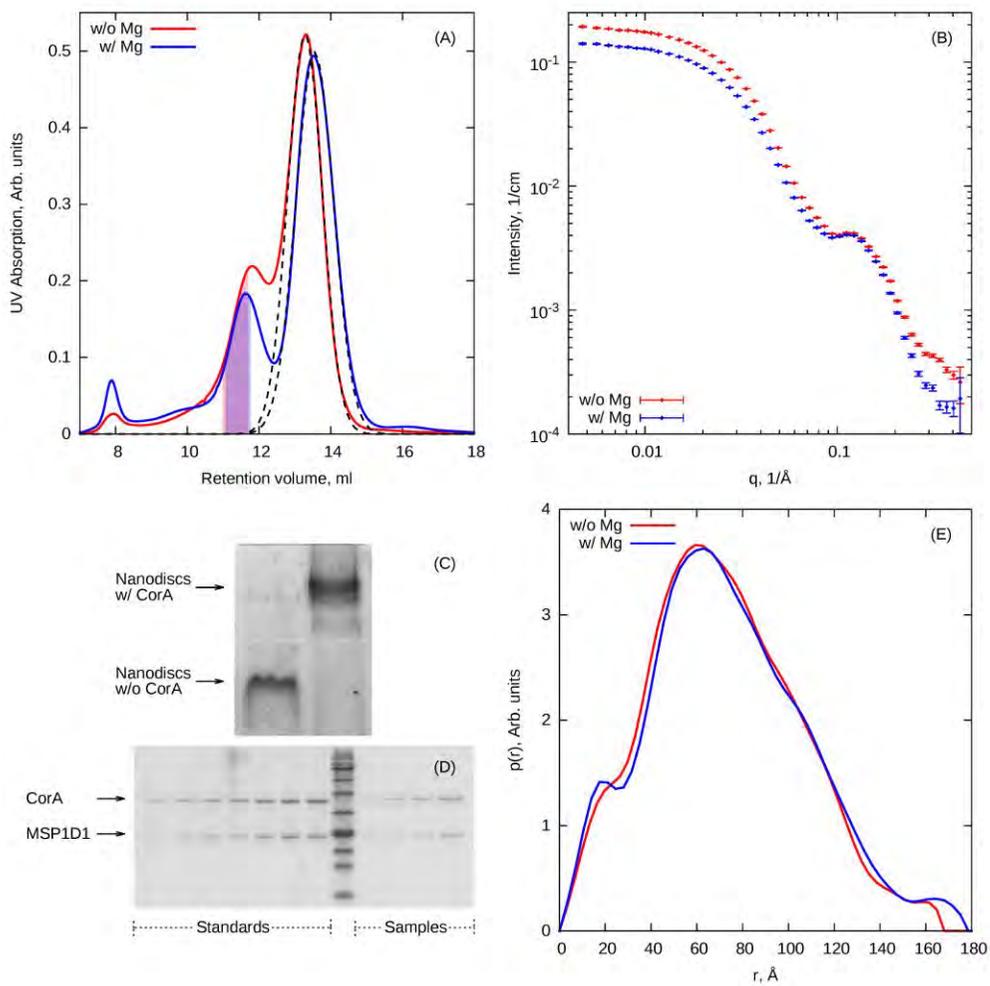
Fig. 2. In (A), full lines: SEC-data from the purification of CorA incorporated in nanodiscs (with and without Mg²⁺). Dashed: Fit of a Gaussian the empty nanodisc peak. Collected CorA loaded fractions highlighted in opaque blue and red. (B), SAXS-data from CorA incorporated nanodisc (with and without Mg²⁺). (C), native PAGE of nanodiscs with and without incorporated CorA. (D), SDS-PAGE of CorA and MSP standards in known concentrations, a protein standard, and CorA loaded nanodisc samples. (E), *p(r)*-distributions determined from the SAXS data in (B).
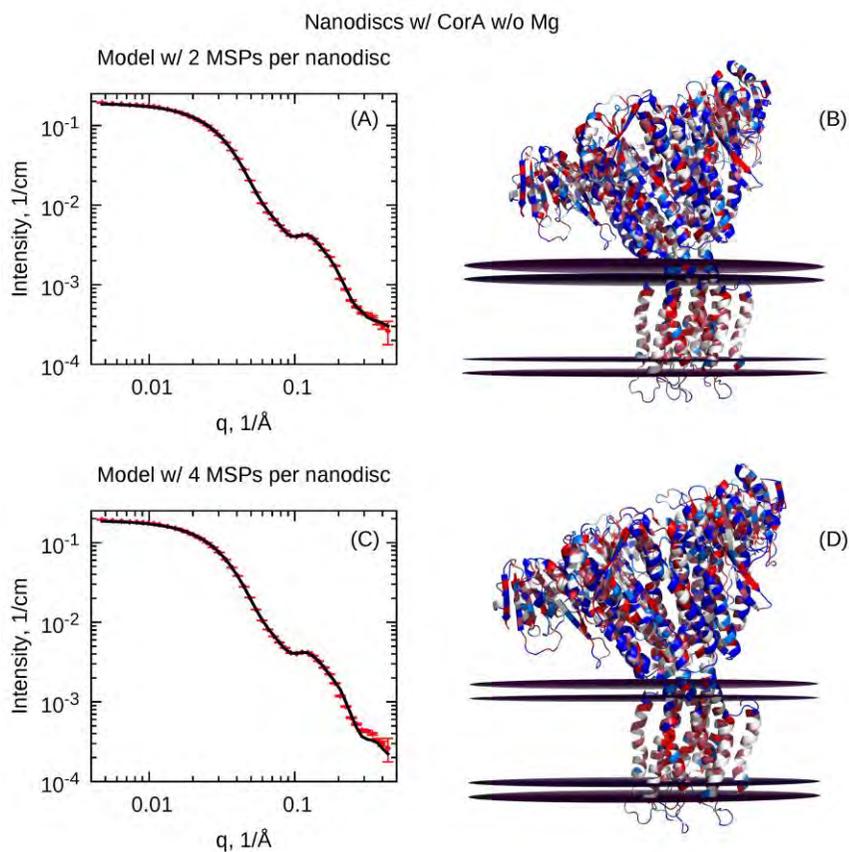
Fig. 3. In (A), SAXS-data from the sample without magnesium are shown along with a fit of a model based on the assumption of two MSP1D1s per nanodisc. For CorA, we use the symmetry-broken structure introduced by Pfoh et al. (Pfoh et al, 2012). In (B), a graphical representation of the model is shown. The protein structure is colored as in Fig. 1. The horizontal surfaces represent the surfaces of the hydrophobic and the hydrophilic parts of the phospholipid bilayer. Similarly in (C) and (D), however, here we use the assumption of four MSP1D1 per nanodisc.
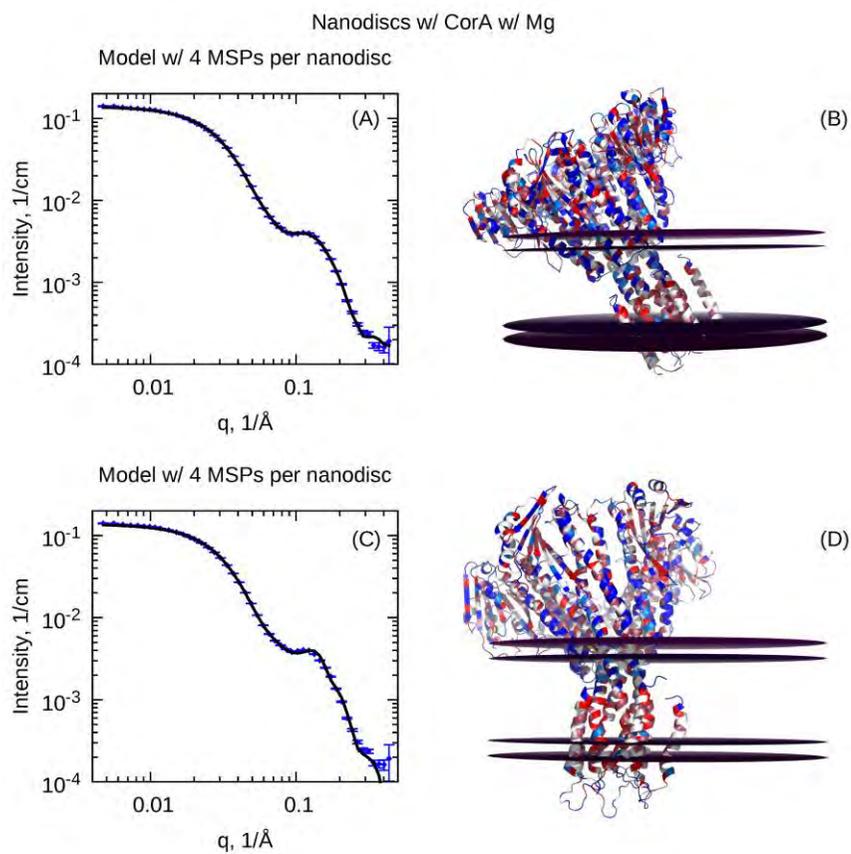
Fig. 4. In (A), SAXS-data from the sample containing magnesium are shown alongside a fit of a model, which assumes four MSPS1D1 per nanodisc. For the CorA itself, here, we use the crystal structure from Lunin et al. (Lunin et al, 2006). In (B), a graphical representation as the ones in Fig. 4 is shown. Similarly in (C) and (D), yet here we assume that the protein structure is described by the previously introduced symmetry-broken structure from Pfoh et al. (Pfoh et al, 2012).
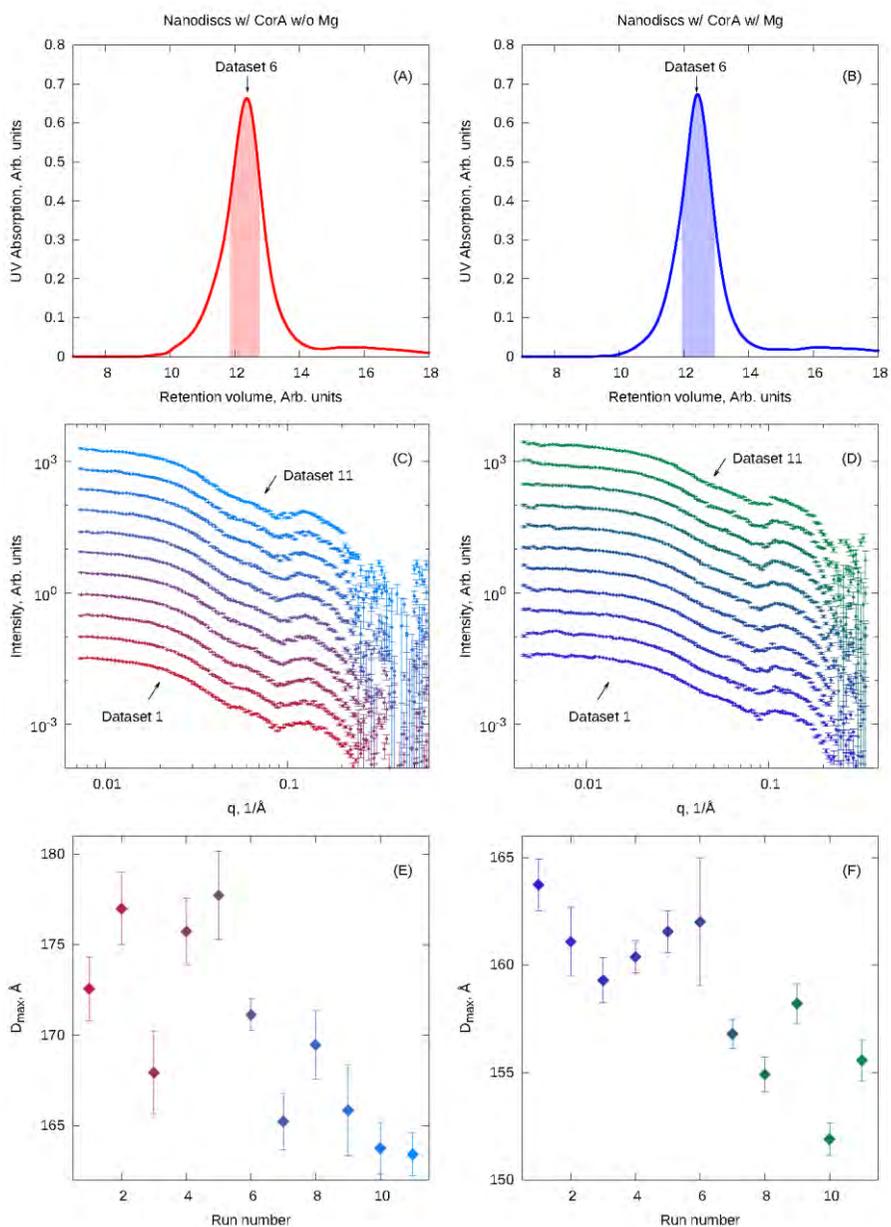
Fig. 5. In (A) and (B), the SEC-data recorded during the SECSAXS-experiment are shown. The fractions, which were exposed in order to generate the data below, are highlighted in opaque blue and red. In (C) and (D), the SAXS-data recorded during the SECSAXS-experiment at the beam-line SWING at Soleil, Paris, are shown. In (E) and (F), values of the largest intra-particle distance, D_{max}, refined from the data above are plotted as a function of run number. Once again, these refinements were done using the BayesApp-software (Hansen, 2012).

Tables

Table 1: Characteristics of the investigated samples. The concentrations and stoichiometries in the two uppermost rows in the table have been computed under the assumption that each nanodisc is constituted of four MSP1D1-belts. The concentrations have been calculated using 254320 M^{-1}cm^{-1} and 21430 M^{-1}cm^{-1} as extinction coefficients for the CorA-pentamer and for MSP1D1, respectively. D_{max} and N_p for the MSP1E3D1-nanodiscs were refined from the peak fraction in the SEC-data in Fig. 5. The quantity N_p is described in detail in the literature (Hansen, 2000).

Characteristic	MSP1D1 w/o Mg	MSP1D1 w/ Mg	MSP1E3D1 w/o Mg	MSP1E3D1 w/ Mg
Absorbance at 280 nm	1.65	1.30	-	-
Concentration of nanodiscs, µM	4.72	3.72	-	-
Phospholipids per nanodisc	198 ± 21	-	-	-
MSPs per nanodisc	3.7 ± 0.6	-	2.0 ± 0.3	2.5 ± 1.1
D_{max}, Å	164 ± 2	174 ± 2	171 ± 0.87	162 ± 2.99
N_p	14.0 ± 0.3	14.6 ± 0.3	7.38 ± 0.02	6.18 ± 0.07
$I(0)$, $1/cm$	0.191	0.140	-	-

Table 2: Values of model parameters refined from SAXS-data. Apart from the reported parameters, a constant background and a term accounting for surface roughness were also refined.

Parameter	2 belts w/o Mg Fig. 3 (A) and (B)	4 belts w/o Mg Fig. 3 (C) and (D)	Straight w/ Mg Fig. 4 (A) and (B)	Bended w/ Mg Fig. 4 (C) and (D)
Axis ratio of the bilayer patch	1.60	1.60	1.00	1.60
Hydrophobic height of bilayer, Å	29.8	29.1	29.9	31.3
Height of belt of MSP1D1, Å	19.9	17.5	17.0	19.0
Number of lipids per disc	262	195	393	233
Volume of MSP1D1, Å^3	26600	32700	35500	32900
Volume of POPC, Å^3	1270	1290	1240	1300
Correction to volume of amino acid residue in CorA	1.03	1.00	1.01	1.00
Horizontal displacement of CorA, Å	16.8	15.8	-6.43	-14.0
Vertical displacement of CorA, Å	-6.80	-5.96	-12.3	-18.2
Rotation of CorA around vertical axis, radians	-0.209	-0.247	-0.671	-0.307
Rotation of CorA around horizontal axis, radians	0.332	0.484	0.541	-0.0493

Supporting Information
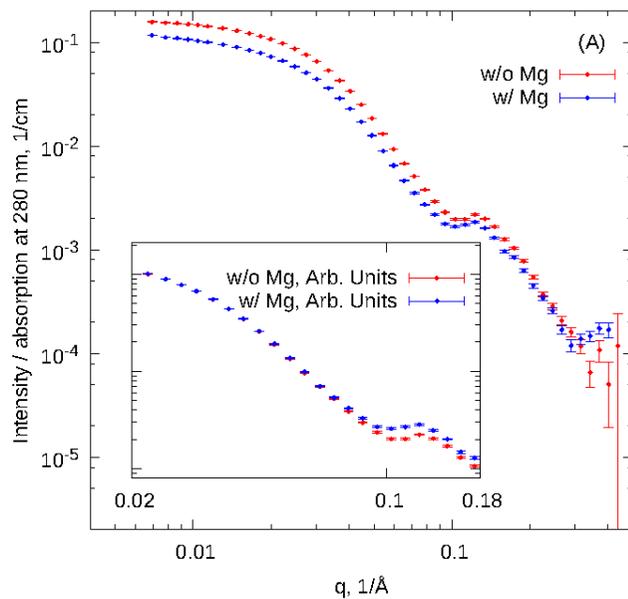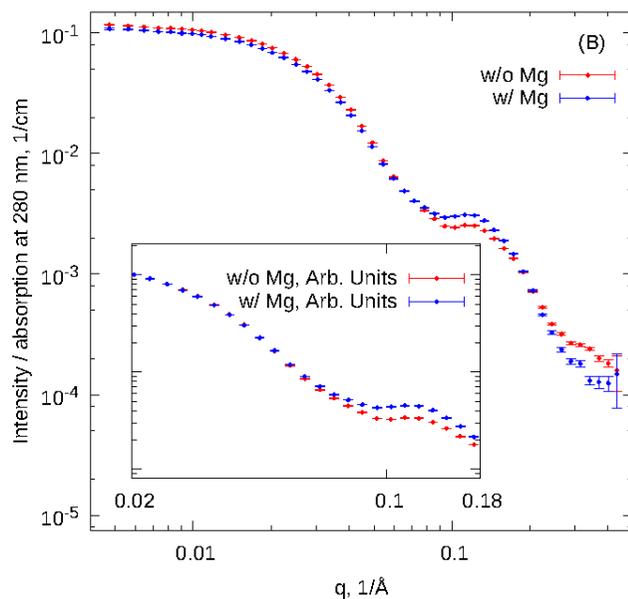
SI Fig. 1. SAXS-data from CorA incorporated nanodisc (with and without Mg²⁺). Inserts illustrate the structural difference in scattering data between the two conformations. (A) For w/Mg data MgCl₂ was added after CorA reconstitution in nanodiscs. (B) MgCl₂ was added from beginning of the reconstitution in nanodiscs.

100

E. Activity and inhibition of disulphide bond forming protein DsbB in nanodiscs

Preface

This project has been a collaboration with people from the Chemistry and Structural Biology Division, Institute for Molecular Bioscience, at The University of Queensland, Australia. The experiments were carried out during my research stay abroad in Professor Jenny Martin's group.

The following draft has been co-written with Patricia Walden and Maria Halili.

Using nanodiscs to improve the stability and activity of integral membrane protein DsbB from *Escherichia coli*

Pie Huda[1], Maria Halili[2], Jennifer L Martin[2] and Patricia M Walden[2*]

1 Structural Biophysics, Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, DK-2100 København Ø, Denmark

2 Chemistry and Structural Biology Division, Institute for Molecular Bioscience, The University of Queensland, Brisbane St Lucia, QLD 4072, Australia

*Correspondence:

Dr Patricia M Walden

Institute for Molecular Bioscience, The University of Queensland

306 Carmody Road, St Lucia, Brisbane, QLD 4072 Australia

Email: p.walden@uq.edu.au

Phone: +61 7 3346 2022

Abstract

Disulphide bond forming protein B (DsbB) is an integral membrane protein and required for the folding machinery of many Gram negative and positive bacteria. In *Escherichia coli*, DsbB is involved in oxidation of its partner protein DsbA to subsequently fold their substrates for function and activity in the periplasm of Gram-negative bacteria. Inactivating these two enzymes will affect folding of proteins associated with virulence vectors. As such, DsbB is a valuable target for the design of novel antimicrobiotic drugs. However, to study DsbB biochemically presents a demanding task, not only because of the immense quantity that is needed for most biochemical assays, also, its stability and function must be retained.

To date, detergents are the means of choice to extract the membrane protein out of its "comfort zone". However, detergents can also be unfavorable for the membrane protein because the micelle's tendency to aggregate, leaving the membrane protein exposed to the harsh extracellular environment. The Nanodisc (ND) has been developed as an alternative and surfactant free approach to solubilise membrane proteins. In this study, we show that ND-embedded DsbB from *Escherichia coli* (EcDsbB) has been produced in substantial amounts for activity assays, and with drastically increased functionalities in two oxidation assays. EcDsbB prepared in NDs will aid in the screening for potential DsbB/DsbA inhibitors and the development of novel antimicrobial drugs.

Introduction

The growing resistance to antibiotics is a concerning problem worldwide. According to the World Health Organization's latest report it is clear that antimicrobial resistance has reached alarming levels in many parts of the world (WHO, 2014). Bacteria express a wide range of virulence factors such as toxins, adhesins, pili, biofilms and proteases in order to establish an infection in a host (Finlay & Falkow, 1997). To accommodate the growing antibiotics resistance, alternatives to the classic treatment are needed. One approach is to block these virulence factors by inactivating the proteins involved.

After synthesis, proteins must fold into their three-dimensional structure to become biologically active, and for numerous proteins the formation of disulphide bonds is crucial for their stability and activity (Rietsch & Beckwith, 1998). Disulphide bond formation takes place between two cysteine residues, enabled by the removal of two protons and two electrons from two participating cysteines, thus catalysing oxidative folding. This is necessary for the function of a number of proteins and is especially common in bacterial virulence factors (Collet & Bardwell, 2002; Heras et al, 2009).

A family of proteins called disulphide bond (Dsb) proteins, found in many prokaryotes, facilitates the formation of these cysteine bonds (Bader et al, 1998). Two central players are the 23 kDa soluble DsbA, and the 20 kDa four helical transmembrane protein DsbB (Bader et al, 2000). DsbA serves as the primary catalyst of disulphide bond formation by transferring a cysteine bond from its active site to the target protein, which leaves DsbA in its reduced form (Kadokura & Beckwith, 2002). DsbB reoxidizes DsbA by transferring electrons to ubiquinone (UQ-1) and further to cytochrome oxidases localized in the membrane (Bader et al, 1999). Inhibiting this oxidative folding machinery may be a possible alternative to antibiotics treatment. It has been shown that *E. coli* mutants lacking Dsb proteins are not able to colonise their host and are thereby rendered non-infectious (Totsika et al, 2009). Inhibiting either DsbA or DsbB could ideally turn infectious bacteria into non-virulent pathogens. With this in mind, Duprez and co-workers were able to synthesize a peptide that is able to inhibit the activity of DsbA by targeting the DsbA/DsbB interface (Duprez et al, 2015).

However, studying membrane proteins such as DsbB is challenging due to their amphipathic nature. Most techniques require solubilisation of the membrane proteins, typically achieved through the use of detergents. However, the environment provided by detergents is less ideal

than the native cell membrane, impacting both structure and function of the protein (Garavito & Ferguson-Miller, 2001). Consequently, alternatives to detergents for studying inhibition of the Dsb machinery are highly desirable.

The use of nanodiscs (NDs) suggests a useful alternative to detergents for studying the DsbA/DsbB interface (Borch & Hamann, 2009). NDs are discoidal lipoprotein particles, which self-assemble from a so-called membrane scaffold protein (MSP) and phospholipids. Each ND consists of two MSPs surrounding a phospholipid bilayer, together holding a diameter of approximately 10 nm (Bayburt et al, 2002). Membrane proteins can be reconstituted into NDs to provide a native-like stable environment (Bayburt & Sligar, 2010) and is superior to detergent micelles for maintaining functional properties of membrane proteins (Lyukmanova et al, 2012).

In the presented study we have successfully incorporated *Escherichia coli* DsbB (EcDsbB) into NDs in large quantities, which enabled us to elucidate the potential of NDs in DsbB/DsbA oxidation and inhibition assays. We compared the ND embedded EcDsbB activity with that of EcDsbB preparations in detergent micelles and crude cell membranes. We were able to show that the EcDsbB-ND complex is highly active in oxidating EcDsbA and that the measured EcDsbB activities in NDs were significantly higher compared to that of micelle-contained EcDsbB. This facilitated us to obtain results comparable with EcDsbB in its native cell membrane environment. Furthermore, activity of EcDsbB could be maintained for substantial longer time than in both detergent micelles and native cell membrane. The study shows that EcDsbB can be prepared in large quantities in NDs, which will contribute immensely to the screening for potential DsbB/DsbA inhibitors and the development of novel antimicrobial drugs.

Results
DsbB can be reconstituted in nanodiscs

EcDsbB embedded in the nanodisc (denoted as EcDsbB:ND) was separated from NDs without EcDsbB (nanodisc without EcDsbB, denoted as MSP1D1(-):ND) by affinity chromatography, and the fraction containing EcDsbB:ND was subsequently purified by size exclusion to obtain a homogeneous and monodisperse sample. A symmetric peak could be obtained from the gel filtration, which indicated that the sample was monodisperse and not aggregated (**Figure 1A**). The entire peak fraction contained both MSP1D1(-) and EcDsbB, demonstrating EcDsbB was successfully inserted into nanodiscs (**Figure 1B**). The final absorbance of EcDsbB:ND at 280

nm was 0.35, corresponding to a concentration of 1.4 μM EcDsbB, assuming each nanodisc contained one EcDsbB.

Increased activity of DsbB in ubiquinone reduction assay

EcDsbA is oxidized by EcDsbB to assist in disulphide bond formation, which is associated with a reduction of UQ-1 by EcDsbB. The activity of EcDsbB can therefore be measured by UQ-1 reduction by monitoring its absorbance at 275 nm. In this assay, we compared EcDsbB activity in nanodiscs to EcDsbB protein prepared in β-DDM detergent micelles.

Our results demonstrate that EcDsbB activity was higher when reconstituted in NDs compared to β-DDM micelles as a higher rate of UQ-1 reduction was observed (**Figure 2**). UQ-1 was half dissociated at a rate of 0.619 ±0.001 minutes in NDs compared to 1.174 ± 0.002 minutes in detergent micelles. This implies a higher degree of UQ-1 was reduced in EcDsbB:NDs, corresponding to an increased rate of EcDsbA oxidation by EcDsbB.

NDs prolongs catalytic ability of EcDsbB

We further analysed the catalytic ability of EcDsbB:NDs with reference to its application in rapid screening inhibitor studies. A synthetic peptide containing two cysteines at the C- and N-terminus with N-terminal europium-DOTA (1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid) and C-terminal methylcoumarin, respectively, was used for fluorescence monitoring in this assay. We compared peptide oxidization catalysed by EcDsbB:NDs to that of EcDsbB embedded in native *E. coli* crude membranes. Peptide oxidation using crude cell membranes of *E. coli* has been studied earlier and is known to retain a significantly higher activity than EcDsbB in β-DDM detergent micelles. In this assay, EcDsbB:ND was able to oxidize the substrate at a similar rate as EcDsbB embedded in crude cell membranes, but with significantly less protein required when using EcDsbB:ND (80 nM) as compared to EcDsbB in cell membranes (1.6 μM) (**Figure 3A and 3B**). The reaction also continued for longer using EcDsbB:ND as an external source of UQ-1 was present, whereas the EcDsbB in crude membrane extract relies on the UQ-1 present in the membrane fraction.

As a further proof of concept, we investigated whether EcDsbB:ND could be used for rapid inhibitor screening using the same peptide oxidation assay and additional inhibitors added to each reaction. Previous studies of EcDsbB in membrane preparations lead to the discovery of a number of inhibitors for EcDsbA and EcDsbB (Duprez et al, 2015; Halili et al, 2015). The

peptide inhibitor PFATCDF is based on the peptide binding sequence of EcDsbA into the active site of EcDsbB and has been shown to bind in the peptide binding groove of EcDsbB, with a Kd 5.9 ± 0.4 μM (Duprez et al, 2015). Using the peptide oxidation assay, we observed that EcDsbB:NDs can be used for inhibitor assays (**Figure 3C**), and we were able to obtain IC_{50} values for the peptide inhibitor of 3.2μM ± 0.5 for EcDsbB:ND and 4μM ± 0.9 for EcDsbB crude membrane.

We noted that significantly less EcDsbB:ND was required to obtain the same result as with EcDsbB embedded in crude membranes. Also the activity of EcDsbB:ND was maintained for a week at 4°C which is substantially longer than what we have observed for micelle contained EcDsbB or crude membrane preparations.

Discussion

EcDsbB is a crucial membrane protein in the oxidation process of many virulent factors in the periplasm of *E. coli*. It is therefore important to screen for inhibitors against this protein to minimize or eliminate bacterial virulence. However, purifying and maintaining a membrane protein such as EcDsbB is an immense challenge to ensure its activity and stability is maintained. The nanodisc provides many advantages over detergent-prepared membrane proteins due to its membrane mimicking properties. They offer a close to native environment to maintain activity of membrane proteins which otherwise become inactive or functionally crippled when kept in detergent micelles (Ding et al, 2015; Lyukmanova et al, 2012; Ranaghan et al, 2011). Nanodiscs have also shown to be able to restore folding of transmembrane domains (Shenkarev et al, 2013). We were able to show that EcDsbB can be successfully incorporated in NDs and obtained in large quantities, thus providing pure and active material for the application in important oxidation and inhibitor assays.

In the UQ-1 reduction assay, the rate of reduced UQ-1 and thus the rate of EcDsbA oxidation could be almost doubled when EcDsbB:ND was used compared to β-DDM detergent micelles. This is a great advantage, because a smaller yield and concentrated sample of EcDsbB:ND was used to obtain a better UQ-1 reduction rate than previously described (Walden et al, 2013). Additionally, in the peptide oxidation assay, the same rate of peptide oxidation was observed in both EcDsbB:ND and EcDsbB prepared in crude cell membranes, using significantly less protein in the EcDsbB:ND samples. The ability to use less protein in assays is advantageous due

to the difficulty in purifying membrane proteins, often resulting in low yields. The use of NDs allows us to conserve protein, as well as greatly improving stability of the protein.

Another advantage of using NDs in this assay was that the assayed sample composition in EcDsbB:ND is known in contrast to using crude cell membranes. In crude cell membranes, multiple unknown cofactors could influence the fluorescence reading and may impact the results of this assay to give false-positive or higher readings. Also, quantification of EcDsbB in crude membranes is challenging, as the exact ratio of EcDsbB to crude cell membranes is difficult to estimate. Furthermore, EcDsbB reconstituted in NDs had a prolonged activity compared to EcDsbB prepared in crude cell membranes. In fact, EcDsbB:ND was stored at 4 ºC for up to a week with no signs of aggregation or loss of activity in the applied assays. The stability of EcDsbB:ND will allow for more time depending assays where the membrane protein activity is not a limiting factor in the future.

As EcDsbB:ND was found to hold a catalytic effect at similar levels to EcDsbB contained in crude cell membrane in the peptide oxidation assay we were interested in applying this system in an inhibition assay. The IC_{50} values we obtained are similar to the Kd value previously reported for this peptide, giving proof that the EcDsbB-NDs can be used for inhibitor screening.

Earlier experiences exploiting nanodiscs for studies of EcDsbB has shown promising results for the use in inhibition assays. When comparing inhibition screening results of EcDsbB in NDs and β-DDM micelles it was possible to identify false positives from β-DDM micelles (Früh et al, 2010). This further demonstrates that NDs can serve as an excellent tool in membrane protein assays. Additionally, the fact that EcDsbB:ND can be stored for longer periods at 4 ºC and shows a higher activity will help to rule out false negatives as an effect of detergent influence and inactivation and will efficiently allow the identification of inhibiting ligands. This will aid in the development of new anti-virulence drugs in the future.

Materials and Methods
Protein expression and purification

MSP1D1 was expressed and purified as earlier described (Ritchie et al, 2009). After purification the N-terminal polyhistidine tag was removed by enzymatic cleavage using Tobacco Etch Virus (TEV) protease. MSP1D1 without the His-Tag (MSP1D1(-)) was further purified by affinity

chromatography using a 5 ml HiTrap HP column (GE Healthcare, Australia). EcDsbA and EcDsbB were expressed and purified as earlier reported (Bader et al, 1998; Duprez et al, 2015).

EcDsbB reconstitution in nanodiscs

1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine (POPC) (Avanti Polar Lipids, Australia) was solubilised in buffer containing 20 mM Tris/HCl pH 7.5, 100 mM NaCl and 100 mM sodium cholate (Sigma Aldrich, Australia) to a final lipid concentration of 50 mM. EcDsbB was mixed with lipids in a 1:260 molar ratio and left for incubation at 4°C for 10 min. MSP1D1(-) was added to the mixture at a molar ratio of 1:260:4 and the mixture was incubated for 30 min at 4°C. 0.5 mL of damp amberlite XAD-2 beads (SUPELCO, USA) were added to the MSP1D1(-):DsbB:POPC mixture and left for incubation at 4°C and continuous stirring for ~3 hours. Amberlite beads were removed by centrifugation at 14,000 RPM and the sample filtrated through a 0.2 μm syringe filter (Whatman, UK). The sample was further applied to a 1 mL HiTrap HP column (GE Healthcare, Australia) that was equilibrated with buffer containing 20 mM Tris/HCl pH 7.5, 100 mM NaCl and 40 mM imidazole. The EcDsbB:ND mixture was eluted from the column with buffer containing 20 mM Tris/HCl pH 7.5, 100 mM NaCl and 400 mM imidazole. The eluted sample was applied onto a Superdex 200 10/300 GL column (GE Healthcare, Australia) which was equilibrated with buffer containing 20 mM Tris/HCl pH 7.5, 0.1 M NaCl. EcDsbB:ND was purified at a flow rate of 0.5 mL/min from the same buffer. Fractions containing EcDsbB:ND were collected and concentrated for redox assays.

Ubiquinone reduction assay

EcDsbB activity in nanodiscs and in β-DDM detergent (Avanti Polar Lipids Inc., USA) micelles were quantified by measuring the capacity of the enzyme to reoxidize the protein EcDsbA. The simultaneous reduction of the cofactor ubiquinone-1 (UQ-1) was monitored at 275 nm. EcDsbA was reduced using 10 mM dithiothreitol (Sigma Aldrich, Australia) for 10 minutes on ice, which was subsequently removed on a PD-10 column pre-equilibrated with degassed assay buffer containing 300 mM sodium phosphate, 100 mM NaCl and 0.03% β-DDM. For detergent solubilised EcDsbB, the reduction of UQ-1 at a concentration of 30 μM was monitored at 275 nm and 30ºC immediately after mixing with 30 μM of EcDsbA and 80 nM of EcDsbB in assay buffer. ND-embedded EcDsbB at a concentration of 20 nM was mixed with same concentrations of UQ-1 and DsbA. The activity of EcDsbB in terms of moles ubiquinone

reduced/moles EcDsbB min−1 was calculated by using the slope of absorption decrease at 275 nm upon reduction of UQ-1 using PRISM®.

Peptide oxidation assay

The assay buffer conatined 50 mM MES, 50 mM NaCl, 2 mM EDTA at pH 5.5. EcDsbB:ND (40, 80 or 140 nM) or EcDsbB in *E. coli* membrane (1.6 μM) were mixed with DsbA (80 nM) and UQ-1 (10 or 100 μM). The peptide substrate CQQGFDGTQNSCK with an N-terminally attached europium-DOTA (1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid) and a C-terminally attached methylcoumarin (AnaSpec, USA) was added to the protein mixture at a concentration of 8 μM. Formation of the disulphide bond in the peptide substrate was monitored via time-resolved fluorescence using excitation wavelength of 340 nm and emission of 615 nm with a delay of 150 ms before reading and 100 ms read time, in a Synergy H1 Multi-mode plate reader (Biotek, USA). The assay was conducted in a 384-well white reaction plate (OptiPlate, Perkin Elmer). Samples containing buffer only, EcDsbA without substrate or EcDsbB without substrate were used as controls. The initial rate of the reaction was calculated over the linear portion of the curve (first 10-20 minutes of reaction).

For the inhibition of EcDsbB using the same assay, the EcDsbA-inhibiting peptide PFATCDF was added to the protein mixture in a two-fold serial dilution starting at 1 mM down to 0.5 μM, diluted in assay buffer. The peptide substrate CQQGFDGTQNSCK was added last to a concentration of 8 μM. The assay was monitored as mentioned above. Reactions were run in duplicates or triplicates per plate, and tested twice on separate occasions. Data was plotted and analysed using GraphPad Prism version 6.0 (GraphPad Software, Inc. USA). For calculating the IC_{50} values, the initial rates of the reaction were calculated over the linear portion of the curve (the first 10 or 20 min of each reaction) and normalized as a percentage relative to no inhibitor present.

Acknowledgements

References

Bader M, Muse W, Ballou DP, Gassner C, Bardwell JCA (1999) Oxidative protein folding is driven by the electron transport system. *Cell* **98:** 217-227

Bader M, Muse W, Zander T, Bardwell J (1998) Reconstitution of a protein disulfide catalytic system. *J Biol Chem* **273:** 10302-10307

Bader MW, Xie T, Yu CA, Bardwell JCA (2000) Disulfide bonds are generated by quinone reduction. *J Biol Chem* **275:** 26082-26088

Bayburt TH, Grinkova YV, Sligar SG (2002) Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Letters* **2:** 853-856

Bayburt TH, Sligar SG (2010) Membrane protein assembly into Nanodiscs. *Febs Letters* **584:** 1721-1727

Borch J, Hamann T (2009) The nanodisc: a novel tool for membrane protein studies. *Biological chemistry* **390:** 805-814

Collet JF, Bardwell JCA (2002) Oxidative protein folding in bacteria. *Mol Microbiol* **44:** 1-8

Ding Y, Fujimoto LM, Yao Y, Plano GV, Marassi FM (2015) Influence of the lipid membrane environment on structure and activity of the outer membrane protein Ail from Yersinia pestis. *Biochim Biophys Acta-Biomembr* **1848:** 712-720

Duprez W, Premkumar L, Halili MA, Lindahl F, Reid RC, Fairlie DP, Martin JL (2015) Peptide Inhibitors of the Escherichia coli DsbA Oxidative Machinery Essential for Bacterial Virulence. *Journal of Medicinal Chemistry* **58:** 577-587

Finlay BB, Falkow S (1997) Common themes in microbial pathogenicity revisited. *Microbiology and Molecular Biology Reviews* **61:** 136-169

Früh V, Zhou Y, Chen D, Loch C, Eiso AB, Grinkova YN, Verheij H, Sligar SG, Bushweller JH, Siegal G (2010) Application of Fragment Based Drug Discovery to Membrane Proteins: Biophysical Identification of Ligands of the Integral Membrane Enzyme DsbB. *Chemistry & biology* **17:** 881-891

Garavito RM, Ferguson-Miller S (2001) Detergents as tools in membrane biochemistry. *The Journal of biological chemistry* **276:** 32403-32406

Halili MA, Bachu P, Lindahl F, Bechara C, Mohanty B, Reid RC, Scanlon MJ, Robinson CV, Fairlie DP, Martin JL (2015) Small Molecule Inhibitors of Disulfide Bond Formation by the Bacterial DsbA-DsbB Dual Enzyme System. *ACS Chemical Biology* **10:** 957-964

Heras B, Shouldice SR, Totsika M, Scanlon MJ, Schembri MA, Martin JL (2009) DSB proteins and bacterial pathogenicity. *Nat Rev Microbiol* **7:** 215-225

Kadokura H, Beckwith J (2002) Four cysteines of the membrane protein DsbB act in concert to oxidize its substrate DsbA. *Embo J* **21:** 2354-2363

Lyukmanova EN, Shenkarev ZO, Khabibullina NF, Kopeina GS, Shulepko MA, Paramonov AS, Mineev KS, Tikhonov RV, Shingarova LN, Petrovskaya LE, Dolgikh DA, Arseniev AS, Kirpichnikov MP (2012) Lipid-protein nanodiscs for cell-free production of integral membrane proteins in a soluble and folded state: Comparison with detergent micelles, bicelles and liposomes. *Biochim Biophys Acta-Biomembr* **1818:** 349-358

Ranaghan MJ, Schwall CT, Alder NN, Birge RR (2011) Green Proteorhodopsin Reconstituted into Nanoscale Phospholipid Bilayers (Nanodiscs) as Photoactive Monomers. *Journal of the American Chemical Society* **133:** 18318-18327

Rietsch A, Beckwith J (1998) The genetics of disulfide bond metabolism. *Annual review of genetics* **32:** 163-184

Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, Sligar SG (2009) *Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs*, SAN DIEGO: ELSEVIER ACADEMIC PRESS INC.

Shenkarev ZO, Lyukmanova EN, Butenko IO, Petrovskaya LE, Paramonov AS, Shulepko MA, Nekrasova OV, Kirpichnikov MP, Arseniev AS (2013) Lipid-protein nanodiscs promote in vitro folding of transmembrane domains of multi-helical and multimeric membrane proteins. *Biochim Biophys Acta-Biomembr* **1828:** 776-784

Totsika M, Heras B, Wurpel DJ, Schembri MA (2009) Characterization of Two Homologous Disulfide Bond Systems Involved in Virulence Factor Biogenesis in Uropathogenic Escherichia coli CFT073. *Journal of Bacteriology* **191:** 3901-3908

Walden PM, Halili MA, Archbold JK, Lindahl F, Fairlie DP, Inaba K, Martin JL (2013) The alpha-Proteobacteria Wolbachia pipientis Protein Disulfide Machinery Has a Regulatory Mechanism Absent in gamma-Proteobacteria. *Plos One* **8**

WHO. (2014) Antimicrobial resistance: global report on surveillance. World Health Organization, France.

Figures

A

B
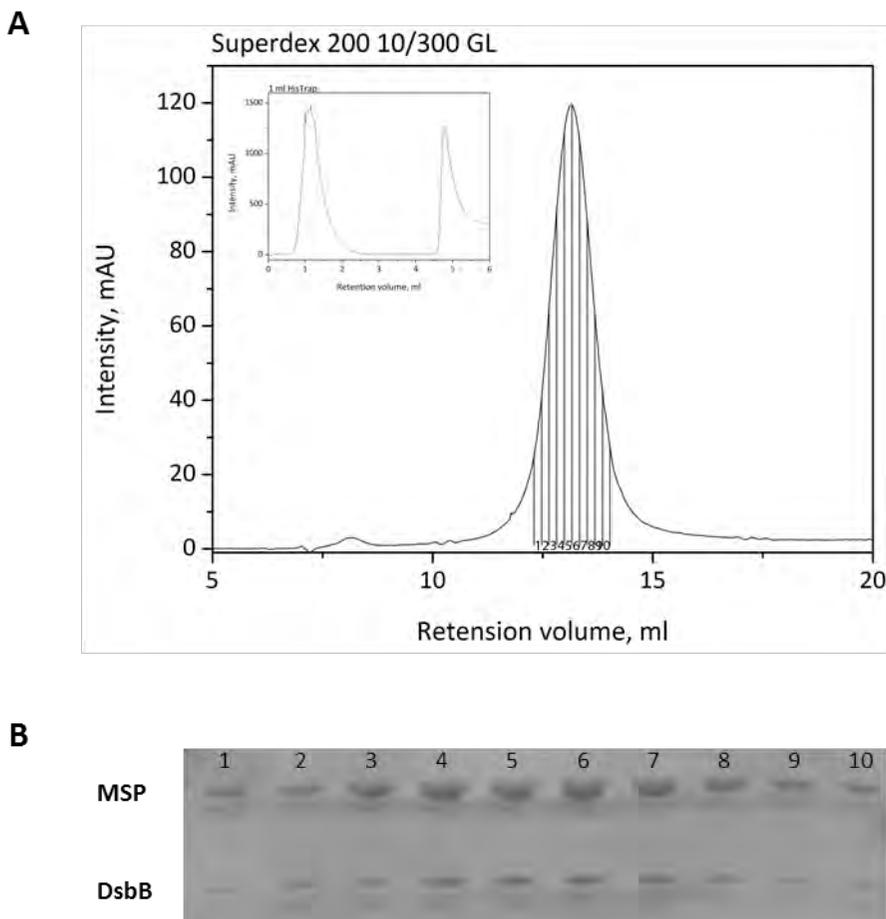
Figure 1. Purification of DsbB nanodiscs. **A)** Gel filtration of EcDsbB:ND using a Superdex 200 10/300 GL column, 200 μL fractions marked 1-10 was applied on SDS gel. Insert, IMAC purification using 1 ml HisTrap column prior to gel filtration. Second peak was collected and applied on the gel filtration column. **B)** SDS PAGE of gel filtration fraction, marked 1-10 in chromatogram on A. Top band row shows MSP1D1(-) and bottom band row corresponds to EcDsbB.
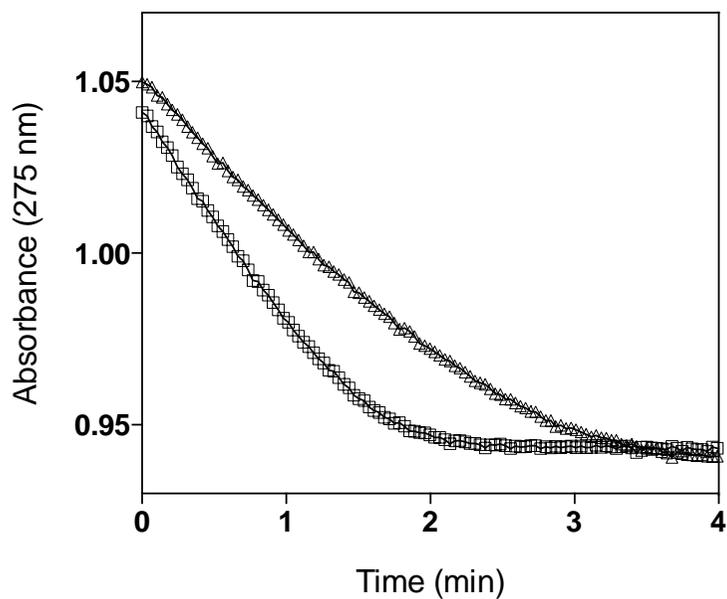
Figure 2. Ubiquinone reduction assay. The UQ-1 reduction assay was performed to measure the ability of EcDsbB to reduce UQ-1. The reduction of UQ-1 either catalyzed by the EcDsbB:ND (square) or EcDsbB prepared in β-DDM micelles (triangle) was monitored at 275 nm over time. The assay was performed in triplicates indicated by the mean (black line).
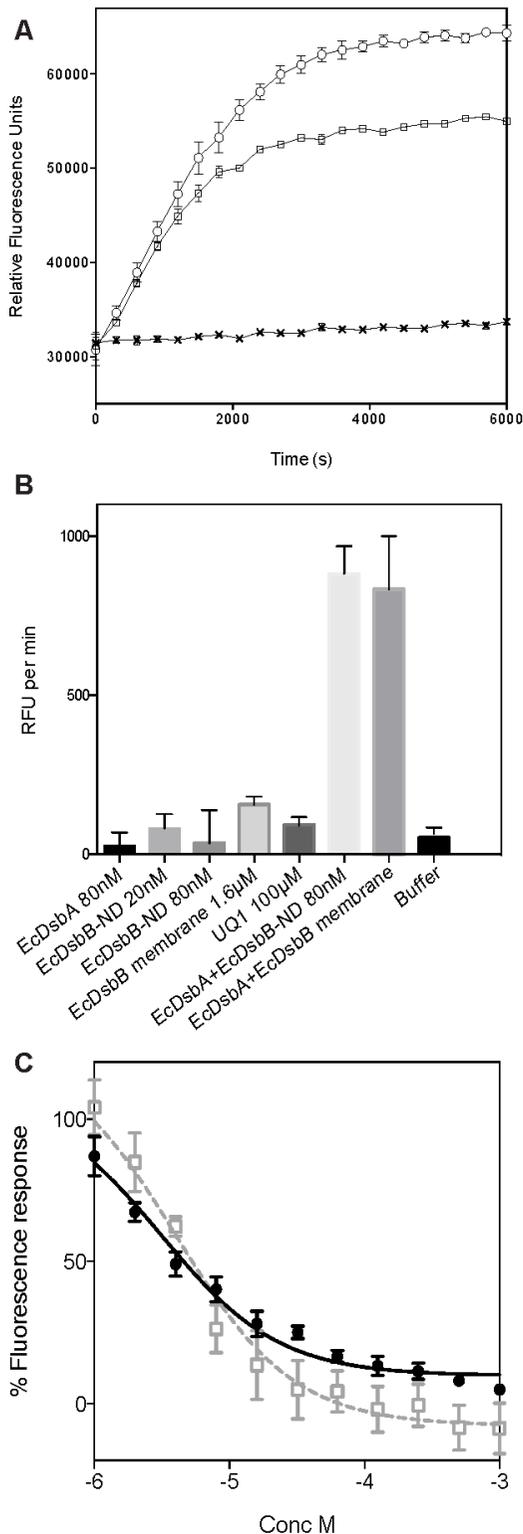
Figure 3. Peptide oxidation assay fluorescence spectra. A) Fluorescence curves showing peptide oxidation of 80 nM of EcDsbB:ND (circle), 1.6 μM of EcDsbB in crude cell membranes (square), and the buffer as blank (cross). **B)** Rates of each component in the peptide oxidation assay. Protein alone (EcDsbA or EcDsbB) or UQ-1 does not cause significant oxidation of the peptide substrate, as compared to the buffer baseline. Only when all components are mixed together does the oxidation reaction occur. EcDsbB-ND at 80nM and EcDsbB in crude membranes at 1.6μM are able to oxidize the peptide substrate at a similar rate.

C) EcDsbB-ND can be used for testing inhibitors. The peptide-based inhibitor PFATCDF was added to the corresponding protein mixtures, and was able to inhibit EcDsbA-mediated oxidation of the substrate. IC_{50} values were calculated to be 3.2μM $\pm$ 0.5 for EcDsbB-ND (filled circles) and 4μM $\pm$ 0.9 for EcDsbB crude membrane (squares). For **A** and **B**, assays were performed in triplicate and shows the mean +/- SEM, and was performed on two separate occasions. For **C**, assays were performed in duplicate and shows the mean $\pm$ SEM, run on two separate occasions.

Chapter 7 - Probing protein dynamics

One project will be presented in this chapter concerning probing structural dynamics of proteins:

F. Developing a microfluidic chip for time resolved SAXS

Preface to the project
This has been a collaborate project together with Grethe Vestergaard Jensen and Søren Skou Nielsen at the Niels Bohr Institute. My responsibility in this project has mainly been sample preparation and participation in data collection together with Grethe and Søren, who were responsible for chip design and manufacturing. In addition to the experiments presented in this report an experiment was carried out at the cSAXS beamline at the Swiss Light Source. However, due to alignment problems no useful data could be obtained.

The following report has been co-written with Grethe Vestergaard Jensen.

F. Developing a microfluidic chip for time resolved SAXS

Abstract

The three-dimensional structure of a protein determines its functionality which makes studying of protein structures fundamental for understanding cellular function and diseases. Proteins are dynamic and their actions linked to structural transitions caused by certain changes in the local environment. Probing the structural dynamics of proteins can give valuable information about structure-function relationship.

A microfluidic mixing chip for time-resolved SAXS studies has been developed, aimed at investigations of structural dynamics of proteins in solution. Whereas ligand-induced structural transitions are an important aspect of the functioning of numerous proteins, only limited knowledge on these processes is currently available. By introducing a microfluidic platform for fast mixing of protein and ligand solutions, which is compatible with micro-focus synchrotron SAXS, it is possible to obtain information on the kinetic processes with sub-millisecond time resolution and a minimal sample consumption.

Introduction

Proteins fulfil most vital functions in the cell. Their three-dimensional structure is essential for their function and their actions are coupled to structural transitions, such as calcium binding of calmodulin inducing a structural change, seen in Figure 1, activating or regulating a number of other proteins by its conformational change (Chin & Means, 2000).

Calmodulin is one of the most prominent Ca^{2+} sensing proteins and is expressed in all eukaryotic cells. Here it participates in signalling pathways regulating crucial processes such as cell growth, cell division and cell movement. Calmodulin contains only 148 amino acids and is highly conserved (Chin & Means, 2000). It adopts very distinct conformations in the presence and absence of Ca^{2+}, as shown in Figure 1, making it a good candidate for structural time resolved studies.

117
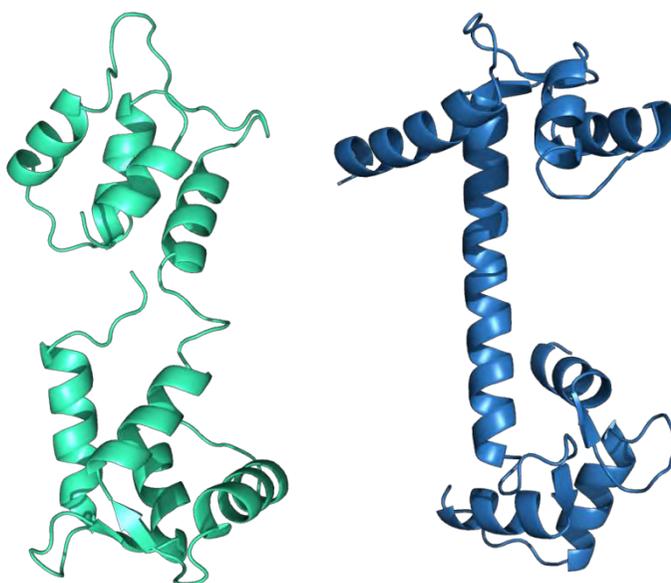
Figure 1. Calmodulin structure in absence and presence of Ca²⁺. Left without Ca²⁺ [pdb entry 1CFD]. Right Ca⁺² bound structure [pdb entry 1CLL].

The structural transition of calmodulin upon binding to calcium has previously been studied using a microfluidic chip and structural changes probed by fluorescence. The rapid mixing technique applied enabled times from 40 μs to 2 ms to be probed, and using a different setup for stopped flow mixing, times from 1 ms to 2 s resolution could be obtained. From this combination of techniques it was possible to identify two distinct events during Ca²⁺ binding at ~490 μs and ~20 ms (Park et al, 2008).

The structural kinetics have subsequently been probed by SAXS, by collecting data after photo-induced release of caged Ca²⁺ from a photo sensitive Ca²⁺ chelator. Times from 0.5 to 180 ms were covered with this method, revealing the presence of a compact intermediate conformational state (Yamada et al, 2012). The advantage of using small angle scattering, over e.g. fluorescence based techniques, is that this kind of structural details about the protein can be obtained in addition to the various kinetic rates.

The action of proteins in form of structural changes varies extensively in time scale and can be observed from picoseconds to milliseconds (van den Bedem & Fraser, 2015). Both method and protein candidates for structural dynamics studies must be carefully evaluated to meet the relevant time frames.

For this study a microfluidic mixing chip has been designed for studying structural dynamic of proteins using SAXS. The chip design allows for easy and cheap manufacturing and is inspired by the work of Pollack and co-workers (Park et al, 2008). Protein and ligand solutions enter the chip through inlet channels in a continuous flow, determined by a set of pumps. The solutions meet at a mixing point and flow in sheets, side by side in a laminar fashion. A structural transition of the protein is triggered by the ligand diffusing into the protein solution sheet. Any

position downstream of the mixing point represents a certain time after mixing. The structural protein transition can therefore be followed time-resolved by probing the structure at different channel positions by SAXS. In this project it is shown that controllable widths of the protein sheet can be formed, down to 5 μm (giving a typical time of ligand diffusion into the centre of the sheet of ca. 4 ms for a Ca^{2+} ligand (Ander et al, 1979)). Flow velocities between 3 and 20 mm/s can be obtained by tuning the pump speeds, resulting in a time resolution down to 0.5 ms for the fastest flow rate for a beam size of 10 μm. The longest time that can be probed is given by the slowest flow rate and the length of the channel (here 2.5 cm), resulting in a time of 5 s

In the effort to make a microfluidic platform for fast mixing of protein and ligand solutions easily accessible to relevant users within structural biology, some reflections were made concerning the chip design. First of all, the manufacturing of the chip must be cheap and easy. This implies that the fabrication method should be generally available and the chip and window materials relatively cheap. Designing the chip to hold a relatively wide channel (> 200 μm) allows for simple manufacturing methods such as milling or casting of the chip, opposed to etching techniques which require clean room facilities. This also facilitates easy mass production increasing prospects for commercialization, which is the essential key to accessibility.
Secondly, low sample consumption is essential as biological samples in most cases are hard to obtain in large quantities. This can be achieved by creating a narrow sample sheet within the channel of the chip. In a wide channel this is only possible with a stable well-controlled sample flow. This means the channel must be well characterized to ensure a stable sample sheet can be maintained. Reliable pumps are crucial for delivering a steady continues flow of solutions to the inlets, down to μl/min precision.

Methods
All chemicals were purchased from Sigma Aldrich.

Proteins and solutions
Bovine serum albumin (BSA) was purchased from Sigma Aldrich and solubilised in 20 mM Tris/HCl pH 7.5, 100 mM NaCl. Concentrations around 5-7 mg/ml were applied. SDS solution for mixing contained 20 mM Tris/HCl pH 7.5, 100 mM NaCl, 40 mM SDS.

Insulin lithocholyl-γ-desB30, obtained from Novo Nordisk, was solubilised in milliQ water and added NaOH to a pH >10 and pH lowered to pH ~7.5 with HCl. The solution was added $Zn(OAc)_2$, phenol, $NaPO_4$ (pH 7.5) and NaCl to final concentrations at 0.1 mM, 32 mM, 0.7 mM and 30 mM respectively. A Protein concentration of 0.6 mM was applied. NaCl solution for mixing contained 10 mM Tris/HCl pH 7.5, 150 mM NaCl, 32 mM phenol.

Calmodulin (CaM) was expressed and purified as earlier described (Singla et al, 2001). In short CaM cDNA from sea Urchin inserted in a pET23d vector was transformed into *E. coli* BL21(DE3). Cells were grown in TB media and expression induced with IPTG. A 5 and 15 ml phenyl-sepharose column equilibrated in 10 mM Tris/HCl pH 7.5, 0.5 M NaCl, 1 mM EDTA and 10 mM Tris/HCl pH 7.5, 0.5 M NaCl, 10 mM $CaCl_2$, respectively. CaM was eluted from the

15 ml column with 10 mM Tris/HCl pH 7.5, 0.5 M NaCl, 50 mM EGTA and dialysed against 10 mM Tris/HCl pH 7.5, 10 mM NaCl, 1 mM EGTA. Calcium solution for mixing contained 10 mM Tris/HCl pH 7.5, 10 mM NaCl, 100 mM $CaCl_2$. Prior to applying CaM solution on the chip, the channel was flushed with 10 mM EGTA. A protein concentration of 11 mg/ml was applied.

Concentrations of all protein solutions were measured by UV absorbance at 280 nm using a nanodrop-1000.

Chip description

The microfluidic chip has five inlets, as shown in Figure 2, where protein and ligand solutions can be applied. For all experiments inlet 3 in the chip was used for the protein sample solution and either inlet 2 and 4 or 1 and 5 were used for ligand solution while the remaining two inlets were kept blocked. The additional inlets are designed to enable introduction of a thin buffer sheet between the protein and ligand solution, preventing uncontrolled mixing at the inlets (Park et al, 2006). The inlets are designed in such a way that buffer laminates top and bottom of the protein sheet, so-called 3D focusing, protecting the windows from the protein sample (Park et al, 2006). This means that the protein sheet does not cover the entire X-ray path length of 1 mm. 1 mm path length was chosen for optimal scattering signal. In the chip protein and ligand solutions enter the inlet channels in a continuous flow, determined by a set of pumps. The solutions meet at a mixing point and flow side by side in a laminar fashion. The mixing occurs passively by diffusion of ligand into the protein sheet. Any position downstream of the channel represents a certain time after the mixing point. The chip can be mounted and aligned on an X-ray beamline, as seen in Figure 3.
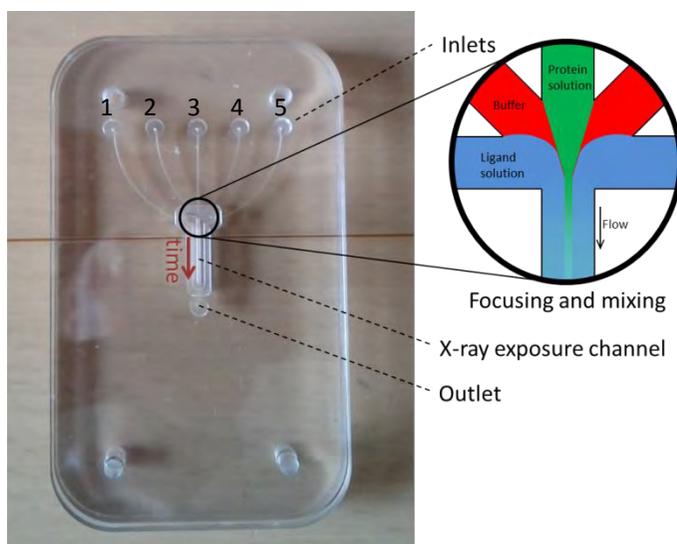
Figure 2. Photo of the microfluidic chip with a schematic illustration of the laminar flow in the microfluidic chip. The protein and ligand solutions meet at the mixing point (circled), followed by diffusion of the ligand into the protein sheet. Different times after the mixing point are sampled by measuring at different positions on the protein sheet, with time increasing downstream the flow direction. Inlets are numbered 1-5.
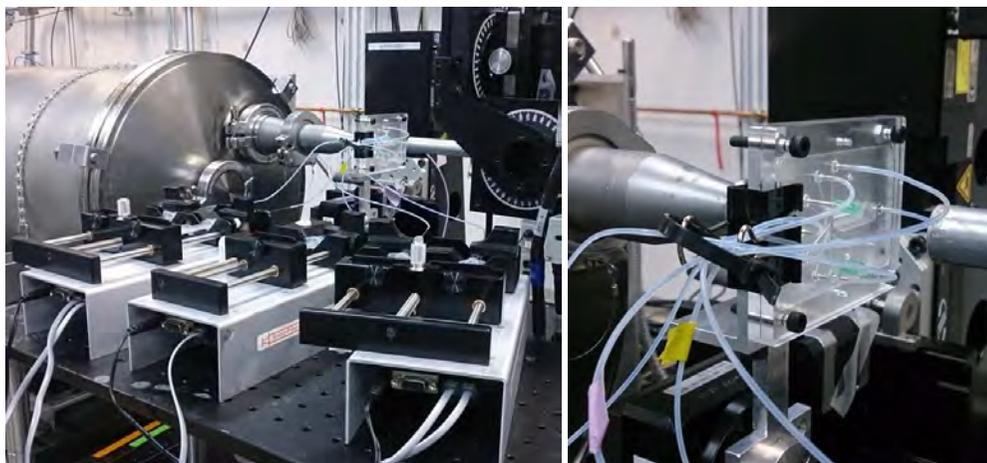
Figure 3. The chip mounted at the BioCAT beamline. Left image show the chip connected to the three syringe pumps holding 5 syringes in total. Right image, a close up of the mounted chip. The chip is placed in a chip holder with an opening around the exposure channel in such a manner that the X-ray beam only passes the sample and the thin chip window.

Two different chip designs were used for the two presented experiments both using step motor syringe pumps (ProSense). The specifications were as followed.

Chip specifications P12
Chip material: Poly(methyl methacrylate) (PMMA)
Window material: 25 μm polystyrene
Fabrication method: milling
Channel width: 500 μm
Channel length: 10 mm
X-ray path length, total: 1000 μm
X-ray path length through protein sheet: 600 μm

Chip specifications BioCAT
Chip material: UV-curing thiol-ene
Window material: 25 μm polystyrene
Fabrication method: moulding
Channel width: 200 μm
Channel length: 25 mm
X-ray path length, total: 1000 μm
X-ray path length through protein sheet: 600 μm

Beamline specifications
Scattering data were recorded at the P12 beamline at PETRAIII (Blanchet et al, 2015) in September 2013 and the BioCAT beamline at APS (Fischetti et al, 2004) in December 2014. All

experiments were conducted at room temperature. Acquisition times are specified in figure legends of presented data in the results section.

P12
Beam size: ~70x40 μm (Full Width Half Maximum (FWHM))
Wavelength: 1 Å
Detector: Pilatus M2
Detector distance: 3 m

BioCAT
Beam size: ~11x6 μm
Wavelength: 1.033 Å
Detector: Pilatus 100k
Detector distance: 2821.6 mm

In both cases the SAXS data were converted to scattering intensity units as a function of scattering vector $\vec{q}$, $I(q)$, where $q = \left| \frac{4\pi sin\theta}{\lambda} \right|$. Background subtraction of the data was performed using the BioXtas raw software (Nielsen et al, 2009).

Results and discussion

The following sections will cover the two experiments carried out at the P12 beamline at PETRAIII and the BioCAT beamline at APS, respectively. The experiment at P12 enabled basic characterizations of the chip, using a relatively large beam. From the experiment at the BioCAT beamline a more detailed characterization of the channel and the protein sheet could be obtained due to a much narrower beam.

P12 – basic characterization of the chip

Water measurements and background measurements
To allow for data calibration at different positions and to test if reproducible data of good quality could be obtained throughout the channel of the chip, different measurements of empty chip and water were measured at different positions downstream the channel. The empty background data were subtracted from the water data, which should result in a flat scattering pattern.
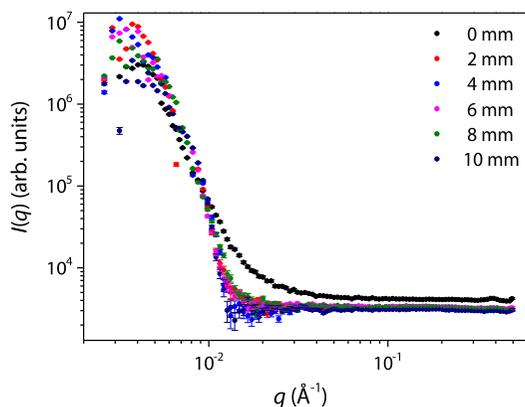
Figure 4. Background subtracted water measurements at different positions downstream the channel. Background measured at each position was used for subtraction. Acquisition time 10 x 0.5 s at each point. Label shows distance from mixing point.

Reproducible background subtracted data can be obtained for different positions in the channel as seen in Figure 4, using backgrounds measured at each position, resulting in acceptable results for $q > 0.01$ Å^{-1}. That is, with this chip data can be collected for $q > 0.01$ Å^{-1}. This is except from the mixing point of the chip, where the intensity for some reason is higher. This may be caused by a scattering contribution from the edge of the channel even though this was not deducible from the 2D detector images. Another reason could be a slightly larger path length in this area of the chip giving a higher intensity. The fact that the channel does not have any sides at the top of the mixing point (since the buffer inlets are coming in from the sides) might also explain the higher intensity, if the tails of the beam cover the entire channel width. This explanation seems plausible based on the results obtained for varying protein sheet width presented below.

Varying protein sheet width
The width of the sample sheet needs to behave predictably for a good control of the mixing and sample consumption. This is controlled by changing pump speed of the different solutions, resulting in expected protein sheet widths, which can be calculated from the flow rates of the different solutions. Different sheet thicknesses of BSA solution were applied in the chip, surrounded by a buffer solution, and SAXS data were collected to test if this was fully controllable. No structural transition of the protein was introduced in this experiment, which was only used to test the performance of the chip. Data were collected at a single position on the chip and normalized by the expected protein sheet widths, which should result in overlapping scattering patterns.
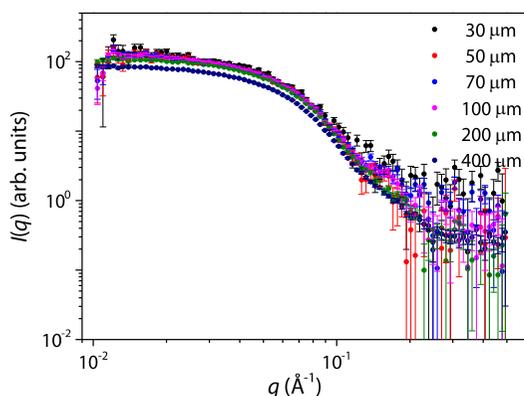
Figure 5. Varying sheet widths of BSA, measured 2 mm downstream the channel from the mixing point. Pump rates for buffer and protein were varied to obtain the different widths. The acquisition time is 10 x 0.5 s for each measurement. Intensities are normalized by the expected protein sheet width. Label shows applied predicted sheet widths.

Reproducible data can be obtained for all widths according to data in Figure 5, showing the scattering patterns normalized by the expected sheet widths. The normalized intensity for the 400 μm sheet is slightly lower than for the other measurements. This is caused by the fact that the 400 μm sheet is wider than the X-ray beam, and the full sheet width is therefore not in the beam. It may be expected that this had occured already at a sheet width of 200 μm, considering the beam width of 70 μm, but as the beam size is specified as FWHM, the total beam width extends beyond this value, possibly with significant intensity outside the 70 μm. The fact that the normalized intensity only decreases for the 400 μm sheet implies that the beam is actually quite wide with significant intensity beyond a width of 200 μm.

Varying flow rate

It is expected that the total flow rate in the channel does not affect the sheet width. In order to be confident about the performance and reproducibility of the pumps, and thereby the control over the flow profile and mixing, this expectation was put to the test. Data were collected at a single point in the chip for different flow rates while keeping constant mixing ratio of BSA and buffer solution. Ideally this should result in overlapping scattering patterns.
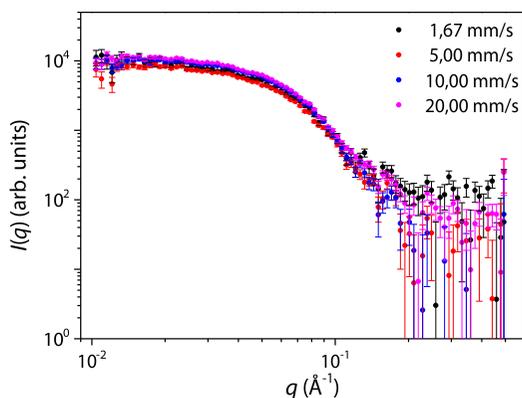
Figure 6. Varying total flow rate (buffer and protein) measured at mixing point and a protein sheet width kept at 50 μm. Acquisition time 10 x 0.5 s at each flow rate.

The intensity was relatively constant at the different flow rates as shown in Figure 6, though a minor increase in intensity was observed for a flow rate of 20 mm/s. This may be caused by the fact that the data were collected close to the mixing point and this can have caused some inaccuracy in controlling the protein sheet width at high flow rates.

Varying distance from mixing point
Ideally and theoretically, the protein sheet will remain centered downstream the channel. To test if this was the case, data were collected downstream the chip by decreasing the distance from the mixing point with 2 mm per measurement. The protein sheet thickness and the flow rate were kept constant at 50 μm and 10 mm/s respectively. For a well-centered protein sheet it is expected that the protein is remained within the beam implying that consistent intensities can be collected down the channel. However, it is expected that protein eventually diffuses out of the sheet over time and will happen faster the narrower protein sheet applied.
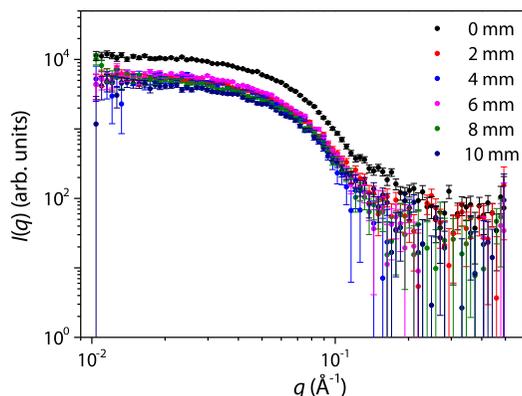
Figure 7. BSA measured downstream the channel. Protein sheet was kept at 50 μm and flow rate at 10 mm/s. Position given in distance from mixing point. Acquisition time 10 x 0.5 s at each point. Label shows distance from mixing point.

From the data in Figure 7 it is noticed that the intensity is higher at the mixing point, as it was observed for the water measurement.

It can be seen that most measurements resulted in similar intensities, as expected, though at 10 mm the intensity decreases slightly. This is probably caused by some protein has diffused far out from the sheet at this point and the total protein sample is no longer within the beam width. However, from the data it cannot be determined if the protein sheet width equal 50 μm as expected, since the applied beam size is several times greater as concluded earlier.

Kinetics studies of BSA unfolding by SDS

Once the ability to control flowrates and mixing ratios was confirmed, the next step was to collect data for a protein system undergoing a structural transformation. Unfolding of BSA by complexation with the surfactant sodium dodecyl sulphate (SDS) is a well-known phenomenon, and the complexes are structurally well-characterized by SAXS. (Santos et al, 2003). Additionally the system is easy accessible and was therefore an ideal candidate for testing the performance of the chip for time resolved studies. BSA solution was applied as the protein sheet while a 40 mM SDS solution was used as surrounding ligand leading to unfolding of BSA as SDS diffuses into the protein sheet. 40 mM SDS was used as this concentration is known to be sufficient for inducing unfolding and form complexes with BSA (Santos et al, 2003). A BSA concentration of approximately 7 mg/ml was applied to acquire high intensity while avoiding effects associated with protein interactions, such as oligomerisation or structure factor contributions to the scattering pattern. Data were collected in the center of the mixing channel at positions downstream from the mixing point. The applied protein sheet width was kept constant at 50 μm, and the total flow rate was fixed at 10 mm/s. As the protein sheet was mixed with a SDS solution, the structural transition associated with the unfolding of the protein was hoped to be observed.
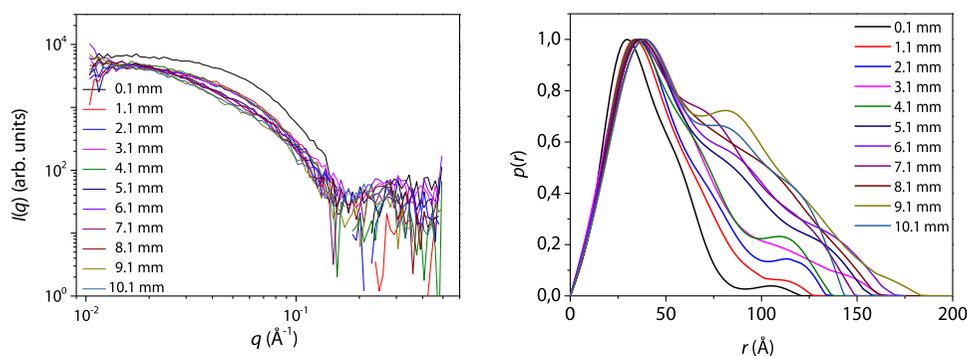
Figure 8. Mixing of BSA and SDS for kinetic studies. left, scattering data collected at different distances from the mixing point, corresponding to different time points after mixing. A 50 μm protein sheet and a 10 mm/s flow rate were applied. Acquisition time 40 x 0.5 s at each point. Right, corresponding *p(r)*-functions.

As seen in Figure 8 the scattering pattern changes with increasing distance from the mixing point, indicating a structural transition corresponding to the unfolding of BSA. Closest to the mixing point (distance 0.1 mm), the intensity is much higher than for the following points, as it was also observed for the measurements on water. Again, this was related to the fact that the

beam tails cover the entire channel width, resulting in additional signal at the mixing point where the channel does not have sides but rather inlets. A further very slight decrease in intensity is also observed with increasing distance from the mixing point, due to dilution of the protein (diffusion out of the beam centre).

To study the kinetics of BSA unfolding, the diffusion of SDS into the BSA sheet must occur faster than the kinetic times wished to probe. Due to the wide beam and the correspondingly wide protein sheet, this is not the case here. The BSA sheet width of 50 μm together with the diffusion constant of SDS (4.4×10^{-10} m²/s (Chen et al, 1995)) implies that it will take an SDS molecule 0.71 seconds to diffuse half the sheet width. At a flow rate of 10 mm/s, this corresponds to a position in the channel at 7.1 mm from the mixing point, almost the entire channel length. This means that the mixing is simply too slow for studying any kinetics and the observed structural transition is an effect of a steadily increasing SDS concentration rather than a kinetic response to an instantaneous increase of concentration. The slow diffusional mixing also means that at any measuring point, some of the protein is unfolded (at the edge of the BSA sheet) whereas some have not yet been mixed with SDS and remains folded (at the centre of the sheet), as illustrated in Figure 9, and the data actually represent a collection of different structures of SDS, folded BSA and unfolded BSA as a function of SDS concentration down the channel. In order to follow an actual kinetic process, it is necessary to obtain faster diffusional mixing, either by reducing the sheet width (requiring a smaller X-ray beam) or by studying a system with a faster diffusing ligand.
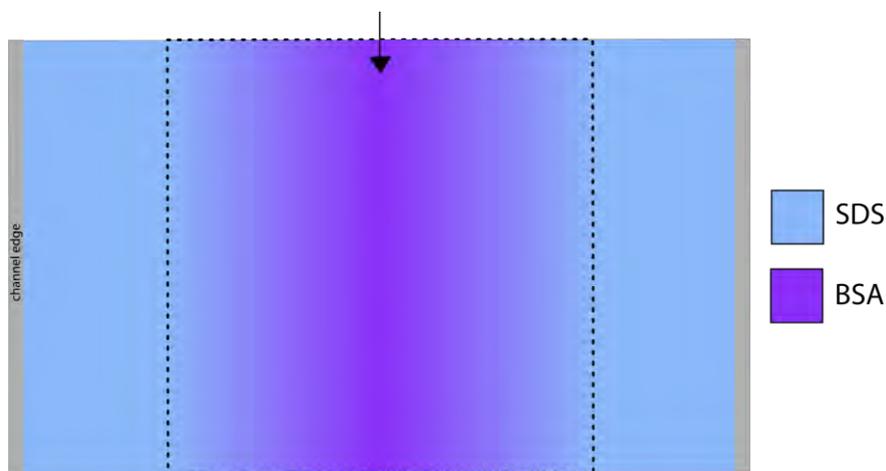
Figure 9. Inhomogeneous mixing in a measuring point of the protein sheet (dashed line indicates beam area). The BSA protein sheet is mixed with SDS. Due to slow mixing, the closer to the center of the sheet (indicated with arrow), less SDS has diffused into the sheet, implying some of the BSA in the center remains unmixed and folded. Channel and beam size are not drawn to scale.

Kinetic studies with insulin and NaCl

To attempt kinetic studies with another protein system (other than BSA-SDS), which has also been well characterized by SAXS, a specific type of insulin was applied. Lithocholyl-γ-desB30 insulin is known to be able to self-assemble into a characteristic structure at certain NaCl

concentrations. This insulin is formulated as hexamers, in the presence of phenol at low NaCl concentration. When exposed to physiological NaCl concentration the hexamers oligomerize in a specific manner into rod shaped particles (Nygaard et al, 2015). As NaCl is a considerable smaller molecule than SDS the diffusion is much faster and it was therefore expected that some kinetics could be observed for this system. Insulin was applied on the chip and mixed with a physiological saline buffer in order to observe this self-assembly. The protein sheet was kept at a constant width of 50 µm. Data were recorded at different positions in the channel at flow rate 10 mm/s.
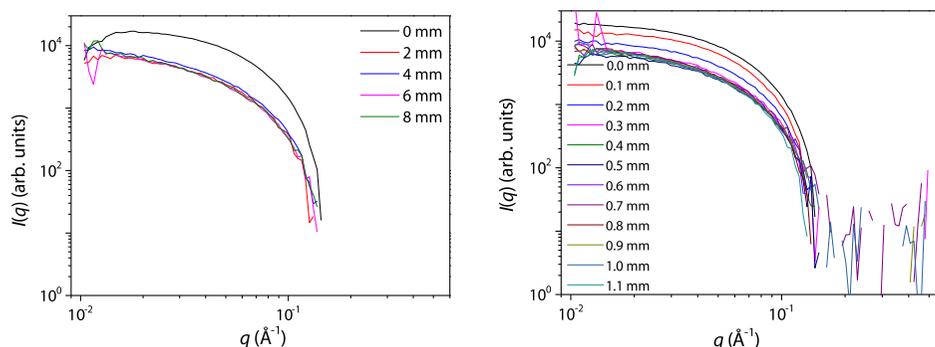
Figure 10. Mixing of insulin and NaCl for kinetic studies. Measurements downstream the channel at distances from the mixing point corresponding to different time points. A 50 µm protein sheet and a 10 mm/s flow rate were applied. Acquisition time 40 x 0.5 s at each point. Left, data points from 0 to 8 mm with 2 mm increment. Right, data points from 0 to 1.1 mm with 0.1 mm increment. Positions in the chip is given in mm distance from the mixing point.

From the mixing study of insulin and NaCl a decrease in intensity and a small change in scattering pattern are observed in the data from 0 to 2 mm down the channel, as seen in Figure 10. It is noticed that the observed change is due to both structural changes and mixing, as the scattering pattern changes both shape and decreases in intensity over time.

Investigating this reaction at shorter time steps, it can be deduced from Figure 10 that most of the reaction happens within the first 1 mm corresponding to 0.1 s. For NaCl it takes a molecule 0.21 seconds to diffuse half the sheet width (NaCl diffusion constant 1.5×10^{-9} m²/s (Vitagliano & Lyons, 1956)). This corresponds to a position in the channel of 2.1 mm from the mixing point at a flow rate of 10 mm/s. Mixing at 1 mm from mixing point is therefore not complete and so no kinetics can be observed at this point. First after 2.1 mm can any potential kinetics be observed, though here would be a mix of structural stages present due to inhomogeneous mixing times. Over subtraction of the data is unfortunately observed in the high q-region. The reason for this is unclear but may be related to combinations of impurities in the chip and low intensities giving bad statistics in the high q-region.

To evaluate the degree of structure change in relation to mixing, radius of gyration R_g and zero angle intensities $I(0)$ obtained from Guinier plots of the data were calculated at each data point. As R_g is related to particle size it was expected to grow as a function of time due to the oligomerization of insulin. Dilution of the protein will cause a decrease in $I(0)$ as the scattering intensity is dependent on the protein concentration.
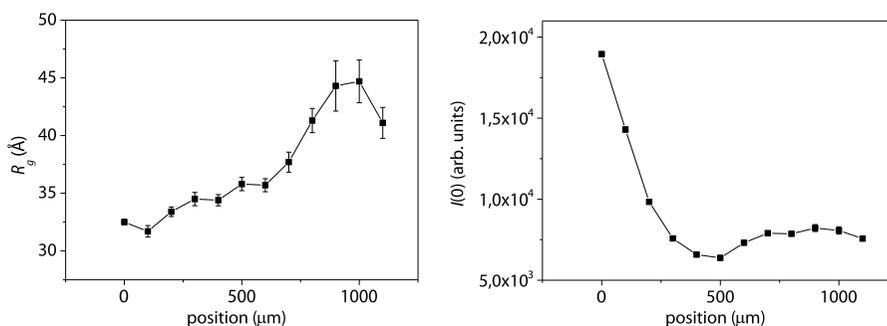
Figure 11. Radius of gyration (left) and zero angle intensities (right) of insulin over time after mixing with NaCl.

It is seen that the R_g in Figure 11 grows over the first 1 mm showing that the particles grow in size, implying some structural change. However, due to the mixing is not complete at this point we have different structures present where insulin in the edge of the sheet have been mixed with NaCl while insulin in the centre remain unmixed, similar to the earlier BSA-SDS study. $I(0)$ decreases at the first time points as dilution occurs decreasing the scattering intensity. It is noticed that the change in scattering pattern at the first time points is mainly caused by dilution due to the large decrease in $I(0)$. Though, the high intensity observed at the first point is again in the mixing point which for the earlier experiments also showed an elevated intensity. R_g does not increase drastically meaning the final oligorimization state is not reached at this short time frame as a rod-like structure of length ca. 300 Å and an R_g of 95 Å is observed for this type of insulin at conditions corresponding to the final state (Nielsen, 2013). Only a minor structural change or growth in particle size is observed, indicating that the further oligomerization takes place at a longer time scale than 0.8 s, which cannot be covered with this chip and might be more suitable to investigate using a stopped-flow approach.

BioCAT – detailed sheet characterization

A much smaller beam was available at the BioCAT beamline, compared to the P12 beamline, which allowed for a more detailed characterization of the protein sheet profiles in the chip. The smaller beam also enabled the use of a much narrower sheet of protein solution without loss of signal, since the width of the sheet covers the size of the beam. The narrow sheet leads to lower sample consumption and much faster mixing. Faster mixing times are enabled using a narrower protein sheet due to the fact that diffusion time scales with squared diffusion distance. A chip holding a narrower channel was applied, compared to the one used at P12, and the buffer consumption was therefore also decreased. Additionally, the channel was longer allowing for data collection at positions corresponding to longer reaction times.

Characterization of the channel

The small beam applied at the BioCAT beamline allowed for a detailed scan of the chip to characterize the channel dimensions and determine the coordinates giving the position in the chip relative to the mixing points and the channel walls. Transmission scans across the chip channel at three different positions along the channel were performed to confirm similar

dimensions along the channel. This was supplemented with scans performed later during the beamtime to evaluate eventual changes in chip position. A transmission scan along the channel was also performed.
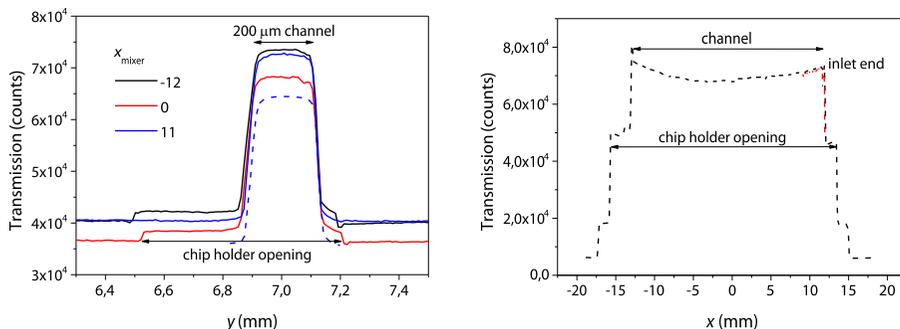
Figure 12. Left, transmission across (y-direction) the chip at three positions over 23 mm downstream (x-direction) the channel, showing the width of the channel. Dashed line is a transmission scan of the channel later in the experiment compared to the first scans. Highest transmission shows the channel width. Right, transmission along the channel (x-direction), showing the height of the channel.

Seen on the obtained transmission scans of the chip in Figure 12, the channel has a well-defined width of 200 μm as determined at each end of the channel and at a point close to the centre. The coordinates for the channel position is the same at all three positions, indicating that the channel is oriented horizontally as expected, without any significant tilt. It can be seen that the chip is in fact not centred in the chip holder, but since the channel is still placed within the holder opening, this will not influencing the measurements. The absolute value of the transmission appears to differ along the channel, which could indicate that the height is not constant (slightly larger path length due to window deformation, increasing with distance from the channel ends), as it is also seen in Figure 12 in the right panel. This might be improved by optimisation of the manufacturing procedure. However, this difference in transmission will not affect the results obtained as all data are normalized by transmission measured in the same position. Furthermore, buffer measurements are performed in the same position to use for subtraction of background scattering, compensating for any difference in background scattering at the different positions.

In addition to the dimensional characteristics, it is noticed that the channel position subjects to minor movements over time. The profile of a transmission scan across the channel, performed at a later time during the beamtime is also shown in Figure 12 and reveals a slight position change compared to the earlier transmission scans. This may be caused by minor inaccuracy from the xy robot on where the chip is mounted or by instabilities of the chip holder. Regular transmission scans were for this reason necessary during the beamtime to inspect potential movements of the channel position in relation to the beam, enabling an adequate determination of the coordinates for each measurement relative to mixing point and chip walls.

Characterization of sheet width

A narrow protein sheet is favourable for fast mixing and low sample consumption. It was therefore interesting to test how narrow a sample sheet could be obtained with the chip, also determining the lowest possible mixing time. As for the tests at P12, a BSA solution together with a buffer solution was applied for testing the performance of the pumps and chip. It was evaluated if sheet widths could be controlled by changing pump speeds. Different BSA protein sheets widths were applied surrounded by buffer and characterized by scanning across the channel width in 5 μm steps. Sheets with expected widths of 5, 10 and 20 μm were formed, and BSA scattering signal was expected within a well-defined width of the channel, when performing the measurements close to the mixing point. No structural change of the protein was induced, as this experiment was only intended to test the performance of pumps and the chip.
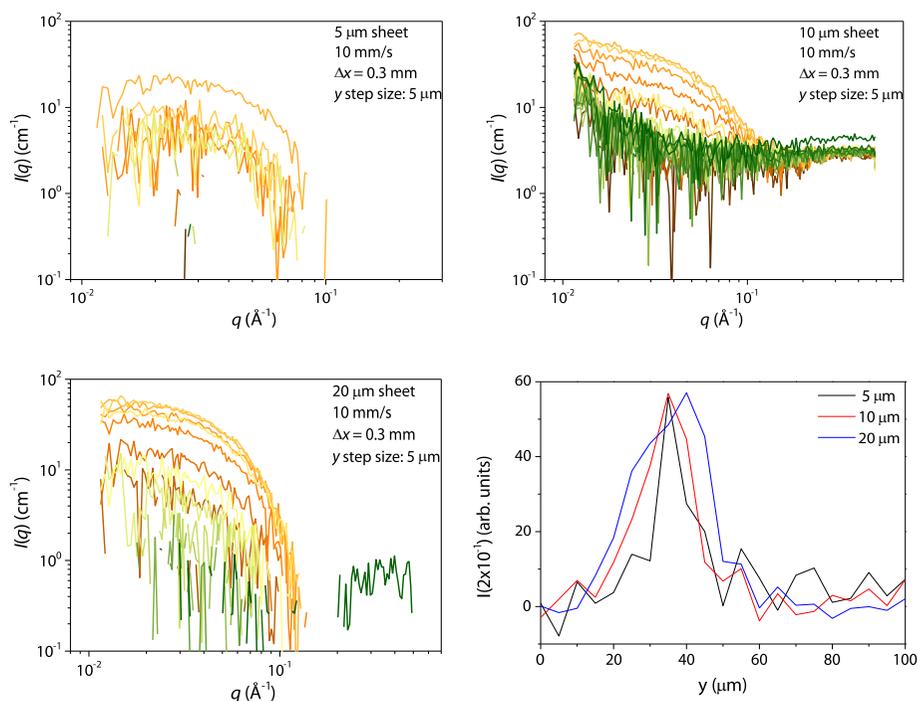
Figure 13. Characterization of different protein sheet widths. Varying BSA flowrate, while keeping the total flow rate at 10 mm/s, to obtain protein sheets corresponding to 5, 10 and 20 μm. Measurements were taking across (y-direction) the channel with 5 μm increments for 100 μm in the center of the channel. Acquisition time 0.2 s at each point. Bottom right show intensity as a function to distance in y-direction of the three different measurements.

From the collected data across the channel presented in Figure 13 it is seen that the protein sheet can more or less be controlled in widths of 5-20 μm. This can limit the protein sample consumption to 1.8 μl/min when using a 5 μm protein sheet and a 10 mm/s flow rate. This also means that fast mixing times can be obtained, of cause depending on the size of the molecules as diffusion rate decreases with increasing molecule size. However, the sheets are not perfectly centered in the channel but shifted approximately 10 μm to one side, implying a slightly higher pressure from one side of the inlets. Very bad background subtractions for many of the

measurements were unfortunately obtained for the 5 and 20 μm protein sheet. This inconsistency may be due to the chip setup as it was noticed that the chip did not remain fixed at the exact positions in relation to the beam. It implies that the chip could have moved in relation to the beam during measurements causing that the sample and background measurements are not in the exact same position, giving rise to over subtracted data.

Unfolding of BSA

To test the performance of the chip for kinetic measurements, the unfolding of BSA by SDS was applied as at P12. Due to the much smaller beam, a narrower protein sheet could be applied, giving much faster mixing times. The faster mixing time would hopefully allow for kinetic studies of the BSA unfolding after complete mixing, as opposed to the results from P12, which mainly showed the structural response upon increased SDS concentration as the mixing progressed with increasing distance from the mixing point. The flow rate was kept at 10 mm/s and the protein sheet width at 10 μm for faster mixing. Data were collected at different positions along the channel, corresponding to different time points from the mixing point.
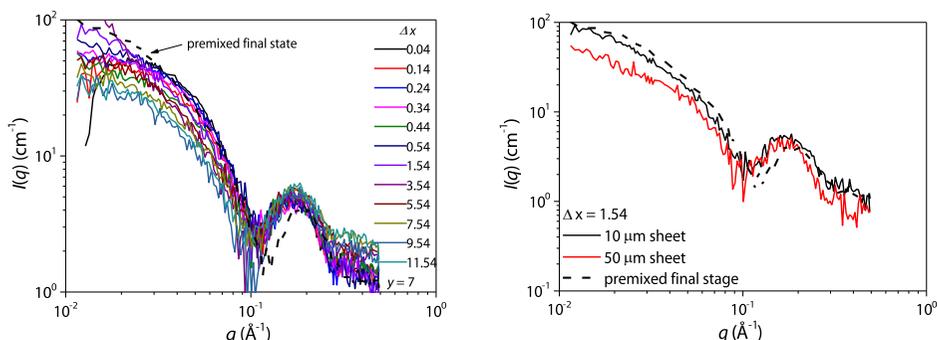
Figure 14. Mixing of BSA and SDS. Left, measurements downstream the channel for a 10 μm protein sheet, with a 10 mm/s flowrate, corresponding to different time points, compared to a premixed final state containing 6 mg/ml BSA and 40 mM SDS (dashed). The distance from the mixing point, Δx, can be translated into a time after the mixing point, Δt by Δt = vΔx, where v is the flow rate, v = 10 mm/s. Right, comparison of 0.15 s after mixing using a 10 vs 50 μm protein sheet. Acquisition time 0.2 s at each point.

The scattering data in Figure 14 show that mixing of BSA and SDS occurs over time. Using a 10 μm sheet, it takes SDS 0.028 s to diffuse half a sheet width, meaning that mixing is reached at approximately 0.28 mm from the mixing point, when using a 10 mm/s flow rate. This implies that any structural transition beyond this point reflects structural kinetics. It can be seen that some structural changes happens over the probed times as the scattering pattern slightly changes. However, the final state of the BSA-SDS complex is not reached, when compared to a premixed sample in Figure 14. The transition to the final state is apparently too slow to be caught within the timescales of the chip.

In Figure 14, data obtained at a certain distance from the mixing point are compared for a 10 and a 50 μm protein sheet. It is seen that the two structures are not identical, which is caused by the fact that mixing is much slower in the wider sheet, estimated to be 0.71 s, corresponding to

132

7.1 mm from the mixing point, and therefore, the mixing is not finished at the position where the data were collected.

Calcium binding of calmodulin

To address a structural transition of a protein due to ligand binding, calcium free calmodulin and a calcium containing buffer was mixed on the chip. Calmodulin's structural dynamics have earlier been investigated both by using a microfluidic mixer and by SAXS and should have a structural transition events at 490 µs and 20 ms after ligand binding (Park et al, 2008). For a 10 µm wide protein sheet, mixing will take approximately 10 ms, assuming a diffusion constant for $CaCl_2$ of 1.2×10^{-9} m^2/s (Hall et al, 1953). However, calcium at lower concentrations will be present with the calmodulin also at earlier times, and we can therefore expect to be able to follow the slowest of the transitions. Flow rate and protein sheet width were varied to find the ideal mixing rate.
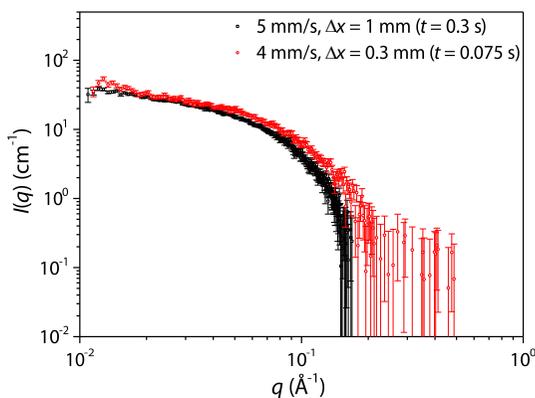
Figure 15. Mixing of calmodulin and calcium with a 50 µm protein sheet. Measurements at two different positions downstream the channel, corresponding to 300 and 75 ms after mixing point. Acquisition time 50x0.2 s at each point for 1 mm and 1s for 0.3 mm.

It was unfortunately not possible to observe any kinetics in this experiment. According to the literature one transition event of calmodulin from its unbound to its calcium bound state should happen within 20 ms. Probing this structural event was not feasible due to diffusion of calmodulin out of the beam, leading to dilution of the scattering signal. From the obtained scattering data in Figure 15 it is seen that low scattering intensity is obtained even for very long acquisition times at only 1 mm from the mixing point, using a relatively wide protein sheet of 50 µm width, which was necessary to decrease the speed of protein diffusion out of the X-ray beam. However, this also leads to unacceptably slow mixing. For $CaCl_2$ it takes 0.26 s to diffuse an average distance of 25 µm corresponding to half a sheet width. Calmodulin is a relatively small protein ~17 kDa and diffuses much faster than for instance BSA ~66 kDa. For SAXS microfluidics we therefore need a larger protein with slow diffusion time to achieve substantial scattering signal throughout the channel to probe a structural transitions.

General method discussion

Using a microfluidic chip for diffusional mixing of protein and ligand for induced structural transition probed by SAXS is challenging and not yet much applied. The described method is still new and unestablished and some development remains for its standard implementation in structural dynamics studies. However it is a very sophisticated method which could give important insights in the structure-function relationship of proteins.

The chips designed for this project meet the criteria for cheap manufacturing and low sample consumption mentioned in the introduction. However, some limitations do apply to this setup. The narrow sample sheet required for fast mixing and low sample consumption does set some restrictions to the protein sample. As diffusion rate is related to size of a molecule, small proteins will obviously diffuse out of the sheet faster than larger proteins. This limits the time resolution which can be obtained for small proteins compared to large proteins. The problem is clearly observed from the data collected of calmodolin, as the protein intensity in the sheet becomes too low for any actual structural changes to be probed. Proteins of a considerable size are therefore better suited for this method, such as BSA, where protein kinetics could be observed when mixed with SDS. A small beam size is necessary to be compatible with the narrow sample sheet in order to obtain an adequate signal compared to background scattering from the surrounding solution/buffer.

Also, the described method shows very promising results and further characterization and developments should be continued. Experiments using larger proteins are under planning.

Conclusion

A microfluidic mixing chip for time-resolved SAXS studies has been developed for investigating structural dynamics of proteins. It has been designed with a motive for uncomplicated and cheap manufacturing enabling easy accessibility to the method. The performance of the chip has been tested delivering satisfactory results. Flowrates can be controlled by syringe pumps, and a controllable protein sheet width down to 5 μm can be obtained, giving mixing times of a few milliseconds, depending on the diffusion constant of the ligand, and assuring limited sample consumption. The limitation for the method is the applied protein size, as small proteins diffuse out of the X-ray beam within the probed time scales. The method is moreover depended on a sophisticated instrument setup able to deliver a high flux micro-focus beam, such as the BioCAT beamline.

Chapter 8 – Final Remarks

Discussion and perspectives

A central focus of this thesis has been to optimise reconstitution of membrane proteins in nanodiscs for small angle scattering studies. Different aspects of sample optimisation were investigated in order to obtain new knowledge about the nanodisc and the reconstitution process. Firstly, the self-assembly of nanodiscs were structurally explored to gain a fundamental understanding of this process. D_2O buffer conditions were investigated with the interest to decrease aggregation of nanodiscs to permit collecting high quality non-aggregated SANS data. Secondly, reconstitution conditions were varied for proteorhodopsin to identify general influences on the nanodisc-membrane protein complexation. These experiences provided valuable information leading to successful reconstitution of CorA, which was structurally characterised. This was also adopted for the successful reconstitution of DsbB, used in protein assays. Finally, to accommodate the need of new methods for studying protein dynamics with SAS, a microfluidic chip was developed and tested. Very promising results were obtained from these experiments towards enabling more sophisticated future structural dynamic studies of membrane proteins.

Many of the projects presented in this thesis demonstrate examples of how nanodiscs can be utilized. The reasons for using nanodiscs, in contrast to detergent micelles, when studying membrane proteins, are many. Firstly, membrane proteins in detergent micelles do not necessarily reflect their native state due to a very unnatural environment. Secondly, many detergents compromise protein activity by destabilising the native structure of the membrane proteins (Bowie, 2001). The lipid bilayer offered by the nanodisc is more native like and stabilising for a membrane protein compared to detergents (Lyukmanova et al, 2012).

From a structural aspect, the size of the detergent-protein aggregate varies highly with type of detergent and membrane protein. Ideally detergents form perfect spheres but can also be found as oblate or rod like structures with large size distributions (Lichtenberg et al, 1983; Lin et al, 1986; Lin et al, 1987). The shape of detergent-protein micelles is less well defined as the size and structure highly depends on the size and shape of the transmembrane domain of the protein (Moller & Lemaire, 1993). This makes the micellar structure around a membrane protein hard to predict theoretically and numerically, whereas the nanodisc on the contrary has proved only minor variance regardless shape of membrane protein.

Contrast variation important for sufficient scattering information

When studying structures of membrane proteins in nanodiscs it is desirable to collect additional scattering contrasts. This is mainly due of the very sparse information obtained of the lipid bilayer from SAXS due to the low contrast of the lipids. A powerful solution is to collect complementary SANS data of the sample in D_2O buffer revealing substantial more information about the lipids and the sample in general. This strengthens the subsequent data analysis of the nanodisc-membrane protein complex by minimizing the freedom in the structure modelling

process (Pedersen et al, 2014; Skar-Gislinge et al, 2010). In theory obtaining a SANS contrast seems straightforward but experiences show that biomolecules tend to coagulate or aggregate in D_2O. In Project C, in this thesis, this problem was addressed, though no sustainable solution was found for stabilising nanodiscs in D_2O. This was aside from finding that the buffer exchange method has a great impact on the sample stability.

Aggregation is an unfortunate problem with serious impacts on the scattering data and in turn data analysis. This is due to the fact that aggregation hampers the scattering data, observed as a slight upturn in the low q-region. The information which can be derived from SANS scattering data of aggregated biomolecules is very sparse if not trivial. This problem has to be solved for SANS experiments to serve as supplementary strategy to SAXS for these types of samples. A strategy differently from the one treated in Project C could be to modify the MSP towards increasing stability in D_2O. Systematically changing the solvent exposed amino acids may result in a combination stabilising the nanodisc. To increase the prospect of such project, collaborations with experts in D_2O and proteins should be initiated. The deuteration laboratory in Grenoble is an excellent proposal for such collaboration, who has proved great expertise within this field (Compton et al, 2014; Maric et al, 2014).

Another approach is to obtain different SAXS contrasts as substitution for a SANS contrast. By adding a contrast agent to the buffer, different contrasts of the investigated protein can be obtained. This has been approached adding sucrose to the solution and a 65% (w/v) sucrose concentration has experimentally shown to match out proteins (Grishaev et al, 2012). This could be useful in nanodisc studies and is already being investigated by Søren Roi Midtgaard at the Niels Bohr institute. He has found that sucrose unfortunately causes disruption of the nanodisc. Replacing sucrose for glucose stabilised the nanodisc but creates a very visible exclusion zone around the nanodisc. However, very promising results have been obtained using cesium iodide and it is currently investigated how much information that can be retrieved at different concentrations (Midtgaard, 2015).

As a comment to contrast variation experiments one may ask if analysing a protein in a D_2O based solution or at 65% sucrose is still a native-like environment? This considering the fact that high D_2O concentrations are poisonous to most organisms and we observe destabilising effects on proteins. This is an interesting reflection when claiming SAS is used for structural characterization of proteins in close to natural environment.

Reconstitution of membrane proteins in nanodiscs

Despite the great advantages of using nanodiscs in membrane protein studies, obtaining quality samples for structural studies remain somewhat challenging. Though, taking advantage of experiences obtained from Project B, D and E will certainly make it less complicated for future projects. The key essentials for successful reconstitutions are high quality starting materials (pure proteins) and practice. Nevertheless, the projects presented in this thesis demonstrated that the nanodisc system still may possess unknown properties. When nanodiscs are prepared without membrane proteins they form from two MSPs. This is also assumed to be the rule for nanodiscs prepared with membrane proteins. However, in Project D this surprisingly showed to depend on the embedded membrane protein. What mechanisms creating this assembly is unclear but

may be related to the size of the transmembrane domain of the membrane protein. Nevertheless, it is to my knowledge not before observed. A recent study of a membrane protein containing a slightly larger transmembrane domain than CorA, namely aquaporin, could not be reconstituted in nanodiscs using MSP1D1, though possible using the slightly larger MSP1E3D1 (Järvå et al, 2015). CorA is a large protein and from the dimensions of its crystal structure it should be able to fit in 10 nm diameter nanodiscs whereas for aquaporin this is more doubtful. This implies that an additional factor other than the actual "fit" within the disc size may have triggered this curious outcome of four MSPs per nanodisc. CorA is believed to hold high flexibility (Pfoh et al, 2012; Rapp et al, 2004) which may have been a contributing factor. It would be interesting to study membrane proteins in a range of sizes to further explore this phenomenon. In addition to that, a self-assembly study similar to the one presented in Project A, including a membrane protein, would be very helpful. Studying the time dependence of how the membrane protein is reconstituted in nanodiscs would be useful to understand what mechanisms govern this process. Though, this would be a very advanced study setting high demands to the data analysis. The project may be a good MD simulation study, implementing some of the knowledge obtained from the self-assembly study in Project A.

Nanodiscs in drug discovery

Nanodiscs have been used in numerous biochemical studies, assaying function, activity and ligand binding of membrane proteins (Bocquet et al, 2015; Laursen et al, 2014; McCoy et al, 2014). The reconstitution of DsbB in Project E clearly showed the force of nanodiscs in membrane protein assays. Though, using the nanodisc for drug based screening is not yet as well established. Due to the fact that many membrane proteins are drug targets, nanodiscs may serve as a useful tool in drug screening. DsbB had before been reconstituted in nanodiscs and applied in the screening of fragment ligands of small molecule drugs (Früh et al, 2010). Here nanodiscs provided a stabilising environment for DsbB during ligand screening and DsbB maintained its ligand binding capacity for multiple rounds of screenings. In addition, a higher signal to noise ratio could be obtained using nanodiscs as well as higher specific binding, when compared to conventional methods (Früh et al, 2010). Nanodiscs have also been used in drug screening of cytochrome P450 3A4. Cytochrome P450s are important membrane proteins in drug-drug interactions that potentially can lead to drug toxicity as they are major contributors in drug metabolism (Guengerich, 2008). Nanodiscs were successfully used for stabilising the membrane protein which normally tends to denature or inactivate when membrane bound or kept in detergents. The nanodisc enabled rapid screening for binding of multiple drug candidates (Das et al, 2009). Recently nanodiscs have been used to discover compounds blocking Alzheimer's disease related toxins. Membrane proteins from the central nervous system synapse were reconstituted in nanodiscs and used for identifying targets of neural binding cell-secreted oligomers implicated in Alzheimer's disease. Subsequently, this allowed screening of antagonists from a small molecule compound library (Wilcox et al, 2015).

In the mentioned studies, together with Project E presented in this thesis, it is shown that nanodiscs provide a stabilising membrane protein solubilising environment which allows for sophisticated drug screening. It enables screenings of membrane proteins usually not possible due to denaturation or inactivation using conventional methods. Many of the studies rely on

surface based techniques, fixing the membrane proteins to surfaces. This is advantageous as it allows for whole library screening containing thousands of molecules using minimal protein material (Cooper, 2004). However, as some proteins tend to inactivate on the classical used surfaces, such as cytochrome P450s (Das et al, 2009), developing techniques that can stabilise membrane proteins on surfaces are needed. It is clearly demonstrated that nanodiscs could accommodate this problem in drug screenings.

Nanodiscs for drug delivery

In addition to drug discovery, nanodiscs have been investigated for their potential use in drug delivery. A few studies suggest that nanodiscs could be useful in the application to low soluble drugs. A glycoprotein from the viral genome of influenza A virus, important in virus transmission, has been inserted in nanodiscs and used for immunization of mice (Bhattacharya et al, 2010). This elicited a broad effective protective immune response which could not be produced by administering the protein alone.

The acetycholine receptor found in skeletal muscle has also been inserted into nanodiscs and used as "antibody traps" for autoantibodies playing a critical role in the autoimmune disease myasthenia gravis (Meriggioli et al, 2011). This significantly decreased autoantibody levels and reduced signs of disease weakness in mice. Interesting when this protein was administered alone, appeared to aggravate the disease.

To further evaluate the potential of using the nanodisc for drug delivery the biodistribution of nanodiscs have been investigated (Huda et al, 2015). Nanodiscs were ^{64}Cu radiolabelled and the biodistribution followed by positron emission tomography (PET) in tumour bearing mice. The nanodiscs showed no signs of inducing any immune response and could be detected in the tumours 48 hours post injection.

The mentioned studies indicate that nanodiscs can act as a potent drug delivery system for membrane protein based drugs and may have great potential for anti-cancer drugs.

In solution protein dynamics

The motivations for studying dynamics of proteins are many, as mentioned several times in this thesis. Most importantly it can boost knowledge about the structure-function relationship of proteins. Spectroscopy techniques are among most commonly used methods for studying protein dynamics due to accessibility and sensitivity (Schmid, 2001). However these methods do not provide information of the structural details of the protein movements.

In solution techniques are preferred as a direct response can be monitored by the addition of a reactant. NMR is a very powerful technique in this context as high resolution structures can be obtained. Flexible protein segments can be visualized in detail and in multiple conformations. This is a highly developed technique used for numerous dynamics studies. For an interesting review on methods in protein dynamics and the advantages of NMR in this relation see (Henzler-Wildman & Kern, 2007). The drawback of NMR is that it is generally best suited smaller proteins as larger proteins subject to crowded correlation spectra and peak broadening (Yee et al, 2014). A few advances can be made to address these problems, allowing for studying proteins up to 100 kDa size. These include, among other things, ^{15}N labelling and perdeuteration of the protein (Kainosho et al, 2006; Tzakos et al, 2006), creating multiple

dimension correlation spectra. Though, this is costly and complex as proteins must be synthesized from modified amino acids or by cells grown in labelled media, for the proteins to obtain these modifications.

Using SAS for studying protein dynamics can be applied for a large range of protein sizes and no modifications of the proteins are needed. The drawback is the somewhat lower, structural resolution compared to NMR and the lack of techniques available for low sample consumption.

Time resolved SAXS/WAXS experiments of proteorhodopsin, bacteriorhodopsin and lately visual rhodopsin show very interesting and promising applications of this technique (Malmerberg et al, 2015; Malmerberg et al, 2011). Here the structural conformation change was induced by a laser due to the light sensitivity of these proteins. While time resolved SAS primarily has been applied for light sensitive proteins, most proteins are not stimulated by light (Reddington et al, 2013). New methods are needed to accommodate structural dynamics studies of non-light sensitive proteins.

The method described in Project F is a sophisticated development for structure dynamic studies using SAXS. This method could become a very valuable tool as very little sample is needed, crucial when working with complicated and fragile samples. It enables studies of proteins stimulated by the presence of a ligand and relies on fast mixing by diffusion. However, the technique is still in the developmental stage and has yet to reveal its full potential. Further advances to the microfluidic chip would be interesting and could include additional techniques than SAXS, such as spectroscopy techniques. This would increase the information which can be obtained from the dynamic study.

Collecting data using low sample consumption is achieved through a high energy micro focus X-ray beam. With the current development of synchrotrons toward high flux and small beam dimensions, not to mention the free electron lasers, are large step forward in this field (Boutet et al, 2012; Chapman et al, 2011). The developments point towards great advances within the next 10 years and will be very interesting to follow.

Combining SAS with MD silulations is a different and promising approach for structural dynamics studies. This was clearly demonstrated in Project D as SAXS data were used to evaluate the MD structure of CorA in its open conformation. The combination of SAS and MD gives the possibility to validate MD structures by experimental determination. In addition to that it enables obtaining structures with a level of detail usually impossible with SAS.

Structure determination using nanodiscs
In this project the focus has been to optimise the nanodisc for SAS, but nanodiscs have also been utilized for structural studies of membrane proteins by other techniques. The developments and experiences obtained in this project may thus also have significance to other methods.

Nanodiscs have been successfully applied in solution NMR structure determination of smaller membrane proteins (Glueck et al, 2009; Raschle et al, 2009; Shenkarev et al, 2010; Susac et al, 2014). A size optimised version of the nanodisc has even been developed for this purpose (Hagn et al, 2013). The advantage of using nanodiscs in NMR studies is that the effective molecular size can be kept fairly constant compared to detergent micelles. This means that the

NMR relaxation properties are kept constant whereas it for detergent micelles varies with the size of the membrane protein (Raschle et al, 2009). It has been found that diffusion rates of nanodiscs with or without membrane protein are nearly identical implying that the hydrodynamic properties of nanodiscs are only affected very little by the embedded membrane protein. This suggests that it may be possible to obtain high quality NMR spectra of large membrane proteins in nanodiscs (Susac et al, 2014). Though, this would only apply for proteins mainly positioned in the bilayer as a large soluble domain inevitable would add to the hydrodynamic radius. Furthermore, it must be kept in mind that studying large membrane proteins still rely on protein modifications as earlier mentioned.

Cryo-EM is a fast developing technique and seems very promising in future structure determination, but is at the moment limited by the need of large symmetrical proteins (Glaeser & Hall, 2011; Kuehlbrandt, 2014). Only a couple of cryo-EM studies using nanodiscs are published, applied in the structure determination of large membrane protein complexes (Frauenfeld et al, 2011; Gogol et al, 2013). It is unclear if nanodiscs hold some special advantages to cryo-EM aside from the fact that membrane proteins can be studied in a native like bilayer environment. However, I am convinced that nanodiscs could be very useful in future cryo-EM studies for this reason alone, as the technique gradually evolves towards the handling of smaller proteins.

In all cases, sample quality and data quality are closely connected regardless of method. SAS is a sensitive method for testing sample quality of different reconstitution conditions and could be used as a precursor for both NMR and cryo-EM studies. This could aid the sample optimisation of membrane proteins in nanodiscs for various structural studies.

Importance of cross-disciplinary collaborations

Reconstitution of membrane proteins in nanodiscs for structural studies is challenging and every new experiment needs optimisation specifically for the membrane protein in focus. It is difficult to imagine the nanodisc becoming a simple tool for structure studies of membrane proteins in the very near future, but the experience and knowhow obtained during this project contribute towards this goal. The developed protocol in Appendix 2 can be used as a guideline, but additional optimisation of parameters, such as salt concentrations and pH, should be considered for the individual membrane proteins. From my own experience, this protocol has been used for successful reconstitution of several membrane proteins. One of the most important factors for success with nanodisc reconstitution for structural studies relies on practise and experience. During my project I have taught several fellow researchers this method. This includes, among other, collaborations with Jeremy Lakey's group from Newcastle University in the reconstitution of *E. coli* ion transporter OmpF. Claus Løland's group at the University of Copenhagen, on *Aquifex aeolicus* amino acid transporter LeuT. Anja Fuglsang's group at the University of Copenhagen, on *Arabidopsis thaliana* AHA2 proton pump. Ozman Mirsa's group from the University of Copenhagen, on *E. coli* dipeptide transporter YbgH. Susanna Törnroth-Horsefield's group from University of Gothenburg, on spinach leaf aquaporin. Obtaining a sample of sufficient quality for small angle scattering studies has only succeeded for the few who spent substantial amount of time practising and developing the method for their specific

membrane protein (Järvå et al, 2015). Several have made good attempt but unfortunately with poor sample quality.

The successful projects were I have been more directly involved have been collaborations with groups who have specialized in certain families of membrane proteins. Collaborations with labs focusing on specific membrane proteins, able to deliver protein for nanodisc reconstitution, are necessary for continuation of this project. Purity of the membrane protein used for reconstitution is crucial for obtaining a nanodisc sample of adequate quality for small angle scattering. It is simply not possible to be self-contained with high quality membrane proteins in the group as we neither have the experience nor facilities for expressing and purifying various membrane proteins. On the other hand we have the experience in nanodisc reconstitution which I have seen failed for structural studies numerous times in other labs. The interplay between membrane protein experts, nanodisc experts and structure modelling experts are for this reason essential for future perspectives of this project.

As a last comment, I am convinced that combining cross-disciplinary techniques is a necessary precedency to study complicated biological systems. A close collaboration between biophysicists and biologists is crucial for success in these types of experiments. From own experiences this means that special effort must be taken from all collaborators to ensure recognition of diverse expectations in such as sample quality and method limitations.

Final conclusions

In this project nanodiscs have been applied and evaluated as a platform for studying membrane proteins in a detergent-free environment. They have been applied for both structural and functional studies of different membrane proteins.

Successful reconstitution of membrane proteins in nanodiscs for structural studies relies heavily on practice and protein purity. Protein purity is crucial to obtain a sample of adequate quality for small angle scattering as this method is very sensitive to minor impurities or partly degradation of the sample. This was for instance showed in Project C where future challenges lies in decreasing aggregation of biomolecules in D_2O. Currently this problem limits the information which can be obtained from SANS data when using D_2O as solvent, leaving the data analysis relying on a SAXS contrast alone, where the lipids are not highly visible.

Nanodiscs are very valuable for functional studies of membrane proteins. Reconstitution of DsbB showed that the protein could be contained in nanodiscs with higher activity and kept for substantial longer time compared to detergents and crude cell membrane. Using nanodiscs, membrane protein assays can be improved and broaden the possibilities in relation to assay design. For structural studies of membrane proteins it is somewhat more challenging. Higher demands are set to the sample quality and quantity and thereby the reconstitution process. However, this can succeed by following some general guidelines and through optimisation to the protein in focus.

New insights regarding the nanodisc were found in this project as it surprisingly showed that nanodiscs are able to form containing 4 MSPs, when reconstituting CorA. The mechanisms causing this assembly are unclear and would be interesting to study more closely. If this could be obtained for other membrane proteins, it would give a better understanding of the interplay between the nanodisc and membrane protein, such as size of the membrane domain or protein flexibility.

In relation to flexible proteins, this project indicated that solution structures are highly relevant to study, as crystal structures of such proteins may not always be the dominating conformation in solution. Crystal structures of flexible proteins may be constrained in the highest ordered conformation required for crystallization. CorA is an interesting example of such a flexible protein where, from our data, it appears that the solution structure is not the dominating found in the crystal. The project also showed that combining nanodiscs and small angle scattering show promising results for future use for studying flexible membrane proteins in solution.

A central goal in this project was to obtain time resolved insight on structural dynamic of membrane proteins. This was a rather ambitious goal where part of the road has now been paved. It has been carried out by means of developing a new method for studying protein dynamic by SAXS using a microfluidic chip for fast mixing. Also, it has been possible to characterise an interesting membrane protein system for such experiment, namely CorA.

However, significantly future efforts are still required to perform an experiment combining these two.

Throughout this project, cross-disciplinarily collaborations have been vital in order to access high quality sample material and sophisticated data analysis, among other things. High level expertise in several subdisciplines was necessary to elucidate the complex nature of the open conformation of CorA. This included experts in the expression and purification of CorA, experiences and expertise in the reconstitution of membrane proteins in nanodiscs and in SAXS data collection. For the data analysis the modelling relied on expertise in MD simulation and state of the art SAXS data analysis of multi contrast particles.

For the time resolved chip study several disciplines were fundamental to reach the level of advances made. Expertises in microfluidiscs and chip manufacturing as well as knowledge about relevant biological systems were needed. Immense support from beamline scientists and beamline technicians were also extremely important in this project for carrying out and customize the desired experiments. These examples clearly show that special efforts must be taken to work together interdisciplinary, also in the future, to strive towards obtaining further knowledge about challenging biomolecular problems generally and membrane proteins specifically.

References

Agre P, MacKinnon R. (2003) The Nobel Price in Chemistry 2003. *"for discoveries cencerning channels in cell membrane"*. The Royal Swedish Academy of Sciences, Sweden.

Alberts B (2004) Membrane Structure. In *Essential cell biology*, second edn, 12. New York; London: Garland Science

Als-Nielsen J, McMorrow D (2011) Small-angle X-ray scattering (SAXS). In *Elements of modern X-ray physics*, second edn. New York: Wiley

Ander P, Leunglouie L, Silvestri F (1979) POLYCATIONS .1. SODIUM, CALCIUM, AND SULFATE ION DIFFUSION-COEFFICIENTS IN AQUEOUS SALT-SOLUTIONS CONTAINING IONENE BROMIDES. *Macromolecules* **12:** 1204-1207

Angelov B, Angelova A, Drechsler M, Lesieur S (2015) Rapid mixing stopped-flow small-angle X-ray scattering study of lipoplex formation at beamline ID02@ESRF. *J Synch Investig* **9:** 105-110

APS. (2015) BioCAT beamline. Advanced Photon Source, http://www.aps.anl.gov/Beamlines/Directory/showbeamline.php?beamline_id=25, Accessed March 10 2015.

Ashcroft FM (2006) From molecule to malady. *Nature* **440:** 440-447

Axel R, Buck LB. (2004) The Nobel Price in Physiology or Medicine 2004. *"Odorant receptors and the organization of the olfactory system"*. The Royal Swedish Academy of Sciences, Sweden.

Bader M, Muse W, Ballou DP, Gassner C, Bardwell JCA (1999) Oxidative protein folding is driven by the electron transport system. *Cell* **98:** 217-227

Bader M, Muse W, Zander T, Bardwell J (1998) Reconstitution of a protein disulfide catalytic system. *J Biol Chem* **273:** 10302-10307

Bader MW, Xie T, Yu CA, Bardwell JCA (2000) Disulfide bonds are generated by quinone reduction. *J Biol Chem* **275:** 26082-26088

Bartesaghi A, Matthies D, Banerjee S, Merk A, Subramaniam S (2014) Structure of beta-galactosidase at 3.2-angstrom resolution obtained by cryo-electron microscopy. *Proceedings of the National Academy of Sciences of the United States of America* **111:** 11709-11714

Bayburt TH, Carlson JW, Sligar SG (1998) Reconstitution and imaging of a membrane protein in a nanometer-size phospholipid bilayer. *Journal of Structural Biology* **123:** 37-44

Bayburt TH, Grinkova YV, Sligar SG (2002) Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Letters* **2:** 853-856

Bayburt TH, Sligar SG (2010) Membrane protein assembly into Nanodiscs. *Febs Letters* **584:** 1721-1727

Berg JM, Tymoczko JL, Stryer L (2012) Lipids and Cell Membranes. In *Biochemistry*, Sixth Edition edn, 12. Basingstoke: W.H. Freeman

Bernado P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *Journal of the American Chemical Society* **129:** 5656-5664

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) PROTEIN DATA BANK - COMPUTER-BASED ARCHIVAL FILE FOR MACROMOLECULAR STRUCTURES. *Journal of Molecular Biology* **112:** 535-542

Bertani G (1951) STUDIES ON LYSOGENESIS .1. THE MODE OF PHAGE LIBERATION BY LYSOGENIC ESCHERICHIA-COLI. *Journal of Bacteriology* **62:** 293-300

Bhattacharya P, Grimme S, Ganesh B, Gopisetty A, Sheng JR, Martinez O, Jayarama S, Artinger M, Meriggioli M, Prabhakar BS (2010) Nanodisc-Incorporated Hemagglutinin Provides Protective Immunity against Influenza Virus Infection. *Journal of Virology* **84:** 361-371

Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, Franke D, Mark D, Zengerle R, Cipriani F, Fiedler S, Roessle M, Svergun DI (2015) Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *Journal of Applied Crystallography* **48:** 431-443

Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, ColladoVides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of Escherichia coli K-12. *Science* **277:** 1453-1462

Bocian W, Sitkowski J, Bednarek E, Tarnowska A, Kawecki R, Kozerski L (2008) Structure of human insulin monomer in water/acetonitrile solution. *Journal of Biomolecular Nmr* **40:** 55-64

Bocquet N, Kohler J, Hug MN, Kusznir EA, Rufer AC, Dawson RJ, Hennig M, Ruf A, Huber W, Huber S (2015) Real-time monitoring of binding events on a thermostabilized human A(2A) receptor embedded in a lipid bilayer by surface plasmon resonance. *Biochim Biophys Acta-Biomembr* **1848:** 1224-1233

Borch J, Hamann T (2009) The nanodisc: a novel tool for membrane protein studies. *Biological chemistry* **390:** 805-814

Bornhorst JA, Falke JJ (2000) Purification of proteins using polyhistidine affinity tags. *Applications of Chimeric Genes and Hybrid Proteins, Pt A* **326:** 245-254

Boutet S, Lomb L, Williams GJ, Barends TRM, Aquila A, Doak RB, Weierstall U, DePonte DP, Steinbrener J, Shoeman RL, Messerschmidt M, Barty A, White TA, Kassemeyer S, Kirian RA, Seibert MM, Montanez PA, Kenney C, Herbst R, Hart P, Pines J, Haller G, Gruner SM, Philipp HT, Tate MW, Hromalik M, Koerner LJ, van Bakel N, Morse J, Ghonsalves W, Arnlund D, Bogan MJ, Caleman C, Fromme R, Hampton CY, Hunter MS, Johansson LC, Katona G, Kupitz C, Liang M, Martin AV, Nass K, Redecke L, Stellato F, Timneanu N, Wang D, Zatsepin NA,

Schafer D, Defever J, Neutze R, Fromme P, Spence JCH, Chapman HN, Schlichting I (2012) High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science* **337:** 362-364

Bouvier ESP, Koza SM (2014) Advances in size-exclusion separations of proteins and polymers by UHPLC. *Trac-Trends in Analytical Chemistry* **63:** 85-94

Bowie JU (2001) Stabilizing membrane proteins. *Current Opinion in Structural Biology* **11:** 397-402

Brennich ME, Nolting J-F, Dammann C, Noeding B, Bauch S, Herrmann H, Pfohl T, Koester S (2011) Dynamics of intermediate filament assembly followed in micro-flow by small angle X-ray scattering. *Lab on a Chip* **11:** 708-716

Bu Z, Callaway DJE (2011) PROTEINS MOVE! PROTEIN DYNAMICS AND LONG-RANGE ALLOSTERY IN CELL SIGNALING. *Adv Protein Chem Struct Biol* **83:** 163-221

Bunney TD, Cole AR, Broncel M, Esposito D, Tate EW, Katan M (2014) Crystal Structure of the Human, FIC-Domain Containing Protein HYPE and Implications for Its Functions. *Structure* **22:** 1831-1843

Cammarata M, Levantino M, Schotte F, Anfinrud PA, Ewald F, Choi J, Cupane A, Wulff M, Ihee H (2008) Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering. *Nature Methods* **5:** 881-886

Carlson JW, Jonas A, Sligar SG (1997) Imaging and manipulation of high-density lipoproteins. *Biophysical Journal* **73:** 1184-1189

Carlsson A, Greengard P, Kandel E. (2000) The Nobel Price in Physiology or Medicine 2000. *"signal transduction in the nervous system"*. The Royal Swedish Academy of Sciences, Sweden.

Carrington JC, Dougherty WG (1988) A VIRAL CLEAVAGE SITE CASSETTE - IDENTIFICATION OF AMINO-ACID SEQUENCES REQUIRED FOR TOBACCO ETCH VIRUS POLYPROTEIN PROCESSING. *Proceedings of the National Academy of Sciences of the United States of America* **85:** 3391-3395

Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U, Doak RB, Maia FRNC, Martin AV, Schlichting I, Lomb L, Coppola N, Shoeman RL, Epp SW, Hartmann R, Rolles D, Rudenko A, Foucar L, Kimmel N, Weidenspointner G, Holl P, Liang M, Barthelmess M, Caleman C, Boutet S, Bogan MJ, Krzywinski J, Bostedt C, Bajt S, Gumprecht L, Rudek B, Erk B, Schmidt C, Hoemke A, Reich C, Pietschner D, Strueder L, Hauser G, Gorke H, Ullrich J, Herrmann S, Schaller G, Schopper F, Soltau H, Kuehnel K-U, Messerschmidt M, Bozek JD, Hau-Riege SP, Frank M, Hampton CY, Sierra RG, Starodub D, Williams GJ, Hajdu J, Timneanu N, Seibert MM, Andreasson J, Rocker A, Joensson O, Svenda M, Stern S, Nass K, Andritschke R, Schroeter C-D, Krasniqi F, Bott M, Schmidt KE, Wang X, Grotjohann I, Holton JM, Barends TRM, Neutze R, Marchesini S, Fromme R, Schorb S, Rupp D, Adolph M, Gorkhover T, Andersson I, Hirsemann H, Potdevin G, Graafsma H, Nilsson B, Spence JCH (2011) Femtosecond X-ray protein nanocrystallography. *Nature* **470:** 73-U81

Chen AD, Wu DH, Johnson CS (1995) DETERMINATION OF THE BINDING ISOTHERM AND SIZE OF THE BOVINE SERUM ALBUMIN-SODIUM DODECYL-SULFATE COMPLEX BY DIFFUSION-ORDERED 2D NMR. *Journal of Physical Chemistry* **99:** 828-834

Chen CH, Tow F, Berns DS (1984) SOLVENT ISOTOPE EFFECT ON THE DIFFERENCES IN STRUCTURE AND STABILITY BETWEEN NORMAL AND DEUTERATED PROTEINS. *Biopolymers* **23:** 887-896

Chin D, Means AR (2000) Calmodulin: a prototypical calcium sensor. *Trends in Cell Biology* **10:** 322-328

Christensen CJ, Nielsen A, Bahnsen A, Brown WK, Rustad BM (1967) HALF-LIFE OF FREE NEUTRON. *Physics Letters B* **B 26:** 11-13

Cleverley RM, Kean J, Shintre CA, Baldock C, Derrick JP, Ford RC, Prince SM (2015) The Cryo-Electron Microscopy Structure of the Cora Channel from Methanocaldococcus Jannaschii in Low Magnesium Conditions. *To be published*

Collet JF, Bardwell JCA (2002) Oxidative protein folding in bacteria. *Mol Microbiol* **44:** 1-8

Compton ELR, Page K, Findlay HE, Haertlein M, Moulin M, Zachariae U, Norman DG, Gabel F, Javelle A (2014) Conserved structure and domain organization among bacterial Slc26 transporters. *Biochemical Journal* **463:** 297-307

Cooper MA (2004) Advances in membrane receptor screening and analysis. *Journal of Molecular Recognition* **17:** 286-315

Covingto.Ak, Paabo M, Robinson RA, Bates RG (1968) USE OF GLASS ELECTRODE IN DEUTERIUM OXIDE AND RELATION BETWEEN STANDARDIZED PD (PAD) SCALE AND OPERATIONAL PH IN HEAVY WATER. *Analytical Chemistry* **40:** 700-706

Cunningham F, Deber CM (2007) Optimizing synthesis and expression of transmembrane peptides and proteins. *Methods* **41:** 370-380

Daley DO, Rapp M, Granseth E, Melen K, Drew D, von Heijne G (2005) Global topology analysis of the Escherichia coli inner membrane proteome. *Science* **308:** 1321-1323

Dalmas O, Cuello LG, Jogini V, Cortes DM, Roux B, Perozo E (2010) Structural Dynamics of the Magnesium-Bound Conformation of CorA in a Lipid Bilayer. *Structure* **18:** 868-878

Dalmas O, Sompornpisut P, Bezanilla F, Perozo E (2014) Molecular mechanism of Mg2+-dependent gating in CorA. *Nature Communications* **5**

Das A, Zhao J, Schatz GC, Sligar SG, Van Duyne RP (2009) Screening of Type I and II Drug Binding to Human Cytochrome P450-3A4 in Nanodiscs by Localized Surface Plasmon Resonance Spectroscopy. *Analytical Chemistry* **81:** 3754-3759

Das R, Baker D (2008) Macromolecular modeling with Rosetta. In *Annual Review of Biochemistry* Vol. 77, pp 363-382.

Deisenhofer J, Huber R, Michel H. (1988) The Nobel Prize in Chemistry 1988. *"for the determination of the three-dimensionel structure of a photosynthetic reaction centre"*. The Royal Swedish Academy of Sciences, Sweden.

Denisov IG, Grinkova YV, Lazarides AA, Sligar SG (2004) Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *Journal of the American Chemical Society* **126:** 3477-3487

Diamond. (2015) B21 Equipment Description. Diamond Light Source, http://www.diamond.ac.uk/Beamlines/Soft-Condensed-Matter/small-angle/B21/description.html, Accessed March 10 2015.

Ding Y, Fujimoto LM, Yao Y, Plano GV, Marassi FM (2015) Influence of the lipid membrane environment on structure and activity of the outer membrane protein Ail from Yersinia pestis. *Biochim Biophys Acta-Biomembr* **1848:** 712-720

Dougherty WG, Cary SM, Parks TD (1989) MOLECULAR GENETIC-ANALYSIS OF A PLANT-VIRUS POLYPROTEIN CLEAVAGE SITE - A MODEL. *Virology* **171:** 356-364

Drew DE, von Heijne G, Nordlund P, de Gier JWL (2001) Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli. *Febs Letters* **507:** 220-224

Dubendorff JW, Studier FW (1991) CONTROLLING BASAL EXPRESSION IN AN INDUCIBLE T7 EXPRESSION SYSTEM BY BLOCKING THE TARGET T7 PROMOTER WITH LAC REPRESSOR. *Journal of Molecular Biology* **219:** 45-59

Duerr UHN, Gildenberg M, Ramamoorthy A (2012) The Magic of Bicelles Lights Up Membrane Protein Structure. *Chemical Reviews* **112:** 6054-6074

Dumon-Seignovert L, Cariot G, Vuillard L (2004) The toxicity of recombinant proteins in Escherichia coli: A comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification* **37:** 203-206

Duprez W, Premkumar L, Halili MA, Lindahl F, Reid RC, Fairlie DP, Martin JL (2015) Peptide Inhibitors of the Escherichia coli DsbA Oxidative Machinery Essential for Bacterial Virulence. *Journal of Medicinal Chemistry* **58:** 577-587

EMBL. (2015) Technical characteristics of beamline P12. The European Molecular Laboratory, http://www.embl-hamburg.de/biosaxs/p12/characteristics.html, Accessed March 10 2015.

Eshaghi S, Niegowski D, Kohl A, Molina DM, Lesley SA, Nordlund P (2006) Crystal structure of a divalent metal ion transporter CorA at 2.9 angstrom resolution. *Science* **313:** 354-357

ESRF. (2015) BM29 beamline specifications. The European Synchrotron http://www.esrf.eu/home/UsersAndScience/Experiments/MX/About_our_beamlines/BM29 /beamline-specifications.html, Accessed March 10 2015.

Fekete S, Beck A, Veuthey J-L, Guillarme D (2014) Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *Journal of Pharmaceutical and Biomedical Analysis* **101:** 161-173

Finlay BB, Falkow S (1997) Common themes in microbial pathogenicity revisited. *Microbiology and Molecular Biology Reviews* **61:** 136-169

Fischetti R, Stepanov S, Rosenbaum G, Barrea R, Black E, Gore D, Heurich R, Kondrashkina E, Kropf AJ, Wang S, Zhang K, Irving TC, Bunker GB (2004) The BioCAT undulator beamline 18ID: a facility for biological non-crystalline diffraction and X-ray absorption spectroscopy at the Advanced Photon Source. *Journal of Synchrotron Radiation* **11:** 399-405

Fisher SJ, Helliwell JR (2008) An investigation into structural changes due to deuteration. *Acta Crystallographica Section A* **64:** 359-367

Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography* **42:** 342-346

Frauenfeld J, Gumbart J, van der Sluis EO, Funes S, Gartmann M, Beatrix B, Mielke T, Berninghausen O, Becker T, Schulten K, Beckmann R (2011) Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. *Nature Structural & Molecular Biology* **18:** 614-621

Früh V, Zhou Y, Chen D, Loch C, Eiso AB, Grinkova YN, Verheij H, Sligar SG, Bushweller JH, Siegal G (2010) Application of Fragment Based Drug Discovery to Membrane Proteins: Biophysical Identification of Ligands of the Integral Membrane Enzyme DsbB. *Chemistry & biology* **17:** 881-891

Garavito RM, Ferguson-Miller S (2001) Detergents as tools in membrane biochemistry. *The Journal of biological chemistry* **276:** 32403-32406

Giladi M, Sasson Y, Fang X, Hiller R, Buki T, Wang Y-X, Hirsch JA, Khananshvili D (2012) A Common Ca2+-Driven Interdomain Module Governs Eukaryotic NCX Regulation. *Plos One* **7**

Gilman AG, Rodbell M. (1994) The Nobel Price in Physiology or Medicine 1994. *"for their discovery of G-proteins and the role of these proteins in signal transduction in cells".* The Royal Swedish Academy of Sciences, Sweden.

Giovannardi S, Lando L, Peres A (1998) Flash photolysis of caged compounds: Casting light on physiological processes. *News in Physiological Sciences* **13:** 251-255

Giovannoni SJ, Bibbs L, Cho JC, Stapels MD, Desiderio R, Vergin KL, Rappe MS, Laney S, Wilhelm LJ, Tripp HJ, Mathur EJ, Barofsky DF (2005) Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* **438:** 82-85

Glaeser RM, Hall RJ (2011) Reaching the Information Limit in Cryo-EM of Biological Macromolecules: Experimental Aspects. *Biophysical Journal* **100:** 2331-2337

Glascock CB, Weickert MJ (1998) Using chromosomal lacI(Q1) to control expression of genes on high-copy-number plasmids in Escherichia coli. *Gene* **223:** 221-231

Glatter O (1977) A new method for the evaluation of small-angle scattering data. *Journal of Applied Crystallography* **10:** 415-421

Glatter O, Kratky O (1982) *Small-angle X-ray scattering*, London: Academic Press Inc.

Glomset JA (1968) PLASMA LECITHIN - CHOLESTEROL ACYLTRANSFERASE REACTION. *Journal of Lipid Research* **9:** 155-167

Glueck JM, Wittlich M, Feuerstein S, Hoffmann S, Willbold D, Koenig BW (2009) Integral Membrane Proteins in Nanodiscs Can Be Studied by Solution NMR Spectroscopy. *Journal of the American Chemical Society* **131:** 12060-12061

Gogol EP, Akkaladevi N, Szerszen L, Mukherjee S, Chollet-Hinton L, Katayama H, Pentelute BL, Collier RJ, Fisher MT (2013) Three dimensional structure of the anthrax toxin translocon-lethal factor complex by cryo-electron microscopy. *Protein Science* **22:** 586-594

Gorter E, Grendel F (1925) On bimolecular layers of lipoids on the chromocytes of the blood. *Journal of Experimental Medicine* **41:** 439-443

Gourdon P, Alfredsson A, Pedersen A, Mahnerberg E, Nyblom M, Widell M, Berntsson R, Pinhassi J, Braiman M, Hansson O, Bonander N, Karlsson G, Neutze R (2008) Optimized in vitro and in vivo expression of proteorhodopsin: A seven-transmembrane proton pump. *Protein Expression and Purification* **58:** 103-113

Graceffa R, Nobrega RP, Barrea RA, Kathuria SV, Chakravarthy S, Bilsel O, Irving TC (2013) Sub-millisecond time-resolved SAXS using a continuous-flow mixer and X-ray microbeam. *Journal of Synchrotron Radiation* **20:** 820-825

Grillo I (2009) Applications of stopped-flow in SAXS and SANS. *Current Opinion in Colloid & Interface Science* **14:** 402-408

Grishaev A, Anthis NJ, Clore GM (2012) Contrast-Matched Small-Angle X-ray Scattering from a Heavy-Atom-Labeled Protein in Structure Determination: Application to a Lead-Substituted Calmodulin-Peptide Complex. *Journal of the American Chemical Society* **134:** 14686-14689

Guengerich FP (2008) Cytochrome P450 and chemical toxicology. *Chemical Research in Toxicology* **21:** 70-83

Guzman LM, Belin D, Carson MJ, Beckwith J (1995) TIGHT REGULATION, MODULATION, AND HIGH-LEVEL EXPRESSION BY VECTORS CONTAINING THE ARABINOSE P-BAD PROMOTER. *Journal of Bacteriology* **177:** 4121-4130

Hagn F, Etzkorn M, Raschle T, Wagner G (2013) Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. *Journal of the American Chemical Society* **135:** 1919-1925

Halili MA, Bachu P, Lindahl F, Bechara C, Mohanty B, Reid RC, Scanlon MJ, Robinson CV, Fairlie DP, Martin JL (2015) Small Molecule Inhibitors of Disulfide Bond Formation by the Bacterial DsbA-DsbB Dual Enzyme System. *ACS Chemical Biology* **10:** 957-964

Hall JR, Wishaw BF, Stokes RH (1953) THE DIFFUSION COEFFICIENTS OF CALCIUM CHLORIDE AND AMMONIUM CHLORIDE IN CONCENTRATED AQUEOUS SOLUTIONS AT 25-DEGREES. *Journal of the American Chemical Society* **75:** 1556-1560

Hansen LH, Knudsen S, Sorensen SJ (1998) The effect of the lacY gene on the induction of IPTG inducible promoters, studied in Escherichia coli and Pseudomonas fluorescens. *Current Microbiology* **36:** 341-347

Hansen S (2000) Bayesian estimation of hyperparameters for indirect Fourier transformation in small-angle scattering. *Journal of Applied Crystallography* **33:** 1415-1421

Hansen S (2012) BayesApp: a web site for indirect transformation of small-angle scattering data. *Journal of Applied Crystallography* **45:** 566-567

Harder T, Scheiffele P, Verkade P, Simons K (1998) Lipid domain structure of the plasma membrane revealed by patching of membrane components. *Journal of Cell Biology* **141:** 929-942

Heimburg T (2007a) The Composition of Biological Membranes. In *Thermal Biophysics of Membranes*, pp 29-40. Wiley-VCH Verlag GmbH & Co. KGaA

Heimburg T (2007b) Membrane Structure. In *Thermal Biophysics of Membranes*, pp 15-27. Wiley-VCH Verlag GmbH & Co. KGaA

Henzler-Wildman K, Kern D (2007) Dynamic personalities of proteins. *Nature* **450:** 964-972

Heras B, Shouldice SR, Totsika M, Scanlon MJ, Schembri MA, Martin JL (2009) DSB proteins and bacterial pathogenicity. *Nat Rev Microbiol* **7:** 215-225

Hochuli E, Bannwarth W, Dobeli H, Gentz R, Stuber D (1988) GENETIC APPROACH TO FACILITATE PURIFICATION OF RECOMBINANT PROTEINS WITH A NOVEL METAL CHELATE ADSORBENT. *Bio-Technology* **6:** 1321-1325

Holtzer A, Emerson MF (1969) Utility of the concept of water structure in the rationalization of the properties of aqueous solutions of proteins and small molecules. *The Journal of Physical Chemistry* **73:** 26-33

Hopkins AL, Groom CR (2002) The druggable genome. *Nature Reviews Drug Discovery* **1:** 727-730

Huda P, Binderup T, Pedersen MC, Midgaard SR, Elema DR, Kjær A, Jensen M, Arleth L (2015) PET/CT based in vivo evaluation of 64Cu labelled nanodiscs in tumor bearing mice. *Plos One* **Accepted**

Hwang S, Shao Q, Williams H, Hilty C, Gao YQ (2011) Methanol Strengthens Hydrogen Bonds and Weakens Hydrophobic Interactions in Proteins - A Combined Molecular Dynamics and NMR study. *Journal of Physical Chemistry B* **115:** 6653-6660

Israelachvili JN (2007a) 16 - Thermodynamic Principles of Self-Assembly. In *Intermolecular and Surface Forces (Second Edition)*, pp 341-365. San Diego: Academic Press

Israelachvili JN (2007b) 17 - Aggregation of Amphiphilic Molecules into Micelles, Bilayers, Vesicles and Biological Membranes. In *Intermolecular and Surface Forces (Third Edition)*, pp 366-394. San Diego: Academic Press

Israelachvili JN, Mitchell DJ, Ninham BW (1976) THEORY OF SELF-ASSEMBLY OF HYDROCARBON AMPHIPHILES INTO MICELLES AND BILAYERS. *Journal of the Chemical Society-Faraday Transactions Ii* **72:** 1525-1568

Janknecht R, Demartynoff G, Lou J, Hipskind RA, Nordheim A, Stunnenberg HG (1991) RAPID AND EFFICIENT PURIFICATION OF NATIVE HISTIDINE-TAGGED PROTEIN EXPRESSED BY RECOMBINANT VACCINIA VIRUS. *Proceedings of the National Academy of Sciences of the United States of America* **88:** 8972-8976

Jonas A, Wald JH, Toohill KLH, Krul ES, Kezdy KE (1990) APOLIPOPROTEIN-A-I STRUCTURE AND LIPID PROPERTIES IN HOMOGENEOUS, RECONSTITUTED SPHERICAL AND DISCOIDAL HIGH-DENSITY-LIPOPROTEINS. *J Biol Chem* **265:** 22123-22129

Jones MN (1999) Surfactants in membrane solubilisation. *International Journal of Pharmaceutics* **177:** 137-159

Jung K-H (2007) The distinct signaling mechanisms of microbial sensory rhodopsins in Archaea, Eubacteria and Eukarya. *Photochemistry and Photobiology* **83:** 63-69

Järvå M, Kynde S, Huda P, Törnroth-Horsefield S, Arleth L (2015) SAXS and SANS investigation of SoPIP2;1 Aquaporin-tetramers in POPC-nanodiscs. Bench-mark of the nanodisc approach to extract structural information about membrane proteins *Submitted*

Kadokura H, Beckwith J (2002) Four cysteines of the membrane protein DsbB act in concert to oxidize its substrate DsbA. *Embo J* **21:** 2354-2363

Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Ono AM, Guntert P (2006) Optimal isotope labelling for NMR protein structure determinations. *Nature* **440:** 52-57

Kang HJ, Lee C, Drew D (2013) Breaking the barriers in membrane protein crystallography. *International Journal of Biochemistry & Cell Biology* **45:** 636-644

Katritch V, Cherezov V, Stevens RC (2013) Structure-Function of the G Protein-Coupled Receptor Superfamily. In *Annual Review of Pharmacology and Toxicology, Vol 53, 2013*, Insel PA (ed), Vol. 53, pp 531-556.

Kawano M, Miida T, Fielding CJ, Fielding PE (1993) QUANTITATION OF PRE-BETA-HDL-DEPENDENT AND NONSPECIFIC COMPONENTS OF THE TOTAL EFFLUX OF CELLULAR CHOLESTEROL AND PHOSPHOLIPID. *Biochemistry* **32:** 5025-5028

Kay LE (2005) NMR studies of protein structure and dynamics. *Journal of Magnetic Resonance* **173:** 193-207

Koch MH, Vachette P, Svergun DI (2003) Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q Rev Biophys* **36:** 147-227

Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiyagarajan P, Sosnick TR, Hasan MZ, Pande VS, Ruczinski I, Doniach S, Plaxco KW (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 12491-12496

Kohn WD, Kay CM, Hodges RS (1997) Salt effects on protein stability: Two-stranded alpha-helical coiled-coils containing inter- or intrahelical ion pairs. *Journal of Molecular Biology* **267:** 1039-1052

Kolli S, Meng X, Wu X, Shengjuler D, Cameron CE, Xiang Y, Deng J (2015) Structure-Function Analysis of Vaccinia Virus H7 Protein Reveals a Novel Phosphoinositide Binding Fold Essential for Poxvirus Replication. *Journal of Virology* **89:** 2209-2219

Kuehlbrandt W (2014) Cryo-EM enters a new era. *Elife* **3**

Kuwamoto S, Akiyama S, Fujisawa T (2004) Radiation damage to a protein solution, detected by synchrotron X-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants. *Journal of Synchrotron Radiation* **11:** 462-468

Kynde SAR, Skar-Gislinge N, Pedersen MC, Midtgaard SR, Simonsen JB, Schweins R, Mortensen K, Arleth L (2014) Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment. *Acta Crystallographica Section D-Biological Crystallography* **70:** 371-383

Labrador A, Cerenius Y, Svensson C, Theodor K, Plivelic T (2012) The yellow mini-hutch for SAXS experiments at MAX IV Laboratory. In *11th International Conference on Synchrotron Radiation Instrumentation*, Vol. Conference Series 424. Lyon, France

Laursen T, Singha A, Rantzau N, Tutkus M, Borch J, Hedegard P, Stamou D, Moller BL, Hatzakis NS (2014) Single Molecule Activity Measurements of Cytochrome P450 Oxidoreductase Reveal the Existence of Two Discrete Functional States. *Acs Chemical Biology* **9:** 630-634

le Maire M, Champeil P, Moller JV (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta-Biomembr* **1508:** 86-111

Lefkowitz RJ, Kobilka BK. (2012) The Nobel Prize in Chemistry 2012. *"for studies of G-protein coupled receptors"*. The Royal Swedish Academy of Sciences, Sweden.

Levantino M, Schiro G, Lemke HT, Cottone G, Glownia JM, Zhu D, Chollet M, Ihee H, Cupane A, Cammarata M (2015) Ultrafast myoglobin structural dynamics observed with an X-ray free-electron laser. *Nature communications* **6:** 6772-6772

Lichtenberg D, Robson RJ, Dennis EA (1983) SOLUBILIZATION OF PHOSPHOLIPIDS BY DETERGENTS - STRUCTURAL AND KINETIC ASPECTS. *Biochimica Et Biophysica Acta* **737:** 285-304

Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Tan S (2005) Comparison of affinity tags for protein purification. *Protein Expression and Purification* **41:** 98-105

Lin TL, Chen SH, Gabriel NE, Roberts MF (1986) USE OF SMALL-ANGLE NEUTRON-SCATTERING TO DETERMINE THE STRUCTURE AND INTERACTION OF DIHEXANOYLPHOSPHATIDYLCHOLINE MICELLES. *Journal of the American Chemical Society* **108:** 3499-3507

Lin TL, Chen SH, Gabriel NE, Roberts MF (1987) SMALL-ANGLE NEUTRON-SCATTERING TECHNIQUES APPLIED TO THE STUDY OF POLYDISPERSE RODLIKE DIHEPTANOYLPHOSPHATIDYLCHOLINE MICELLES. *Journal of Physical Chemistry* **91:** 406-413

Lindblom G, Rilfors L (1992) NONLAMELLAR PHASES FORMED BY MEMBRANE-LIPIDS. *Advances in Colloid and Interface Science* **41:** 101-125

Lu P, Bai X-c, Ma D, Xie T, Yan C, Sun L, Yang G, Zhao Y, Zhou R, Scheres SHW, Shi Y (2014) Three-dimensional structure of human gamma-secretase. *Nature* **512:** 166-170

Luecke H, Schobert B, Richter HT, Cartailler JP, Lanyi JK (1999) Structure of bacteriorhodopsin at 1.55 angstrom resolution. *Journal of Molecular Biology* **291:** 899-911

Lunin VV, Dobrovetsky E, Khutoreskaya G, Zhang RG, Joachimiak A, Doyle DA, Bochkarev A, Maguire ME, Edwards AM, Koth CM (2006) Crystal structure of the CorA Mg2+ transporter. *Nature* **440:** 833-837

Lusk JE, Williams RJP, Kennedy EP (1968) MAGNESIUM AND GROWTH OF ESCHERICHIA COLI. *J Biol Chem* **243:** 2618-2624

Lyukmanova EN, Shenkarev ZO, Khabibullina NF, Kopeina GS, Shulepko MA, Paramonov AS, Mineev KS, Tikhonov RV, Shingarova LN, Petrovskaya LE, Dolgikh DA, Arseniev AS, Kirpichnikov MP (2012) Lipid-protein nanodiscs for cell-free production of integral membrane proteins in a soluble and folded state: Comparison with detergent micelles, bicelles and liposomes. *Biochim Biophys Acta-Biomembr* **1818:** 349-358

Mabanglo MF, Hast MA, Lubock NB, Hellinga HW, Beese LS (2014) Crystal structures of the fungal pathogen Aspergillus fumigatus protein farnesyltransferase complexed with substrates and inhibitors reveal features for antifungal drug design. *Protein Science* **23:** 289-301

Makhatadze GI, Clore GM, Gronenborn AM (1995) SOLVENT ISOTOPE EFFECT AND PROTEIN STABILITY. *Nature Structural Biology* **2:** 852-855

Makrides SC (1996) Strategies for achieving high-level expression of genes in Escherichia coli. *Microbiological Reviews* **60:** 512-538

Maleknia SD, Ralston CY, Brenowitz MD, Downard KM, Chance MR (2001) Determination of macromolecular folding and structure by synchrotron X-ray radiolysis techniques. *Analytical Biochemistry* **289:** 103-115

Malmberg E, Bovee-Geurts PHM, Katona G, Deupi X, Arnlund D, Wickstrand C, Johansson LC, Westenhoff S, Nazarenko E, Schertler GFX, Menzel A, de Grip WJ, Neutze R (2015) Conformational activation of visual rhodopsin in native disc membranes. *Science Signaling* **8**

Malmberg 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, Neutze R (2011) Time-Resolved WAXS Reveals Accelerated Conformational Changes in Iodoretinal-Substituted Proteorhodopsin. *Biophysical Journal* **101:** 1345-1353

Maric S, Skar-Gislinge N, Midgaard S, Thygesen MB, Schiller J, Frielinghaus H, Moulin M, Haertlein M, Forsyth VT, Pomorski TG, Arleth L (2014) Stealth carriers for low-resolution structure determination of membrane proteins in solution. *Acta Crystallographica Section D-Biological Crystallography* **70:** 317-328

Markwick PRL, Malliavin T, Nilges M (2008) Structural Biology by NMR: Structure, Dynamics, and Interactions. *Plos Computational Biology* **4**

Martin RB (1990) BIOINORGANIC CHEMISTRY OF MAGNESIUM. *Metal Ions in Biological Systems* **26:** 1-13

MaxLab. (2015) BEAMLINE I911-SAXS. The MAX IV Laboratory, https://www.maxlab.lu.se/node/35, Accessed March 10 2015.

McCoy JG, Rusinova R, Kim DM, Kowal J, Banerjee S, Cartagena AJ, Thompson AN, Kolmakova-Partensky L, Stahlberg H, Andersen OS, Nimigean CM (2014) A KcsA/MloK1 Chimeric Ion Channel Has Lipid-dependent Ligand-binding Energetics*. *J Biol Chem* **289:** 9535-9546

McCusker EC, Bagneris C, Naylor CE, Cole AR, D'Avanzo N, Nichols CG, Wallace BA (2012) Structure of a bacterial voltage-gated sodium channel pore reveals mechanisms of opening and closing. *Nature Communications* **3**

McDermott A (2009) Structure and Dynamics of Membrane Proteins by Magic Angle Spinning Solid-State NMR. In *Annual Review of Biophysics* Vol. 38, pp 385-403.

Mei X, Atkinson D (2011) Crystal Structure of C-terminal Truncated Apolipoprotein A-I Reveals the Assembly of High Density Lipoprotein (HDL) by Dimerization. *J Biol Chem* **286:** 38570-38582

Meilleur F, Contzen J, Myles DAA, Jung C (2004) Structural Stability and Dynamics of Hydrogenated and Perdeuterated Cytochrome P450cam (CYP101)†. *Biochemistry* **43:** 8744-8753

Meriggioli MN, Sheng JR, Bhattacharya P, Stowell MHB, Grimme S, Artinger M, Prabhakar BS (2011) In Vivo Adsorption of Autoantibodies in Myasthenia Gravis Using Nanodisc-Incorporated Acetylcholine Receptor. *Neurology* **76:** A286-A286

Mertens N, Remaut E, Fiers W (1995) TIGHT TRANSCRIPTIONAL CONTROL MECHANISM ENSURES STABLE HIGH-LEVEL EXPRESSION FROM T7 PROMOTER-BASED EXPRESSION PLASMIDS. *Bio-Technology* **13:** 175-179

Midtgaard S. (2015) SAXS constrast variation. Personal communication.

Midtgaard SR, Pedersen MC, Kirkensgaard JJK, Sorensen KK, Mortensen K, Jensen KJ, Arleth L (2014) Self-assembling peptides form nanodiscs that stabilize membrane proteins. *Soft Matter* **10:** 738-752

Miroux B, Walker JE (1996) Over-production of proteins in Escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology* **260:** 289-298

Molina R, Stella S, Redondo P, Gomez H, Marcaida MJ, Orozco M, Prieto J, Montoya G (2015) Visualizing phosphodiester-bond hydrolysis by an endonuclease. *Nature Structural & Molecular Biology* **22:** 65-72

Moller JV, Lemaire M (1993) DETERGENT BINDING AS A MEASURE OF HYDROPHOBIC SURFACE-AREA OF INTEGRAL MEMBRANE-PROTEINS. *J Biol Chem* **268:** 18659-18672

Moncrief MBC, Maguire ME (1999) Magnesium transport in prokaryotes. *Journal of Biological Inorganic Chemistry* **4:** 523-527

Moomaw AS, Maguire ME (2008) The Unique Nature of Mg2+ Channels. *Physiology* **23:** 275-285

Mouritsen OG, Bloom M (1984) MATTRESS MODEL OF LIPID-PROTEIN INTERACTIONS IN MEMBRANES. *Biophysical Journal* **46:** 141-153

Neely A, Hidalgo P (2014) Structure-function of proteins interacting with the alpha(1) pore-forming subunit of high-voltage-activated calcium channels. *Frontiers in Physiology* **5**

Neher E, Sakmann B. (1991) The Nobel Price in Physiology or Medicine 1991. *"for their discoveries concerning the function of single io channels in cells".* The Royal Swedish Academy of Sciences, Sweden.

Nielsen KRR (2013) SAXS-studier af insulinoligomerer i opløsning. Bachelor Thesis, Niels Bohr Institute, University of Copenhagen,

Nielsen SS, Toft KN, Snakenborg D, Jeppesen MG, Jacobsen JK, Vestergaard B, Kutter JP, Arleth L (2009) BioXTAS RAW, a software program for high-throughput automated small-

angle X-ray scattering data reduction and preliminary analysis. *Journal of Applied Crystallography* **42:** 959-964

NIST. (2015) Neutron scattering lengths and cross sections. NIST Center for Neutron Research, http://www.ncnr.nist.gov/resources/n-lengths/, Accessed April 23 2015.

Nour-Eldin HH, Hansen BG, Norholm MHH, Jensen JK, Halkier BA (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Research* **34**

Nygaard J, Høiberg-Nielsen R, Malik L, Hoeg-Jensen T, Arleth L (2015) Structural insight into the self-assembly of a pharmaceutically optimized insulin analogue as obtained by Small-Angle X-ray Scattering. *In preparation*

Olesen C, Picard M, Winther A-ML, Gyrup C, Morth JP, Oxvig C, Moller JV, Nissen P (2007) The structural basis of calcium transport by the calcium pump. *Nature* **450:** 1036-U1035

Onuchic JN, LutheySchulten Z, Wolynes PG (1997) Theory of protein folding: The energy landscape perspective. *Annual Review of Physical Chemistry* **48:** 545-600

Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289:** 739-745

Palombo I, Daley DO, Rapp M (2012) The Periplasmic Loop Provides Stability to the Open State of the CorA Magnesium Channel. *J Biol Chem* **287:** 27547-27555

Park HY, Kim SA, Korlach J, Rhoades E, Kwok LW, Zipfell WR, Waxham MN, Webb WW, Pollack L (2008) Conformational changes of calmodulin upon Ca2+ binding studied with a microfluidic mixer. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 542-547

Park HY, Qiu XY, Rhoades E, Korlach J, Kwok LW, Zipfel WR, Webb WW, Pollack L (2006) Achieving uniform mixing in a microfluidic device: Hydrodynamic focusing prior to mixing. *Analytical Chemistry* **78:** 4465-4473

Parks TD, Leuther KK, Howard ED, Johnston SA, Dougherty WG (1994) RELEASE OF PROTEINS AND PEPTIDES FROM FUSION PROTEINS USING A RECOMBINANT PLANT-VIRUS PROTEINASE. *Analytical Biochemistry* **216:** 413-417

Patra M, Salonen E, Terama E, Vattulainen I, Faller R, Lee BW, Holopainen J, Karttunen M (2006) Under the influence of alcohol: The effect of ethanol and methanol on lipid bilayers. *Biophysical Journal* **90:** 1121-1135

Paulson JC, Colley KJ (1989) GLYCOSYLTRANSFERASES - STRUCTURE, LOCALIZATION, AND CONTROL OF CELL TYPE-SPECIFIC GLYCOSYLATION. *J Biol Chem* **264:** 17615-17618

Payandeh J, Pai EF (2006) A structural basis for Mg2+ homeostasis and the CorA translocation cycle. *Embo J* **25:** 3762-3773

Payandeh J, Pfoh R, Pai EF (2013) The structure and regulation of magnesium selective ion channels. *Biochim Biophys Acta-Biomembr* **1828:** 2778-2792

PDB. (2015a) Histogram of expression organisms. Protein Data Bank, http://www.pdb.org/pdb/statistics/histogram.do?mdcat=entity_src_gen&mditem=pdbx_host _org_scientific_name&numOfbars=50&name=Expression%20Organism%20(Gene%20Source)#, Accessed Feb 10 2015.

PDB. (2015b) PDB Statisics on Proteins. Protein Data Bank, http://www.pdb.org/pdb/results/results.do?qrid=E6BBA33A&tabtoshow=Current, Accessed March 1 2015.

Pedersen JS (1997) Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting. *Advances in Colloid and Interface Science* **70:** 171-210

Pedersen JS. (2015) Photon flux from SAXS home sources. Personal communication.

Pedersen MC, Arleth L, Mortensen K (2013) WillItFit: a framework for fitting of constrained models to small-angle scattering data. *Journal of Applied Crystallography* **46:** 1894-1898

Pedersen MC, Hansen SL, Markussen B, Arleth L, Mortensen K (2014) Quantification of the information in small-angle scattering data. *Journal of Applied Crystallography* **47:** 2000-2010

Pernot P, Round A, Barrett R, Antolinos ADM, Gobbo A, Gordon E, Huet J, Kieffer J, Lentini M, Mattenet M, Morawe C, Mueller-Dieckmann C, Ohlsson S, Schmid W, Surr J, Theveneau P, Zerrad L, McSweeney S (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *Journal of Synchrotron Radiation* **20:** 660-664

Pfoh R, Li A, Chakrabarti N, Payandeh J, Pomes R, Pai EF (2012) Structural asymmetry in the magnesium channel CorA points to sequential allosteric regulation. *Proceedings of the National Academy of Sciences of the United States of America* **109:** 18809-18814

Polayes DA, Parks TD, Johnston SA, Dougherty WG (1998) Application of TEV Protease in Protein Production. *Methods in molecular medicine* **13:** 169-183

Pollack L, Tate MW, Darnton NC, Knight JB, Gruner SM, Eaton WA, Austin RH (1999) Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle x-ray scattering. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 10115-10117

Pollock NL, McDevitt CA, Collins R, Niesten PHM, Prince S, Kerr ID, Ford RC, Callaghan R (2014) Improving the stability and function of purified ABCB1 and ABCA4: The influence of membrane lipids. *Biochim Biophys Acta-Biomembr* **1838:** 134-147

Promega. (2014) Single Step (KRX) Competent Cells. Promega Corporation.

PSI. (2015) cSAXS. Poul Scherrer Institut, http://www.psi.ch/sls/csaxs/csaxs, Accessed March 10 2015.

Ranaghan MJ, Schwall CT, Alder NN, Birge RR (2011) Green Proteorhodopsin Reconstituted into Nanoscale Phospholipid Bilayers (Nanodiscs) as Photoactive Monomers. *Journal of the American Chemical Society* **133:** 18318-18327

Rapp M, Drew D, Daley DO, Nilsson J, Carvalho T, Melen K, De Gier JW, Von Heijne G (2004) Experimentally based topology models for E-coli inner membrane proteins. *Protein Science* **13:** 937-945

Raschle T, Hiller S, Yu T-Y, Rice AJ, Walz T, Wagner G (2009) Structural and Functional Characterization of the Integral Membrane Protein VDAC-1 in Lipid Bilayer Nanodiscs. *Journal of the American Chemical Society* **131:** 17777-17779

Rasmussen SGF, Choi H-J, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VRP, Sanishvili R, Fischetti RF, Schertler GFX, Weis WI, Kobilka BK (2007) Crystal structure of the human beta(2) adrenergic G-protein-coupled receptor. *Nature* **450:** 383-U384

Reddington SC, Rizkallah PJ, Watson PD, Pearson R, Tippmann EM, Jones DD (2013) Different Photochemical Events of a Genetically Encoded Phenyl Azide Define and Modulate GFP Fluorescence. *Angewandte Chemie-International Edition* **52:** 5974-5977

Rietsch A, Beckwith J (1998) The genetics of disulfide bond metabolism. *Annual review of genetics* **32:** 163-184

Rigaud JL, Levy D, Mosser G, Lambert O (1998) Detergent removal by non-polar polystyrene beads - Applications to membrane protein reconstitution and two-dimensional crystallization. *European Biophysics Journal with Biophysics Letters* **27:** 305-319

Riske KA, Dobereiner HG, Lamy-Freund MT (2003) Comment on "Gel-Fluid transition in dilute versus concentrated DMPG aqueous dispersions". *Journal of Physical Chemistry B* **107:** 5391-5392

Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, Sligar SG (2009) *Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs*, SAN DIEGO: ELSEVIER ACADEMIC PRESS INC.

Robertson EMT, Ubbelohde AR (1939) Structure and thermal properties associated with some hydrogen bonds in crystals I. The isotope effect. *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences* **170:** 0222-0240

Robinson RA, Paabo M, Bates RG (1969) DEUTERIUM ISOTOPE EFFECT ON DISSOCIATION OF WEAK ACIDS IN WATER AND DEUTERIUM OXIDE. *Journal of Research of the National Bureau of Standards Section a-Physics and Chemistry* **A 73:** 299-+

Rouser G, Siakotos AN, Fleische.S (1966) QUANTITATIVE ANALYSIS OF PHOSPHOLIPIDS BY THIN-LAYER CHROMATOGRAPHY AND PHOSPHORUS ANALYSIS OF SPOTS. *Lipids* **1:** 85-&

Russ AP, Lampel S (2005) The druggable genome: an update. *Drug Discovery Today* **10:** 1607-1610

Santos SF, Zanette D, Fischer H, Itri R (2003) A systematic study of bovine serum albumin (BSA) and sodium dodecyl sulfate (SDS) interactions by surface tension and small angle X-ray scattering. *Journal of Colloid and Interface Science* **262:** 400-408

Schmid F-X (2001) Biological Macromolecules: UV-visible Spectrophotometry. In *eLS*. John Wiley & Sons, Ltd

Schmidt T, Bergner A, Schwede T (2014) Modelling three-dimensional protein structures for applications in drug design. *Drug Discovery Today* **19:** 890-897

Schrödinger L. (2015) The PyMOL Molecular Graphics System. pymol.org.

Seddon AM, Curnow P, Booth PJ (2004) Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta-Biomembr* **1666:** 105-117

Shenkarev ZO, Lyukmanova EN, Butenko IO, Petrovskaya LE, Paramonov AS, Shulepko MA, Nekrasova OV, Kirpichnikov MP, Arseniev AS (2013) Lipid-protein nanodiscs promote in vitro folding of transmembrane domains of multi-helical and multimeric membrane proteins. *Biochim Biophys Acta-Biomembr* **1828:** 776-784

Shenkarev ZO, Paramonov AS, Lyukmanova EN, Shingarova LN, Yakimov SA, Dubinnyi MA, Chupin VV, Kirpichnikov MP, Blommers MJJ, Arseniev AS (2010) NMR Structural and Dynamical Investigation of the Isolated Voltage-Sensing Domain of the Potassium Channel KvAP: Implications for Voltage Gating. *Journal of the American Chemical Society* **132:** 5630-5637

Shenoy SR, Jayaram B (2010) Proteins: Sequence to Structure and Function - Current Status. *Current Protein & Peptide Science* **11:** 498-514

Shih AY, Denisov IG, Phillips JC, Sligar SG, Schulten K (2005) Molecular dynamics simulations of discoidal bilayers assembled from truncated human lipoproteins. *Biophysical Journal* **88:** 548-556

Shih AY, Freddolino PL, Sligar SG, Schulten K (2007) Disassembly of nanodiscs with cholate. *Nano Letters* **7:** 1692-1696

Singer SJ, Nicolson GL (1972) FLUID MOSAIC MODEL OF STRUCTURE OF CELL-MEMBRANES. *Science* **175:** 720-731

Singla SI, Hudmon A, Goldberg JM, Smith JL, Schulman H (2001) Molecular characterization of calmodulin trapping by calcium/calmodulin-dependent protein kinase II. *J Biol Chem* **276:** 29353-29360

Skar-Gislinge N, Arleth L (2011) Small-angle scattering from phospholipid nanodiscs: derivation and refinement of a molecular constrained analytical model form factor. *Physical Chemistry Chemical Physics* **13:** 3161-3170

Skar-Gislinge N, Høiberg-Nielsen R, Arleth L (2015) What determines the shape and stoichiometry of self-assembled phospholipid nanodiscs: Speed of self-assembly process, initial lipid:MSP stoichiometry or detergent type? *In preparation*

Skar-Gislinge N, Simonsen JB, Mortensen K, Feidenhans'l R, Sligar SG, Moller BL, Bjornholm T, Arleth L (2010) Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. *Journal of the American Chemical Society* **132:** 13713-13722

Skou JC. (1997) The Nobel Price in Chemistry 1997. *"for the first discovery of an ion-transporting enzyme, Na + , K + -ATPase".* The Royal Swedish Academy of Sciences, Sweden.

Sligar SG. (2015) Nanodisc Technology. http://sligarlab.life.uiuc.edu/nanodisc.html, Accessed March 20 2015.

Soleil. (2015) SWING beamline. Soleil Synchrotron, http://www.synchrotron-soleil.fr/portal/page/portal/Recherche/LignesLumiere/SWING, Accessed March 10 2015.

Soper AK, Benmore CJ (2008) Quantum differences between heavy and light water. *Physical Review Letters* **101**

Stevens TJ, Arkin IT (2000) Do more complex organisms have a greater proportion of membrane proteins in their genomes? *Proteins-Structure Function and Genetics* **39:** 417-420

Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **41:** 207-234

Studier FW, Moffatt BA (1986) USE OF BACTERIOPHAGE-T7 RNA-POLYMERASE TO DIRECT SELECTIVE HIGH-LEVEL EXPRESSION OF CLONED GENES. *Journal of Molecular Biology* **189:** 113-130

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW (1990) USE OF T7 RNA-POLYMERASE TO DIRECT EXPRESSION OF CLONED GENES. *Methods in enzymology* **185:** 60-89

Su X-D, Zhang H, Terwilliger TC, Liljas A, Xiao J, Dong Y (2015) Protein crystallography from the perspective of technology developments. *Crystallography Reviews* **21:** 122-153

Susac L, Horst R, Wuethrich K (2014) Solution- NMR Characterization of Outer- Membrane Protein A from E. coli in Lipid Bilayer Nanodiscs and Detergent Micelles. *Chembiochem* **15:** 995-1000

Svergun DI, Koch MHJ (2003) Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics* **66:** 1735-1782

Sweetman G, Trinei M, Modha J, Kusel J, Freestone P, Fishov I, JoseleauPetit D, Redman C, Farmer P, Norris V (1996) Electrospray ionization mass spectrometric analysis of phospholipids of Escherichia coli. *Mol Microbiol* **20:** 233-234

Takahashi T, Yan B, Mazur P, Derguini F, Nakanishi K, Spudich JL (1990) COLOR REGULATION IN THE ARCHAEBACTERIAL PHOTOTAXIS RECEPTOR PHOBORHODOPSIN (SENSORY RHODOPSIN-II). *Biochemistry* **29:** 8467-8474

Tartof, Hobbs. (1987) Improved media for growing plasmid and cosmid clones. *Bethesda Res. Lab. Focus*, Vol. 9:12, p. 16.

Tickle I, Sharff A, Vinkovic M, Yon J, Jhoti H (2004) High-throughput protein crystallography and drug discovery. *Chemical Society Reviews* **33:** 558-565

Tomita K, Blout ER, Loze CD, Rich A (1962) EFFECT OF DEUTERATION ON GEOMETRY OF ALPHA-HELIX. *Journal of Molecular Biology* **4:** 83-92

Totsika M, Heras B, Wurpel DJ, Schembri MA (2009) Characterization of Two Homologous Disulfide Bond Systems Involved in Virulence Factor Biogenesis in Uropathogenic Escherichia coli CFT073. *Journal of Bacteriology* **191:** 3901-3908

Tzakos AG, Grace CRR, Lukavsky PJ, Riek R (2006) NMR techniques for very large proteins and RNAs in solution. In *Annual Review of Biophysics and Biomolecular Structure* Vol. 35, pp 319-342.

Uhlen M, Fagerberg L, Hallstroem BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjoestedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA-K, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist P-H, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Ponten F (2015) Tissue-based map of the human proteome. *Science* **347:** 394-404

van den Bedem H, Fraser JS (2015) Integrative, dynamic structural biology at atomic resolution-it's about time. *Nature methods* **12:** 307-318

Vitagliano V, Lyons PA (1956) DIFFUSION COEFFICIENTS FOR AQUEOUS SOLUTIONS OF SODIUM CHLORIDE AND BARIUM CHLORIDE. *Journal of the American Chemical Society* **78:** 1549-1552

Voet D, Voet JG (2004) Lipids and Membranes. In *Biochemistry*, 3rd edn, 12. United States of America: Wiley

Walden PM, Halili MA, Archbold JK, Lindahl F, Fairlie DP, Inaba K, Martin JL (2013) The alpha-Proteobacteria Wolbachia pipientis Protein Disulfide Machinery Has a Regulatory Mechanism Absent in gamma-Proteobacteria. *Plos One* **8**

Ward AB, Sali A, Wilson IA (2013) Integrative Structural Biology. *Science* **339:** 913-915

Wennerstrom H, Lindman B (1979) MICELLES - PHYSICAL-CHEMISTRY OF SURFACTANT ASSOCIATION. *Physics Reports-Review Section of Physics Letters* **52:** 1-86

Whiles JA, Deems R, Vold RR, Dennis EA (2002) Bicelles in structure-function studies of membrane-associated proteins. *Bioorganic Chemistry* **30:** 431-442

White S, Laboratory. (2015) Membrane Proteins of Known 3D Structure. Structural Biology Knowledgebase http://blanco.biomol.uci.edu/mpstruc/, Vol. March 1 2015.

Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89:** 392-400

Whitten AE, Jacques DA, Hammouda B, Hanley T, King GF, Guss JM, Trewhella J, Langley DB (2007) The structure of the KinA-Sda complex suggests an allosteric mechanism of histidine kinase inhibition. *Journal of Molecular Biology* **368:** 407-420

WHO. (2014) Antimicrobial resistance: global report on surveillance. World Health Organization, France.

Wilcox KC, Marunde MR, Das A, Velasco PT, Kuhns BD, Marty MT, Jiang H, Luan C-H, Sligar SG, Klein WL (2015) Nanoscale Synaptic Membrane Mimetic Allows Unbiased High Throughput Screen That Targets Binding Sites for Alzheimer's-Associated A beta Oligomers. *Plos One* **10**

Williams T, Kelley C. (2015) Gnuplot: an interactive plotting program. gnuplot.sourceforge.net.

Wulff M, Schotte F, Naylor G, Bourgeois D, Moffat K, Mourou G (1997) Time-resolved structures of macromolecules at the ESRF: Single-pulse Laue diffraction, stroboscopic data collection and femtosecond flash photolysis. *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment* **398:** 69-84

Yamada Y, Matsuo T, Iwamoto H, Yagi N (2012) A Compact Intermediate State of Calmodulin in the Process of Target Binding. *Biochemistry* **51:** 3963-3970

Yang AS, Honig B (1993) ON THE PH-DEPENDENCE OF PROTEIN STABILITY. *Journal of Molecular Biology* **231:** 459-474

Yee AA, Semesi A, Garcia M, Arrowsmith CH (2014) Screening Proteins for NMR Suitability. In *Structural Genomics and Drug Discovery: Methods and Protocols*, Anderson WF (ed), Vol. 1140, pp 169-178.

Zachowski A (1993) PHOSPHOLIPIDS IN ANIMAL EUKARYOTIC MEMBRANES - TRANSVERSE ASYMMETRY AND MOVEMENT. *Biochemical Journal* **294:** 1-14

Zhang X, Chien EY, Chalmers MJ, Pascal BD, Gatchalian J, Stevens RC, Griffin PR (2010) Dynamics of the beta2-adrenergic G-protein coupled receptor revealed by hydrogen-deuterium exchange. *Anal Chem* **82:** 1100-1108

Appendix 1. Nanodisc assembly

The following protocol is based on personal experience and Sligar's *PROTOCOLS FOR PREPARATION OF NANODISCS* (a link to this protocol is found under recommended reading).

Different versions of MSPs are available assembling into different disc sizes. MSP1D1 an MSP1E3D1 are the most common giving disc sizes on approximately 10 and 13 nm in diameter respectively (Denisov et al, 2004).

MSP can expressed and purified with high yields (>100 mg/L culture) as described in (Ritchie et al, 2009).

The polyhistidine tag on MSP can be removed by enzymatic cleavage, if desired, using TEV protease following purification by affinity chromatography. MSP without his tag is usually referred to as MSP(-).

Assembly of nanodiscs

Phospholipids are dissolved in sodium cholate buffer (20 mM Tris/HCl, 0.1 M NaCl, 100 mM sodium cholate, pH 7.5) to a concentration of 0.05 M. Recommended is 5 mg powdered POPC (Avanti Polar Lipids) for a standard preparation.

MSP (recommended concentration 5-10 mg/ml) in Tris buffer (20 mM Tris/HCl, 0.1 M NaCl, pH 7.5) is mixed with phospholipids in cholate buffer in a suitable molar ratio (see Table A) and shaken for a few minutes at recommended temperature (Table A).

Table A. MSP to lipid molar mixing ratio of selected phospholipids and recommended incubation temperatures.

Lipids to MSP	POPC	DMPC	DPPC
MSP1D1	65	80	90
MSP1D1(-)	65	80	90
MSP1E3D1	130	150	170
MSP1E3D1(-)	130	150	170
Temperature	4°C	25°C	37°C

To initiate the self-assembly process the mixture is added ~0.5 ml biobeads (BioRad) or Amberlite XAD-2 (SULPECO) to 1 ml sample and left for continuous shaking for 3 hours at recommended temperature. The biobeads are removed by centrifugation and filtration through a 0.22 μm syringe filter.

The nanodiscs can then be purified by gelfiltration using a Superdex 200 30/100 (GE) column with a 0.5 ml/min flow rate using Tris buffer as eluent and collected at approximately 13-14 ml. A chromatogram of such purification is shown in Figure A. If desired, nanodiscs can be

concentrated using a 10 kDa cut off spin filtration column (Amicon) to a concentration of approximately 10 mg/ml.
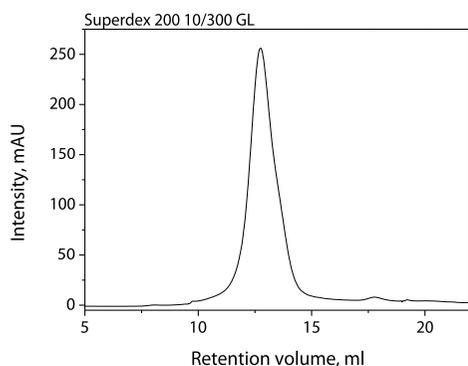
Figure A. Chromatogram of MSP1D1(-) nanodisc purification using a Superdex 200 10/300 GL column.

Step-by-step example

1. Solubilise 5 mg powdered POPC in 132 μl cholate buffer (20 mM Tris/HCl, 0.1 M NaCl, 100 mM sodium cholate, pH 7.5).
2. Mix 130 μl POPC solution with 410 μl MSP1D1 (6 mg/ml) in Tris buffer (20 mM Tris/HCl, 0.1 M NaCl, pH 7.5) and shake for a few minutes at 4°C.
3. Add ~250 μl Amberlite XAD-2 and leave for 3 hours incubation at 4°C under continuous stirring.
4. Remove Amberlites by centrifugation at 1 min 5000 rpm followed by filtration using a 0.22 μm syringe filter.
5. Apply the sample on a Superdex 200 10/300 GL column equilibrated in Tris buffer and collect the nanodisc fraction.

Recommended reading:

PROTOCOLS FOR PREPARATION OF NANODISCS
http://sligarlab.life.uiuc.edu/objects/pdfs/Nanodisc_Protocols.pdf

Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. Skar-Gislinge et al., Journal of the American Chemical Society, 2010

What determines the shape and stoichiometry of self-assembled phospholipid nanodiscs: Speed of self-assembly process, initial lipid:MSP stoichiometry or detergent type? Skar-Gislinge et al., in preparation, 2015, arleth@nbi.ku.dk

Appendix 2. Nanodisc assembly with membrane proteins

The following protocol is based on personal experience. It is recommended to practice the preparation of nanodiscs without membrane proteins, as described in Appendix 1, before starting reconstitution of membrane proteins in nanodiscs.

It is highly recommended to include an affinity tag on the membrane protein for reconstitution to be able to separate discs without membrane protein from the membrane protein loaded discs.

Different versions of MSPs are available assembling into different disc sizes. MSP1D1 an MSP1E3D1 are the most common giving disc sizes of approximately 10 and 13 nm in diameter respectively (Denisov et al, 2004).

MSP can be expressed and purified with high yields (>100 mg/L culture) as described in (Ritchie et al, 2009).

The polyhistidine tag on MSP should be removed by enzymatic cleavage, if the applied membrane protein has a His tag. This can be facilitated by enzymatic cleavage using TEV protease following purification by affinity chromatography.

Assembly of nanodiscs

Phospholipids are dissolved in sodium cholate buffer (20 mM Tris/HCl, 0.1 M NaCl, 100 mM sodium cholate, pH 7.5) to a concentration of 0.05 M. Recommended is powdered POPC (Avanti Polar Lipids). If a lipid mixture is desired these must be separately dissolved in chloroform, mixed in the needed ratio and dried to a lipid film, before dissolving in cholate buffer.

Phospholipid in sodium cholate buffer is then mixed with membrane protein of interest in a molar ratio of e.g. 650:1 or 260:1 lipids:protein (see Table A in Appendix 1 for guidance on lipid to MSP ratio) and incubates a few minutes at the recommended temperature (see Table A, Appendix 1). MSP (recommended concentration 7-10 mg/ml) in Tris buffer (20 mM Tris/HCl, 0.1 M NaCl, pH 7.5) is then added in a molar ratio 65:1 lipids:MSP and left for continues shaking for a few minutes minutes at recommended temperature.

To initiate the self-assembly process the mixture is added ~0.5 ml biobeads (BioRad) or Amberlite XAD-2 (SULPECO) to 1 ml sample and left for continuous shaking for 3 hours at recommended temperature. The biobeads are removed by centrifugation and filtration through a 0.22 μm syringe filter.

The sample is then purified by affinity chromatography suitable for the tag on the membrane protein e.g. Ni-NTA, for removing nanodiscs without membrane protein. Following this, nanodiscs containing membrane protein can be purified by gelfiltration using a Superdex 200

30/100 (GE) column with a 0.5 ml/min flow rate using Tris buffer as. A chromatogram of both purification by affinity chromatography and following gelfiltration is shown in Figure B. The nanodiscs can be concentrated using a 10 kDa cut off spin filtration column (Amicon) to a concentration of approximately 10 mg/ml.
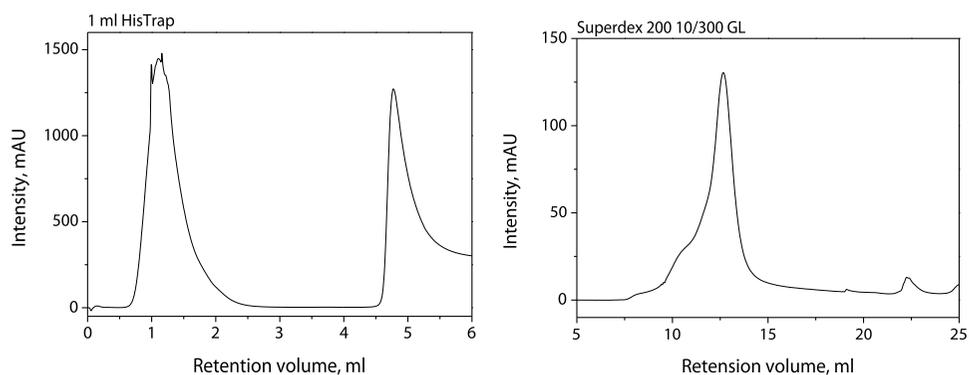
Figure B. Purification of EcDsbB in nanodiscs. Left, affinity chromatography using a HisTrap column. 400 mM imidazole is applied from approximately 4.5 ml for elusion. Right, gel filtration using a Superdex 200 column.

Step-by-step example

1. Solubilise 5 mg powdered POPC in 132 µl Cholate buffer (20 mM Tris/HCl, 0.1 M NaCl, 100 mM sodium cholate, pH 7.5).
2. Mix 130 µl POPC solution with 100 µl EcDsbB (5 mg/ml) and incubate for ~10 min at 4°C.
3. Add 315 µl MSP1D1(-) (7 mg/ml) in Tris buffer (20 mM Tris/HCl, 0.1 M NaCl, pH 7.5) and shake for ~10 minutes at 4°C.
4. Add ~250 µl Amberlite XAD-2 and leave for 3 hours incubation at 4°C under continuous stirring.
5. Remove Amberlites by centrifugation at 1 min 5000 rpm followed by filtration using a 0.22 µm syringe filter.
6. Apply the sample on a 1 ml HisTrap column equilibrated in wash buffer (20 mM Tris/HCl, 0.1 M NaCl, 40 mM imidazole pH 7.5) and wash until a steady baseline.
7. Elute EcDsbB nanodiscs with elusion buffer (20 mM Tris/HCl, 0.1 M NaCl, 400 mM imidazole pH 7.5).
8. Dilute eluate with 2-3 volumes of Tris buffer and concentrate to <500 ul using a 10 kDa cut off spin filter.
9. Apply the sample on a Superdex 200 10/300 GL column equilibrated in Tris buffer and collect the nanodisc fraction.

Recommended reading:

Membrane protein assembly into Nanodiscs, Bayburt & Sligar, Febs letters, 2010

Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment, Kynde et al., Acta Crystallographica Section D-Biological Crystallography, 2014

Appendix 3. Optimising overexpression of membrane proteins in *E. coli*

Preface

During my PhD project I spent 3 months in Morten Nørholm's group at the Novo Nordisk Foundation Center for Biosustainability (CBS). The aim was to improve skills in molecular biology techniques, such as cloning, PCR and protein expression, skills I could not have developed in the structural biophysics group. This has been very valuable for me and has without any doubt helped me in many situations during my PhD project. The project was also intended to identify and optimise expression of a good membrane protein candidate for other projects; however, this was not possible within the short time frame of the project. The following report summarizes the essence of my project at CBS.

Introduction

Membrane proteins constitute about 30% of a cell's proteome regardless of the organism (Stevens & Arkin, 2000). Unfortunately, only a small fraction of these has been studied in detail due to the fact that membrane proteins are far more challenging to handle than soluble proteins. They account for only 2.5% of the total number of protein structures found in the Protein Data Bank (PDB) (PDB, 2015b). The absence of three dimensional structures reduces knowledge about functionality as this property is highly related to structure (Katritch et al, 2013). Obtaining the quantities of protein material needed for the experiments are a common limiting factor. Overexpression of the needed membrane protein is necessary but is less straight forward than for soluble proteins as these are more likely to have a toxic effect on the host cells (Miroux & Walker, 1996). Different factors can be varied to optimise the expression of membrane proteins in *E. coli*, such as choice of expression strain, induction temperature and growth media (Cunningham & Deber, 2007).

Expression strains

Strains optimised for overexpression of membrane proteins have been developed, including the so-called "Walker strains". The strains C41(DE3) and C43(DE3), both derived from the K12 derived BL21(DE3) have been optimised for overexpression of membrane proteins by random mutation selection (Miroux & Walker, 1996). The T7 gene encoded in BL21(DE3) is regulated by a very sensitive lac promoter and in the presence of lactose T7 RNA polymerase is transcribed (see Chapter 3 – Methods in protein isolation for more details).

One approach to control the T7 RNA polymerase is to introduce a plasmid coding for T7 lysozyme, referred to as pLysS. T7 lysozyme inhibits the T7 RNA polymerase, used as expression polymerase by the very commonly applied pET vectors (Studier & Moffatt, 1986). The small amount pLysS provides, reduces the basal activity of the polymerase, ensuring expression of the target protein will not occur prior to induction (Studier et al, 1990). A different method to control the T7 RNA polymerase is to use a more controllable promoter than the lac promoter found in BL21(DE3) strains. In the KRX strain, another K12 derivative, the T7 RNA polymerase is tightly controlled by rhamnose. The T7 RNA polymerase abundance can be

precisely controlled by rahmnose concentration providing precise control of target protein expression (Promega, 2014).

Growth media

The most widely used media for bacterial growth is the minimal nutrition media lysogenic broth (LB) developed in the early 1950s (Bertani, 1951). Also commonly used is terrific broth (TB), a more nutritionally rich media developed in the 1980s to increase plasmid yield (Tartof & Hobbs, 1987). These bacterial media has been used for decades containing less specified nutrients, such as yeast extract. A more modern and nutritionally controlled media is the so-called auto induction (AI) media. This media has been formulated based on thorough research and has the great advantage that cells can be grown to saturation without the inconvenience of induction with IPTG. It has been found useful for the expression of many different proteins, including membrane proteins, for expressing high levels of target protein. The cells can be grown to a much higher density than in conventional media such as LB (Studier, 2005).

Induction temperature

The optimal growth temperature for *E. coli* is 37°C i.e. the temperature of the human body. However, heat may lead to protein aggregation and formation of inclusion bodies, reducing the protein yield and leads to cellular stress. Expression of target protein at lower temperatures can prevent or minimize these events and a higher yield can be obtained (Cunningham & Deber, 2007).

Monitoring protein expression using GFP

An easy method for rapid optimisation of membrane protein expression is fusing GFP to the N-terminal of the protein. This allow for direct monitoring of the protein expression level on the cell culture and avoids the hassle of purification when screening expression of a large number of proteins (Drew et al, 2001). Correct folding of the fused membrane protein will result in folding of GFP. This also means that if a fusion protein gets included in inclusion bodies it will not fluorescent. The concentration of expressed target protein can thereby be directly monitored by fluorescent readout. Expression vectors containing the GFP gene are available for fusion protein expression, such as pWaldo (Drew et al, 2001).

DNA cloning by uracil excision

To express a protein using an expression vector, DNA coding for the protein of interest must first be inserted into the plasmid. This is usually done by PCR-based DNA cloning techniques. A simple method for this is the so-called USER™ (uracil-specific excision reagent) cloning technique. The technique relies on stabile hybridization of complementary 3' overhangs of 8 nucleotides created in the ends of PCR amplified DNA and expression vector. Overhangs are created on the PCR fragments by placing a single uracil in the primers used for amplification and a subsequent treatment of the PCR product with USER™ enzyme mix. This contains DNA glycosylase and DNA glycosylase-lyase Endo VII which remove the two singe uracils which in turn dissociates the single stranded fragments upstream of the cleavage site. The overhangs are designed to precisely overlap a cloning site in the vector and after digestion of the cloning site, hybridization can occur (Nour-Eldin et al, 2006).

Methods summary

Protein expression optimisation

DNA coding for three different *E. coli* membrane proteins NhaA, ClC and GalP were cloned into pWaldo by uracil excision. The hybridized plasmids were transformed into NEB5α and grown over night on LB agar plates at 37°C. Correct insertion of DNA in the plasmid was confirmed by colony PCR. Overnight cultures were prepared in LB media and plasmids were harvested the following day and send for sequencing.

Plasmids containing DNA coding for the different membrane proteins were transformed into three different expression strains, BL21LysS, C43 and KRX, grown on LB agar plates and overnight cultures prepared from colonies in LB media. Expression media was inoculated with the overnight cultures.

The following parameters were varied to find optimal expression conditions:
-   Induction temperature 37°C or 20°C
-   Induction OD_{600} 0.3 or 0.6
-   Induction time 4 hours or over-night (ON)
-   Expression media TB or AI

All cultures were grown in 24-well plates in 3 ml volumes stirring at 250 rpm and at 37°C until induction and induced with 0.4 mM IPTG. All conditions were tested in triplicates.

OD and GFP fluorescence were read directly on cultures using a plate reader, using the respective media as background.

In gel fluorecense and western blot

C43 cells expressing GFP and NhaA, ClC and GalP as fusion proteins in TB media were harvested and lysed using lysozyme and benzonase. Lysed cells were run on SDS PAGE and expression detected by in-gel fluorescence. A western blot was subsequently made from the gel using a primary antibody against GFP.

[35S]-Met labelling

GFP was cloned out of the plasmids and after hybridization the plasmids were transformed into NEB5α and grown over night on LB agar plates at 37°C. Correct insertion of DNA in the plasmid was confirmed by colony PCR. Overnight cultures were prepared in LB media and plasmids harvested the following day and send for sequencing.

The plasmids were transformed into C43, grown on LB agar plates and overnight cultures were prepared from colonies in LB media. LB media was inoculated with overnight culture and grown for ~2 hours. Cells were gently spun down and the media exchange to M9 (minimum salt, methionine free) media and grown to OD_{600} 0.2 and induced with IPTG for 10 min before adding rifamphicilin and incubated for 10 min, to inactivate basic protein synthesis. The culture was added [35S]-methionine and incubated further in 3 minutes before cooled on ice and cells harvested. Cells were lysed and run on SDS PAGE. The protein could thereby be traced by the radioactive marker.

The gel was dried and incubated with a phosphoimager plate and a picture developed.

Result summery

Varying expression temperature

Decreasing the expression temperature from 37˚C of certain proteins has earlier shown useful to boost protein expression. For this reason, NhaA and ClC were expressed in two different strains BL21(DE3)pLysS and C43 at two different temperatures, 20˚C and 37˚C. This was either done by 4 hours or over-night expression in TB or AI media. The protein expression was followed by detecting the GFP fluorescence intensity.
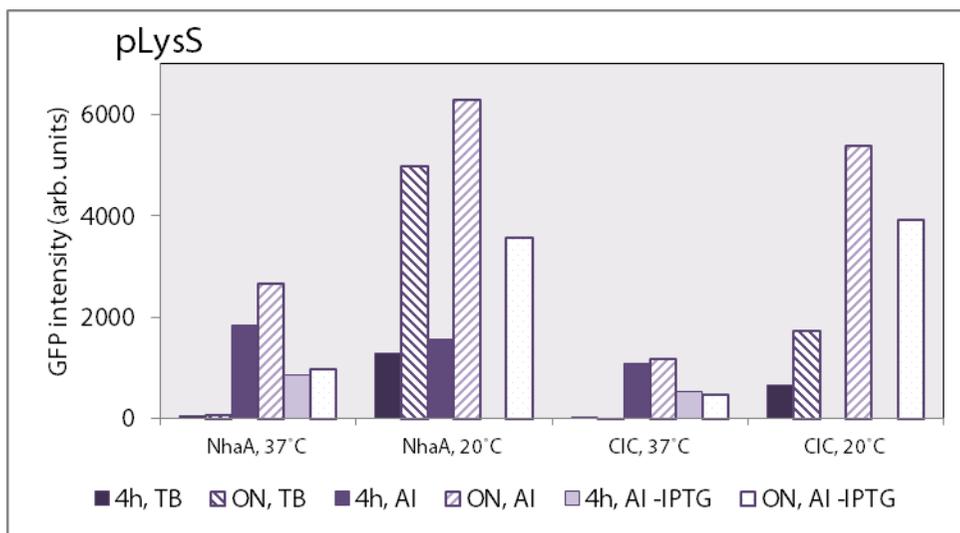
Figure I. Expression of NhaA and ClC in BL21(DE3)pLysS at 37 and 20˚C. Colour codes show media and type of induction. ON means over-night expression, 4h 4 hours expression. TB and AI is type of media and –IPTG is expression by auto induction without IPTG. Protein and expression temperature is listed below each series.
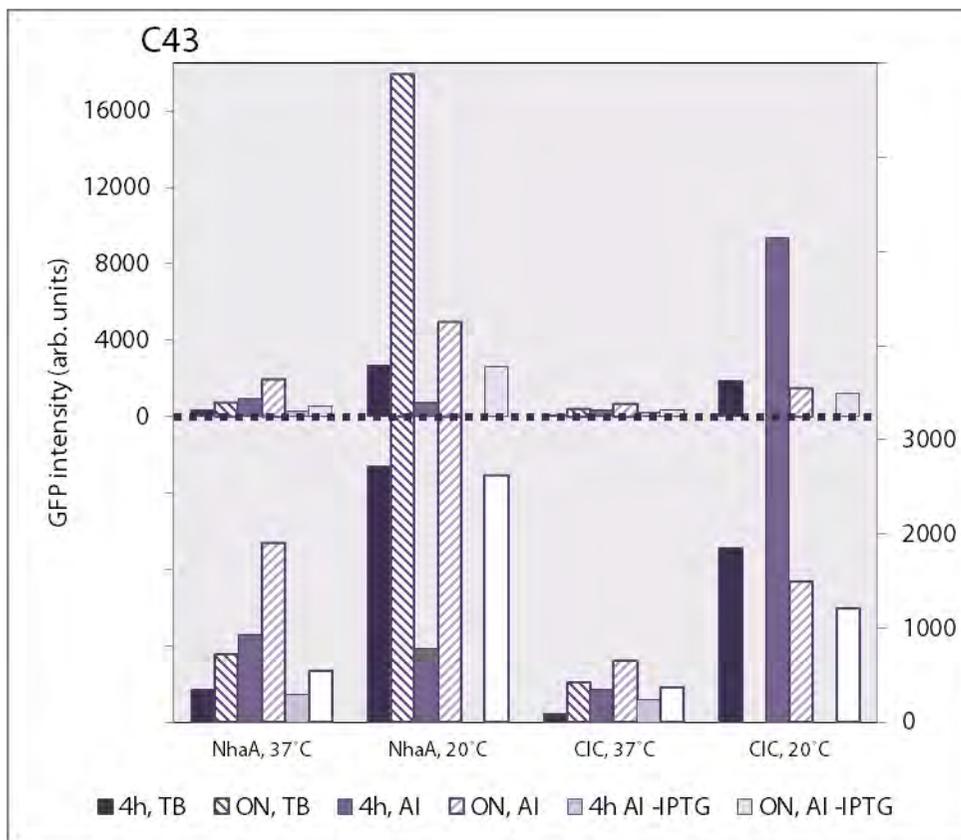
Figure II. Expression of NhaA and ClC in C43 at 37 and 20˚C. Colour codes show media and type of induction. ON means over-night expression, 4h 4 hours expression. TB and AI is type of media and –IPTG is expression by auto induction without IPTG. Protein and expression temperature is listed below each series. Data is viewed on two scales, top diagram show the full scale while the bottom diagram only shows low GFP intensities.

Expression in pLysS generally gives a higher expression at 20˚C using over night induction as from Figure I. TB media appear not to be very applicable at 37˚C whereas for 20˚C over-night induction gives higher yield than 4 hours induction. AI media and over-night induction at 20˚C gives the highest expression for both proteins. Over-night auto induction gives a much better result at 20˚C than 37˚C. Auto induction for 4 hours gives no expression at 20˚C while little expression at 37˚C.

When using C43 as expression strain the difference between 20˚C and 37˚C expression is more clear as seen in Figure II. Again 20˚C generally gives a higher expression for both proteins. TB over-night expression gives the higest expression for NhaA while 4 hours expression in AI is better for ClC. Auto induction for 4 hours does, as with pLysS, not show any expression and minor expression at 37˚C. Over-night expression appear to be better for NhaA than 4 hours and the other way around for ClC.

Including more variables

As expression temperature at 20˚C appeared to give highest yield for NhaA and GalP, this was used for the remaining experiments. Also, auto induced expression for 4 hours at 20˚C did not not show any expression and was not further investigated.

NhaA, ClC and GalP were expressed, now with more varying parameters. Induction OD_{600} was varied between 0.3 and 0.6 for the non auto induced cultures. Expression was carried out in three different strains which earlier have shown good expression of membrane proteins BL21(DE3)pLysS and C43, used in the earler atudy, and KRX, which was added different concentrations of rhamnose at induction in addition to IPTG. TB and AI were againg compared as expression media as well as induction time 4 hours vs over night expression.
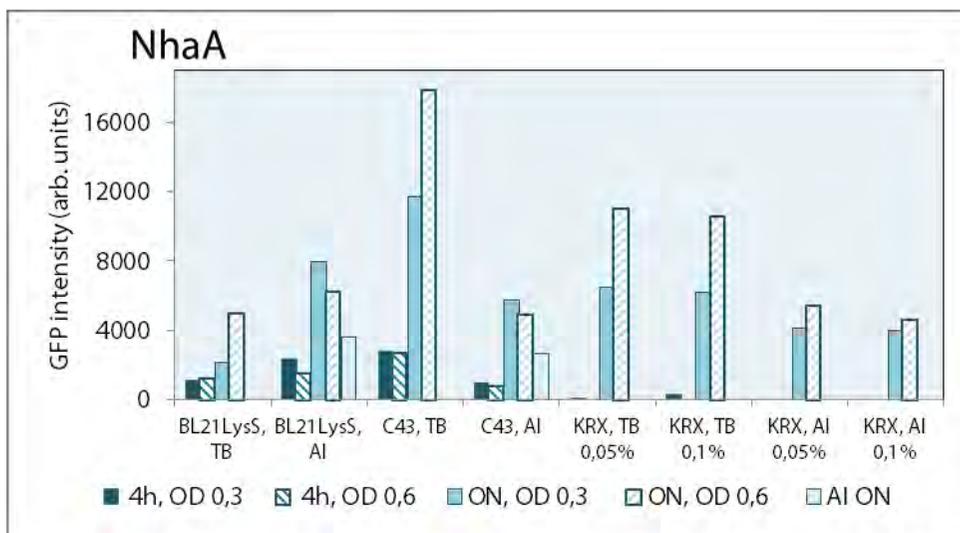
Figure III. NhaA expression in three strains varying induction OD, time and media. 4h is 4 hours induction, ON is over-night induction and AI ON is over-night auto induction. Expression strain and media is listed below each series. For the KRX strain was as addition to IPTG added 0.05 or 0.1% rhamnose.
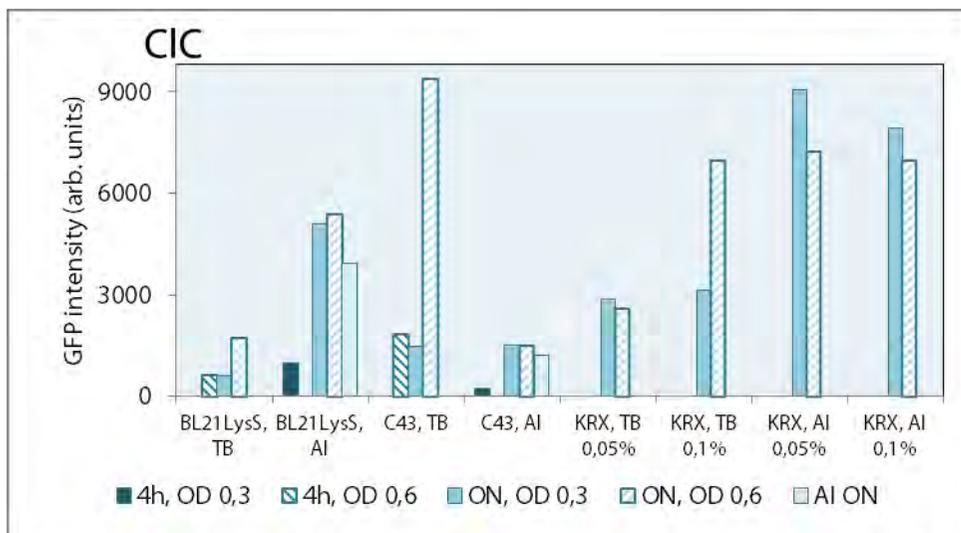
Figure IV. ClC expression in three strains varying induction OD, time and media. 4h is 4 hours induction, ON is over-night induction and AI ON is over-night auto induction. Expression strain and media is listed below each series. For the KRX strain was as addition to IPTG added 0.05 or 0.1% rhamnose.
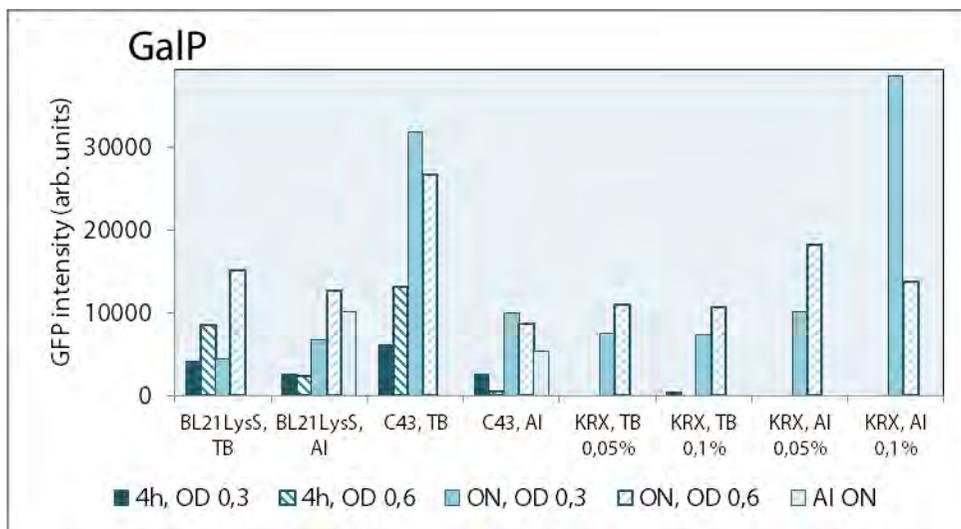
Figure V. GalP expression in three strains varying induction OD, time and media. 4h is 4 hours induction, ON is over-night induction and AI ON is over-night auto induction. Expression strain and media is listed below each series. For the KRX strain was as addition to IPTG added 0.05 or 0.1% rhamnose.

As seen from the optained GFP intensiteis from expression of the three proteins in Figure III, Figure IV and Figure V there is quite a difference in expression patterns. A mutal tendency for high expression in C43 using TB media and over-night expression is found. Interestingly, AI media show a general low expression for C43. An induction OD of 0.3 is preferred for NhaA

and ClC, while 0.6 is preferred for GalP. Again it is shown that over-night expression at 20°C gives a significantly higher yield than 4 hours expression.

For NhaA C43 in TB gives the heighest yield as seen in Figure III while KRX in TB also gives a relatively high expression.

ClC has highest expression C43 using TB media, though expression yields in KRX in AI media are close to the C43 level.

GalP show highest yield using KRX in AI media using 0.1% rhamnose at induction, though high yiels can also be obtain using C43 in TB media.

BL21(DE3)pLysS show the lowest expression of all strains for all proteins.

In gel fluorecense and western blot

To ensure detected fluoresence in the previous experiments were otained from fusion proteins and not disassociated GFP, cultures were run on SDS PAGE and detected for in gel fluoresence. Furthermore, a western blot was made from this gel using a primary antibody against GFP. The sizes of the applied proteins are: GFP 26.9 kDa, while NhaA is 41.2 kDa, ClC 50.2 kDa and GalP 50.8 k Da
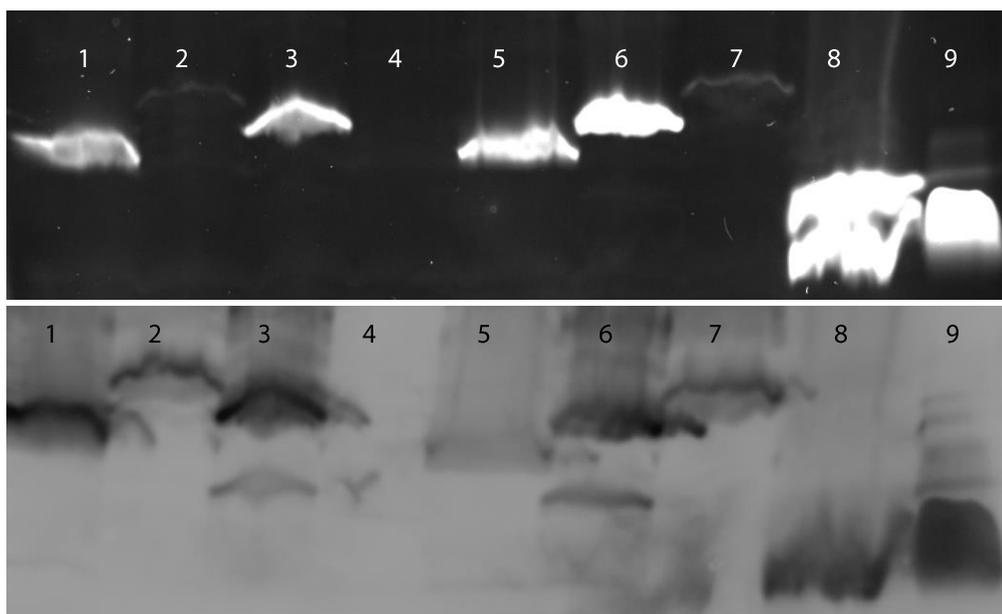
Figure VI. In gel fluorescence (top) and western blot (bottom) of cultures. NhaA in well 1 and 5. ClC in well 2 and 7. GalP in well 3 and 6. GFP in well 8 and GFP control in well 9.

From the in gel fluorescence and western blot in Figure VI it can be seen that the expression products are in sizes corresponding to the respective fusion proteins. All proteins have GFP fused, as no bands are detected in a size corresponding to GFP only. This confirm the previous expression studies.

[35S]-Met labelling

To detect expression of NhaA, ClC and GalP, without fused GFP, they were expressed in C43 and labelled with [35S]-methionine. The cells were run on SDS PAGE and a phosphoimage developed.
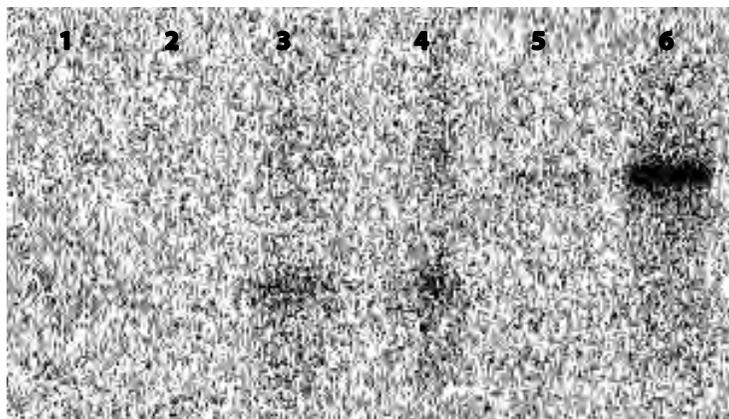
Figure VII. [35S]-Met labelling. ClC in well 1 and 2, NhaA in well 3 and 4 and GalP in well 5 and 6.

As seen on Figure VII, ClC could not be detected by the radiolabelling, which may be due to a low expression level or plasmid error. However, both NhaA and GalP were detected suggesting good expression. This experiment should be complemented with western blot to confirm expression.

GalP

Because of the high expression level of GalP, large scale expression at optimal conditions and following purification was attempted. Nevertheless, it was not possible to purify any GalP despite several attempts. Sequencing of the plasmid revealed an error in the protein sequence which may have caused this and a correction of this should be attempt.

Conclusion

Not surprisingly, lowering the temperature post induction and incubation over-night results without question the best yield for the membrane proteins.

For all the three membrane proteins, C43 grown in TB had high GFP counts. For NhaA and ClC, an induction OD of 0.3 was preferred while an OD of 0.6 was preferred for GalP.

It was not possible to optimise GalP expression in the desired construction within the available time frame.