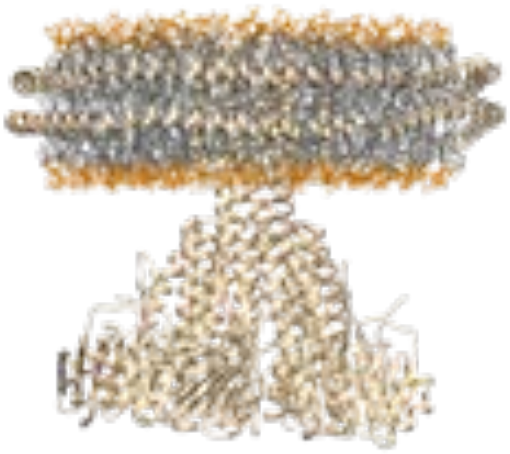
UNIVERSITY OF COPENHAGEN
NIELS BOHR INSTITUTE

PhD Thesis
Nicolai Tidemand Johansen

Towards an Optimal Basis for Structural Analysis of Membrane Proteins in Solution by Small-Angle Scattering

Optimized Membrane Protein Carrier Systems and Online Size-Exclusion Chromatography Coupled to Small-Angle Neutron Scattering



Principal supervisor: Professor Lise Arleth
Co-supervisor: Professor Thomas Günther Pomorski

November 30, 2018

PREFACE

This thesis was carried out at the Niels Bohr Institute and the Department of Plant and Environmental Sciences at the Faculty of Science at the University of Copenhagen under the supervision of Professor Lise Arleth and co-supervision of Professor Thomas Günther Pomorski in the period between December 2015 and November 2018. The work has involved numerous beam times at international facilities around Europe and Australia together with a stay abroad at the University of Ruhr, Germany, in the lab of Thomas Günther Pomorski.

The thesis has been submitted for the degree of Doctor of Philosophy at the University of Copenhagen. It was financed partly by the Niels Bohr Institute and partly by bioSYNergy KU2016.

The main focus of this thesis has been to study membrane protein structure under solution conditions using small-angle scattering methods. A large part of the work has been focused on the optimization of soluble membrane protein carrier systems, in particular nanodiscs via systematic studies of their self-assembly process, genetic engineering of belt proteins and contrast-optimization. Furthermore, a novel small-angle neutron scattering setup was explored to obtain aggregation-free data.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Lise Arleth for giving me the opportunity to join her group and achieve my Ph.D. degree. I thank her and my co-supervisor Thomas G. Pomorski for great supervision throughout the project.

I would also like to thank present and former colleagues at the Niels Bohr Institute. I would like to give a special thanks to Søren Roi Midtgaard for introducing me to various wet labs in Copenhagen and various X-ray and neutron beam lines around the world during the initial year of my thesis. I would also like to give special thanks to Frederik Grønbæk Tidemand, Andreas Haahr Larsen and Martin Cramer Pedersen for good collaborations on various projects.

I would also like to thank present and former colleagues at the Department of Plant and Environmental Sciences, especially Anja Fuglsang and Anette Lund, for letting me work in their lab and for providing helpful day-to-day advise.

Finally, I would like to thank my girlfriend Stine for her support and patience, especially during my many travels for experiments around the world.

ABSTRACT

Small-angle neutron scattering (SANS) is an important method in many areas of soft matter sciences due to the possibility for contrast variation by hydrogen/deuterium substitution. However, compared to small-angle X-ray scattering (SAXS), SANS remains a less explored method in biological sciences, mainly due to higher sample requirements and more restricted access to neutrons. In this thesis, the overall aim was to develop approaches to apply SANS and contrast variation to study the structure of membrane proteins (MPs) in solution. To be kept in solution, MPs must be stabilized by amphipathic carrier systems, *e.g.* detergent micelles or phospholipid nanodiscs. Detergent micelles have the advantage of versatile and easy sample handling, but they often destabilize MPs. Nanodiscs have the advantage of providing a lipid bilayer in a well-defined particle, but sample preparation is demanding and not well-understood. Within the major aim, four sub-aims were pursued; (i) understanding the self-assembly of nanodiscs, (ii) developing stability-optimized nanodiscs, (iii) testing size-exclusion chromatography (SEC)-SANS on MP samples, and (iv) developing contrast-optimized detergent micelles and nanodiscs. For objective (i), SAXS was utilized to investigate the structures of nanodiscs formed under different conditions. It was found that nanodisc self-assembly is a fast process independent of detergent removal rate, whereas the lipid stoichiometry and choice of reconstitution detergent are determining factors and important to optimize. In objective (ii), a novel type of nanodisc was developed by introduction of extra negative surface charges together with covalent circularization by sortase A. Combined SAXS, light scattering, circular dichroism spectroscopy and SEC analysis showed that the nanodisc had the expected structure but with drastically increased stability. The bacterial magnesium transport protein CorA could readily be incorporated in these nanodiscs. For objective (iii), a new SEC-SANS setup was benchmarked on a number of nanodisc samples. Despite using relatively dilute samples, data of sufficient quality for structural analysis were obtained. Finally, in objective (iv), uniformly contrast-matched micelles and nanodiscs were utilized to obtain an optimal basis for structural analysis of MPs inside. Using SEC-SANS, high quality SANS data were obtained on CorA. Surprisingly, however, the obtained data could not be explained by current structural models for CorA. On one hand this suggests that more advanced modeling efforts are required, but it also highlights the importance of structural validation by solution-methods, including SAXS and SANS. Overall, the studies of this theses have provided useful insights to the application of SANS to complex biological samples.

RESUMÈ

Småvinkel neutron-spredning (SANS) er en vigtig metode i mange områder inden for materialevidenskab på grund af muligheden for kontastvariation ved hydrogen/deuterium-udskiftning. Sammenlignet med småvinkel røntgen-spredning (SAXS) er SANS dog mindre anvendt inden for biologiske videnskabsområder, da metoden har højere krav til prøveforberedelse og det generelt er sværere at få adgang til neutronfaciliteter. I denne afhandling var det overordnede mål at udvikle nye måder at anvende SANS og kontrastvariation til at undersøge strukturen af membranproteiner (MPs) i opløsning. For at kunne holde MPs i opløsning skal de stabiliseres af et amfipatisk system, som for eksempel detergentmiceller eller nanodisce. Detergentmiceller har fordelen at de er alsidige of gør prøveforberedelse simpel, men de er dog ofte destabiliserende for MPs. Nanodisce har fordelen at de indeholder et lipid bilag i en veldefineret partikel, men dog er prøveforberedelse krævende og ikke velforstået. Inden for det overordnede mål blev der sat fire delmål; (i) at opnå en bedre forståelse af selvansamlingen af nanodisce, (ii) at udvikle optimerede nanodisce med højere stabilitet, (iii) at teste koblet størrelseskromatografi (SEC)-SANS på membranproteinprøver, (iv) at udvikle kontrastoptimerede detergentmiceller og nanodisce. Til at opnå delmål (i) blev SAXS anvendt til at undersøge strukturen af nanodisce formet ved forskellige betingelser. Det blev observeret, at selvansamlingen af nanodisce var en hurtigt process, der var uafhængig af hastigheden hvormed detergenten blev fjernet, mens lipid-støkiometrien samt hvilken type detergent der blev brugt havde stor indflydelse og var vigtige at optimere. Til at opnå delmål (ii) blev en ny type nanodisc udviklet med ekstra negative overfladeladninger samt kovalent cirkularisering ved hjælp af enzymet sortase A. Med en kombination af SAXS, lysspredning, cirkulær dikroisme spektroskopi og SEC blev det vist at denne nanodisc havde den forventede struktur, men med meget forbedret stabilitet. Det bakterielle magnesium-transportprotein CorA kunne uden besvær indsættes i denne nanodisc. Til at opnå delmål (iii) blev et ny SEC-SANS set-up testet på en række nanodiscprøver. Selvom disse prøver var i forholdsvis lav koncentration kunne der måles data af en god kvalitet, der tillod strukturanalyse. Til at opnå delmål (iv) anvendtes kontrastoptimerede detergentmiceller og nanodisce til at opnå en optimal situation for strukturanalyse af MPs siddende inde i disse systemer. Ved brug af SEC-SANS blev der målt data af høj kvalitet for CorA. Det kom dog som en overraskelse at disse data ikke kunne forklares med nuværende strukturelle modeller for CorA. Dette betyder, at mere avancerede modelleringsværktøjer påkrævet til at forklare data, men fremhæver samtidig vigtigheden af strukturvalidering på prøver i opløsning, som netop kan gøres med SAXS og SANS. Samlet set har disse studier givet værdifuld indsigt i brugen af SANS til at studere komplekse biologiske prøver.

THESIS OUTLINE

1 MOTIVATION

Membrane proteins (MPs) remain exceptionally difficult to study compared to soluble proteins. As reflected by the low number of unique structures (Stephen H. White 2018), MP crystallization is challenging and calls for other structural methods to be applied. Several methods can provide structural information on MPs (and biomolecules in general), but each method has specific advantages and draw backs. For instance, it is difficult to go above approximately 30 kDa in solution NMR (sNMR) and below approximately 100 kDa in cryoEM. The average eukaryotic protein is approximately 40 kDa and the average bacterial protein approximately 30 kDa (Brocchieri and Karlin 2005), which highlights that there is also a need for alternatives to these methods. The term ”integrative structural biology” covers a powerful approach to refine structural modelss based on data from several methods, and is under development these years (Ward et al. 2013). Structural data can be divided into different subcategories ranging from atomic structures and atomic distances to overall size and shape and solvent accessibility (Sali et al. 2015). The atomic resolution is obtained by crystallography, NMR, cryoEM and molecular simulations, whereas techniques such as Förster resonance energy transfer, double electron-electron resonance and electron paramagnetic resonance spectroscopy can provide large atomic distances between fluorescent or magnetic labels. Solvent accessibility as probed by hydrogen/deuterium (H/D) exchange mass spectrometry or NMR can provide information about dynamics and degree of burial. The overall size and shape, along with pair-distance distributions, are readily obtained by small-angle scattering (SAS). SAS has thus emerged as an important method for gaining information on overall molecular size and shape. There are two major reasons for this. First, SAS can be used to study a biomolecules in a very wide range of sizes (molecular weights ranging from few kDa to MDa autociteSvergun2003) and is thereby complementary to NMR and CryoEM techniques that are usually limited by the size of the biomolecules under investigation. Second, samples are measured in solution under presumably more native-like conditions than in a crystal or a cryo-cooled samples, and even intrinsically disordered regions that inhibit crystallization and are not resolved in cryoEM will contribute to the measured SAS intensity.

SAS suffers from the orientational averaging of molecules in solution, and SAS data is consequently of relatively low information content (D. I. Svergun and Koch 2003). However, the information content can be increased by contrast-variation studies, either by combining small-angle X-ray scattering (SAXS) and

small-angle neutron scattering (SANS), or using H/D substitution in SANS. While SAXS is now a well-established method in the biological community facilitated by powerful synchrotron sources and automated fast setups for measurements and data processing of hundreds of samples, SANS is less broadly applied. The main reasons for this are that access to neutrons is limited, sample requirements are typically high and data collection and processing are not fully automated. New sample environments (Jordan et al. 2016) and high-flux sources, in particular the upcoming European Spallation Source (ESS), are being developed to ensure better prerequisites for bioSANS-users in the future. At the same time, there is a general need to become better at handling MPs in solution to obtain high quality samples and thereby optimal, aggregation-free SAS data. Thus, the main motivation for this thesis was to continue the development of approaches to study solution structures of MPs by SAXS and SANS, using well-defined and/or contrast-optimized carrier systems, including nanodiscs based on membrane scaffold proteins (MSP) and n-dodecyl-β-D-maltoside (DDM) micelles. The specific aims and objectives are given in the following section.

2 AIMS

The original working title of this thesis was ”Neutron and X-ray based structural investigations of the interplay between membrane proteins and the surrounding lipid membrane”. The overall aim was to continue developments and solve specific problems related to small-angle neutron and X-ray scattering approaches to investigate MP structure in solution. To achieve this, four major experimental aims were set:

Aim I The first aim was to optimize nanodisc samples to obtain aggregation-free SANS data in D_2O. To achieve this, three specific objectives were proposed: (i) Conduct experiments to elucidate the mechanism and/or kinetics of the nanodisc aggregation in D_2O, (ii) introduce mutations in the MSPs to reduce or eliminate aggregation in D_2O, and (iii) carry out similar experiments as above together with stability measurements on simpler samples, *e.g.* soluble globular proteins, to obtain a general understanding of differential sample properties in D_2O.

Aim II The second aim was to investigate MP structure in contrast optimized carrier systems using SANS. To achieve this, four objectives were proposed: (i) Obtain deuterated nanodisc components, *i.e.* lipids and belt protein, and deuterated detergents via collaborations with international deuteration facilities, and reconstitute and measure model MPs in these systems by SANS, (ii) vary the lipid composition in nanodiscs to probe its effect on MP structure, (iii) explore a new size-exclusion chromatography (SEC)-SANS setup for obtaining aggregation-free data, and (iv) combine the SANS data with SAXS data to model the interplay between the MP and the lipids in the nanodisc.

Aim III The third aim was to complement structural data from SAXS and SANS with other biophysical experiments to gain further insight to *e.g.* structural dynamics and stability of the carrier systems together with the MPs inside.

Aim IV Apply the developed methodologies to investigate MPs of interest to the bioSYNergy network at the University of Copenhagen.

Besides the experimental aims listed above, this thesis aims to provide an overview of current carrier system technologies related to MP stabilization and how SAS can be a valuable tool for investigating MP structure.

3 CONTENT

The thesis is divided in two parts. The first part compromises three chapters. Chapter 1 is an introduction and review of common carrier system used for MP handling, Chapter 2 is an introduction to the basics of SAS techniques, and Chapter 3 presents the major experimental results and gives a final conclusion to the thesis work. Further details are given below.

Chapter 1: Membrane protein carrier systems

This chapter introduces the expanding array of carrier systems developed to handle MPs in solution for structural and functional studies. The aim is to give an overview of current advantages and draw backs of these systems in relation to functional and structural studies of MPs in solution. The chapter was written as part of an ongoing literature study and is planned to be included in a review article. Of the presented systems, detergent micelles and MSP nanodiscs were used extensively in this thesis; detergent micelles for providing easy sample handling and nanodiscs for obtaining structurally well-defined samples with a lipid environment.

Chapter 2: Small-angle X-ray and neutron scattering

This chapter introduces the basic theory, instrumentation, data analysis, and challenges in sample preparation in regards to SAS techniques. The aim is to describe how SAS, and particularly SANS, can contribute to structural studies of MPs in solution. Alongside this, the focus is on current challenges towards obtaining aggregation-free data and what prospect the future might hold for making SANS more widely available to the biological community.

Chapter 3: Results and conclusion

The main results of the thesis are presented and discussed. The chapter is divided in two main sections,

one focused on specific nanodisc developments and the other on developments for SANS studies on MPs in solution. It is attempted to present the results in a logical order towards to common aim of obtaining high quality, aggregation-free SANS data on MP in solution. After the two head sections, this chapter provides an outlook to future studies. Finally, the conclusion of the thesis directly addresses the aims and objectives that were set in section 2.

The second part compromises five chapters that present the experimental work in subdivided studies and in manuscript formats. These chapters include two accepted articles, one manuscript currently under revision, a manuscript in preparation and a general research report not yet ready for publication. Further details are given below, including the author's (NTJ) specific contribution to each work.

Paper I: "Comprehensive Study of the Self-Assembly of Phospholipid Nanodiscs: What Determines Their Shape and Stoichiometry?"

This article investigates the self-assembly of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC)-MSP1D1 nanodiscs under varying conditions by modulating the POPC:MSP1D1 stoichiometry, choice of detergent, and detergent removal speed. The article was published in Langmuir in October 2018.

NTJ contribution: Founding and planning of the subset of experiments related to samples listed under sample preparation. Sample preparation: Protein production and purification, nanodisc reconstitution at different detergent removal rates and from ternary mixed micelles, preparation of mixed micelles. Experiments: SEC and light scattering measurements. Data analysis: Light scattering data analysis. Writing of the manuscript.

Paper II: "Introducing SEC-SANS for studies of complex self-organized biological systems"

The feasibility of SEC-SANS was investigated on three different nanodisc samples; empty POPC-MSP1D1 nanodiscs, PR-loaded POPC-MSP1D1 nanodiscs, and CorA-loaded POPC-MSP1E3D1 nanodiscs. The article is accepted for publication in the special issue *Neutrons in biology* of Acta Crystallographica section D, which, at the time of writing this thesis, is in print.

NTJ contribution: Founding and planning of the project. Sample preparation: Protein production and purification, nanodisc reconstitution. Experiments: SAXS and SANS measurements. Data analysis: Initial data processing. Writing of the manuscript.

Paper III: "Circularized and solubility-enhanced MSPs facilitate simple and high yield production of stable nanodiscs for studies of membrane proteins in solution"

A new MSP was designed with a high abundance of negatively charged amino acids proposed to increase

solubility and made compatible with circularization by the enzyme sortase A. High yield production was readily achieved and conformationally stable nanodiscs could be obtained. The manuscript was submitted to the FEBS journal and has been through a first round of review. The present version has been re-submitted as of ultimo November 2018.
NTJ contribution: Founding and planning of the project. Sample preparation: Protein production and purification, nanodisc reconstitution. Experiments: SAXS, SEC, circular dichroism spectroscopy (CD) and light scattering measurements, sodium dodecyl sulfate poly acryl amide gel electrophoresis (SDS-PAGE) experiments. Data analysis: Analysis of SDS-PAGE, SEC, light scattering, and CD data. Writing of the manuscript.

Paper IV: "Small-angle neutron scattering show that the solution structures of the bacterial Mg^{2+}-channel CorA are overall similar with and without Mg^{2+} bound"

This manuscript describes the development of optimized stealth nanodiscs (sNDs) made with chemically synthesized deuterated POPC and bacterially produced deuterated circularized MSP. CorA was measured in these nanodiscs as well as in stealth DDM (sDDM) micelles by both standard cuvette SANS and SEC-SANS in D_2O. CorA is proposed to undergo a large conformational change upon binding Mg^{2+}, but surprisingly, the SANS revealed no difference in scattering intensities between samples measured with and without Mg^{2+} present. Furthermore, fitting of the currently available high resolution structures of CorA to the SANS data showed that while the overall size of the protein was as expected, there were significant systematic discrepancies on shorter length scales. EM data and activity assays verified the structural and functional integrity of CorA in the conditions used for SANS. Based on these observations, the solution structures of CorA are proposed to be overall similar with and without Mg^{2+}. Modeling efforts are undergoing to produce a structural model of CorA that fit the SANS data. Thus, this work is currently a manuscript in preparation.
NTJ contribution: Founding and planning of the project. Sample preparation: Protein production and purification, nanodisc reconstitution, electron microscopy (EM) samples, reconstitution in large uni-lamellar vesicles (LUVs). Experiments: SAXS, SANS, SEC, EM and activity measurements. Data analysis: Analysis and modeling of SAXS and SANS data, analysis of SEC-SANS data and activity assays. Writing of the manuscript.

Report I: "Heavy water effect on proteins and nanodiscs?"

Small soluble proteins and nanodiscs are investigated in D_2O under the hypothesis that this slightly altered solvent compared to H_2O has a negative influence on sample quality in SANS. A report is presented with data on thermal and temporal protein and nanodisc stability, respectively, in D_2O and these data are com-

pared to identical experiments in H_2O.
NTJ contribution: This report only presents work done by NTJ.

Besides the work presented in this thesis, the author has had minor contributions to other projects that have resulted in published articles.

- Larsen, Andreas N. *et al.* Dimeric peptides with three different linkers self-assemble with phospholipids to form peptide nanodiscs that stabilize membrane proteins, Soft Matter **12**, 5937-5949, (2016)

  NTJ contribution: CD measurements and analysis. Writing the part of the manuscript related to these data.

- Larsen, Andreas N. *et al.* Small-angle neutron scattering studies on the AMPA receptor GluA2 in the resting, AMPA-bound and GYKI-53655-bound states, IUCrJ **5**, 780-793, (2018)

  NTJ contribution: Sample preparation at SANS beam times and concentration and pH measurements. Writing of a short passage related to pH effects.

Furthermore, the author has aided in a number of collaborations on nanodisc preparation and SAXS and SANS data collection. One of these efforts were included in Jennifer Roche's PhD thesis.

- Roche, Jennifer V. Protein-Protein Interactions in Human Aquaporin Regulation. Lund : Lund University, Faculty of Science, Department of Chemistry, 2018. 180 p.

  NTJ contribution: Assistance with production of MSP, nanodisc reconstitution and SAXS measurements.

CONTENTS

ABBREVIATIONS

Acronym	Description
APol	Amphipol
BR	Bacteriorhodopsin
CAC	Critical aggregate concentration
CD	Circular dichroism spectroscopy
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CHAPSO	3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate
CMC	Critical micelle concentration
CMP	Contrast-match point
cMSP	Circularized MSP
cND	Circularized nanodisc
CorA	Magnesium transport protein CorA from *Thermogata maritima*
CryoEM	Cryo electron microscopy
csND	Circularized and solubility-enhanced nanodisc
CV	Column volume
D_{max}	Maximum pair-distance
DAPol	Deuterated amphipol
DDM	n-dodecyl-β-D-maltoside
DHPC	1,2-dihexanoyl-sn-glycero-3-phosphocholine
DIBMA	Diisobutylene-maleic acid
DIBMALP	Diisobutylene-maleic acid lipid particle
DLS	Dynamic light scattering
DM	n-decyl-β-D-maltoside
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DPC	Dodecyl-phosphocholine
DSC	Differential scanning calorimetry
DTT	Dithiothreitol
EM	Electron microscopy

ESRF	European Synchrotron Radiation Facility
eSrt	Evolved sortase A from *Staphylococcus aureus*
eV	Electron volt
FN3	Second Fibronectin type-III module of NCAM2
GPCR	G-protein coupled receptor
HDL	High-density lipoprotein
IDR	Intrinsically disordered region
IEC	Ion-exchange chromatography
IFT	Inverse Fourier transform
ILL	Institut Laue-Langevin
IM7	Colicin-E7 immunity protein
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IUPAC	International Union of Pure and Applied Chemistry
LDAO	Lauryldimethylamine-n-oxide
LMNG	Lauryl maltose neopentyl glycol
LUV	Large uni-lamellar vesicle
MP	Membrane protein
MS	Mass spectometry
MSP	Membrane scaffold protein
NMR	Nuclear magnetic resonance
NSP_r	Reverse nanodisc scaffold protein
OD	Optical density
OG	n-octyl-β-D-glucoside
PC	Phosphocholine
PDC	Protein-detergent complex
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
R_g	Radius of gyration
SANS	Small-angle neutron scattering
SapA	SaposinA
SAS	Small-angle scattering
SAXS	Small-angle X-ray scattering
sDDM	Stealth DDM
SDS	Sodium dodecyl sulfate

SDS-PAGE	Sodium dodecyl sulfate poly acryl amide gel electrophoresis
SEC	Size-exclusion chromatography
SEC-SANS	Size-exclusion chromatography coupled to small-angle neutron scattering
SEC-SAS	Size-exclusion chromatography coupled to small-angle scattering
SEC-SAXS	Size-exclusion chromatography coupled to small-angle X-ray scattering
SLD	Scattering length density
SLS	Static light scattering
SMA	Styrene-maleic acid
SMALP	Styrene-maleic acid lipid particle
SMI	Styrene-maleimide
SMILP	Styrene-maleimide lipid particle
sND	Stealth nanodisc
sNMR	Solution NMR
ssNMR	Solid-state NMR
T_M	Transition melting temperature
TEV	*Tobacco etch* virus
TMD	Transmembrane domain
TOF	Time-of-flight

Part I

INTRODUCTION AND EXPERIMENTAL OUTCOME

1

MEMBRANE PROTEIN CARRIER SYSTEMS

1.1 INTRODUCTION

Structural and biophysical characterization of membrane proteins (MPs) is one of the major challenges in biology. Under native conditions, MPs reside in highly complex biological membranes, but for *in vitro* studies, they must generally be isolated and purified in presence of much simpler agents, while retaining activity and stability. A plethora of such agents, here termed ”carrier systems”, have been developed to facilitate MP research.

Following a general introduction to biological membranes and MPs, this chapter aims to present an overview of current state-of-the-art MP carrier systems together with a qualitative consideration of their advantages and disadvantages related to structural and functional studies of MPs in solution. In this thesis, detergent micelles and nanodiscs were thoroughly used for stabilizing a number of MPs with the goal of obtaining structural information from small-angle scattering (SAS) experiments. These systems together with their use in SAS studies are given special attention.

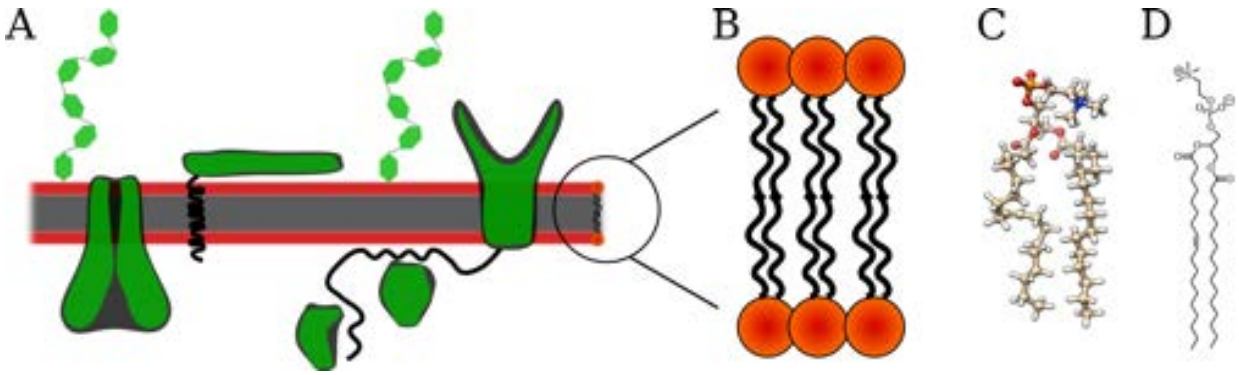


Figure 1 Biological membranes. A: A schematic drawing of a biological membrane, which is a complex medium compromising an overall lipid bilayer structure with membrane proteins and carbohydrates embedded and/or attached. **B:** Zoom on the internal organization of the lipid bilayer, showing the hydrophilic lipid head-groups in red, and the hydrophobic lipid tail-groups in black. **C:** Molecular representation of a POPC lipid, shown in ball and stick representation. **D:** Chemical structure of POPC.

1.1.1 *Biological membranes*

In all living organisms, the cell and cellular departments, including organelles, are surrounded by semipermeable, selective barriers called biological membranes that are highly complex mediums of different lipids, proteins and carbohydrates (Figure 1A) (Nicolson 2014; Robertson 2018). As explained in Section 1.1.1.1, the overall structure of a biological membrane is a lipid bilayer. In eukaryotes, the major constituents are glycerophospholipids (phospholipids) (Figure 1B), of which phosphatidyl-choline (PC) lipids, such as 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), are the most prevalent (Figure 1C and D) (Meer et al. 2008). The molecular composition of biological membranes vary between different cell types and different organelles, and already at the lipid level, the complexity is enormous with thousands of different lipids (Sud et al. 2007). Furthermore, membranes are both laterally and vertically heterogenous. In the lateral dimension, they can form so-called lipid rafts that are local transient lipid domains governed by specific lipid-lipid or MP-lipid interactions (Sezgin et al. 2017), and in the vertical dimension, they can have asymmetric distributions of lipids between the two leaflets of the lipid bilayer (Kiessling et al. 2009). Cholesterol is another common lipid involved in many biological processes, including domain formation and thereby also regulation of MPs (Grouleff et al. 2015; Sezgin et al. 2017). While most biochemical work was previously focused on understanding MPs alone, the role of biological membranes to participate specifically in biochemical processes has started to attract much attention, including the interplay between MPs and lipids, lipids as messenger chemicals in signalling and lipids as bioactive molecules, (Meer et al. 2008; Contreras et al. 2011; Cournia et al. 2015; Harayama and Riezman 2018). Furthermore, as pointed out by Heberle *et al.*(Heberle and Pabst 2017), eukaryotic cells are packed with membrane surfaces, which hint that many biochemical processes occur in the vicinity of biological membranes. As a result, new research fields have emerged, such as lipidomics (Yang and Han 2016; Lydic and Goo 2018). Although a typical bacterial membrane does not exist, bacterial membranes generally have much different composition than eukaryotic membranes. Notably, bacteria don't have cholesterol and PC lipids are rare (Sohlenkamp and Geiger 2016).

1.1.1.1 *Organization of lipids*

Phospholipids are commonly described by a head-tail structure (Figure 1B) that is amphiphilic, *i.e.* has both hydrophilic and hydrophobic properties. Common for all phospholipids, the hydrophilic head-group consists of a negatively charged glycerophosphate-group and, and specific for PC lipids, a positively charged choline group as well, which makes PC lipids zwitterionic (Figure 1C and D). The hydrophobic tail-group consists of two or more hydrocarbon chains. In the case of POPC, these are a palmitoyl (C_{16}) chain at position 1 and a oleoyl (mono-unsaturated C_{18}) chain at position 2. Cardiolipin, which is prevalent in many

bacteria and in the mitochondria of eukaryotes, is an example of a phospholipid that contains four hydrophobic tail-groups, and it furthermore also has two glycerophosphate-groups in the headgroup (Schlame 2008).

The overall bilayer structure of biological membranes can be mimicked in solution by pure PC lipids that self-assemble into liposomes or vesicles (Lasic 1988; Kučerka et al. 2011). Self-assembly is by definition a process, where a system of disordered molecules forms ordered structures guided by specific intermolecular interactions between the molecules themselves, but not directed by an external source (Palmer et al. 2007). The major driving force for self-assembly of lipids and other amphiphilic molecules is the hydrophobic effect as introduced by Charles Tanford (Tanford 1980): It describes the tendency of hydrophobic and hydrophilic molecules to separate, resulting in packing of the hydrophobic tail-groups against each other and with the hydrophilic headgroups facing the aqueous solvent (Figure 2). It turns out that the higher order structure of head-tail molecules can be predicted from their packing parameter as proposed by Jacob Israelachvili (Israelachvili 1991) and given by

$$P = \frac{V}{A \cdot l} \tag{1}$$

where V is the volume of the tail-group, A is the cross sectional area of the headgroup, and l is the length of the tail-group. A packing parameter below unity indicates a cone shape, which leads to spherical or elongated structures called micelles (Figure 2A), whereas packing parameters close to unity indicate cylindrical overall shape and leads to bilayer structures (Figure 2A). Saturated and mono-unsaturated phospholipids, such as 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) and POPC, have packing parameters close to unity (Israelachvili 1991).

Self-assembly displays a high degree of cooperativity and occours in a narrow concentration range (Wennerström 1979). For micelle forming agents, this is referred to as the critical micelle concentration (CMC) (Figure 2B). For phospholipids, the onset of self-assembly is preferentially termed the critical aggregate concentration (CAC), indicating that the formed structures are not micelles. Many biologically relevant lipids have negligible solubility as monomers and thereby equally low CACs (Israelachvili 1991). Thus, the CAC is often assumed to be zero, implying that all lipids are found in ordered aggregates.

1.1.2 *Membrane proteins*

Because lipid bilayers have a hydrophobic interior, passive diffusion of polar and charged substances, including water, ions and proteins, is very energetically unfavorable. The selective permeability of biological membrane originates from the molecules that are embedded in the lipid matrix (Nicolson 2014). Integral MPs reside in biological membranes, where they mediate highly regulated processes between different cellular compartments and between the in- and outside of the cell (Nicolson 2014; Robertson 2018). MPs are

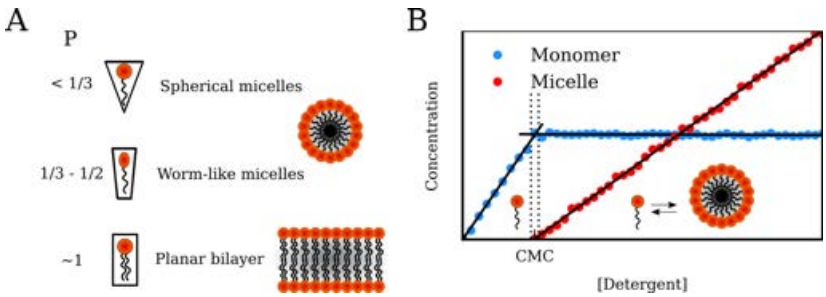


Figure 2 Packing parameter and CMC. A: Illustration of the packing parameters, P, of head-tail amphiphiles leading to micelles (1/3 - 1/2) or bilayers (1). B: Principle of the critical micelle concentration (CMC). The detergent CMC is indicated by the narrow concentration range within the dotted lines. Above CMC, detergent monomers are in equilibrium with detergent micelles.

typically divided in transmembrane proteins, which have a domain spanning the entire thickness of the membrane, and monotopic MPs that are only attached to one side of the membrane. Because the interior of the membrane is hydrophobic, it follows that the transmembrane domains (TMDs) of MPs are also hydrophobic. Regions protruding form the membrane are hydrophilic, and overall MPs are amphiphilic. Proteins interacting temporarily with membranes are sometimes referred to as peripheral membrane proteins. However, the focus will be solely on integral MPs in this thesis.

MPs constitute an estimated 25-35 % of the eukaryotic genome (Fagerberg et al. 2010). It is estimated that proteins constitute around 50% of the mass of biological membranes (Alberts et al. 2002), and like lipids, MPs can segregate and form concentrated domains within the membrane (Poveda et al. 2008). Their biological importance is reflected in the observation that MPs are the target of an estimated 60 % of all drugs on the market (Overington et al. 2006). Gaining insight to MP structure and function is thus of high potential impact, and this was acknowledged with the 2012 Nobel Prize that was given to Lefkowitz and Kobilka for their studies on an important family of MPs called G-protein coupled receptors (GPCR) (Kobilka 2013). However, the number of resolved MP structures is markedly lower than for water-soluble proteins: Around 800 unique high-resolution MP structures have been solved (Stephen H. White 2018), while the total number of unique high-resolution protein structures in the Protein Data Bank (PDB) (Berman et al. 2000) is around 50000, depending on the threshold set for sequence similarity. This clearly reflects the challenges involved with obtaining high quality MP samples.

1.1.2.1 *Challenges in MP research*

Structural and functional characterization of MPs remains challenging compared to water-soluble proteins. Due to several bottlenecks in sample preparation, obtaining a stable MP sample in workable amounts is commonly a laborious process. MPs are often present at low level in their host organism and high yield heterologous expression in *E. coli* is difficult and does not provide potentially vital post-translational modifications (Wagner et al. 2007). Several expression organism might be tested, and incorporation of reporter proteins, such as green fluorescent protein (Bjørkskov et al. 2017), can help optimizing expression condi-

tions. Only if a satisfactory yield is obtained, purification schemes can be applied to isolate the MP. Typically, the first step involves the isolation of the expression organism's membrane, wherein the MP is usually stable. However, the following purification steps often require the MP to be kept in solution, and thus, MPs must be isolated from the highly complex biological membranes and into simpler model membrane mimicking particles (MP carriers).

Historically, detergent micelles (Section 1.2.1) have been the most widely used platform to extract MPs from biological membranes. Detergents can partition in the lipid bilayer where the MP resides, and provided that the relative concentration of detergent to lipid is high enough, bilayers will disrupt and transform into mixed micelles of lipid and detergent (Section 1.2.1.1 and Figure 5) (Garavito and Ferguson-Miller 2001; Rigaud 2002; Seddon et al. 2004; Lichtenberg et al. 2013). Subsequent purification steps are usually carried out in buffer containing the detergent above its CMC, and native lipids will be washed away from the MP eventually, unless tightly bound.The general problem with detergents is that they in general destabilize MPs. Finding a suited detergent is non-trivial and often requires huge screening efforts (Champeil et al. 2016).

The first high resolution MP structure came out in 1985 (Deisenhofer et al. 1985) and in the following twenty years, the growth of unique MP structures was exponential (Stephen H White 2004). This predicted exponential growth has since slowed down (Stephen H. White 2018), reflecting that MP crystallization is notoriously difficult compared to soluble proteins (Baker 2010). It requires control of both the MP and the carrier system, which all together will often be too dynamic to crystallize. Thus, other methods are in many cases required to study MP structure (Section 1.3.4).

However, despite all of the challenges associated with MP handling, the number of solved MP structures is still steadily growing each year (Shimizu et al. 2018). This reflects a high interest in the field and cocomitant buildup of expertise and method development. As part of this expertise, new and better carrier systems have been developed (Section 1.2) and used in combination with alternative methods to crystallography, such as cryo electron microscopy (cryoEM) (Mio and Sato 2018).

1.2 OVERVIEW OF MP CARRIER SYSTEMS

Detergents are the most widely used molecules for stabilizing MPs (Seddon et al. 2004; Stetsenko and Guskov 2017). Despite their popularity, it is well-acknowledged that detergent micelles are not good mimics of biological membranes, and the continuous need for studying more delicate MPs, often not stable in classic detergent micelles, have led researchers to develop better alternatives. During the last two decades, and with an increasing pace, different membrane mimicking particles have been developed for eliminating the bottle neck of obtaining stable MPs. Common for all of these, including detergent micelles, is that they are amphipathic, having a hydrophobic interior to accommodate the MP, and a hydrophilic exterior

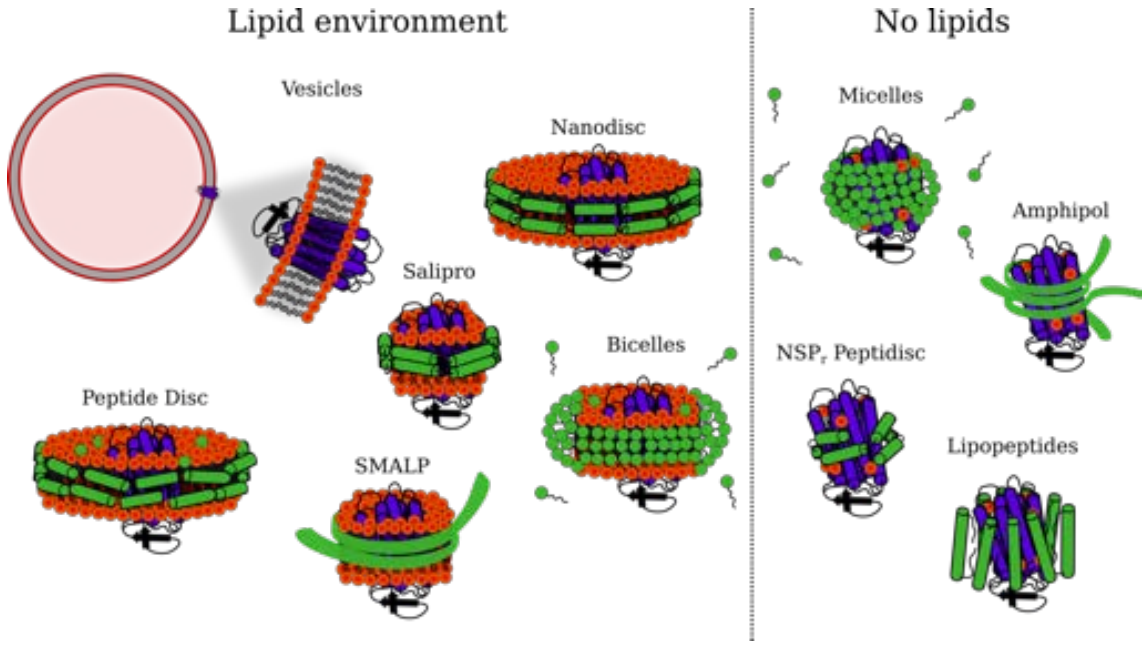


Figure 3 Overview of common membrane protein carrier systems. Common MP carrier systems can overall be divided into those that accommodate a lipid environment for the MP, and those that bind directly to the MP. For the latter, tightly bound lipids might still be present. Lipid environments are found in vesicles (Section 1.2.2), (MSP) nanodiscs (Section 1.2.7), Salipro (Section 1.2.6), peptide discs (Section 1.2.5.1), SMALP (Section 1.2.4.2) and bicelles (Section 1.2.3). Overall lipid-free environments are found in detergent micelles (Section 1.2.1), amphipol (Section 1.2.4.1), NSP$_r$ peptidiscs (Section 1.2.5.1) and lipopeptides (Section 1.2.5.2). See individual sections for abbreviations.

to remain soluble in aqueous solution. An overview of common systems are shown in Figure 3. Overall, these carrier systems can be divided into those that are designed to accommodate a lipid environment and those designed to bind directly to the MP (no lipids). Despite the latter being described as lipid-free, some tightly bound lipids originating from the biological membrane from which the MP was extracted can still be present. Figure 3 as well as the remaining figures in this chapter reflect that such tightly bound lipids might be present.

In this section, the specific carrier systems are introduced. This part was written as part of a literature study and is planned to be included in a review article. It first describes the commonly applied detergent micelle system and lipid vesicle system that provide fundamentally different environments for MPs. Hereafter, the other systems are introduced in the order of mixed detergent-lipid bicelles (Section 1.2.3), to polymer-based systems (Section 1.2.4), to peptide-based (Section 1.2.5) and saposin-based (Salipro, Section 1.2.6) systems. The section ends with introduction of the membrane scaffold protein (MSP)-based nanodisc system (Section 1.2.7) that was used as the major system for MP stabilization in this thesis.

1.2.1 *Detergent micelles*

The International Union of Pure and Applied Chemistry (IUPAC) defines a detergent as a surfactant with cleaning properties in dilute solution; in turn, a surfactant is defined as a substance which lowers the surface

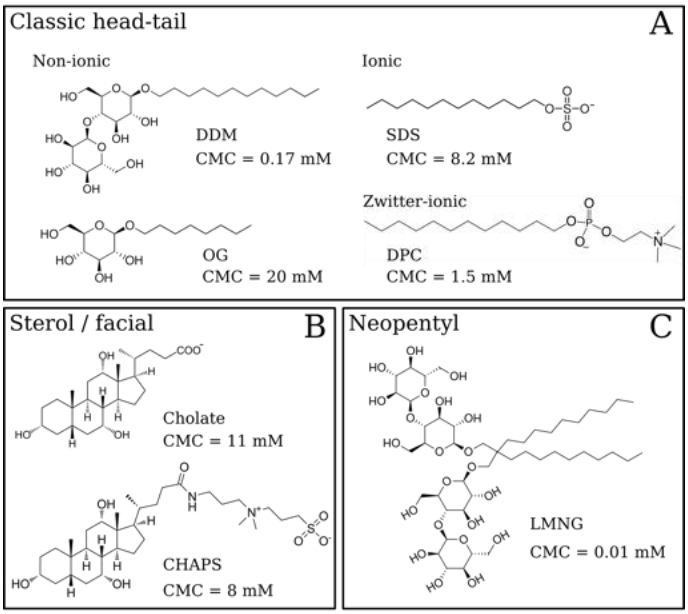


Figure 4 Detergent classification. Examples of detergents used for MP handling together with their estimated CMCs. **A:** Classic head-tail detergents. These are commonly classified based on the charge state of the headgroup. **B:** Sterol-based facial detergents. These contain a relatively planar sterol moiety with polar hydroxy groups facing the same side. **C:** Neopentyl class. The structure of a neopentyl detergent is equivalent to two head-tail detergents that have been linked at a tertiary carbon in the tail groups; in the case of Lauryl maltose-neopentyl glucol (LMNG) these are two n-dodecyl-β-D-maltoside (DDM) molecules. References for approximate CMC values in pure water: DDM (VanAken et al. 1986), n-octyl-β-D-glucoside (OG) (Lorber et al. 1990), sodium dodecyl sulfate (SDS) (Cifuentes et al. 1997), dodecyl phosphocholine (DPC) (Anatrace 2018), cholate (Garidel et al. 2000), 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) (Hjelmeland et al. 1983), LMNG (Chae et al. 2010).

tension of the medium in which it is dissolved, and/or the interfacial tension with other phases (A. D. McNaught and A. Wilkinson 1997). This implies that detergents are amphiphilic, *i.e.* having both hydrophilic and hydrophobic properties. In biochemistry and biophysics, the word detergent is commonly used for small amphiphilic micelle-forming molecules (Garavito and Ferguson-Miller 2001).

A vast amount of detergents have been developed and are used broadly in industry and household. These small compounds can be synthesized with relative ease and in high purity. Here, it is out of the scope to introduce them all, and only a few interesting classes of relevance for MP handling are covered. The following section is split to first cover standard head-tail detergents, which have defined the understanding of detergents and micelles, and later less conventional variants that are not fully described by the formalism used for head-tail detergents.

1.2.1.1 *Classic head-tail detergents*

Classic head-tail detergents can be classified by the charge-state of their headgroup and the type of tail-group (Figure 4) (Seddon et al. 2004). It is commonly found that anionic alkyl-tail detergents, such as sodium dodecyl sulfate (SDS), are denaturing for membrane proteins, while nonionic alkyl-tail detergents, such as n-dodecyl-β-D-maltoside (DDM) or n-octyl-β-D-glucoside (OG), are generally mild and non-denaturing. In relation to functional studies, zwitterionic detergents are generally more deactivating than the non-ionic

detergents (Seddon et al. 2004), but nevertheless, foscholines, such as dodecyl-phosphocholine (DPC) (Figure 4), are very effective for extracting MPs and have gained substantial popularity in nuclear magnetic resonance spectroscopy (NMR) studies (Chipot et al. 2018).

Most common head-tail detergents have packing parameters well below unity and self-assemble into micelles above their CMC (Figure 2B). The CMCs of common detergents are typically quite high (see Figure 4) compared to phospholipids (1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) has CAC = 0.5 nM (Smith and Tanford 1972)), and it is therefore important to consider the equilibrium between free and micelle bound states. The CMC depends on a number of properties, including the hydrophobicity of the tail-group and the polarity and charge state of the headgroup. CMC decreases with the hydrophobicity of tail-groups because of stronger intermolecular interactions between tail-groups and a higher penalty of solvation, whereas CMC increases for charged headgroups, which repel each other and thus unfavor packing in micelles (Oliver et al. 2013). Addition of salt screens charged headgroups and lower CMC significantly (Chattopadhyay and Harikumar 1996). Steric hindrance, related to the packing parameter, of the detergent molecules will determine the overall shape of the micelle and can range from spherical to worm-like (Figure 2A) (Stuart and Boekema 2007). The aggregation number, *i.e.* the number of detergent molecules in a micelle, is determined by a combination of all three properties mentioned above and increases with hydrophobicity of the tail-groups and decreases with electrostatic repulsion and steric hindrance of the headgroup (Oliver et al. 2013).

The extraction of MPs from biological membranes with detergents is termed solubilization and proceeds by a commonly accepted three-stage model (Figure 5A) (Lichtenberg et al. 2013). The detergency of a detergent is given by the ratio of detergent needed per lipid (R_{sat} and R_{sol}) for saturating and solubilizing lipid membranes, respectively (Figure 5B) (Vargas et al. 2015). Both detergent partitioning into vesicles as well as lipid partitioning into micelles govern this equilibrium (Vargas et al. 2015). For classic head-tail detergents, it is primarily governed by the intrinsic curvature of the specific lipids and the detergent and can be predicted from their CMC and water-to-bilayer partition coefficients (Heerklotz and Seelig 2001).

Micellar structures are often pictured as spherical and monodisperse for simplicity (Figure 2), but this is seldom the case and depending on the detergent type and the experimental conditions, they can adapt many different structures and should rather be described as overall prolate or oblate (Oliver et al. 2013). Micelles are highly dynamic, exchange monomers with the solvent rapidly, and can adapt to environmental structural fluctuations, as is the case going from pure detergent micelles to protein-detergent complexes (PDCs) (Garavito and Ferguson-Miller 2001). By incorporation of a MP, a new interface is introduced between detergent and protein, which is upheld by hydrophobic and electrostatic interactions that are specific to the MP (Wolfe et al. 2017). Detergent micelles stabilize MPs in solution by covering their hydrophobic TMD (Figure 5A) (Garavito and Ferguson-Miller 2001; Berthaud et al. 2012). The detergent aggregation

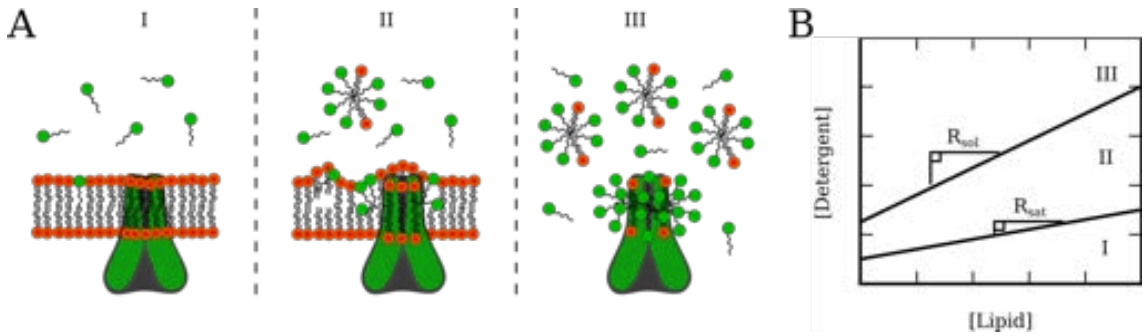


Figure 5 Three-stage model for membrane solubilization. **A:** The three-stage model of membrane solubilization involves membrane partitioning by the detergent molecules until saturation (I), over-saturation of the membrane and appearance of mixed micelles (II), and completely solubilization of the membrane into mixed micelles (III). The three depictions are representations of evolving processes as described by Lichtenberg et al. (Lichtenberg et al. 2013). **B:** The ratio of detergent required to saturate the membrane is termed R_{sat}, while the ratio required to solubilize the membrane is termed R_{sol}.

number in PDCs depends on the MP and will obviously be different, and often higher, from the aggregation number in MP-free micelles (Chaptal et al. 2017). Furthermore, detergent behaviour is influenced by several environmental effects, such as ionic strength, pH, counter ions, concentration and temperature (Israelachvili 1991). Overall, detergent micelles and PDCs are expected to be quite dynamic and polydisperse structures.

In the context of MP handling, the main drawback is that detergent micelles are structurally very different from lipid bilayers and therefore destabilizing for MPs. Furthermore, detergents are added above their CMC, and free micelles will be present. It has been suggested that detergents are denaturing because of the exact same properties that make them so effective at solubilizing membranes: Detergents have the ability to disrupt native interactions, including protein-protein interactions and protein-lipid interactions. Thus, the presence of free micelles can be thought of as a ”hydrophobic sink” that can absorb molecules, including lipids, that were otherwise natively associated to MPs (Jean-Luc Popot 2010). As mentioned in the beginning of this section, the terminology introduced so far has mainly been applicable to head-tail detergents. In the following, some less conventional detergents are described.

1.2.1.2 *Other detergents*

Bile salt acid derivatives, such as cholate or CHAPS (Figure 4), are generally mild, but are not between the most popular detergents for MP handling (Seddon et al. 2004; Stetsenko and Guskov 2017). This class of detergents is structurally very different from more commonly used head-tail detergents. In cholate and CHAPS, polar hydroxy-groups are present throughout the sterol moiety and far from the charged carboxyl-group (Figure 4). Furthermore, these hydroxy groups sit on the same side of the sterol moiety, which gives the detergent an overall facial character. This has been exploited to develop complex, substituted variants with more specific properties (Q. Zhang et al. 2007). Due to the overall different structure to the classic head-tail detergents, facial amphiphiles are not expected to form standard micelles. Cholate, for instance, likely forms either dimers or small assemblies of dimers (Maslova and Kiselev 2018). Although their CMCs

are usually relatively high, facial detergents are generally highly efficient at solubilizing lipid bilayers. A qualitative explanation for this is their tendency to favour discoidal mixed micelles, which accomodate a bilayer-like packing of lipids and thereby favorizes lipid partitioning into the mixed micelles (Vargas et al. 2015). Furthermore, they are widely used for preparation of other MP carrier systems, such as nanodiscs (Section 1.2.7) and bicelles (Section 1.2.3).

Maltose-neopentyl glucols (MNG) is another interesting class of detergents. MNGs have an overall head-tail structure, but with a more complex organization than classic detergents. This is exemplified by lauryl-MNG (LMNG), which is basically two DDM molecules linked at a tertiary carbon in their tail-groups. The LMNG detergent is not well-studied, but a single study has indicated that it form unconventional and highly polydisperse micelles (Chaptal et al. 2017). With an estimated value of 10 nM, the CMC of LMNG approaches the range of CAC values for phospholipids ($\sim$0.5 nM). Thus LMNG is expected to be very tightly bound to MPs, and it might be possible to carry out MP purification steps in detergent-free buffers. Finally, LMNG induces markedly increased MP stabilization compared to DDM and has furthermore been used for crystallization of difficult GPCRs (Chae et al. 2010).

Fluorinated detergents have the same overall chemical structure as conventional detergents, but with F exchanged for H in the tail-groups. They do not partition well into hydrogenated lipid bilayers and hydrogenated lipids are thought to also partition poorly into micelles of fluorinated detergents (Jean-Luc Popot 2010; Vargas et al. 2015). Furthermore, fluorinated detergents might have low affinity for MPs and thereby being less efficient at competing for stabilizing protein-protein interactions, which is expected to alleviate the denaturing effect of free micelles (Jean-Luc Popot 2010). However, some short-chained fluorinated detergents were recently shown to solubilize POPC vesicles and MPs from native membranes very efficiently (Boussambe et al. 2018). This is counterintuitive given the poor miscibility of hydrocarbons and fluorocarbons (Jean-Luc Popot 2010; Vargas et al. 2015), suggesting that the mechanism might be different from standard detergents. Overall, the use of fluorinated detergents for MP handling holds great potential, but lacks a broader understanding of the mechanism of solubilization compared to standard detergents.

The many different classes of detergents (only a few were presented here) makes it difficult to generalize their properties for MP extraction and stabilization. However, detergent micelles do not form the fundamental bilayer structure that is found in biological membranes. To get that property, other systems must be applied.

1.2.2 *Vesicles*

Vesicles are large lipid bilayers that enclose a volume of liquid (Lasic 1988). The lipid bilayer of vesicles is commonly seen as the best among the simple, soluble model systems for biological membranes. They can

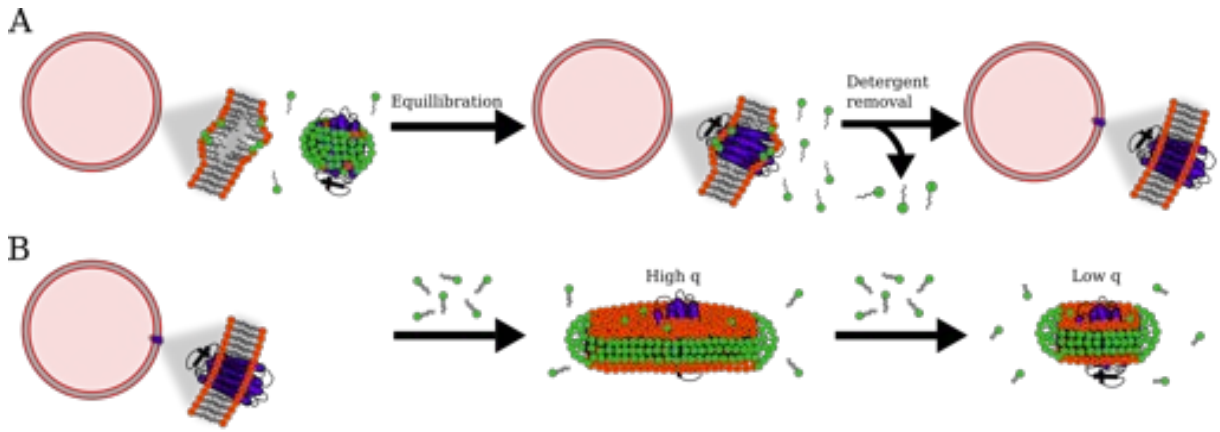


Figure 6 MP reconstitution in vesicles and bicelles. A: Detergent-assisted MP insertion in vesicles. Preformed vesicles are mixed with a MP solubilized in detergent at a concentration close to R_{sat}. The MP will spontaneously participate in the lipid bilayer (usually with random orientation). Detergent removal leaves the MP reconstituted in pure lipid vesicles. **B:** Bicelle formation by "q-titration". MP-loaded vesicles of a bicelle compatible lipid, are titrated with a bicelle compatible detergent whereby bicelles form. The size of the bicelle decreases with the molar ratio of detergent to lipid.

be formed with a single (uni-lamellar) or with multiple lipid bilayers (multi-lamellar) per vesicle and in a wide variety of sizes using many different techniques (Lasic 1988). Uni-lamellar vesicles can be obtained by extrusion, where a suspension of multi-lamellar vesicles, obtained *e.g.* by hydrating a lipid film, is passed multiple times through a polymeric membrane with well-defined pore size. The size of the vesicle depends on the pore size, and typically ranges from 100 nm to 1 μm (Lasic 1988). Small vesicles will have higher curvature than larger vesicles, and this parameter might affect the structure and function of MPs (Epand et al. 2015). Giant uni-lamellar vesicles that are on the order or several μm in diameter, thus approaching the size of cells, can be obtained by other means (Jørgensen et al. 2017).

MPs can, among others approaches, be reconstituted into preformed vesicles by detergent-assisted insertion, where the lipid bilayer is destabilized by addition of detergent to R_{sat} followed by removal of the detergent (Figure 6A) (Rigaud 2002). Detergent removal with polymeric beads that contain a hydrophobic interior (Rigaud et al. 1998) or by dialysis is possible because the concentration of free lipids (CAC) is generally neglible (Section 1.1.1.1). Using this technique presumably provide more monodisperse samples compared to *e.g.* preparation from mixed-micelles and detergent depletion or sonication techniques (Wang and Tonggu 2015). By detergent-assisted reconstitution, MPs are usually inserted in random orientation in the vesicles, but including charged lipids can for some MPs favor controlled direction of insertion (Tunuguntla et al. 2013).

An advantage of vesicles in the context of functional studies is compartmentalization, which makes it possible to measure vectorial function (Rigaud 2002). The inside of the vesicle is effectively separated from the outside solvent, which can be exchanged to create chemical gradients and thereby conduct functional studies on MP transporters and channels. However, because of their large size, and often low abundance of MP compared to lipid, vesicles are generally not suited for structural studies. Furthermore, large vesicles

are dense and require magnetic stirring or alike to effectively to be kept in solution, *i.e.* it can be discussed whether they are actually in solution or better can be thought of as a pseudo-phase.

Efforts have been made towards generation of vesicle-like structures with higher stability than pure lipid bilayers. This can be achieved with di- or tri-block co-polymers to generate so-called polymersomes or hybrid vesicles made with mixtures of such polymers and lipids (Beales et al. 2017). MPs can retain their structure and function in polymersomes (Garni et al. 2017), and thus, the technology has wide industrial applications, *e.g.* for making selective membranes (González-Pérez et al. 2018). By the same arguments as for pure lipid vesicles, polymersomes and hybrid vesicles are also not generally suited for structural studies on MPs.

In general, there is a need for small carrier systems that have the stabilizing properties of lipid bilayers, but the size and versatility of detergent micelles to perform structural studies of MPs.

1.2.3 *Bicelles*

Bicelles bridge between detergent micelles and pure lipid vesicles. They are formed when mixing certain long-chain lipids with detergents or short-chain PC, such as 1,2-dihexanoyl-sn-glycero-3-phosphocholine (DHPC), or certain detergents (Sanders and Prosser 1998). The mixture of DMPC and DHPC has been the most popular choice for bicelle formation (Dürr et al. 2012): DMPC lipids form a central bilayer stabilized on the rim by DHPC that effectively functions as a detergent (Figure 6B) (Luchette et al. 2001). Facial amphiphiles, in particular 3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO), and more complex versions derived from cholate have also been proven well-suited for generation of stable bicelles (Matsui et al. 2016).

The main determinant of bicelle morphology is the molar ratio, q, of long and short-tail lipids or detergents; the higher the q, the larger the size of the bicelle (Dürr et al. 2012). Bicelle morphology is sensitive to environmental effects, including temperature, hydration level, and ionic strength (Nieh et al. 2002; Yamamoto et al. 2009). It is thus important to be aware of many restrictions in the phase diagram for bicelle formation. Bicelles coexist with free short-tail PC or detergent in a concentration termed the critical bicelle concentration in analogy with CMC (Beaugrand et al. 2014). Thus, bicelle structure is also sensitive to dilution. Finally and although still termed bicelles, it has been shown that DMPC-DHPC mixtures are better described as mixed micelles below q = 1 (Beaugrand et al. 2014), *i.e.* particles without the characteristics of a lipid bilayer.

MP reconstitution in bicelles can follow different schemes (Dürr et al. 2012). For instance, can MPs already reconstituted in vesicles (Figure 6A) be titrated with a suitable detergent/DHPC in a so-called q-titration, whereby the optimal q can be investigated (Figure 6B) (Son et al. 2012). Alternatively, detergent-

solubilized MPs can be mixed with vesicles, similar to the method for reconsitution in vesicles (Figure 6A), but with a bicelle compatible detergent above its R_{sol}. It is important to consider that detergent removal is not applicable for bicelles, and thus, it is necessary to purify MPs in suitable detergents to carry out the latter method. It is also possible to refold MPs into bicelles (Lyukmanova et al. 2012).

While bicelles or q>1 can provide a lipid bilayer environment, they coexist with free detergent or short-chain lipids. The presence of free detergent was mentioned as a drawback of using detergent micelles for MP handling in Section 1.2.1, and although the amount of free detergent co-existing with bicelles is probably lower than for micelles, it is desired to use systems with low CMC/CAC that are not expected to have this property.

1.2.4 *Amphiphilic co-polymers*

Amphiphilic co-polymers generally have low CAC and have become extremely popular in MP research. The two most prevalent systems are amphipols (APols) (Tribet et al. 1996) and styrene-maleic acid (SMA) co-polymers (Knowles et al. 2009), which will be adressed in more detail in the following sections. Amphipols were designed to eliminate the denaturing properties of detergent micelles (J.-L. Popot et al. 2011), whereas SMA has been mainly explored for its ability to capture MPs together with patches of native membrane, without the aid of detergent, in so-called SMA lipid particles (SMALP) (Dörr et al. 2014). The common advantages and drawbacks of polymers are related to their large scale synthesis that on one hand is inexpenssive, but on the other hand gives rise to some polydispersity in the final product in terms of length and composititon (Giusti et al. 2014; Dörr et al. 2016). Furthermore, one of the two building blocks in the most common co-polymers, including SMA and amphipols, often contains a carboxylate-group. Carboxylates can chelate divalent metal ions and they are partly protonated at neutral or slightly acidic pH, which leads to lowered solubility. Thus, many co-polymers are incompatible with these conditions and thereby not adequate for studying *e.g.* magnesium transporters (Maguire 2006). Recently, a wide array of polymers with different functionality are being developed to ensure compatibility with different applications (Simon et al. 2018). However, this section only introduces the most prominent and well-understood systems.

1.2.4.1 *Amphipols*

APols were proposed by Jean-Luc Popot and co-workers (Tribet et al. 1996) and their use in MP research has been reviewed on several occasions (see *e.g.* (J.-L. Popot et al. 2011; Jean-Luc Popot and Engelman 2016; Zoonens and Jean-Luc Popot 2014; Le Bon et al. 2018)). APols are short amphiphilic co-polymers that contain multiple hydrophobic and hydrophilic groups and come in a wide variety of versions (J.-L. Popot et al. 2011; Zoonens and Jean-Luc Popot 2014). The first APols were based on polyacrylate templates

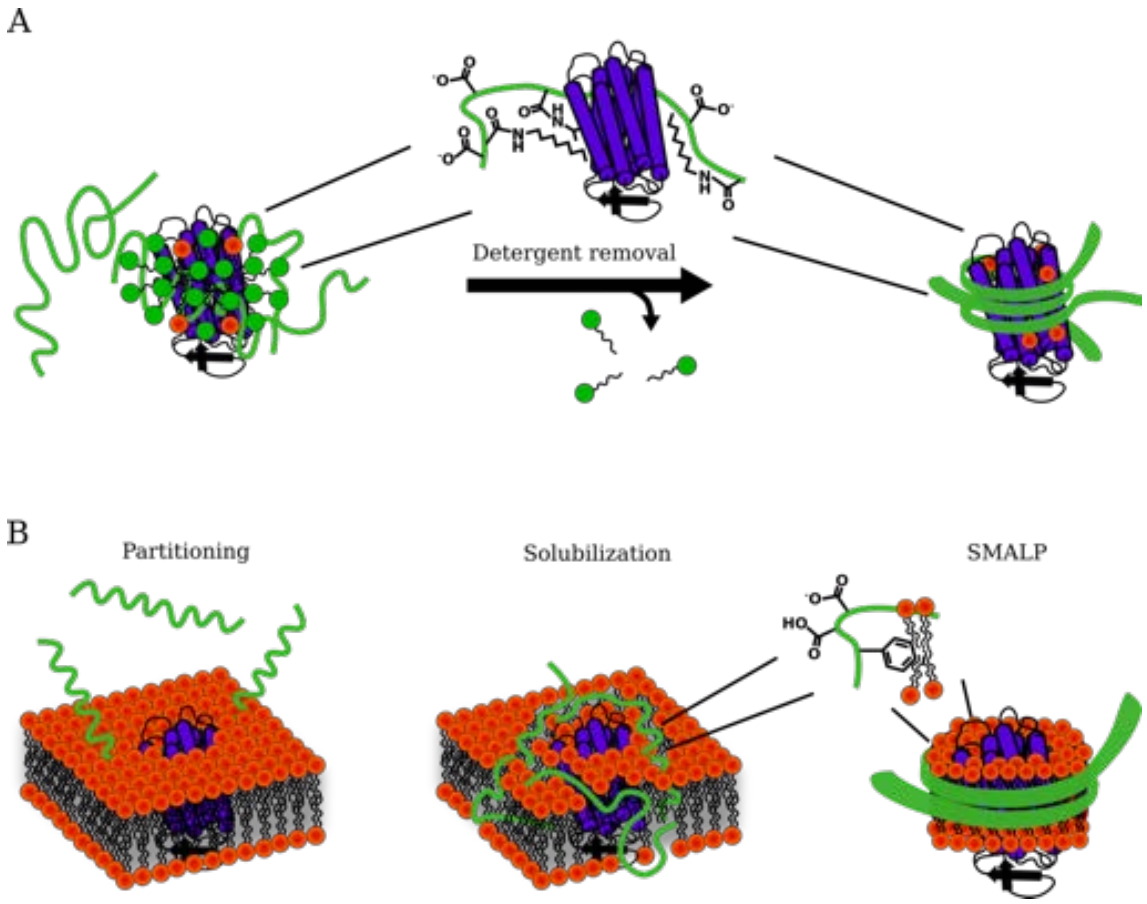


Figure 7 Amphiphatic co-polymers. A: MP reconstitution into amphipols from a detergent solubilized step. Amphipols compete with detergent binding to the MP and eventually obtains several interactions-sites per polymer to give high avidity. **B:** MP solubilization from a lipid bilayer using the styrene maleic-acid (SMA) polymer in three stages yielding the final SMALP containing both the MP and lipids.

with hydrophobic groups, for instance octyl, grafted onto a portion of these (Figure 7A) (Tribet et al. 1996). The best characterized APol is A8-35, whose name indicates that this polymer is anionic, has an average molecular weight of 8 kDa, and has 35 % of the template carboxylates ungrafted. APol A8-35 is well studied with and without MPs incorporated. The free polymer collapses into well-defined particles resembling detergent micelles in solution at slightly basic pH ($\geq$ 7.5), compromising an average of four polymers and thereby 75-80 octyl groups, which is close to that found in OG micelles (Gohon et al. 2006). The CAC of A8-35 has been estimated to 2mg/l, which for all practical applications is negligible (J.-L. Popot et al. 2011).

APols are generally not capable of solubilizing membranes, although they interact with and can permeabilize membranes (Champeil et al. 2000). Functional reconstitution of MPs thus follows a common scheme for many MP carriers, where the MP is first solubilized in detergent, then incorporated in APols followed by detergent removal (Figure 7A) (Gohon et al. 2008). Compared to a detergent molecule, an APol has several interaction sites with an MP, providing binding avidity and causing slow exchange rates. The increased

stability of MPs in APols is presumably due to abscence of free surfactant and thereby no denaturing "hydrophobic sink" effect (J.-L. Popot et al. 2011).

The main drawback of amphipols is the general need for detergent in sample preparation. However, APols have proven good at refolding MPs from a denatured state in the absence of detergent (Jean-Luc Popot and Engelman 2016). The high abundance of negative charges in A8-35 is inhibitory for applications in cell-free synthesis of MPs, but non-ionic homopolymeric APols are highly efficient for this purpose (Bazzacco et al. 2012). A further advantage of APols is that deuterated precursor molecules can be used in the synthesis to produce partially or perdeuterated polymers (Gohon et al. 2006; Giusti et al. 2014), which are desired for NMR and small-angle neutron scattering (SANS) studies (Section 1.3.4).

1.2.4.2 *SMALP*

SMA is a co-polymer synthesized from styrene and maleic anhydride followed hydrolysis (Dörr et al. 2016). The co-polymer, which has a yearly global production in several tons, has wide applications in industry and science, but it applicability for stabilizing MPs was first explored a decade ago (Knowles et al. 2009). SMA and other related co-polymers, such as diisobutylene-maleic acid (DIBMA) and styrene-maleimide (SMI), are capable of directly solubilizing lipid bilayers, both synthetic and native, producing discoidal lipid particles (SMALP, DIBMALP and SMILP, respectively) (Figure 7) (Tonge and Tighe 2001; Oluwole et al. 2017; Hall et al. 2018). Thus, MPs can be extracted from biological membranes together with their native environment without the aid of detergents (Dörr et al. 2016). The resulting lipoprotein particles have consequently been called native nanodiscs (Dörr et al. 2014). The use of SMALP and derivatives for MP handling has been reviewed elsewhere (see *e.g.* (Dörr et al. 2016; Simon et al. 2018)).

The mode of action in membrane solubilization by SMA has been described by a three-stage model very similar to the one shown for conventional detergents in Figure 5: First, SMA binds to the membrane, then it inserts into and destabilizes the membrane, and finally SMALPs are formed (Figure 7B) (Scheidelaar et al. 2015). The binding event is driven mainly by hydrophobic interactions and modulated by electrostatic interactions, whereas the full insertion and destabilization of the membrane depends on acyl chain packing and lateral pressure of the membrane (Scheidelaar et al. 2015). The high efficiency of SMA to solubilize lipid bilayers reflects that SMALPs favour discoidal structures that accomodate the bilayer packing of lipids, *i.e.* the energy penalty of transferring lipids from the bilayer to the SMALP is low (Vargas et al. 2015).

The most common SMA polymers are termed SMA(3:1) or SMA(2:1) given their ratio of styrene to maleic acid groups (Dörr et al. 2016). The average weight of SMA polymers used in studies of MPs has been 7.5 kDa to 10 kDa and with a polydispersity index in the range 2.0 to 2.5 (Dörr et al. 2016). The more recently explored DIBMA is larger and has an average molecular weight of 37 kDa (Oluwole et al. 2017). SMALPs are widely studied by electron microscopy (EM) and found to have an average size

of 13 nm (Cuevas Arenas et al. 2016), whereas DIBMALPs are accordingly larger at approximately 18 nm (Oluwole et al. 2017). A more detailed structural description of SMALPs (SMA(2:1)) was pursued by SANS, but the proposed discoidal core-shell model did not fit the data well (Jamshad et al. 2015). Thus, the detailed structure of SMALPs remain largely unresolved and it is unclear to what extend the polymer inserts in the proposed lipid bilayer (Cuevas Arenas et al. 2016).

It is thus important to consider the physico-chemical properties of the lipids in SMALP and other related co-polymer particles. For SMALP, the styrene groups are believed to strongly interact with lipids, in effect disordering them to a state far from that in pure bilayers (vesicles) (Cuevas Arenas et al. 2016). This is further evident from the extremely fast lipid exchange rates observed for SMALPs (Cuevas Arenas et al. 2017), which gives an overall picture of highly dynamic particles. In fact, SMALPs are so dynamic that they can even exchange MPs (Broecker et al. 2017). Exchanging the phenyl-groups of SMA to isobutyl in DIBMALP releviate some of these effects, but with the consequence of a reduced lipid bilayer solubilization efficiency (Oluwole et al. 2017). Increasing the size of the SMA polymer also decreaes exchange rates, but at the same time reduces solubilization efficiency (Domínguez Pardo et al. 2018). Despite the fast dynamics, SMALPs and derivatives are in general found to stabilize MPs efficiently compared to detergents (Dörr et al. 2016).

Co-polymer based systems are inexpensive to produce, but the relatively random nature of their synthesis gives rise to a polydisperse product. Other polymers, such as peptides, can be synthesized with a high degree of control, leading to pure products. These are introduced next.

1.2.5 *Peptides*

Peptide secondary structure, α helix or β sheet, provides a great scaffold for designing amphiphilic molecules as amino acids point outwards in a highly predictable manner. A wealth of different amphiphilic peptides have been explored for their ability to stabilize MPs. Unlike the co-polymers, peptides are synthesized in a highly controlled manner, which yields a product of greater purity (Jaradat 2018). The cost of synthesis depends on the length of the peptide and is generally much more expensive than for co-polymers. An advantage of common MP stabilizing peptides is that they are zwitterionic because they contain several different amino acids. Thus, peptides are in general compatible with a wide range of pH and additives. This section will focus on amphipathic α-helical peptides consisting only of amino acids, but will also include a few examples of further chemically modified peptides, so-called lipopeptides.

1.2.5.1 *Apo-A1 peptides*

A number of amphiphilic peptides mimicking the helical structures found in the lipoproteins, such as Apo-A1 of the human body (Rothblat and Phillips 2010) have been synthesized and investigated for their ability to stabilize MPs. One thoroughly studied peptide of this class was designed from the 18 amino acid consensus sequence of the helical repeats in Apo-A1 and hence named 18A (Saito et al. 2004). Mixed with lipids, 18A makes discoidal particles with a central lipid bilayer stabilized on the rim by 18A (Midtgaard et al. 2014). These particles are here termed peptide discs. Bi-helical peptides consisting of two 18A sequences separated by a single amino acid linker, such as proline (Saito et al. 2004; Kariyazono et al. 2016), also form peptide discs (A. N. Larsen et al. 2016), and a 14 amino acid peptide (truncated 18A) was used for generating large, magnetically alignable peptide discs, termed macrodiscs (Park et al. 2011). From studies of Apo-A1 helix 10 (22 amino acids) (Imura et al. 2014) and truncated versions, it has been suggested that a minimum of 14 amino acids is required for the ability to form peptide discs (Imura et al. 2014).

A notable observation for peptide discs is that they self-assemble to relatively well-defined structures that however evolve to larger species relatively quickly after size-exclusion chromatography (SEC) purification (Midtgaard et al. 2014; A. N. Larsen et al. 2016). Increased stability was achieved using a modified self-polymerizing 18A peptides, indicating a relationship between peptide length and stability of the formed peptide discs (Kondo et al. 2016). In line with this, peptide discs made with bi-helical 18A had increased stability over 18A peptide discs (Midtgaard et al. 2014; A. N. Larsen et al. 2016). This would indicate that short peptides are highly dynamic and exchange rapidly from lipid-bound to lipid-free states, which has also been indicated by lipid exchange studies (Miyazaki et al. 2010).

The latest addition to the ensemble of Apo-A1 mimicking peptides was generated from the bi-helical 18A with a proline linker. It had its amino acid sequence reversed in order to increase solubility and was consequently named reverse nanodisc scaffold protein (NSP_r) (Carlson et al. 2018). NSP_r can be used for direct stabilization of MPs, *i.e.* without added lipids, into particles termed peptidiscs. However, the use of the nanodisc and peptidisc terminology is somewhat confusing: The term nanodisc (and disc) has so far been used for discoidal particles containing lipids, which the peptidisc is not expected to resemble.

18A peptides can solubilize membranes (Miyazaki et al. 2010), but MP reconstitutions in peptide systems have in most cases been detergent-assisted. For 18A and bihelical 18A peptide discs, MP insertion was achieved by simply mixing a detergent solubilized MP into a solution of peptide discs followed by SEC purification (Midtgaard et al. 2014; A. N. Larsen et al. 2016). Notably, it was not possible to remove detergent by convential methods of dialysis or detergent absorbing beads, as these procedures would also remove the peptide. Together with the observed increase in size after purification by SEC, this suggest that these peptides have an appreciated CAC. For lipid-free reconstitution, detergents could be readily exchanged

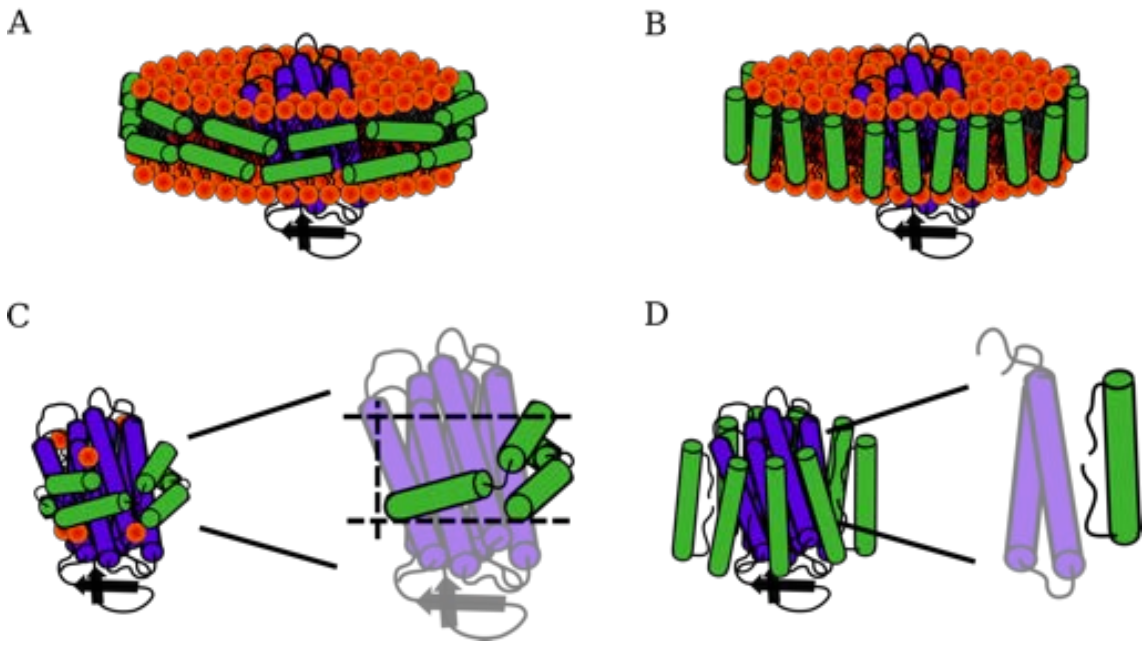


Figure 8 Peptide-based carriers. A: Illustration of a MP incorporated in a peptide disc with the peptides oriented mainly orthogonal to the bilayer normal. **B:** Illustration of the "picket-fence" model with the peptides oriented mainly parallel to the bilayer normal. **C:** Illustration of the suggested tilted orientation for the bihelical reverse nanodisc scaffold protein (NSP_r) peptide. In the zoom-in structure, punctuated lines are shown for comparison to orthogonal and parallel orientation to the bilayer normal **D:** Model of lipopeptides surrounding a MP. The zoom-in structure shows the probable orientation of the alkyl-tails grafted to the peptide.

for NSP_r, *e.g.* on an affinity column (Carlson et al. 2018). This suggest the opposite that NSP_r has a very low CAC, and thus indicates that the CAC is related to the presence or absence of lipids.

The understanding of the orientation of amphilic α-helical peptides in peptide discs vary from a "picket-fence" model with the peptide being mostly parallel to the bilayer normal (Figure 8B) (Islam et al. 2018) to a model with the peptide being mostly orthogonal to the bilayer normal (Figure 8A) (Midtgaard et al. 2014; A. N. Larsen et al. 2016). Notably, the later orientation was confirmed by NMR for the 18A peptide (Mishra et al. 2006), and it is clearly understood that bi-helical peptides (37 amino acids) are too long to maintain hydrophobic contacts throughout the peptide in the "picket-fence" model (Figure 8C). Thus, the "picket-fence" model is unlike to be correct for most peptides. In the orthogonal orientation (Figure 8A), the MP:peptide ratio is expected to vary discretely by a factor of two. However, this was not the case for the lipid-free NSP_r peptidisc as evaluated by mass spectometry (Carlson et al. 2018). The authors suggested a tilted orientation for NSP_r, which would also be highly adaptable to the varying length of the TMDs of different MPs (Figure 8C). Indeed, NSP_r was shown to stabilize a number of different MPs with varying sizes compared to detergent micelles (Carlson et al. 2018).

1.2.5.2 *Lipopeptides*

Lipopeptides are alkyl-modified peptides that form micellar structures in solution and can be used to stabilize MPs (Privé 2009). The stabilization is thought to originate from a preferred cylindrical packing of

the alkyl-tails towards the MP (Figure 8C) (McGregor et al. 2003), which suggests that lipopeptide detergents pack according to the "picket-fence" model (Privé 2009). The grafted alkyl-chains originate from the ends of the peptide and point toward the center of the peptide, making a structural interface resembling that of lipid tails in a bilayer (Figure 8C). While the α-helix has been widely used for generating lipopeptides (McGregor et al. 2003; Sauer et al. 2006; Privé 2009), a single study reports on β-strand lipopeptides for MP stabilization (Tao et al. 2013). These peptides were chemically synthesized with octyl groups in both ends to function as the hydrophobic moiety stabilizing incorporated MPs. Lipopeptides have not been used for direct solubilization of MPs, seemingly because synthesis is expensive (Privé 2009). Thus, an initial solubilization step in cheaper detergents is commonly used.

Expensive chemical synthesis is obviously problematic for practical purposes. If long peptides are required, it will at some point make more sense to switch to biological synthesis, which is even less error prone than chemical synthesis. The following section describes the use of amphipathic proteins that can be expressed in *E. coli* for handling MPs.

1.2.6 *Salipro*

Saposin-like proteins are a family of lipid-interacting proteins conserved throughout eukaryotic species (Bruhn 2005). Saposin A (SapA), a member of this family, contains three disulfides, which stabilize the overall helix-turn-helix topology (Popovic et al. 2012). At neutral pH, SapA is a soluble and compact protein with a hydrophobic core (Figure 9A), but at pH 4.8, the protein is only soluble in the presence of lipids, likely because it partially unfolds and expose hydrophobic residues (Figure 9B) (Popovic et al. 2012; Chien et al. 2017). A crystal structure of two SapA chains in complex with 40 lauryldimethylamine N-oxide detergent molecules has been solved with all acyl chains resolved (Popovic et al. 2012).

SapA is capable of making soluble lipoprotein complexes (termed Salipro) down to sizes of about 3.2 nm when mixed with liposomes (Figure 9B) (Popovic et al. 2012). When reconstituting Salipro particles from mixed micelles of lipids and detergent, followed by detergent removal, increasing the lipid:SapA ratio leads to larger, but also somewhat polydisperse particles (Chien et al. 2017). The stability, size and stoichiometry of Salipro are pH-dependent, producing small but stable assemblies at physiological relevant pH 4.8, but larger and more unstable assemblies at neutral pH (Li et al. 2016).

MPs can be incorporated in Salipro using detergent-assisted self-assembly (Figure 9C) (Frauenfeld et al. 2016). Like the peptides discussed previously (Section 1.2.5.1), Salipro can adapt in size to the incorporated MP, using the MP and a layer or two of lipids around as scaffold to make a compact assembly (Frauenfeld et al. 2016; Flayhan et al. 2018). A drawback lies in the several cysteins of SapA, which make production in *E. coli* and purification more complicated than for other amphipathic proteins.

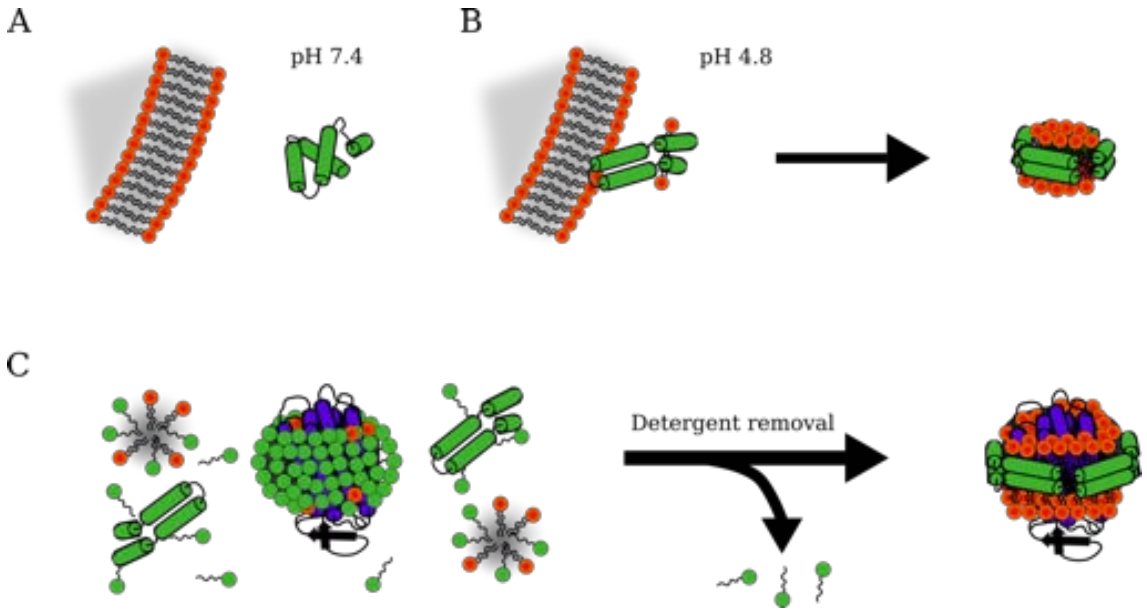


Figure 9 Saposin A lipid particles (Salipro). A: SapA is a compact soluble protein at standard physiological pH and does not form lipid particles. **B:** At pH 4.8, the physiologically relevant pH of lysosomes, SapA is thought to partially unfold, interact with lipid bilayers and form small lipid-protein particles. **C:** In presence of detergent, SapA unfolds to the open state. Adding detergent-lipid mixed micelles to SapA, followed by detergent removal, allow SapA to incorporate lipids to form larger lipid-protein particles.

1.2.7 *Nanodiscs*

Apo-A1 is the major constituent in high-density lipoproteins in the human reverse cholesterol transport system (Davidson and Thompson 2007; Rothblat and Phillips 2010), and has been shown to form discoidal particles with lipids in solution (Nakano et al. 2009; Midtgaard et al. 2015). Nanodiscs are discoidal particles made with amphiphilic membrane scaffold proteins (MSPs) that are derivatives of Apo-A1 to compromise only its lipid-binding helical-repeat domain (Bayburt et al. 2002). The nanodisc field, which compromises an overwhelming number of structural and functional studies, has been reviewed on several occasions recently (see *e.g.* (Ilia G. Denisov and Sligar 2016; Ilia G. Denisov and Sligar 2017; Rouck et al. 2017). MSPs are typically expressed in high yield in *E. coli* with reported levels of about 250 mg per liter bacterial culture (Bayburt et al. 2002) and in general have good solubility in aqueous solutions. The crystal structure of a truncated Apo-A1 shows that MSP alone likely folds on itself or in dimers in solution to shield hydrophobic regions (Mei and Atkinson 2011). As mentioned previously (section 1.2.4.2 and section 1.2.5.1), nanodisc is now a widely used term for the disc-like particles obtained with MSP, peptides, or SMA, but here it will be used solely for those prepared with MSPs.

1.2.7.1 *Nanodisc assembly*

Nanodiscs are generated from detergent-lipid mixed micelles and MSP, forming discoidal particles consisting of a central lipid bilayer patch stabilized on the rim by the MSP (Bayburt et al. 2002; Ritchie et al. 2009). A MP can be readily mixed into this reconstitution mixture and thereby led into the final nanodisc

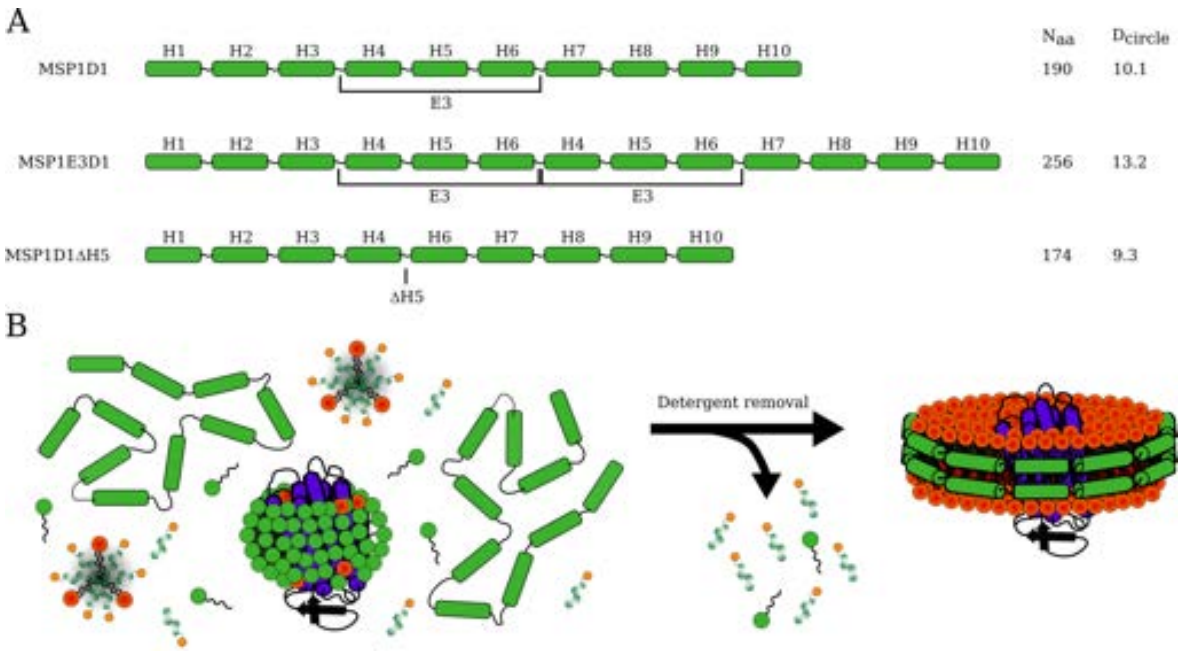


Figure 10 Membrane scaffold proteins and nanodiscs. **A:** Schematic representations of some common MSPs, showing the number of amphipathic α-helices in each construct. For each construct, the number of amino acids (N_{aa}) as well as the diameter of a circle (D_{circle}) with the theoretic circumference of the specific MSP are given. **B:** MP reconstitution into nanodiscs is achieved by mixing detergent solubilized MP, mixed micelles (commonly cholate as the detergent component) and MSP. Upon detergent removal, nanodiscs are formed by self-assembly.

(Figure 10B) (Ritchie et al. 2009; Bayburt and Sligar 2010). The MP reconstitution is often done in excess of MSP and lipids, whereby both empty and loaded nanodiscs are produced. Thus, the MP must either contain an affinity tag so that it can be isolated by affinity chromatography or be sufficiently large that loaded nanodiscs can be separated from empty nanodiscs by SEC (Ritchie et al. 2009). The standard protocol for nanodisc reconstitution is compatible with a wide variety of MPs, but the specific thermodynamics remain poorly understood (Ilia G. Denisov and Sligar 2017). Thus, reconstitution in nanodisc is empirical and often requires tuning of experimental parameters such as the temperature, the lipid:MSP stoichiometry and the choice of detergent and lipid types. Gaining more general knowledge about the nanodisc assembly process will be crucial to optimize the incorporation of MPs in nanodiscs. Such efforts were carried out as part of this thesis and are described in paper I. The main conclusions from this work were that (i) not all detergents, and especially not the popular DDM, were suited for making nanodiscs, unless in combination with cholate, (ii) that it was best to add a little excess amount of lipid to avoid lipid-poor nanodiscs, and (iii) that nanodisc self-assembly is fast and likely starts from pre-aggregates that have more or less the the stoichiometry of MSP to lipid found in the formed nanodisc after detergent removal. With that knowledge, it can be further generalized that MPs should (i) be compatible with the presence of cholate, and (ii) be equilibrated sufficiently long time in the reconstitution mixture to mix into the MSP-lipid-detergent pre-aggregates before removing detergent.

Finally, it is noted that MSPs are not able to solubilize lipid bilayers (Scheidelaar et al. 2015), but with the addition of adequate amounts of detergent, they have even been used for extracting MPs from their native membranes into nanodiscs (Marty et al. 2013; Yeh et al. 2018). Furthermore, nanodiscs have proven well-suited, and better than other common systems, for cell-free expression of MPs without detergent present (Lyukmanova et al. 2012).

1.2.7.2 *Nanodisc size*

The most utilized MSP for generation of nanodiscs has been the so-called MSP1D1. MSP1D1 was directly derived from human Apo-A1 by truncation of the 110 amino acid N-terminal domain as well as 11 extra residues from the first helix shown not to be involved in lipid binding (Figure 10A) (I. G. Denisov et al. 2004). MSP1D1-based nanodiscs have a slightly elliptical cross-section of roughly 10 nm in diameter (Skar-Gislinge et al. 2010). Another popular MSP was generated from insertion of three extra helices inside MSP1D1, yielding the construct MSP1E3D1, which lead to nanodiscs of roughly 12 nm in diameter (Figure 10A) (I. G. Denisov et al. 2004). Truncated MSP1D1, for instance MSP1D1ΔH5 (Figure 10A), yields nanodiscs of roughly 8 nm in diameter, which are compatible with solution NMR (sNMR) (Hagn et al. 2013). Because of this, MSP1D1ΔH5 nanodiscs facilitated the first high-resolution structure of a nanodisc as determined in a recent NMR-based study (Bibow et al. 2017). This structure showed that each nanodisc contained two MSPs oriented as anti-parallel belts surrounding a lipid bilayer and laying orthogonal to the bilayer normal. Furthermore, the two copies of the MSPs were stabilized by specific ionic interactions between side chains (Bibow et al. 2017).

It turns out that the nanodisc circumference is directly proportional to the length of the MSP (I. G. Denisov et al. 2004; Grinkova et al. 2010). Thus, by tuning the length of the MSP, the space available in the nanodisc can be controlled for incorporation of MPs in a controlled oligomeric state (Nasr et al. 2017). Another consequence is that nanodiscs have a maximum loading capacity for lipids (I. G. Denisov et al. 2004). This effect was studied systematically for MSP1D1 nanodiscs in paper I, finding an optimal POPC:MSP stoichiometry of approximately 70:1; starting from a stoichiometry well below this resulted in lipid-poor and more elliptical nanodiscs, whereas starting well above it resulted in the formation of larger undefined particles. Thus, the lipid:MSP stoichiometry is often optimized to ensure that nanodisc are made at this maximum capacity (Bayburt et al. 2002; I. G. Denisov et al. 2004; Grinkova et al. 2010).

1.2.7.3 *Stability*

MSPs are of well-defined composition and interact with lipids in a highly cooperative manner between the many helical-repeats (Ilia G. Denisov and Sligar 2017). Compared to lipid-particles made with short peptides or polymers, nanodiscs are more dynamically restrained, and the MSP belts are unlikely to exchange

over a period of days to weeks (Ilia G. Denisov and Sligar 2017). Thus, nanodiscs are capable of retaining a very well-defined and narrow size distribution. Differential scanning calorimetry (DSC) studies have shown that the stability of lipids in the bilayer depends on nanodisc size, explained by the presence of destabilized rim-lipids in near vicinity of the MSP and more bulk-like lipids in the core of the bilayer (Ilia G. Denisov et al. 2005). An interesting observation is that lipid exchange dynamics are higher in nanodiscs compared to vesicles (Nakano et al. 2009). Notably, though, the lipid exchange rate in nanodisc is still substantially slower than for SMALP and peptide discs (Miyazaki et al. 2010; Cuevas Arenas et al. 2017; Domínguez Pardo et al. 2018). Taken together, the nanodisc is overall less dynamic than most other soluble carrier systems and likely resembles the lipid bilayer in vesicles better. It has even been suggested that the flat bilayer in a nanodisc might be more native-like than the lipid bilayer in a vesicle that have has curvature due to the overall spherical structure of the vesicle (Mörs et al. 2013).

The narrow size distribution of nanodiscs is only valid up to $\sim 30^{\circ}$C. At higher temperatures, nanodiscs evolve into several and larger species over time (Ilia G. Denisov et al. 2005). DSC experiments on DMPC nanodiscs yields a broad transition, but after being heated to high temperature, the transition is sharp and reminiscent to that of pure DMPC bilayers (Shaw et al. 2004). This suggest that the nanodisc structure might be a kinetically stable structure, but not the thermodynamically most stable structure. Consequently, thermodynamic analysis as well as structural and functional studies at physiologically relevant temperatures can be complicated. Within the last two years, optimized MSPs were developed to increase the thermal resistence of nanodiscs as well as gaining better control of MP reconstitution: A major upgrade resulted from covalent linkage of the MSP's N- and C-termini by an enzymatic reaction with the transpeptidase sortase A (Nasr et al. 2017). Covalently linked MSPs were termed circularized MSPs (cMSPs) and nanodiscs made with cMSPs were in turn termed circularized nanodiscs (cND) (Nasr et al. 2017), not to be confused with the close to circular shape of a nanodisc. cNDs have markedly improved stability and remain well-defined in size at physicological temperature (Yusuf et al. 2018)- Furthermore, they allow for tighter control of the oligomerization states of inserted MPs (Nasr et al. 2017). Circularization also facilitated the development of a 50 nm wide nanodisc (Nasr et al. 2017), which until then had been hindered by inherent low stability and collapse into spherical aggregates of nanodiscs made with standard MSPs (Ilia G. Denisov and Sligar 2017). Circularization together with genetic engineering of MSPs were explored to obtain stable and more soluble nanodiscs in paper III.

1.3 COMPARISON OF PREVALENT CARRIER SYSTEMS

During the introduction of carrier systems in the previous section, specific advantages and drawbacks were mentioned. In Table 2, these are presented collectively to allow for easier comparison of the different

systems. As this thesis mainly involved the use of nanodiscs and detergent micelles for MP handling, the following sections will have these systems as reference and main focus for comparisons.

System	Advantages	Drawbacks	Comments
Detergent micelles	Small molecules: - Relatively simple synthesis - Feasible for deuteration Solubilize membranes Used in MP crystallization Dynamic: Adaptable in size	Free micelles Dynamic: Structurally polydisperse Sensitive to environment Generally deactivating or denaturing	It is difficult to consider all detergents under one header. The properties of detergents are MP specific.
APol A8-35	Inexpenssive synthesis Comercially available Deuterated versions comercially available High thermal stability Adaptable in size	Polymerization and random grafting in synthesis: - Polydispersity - Batch variation Detergent in sample preparation Likely restrain MP dynamics Incompatible with low pH and divalent metal ions	Incompatibilities can be fixed by changing the nature of side groups, as for sulfonated APols or nonionic glucose-based APols (J.-L. Popot et al. 2011).
SMALP	Inexpenssive synthesis Comercially available High thermal stability Detergent-free sample preparation Native lipids and complex partners Dynamic: Adaptable in size	Polymerization synthesis: - Polydispersity - Batch variation Dynamic: Polydisperse Incompatible with low pH and divalent metal ions Some interference with UV absorption	Incompatibilities and UV interference can be fixed by changing chemical nature of side groups, as for SMILP (Hall et al. 2018) and DIBMALP (Oluwole et al. 2017).
Peptide disc	Well-defined synthesis: pure product Lipid bilayer Dynamic: Adaptable in size	Dynamic: - Polydisperse - Unstable over time Detergent in sample preparation Interference with UV absorption	Longer peptides reduce dynamics (A. N. Larsen et al. 2016; Kondo et al. 2016).
NSP_r (peptidisc)	Well-defined synthesis: pure product Adapts to size of MP	Detergent in sample preparation Interference with UV	
Lipopeptides	Adapts to size of MP Acyl-chains ordered towards the MP	Expensive synthesis Detergent in sample preparation Interference with UV	
Salipro Stabilize MP with thin layer of lipids	Adaptable to MP size Six Cys in SapA: Requires control with reducing agents Detergent in sample preparation Interference with UV absorption		
Nanodisc (MSP-based)	High expression of most MSPs in *E. coli* Well-defined size: Tunable by MSP length Lipid bilayer Control of MP oligomerization state	Detergent in sample preparation Low expression of long MSPs in *E. coli* Unstable at physiologically relevant temperature Interference with UV absorption	Circularization of MSP increases nanodisc stability and enables studies at physicological temperatures (Nasr et al. 2017; Yusuf et al. 2018). Circularized nanodisc gives even better control of MP oligomeric state (Nasr et al. 2017).
Bicelles	Simple components Tuneable size Used for MP crystallization	Narrow phase-diagram: - Very sensitive to environmental effects Dynamic: Polydisperse Limited number of compatible lipids and detergents	Low-q ($q < 1$) bicelles behave as mixed micelles (Beaugrand et al. 2014).
Vesicles	True lipid bilayer Compartmentalization: - Chemical gradients possible	Very big (pseudo-phase) Low MP loading	The drawbacks mainly relate to structural studies of MPs.

Table 2 Advantages and drawbacks of MP carrier systems. Some parameters both give rise to advantages and drawbacks, for instance fast dynamics, which causes some systems to be widely adaptable towards MPs of different size, but at the same time causes inherent polydispersity.

1.3.1 *Lipids or no lipids?*

A notable property of lipid bilayers is their extreme lateral pressure profile, varying with up to hundreds of atmospheres over the approximately 5 nm wide bilayer (Cantor 1997). This property may not be different in micelles due to the fundamentally different packing of detergents. This is among other reasons why lipid bilayer structures as vesicles are commonly seen to represent ideal model mimics for biological membranes, whereas detergent micelles are commonly seen to represent the other limit on a scale of biological relevance.

Howecver, micelles provide the most simple and versatile sample handling, where vesicles have limited practical applications. The remaining soluble MP carriers have in general been designed to be more stabilizing for MPs than micelles, while keeping the versatility of micelles that is not available with vesicles. As exemplified *e.g.* by APols, a lipid bilayer is not an absolute necessity for stabilization of MPs, (Jean-Luc Popot and Engelman 2016) but many biochemical processes are dependent on protein-membrane interactions, (Lemmon 2008) which can not be studied in lipid-free systems. Overall, lipid bilayer systems are desired to ensure an environment that to the best extend possible mimics that of biological membranes. The lipid-free systems are rather explored due to application specific advantages, including that they are inherently less complex and make small complexes with MPs.

1.3.2 *Considerations for sample preparation*

The ease of sample preparation is an important parameter to consider. All systems but SMALP and derivatives make use of detergents to isolate MPs. This is probably one of the a main reasons why standard detergents (Section 1.2.1.1) remain so popular despite their well-established denaturing effect on MPs. Detergent free sample preparation is desired, and thus, SMALPs excel in this category when compared to other systems (Dörr et al. 2014). As a recent example, a MP in a native trimeric state was extracted from native membranes into SMALP and solved to 8.2 Å resolution by cryoEM (Parmar et al. 2018). Thus, the ease of sample preparation with SMALP can be similar to MP extraction by detergents (Postis et al. 2015). However, it is often desired to be able to control the lipid composition to test the specific influence of different types of lipids on MP structure or activity with reproducible results. In these cases, native lipids must be removed by *e.g.* detergents and samples afterwards replenished with synthetically pure lipids.

Ease of sample preparation can also be assessed from the carrier system's adaptability towards MPs of different size. Again, a detergent micelle is a highly flexible carrier that can adapt to any MP (Garavito and Ferguson-Miller 2001). Other small and dynamic proteins (Salipro) (Frauenfeld et al. 2016), polymers (J.-L. Popot et al. 2011; Dörr et al. 2016), and peptides (Carlson et al. 2018) are also generally adaptable towards MPs of different size. Only the nanodisc, where the relatively long MSP determines the overall

well-defined size, is an exception 1.2.7.2. This implies that there are more considerations to be made about optimal MSP length, lipid stoichiometry and MP stoichiometry in reconstitution of nanodiscs. As such, nanodisc reconstitution is commonly rated more difficult than for other systems.

Ease of sample preparation is of course important, but in the end, the formed MP-loaded carrier system must be compatible with a given experimental technique. In the following sections, the advantages and drawbacks are compared in relation to functional and structural studies of MPs.

1.3.3 *Functional studies*

The major advantage of small carrier systems is that they allow for investigating MPs by the same methods as used for soluble proteins. With its well-defined size, the nanodisc is especially advantageous for biophysical studies because the MP oligomerization state can be controlled, which allows for quantitative and reproducible measurements (Ilia G. Denisov and Sligar 2017). The general disadvantage of the small carrier systems is that they do not provide compartmentalization as in vesicles, which is important to study the vectorial function of many MPs (Alberts et al. 2002).

Another parameter to consider is the dynamics of the MP in the carrier system. In the case of the stable β-barrel MP OmpX, its inherent flexibility was found to be dramatically reduced in DPC micelles compared to both nanodiscs and bicelles made with DMPC lipids (Frey et al. 2017). Furthermore, molecular dynamics simulations have indicated that APols also restrain dynamics in OmpX compared to both DHPC micelles and DOPC bilayers (Perlmutter et al. 2014). Restrained dynamics are expected to reduce activity of MPs, and thus, carrier systems exhibiting these effects should be avoided for functional studies.

Obviously, a carrier system should be compatible with the desired experimental conditions: The native conditions for many MPs involve slightly acidic pH, such as in lysosomes, (Winchester 2001) or presence of Mg^{2+}, which is the most abundant ion in the organism (Jahnen-Dechent and Ketteler 2012). Thus, many amphiphilic polymers should be used with caution due to incompatibilities with these conditions (Section 1.2.4). Finally, the carrier system should ideally be compatible with physiologically relevant temperatures. In the case of nanodiscs, circularization should be considered to retain the well-defined size at this temperature, as discussed in Paper III of this thesis. Many of the dynamic carrier systems, especially peptide discs (Section 1.2.5.1) and some bicelles (Section 1.2.3), are also unstable at elevated temperature and should be used with caution.

In terms of concentration, the sample requirements for functional studies depend on the specific method. Activity assays can typically be carried out at quite low concentrations, whereas the concentration needed for binding studies depend on the strength of interaction. Overall, the requirements are typically lower than

for structural studies, though. This is one of the major reasons why functional studies are more abundant than structural studies, which is *e.g.* the case in the nanodisc field (Ilia G. Denisov and Sligar 2017).

1.3.4 *Structural studies*

X-ray crystallography remains the standard method for structural investigation with atomic scale resolution and is not limited by molecular size. The PDB is dominated by structures solved by this method, but many MPs are difficult to crystallize. Most of the presented carrier systems are in themselves quite dynamic particles and do not aid in crystallization. This necessitates other structural methods, including the high- to medium-resolution methods of NMR and cryoEM, and low-resolution methods, such as small-angle X-ray scattering (SAXS) and SANS. A powerful approach is to combine the synergistic advantages of all methods to describe full-length models in better detail than what could be achieved with the methods individually. MPs in carrier systems are inherently complex systems, especially those with lipids present, and fall into the category of samples relevant for such an integrative structural biology approach (Ward et al. 2013; Sali et al. 2015).

This section provides a brief overview of the main structural methods used to characterize MPs in solution, *i.e.* NMR (Section 1.3.4.1), cryoEM (section 1.3.4.2) and SAS (section 1.3.4.3), with focus on sample requirements. Structural investigations of MPs by SAS was the one of main focuses of this thesis: This section provides specific examples of such studies, whereas a more detailed description of the method is given in Chapter 2.

1.3.4.1 *NMR*

NMR data provide both structural and dynamic information at atomic resolution (Kwan et al. 2011). The field of NMR is divided in two regimes, namely solution NMR (sNMR) and solid-state NMR (ssNMR). These measure the same fundamental probes, namely magnetically active nuclei, but requirements for samples and equipment are somewhat different as explained in this section. A closer inspection of sample conditions used in sNMR and ssNMR reveal that it is a massive challenge to study MPs (Table 3).

The spectral resolution of sNMR depends on the molecular weight of the particle under investigation. High mass causes slow molecular tumbling, which in turn leads to peak broadening. Small carrier systems have been used for sNMR studies of MPs. These include amphipols (Planchard et al. 2014), Salipro (Chien et al. 2017), micelles (mostly DPC (Chipot et al. 2018)), low q bicelles (Dürr et al. 2012), and small nanodiscs (Hagn et al. 2013). Studies of nanodiscs illustrate the current size limitations of sNMR, as high quality data could only be obtained for the truncated, small versions, such as MSP1D1ΔH5 nanodisc (Hagn et al. 2013; Bibow et al. 2017). These small nanodiscs have gained much popularity for sNMR studies (Puthen-

veetil et al. 2017). To obtain good signal-to-noise, high μM sample concentrations are typically needed (Table 3), which requires the concentration of samples with a risk of forming heterogenous larger particles. Circularized nanodiscs are more conformationally stable than standard nanodiscs (Yusuf et al. 2018) and might aid NMR studies at such high concentrations and elevated temperatures. Indeed, well-resolved spectra were obtained for a MP inside circularized MSP1D1 nanodiscs (Nasr et al. 2017). The success of cNDs highlights that the properties of the carrier system are important to obtain optimal data. Furthermore, sNMR is often carried out at low pH to suppress hydrogen exchange and at elevated temperature (Table 3) to enhance tumbling. Furthermore, ^{15}N, ^{13}C and deuterium-labeling is necessary for structure determination of large molecular complexes, including MPs in nanodiscs (Hagn et al. 2013). Thus, carrier systems should preferably be inexpensive to synthesize and readily deuterated, and at the same time compatible with low pH and physiological temperatures.

Technique	**Protein**	**MW, kDa**	**Carrier**	**Lipid**	**[P], μM**	**T, °C**	**Notes**	**Ref.**
sNMR	OmpX	16.4	ΔH5 ND	DMPC DMPG (3:1)	1000	45	U-^{2}H, U-^{13}C and U-^{15}N, d-lipids	Hagn et al. 2013
sNMR	VDAC1	30.7	1D1 cND	DMPC DMPG (3:1)	500	45	U-^{2}H, U-^{15}N d-lipids	Nasr et al. 2017
ssNMR	fd coat	5.2	SMA macrodisc	DMPC	230	40	U-^{15}N.	Radoicic et al. 2018
ssNMR	CXCR1	35.1	Liposome	DMPC	2000	25	U-^{15}N, U-^{13}C, 11 kHz MAS	Park et al. 2012

Table 3: Examples of MP samples studied by NMR techniques. Abbreviations: solution NMR (sNMR), ssNMR (solid-state NMR), molecular weight (MW), MSP1D1ΔH5 nanodisc (ΔH5 ND), circularized MSP1D1 nanodisc (1D1 cND), 1,2-dimyristoyl-sn-glycero-3-phosphorylglycerol (DMPG), protein concentration (**[P]**), uniform labeling (U-), deuterated lipids (d-lipids).

For ssNMR, sample requirements are somewhat different from sNMR. In particular, large carrier systems containing lipid bilayers are desired as these can be aligned in the magnetic field (Hansen et al. 2015). This has been explored for large peptide discs (macrodiscs) and bicelles (Park et al. 2011) and lately SMALPs (Table 3) (Radoicic et al. 2018). It should in principle also be possible for large nanodiscs (Nasr et al. 2017), which however appear to be difficult to produce at the required concentrations for ssNMR. Non-aligned samples can also be utilized for ssNMR, such as MPs in liposomes. Typically, these samples are pelleted by centrifugation into a highly specialized sample holder and rotated at tens to hundreds of kHz at

the magic angle relative to the magnetic field (54.7°) to simulate molecular tumbling and thereby achieve better resolved peaks (Polenova et al. 2015). ssNMR is in general still a technique undergoing massive developments in hardware and requires a high level of expertise. This means that structural studies of MPs are few, but in theory, the technique has no apparent size limitation and can thus complement sNMR on large systems (Hansen et al. 2015).

Overall, sNMR is still limited to relatively small MPs and require high sample concentration, whereas ssNMR requires even higher sample concentrations and highly specialized equipment. Furthermore, deuterium-labeling of MPs will at some point become a bottleneck as not all MPs can be produced in simple organism that can survive in deuterated media (Dunne et al. 2017). Thus, large structures are rather investigated by other structural methods, as explained in the following sections.

1.3.4.2 *CryoEM*

The number of cryoEM structures in the PDB is rising exponentially. CryoEM makes it possible to solve structures of non-crystallizable samples to near-atomic resolution, and it is expected to set new standards for the future of structural biology (Wyhe 2016). It is therefore not surprising that most of the carrier systems have been exploited for obtaining MP structures by cryoEM recently, and with great success (Mio and Sato 2018). These include APols (S. Zhang et al. 2017), SMALPs (Parmar et al. 2018), Salipro (Frauenfeld et al. 2016), lipopeptides (Tao et al. 2013), and nanodiscs (Table 4). Detergents are also still popular for cryoEM, with low CMC detergents such as DDM being the most popular choice (Thonghin et al. 2018).

CryoEM relies on heavy image treatment to align particles in different angles and directions, and in the end refinement into a 3D model (Thonghin et al. 2018). Thus, it is beneficial to have well-defined (and large) particles that can be distinguished. The carrier system can have a positive influence on this, with nanodiscs as an example providing a relatively well defined rim around the membrane protein with good contrast (Roh et al. 2018). As opposed to NMR, sample volume and concentration requirements are typically low in cryoEM. As an example, only 3μl with a concentration of approximately 1μM of the Yeast Vacuolar ATPase V_o Proton Channel inside MSP1E3D1 nanodisc was required to obtain adequate grids (Table 4, PDB ID: 6C6L) (Roh et al. 2018). However, if the bilayer of the nanodisc is large compared to the MP, the MP can move around, which consequently makes determination of high-resolution structures difficult (Matthies et al. 2016). To elaborate on this, it is striking that deposited cryoEM structures of MPs in nanodiscs originate from samples where the MP has similar or slightly smaller diameter of its TMD to that of the nanodisc (Table 4). Likewise, size-adaptable carrier systems are promising for cryoEM. Salipro is an example, where only a layer or two of lipids are preserved around the MP, providing a lipid interior, but a compact carrier (Frauenfeld et al. 2016).

PDB ID	MP	MW, kDa	MSP	d_{MP}, Å	d_{ND}, Å	Res., Å	Ref.
6MGV	OSCA1.2	178	MSP2N2	10.5	14,5	3.1	Saotome et al. 2017
6C6L	ATPase V_o	353	MSP1E3D1	11.5	11	3.5	Roh et al. 2018
6EBK	K_V channel	388	MSP1E3D1	9.5	11	3.3	Matthies et al. 2018
6CP6	ATP synthase	573	MSP1E3D1	9.8	11	3.6	Yoshikawa et al. 2016
6FFC	ABC transporter	148	MSP1D1	5.5	8	3.6	S. M. Jackson et al. 2018
6EZN	OST complex	292	MSP1E3D1	8	11	3.3	Wild et al. 2018

Table 4: Examples of MP structures inside nanodiscs solved by cryoEM. Abbreviations: Protein Data Bank (PDB), molecular weight (MW), membrane scaffold protein (MSP), approximate diameter of membrane protein (d_{MP}), approximate diameter of nanodisc (d_{ND}), resolution (Res.).

For near-atomic resolution, cryoEM is still limited to relatively large proteins around 100 kDa or larger (Mio and Sato 2018). For MPs in nanodiscs, cryoEM structures deposited this year are all well-above this limit (Table 4). Furthermore, all of these structures are assymetric with relatively large extra- or intra-cellular domains, which also facilitates particle alignment. It is important to be aware that the carrier system will also affect the effective contrast and make it difficult to discern between carrier system and MP in certain regions (Thonghin et al. 2018). A comparative study showed that optimal resolution was achieved with lauryldimethylamine-n-oxide (LDAO), due to its small micelle and density close to amorphous water, providing optimal contrast (Schulz et al. 2017). Although some carrier systems can provide better contrast in cryoEM, they may at the same time compromise MP activity. LDAO, for instance, is not expected to stabilize MPs very well in general (Carlson et al. 2018). It is obviously paramount to have a functionally relevant and active MP, and thus, activity and stability should always be confirmed. It can be discussed whether the cryo-cooling conditions allows for classifying cryoEM as a solution-based structural method. In any case, it will often be desired to complement cryoEM data with structural data obtained from samples in solution at ambient temperature, which can *e.g.* be done with SAS as described in the following section.

1.3.4.3 *SAS*

SAXS and SANS are low-resolution structural methods suitable for investigating all types of soluble molecules (D. I. Svergun and Koch 2003). As opposed to NMR and cryoEM, these techniques have no apparent size limitations, but suffer from orientational averaging and low information content in general (D. I. Svergun and Koch 2003). SAS allows for determining structures of molecules and molecular assemblies under

solution conditions and in fairly low concentrations, and the methods have been widely applied to describe the structures of detergent micelles (Oliver et al. 2013), APols (Gohon et al. 2006), bicelles (Luchette et al. 2001), SMALP (Jamshad et al. 2015; Hall et al. 2018), nanodiscs (Skar-Gislinge et al. 2010; Huda et al. 2015), peptide discs (Midtgaard et al. 2014; A. N. Larsen et al. 2016), Salipro (Flayhan et al. 2018), and lipid bilayers (Kučerka et al. 2011). SAS is thus a highly relevant method for structural verification of these carrier systems in general. Chapter 2 provides more specifics on these methods, while this section describes them in relation to structural investigations of MP in carrier systems.

In SAS, it is important to be aware that everything in the sample contribute to the scattering signal in biological applications. The scattering intensity is proportional to the squared volume of a particle (see section 2.2.1.3), and thus, to refine a structural model of single particles, a monodisperse sample with only one structural species is desired. Polydisperse or structurally disperse samples require the introduction of more fitting parameters to already relative complex models given the information content in the data. Thus, the inherent polydispersity of the dynamic carrier systems, including peptide discs, SMALPs and detergent micelles, poses a general problem for SAS studies. For this reason, the well-defined nanodisc is a much better choice, and circularized nanodiscs are particularly interesting due to their superior homogeneity over standard nanodiscs (Nasr et al. 2017; Miehling et al. 2018).

As mentioned above, free detergent micelles will contribute to the scattering signal in SAS, which in general makes it nontrivial to separate the relevant signal from the MP-loaded micelle. However, it has been shown that by using online SEC coupled to SAXS (SEC-SAXS), subtraction of the scattering intensity from free DDM micelles was possible and allowed for extracting the relevant scattering intensity from MPloaded DDM micelles alone (Berthaud et al. 2012). From these data, the DDM torus around aquaporin-0 could be modelled (Berthaud et al. 2012). It is noteworthy, though, that this approach asummes the amount and size of free micelles to be constant throughout column elution, although MPs are know to alter the dynamics of micelles (Garavito and Ferguson-Miller 2001). Thus, detergent micelles should at least be used with caution, but generally, less polydisperse systems are desired. Recently, different MPs were measured in Salipro using SAXS, which yielded small and relatively homogeneous particles that allowed for *ab initio* modeling of the entire MP-carrier complex (Flayhan et al. 2018). Salipro thus has some clear advantages over detergent micelles for SAXS studies, but a more detailed understanding of the system, in particular its dynamics, is at the same time lacking.

Even with Salipro, the MP and carrier system might be to complex that a single SAXS measurement provides enough information to separate the signal from the MP. However, more information can be obtained by performing contrast variation (See Section 2.2.2) (Zaccai et al. 2016). One approach is to simply combine SAXS and SANS measurements, where protein and lipids or detergent have inherently different contrasts. Bacteriorhodopsin inside a nanodisc was modeled from combined SAXS and SANS using a hy-

brid approach of an analytical description of the nanodisc (Section 2.3.3) combined with a bead model of the MP inside (Kynde et al. 2014). The protein was fitted to be slightly tilted and displaced from the center of the membrane. Notably, though, more than one solution fitted the data, indicating the need for even more data and/or a simpler system.

A simpler system can be achieved by masking the signal of specific components by doing SANS measurements at their respective contrast-match point (CMP) (Section 2.2.2). The CMP can further be manipulated by selective deuteration to eliminate signal from different components at the same time. Many examples of such approaches have been reviewed (Breyton et al. 2013) and a few newer cases are provided here and summarized in (Table 5). Using commercially available deuterated OG detergent and employing multiple contrasts, the solution structure of the β-barrel FhaC together with its surrounding micelles could be modelled (Gabel et al. 2014). Interestingly, this study used molecular modeling to generate an ensemble of structures to fit the SANS data, which included some structures with detergent-free patches on the MP surface (Gabel et al. 2014). It has later been suggested that given information about the number of detergents surrounding a MP and SANS measurements at two different contrast, *ab initio* modeling of MPs is feasible (Koutsioubas 2017). However, this remains a theoretical study. Some APols are readily deuterated (DAPols), allowing for contrast variation SANS measurements of trapped MPs. However, measurements of bacteriorhodopsin (BR) in DAPols A8-35 revealed the presence of some free DAPol micelles, and furthermore that BR was not solely in a monomeric state (Gohon et al. 2008). This highlights that while contrast variation can provide more information, the sample still needs to be structurally monodisperse to allow trivial data analysis.

When the structure of the MP is the sole focus, the most simple system can be achieved by manipulating the contrast of the carrier system by selective deuteration to completely eliminate its contribution to the scattering intensity (Section 2.2.2). The term CMP only describes elimination of the forward scattering intensity (intensity at zero angle $I(0) = 0$), but it is desired to uniformly contrast match the carrier system, *i.e.* throughout the entire scattering intensity profile ($I(q) = 0$). For lipids and detergents, which do not have a uniform scattering length density (SLD) (Section 2.2.2), this requires selective deuteration of head and tail groups, respectively. Stealth nanodiscs (sND) composed of selectively deuterated phospholipids (Maric et al. 2015) and deuterated MSP were proposed to fulfill this requirement (Maric et al. 2014). Notably, the sND was contrast-matched to D_2O, where the contrast of hydrogenated MPs is optimal and the incoherent background intensity from hydrogen is minimal (Maric et al. 2014). Recently, the structure of the lipid flippase MsbA was modeled in different conformational states to SANS data obtained with a similar sND system (Josts et al. 2018). These structural changes would have been difficult to observe had the nanodiscs not been uniformly contrast-matched.

Carrier	Deuteration	MP	$CMP_{carrier}$	Note for carrier	Ref.
Nanodisc	d-MSP1D1∗, d-PC∗	-	∼100		Maric et al. 2014
Nanodisc	d-MSP1D1∗, d-PC∗	MsbA	∼100	Same as (Maric et al. 2014)	Josts et al. 2018
OG	d17-OG†	FhaC	90‡		Gabel et al. 2014
OG	d-OG	LamB	103	Head: 52% D, Tail: 94% D	Midtgaard et al. 2018
DDM	d25-DDM†	-	49	Mix d-DDM h-DDM (43:57)	Oliver et al. 2017
DDM	d25-DDM†, d-MmIAP∗	MmIAP	49	Same as (Oliver et al. 2017)	Naing et al. 2018
DDM	d-DDM	BR, PSI, GluA2, SERCA	102	Head: 57% D, Tail: 89% D	Midtgaard et al. 2018
DDM	d-DDM	GluA2	102	Same as (Midtgaard et al. 2018)	A. H. Larsen et al. 2018
APol A8-35	d-octyl, d-isopropyl	BR	85		Gohon et al. 2008
APol A8-35	Perdeuteration	-	120.5		Giusti et al. 2014

Table 5: Overview of deuterated carriers used in SANS. ∗ Biological deuteration, † commercially available, ‡ not uniformly contrast-matched.

Following the idea of uniformly contrast-matched carriers (Maric et al. 2014), such an approach has also been developed for DDM micelles. Mixing of commercially available hydrogenated DDM and commercially available deuterated DDM (d25-DDM) were uniformly contrast-matched at 49% D_2O (Oliver et al. 2017). Notably, 49% D_2O is close to the CMP of most proteins (42%) and the incoherent background intensity from 51% H_2O is quite high, suggesting the need for deuterated MPs to obtain sufficient signal (Oliver et al. 2017) This was recently achieved for the aspartyl protease MmIAP to allow accurate determination of overall size together with *ab initio* modeling (Naing et al. 2018). However, as discussed for NMR (Section 1.3.4.1), it will at some point become unfeasible to deuterate MPs, which makes the above approach less useful than approaches, where the carrier system is contrast-matched to a D_2O content far from the CMP of MPs.

A more elegant solution was recently proposed, where DDM and OG detergents were selectively deuterated in head and tail-groups, respectively, to be uniformly contrast-matched to D_2O (Midtgaard et al. 2018). Thereby, hydrogenated MPs could be readily measured in these systems, allowing for direct modeling of MP structure, as demonstrated for a large set of different MPs (Midtgaard et al. 2018). The production of these stealth detergents has yet to be commercialized, but it is a clear advantage that MPs do not need to be deuterated. Stealth DDM (sDDM) was used in this thesis (Paper IV) and was obtained partly on a collaboration basis with the National Deuteration Facility at ANSTO, Australia.

Overall, deuteration strategies make SANS a highly relevant method for studying MPs under solution conditions, and systems containing bilayers could in principle be used to study MP-lipid interactions. When the scope is limited to modeling of MP structure, Salipro provide an intriguing example of making relatively homogenous particles with few lipids present that might be modeled using a single SAXS contrast.

1.3.5 *Outlook*

MP carrier systems have been developed for the main reason of stabilizing MPs in solution compared to conventional detergents. Most of these systems have proven useful for functional and structural studies of MPs which have led to an acceleration of the number of MP related studies. This chapter aimed to give an overview current advantages and drawbacks of the most popular carrier systems.

Nanodiscs and other lipid-bound carrier systems provide more native-like environments and backbone dynamics are observed to increase in these compared to detergent. On the other hand, detergent micelles can be small and facilitate NMR and cryoEM studies. Thus, there will often be a trade-off between resolution and native dynamics. NMR on MPs require labeled and highly concentrated samples, while cryoEM is still limited to relatively large and asymmetric structures and do not resolve flexible domains. Given the fact that a MP in a carrier system is a multi-component system, SAS and contrast variation is well-suited for structural characterization at low resolution. Integrated analysis from a combination of all methods might provide a further optimized basis for structural refinement: sNMR can provide dynamics and structure determination at atomic resolution for small MPs or truncated MPs, ssNMR can provide dynamics and structure determination of larger MPs together with bilayer interactions, cryoEM can provide structure determination of large MPs, excluding disordered regions, at near-atomic resolution, and SAS can provide overall shapes and sizes of entire MP-carrier complexes. Both SANS and NMR benefit from deuteration, and it is therefore expected that the availability of deuterated carrier systems can greatly advance structural studies of MP in solution.

Finally, nanodiscs have become very popular as carrier for structural analysis by sNMR, cryoEM and SAS as they provide a lipid bilayer environment in a very well-defined and stable particle. However, sample

preparation is expected to be more difficult than for other systems. In this thesis, the reconsitution process of nanodiscs was thoroughly investigated (Paper I) followed by a suggesting for optimization to achieve more stable and soluble nanodiscs (Paper III). Furthermore, nanodiscs were studied by combined SEC-SAXS and SEC-SANS, showing the feasibility of these novel methods for obtaining optimal, aggregation-free data (Paper II). Finally, uniformly contrast-matched carriers, sDDM micelles and sNDs, were applied to study the magnesium transport protein CorA using SANS (Paper IV).

2

SMALL-ANGLE X-RAY AND NEUTRON SCATTERING

2.1 INTRODUCTION

In this chapter, the experimental technique small-angle scattering (SAS) is introduced. SAS can be performed using X-rays (SAXS) or neutrons (SANS) and typically provides structural resolution in the range 10 Å to 1000 Å. SAXS and SANS rely on the same basic theory, but because X-rays interact with electrons and neutrons interact with nuclei, the methods are sensitive towards different structural probes and can be highly complimentary. Here, the basic theory and concepts are described, with the focus on applications to bio-molecules (bioSAXS and bioSANS) and considerations about sample handling. Finally, new experimental facilities and their opportunities for future SAS studies are touched upon.

2.1.1 *X-ray and neutron fundamentals*

2.1.1.1 *X-rays*

X-rays are electromagnetic radiation (photons), and those applied for structural studies typically have wavelengths in the Angstrom (1 Å = 10^{-10} m) region. In synchrotrons, X-rays are produced from electrons that are accelerated in magnetic devices (Als-Nielsen and McMorrow 2011), whereas at home sources, X-rays can be produced *e.g.* from microfocused sealed-tube sources (Krause et al. 2015) or liquid-metal-jet sources (Hemberg et al. 2003). The energy of a photon is given by

$$E = \frac{hc}{\lambda} \tag{2}$$

where h is Planck's constant, c is the speed of light and λ is the wavelength. X-rays with a wavelength of 1 Å , typical for bioSAXS experiments, thus have an energy of 12.4 keV (kilo electron volts). Due to the electromagnetic nature, X-rays might exert a force on an electric charge, and their interaction is therefore mainly with the electrons of atoms.

2.1.1.2 *Neutrons*

Neutrons are neutral particles located in the nuclei of atoms. At neutron scattering facilities, neutrons are produced either in a nuclear reactor or by a spallation source. In the reactor, neutrons are generated from a chain reaction of radioactive uranium, whereas the spallation source utilizes guided collision of high-energy protons with a heavy metal to generate neutrons (Fitter et al. 2006). Unlike photons, neutrons have a rest-mass, m. If the velocity, v, of a neutron is much less than the speed of light, c, all of its energy is kinetic and given by

$$E = \frac{1}{2}mv^2 \tag{3}$$

The relation between energy and wavelength is given by the De Broglie relationship

$$\lambda = \frac{h}{\sqrt{2Em}} \tag{4}$$

For both types of sources, the generated neutrons are fast with energies on the order of MeV and must be slowed down. For SANS on bio-molecules, a wavelength of $\sim$5 Å is typically used, corresponding to a neutron energy on the order of meV. Netrons are neutral particles and interact mainly with the nuclei of atoms.

2.2 SCATTERING

The interaction phenomena of scattering can be described by the same basic theory for X-rays and neutrons (Als-Nielsen and McMorrow 2011). To arrive at relatively simple expressions, a number of approximations are applied. The kinematical approximation states that scattering is weak, or in other words that the probability of scattering a photon or neutron from the incoming beam is low. This implies that multiple scattering can be neglected, which holds true for samples with high transmission. In practice, high transmission is achieved for sufficiently dilute samples or sufficiently thin samples. Furthermore, the far-field approximation applies when the source and detector are placed relatively far from the scattering volume such that the incoming and scattered beams can be described as plane waves.

2.2.1 *Scattering geometry*

To depict a scattering event, a plane wave of X-rays or neutrons with wavelength λ is described by a wave vector $\boldsymbol{k}$ with length $|\boldsymbol{k}| = 2\pi/\lambda$ (Figure 11). The wave interacts with a particle and a scattered wave, with

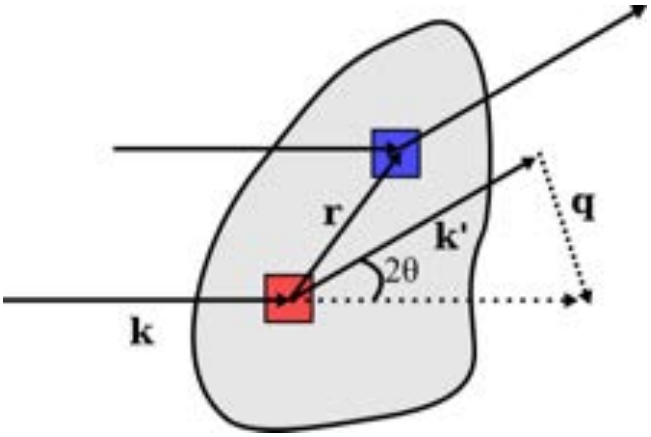


Figure 11 Small-angle scattering geometry. An incoming wave with wave vector $\boldsymbol{k}$ interacts with a particle and a scattered wave with wave vector $\boldsymbol{k'}$ is produced. The scattering vector $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k'}$ is related to the angle between the incoming and the scattered wave by $q = |\boldsymbol{q}| = 4\pi sin(\theta)/\lambda$. The scattering originating from two or more scattering centers in the particle separated by a vector $\boldsymbol{r}$ (red and blue squares) has a phase difference of $\boldsymbol{q} \cdot \boldsymbol{r}$.

wave vector $\boldsymbol{k'}$, is produced, with the scattering vector defined as $\boldsymbol{q} = (\boldsymbol{k} - \boldsymbol{k'})$. In elastic scattering (no exchange of energy), $|\boldsymbol{k}| = |\boldsymbol{k'}|$, and from the scattering triangle in Figure 11, it follows that

$$q = |\boldsymbol{q}| = 2|\boldsymbol{k}|sin\theta = \frac{4\pi sin\theta}{\lambda} \tag{5}$$

where θ is half the angle between the incoming and the scattered wave. It can further be shown that the scattering originating from two or more scattering centers separated by a distance, as given by the vector $\mathbf{r}$ (Figure 11), in a particle has a phase difference of $(\boldsymbol{k} - \boldsymbol{k'}) \cdot \boldsymbol{r} = \boldsymbol{q} \cdot \boldsymbol{r}$, which gives rise to an angle-dependent interference term (Als-Nielsen and McMorrow 2011).

2.2.1.1 *Scattering properties*

The strength of the interaction with the incoming beam is given by the scattering cross section of the sample. In an experiment, the intensity of the scattered beam, I_{scat}, is recorded by a detector (Figure 13). Normalizing with the flux of the incoming beam, Φ_0, and the size of the detector (which subtends a solid angle $\Delta\Omega$) gives the differential scattering cross section

$$\frac{d\sigma}{d\Omega} = \frac{I_{scat}}{\Phi_0 \Delta\Omega} \tag{6}$$

The differential scattering cross-section is thus a measure of the efficiency of a scattering process, *i.e.* the probability of scattering. The scattering potency of an atom is described by its scattering length, b. At zero angle, all scatterers (electrons or nuclei) scatter in phase and the differential scattering cross-section for each atom is related to the scattering length by

$$\frac{d\sigma}{d\Omega} = |b|^2 \tag{7}$$

	b, cm		
Atom	X-ray	$SANS_{coh}$	$SANS_{inc}$
H	2.82	-3.740	25.27
D	2.82	6.671	4.04
^{12}C	16.92	6.646	
^{14}N	19.74	9.360	
^{16}O	22.56	5.803	
^{31}P	42.30	5.130	
^{32}S	45.12	2.847	

Table 6 Scattering lengths of common elements in biology. b is the scattering length in units of cm, *coh* denotes the coherent scattering length, and *inc* denotes the incoherent scattering length.

For X-rays, the scattering length of an electron is given by the Thomson radius ($2.82 \cdot 10^{-13}$ cm), independently of the atom to which it belongs. Thus, the scattering length increases linearly with the number of electrons in an atom. For neutrons, there is no such simply dependency, and scattering lengths have been obtained empirically (As tabulated in (Sears 1992)). The neutron scattering lengths are typically divided into coherent and incoherent contributions. Coherent scattering contains all the structural information, while the incoherent scattering gives rise to a flat background. The scattering lengths of the most common atoms in biology are given in Table 6. Incoherent scattering lengths are only provided for H and D, which are often exchanged for one another in SANS. Furthermore, the incoherent scattering from other common elements in biology are negligible compared to H and D. Notably, hydrogenated molecules, such as H_2O, gives rise to a high incoherent signal.

2.2.1.2 *The small-angle scattering form factor amplitude*

Due to the interference of scattered waves, the scattering amplitude changes with angle of detection and the distances between atoms in a particle. Each atom contributes with its scattering length, b, and a phase factor $e^{i\mathbf{q}\cdot\mathbf{r}}$, which leads to a general expression for the scattering amplitude for a particle of n atoms

$$A(\boldsymbol{q}) = b_1 e^{i\boldsymbol{q}\cdot\boldsymbol{r_1}} + b_2 e^{i\boldsymbol{q}\cdot\boldsymbol{r_2}} + \ldots + b_n e^{i\boldsymbol{q}\cdot\boldsymbol{r_n}} = \sum_{j=1}^{n} b_j e^{i\boldsymbol{q}\cdot\boldsymbol{r_j}} \tag{8}$$

where $\boldsymbol{r}$ is a vector from an arbitrary origin to the atom. SAS is typically measured at low angles with $q_{max} \sim 0.5$ Å^{-1}, where the resolution is not atomic. Thus, a particle can be described by its scattering length density (SLD), ρ, which is the sum of the scattering lengths from individual atoms in the particle divided by the particle volume. For particles in solution, the solvent gives rise to scattering and has to be taken into account as well. In effect, the scattering potency of a particle in solution is given by its excess SLD to the solvent as

$$\Delta\rho(\boldsymbol{r}) = \rho(\boldsymbol{r}) - \rho_{solution} \tag{9}$$

The excess SLD is also termed the contrast. Assuming a continuous excess SLD function, $\Delta\rho(\boldsymbol{r})$, equation 8 can be expressed as an integral over the volume, V, of a particle

$$A(\boldsymbol{q}) = \int_V \Delta\rho(\boldsymbol{r}) e^{i\boldsymbol{q}\cdot\boldsymbol{r}} d\boldsymbol{r} \tag{10}$$

$A(\boldsymbol{q})$ is termed the form factor amplitude and recognizes as the Fourier transform of the excess SLD function, $\Delta\rho(\boldsymbol{r})$.

2.2.1.3 *The small-angle scattering form factor intensity*

The experimentally measured quantity is the scattering intensity and not the amplitude. The form factor intensity is given as the product of the form factor amplitude with its complex conjugate $A(\boldsymbol{q})^*$

$$I(\boldsymbol{q}) = A(\boldsymbol{q})A(\boldsymbol{q})^* = A^2(\boldsymbol{q}) \tag{11}$$

The scattering intensity is effectively a measure of the scattering cross section (cm^2) per unit volume (cm^3), which gives the macroscopic scattering cross section in units of cm^{-1}.

In solution, particles rotate freely and are probed in all orientations. This implies that the particle signal is orientationally averaged, and the scattering vector $\boldsymbol{q}$ can be replaced by its modulus, q (equation 5). For sufficiently dilute particles and with orientational averaging applied, the scattering intensity of a sample of n identical particles, each of volume V, is given by

$$I(q) = n\Delta\rho^2 V^2 P(q) \tag{12}$$

where the normalized form factor intensity $P(q)$ has been introduced. $P(q)$ encodes the structural information about the size and shape of the particle and is given by

$$P(q) = \frac{\langle (\int_V \Delta\rho(\boldsymbol{r}) e^{i\boldsymbol{q}\cdot\boldsymbol{r}} d\boldsymbol{r})^2 \rangle}{M} \tag{13}$$

where M is the total scattering mass $M = (\int_V \Delta\rho(\boldsymbol{r}) d\boldsymbol{r})^2$ and the angle brackets indicate orientational averaging. It follows that $P(q)$ is normalized such that $P(0) = 1$. Because $I(q)$ is the square of $A(q)$, all information about phases is lost. Together with the orientational averaging, this has the consequence that the data can not be directly transformed into the SLD function, $\Delta\rho(\boldsymbol{r})$.

$I(q)$ is related to the commonly used pair-distance distribution function, $p(r)$ (Otto Glatter 2002) by

$$I(q) = 4\pi \int_0^\infty p(r) \frac{sin(qr)}{qr} dr \tag{14}$$

where r is the pair-distance between two scattering centers. Like $P(q)$, the $p(r)$-distribution contains information about the size and shape of the particle, but in real-space coordinates (Section 2.3.2.3). In practice, $I(q)$ can not be directly transformed to obtain the $p(r)$-distribution because it would require the collection of data from $q = 0$ to $q = \infty$, which is not experimentally possible.

For concentrated samples, interparticle interactions, *i.e.* repulsion or attraction, will contribute to the scattering intensity, and another function, the so-called effective structure factor $S(q)$, must be included in equation 12. The scattering intensity of samples with interparticle interactions present is given by

$$I(q) = n\Delta\rho^2 V^2 P(q) S(q) \tag{15}$$

$S(q)$ depends on the shape, size and internal orientation of particles and is non-trivial to calculate. Analytical expressions have been derived for certain particles, *e.g.* hard spheres, which in some cases can be used as approximation (J. S. Pedersen 2002). In practice, samples are often measured in a dilution series to find the optimal concentration where interparticle interactions are eliminated (Skrypnychuk et al. 2016).

Finally, if a sample contains several different species, the total scattering intensity will effectively be the sum-weighted contribution from the individual species, where the different species in the sample have specific volumes, contrasts and form factors (Jeffries et al. 2016). Unless the stoichiometry or shapes of the species are relatively well-known, it becomes difficult to extract meaningful structural information from the scattering intensity.

2.2.2 *Contrast variation*

From the definition of contrast (equation 9), it follows that this parameter can be varied in scattering experiments by changing the SLD of the molecule or the solvent, which for bio-molecules is usually an aqueous buffer. This is commonly used to study multi-component systems (Zaccai et al. 2016), including nanodiscs (Skar-Gislinge et al. 2010) and membrane proteins (MPs) in nanodiscs (Kynde et al. 2014). SAXS and SANS can provide complementary information due to their inherently different contrast situations of some molecules, including protein and lipid in the nanodisc; the more electron-rich protein belts and lipid head groups contribute most to the scattering signal in SAXS (Figure 12B), whereas the hydrogen rich lipid tail groups contribute most to the scattering signal in SANS in D_2O (Figure 12A). Table 7 lists the contrast for some of the molecules used in this thesis.

Contrast variation in SAXS is achieved by adding or removing electrons, which implies substitution of atoms or molecules. The SLD of the buffer can be tuned by adding electron-rich salts or osmolytes such as sugars (Ballauff 2001). However, to obtain an appreciated effect, these should be added in very high concentrations, which is expected to change the activity of the solvent and might perturb the structure of

Molecule	SAXS in H_2O $\Delta\rho{\cdot}10^{-14}$, cm·Å^{-3}	SANS in D_2O $\Delta\rho{\cdot}10^{-14}$, cm·Å^{-3}
$POPC_{head}$	5.10	-4.50
$POPC_{tail}$	-1.65	-6.17
DDM_{head}∗	5.20	-3.78
DDM_{tail}	-1.53	-6.78
Protein†	2.80	-3.30

Table 7 Excess scattering length densities of some bio-molecules. DDM: n-dodecyl β-D-maltoside. POPC: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine.
∗ Assuming H/D exchange in hydroxyl groups.
† Assuming overall 90% H/D exchange for labile hydrogens.

the bio-molecule. Changing the SLD of the bio-molecule would in most cases not be feasible as it would require a dramatic change of its chemical composition.

In SANS, contrast variation is readily achieved by hydrogen/deuterium (H/D) substitution: H and D have identical volumes, but dramatically different scattering lengths (Table 6) and thereby different SLDs as well. The SLD of the solvent varies linearly with the deuterium content, *i.e.* the percentage of D_2O (Figure 12A). A common strategy is to match the contrast between molecules, such as proteins, and the solvent, which causes elimination of the scattering intensity, as seen from equation 12. The so-called contrast-match point (CMP) for a molecule is given by the % of D_2O required to eliminate its forward scattering intensity ($I(0)$=0). For particles with homogenous SLD, $I(q)$ will also be zero at the CMP. The CMP is usually determined by a contrast-variation series, where *e.g.* a protein is measured in a series of buffers with varying D_2O content (Zaccai et al. 2016). It is generally found that proteins have CMPs around 42% D_2O. Thus, at this level of D_2O, only the lipids would be visible in a nanodisc (Figure 12A). It is thereby clear that contrast-matching can be used to study molecular complexes between proteins and other types of molecules with different SLDs, such as lipids, in better detail. The same holds true for *e.g.* complexes of protein and DNA or RNA, which have CMPs in the range 60% to 70% D_2O (Gabel 2015; Zaccai et al. 2016).

The other common strategy in SANS is to produce deuterated molecules, which have increased SLDs and thus increased CMPs. In this way, molecular complexes of the same types of molecules, such as two proteins, can be investigated in better detail given that one protein is deuterated and the other not (Whitten et al. 2007; Appolaire et al. 2014). Likewise, MPs can be studied by SANS without interfering scattering from the carrier system (Maric et al. 2014; Midtgaard et al. 2018; Naing et al. 2018; Josts et al. 2018). For MPs, it is especially desired to produce deuterated carrier systems that uniformly match the SLD of D_2O, as exemplified by contrast-matched nanodiscs (Figure 12C) (Maric et al. 2014; Josts et al. 2018). First, in 100% D_2O, the contrast of a hydrogenated MP is good (Figure 12A), meaning that there is no need to produce deuterated MPs. This is of great practical importance, as MP expression and purification is difficult and not all expression organisms can grow in deuterated media (Dunne et al. 2017). Second, because there is 0% H_2O in the sample (ignoring exchange of labile hydrogens in proteins), the background due to incoherent

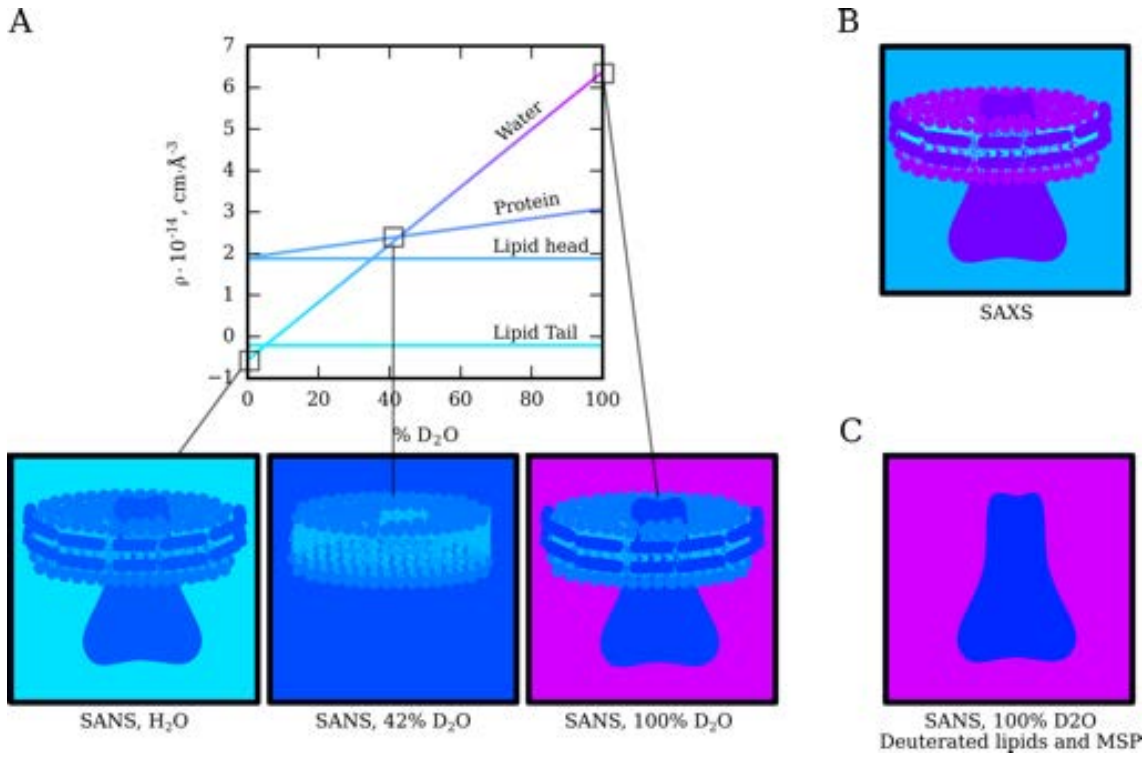


Figure 12 Contrast and contrast variation. **A:** Principle of contrast variation in SANS by tuning the D_2O content of the solvent exemplified by a MP in a nanodisc. The SLD is represented by the color ranging from light blue to purple, as shown for water in the graph. Molecules (or part of molecules) that match the color of the solvent contribute with zero scattering intensity. **B:** SAXS provides an inherent complimentary contrast to SANS. Note that the lipids have negative contrast. **C:** Using selective deuteration, it is possible to contrast-match specific molecules. In this example, the nanodisc is contrast-matched to D_2O in SANS, whereby only the MP contributes to the scattering intensity.

scattering from hydrogen (see Table 6) is minimal (Jacrot 1976). Carrier systems that uniformly match at 100% D_2O is a new and interesting approach for studying MPs in solution and requires selective deuteration strategies (Section 1.3.4.3) (Maric et al. 2014; Josts et al. 2018; Midtgaard et al. 2018). The use of such uniformly contrast-matched carrier systems for studying the solution structure of the MP CorA was applied in Paper IV.

2.2.3 *Experimental setup*

In principal, experimental setups for classical SAXS and SANS are quite similar. X-rays and neutrons generated from a suitable source are first made monochromatic (Figure 13). For X-rays, this can *e.g.* be achieved with double crystals (Blanchet et al. 2015) or double multilayers (Pernot et al. 2013). For neutrons, a velocity selector can be utilized given the relationship between wavelength and velocity (equation 3 and equation 4).

The velocity-dependent energy of neutrons can also be exploited for so-called time-of-flight (TOF) SANS using a pulsed source (Ostanevich 1988). In this technique, the detector reads out scattering intensity as a function of time as fast neutrons (short wavelength) will arrive at the detector before slow neutrons (long

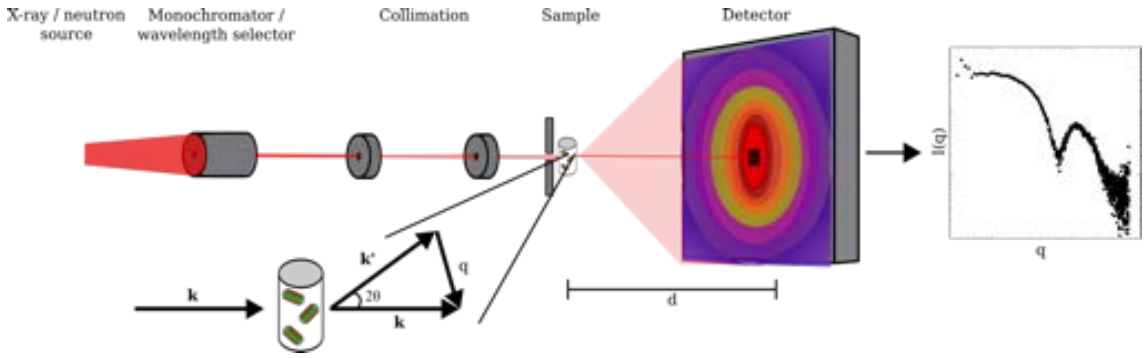


Figure 13 A simplified experimental setup for a SAS experiment. The instrumentation for SAXS and SANS can be described by the same basic setup. X-rays or neutrons are generated by a source. The desired wavelength or distribution of wavelengths are selected by a monochromator. The monochromatized beam is then collimated and focused on the sample. The scattering intensity from the sample is collected at a detector placed a distance, d, from the sample. The direct beam is blocked by a beam stop. Due to orientational averaging of molecules in solution, the scattering pattern is centrosymmetric on the detector and can be azimuthally averaged into a 1D curve of $I(q)$ vs q.

wavelength). Following the relation between the scattering vector q and the wavelength (equation 5), a wide q-range can be obtained for a single sample-to-detector distance in TOF-SANS. Many spallation sources are pulsed by nature, whereas continuous sources have to be chopped. TOF-SANS was not utilized in this thesis, but will be an interesting feature to explore at the upcoming ESS facility (section 2.4.3).

After the monochromator, or velocity selector, a section of slits and mirrors is used to collimate the beam and focus it on the sample (Figure 13). The scattering from the sample is recorded on a detector placed with some defined distance to the sample and in a flight tube that is usually kept under high vacuum to eliminate air scattering. Changing the sample to detector distance allows for measuring more angles to cover a larger q-range, which is typically needed in bioSANS. For some bioSAXS instruments, however, the distance is fixed, because a q-range relevant for most bio-molecules can be achieved in a single setting. For the dilute samples in bioSAXS and bioSANS, most of the incoming beam passes through the sample, and the detector is thus protected by a beam stop (Figure 13, black square on the detector). Due to the orientational averaging in solution SAS, the scattering pattern measured on the detector is centrosymmetric. It can thus be azimuthally averaged to yield a one-dimensional scattering curve with $I(q)$ vs q (Figure 13). Because data is measured as a function of q and not the angle, data obtained with different wavelengths and/or sample-to-detector distance are directly comparable. In practice, and as seen on the plot in Figure 13, the intensities at very low-q are often not well-determined. These are the intensities recorded very close to the beam stop that can be extra sensitive to *e.g.* the alignment of the instrument. Obviously, these data points are systematic outliers and should be discarded before data analysis.

2.3 ANALYSIS OF SAS DATA

2.3.1 *Absolute calibration*

Absolute calibration refers to measuring the scattering intensity in units of cm^{-1}, which is equal to the macroscopic differential scattering cross section (Section 2.2.1.3). By absolute calibrating the intensities, information is obtained on the total amount of scattering, which can be directly related to the number of electron/nuclei in the sample. This information can be used as a constraint in model-fitting or for determination of molecular weight as explained later in this section.

There are several options for absolute calibration. In this thesis, H_2O was used as the main calibration for SAXS intensity given its angle-independent (isotropic) scattering signal. After subtraction of scattering from the sample capillary, the flat part of the scattering signal from H_2O at 20°C is given by $I(q)_{H_2O} = 1.632 \cdot 10^{-2}$ cm^{-1} (Orthaber et al. 2000). Alternatively, the absolute scale can be determined from a protein standard (Skou et al. 2014). In this case, it is of great importance that the protein is completely monodisperse, which is unfortunately seldom the case for the widely used protein standard bovine serum albumin. For SANS, absolute intensity calibration can be achieved by measurement of the direct beam intensity or using secondary standards like water (Jacrot 1976). For the work in this thesis, absolute calibration was obtained with measurements of the primary intensity of the incoming beam using mechanical attenuation (at D22 at ILL and at QUOKKA at ANSTO) or by plexiglass as a secondary standard (at KWS-1 at FRM2), which like water has isotropic scattering.

2.3.2 *Model-independent analysis*

The orientational averaging makes SAS a technique of relatively low information content. However, it allows for the immediate determination of a few very useful overall parameters related to molecule size and shape.

2.3.2.1 *Guinier analysis*

In the Guinier approximation (Guinier and Fournet 1955), the scattering intensity at low angles is directly related to the radius of gyration, R_g, of the molecule by

$$I(q \sim 0) \simeq I(0) e^{-R_g^2 q^2/3} \tag{16}$$

where $I(0)$ is termed the forward scattering. R_g is the root mean square distances from scattering centers to the center-of-mass in a molecule weighted by the SLD distribution. A non-aggregated sample should reveal a linear relationship between $log(I(q))$ and q^2 within the range $q < 1.3/R_g$ beyond which the approximation breaks down (D. I. Svergun and Koch 2003). This flat low-q region is also sometimes termed the Guinier-region. Data should preferably be collected down to a q_{min}, where a Guinier-region is obtained.

2.3.2.2 *Molecular weight*

If the scattering data is measured on absolute scale, the molecular weight of the sample can be determined from the forward scattering, *e.g.* determined from Guinier analysis. Using the partial specific volume of the molecule, $\widetilde{v}$, and the solute concentration in weight per volume, c, equation 12 can be conveniently rewritten (Orthaber et al. 2000) to

$$I(q) = c(\Delta\rho\widetilde{v})^2 \frac{M}{N_A} P(q) \tag{17}$$

where M is the molecular weight in kDa and N_A is Avogadro's number. From equation 17, it is evident that the molecular weight of a molecule is proportional to $I(0)$ and can be calculated by

$$M = \frac{I(0)N_A}{c(\Delta\rho\widetilde{v})^2} \tag{18}$$

where $I(0)$ is *e.g.* estimated from Guinier analysis (equation 16) or from $p(r)$-distribution refinement (Section 2.3.2.3). Many proteins have the same effective value of the partial specific volume, $\widetilde{v}$ = 0.7425 cm^3/g (Mylonas and D. I. Svergun 2007).

2.3.2.3 *Pair-distance distributions*

Equation 14 shows the relation between the scattering intensity and the $p(r)$-distribution. By applying indirect Fourier transformation (IFT), a $p(r)$-distribution can fitted to the scattering intensity (O. Glatter 1977). In very basic principle, the IFT method use a linear combination of smooth functions constrained by a maximum pair-distance (D_{max}) to fit the data. The $p(r)$-distribution shows the relative abundance of pair-distances in the molecule and is represented on real-space coordinates. Figure 14 shows a couple of examples of $p(r)$-distribution, refined with data from protein and nanodisc samples. Globular proteins gives rise to bell-shaped or slightly skewed distributions (Figure 14, blue and green), whereas it is more complicated for the nanodisc (Figure 14, red and yellow). This is because the distances are weighted by the contrast, and $p(r)$-distributions for multi-contrast molecular complexes, such as nanodiscs can be difficult to interpret. An example is the nanodisc that consists of protein (positive SAXS contrast) and lipids (negative SAXS contrast in tail-groups), which gives rise to negative values in the $p(r)$-distribution (Fig-

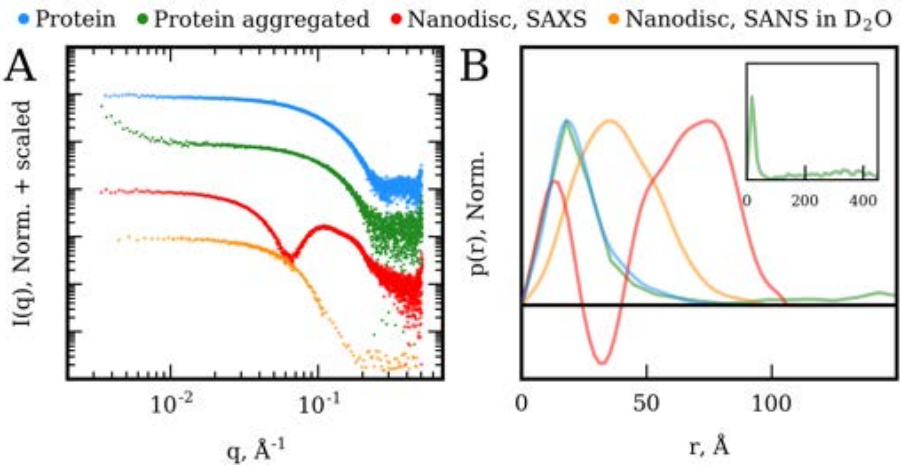


Figure 14 Examples of $p(r)$-distributions. **A:** SAXS data from two samples of the same protein (IM7, see report I) (blue and green), together with SAXS (red) and SANS (orange, in D_2O) data obtained from a nanodisc sample. One SAXS data set (green) indicates an aggregated sample, as seen by the upturn at low-q. **B:** $p(r)$-distributions obtained by IFT. The inset shows the entire distribution for the aggregated sample in A.

ure 14). In SANS, the nanodisc has an overall more homogenous contrast and yields an entirely positive $p(r)$-distribution (Figure 14). From the IFT, $I(0)$ and R_g can be extracted. Unlike the Guinier approximation, $p(r)$-distributions use data of the entire q-range collected and are thereby likely to give better estimates for these values in samples containing small amounts of impurities (Putnam et al. 2007). Furthermore, the maximum molecular distance D_{max} can be extracted as the r-value where the $p(r)$-distribution goes to zero. However, an accurate determination of this parameter can be difficult and is affected by the choice of algorithm for refining the IFT (Putnam et al. 2007). Finally, the $p(r)$-distribution can be used to evaluate whether a sample contains aggregation, as discussed later (Section 2.4.1).

2.3.3 *Modeling*

Because of the orientational average in solution and concomitant loss of three dimensional (3D) information as well as the loss of information about phases, it is not possible to obtain 3D models directly from the SAS data. Instead, an indirect approach is applied, where a model is developed and fitted to the data. A few approaches are introduced in this section.

2.3.3.1 *Geometrical models*

With the inherent low resolution in SAS, geometrical shapes are often applied as models and fitted to the data. A good example is the model developed for a nanodisc that consist of a number of cylindrical objects (Skar-

Figure 15 Geometric model of a nanodisc. The overall geometric model of a nanodisc is described by a hallow cylinder representing the MSP belt (green), two cylinders representing the lipid head groups (red), two cylinders representing the methylenes of the lipid tail groups (grey) and a cylinder representing the methyls of the lipid tail-groups. The cylinders are parameterized with an axis ratio that allows them to be elliptical, as seen from the top-view.

Gislinge et al. 2010; Skar-Gislinge and Arleth 2011). The total form factor amplitude of the nanodisc is given by the sum of the form factor amplitudes from individual components (Figure 15)

$$A(\boldsymbol{q})_{MSPbelt} + A(\boldsymbol{q})_{lipidheads} + A(\boldsymbol{q})_{lipidmethylenes} + A(\boldsymbol{q})_{lipidmethyls} = A(\boldsymbol{q})_{nanodisc} \tag{19}$$

Such a model contains a large number of fitting parameters, but many of these are constrained with *a priori* information, allowing for a physically meaningful fit. For the nanodisc model, for instance, the molecular composition is well-known along with the associated partial specific volumes of the different components. Thus, the cylindrical objects that describe the low-resolution structure of the nanodisc can be endowed with the SLD densitites of the molecules they are meant to represent (Skar-Gislinge et al. 2010; Skar-Gislinge and Arleth 2011). When more contrasts are measured, a model can be refined with better certainty as the model has to match the different contrasts (M. C. Pedersen et al. 2014). For the development and validation of the nanodisc model, SAXS data were complemented by SANS data at 42% D_2O (protein matched) and at 100% D_2O (Skar-Gislinge et al. 2010). The nanodisc model along with a few other analytical models for geometrical shapes can be fitted in the WillItFit software (M. C. Pedersen et al. 2013).

2.3.3.2 *Modeling of proteins*

For many proteins, crystal structures are available and can be fitted to the data because the atomic distances are well-known. Assuming rigid structures, *e.g.* protein-protein complexes can be modeled using so-called rigid body modelling, for instance in the program SASREF (Petoukhov and D. I. Svergun 2005). In the case of a protein of unknown structure, *ab initio* modeling software has been developed to determine the overall shape of proteins by bead-modelling, such as DAMMIF (Franke and D. I. Svergun 2009). In bead-modeling, the data are in principle over-fitted, but the algorithms are commonly observed to reproduce overall similar structures from multiple runs. Bead modeling of multi-contrast data sets is available in the program MONSA (D. Svergun 1999). As standard practice, the results of several runs are averaged and

filtered for occupancy to produce a consensus bead-model. Importantly, it is only the bead-models from the individual runs, and not necessarily the consensus model, that fit the data.

To fit MPs into nanodisc, a hybrid approach of a geometrical model of the nanodisc can be combined with bead modeling of the MP (Kynde et al. 2014). This model is also availbale in WillItFit (M. C. Pedersen et al. 2013).

2.3.3.3 *Aggregates in models*

Data from slightly aggregated samples are particularly problematic in regards to refining analytical models containing a large number of fitting parameters, such as the nanodisc model (Skar-Gislinge et al. 2010). A tempting fix is to truncate data at a q-value where the upturn is not present. With this procedure, it is assumed that eventual aggregates neither alter the finer structure of the particles nor smear the signal at the corresponding high q-values. This assumption is obviously not always valid and truncation should be done with great caution.

For MPs in detergent micelles or other dynamic carriers, oligomerization or aggregation over time can be problematic. For the MP GluA2 measured in sDDM micelles, random higher order oligomers were modeled as mass fractals to minimize misinterpreation of the scattering data towards a simple structural hypothesis about the native tetrameric state (A. H. Larsen et al. 2018). The same approach was applied for the MP SERCA in sDDM micelles, where the high-q data was of high quality and allowed for the distingushing between two structural models (Midtgaard et al. 2018). Thus, given a sufficiently simple hypothesis, *e.g.* about a large conformational change in a protein, aggregation is not necessarily detrimental to the subsequent data analysis. Of course, for modeling of single particles, it is preferred to measure monodisperse particles and obtain a single, high quality curve without the need to introduce extra fitting parameters to describe *e.g.* fractals.

Aggregates forming in biological samples over time can in some cases be fairly ordered. An example is protein fibrillation that refers to the transition of monomeric protein into long, well-ordered molecular assemblies. For these systems, it is an advantage that scattering intensities from different species are additive, which makes it possible to deconvolute the overall scattering into individual signals along with their relative populations (Herranz-Trillo et al. 2017). A new ATSAS program, DAMMIX (Konarev and D. I. Svergun 2018), allows for modeling of an unknown intermediate structure, when models or experimental data are available for the initial and end states.

2.3.3.4 *Resolution*

The scattering vector, q, has units of $Å^{-1}$ and can in principle be converted to a real space distance, d, by Bragg's law

$$d = \frac{2\pi}{q} \tag{20}$$

SAS data obtained to $q_{max} = 0.63 Å^{-1}$ thus gives a nominal resolution of 10 Å. However, the resolution is not discrete with this value, and SAS is sensitive to much smaller changes of molecular distances, as long as the distance is longer than the resolution defined by q_{max}. For instance, SAXS was sensitive to changes in the thickness of a lipid bilayer of only 2 Å (Andersen et al. 2011). This was exploited in Paper I to measure the size and shape of nanodiscs prepared under different conditions using SAXS.

2.4 CHALLENGES IN SMALL-ANGLE SCATTERING

On the cost of resolution and low information content owing to orientational averaging, sample preparation in SAS is much easier than for crystallography as there is no need for obtaining a well-diffracting crystal. As the sample is measured in solution in SAS, it is relatively easy to screen a large number of conditions, such as pH, ionic strength, and temperature (Skrypnychuk et al. 2016). However, the effort of adequate sample preparation for SAS experiments should not be underestimated as everything present in the sample will contribute to the scattering intensity as explained in Section 2.2.1.3 (Jeffries et al. 2016). Thus, to obtain high quality data, highly pure and homogeneous samples are needed. Furthermore, due to the relatively low signal-to-noise ratio between sample and buffer in SAXS, and the weak neutron flux in SANS, sample concentrations of 0.5-10 mg/ml (depending on the size of the molecules) are typically required to obtain sufficient signal (Jeffries et al. 2016). As explained in section 2.2.1.3, it is also important to avoid interparticle interactions when the goal is to characterize single-particle structure. Therefore, concentration series are often measured in SAXS to find the optimal concentration. If interparticle interactions are only eliminated at low concentration, the high-q data will be of low signal-to-noise ratio. In some cases, the high-q part can then be recorded in a more concentrated sample and subsequently merged with the low-q data measured in low concentration (Skou et al. 2014).

2.4.1 *Aggregation-free data*

A particular concern is the presence of aggregates in the sample. Many bio-molecules become heterogeneous over time after purification, for instance by slight aggregation induced by storage conditions. It can

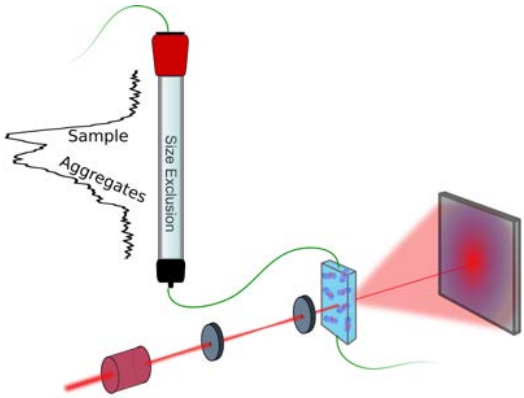


Figure 16 Principle of SEC-SAS. The sample is purified by SEC directly into the capillary/flow-cell that is illuminated with X-rays/neutrons.

be very difficult to test for sufficient sample stability until at a SAXS instrument. Data from slightly aggregated samples will have relatively more scattering at low angles due to the presence of very large particles, and the scattering curve will often feature an upturn at low-q (Figure 14A, green). Qualitatively, the corresponding $p(r)$-distribution (Figure 14B, green) can serve as a guide to evaluate sample homogeneity: The lack of a smooth descend to D_{max} as well as a much larger D_{max} than expected for the given structure would in most cases indicate the presence of larger or aggregated particles. Aggregates will also show in Guinier-plots, or log-log plots, that both deviate from linearity at low-q.

2.4.1.1 *Size-exclusion chromatography (SEC) combined with SAXS or SANS*

Before a SAS measurement, size-exclusion chromatography (SEC) is usually applied to obtain homogeneous samples. To eliminate the eventual effects of storage between purification and SAS measurement, set-ups of online SEC coupled to SAS (SEC-SAS) (Figure 16) have been installed at most bioSAXS beamlines (Mathew et al. 2004; David and Pérez 2009; Pernot et al. 2013; Blanchet et al. 2015; Ryan et al. 2018) and lately at the D22 SANS instrument at ILL, which is described in Paper II (Jordan et al. 2016). With SEC-SAS, molecules can be measured as they are purified, which is highly efficient for eliminating issues with interfering aggregates (Jeffries et al. 2016). SEC-SAXS has become very popular over the last few years, facilitating development of dedicated software for data analysis (Panjkovich and D. I. Svergun 2018) as well as new developments for the use of other direct purification methods coupled with SAXS, such as ion-exchange chromatography (Brennich et al. 2017).

A main disadvantage of the method is the time required per measurement, which is equivalent to the time a SEC run takes and often amounts to between 20 min and 2 hours. For SANS, where more than one detector setting is usually required to cover a broad q-range, equally many SEC runs are also required, causing higher sample consumption. Furthermore, UV absorption is sometimes measured in the HPLC, and the sample might be diluted before reaching the SAXS capillary, which makes concentration estimation difficult. This has been recognized as a major source of error and efforts are made towards better solutions: At the SWING

beamline at Soleil, UV absorption is detected on the SAXS capillary (David and Pérez 2009), whereas at the SAXS/WAXS beamline at the Australian Synchrotron, the UV detector has been placed as close to the capillary as possible (Ryan et al. 2018). The P12 beamline at the Deutsches Elektronen-Synchrotron (DESY) now features a fully integrated systems for combined HPLC, SAXS, UV absorption, light scattering, and refractive index measurements to allow a thorough sample characterization and good concentration estimation (Graewert et al. 2015). During the coming shutdown at the European Synchrotron Radiation Facility (ESRF), BM29 will be upgraded with a better solution for estimating concentration adequately. A combined SEC-SAXS and SEC-SANS study is described in paper II, and SEC-SANS was also utilized to study CorA in paper IV. Overall, these methods provided superior, aggregation-free data and were thus important to achieve data of sufficient quality for a number of MP samples throughout this thesis.

2.4.2 *Other considerations*

2.4.2.1 *Radiation damage in SAXS*

With the extremely high flux at modern synchrotrons, the radiation dose imposed on samples is high enough to cause radical formation, which leads to sample deterioration (Kuwamoto et al. 2004). This phenomena is termed radiation damage and is a major concern in SAXS experiments (Hopkins and Thorne 2016). At bioSAXS beamlines, a first solution to alleviate radiation damage is to flow the sample during measurement, such that the absorbed dose is spread out on a larger volume (Pernot et al. 2013; Blanchet et al. 2015). This is done both in robot/static measurements and in SEC-SAXS. Data frames are usually collected in short time bins, *e.g.* 50 ms at P12 (Blanchet et al. 2015) or 1 s at BM29 (Pernot et al. 2013), which allows for comparison of frames to evaluate whether a temporal signal evolution in the scattering curves is present, as this would indicate fouling of the sample capillary caused by radiation damage. Furthermore, buffer frames measured before and after the sample frames are compared and should match for a sample without radiation damage.

Often, flowing of the sample is not sufficient. Addition of radiation scavengers, such as glycerol (Kuwamoto et al. 2004), uridine (Crosas et al. 2017), or reducing agents such as dithiothreitol (DTT) or ascorbic acid (Jeffries et al. 2015), can raise the critical dose. However, these compounds are often added in high amounts, glycerol typically around 5 % v/v, and may interact with the sample. Furthermore, they may change the electron density of the solvent and thereby the contrast, which should be included in model calculations. A last resort to avoid radiation damage is to attenuate the beam intensity with e.g. aluminum plates. While this is very effective, it may seem counter-productive to not utilize the full intensity of modern synchrotrons. However, if good data can be obtained with slightly lower flux, it is worth doing so compared to adding

high concentrations of radiation scavengers. In this thesis, this was particularly needed for SEC-SAXS experiments, where the maximum flow-rate is limited by the SEC-column.

2.4.2.2 *D_2O in SANS*

The possibilty of contrast variation using H/D substitution is a major advantage of SANS. Although chemically very alike, H and D are however not identical and substitution might affect physiochemical properties of biological samples. For SANS, it is particularly problematic that protein solubility is markedly lower in D_2O than in H_2O (Gripon et al. 1997). This could lead to formation of aggregates in this solvent, which indeed has been a major bottleneck for obtaining good SANS data on nanodiscs and MP samples previously (Midtgaard 2013; Huda 2015). Thus, it is good practice to characterize samples in D_2O to make sure they are conformationally stable and monodisperse before an expensive SANS experiment (Jeffries et al. 2016). Often a well-resolved peak in SEC (in D_2O) is a good first indication of a well-behaved sample, and its structural integrity can further be verified *e.g.* by SAXS. Because SANS measurement times can be very long, SAXS should preferably be measured on a couple of time points to evaluate the temporal stability of the sample (Jeffries et al. 2016). In this thesis, proteins and nanodisc samples were investigated in D_2O in an attempt to gain a better understanding of the increased aggregation propensity observed in this solvent (Report I). In Paper IV, careful control experiments were carried out in D_2O to show that the MP CorA was structurally and functionally intact under the conditions used for SANS.

2.4.3 *New large-scale facilities*

SAXS has been widely used to study bio-molecules over the last few decades (D. I. Svergun and Koch 2003). The application of SAXS to bio-molecules has been facilitated by the development of synchrotrons during the last decades with exceptional flux of X-rays (Als-Nielsen and McMorrow 2011). Today, large-scale facilities over the world host dedicated bioSAXS beamlines with automatic sample changers. Performing a SAXS experiment is usually straight-forward and several hundreds of samples can be measured over the course of an 8 hour shift. Furthermore, a large amount of user-friendly software has been developed for SAXS, and experiments and data analysis can be carried out by non-expert users also. SANS for bio-molecules is also gaining popularity, but it remains a more exclusive technique. Sample requirements are higher than for SAXS as the flux at neutron sources is low, and measurement times are accordingly long. For comparison, the flux at the D22 instrument at ILL is on the order of 10^8 neutrons$\cdot$cm$^{-2}\cdot$s^{-1}, while it is 10^{12}-10^{14} photons$\cdot$cm$^{-2}\cdot$s^{-1} at bioSAXS beamlines, *e.g.* at P12 at DESY (Blanchet et al. 2015) and BM29 at ESRF (Pernot et al. 2013). The currently available data reduction software for SANS is not as user-friendly as the SAXS software and there is a general lack of supporting facilities for sample handling.

All these things combined indicate that SANS beamlines, in general, have not yet been fully streamlined for bioSANS, as is the case for bioSAXS beamlines at synchrotrons.

New facilities are under construction in Lund, Sweden, namely the MAX IV fourth-generation synchrotron X-ray source and the ESS neutron spallation source. At ESS, the LoKI instrument is planned to feature TOF-SANS (Section 2.2.3) combined with multiple detectors to cover a dynamic q-range over three orders of magnitude (A. J. Jackson and Kanaki 2014). This will allow the recording of a full q-range without moving detectors, leading to faster experiments. This mode is highly desirable for SEC-SANS, as the present need for running two or more identical samples to cover a broad q-range is eliminated. The pulsed source at the ESS is expected to deliver a neutron flux at the LoKI instrument that is 20 times higher than the present best at D22 at ILL. In the proposed plan for the LoKi instrument, it is also recognized that there is a general need to be able to measure smaller sample volumes in high throughput and with an automated data processing and (initial) analysis pipeline (A. J. Jackson and Kanaki 2014). This will, in combination with the high flux, facilitate the measurement of tens to hundreds of samples during a beam time and with the possibility of easily evaluating data quality during measurements. In this way, the user has a chance to improvise and optimize samples during the experiment. In regards to this, it is highly important that the plan of a nearby sample preparation facility is executed.

At present, the D22 instrument at ILL is most likely the best place internationally to carry out bioSANS experiments. Not only has the beamline been streamlined to measurements of biological samples, it is also located next to the European Molecular Biology Laboratory (EMBL) for sample preparation and the ESRF for complementary bioSAXS experiments. Samples prepared at the EMBL can be split in two and taken to the ILL and ESRF for SANS and SAXS, respectively, to make sure the exact same parent sample are probed with the two techniques. This eliminates variation in sample preparation and allows better comparison and/or combination of the data. Situated close to the ESS, the MAX IV synchrotron will feature a multi-purpose SAXS beamline (CoSAXS) with both a robot sample changer for standard solution SAXS and SEC-SAXS available (MAXIV 2018). It is the hope that these facilities in close vicinity should facilitate the same synergy that can be experienced at the ILL and ESRF.

2.5 OUTLOOK

SAXS and SANS can provide low-resolution information on samples in solution and provide internally complementary information as given by the contrast of the molecules in the sample. SAXS and SANS are considered as complementary low-resolution techniques to the high-resolution techniques of crystallography, NMR and cryoEM. Compared to crystallography, sample preparation is relatively straight-forward as molecules are investigated in solution, and compared to sNMR and cryoEM (Section 1.3.4), there is no

apparent size limitations. However, it is very important to prepare well-defined samples as aggregates will contribute significantly to the overall scattering.

The development of SEC-SAXS and SEC-SANS has effectively eliminated the issues related to sample aggregation (or degradation). These techniques were applied in paper I and paper II to study nanodisc samples with or without MP incorporated. For paper II, it was suggested by a reviewer that SEC-SANS is presently only feasible at D22 owing to the high flux and the state-of-the-art low background detector. However, the technique should also soon become implemented at the SANS instruments at FRM2 at MLZ in Munich. Like SEC-SAXS has made it possible to study a wider range of unstable biological samples, it can also be speculated that the increased availability SEC-SANS will facilitate the use of SANS on a wider array og biological samples and in a wider scientific community than at present. Overall, new large scale neutron facilities together with the availability of selectively deuterated molecules and automated sample handling will help to establish SANS in the biological communities and not leave it as a niche method.

3

RESULTS AND CONCLUSION

This chapter is divided into two sections that present an excerpt of the main results from the papers in this thesis (Part II) together with a few illustrative preliminary data sets. The first section (Section 3.1) presents work that was done to gain a better understanding of the self-assembly process of the nanodisc carrier system (Section 3.1.1) together with protein-engineering efforts to make more stable and less aggregation prone nanodiscs (Section 3.1.2). The second section (Section 3.2) presents work that was done to obtain an optimal basis for structural analysis of membrane proteins (MPs) in solution by small-angle scattering (SAS). This work includes studies on nanodisc and protein stability as well as nanodisc aggregation propensity in D_2O (Section 3.2.1), the use of size-exclusion chromatography (SEC) coupled to small-angle scattering (SEC-SAS) for MP samples and the development of stealth MP carrier systems (Section 3.2.3.1).

3.1 DEVELOPING THE NANODISC AS AN MP CARRIER

Recently, when the nanodisc field was reviewed, it was stated that despite the general success of the technology, the processes of nanodisc self-assembly and reconstitution of MPs into nanodiscs were still not well-understood (Ilia G. Denisov and Sligar 2017). Furthermore, the number of structural studies of MPs in nanodiscs is still relatively low. One explanation for this is that it is generally difficult to obtain a structurally homogenous sample of a MP inside a nanodisc. Functional studies of MPs in nanodiscs, which are more abundant, do not require the same level of homogeneity and can often be performed at low concentrations. Thus, more knowledge about the nanodisc reconstitution process is expected to facilitate an increase in the number of structural studies of MPs in the future.

3.1.1 *Understanding the nanodisc self-assembly*

Reconstitution of MPs in nanodiscs (Section 1.2.7) is expected to be more difficult than for the more dynamic carrier systems, such as detergents (Section 1.2.1), SMALP (Section 1.2.4.2) or peptide discs

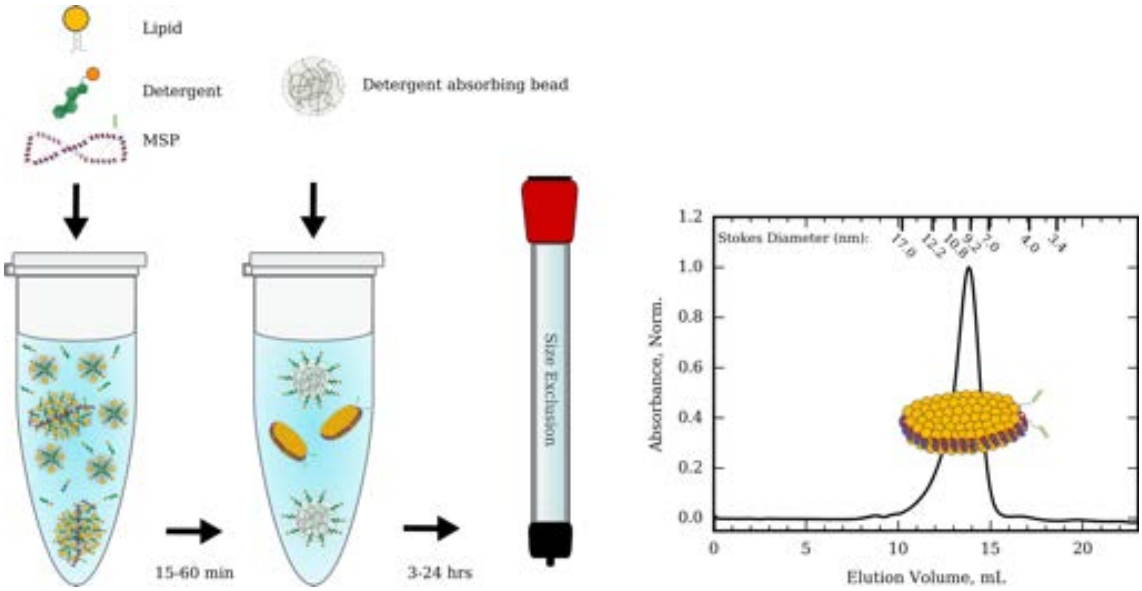


Figure 17 Overview of nanodisc reconstitution. In the standard scheme for nanodisc reconstitution, mixed micelles of lipid and detergent are mixed with membrane scaffold protein (MSP) and incubated for a short while. In a second step, detergent absorbing beads are added and nanodisc form by self-assembly. Monodisperse nanodisc are separated from eventual aggregates by SEC. The size of the nanodisc depends on the length of the MSP belt protein. Nanodiscs made with MSP1D1 are approximately 10 nm wide. The figure is adapted from Paper I.

(Section 1.2.5.1) that can accommodate different stoichiometries and adapt to MPs of different size (Section 1.3.2). In the nanodisc, the membrane scaffold protein (MSP) puts a tight constraint of the available space for phospholipid loading and MP loading, and reconstitution parameters must often be fine-tuned empirically to obtain good quality samples. In Paper I, the self-assembly process of empty nanodiscs made from MSP1D1 and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) (Figure 17) was systematically modulated to gain insight to the factors that determine the size and shape of nanodiscs.

3.1.1.1 *The choice of reconstitution detergent is a major determining factor for nanodisc size and shape*

Given the individual properties of reconstitution detergents (Section 1.2.1) related to *e.g.* size, shape and charge, the structure of POPC-detergent mixed micelles will depend on the type of detergent. Cholate is used as the standard detergent for the classical nanodisc reconstitution protocol (Ritchie et al. 2009), but seldom for MP stabilization. It is thus of great importance to probe the effect of using other types of detergents in the nanodisc reconstitution protocol.

In Paper I, SAXS was used to evaluate how the choice reconstitution detergents affected the size and shape of the formed nanodiscs. Standard nanodiscs prepared with MSP1D1, POPC and cholate gives rise to a SAXS pattern with a local minimum at $q = 0.065$ Å^{-1} and a double-bump between $q = 0.1$ Å^{-1} and $q = 0.2$ Å^{-1} (Figure 18A, red data). These nanodiscs are slightly elliptical and contain roughly 125 POPC lipids (Figure 18B). Interestingly, changing the type of detergent had a major impact on nanodisc size, shape and lipid stoichiometry. A particular highlight was that nanodiscs prepared with n-dodecyl-β-D-maltoside

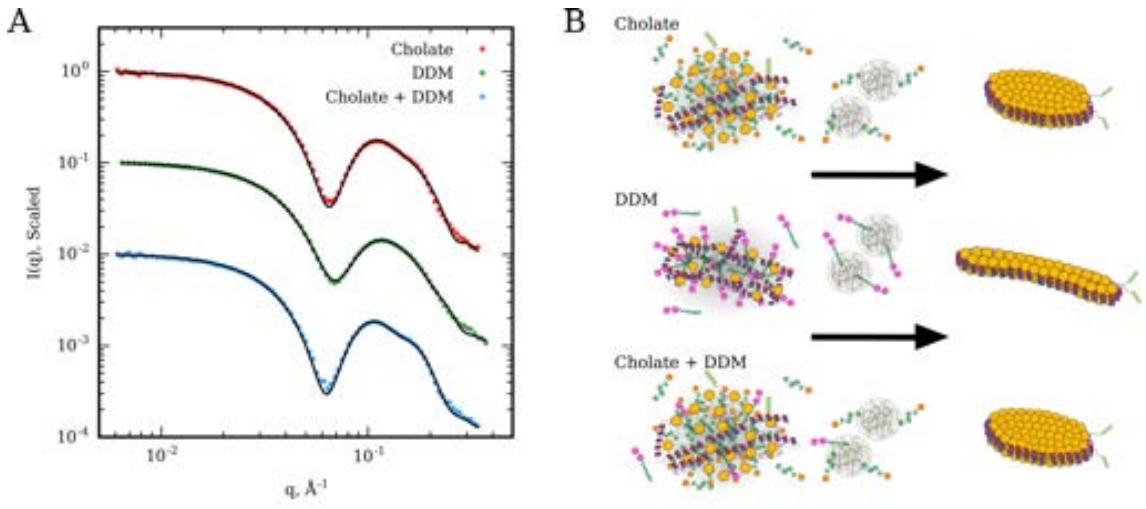


Figure 18 Effect of DDM on nanodisc assembly. A: SAXS curves obtained from nanodisc samples prepared with cholate, DDM or both. The SAXS data from the cholate sample serves as a reference for an optimal nanodisc. **B:** Schematic representations of the formation of nanodiscs. Based on the SAXS data in **A**, nanodiscs prepared with DDM yield elongated and lipid-poor nanodiscs, whereas a mix of cholate and DDM yield identical nanodiscs to those obtained from cholate alone. The figure is adapted from Paper I.

(DDM) were very different from nanodiscs prepared with cholate. This was clearly visible in the SAXS data, where the local minimum was shifted slightly to higher q and only a single bump was present between $q = 0.1$ Å^{-1} and $q = 0.2$ Å^{-1} (Figure 18A, green data). Furthermore, model fitting to these data showed that nanodiscs prepared with DDM were elongated and lipid-poor (83±1 POPC per nanodisc) (Figure 18B). This is problematic given that DDM is the most widely used detergent for MP handling (Stetsenko and Guskov 2017). Nanodiscs were then reconstituted with both cholate and DDM, where the latter was added to a concentration three times above its critical micelle concentration (CMC) to simulate a typical concentration used for MP storage. The SAXS data obtained from these nanodiscs were indistinguishable from nanodiscs prepared with cholate alone (Figure 18A). From this observation, it is concluded that the beneficial effects of cholate with respect to nanodisc reconstitution dominate the negative effect of DDM. Furthermore, this shows that the nanodisc platform is indeed suitable for insertion of MPs solubilized in DDM, as long as the MP is tolerant to the presence of 20 mM cholate. The same overall results were obtained for another popular detergent for MP handling, n-octyl-β-D-glucoside (OG).

3.1.1.2 *Nanodisc self-assembly is fast and likely proceeds from a pre-aggregate with the correct stoichiometry*

Part of the study in Paper I was inspired by studies of vesicle self-assembly (Lasic 1988; Lévy et al. 1990; Egelhaaf and Schurtenberger 1999). Like nanodiscs, vesicles can be formed from detergent depletion of mixed micelles. The rate of detergent removal is a critical parameter for controlling the size of the final vesicles as intermediate bilayer fragments are stabilized by residual detergent and allowed longer time to

grow (Lasic 1988). In effect, increased detergent removal rate leads to small vesicles and *vice versa*. For nanodiscs, the detergent removal rate was investigated by varying the amount of detergent absorbing beads added to the samples. In contrast to vesicles, the final nanodisc structure was completely insensitive to this parameter. This indicates that nanodisc formation is fast and different from the mechanism of vesicle formation.

Modulating the lipid:MSP ratio has been shown to have a major effect on the size of formed nanodiscs, as commonly evaluated by SEC (Bayburt et al. 2002; I. G. Denisov et al. 2004; Grinkova et al. 2010). In Paper I, this effect was further investigated for MSP1D1-POPC nanodiscs by a more detailed structural analysis by SAXS. First, it was found that the POPC:MSP1D1 stoichiometry saturated at a value of approximately 60:1 if the reconstitution was initiated with a POPC:MSP1D1 ratio of 70:1 or above. Starting below 70:1 resulted in the formation of "under-loaded", *i.e.* lipid-poor, nanodiscs. "Under-loaded" nanodiscs were characterized by a low lipid:MSP stoichiometry and a high axis ratio, *i.e.* an overall elongated shape, and the obtained SAXS pattern resembled that obtained for the nanodiscs prepared with DDM (Figure 18A, green data). That "under-loading" is possible indicates that the MSP is flexible and capable of adapting its shape according to the number of lipids inside. Starting from a stoichiometry above this optimum, nanodiscs of the expected shape were still formed, but together with larger species containing MSP and lipids, as evaluated by SEC. Overall, it was concluded that nanodiscs with a POPC:MSP ratio of approximately 60:1 are obtained if the stoichiometry in the reconstitution mixture is at or above the optimum of 70:1.

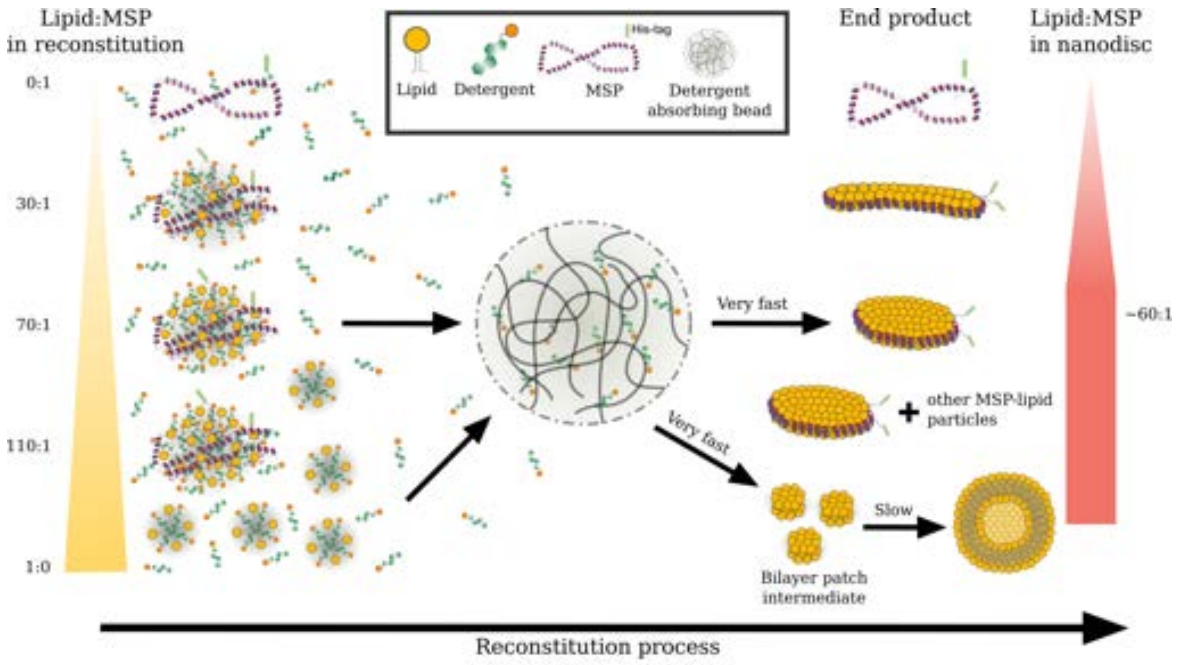


Figure 19 Overview of nanodisc self-assembly. Lipid-detergent mixed micelles are mixed with MSP in a given ratio. Starting at an optimal lipid:MSP ratio of 70:1, nanodisc are formed upon detergent removal independent of the rate of removal. Starting from below this optimum, "under-loaded" nanodiscs are formed. Starting above the optimal ratio leads to nanodiscs together with larger species. In the extreme case with no MSP present, vesicles are formed. Because of a bilayer patch intermediate, the size of the formed vesicles depend on the detergent removal rate. The figure is adapted from Paper I.

Figure 19 summarizes the reconstitution process based on the above observations. Here, it is proposed that nanodisc self-assembly likely starts from a pre-aggregate structure of MSP, lipid and detergent that has the overall same lipid:MSP stoichiometry as the nanodisc formed after detergent removal. The MSP puts a constraint for the maximum number of lipids to go inside the nanodisc, but it is also highly flexible towards loading fewer lipids, which leads to the elongated shape of the lipid-poor nanodiscs.

The initial stoichiometry in the nanodisc reconstitution mixture is commonly based on mixing lipids in a well known concentration with MSP, whose concentration has been estimated by UV absorption measurements at 280 nm and a molar extinction coefficient calculated from its primary sequence (Gasteiger et al. 2005). However, this concentration estimation is associated with a number of systematic errors. First, MSPs are produced in *E. coli* and extracted using standard protein purification schemes: Obtaining 100% purity is for all practical purposes impossible. Second, the molar extinction coefficient is only truly valid for completely unfolded proteins. It is therefore recommended to add a small excess of lipids, because a stoichiometry above the optimum lipid:MSP ratio will result in mainly nanodisc of the correct size and shape.

3.1.2 *Optimization of nanodiscs*

There are some pitfalls of the nanodisc as a MP carrier, in particular its conformational instability at elevated temperatures. Heated to physiologically relevant temperature, the nanodisc structure disintegrates and other, larger MSP-lipid aggregates are formed (Ilia G. Denisov et al. 2005). This can be illustrated with experimental SEC data obtained for MSP1D1-POPC nanodiscs: After reconstitution and detergent removal, the SEC purification of nanodiscs reveals a symmetric and well-resolved peak (Figure 20A). Incubation of the sample collected from this SEC (Figure 20A, blue area) for three days at 40°C leads to the formation of several species (Figure 20B), including a relatively well-defined peak at around 11 ml. The species from that peak retains its overall size (Figure 20C), *i.e.* it doesn't re-equilibrate back to the nanodisc structure. This implies that the nanodisc is not the thermodynamically most stable structure. But this is also not surprising as the lipids are not expected to be fully matched to the hydrophobic height of the MSP belt (Skar-Gislinge et al. 2010): It is highly likely that the formed structures are larger bilayer-like structures in which the lipids are more adequately packed. However, this effect is problematic in relation to structural (and functional) studies of MPs inside nanodisc. In this section, the development of stable and less aggregation prone nanodiscs is described.

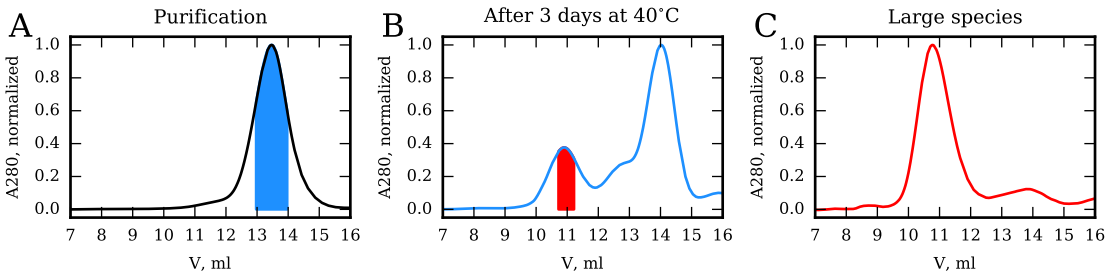


Figure 20 Nanodiscs form larger species at elevated temperature. MSP1D1-POPC nanodiscs were prepared by the standard protocol (Ritchie et al. 2009). **A:** Initial SEC purification of nanodiscs. The blue area indicates the collected sample of monodisperse nanodiscs. **B:** The sample from A was incubated for three days at 40°C and analyzed by SEC. A sample (red area) was collected from the peak at 11 ml. **C:** The sample from B was analyzed by SEC, where it retained its elution volume. SEC was performed on a Superdex 200 10/300 GL column in 20 m*M* TrisHCl pH 7.5, 100 m*M* NaCl.

3.1.2.1 *Increasing the conformational stability of nanodiscs by mutation and circularization*

In this thesis, stabilization of nanodiscs was initially explored by two different approaches. First, MSP1D1 was mutated to contain more negative surface charges to achieve repulsion between nanodiscs and thereby decrease aggregation propensity and enhance solubility (Report I). Specifically, the mutations were all Asn to Asp, all Gln to Glu, and all Thr to Ser, which have been shown to increase the solubility of small, globular proteins (Trevino et al. 2007; Kramer et al. 2012). Second, covalent circularization of the N- and C-termini (known to increase protein stability (Antos et al. 2009)) of MSP by sortase A was made possible by insertion of relevant N- and C-terminal tags. During this work, another group developed circularized nanodiscs and showed that they had the expected increased stability (Nasr et al. 2017). Following this development, a MSP containing both the solubility-enhancing mutations and circularization tags was designed (Paper III). To accommodate the relatively large bacterial magnesium transport protein CorA (200 kDa), it was designed based on MSP1E3D1, which makes nanodiscs at about 12 nm in diameter (I. G. Denisov et al. 2004). The designed MSP is named His-tagged linear and soluble MSP1E3D1 (His-lsMSP1E3D1) and is shown in a sequence alignment with MSP1E3D1 in Figure 21A. The circularization process involves three major steps (Figure 21B). First, His-lsMSP1E3D1 is cleaved with TEV protease to yield linear-soluble MSP1E3D1 (lsMSP1E3D1), which has an N-terminal Gly residue that is needed for circularization. Second, lsMSP1E3D1 is circularized by evolved sortase A (eSrt), utilizing the N-terminal Gly and the C-terminal LPETG motif, to yield circularized-soluble MSP1E3D1 (csMSP1E3D1). Finally, csMSP1E3D1 is separated from non-circularized substrates and enzymes by reverse affinity chromatography. The intermediates and product of this reaction were readily visualized with SDS-PAGE (Figure 21D). The final product was verified by mass spectrometry as well, and this analysis showed the expected mass of csMSP1E3D1 (Figure 21E) together with the presence of the unique linker-peptide owing to covalent circularization (Figure 21F).

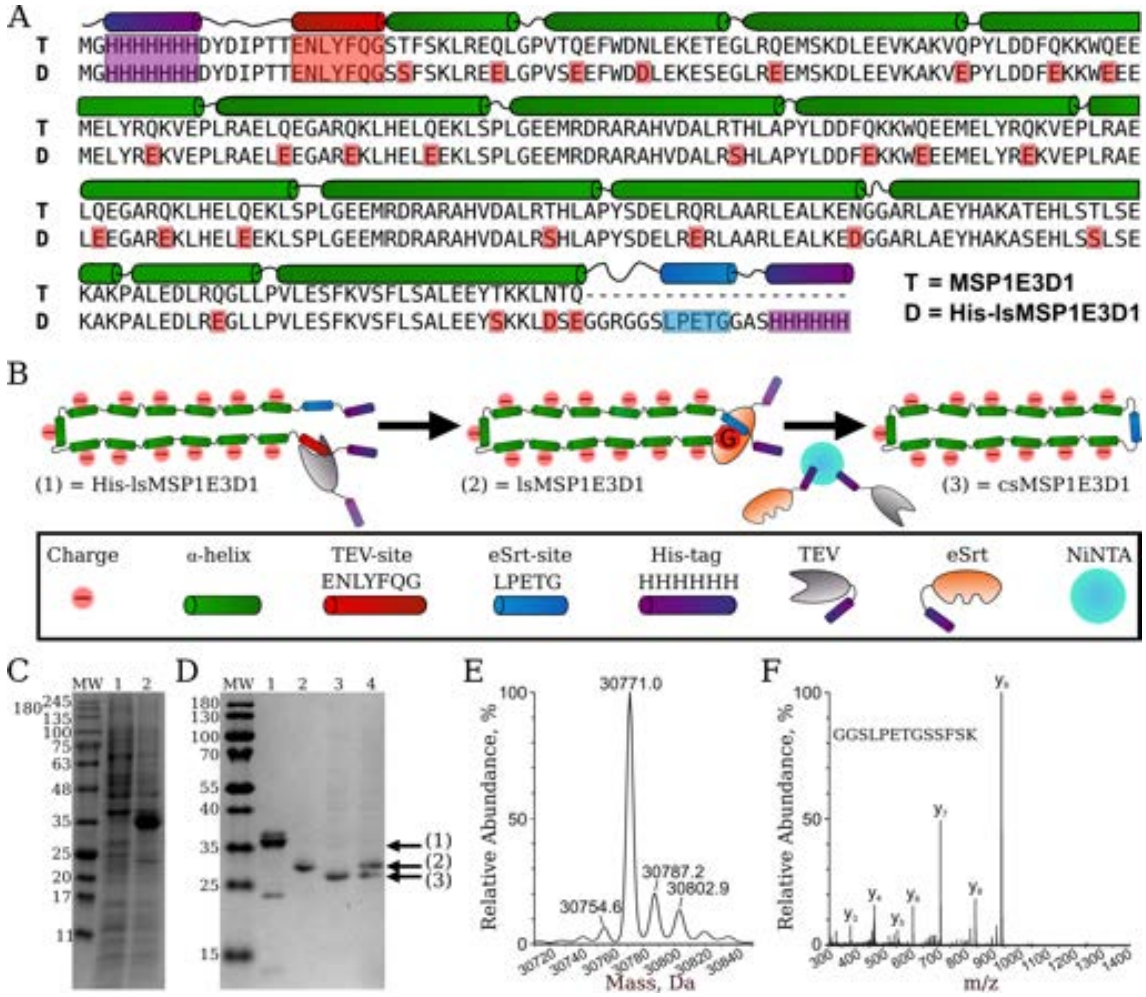


Figure 21 Solubility- and stability enhanced nanodiscs. **A**: Sequence alignment of the designed (D) His-lsMSP1E3D1 to the template (T) MSP1E3D1. Important sequence elements (see legend in **B**) are highlighted. **B**: Schematic of the circularization reaction. His-lsMSP1E3D1 (1) is cleaved by TEV protease and becomes lsMSP1E3D1 (2). lsMSP1E3D1 undergoes circularization by evolved sortase A (eSrt) to become csMSP1E3D1 (3). **C**: SDS-PAGE analysis of protein expression in *E. coli*. MW = molecular weight, lane 1 = total cell extract before expression, lane 2 = total cell extract after expression. **D**: SDS-PAGE analysis of protein purification and circularization. Lane 1 = Affinity purification of His-lsMSP1E3D1, lane 2 = lsMSP1E3D1, lane 3 = csMSP1E3D1, and lane 4 = lsMSP1E3D1 and csMSP1E3D1 together. **E**: Intact mass analysis by mass spectrometry. The mass of 30771 corresponds to csMSP1E3D1. The two heavier side peaks corresponds to oxidized csMSP1E3D1 (Paper III). **F**: Identification of the unique peptide of csMSP1E3D1 resulting from covalent circularization. The figure was adapted from Paper III.

3.1.2.2 *Expression, circularization and purification of csMSP1E3D1 is straight-forward with high yield*

Because reconstitution of MPs in nanodiscs is often difficult, it is desired to be able to readily obtain high yield of MSP to be able to carry out multiple tests experiments without being limited by the availability of the MSP. In that way the focus can be put on MP expression and purification that is more difficult. In general, MSPs express very well in *E. coli*, with reports of up to 250 mg per liter culture (Bayburt et al. 2002), and can be purified in good yield. However, the extra purification steps involved with the circularization process compared to standard MSPs are expected to have a negative influence on the final yield. In Paper III, His-lsMSP1E3D1 was found to be highly overexpressed in *E. coli* (Figure 21C). Furthermore, the charged mutations had a positive effect on its solubility, which made affinity purification, circularization, and ion-

exchange chromatography straight-forward with high yield. The final yield of csMSP1E3D1 was 75 mg per liter of bacterial culture. As a standard sample preparation of 500 μL csND requires approximately 2 mg csMSP1E3D1, this amount allows for several testings and optimization of sample preparation. Furthermore, this was a clear improvement to previously described circularized MSPs, where high amounts of the detergent Triton X-100 was required in the circularization reaction to achieve only modest yield (approximately 12 mg per liter culture for circularized MSP1E3D1) (Yusuf et al. 2018).

3.1.2.3 *Circularized and mutated nanodiscs retain the expected nanodisc structure*

Nanodisc could be readily obtained following the standard protocol (Figure 17) with POPC lipids, cholate and csMSP1E3D1 (csND) or lsMSP1E3D1 (lsND), respectively, as evaluated by well-defined SEC-peaks (Figure 22A). In Paper III, their structures were investigated in detail using SAXS and modeling. Overall, the structures of csND and lsND were quite similar, as readily seen from the SAXS data (Figure 22B). The pair-distance distributions ($p(r)$-distributions) refined from the SAXS data revealed that the lsND was just slightly larger than the csND (Figure 22B, insert). Furthermore, the modeling showed that the lsND contained slightly more lipids than the csND, which is in good agreement with the idea that the flexibility of the MSP is a determinant for the maximum lipid capacity in nanodiscs, as discussed in both Paper I and Paper III.

Notably, the above observations are conflicting with previous studies on cNDs. First, the diameter of other similar cNDs were systematically found to be 1-2 nm larger than non-circularized nanodiscs as evaluated by electron microscopy (EM) (Miehling et al. 2018). Second, non-circularized nanodiscs were generally found to be more polydisperse than cNDs, also evaluated by EM (Nasr et al. 2017; Miehling et al. 2018). At this point, more comparative experiments, *i.e.* EM data on the csND and the lsND, are required to make an optimal comparison. However, it is intriguing to speculate that the extra negative surface charges on the lsND and the csND simply have a positive effect on their structural dispersity.

3.1.2.4 *csNDs are conformationally stable at physiologically relevant temperature*

Paper III also describes the stability of the csND compared to the lsND and MSP1E3D1 nanodiscs. Using a combination of static light scattering (SLS) (Figure 23A) and SEC (Figure 23B), csNDs were shown to remain homogeneous at 37°C over a period of at least three days. Interestingly, these data also showed that the increased conformational stability was governed by both the charged mutations and the covalent circularization. Specifically, this was evaluated by the appearance of larger species at elevated temperature for lsNDs and MSP1E3D1 nanodiscs, with the latter already disintegrating significantly at 37°C, as evaluated by SEC (Figure 23B). Based on these observations, an energy landscape was proposed for the nanodisc (Figure 23). In this schematic, an intermediate state of MSP-lipid aggregates is labeled ”fused disc” under

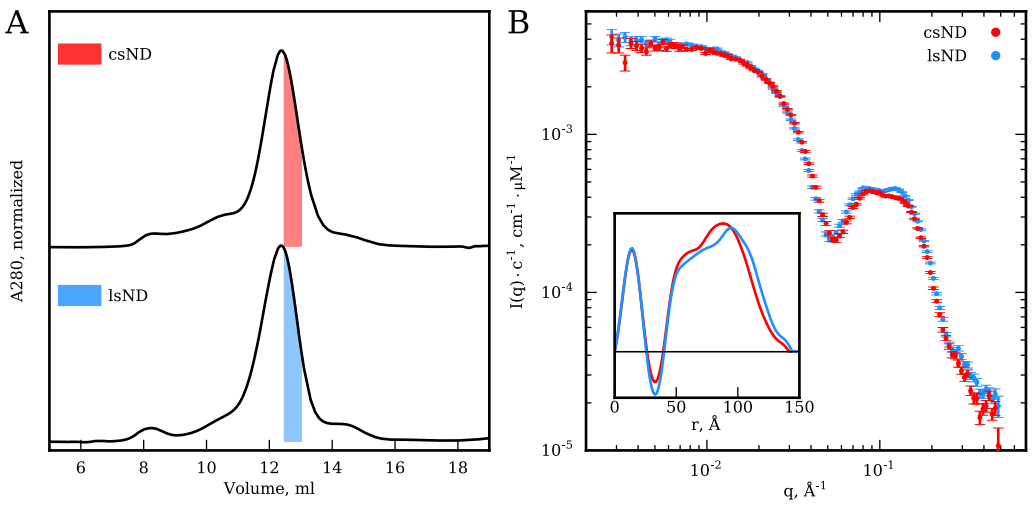


Figure 22 Structural characterization of optimized nanodiscs. **A:** SEC purification of nanodiscs made from POPC and csMSP13D1 (csND, red) or POPC and lsMSP1E3D1 (lsND, blue). **B:** SAXS data collected on the SEC fractions indicated in **A**. $p(r)$-distributions obtained by indirect Fourier transform (IFT) are shown in the insert. The figure was adapted from Paper III.

the assumption that large MSP-lipid aggregates can be formed by a fusion process. Furthermore, the lowest energy minimum is described as a vesicle-like structure, where lipids are optimally packed, but still with some MSP interaction. It should be noted that both of these structures were not experimentally investigated. Interestingly, the intermediate particles (fused discs) were not observed for the csND even after incubation at 50°C for three days (Figure 23B). It is indeed difficult to imagine a fusion process of two circularized nanodiscs because the MSP remains covalently closed at the rim of the lipid bilayer. At 50°C, some very large aggregates, *i.e.* vesicle-like structures, were observed for the csND, although at fairly low abundance, as evaluated by SEC (Figure 23B). Not surprisingly, this shows that nanodiscs can disintegrate and aggregate via other mechanisms than fusion.

The high conformational stability of csNDs is an obvious advantage for functional and structural studies of nanodiscs and MPs inside nanodiscs at physiologically relevant temperatures. An example is the study of lipid exchange, which has previously been determined for nanodiscs at temperatures ranging from 27°C to 42°C (Nakano et al. 2009). These samples were thus highly unlikely to remain homogeneous throughout the experiment, which could have had an effect on the extracted thermodynamic parameters. Stable csNDs should provide an optimal bases for studying lipid exchange in nanodiscs at these temperatures, and preliminary data obtained at KWS-1 at MLZ, Munich during this thesis indeed looked promising. However, these data are not included as more neutron beam time is required to obtain a full data set.

50°C is for many applications well-above the desired value as many MPs are expected to be significantly destabilized at this temperature. It is noteworthy, though, that recent NMR studies on MPs in nanodiscs were carried out at 45°C (Section 1.3.4.1) (Hagn et al. 2013; Nasr et al. 2017). Obviously, the csND studied in Paper III is too large for NMR, but truncated versions of circularized and solubility-enhanced MSPs have since been developed in our group. The preliminary results in Figure 27B show that csNDs down to 8 nm in

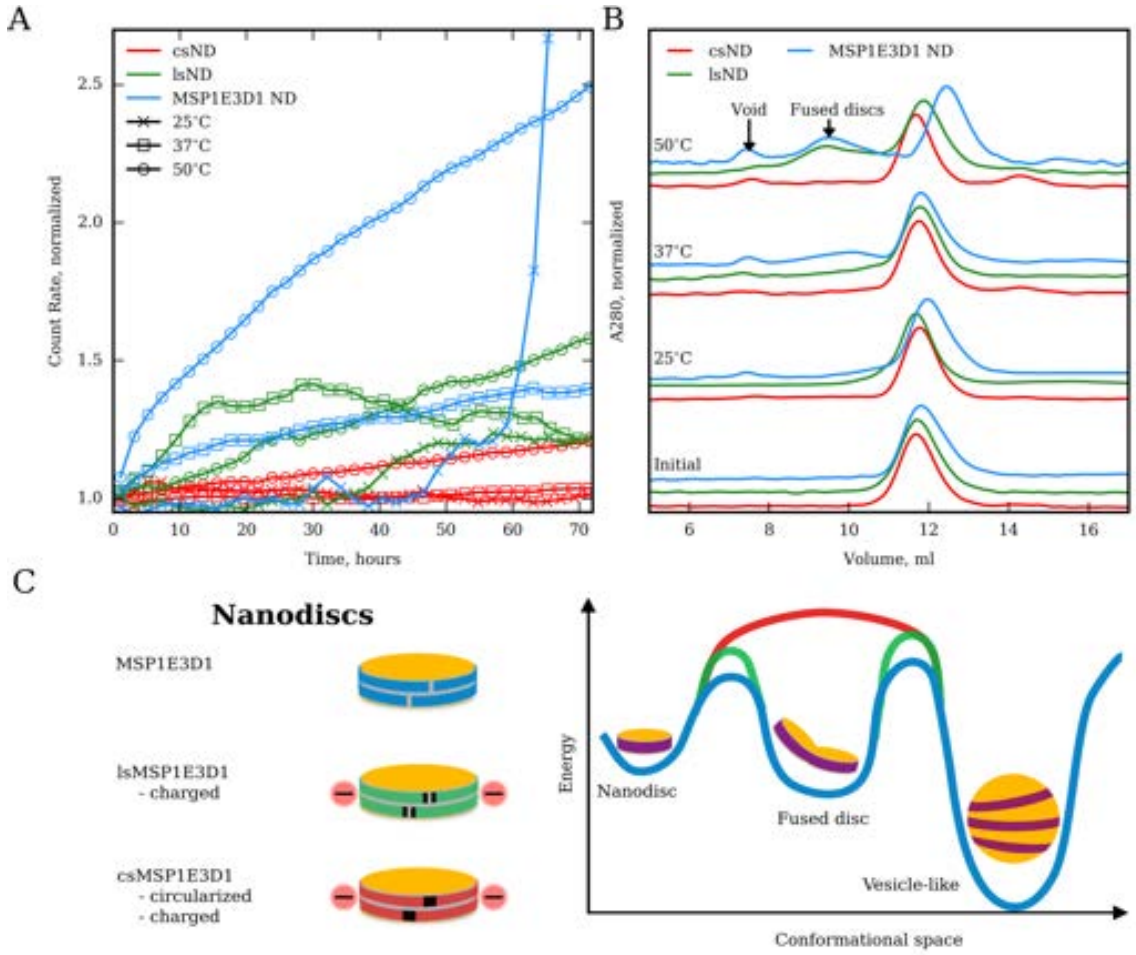


Figure 23 Temporal stability of nanodiscs. A: The structural integrity of three types of nanodiscs was monitored by SLS at 25°C, 37°C and 50°C. **B:** SEC was performed before (initial) and after the SLS measurements at the indicated temperatures. Arrows indicate the approximate elution volume of two species that formed besides the main species in the main peak. **C:** Proposed nanodisc energy-landscape based on the results in **A** and **B**. The csND (red) can not aggregate by disc-fusion and has the overall largest energy barrier to an aggregated, vesicle-like state.

diameter (corresponding to MSP1D1ΔH5 nanodiscs, Section 1.2.7.2) can be readily made, which are very likely to be compatible with NMR.

Overall, the studies focused on nanodisc development in this thesis have provided important insight to the nanodisc reconstitution procedure, together with the generation of a novel stability- and solubility-enhanced nanodisc. Overall, these efforts were made to facilitate structural studies of MPs in solution by SANS, which are described in the following section.

3.2 OBTAINING HIGH QUALITY SANS DATA ON DIFFICULT SAMPLES

The overall aim of this thesis was to obtain high quality SANS data on MPs in solution. A specific objective was to obtain monodisperse nanodisc samples in D_2O-based buffers, which had proved difficult by the standard methods used for preparing nanodisc samples in H_2O-based buffers (Huda et al. 2015). To

solve this D_2O problem, three major approaches were followed. First, the "heavy water effect" (HWE) was investigated under the hypothesis that D_2O induced aggregation could be a temporal phenomena proceeding after sample purification due to altered sample stability. Second, based on findings from the first approach, solubility-enhanced nanodiscs were proposed, which ultimately led to the development of the csND (Section 3.1.2). Finally, the feasibility of SEC-SANS for nanodisc samples was tested as a method to avoid aggregation. SEC-SANS was also later applied to obtain aggregation-free data on MPs in detergent micelles.

3.2.1 *Heavy water effects?*

Many biophysical and structural methods use D_2O to enhance signal. This is most often done under the assumption that H_2O and D_2O are chemically very alike. However, the stability, among other physiochemical properties, of macromolecules is governed by subtle balances between enthalpy and entropy and is highly dependent on the energy of solvation. Thus, it is not surprising that exchanging the light H atom for the heavier D will have some consequences, *i.e.* that there is a HWE.

However, it turned out to be quite difficult to probe a HWE on nanodisc (MSP1D1, POPC) samples. From the beginning of the project, SEC was utilized as the standard method to exchange solvents and the previously observed aggregation issues were effectively eliminated. Thus, to illustrate the previous problems, a data set taken out of context, but resembling previous data sets in D_2O, is shown in Figure 24A. Nanodisc aggregation is readily visualized by a steep low-q region compared to a monodisperse nanodisc. The calculated $p(r)$-distribution clearly shows that larger particles are present in the sample (Figure 24A, insert). Using a desalting column to exchange buffer, the SAXS profiles and corresponding $p(r)$-distribution were identical in H_2O and D_2O (Figure 24A), suggesting this method was mild and did not result in the formation of aggregates.

Returning to the aggregated samples, it seems that the high-q regions of the monodisperse and aggregated data sets in Figure 24A are similar, which in turn indicates that the finer nanodisc structure is preserved in the aggregated sample. Thus, D_2O was not expected to disintegrate the nanodisc, but rather to cause residual aggregation. To be firmly convinced that the HWE was minor, a large number of experiments were carried out to compare the behavior of nanodisc as well as simpler globular proteins in H_2O and D_2O.

3.2.1.1 *Protein stability increases linearly with D_2O content*

The literature holds many examples of differential behavior of soluble proteins in D_2O and H_2O. The overall consensus observations of the HWE on proteins are increased conformational stability (Efimova et al. 2007), reduced flexibility, a more compact structure (Cioni and Strambini 2002), and reduced solubility (Gripon

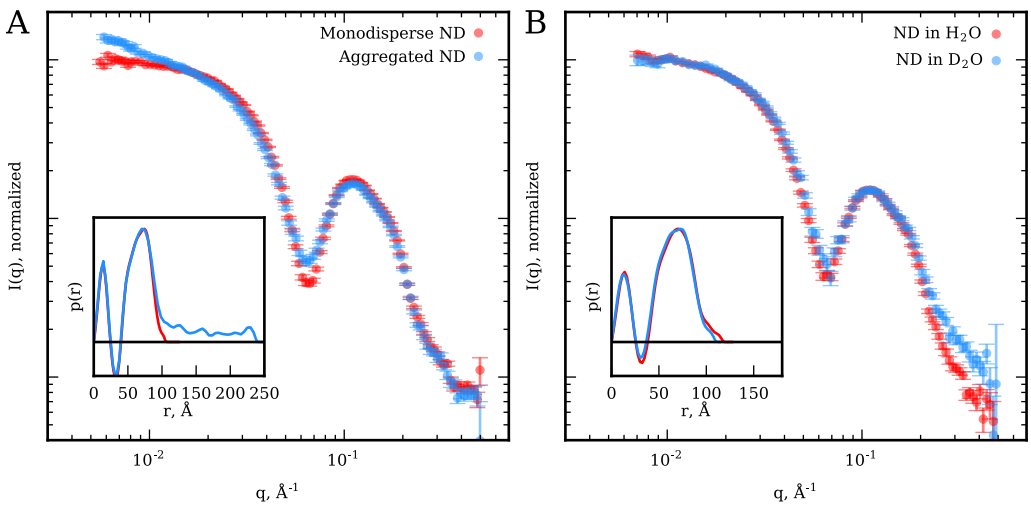


Figure 24 Aggregated nanodiscs. A: Representative SAXS data sets are shown for a monodisperse nanodisc (red) and an aggregated nanodisc (blue). The aggregated data set was obtained out of context, but resembles the overall trend of SAXS data obtained from nanodiscs in D_2O previous to this thesis. **B:** SAXS data obtained from a nanodisc sample in H_2O and from the same sample after exchange into D_2O buffer. Panel B was adapted from Paper I.

et al. 1997). In Report I, the HWE in terms of stability was probed on a set of soluble, globular proteins and nanodiscs using circular dichroism spectroscopy and differential scanning calorimetry. For the set of globular proteins, a clear increase of transition melting temperature (T_M) was observed and had a linear correlation to the amount of D_2O in the buffer (Figure 25). Nanodisc stability could not be readily quantified by thermal unfolding, in part because the secondary-structure stability was too high, and in part because the nanodisc structure disintegrated at elevated temperature, as illustrated for a similar nanodisc sample in Figure 20. Qualitatively, however, these results indicated that there was little to no cause to be concerned that the observed aggregation in D_2O was because of decreased stability.

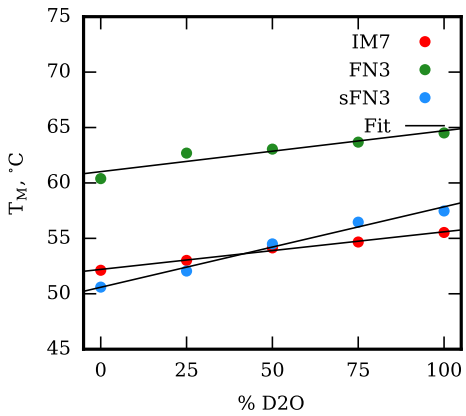


Figure 25 T_M of soluble proteins increases linearly with D_2O content in the solvent. T_M values were determined from thermal unfolding of three soluble, globular proteins. Unfolding of IM7 and sFN3 was monitored by differential scanning calorimetry. Unfolding of FN3 was monitored by circular dichroism spectroscopy. The figure was adapted from Report I.

3.2.1.2 *An increased aggregation propensity can not be appreciated for dilute nanodisc samples in D_2O*

The increased stability is somewhat contradictory to the increased aggregation propensity of nanodiscs in D_2O reported previously (Huda et al. 2015). In Report I, the problem was thus approached by a more practical and qualitative study in which nanodisc homogeneity was followed over long periods of time at 4°C, 20°C, and 40°C by SLS and/or SEC, similar to the studies on optimized nanodiscs (Figure 23A and B). The results obtained in H_2O and D_2O buffers were strikingly similar and showed that nanodiscs remained homogenous for several days at relatively low temperatures (4°C and 20°C). At 40°C, nanodiscs disintegrated and formed other aggregate structures, but again indistinguishably between H_2O and D_2O buffers. Thus, these results show that nanodisc samples prepared by SEC (or alike) in D_2O are qualitatively identical to nanodiscs prepared in H_2O buffers.

3.2.1.3 *Solubility-enhanced nanodiscs facilitate SANS experiments in D_2O*

Previously, solvent exchange was mainly carried out by serial concentration and dilution steps. Concentration was performed in spin filters, which allow solvent to go through but retain molecular species above a given size threshold. With this preparation technique, molecular species thus end up in a very high local concentration near the filter, which might be the main driver of aggregation. This lead to the hypothesis that increased solubility could alleviate D_2O induced aggregation. Negative surface charges have a strong positive influence on protein solubility (Trevino et al. 2007; Kramer et al. 2012), and thus, solubility-enhanced nanodiscs were proposed, including the csND (Section 3.1.2). Therefore, the focus was shifted from the standard MSP1D1 nanodisc to csND samples. Using SEC purification, monodisperse samples of empty csNDs were readily obtained in D_2O as evaluated by SANS (Figure 26, green data). The same was true for CorA-loaded csNDs (Figure 26, red and blue data), showing a general applicability of csNDs and SEC purification to obtain high quality MP samples for SANS. Furthermore, these samples were measured over a period of several hours to a day by static cuvette-SANS at 10°C without aggregating. Thus, the high conformational stability of these samples make them compatible with the long exposure times required for dilute or low-contrast samples in SANS.

3.2.2 *SAXS/SANS coupled to online SEC is an optimal method to obtain aggregation-free data*

Recently, a SEC-SANS setup was installed at D22 at the ILL, Grenoble, France, which proved feasible for measuring a set of soluble proteins despite the markedly lower flux of neutrons compared to X-rays at synchrotrons (Jordan et al. 2016). During this thesis, this setup was first tested for its feasibility for dilute MP samples (Paper I), and later utilized to obtain high quality data on important samples (Paper IV).

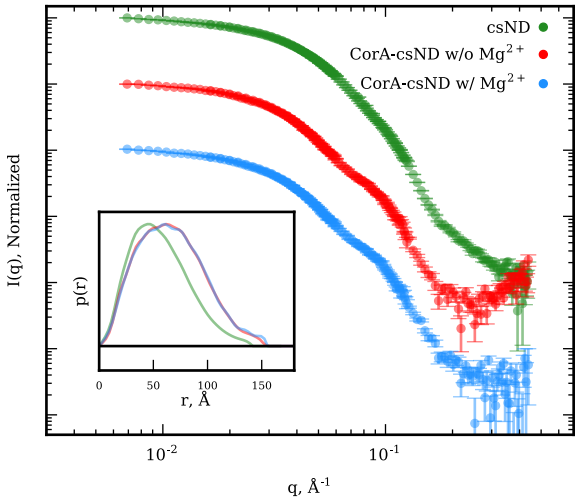


Figure 26 SANS data of csND and CorA-loaded csND measured in 100% D_2O buffer. The data was collected at FRM2, Munich, Germany and samples were prepared as described in Paper III but with SEC purification in D_2O buffer. Data are normalized and scaled for clarity. The upturn in the intensity at high-q was an artifact from parasitic scattering arising from a neighboring instrument.

3.2.2.1 *SEC-SANS is feasible for studying empty and MP-loaded nanodisc samples*

Alongside the development of the csND for SANS (Section 3.1.2.1 and 3.2.1.2), the feasibility of SEC-SANS was investigated on a number of nanodisc (MSP1D1/MSP1E3D1, POPC) samples, with and without MPs inside. Paper II describes the first of these efforts, carried out in close collaboration with Anne Martel and Lionel Porcar during development of the SEC-SANS as a fully integrated set-up at D22 (Figure 27A). The SEC-SANS measurements were carried out with slow flow rates and data were binned in 30 s frames to obtain better statistics. For empty nanodisc samples in typical concentrations, data of surprisingly good quality were obtained. The data were measured to very low q-values (0.004 $Å^{-1}$) and in all cases had well-defined Guinier regions, which indicated aggregate-free samples (Section 2.3.2.2). In the high-q region, the signal-to-noise ratio limited information to be obtained to approximately $q_{max} = 0.15$ $Å^{-1}$. Proteorhodopsin-loaded nanodisc samples were more difficult to study because several species were present in the sample prior to the SEC (Figure 27C). A combined set of frames with the overall expected size (as deduced from the $p(r)$-distributions) of a proteorhodopsin-loaded nanodisc could, however, be extracted, although the two major peaks were not completely resolved. To obtain a basis for model fitting in Paper II, complimentary SEC-SAXS data were collected on identical samples. Importantly, the quality of the SANS data was sufficiently high to impact the global model fits to the combined SEC-SANS and SEC-SAXS data set.

3.2.2.2 *SEC-SAXS and SEC-SANS reveal structural dispersion in nanodiscs*

In addition to the obvious benefits from aggregation-free data, the SEC-SANS (and SEC-SAXS) studies provided another interesting aspect to the understanding of the nanodisc structure. In both Paper I and

Paper II, it was shown that nanodiscs have a size distribution, related to their lipid:MSP stoichiometry, which is convoluted in the resolution of the SEC peak. In other words, nanodiscs are loaded with different amount of lipid, with the larger lipid-rich nanodiscs eluting from the SEC column before smaller lipid-poor nanodiscs. Thus, the size of nanodiscs was found to slightly but systematically decrease over the SEC-peak. Furthermore, in Paper II, the investigated SAXS/SANS frames had significantly smaller R_g and D_{max} values than nanodisc samples measured on other occasions with static SANS/SAXS (see *e.g.* Figure 24B) although they had been extracted from the same part of the SEC-peak. This indicates that nanodisc samples extracted from the middle or right hand side of a SEC-peak eventually equilibrate to a distribution of larger and smaller nanodiscs. In preliminary studies, this effect was probed for nanodiscs of different sizes (Figure 27B): This effect was seemingly more pronounced for larger nanodiscs, and almost not present for small 8 nm wide nanodiscs as determined by the decrease in R_g over the SEC-peak. Due to the multi-contrast situation in nanodiscs, R_g values are not easily interpreted and data are being modeled to gain a more detailed understanding of the phenomena together with data from lipid exchange studies.

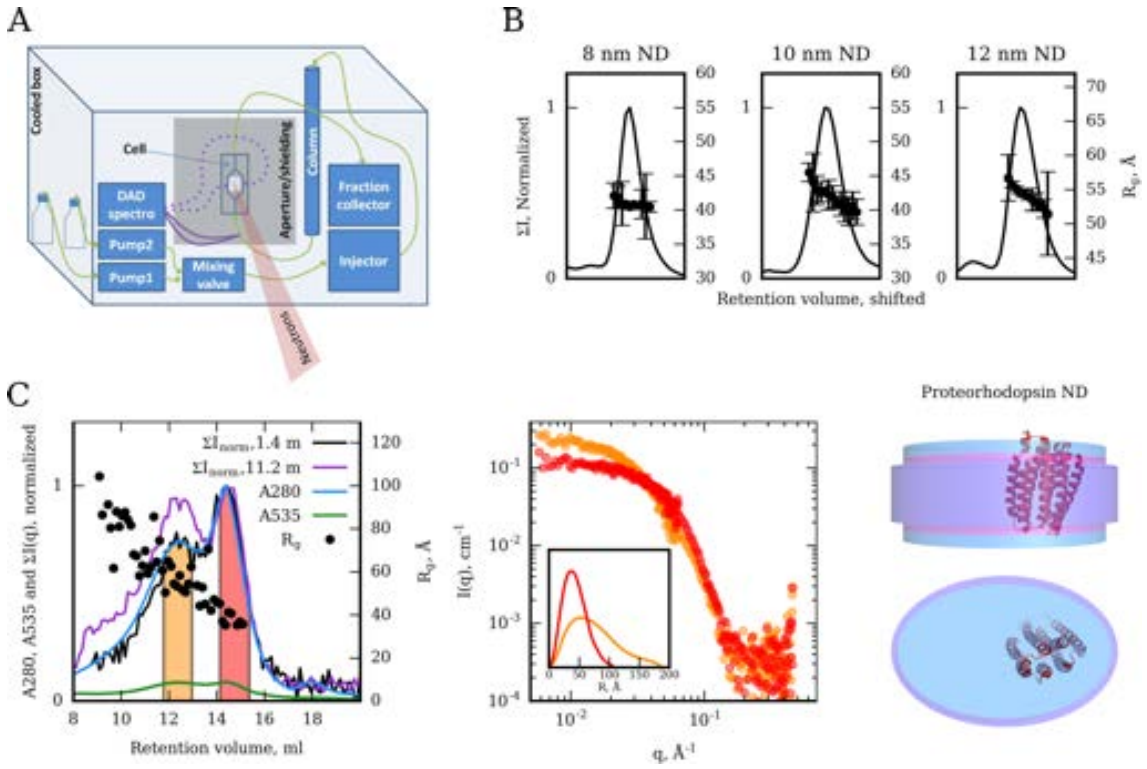


Figure 27 SEC-SAS. A: Schematization of the dedicated SEC-SANS setup at D22 at ILL at time of data collection for Paper I. **B:** SEC-SANS data obtained on proteorhodopsin-loaded nanodiscs. Left plot: Scattergrams and chromatograms together with initial estimations of R_g values. **C:** SEC-SAXS data on nanodiscs of varying size obtained at BM29 at ESRF. R_g values were estimated for each data frame over the peak. Panel **A** and **B** were adapted from Paper I.

3.2.2.3 *Limitations of the method*

Model refinement of the proteorhodopsin-loaded nanodisc provided two important lessons for SEC-SAS. First, this method can not be used for any sample; aggregates or larger species should be separable from the

species of interests on the column. In the case of the proteorhodopsin-loaded nanodisc, the larger particles present in the broad peak at 12 ml (Figure 27C) could perhaps be avoided by utilizing circularized nanodiscs, leading to a more convincing chromatogram where the nanodisc species is fully resolved. Second, some parameters of the Proteorhodopsin-nanodisc model (Figure 27) were refined to relatively nonphysical values, suggesting systematic errors in the data. The concentration estimation in the SEC-SAXS experiment was not ideal and was thereby expected to influence the model a lot. However, fitting of the concentration introduced too much freedom in the model. This shows the importance of good concentration estimations in SEC-SAXS, which is well-acknowledged (David and Pérez 2009; Graewert et al. 2015; Ryan et al. 2018), but not always available (Section 2.4.1.1). Another drawback is that SEC-SAS is very time consuming, and SEC-SANS do not provide optimal data statistics. Thus, if stable homogenous samples can be obtained by offline SEC, static measurements are recommended. That also allows for better measurements of concentration and for adding ligands or diluting the sample during an experiment. However, for many of the samples measured in this thesis, SEC-SAS had a major positive impact on data quality. Based on the expertise gained through the work done in Paper II, SEC-SANS was applied to study more difficult samples and proved to be a valuable tool for obtaining data from monodisperse species, *e.g.* CorA in detergent micelles (Section 3.2.3.2).

3.2.3 *Structural studies of CorA*

Following previous work (Huda et al. 2015), the pentameric bacterial magnesium transport protein CorA from *Thermogata maritima* was studied throughout this thesis. Besides being highly interesting on its own for reasons mentioned below, CorA was a good model MP which could relatively easily be inserted into nanodiscs. Thus, it was used to evaluate *e.g.* the ability of the newly developed csNDs to accommodate MPs in Paper III.

The structure of CorA has been thoroughly investigated, including several crystal structures (Figure 28A), and complimented by a wide variety of biophysical techniques. Notably, all crystal structures were obtained in presence of divalent cations and in a closed, highly symmetric state. During the initial period of this thesis, cryoEM structures emerged and provided models for open, Mg^{2+}-free, conformations of CorA (Matthies et al. 2016). The two major open conformations deposited in the PDB (*3JCG*, Figure 28A, and *3JCH*) are sufficiently different from the closed state that their theoretic scattering patterns are discernible, most strikingly at $q \approx 0.1$ Å (Figure 28B). Thus, the project was carried out under the hypothesis that different structural states could be observed by SAS. However, after much effort, scattering data obtained on CorA in nanodiscs was observed to be invariable between samples that were Mg^{2+}-free or in excess of Mg^{2+}, respectively. This was the case both in SAXS and SANS (SANS data set shown in Figure 26). In this

case, the nanodisc proved disadvantageous given its high contribution to the overall scattering, which could potentially be shielding the structural rearrangement of CorA inside it. The strategy was thus set towards contrast-optimization in SANS. The data discussed in the following sections are presented in Paper IV and build in different ways on the results obtained throughout Paper I, Paper II, Paper III and Report I.

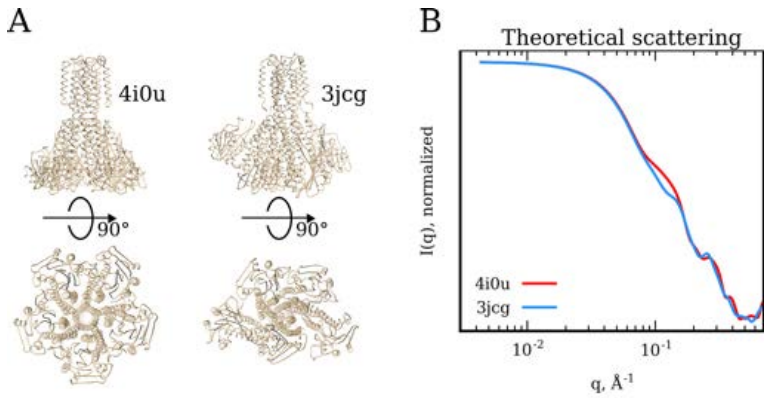


Figure 28 Structural states of CorA. A: Representative structures of the closed (PDB ID: *4I0U*) and open (PDB ID: *3JCG*) CorA, respectively, shown in side and top views. **B:** Theoretical SANS for the two structures shown in **A** (including a water layer 6% denser than bulk).

3.2.3.1 *Development of a stealth nanodisc with chemically pure lipids*

Deuterated carrier systems matched to the scattering length density (SLD) of D_2O (stealth carriers; see principle in Section 2.2.2) provide an optimal contrast situation for MPs in SANS. Here, two such systems were investigated in parallel, namely stealth DDM (sDDM) (Midtgaard et al. 2018) and stealth nanodiscs (sND) (Josts et al. 2018). sDDM was used as described in the literature, while novel sNDs were developed in this thesis. Previous studies on sNDs made with biologically deuterated PC lipids show that complete contrast-match was not obtained (Maric et al. 2014; Josts et al. 2018), and these lipids were furthermore difficult to produce and purify (Maric et al. 2015). To improve on this, chemically synthesized and pure, deuterated POPC (dPOPC) was obtained from the NDF at ANSTO, Australia. In a first attempt, deuterated MSP1D1 (Maric et al. 2014) was utilized to prepare the sNDs. However, sample preparation, in part due to low amounts of the very costly dPOPC, proved difficult and did not allow collection of good quality data. Instead, a deuterated version of the optimized MSP described in Paper III and Section 3.1.2.1 (d-csMSP1E3D1) was obtained from the deuteration lab at ILL (Dunne et al. 2017). sNDs prepared with the chemically pure dPOPC and d-csMSP1E3D1 were well matched to the SLD of D_2O as evaluated by a contrast-variation series measured with SANS (Figure 29A and 29B). Importantly, the sND was uniformly contrast-matched ($I(q) \sim 0$) to D_2O (Figure 29A, insert). Furthermore, a complimentary SAXS data set verified that the sND had the expected nanodisc structure (Figure 29C). The ease of producing csMSP1E3D1, also in the deuterated version, together with the availability of chemically pure dPOPC holds great promise for the sND as a platform for future SANS studies of MPs.

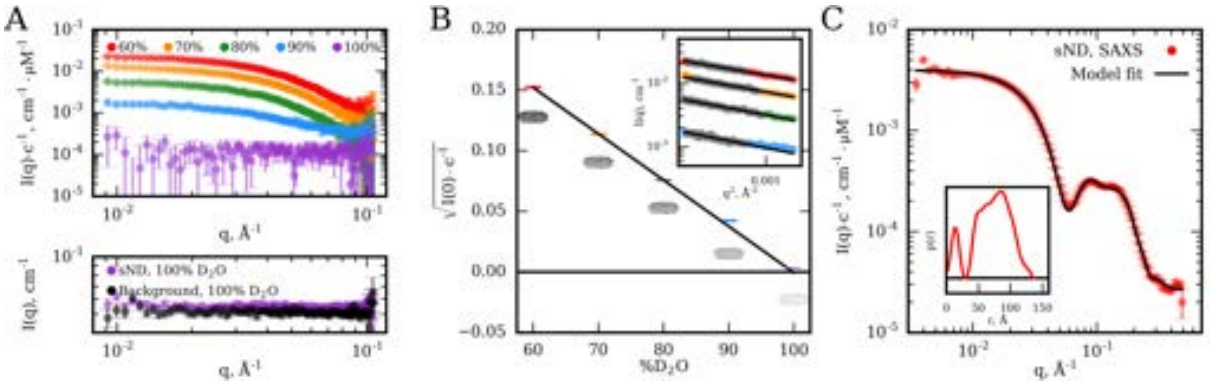


Figure 29 Stealth nanodiscs. **A:** Contrast variation series of sNDs prepared with chemically synthesized dPOPC and d-csMSP1E3D1. The lower panel shows the signal from the sND at 100% D_2O before background subtraction, together with the corresponding background measurement. **B:** Plot of $\sqrt{I(0) \cdot c^{-1}}$, which is proportional to the contrast of the sND, vs the D_2O content. The intercept with the primary axis corresponds to the contrast-match point. $I(0)$ was determined from Guinier analysis (insert). **C:** SAXS data from the sND. The figure was adapted from Paper IV.

3.2.3.2 *The solution structures of CorA in presence and abscence of Mg^{2+} are highly similar*

In Paper IV, CorA was measured by SANS in both sDDM and sNDs. Sample preparation of CorA in sNDs followed the standard protocol for nanodiscs reconstitution (Figure 17), but with dPOPC and d-csMSP1E3D1 followed by SEC in D_2O buffer (Figure 30A). For sDDM, sample preparation was simpler and only required a SEC step into D_2O buffer containing the sDDM (Figure 30B).

When investigating CorA in sDDM micelles, it was found that the SANS data obtained for samples with and without Mg^{2+} present were identical. This was highly surprising given the magnitude of the proposed structural rearrangement between Mg^{2+}-bound and free states in CorA (Figure 28B). However, these SANS data also indicated that both CorA pentamers and dimers of CorA pentamers were present in the samples, which could potentially have smeared out the signal change from open and closed conformations. By later utilizing SEC-SANS at D22, SANS data from solely CorA pentamers could be obtained.

The degree of signal elimination by using stealth carriers is exemplified by comparing the SEC-SANS data obtained on CorA in sNDs (Figure 30C, green data) to SANS data obtained on CorA in hydrogenated csNDs (Figure 30C, red data). The SANS data obtained on CorA in sDDM was overall similar to that obtained in sNDs, but with slightly less intensity at low-q (Figure 30C, blue data). With such efficient signal elimination of the carrier system, the obtained data were highly surprising: The major finding in Paper IV was that the SANS data sets collected on CorA samples with and without Mg^{2+} present were identical (Figure 30D). This was the case both in sDDM and sNDs and using SEC-SANS. Thus, the lack of a visible change in scattering intensity could not be ascribed to the presumably non-native environment in the sDDM micelles compared to the lipid bilayer environment in the nanodisc (Section 1.2.1).

Given that Mg^{2+} is the most hydrated divalent cation in biology and interacts with CorA via its first hydration shell, concerns were raised about a potential HWE. This was investigated by complimentary electron microscopy measurements of CorA in DDM micelles and fluorometric activity assays in large uni-

lamellar vesicles of POPC to verify the structural and functional integrity of CorA in D_2O. The results of these studies supported that CorA had a pentameric structure and was active in D_2O. Thus, it is concluded that the SANS data firmly show that the structural changes in CorA upon Mg^{2+} binding and release are smaller than suggested from cryoEM structures: The solution structures of CorA are overall similar with and without Mg^{2+} present. Also surprising, however, none of the currently available structural models of CorA could fit the data adequately throughout the recorded q-range (Figure 30D, fits).

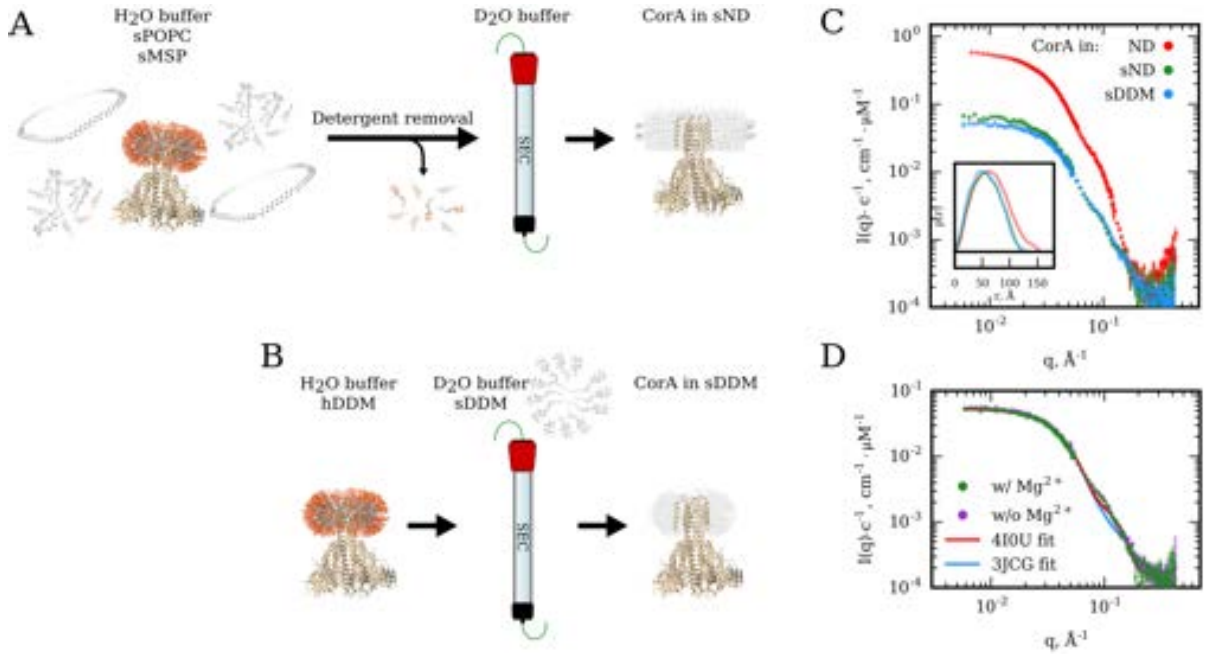


Figure 30 CorA in stealth carriers. **A:** Overview of sample preparation of CorA in sNDs. CorA solubilized in hydrogenated DDM is mixed with dPOPC solubilized in cholate and d-csMSP1E3D1. After detergent removal, the sample is purified by SEC in a D_2O-based buffer. **B:** Overview of sample preparation of CorA in sDDM micelles. CorA solubilized in hydrogenated DDM is purified by SEC in a D_2O-based buffer containing sDDM, using a slow flow rate (< 0.3 ml/min) to ensure full exchange into sDDM. **C:** Comparison of SANS data obtained on CorA in hydrogenated csNDs (red), CorA in sNDs (green) and CorA in sDDM (blue). The insert shows $p(r)$-distributions refined from the SANS data. **D:** Comparison of SANS data obtained on CorA in sDDM with (green) and without (purple) Mg^{2+} added. Solid lines are the fits of the closed crystal structure (PDB ID *4I0U*, red) and an open cryoEM structure (PDB ID *3JCG*, blue).

3.2.3.3 *Contrast variation and SEC-SAS provides an optimal basis for structural analysis of MP solution structures*

DDM micelles are expected to be highly dynamic and thereby somewhat polydisperse (Section 1.2.1), while the csND is fairly large compared to CorA in order to accommodate a lipid bilayer (Paper III). Using stealth carriers, only the signal originating from MP inside is measured with SANS in D_2O. Thus, the use of contrast-matched carrier systems for SANS makes data analysis by well-established programs for structural analysis of proteins, such as the ATSAS package, possible. Recently, *ab initio* bead-modellig was used for a large number of MPs measured in sDDM (Midtgaard et al. 2018). In Paper IV, bead-models reminiscent of the expected structure of CorA could only be obtained when applying five-fold symmetry as a constraint.

Furthermore, rigid-body modeling revealed nonphysical structures (data not shown). The reasons for this remain unclear, but future analysis will predominantly be focused on other approaches. Currently, efforts are made towards combining the obtained SANS data with molecular dynamics simulations to retrieve an ensemble of structures that describe the data.

Although the structural states of CorA in solution remain elusive, the general work related to optimizing samples for SANS together with the development of SEC-SANS holds great promise for future structural studies of MPs in solution. In particular, the combination of SEC-SANS and sDDM provides a straightforward technique to study many different MPs under solution conditions and obtain good, aggregation-free data. However, obtaining data with good signal-to-noise ratio at high-q might in general be difficult with SEC-SANS due to the maximum loading capacity of standard SEC columns. If this data is needed, for instance if a lipid environment has been well-established to be important for MP activity, the stable sND developed in Paper IV provides a better, but more expensive, alternative.

3.2.3.4 *Other carrier systems?*

It might be interesting to test other carrier systems that bind directly to the MP, thus forming a more stable environment than DDM micelles and a smaller particle than the nanodisc. Because of these properties, Salipro (Section 1.2.6) have been used by others to study MPs in solution using only a SAXS contrast (Flayhan et al. 2018). Other carrier systems should have similar properties, such as the NSP$_r$ peptidisc (Section 1.2.5.1), lipopeptides (Section 1.2.5.2) and perhaps some specific versions of amphipols (Section 1.2.4.1). Finally, an MSP with a tailor-suited size to a specific MP could be designed. As an example, the transmembrane domain of CorA is around 4 nm in diameter and should thus fit into much smaller nanodiscs than the 12 nm wide csND. Previous studies have shown that CorA does not incorporate well into MSP1D1-based nanodiscs and form particles larger than expected, likely because the MSP is flexible and allows more than one pentamer to insert. However, with the csNDs in different size briefly mentioned in Section 3.1.2.4, such larger particles should not be possible to form, and it might thus be possible to obtain well-defined small CorA-loaded nanodiscs.

3.3 THE CONTRIBUTION OF SAS IN STUDYING MP STRUCTURE

”Integrative structural biology” is the popular term for combining several different structural and biophysical techniques to refine models of macromolecules (Sali et al. 2015). Here, SAS can provide low-resolution distance constraints together with the overall shape and size of molecules in solution. The solution conditions and the fact that all excess SLD contribute to the intensity must be further highlighted as an advantage of SAS. The data presented on CorA in sDDM and sNDs (Paper IV) highlight that crystal structures are often

only representing a single, very stable state of a protein. The same might to some extent be true in cryoEM, which rely on aligning particles and making sequential selection of models. In particular, it is a cause of concern that five-fold symmetry was applied in model refinement for CorA to obtain better resolution (Matthies et al. 2016). Structural characterization by SAS is important to validate these models in solution.

Intrinsically disordered protein regions (IDRs) are common and have gained much attention (Lee et al. 2014), but have not been explored in this thesis. Notably, many MPs posses intra- and extracellular IDR's, *e.g.* G-protein coupled receptors (GPCRs) (Rasmussen et al. 2007). These MPs are very difficult to crystallize and cryoEM is not expected to resolve the highly flexible IDRs. NMR, on the other hand, is optimal to study IDRs, but the inherent size limitation of the technique makes it difficult to the full length MPs stabilized in carrier systems. Consequently, SAS is an ideal complimentary method to combine high resolution sub-models into a full length model. This, of course, also applies when adding interaction partners. Here, selective deuteration strategies and SANS can provide crucial information to refine the relative orientation and position of the different molecules involved in the complex. The developed contrast-optimized MP carriers, as demonstrated in Paper IV, should provide the optimal basis for studying difficult MPs by SANS.

From a practical point of view, it is possible to test different solvent conditions with relative ease using SAXS. For instance, it might be important to know the oligomerization state of a protein as function of concentration, ionic strength, pH, and temperature. It could further be expanded to validate MPs in solution when stabilized in a wide array of different carrier systems. As a first study, MP activity could be correlated with structural integrity in specific carrier systems. When promising conditions are found, the analysis could be expanded with SANS to obtain even better structural models. Thus, SAXS and SANS have general applicability for the development of new and better carrier systems, like it has been shown in Paper III for the csND.

3.4 CONCLUSION

In the conclusion, the specific aims listed in the Thesis Outline that framed the overall work packages of this thesis are addressed.

For Aim I, the most important objective was to circumvent aggregation of nanodisc samples in D_2O. In Report I, the "heavy water effect" (HWE) was probed via temporal studies of aggregation together with stability measurements. These studies showed that there was not an appreciated effect from D_2O under the investigated conditions. Also in Report I, the second objective of Aim I was fulfilled by introducing solubility-enhancing mutations in MSP1D1 (sMSP1D1). Learning from that study that sMSP1D1 could form nanodisc equally well to standard MSP1D1 led to the development of solubility- and stability-enhanced nanodiscs (csND) in Paper III. The stability-enhancement was achieved via covalent circularization with

sortase A, and had the effect that the csND remained monodisperse at physiological temperature. The last objective of Aim I was to gain a general understanding of the HWE through studies of soluble proteins. While preliminary data were reported in Report I, complimentary and more elucidating experiments are needed to gain a basis for understanding the HWE.

Aim II was focused on obtaining good SANS data of MPs in solution. The objectives for this aim followed very much the learning outcome for Aim I, *i.e.* whether a solution was found to the problem of nanodisc aggregation in D_2O. The first objective for Aim II was partly achieved via three successful proposals to the National Deuteration Facility at ANSTO to secure stealth DDM (sDDM) and stealth POPC lipids (sPOPC), together with a successful proposal to the deuteration lab at ILL in Grenoble for a deuterated version of His-lsMSP1E3D1 that was designed in Paper III. With sPOPC and stealth csMSP1E3D1, a novel stealth nanodisc (sND) was developed and verified to be uniformly contrast-matched to D_2O. The MP CorA was measured both in sDDM and in sNDs, using both standard SANS and SEC-SANS (Paper IV), thereby achieving objective (i) and (iii) of Aim II. The feasibility of SEC-SANS for MP samples was verified in Paper II. The two last objectives remain as ongoing work: It was desired to obtain stealth negatively charged lipids to study their effect of MP structure, and it was desired to model CorA with a combination of SANS and SAXS data. Negatively charged lipids were more difficult to synthesize at the desired deuteration levels and the initial studies presented in Paper IV were thus better carried out with only sPOPC. Now that the sND is established and proven to work, the aim could be set for varying the lipid composition. For CorA, data analysis is still ongoing at the time of writing this thesis.

Aim III was to complement structural data from SAXS and SANS with other biophysical experiments. This was pursued *ad hoc* to answer specific questions throughout the thesis work. The most important results were achieved for CorA, where its structural integrity was verified by negative-stain electron microscopy in D_2O and its functional integrity by activity assays in D_2O, as reported in Paper IV. MPs from the bioSYNergy collaboration mentioned in Aim IV were not measured in the developed stealth carrier systems during this thesis. However, MPs from new and other collaborations were measured in sDDM with great success.

In conclusions, this thesis presents the development and/or use of novel and elegant stealth carrier systems for MP structure determination by SANS. The sDDM combined with SEC-SANS produced excellent results and has a large potential for studying other MPs structurally by SANS. The sND is more difficult because of expensive lipid synthesis combined with a more difficult reconstitution protocol, but provide more well-defined samples that are also suited for static SANS measurements. Paper I provides a comprehensive study of the nanodisc reconstitution process, and the learning outcome of that study will hopefully lead to better success ratios of MP reconstitution endeavors in the future. The problem of obtaining aggregation-free SANS data was approached from different angles, from understanding the HWE (Paper I), to using SEC-SANS (Paper II) and developing more conformationally stable and soluble nanodiscs (Paper III). All of these

approaches were combined into studying the MP CorA in Paper IV. Surprisingly, the obtained SANS data were not immediately in line with current structural models from cryoEM that suggest large conformational transitions in CorA. While data analysis of the SANS data is still ongoing, the initial qualitative observations suggest no such large structural transitions and highlight the relevance of SAS techniques for validating structures obtained by crystallography or cryoEM.

BIBLIOGRAPHY

A. D. McNaught and A. Wilkinson (1997). *IUPAC. Compendium of Chemical Terminology; "The Gold Book"*. Blackwell Scientific Publications, Oxford. arXiv: ISBN:0-9678550-9-8..

Alberts, Bruce et al. (2002). *Molecular Biology of The Cell, 4th edition*. New York: Garland Science.

Als-Nielsen, Jens and Des McMorrow (Mar. 2011). *Elements of Modern X-ray Physics*. 2nd editio. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Anatrace (2018). *https://www.anatrace.com/Products/Lipids/FOS-CHOLINE/F308S*.

Andersen, Heidi D. et al. (Feb. 2011). "Reconciliation of opposing views on membrane–sugar interactions". In: *Proc. Natl. Acad. Sci.* 108.5, pp. 1874–1878.

Antos, John M et al. (June 2009). "A Straight Path to Circular Proteins". In: *J. Biol. Chem.* 284.23, pp. 16028–16036.

Appolaire, Alexandre et al. (Nov. 2014). "Small-angle neutron scattering reveals the assembly mode and oligomeric architecture of TET, a large, dodecameric aminopeptidase". In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70.11, pp. 2983–2993.

Baker, Monya (June 2010). "Making membrane proteins for structures: a trillion tiny tweaks". In: *Nat. Methods* 7.6, pp. 429–434.

Ballauff, M. (May 2001). "SAXS and SANS studies of polymer colloids". In: *Curr. Opin. Colloid Interface Sci.* 6.2, pp. 132–139.

Bayburt, Timothy H., Yelena V. Grinkova, and Stephen G. Sligar (Aug. 2002). "Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins". In: *Nano Lett.* 2.8, pp. 853–856.

Bayburt, Timothy H. and Stephen G. Sligar (May 2010). "Membrane protein assembly into Nanodiscs". In: *FEBS Lett.* 584.9, pp. 1721–1727. arXiv: 15334406.

Bazzacco, Paola et al. (Feb. 2012). "Nonionic Homopolymeric Amphipols: Application to Membrane Protein Folding, Cell-Free Synthesis, and Solution Nuclear Magnetic Resonance". In: *Biochemistry* 51.7, pp. 1416–1430.

Beales, Paul A. et al. (Feb. 2017). "Durable vesicles for reconstitution of membrane proteins in biotechnology". In: *Biochem. Soc. Trans.* 45.1, pp. 15–26.

Beaugrand, Maïwenn et al. (June 2014). "Lipid Concentration and Molar Ratio Boundaries for the Use of Isotropic Bicelles". In: *Langmuir* 30.21, pp. 6162–6170.

Berman, H M et al. (Jan. 2000). “The Protein Data Bank.” In: *Nucleic Acids Res.* 28.1, pp. 235–42.

Berthaud, Alice et al. (June 2012). “Modeling Detergent Organization around Aquaporin-0 Using Small-Angle X-ray Scattering”. In: *J. Am. Chem. Soc.* 134.24, pp. 10080–10088.

Bibow, Stefan et al. (Feb. 2017). “Solution structure of discoidal high-density lipoprotein particles with a shortened apolipoprotein A-I”. In: *Nat. Struct. Mol. Biol.* 24.2, pp. 187–193.

Bjørkskov, Frederik Bühring et al. (Dec. 2017). “Purification and functional comparison of nine human Aquaporins produced in Saccharomyces cerevisiae for the purpose of biophysical characterization”. In: *Sci. Rep.* 7.16899, pp. 1–21.

Blanchet, Clement E. et al. (Apr. 2015). “Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY)”. In: *J. Appl. Crystallogr.* 48.2, pp. 431–443.

Boussambe, Gildas Nyame Mendendy et al. (Sept. 2018). “Fluorinated diglucose detergents for membrane-protein extraction”. In: *Methods* 147.January, pp. 84–94.

Brennich, Martha E., Adam R. Round, and Stephanie Hutin (Jan. 2017). “Online Size-exclusion and Ion-exchange Chromatography on a SAXS Beamline”. In: *J. Vis. Exp.* 119, pp. 1–9.

Breyton, Cécile et al. (July 2013). “Small angle neutron scattering for the study of solubilised membrane proteins”. In: *Eur. Phys. J. E* 36.7, p. 71.

Brocchieri, Luciano and Samuel Karlin (June 2005). “Protein length in eukaryotic and prokaryotic proteomes.” In: *Nucleic Acids Res.* 33.10, pp. 3390–400.

Broecker, Jana, Bryan T. Eger, and Oliver P. Ernst (Feb. 2017). “Crystallogenesis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs”. In: *Structure* 25.2, pp. 384–392. arXiv: 0411586 [cond-mat].

Bruhn, Heike (July 2005). “A short guided tour through functional and structural features of saposin-like proteins”. In: *Biochem. J.* 389.2, pp. 249–257.

Cantor, Robert S. (Mar. 1997). “Lateral Pressures in Cell Membranes: A Mechanism for Modulation of Protein Function”. In: *J. Phys. Chem. B* 101.10, pp. 1723–1725.

Carlson, Michael Luke et al. (Aug. 2018). “The Peptidisc, a simple method for stabilizing membrane proteins in detergent-free solution”. In: *Elife* 7, pp. 1–23.

Chae, Pil Seok et al. (Dec. 2010). “Maltose–neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins”. In: *Nat. Methods* 7.12, pp. 1003–1008.

Champeil, Philippe et al. (Oct. 2016). “A robust method to screen detergents for membrane protein stabilization, revisited”. In: *Anal. Biochem.* 511, pp. 31–35.

Champeil, Philippe et al. (June 2000). “Interaction of Amphipols with Sarcoplasmic Reticulum Ca 2+ -ATPase”. In: *J. Biol. Chem.* 275.25, pp. 18623–18637.

Chaptal, Vincent et al. (Dec. 2017). “Quantification of Detergents Complexed with Membrane Proteins”. In: *Sci. Rep.* 7.1, p. 41751.

Chattopadhyay, Amitabha and K.G. Harikumar (Aug. 1996). “Dependence of critical micelle concentration of a zwitterionic detergent on ionic strength: implications in receptor solubilization”. In: *FEBS Lett.* 391.1-2, pp. 199–202.

Chien, Chih-Ta Henry et al. (Oct. 2017). “An Adaptable Phospholipid Membrane Mimetic System for Solution NMR Studies of Membrane Proteins”. In: *J. Am. Chem. Soc.* 139.42, pp. 14829–14832.

Chipot, Christophe et al. (Apr. 2018). “Perturbations of Native Membrane Protein Structure in Alkyl Phosphocholine Detergents: A Critical Assessment of NMR and Biophysical Studies”. In: *Chem. Rev.* 118.7, pp. 3559–3607.

Cifuentes, Alejandro, Jose L Bernal, and Jose C. Diez-Masa (Oct. 1997). “Determination of Critical Micelle Concentration Values Using Capillary Electrophoresis Instrumentation”. In: *Anal. Chem.* 69.20, pp. 4271–4274. arXiv: S0003-2700(97)00696-3.

Cioni, Patrizia and Giovanni B. Strambini (June 2002). “Effect of Heavy Water on Protein Flexibility”. In: *Biophys. J.* 82.6, pp. 3246–3253.

Contreras, F.-X. et al. (June 2011). “Specificity of Intramembrane Protein-Lipid Interactions”. In: *Cold Spring Harb. Perspect. Biol.* 3.6, a004705–a004705.

Cournia, Zoe et al. (Aug. 2015). “Membrane Protein Structure, Function, and Dynamics: a Perspective from Experiments and Theory”. In: *J. Membr. Biol.* 248.4, pp. 611–640.

Crosas, Eva et al. (Jan. 2017). “Uridine as a new scavenger for synchrotron-based structural biology techniques”. In: *J. Synchrotron Radiat.* 24.1, pp. 53–62.

Cuevas Arenas, Rodrigo et al. (Dec. 2017). “Fast Collisional Lipid Transfer Among Polymer-Bounded Nanodiscs”. In: *Sci. Rep.* 7.1, p. 45875.

Cuevas Arenas, Rodrigo et al. (2016). “Influence of lipid bilayer properties on nanodisc formation mediated by styrene/maleic acid copolymers”. In: *Nanoscale* 8.32, pp. 15016–15026.

David, G. and J. Pérez (Oct. 2009). “Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline”. In: *J. Appl. Crystallogr.* 42.5, pp. 892–900.

Davidson, W. Sean and Thomas B. Thompson (Aug. 2007). “The Structure of Apolipoprotein A-I in High Density Lipoproteins”. In: *J. Biol. Chem.* 282.31, pp. 22249–22253.

Deisenhofer, J. et al. (1985). “Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3Å resolution.” In: *Nature* 318.6047, pp. 618–24.

Denisov, I. G. et al. (Mar. 2004). “Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size”. In: *J. Am. Chem. Soc.* 126.11, pp. 3477–3487. arXiv: 1502.06479.

Denisov, Ilia G. and Stephen G. Sligar (June 2016). “Nanodiscs for structural and functional studies of membrane proteins”. In: *Nat. Struct. Mol. Biol.* 23.6, pp. 481–486.

— (Mar. 2017). “Nanodiscs in Membrane Biochemistry and Biophysics”. In: *Chem. Rev.* 117.6, pp. 4669–4713. arXiv: 15334406.

Denisov, Ilia G. et al. (Aug. 2005). “Thermotropic Phase Transition in Soluble Nanoscale Lipid Bilayers”. In: *J. Phys. Chem. B* 109.32, pp. 15580–15588. arXiv: NIHMS150003.

Domínguez Pardo, Juan J. et al. (July 2018). “Membrane Solubilization by Styrene-Maleic Acid Copolymers: Delineating the Role of Polymer Length”. In: *Biophys. J.* 115.1, pp. 129–138.

Dörr, Jonas M. et al. (Dec. 2014). “Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K + channel: The power of native nanodiscs”. In: *Proc. Natl. Acad. Sci.* 111.52, pp. 18607–18612. arXiv: arXiv:1408.1149.

Dörr, Jonas M. et al. (Jan. 2016). “The styrene–maleic acid copolymer: a versatile tool in membrane research”. In: *Eur. Biophys. J.* 45.1, pp. 3–21.

Dunne, O. et al. (July 2017). “Matchout deuterium labelling of proteins for small-angle neutron scattering studies using prokaryotic and eukaryotic expression systems and high cell-density cultures”. In: *Eur. Biophys. J.* 46.5, pp. 425–432.

Dürr, Ulrich H. N., Melissa Gildenberg, and Ayyalusamy Ramamoorthy (Nov. 2012). “The Magic of Bicelles Lights Up Membrane Protein Structure”. In: *Chem. Rev.* 112.11, pp. 6054–6074.

Efimova, Y. M. et al. (Feb. 2007). “Stability of globular proteins in H2O and D2O”. In: *Biopolymers* 85.3, pp. 264–273.

Egelhaaf, S. U. and P. Schurtenberger (Mar. 1999). “Micelle-to-Vesicle Transition: A Time-Resolved Structural Study”. In: *Phys. Rev. Lett.* 82.13, pp. 2804–2807.

Epand, Richard M. et al. (Jan. 2015). “Membrane curvature modulation of protein activity determined by NMR”. In: *Biochim. Biophys. Acta - Biomembr.* 1848.1, pp. 220–228.

Fagerberg, Linn et al. (Feb. 2010). “Prediction of the human membrane proteome”. In: *Proteomics* 10.6, pp. 1141–1149.

Fitter, Jörg, Thomas Gutberlet, and John Katsaras (2006). *Neutron Scattering in Biology*. Ed. by Elias Greenbaum. Biological and Medical Physics, Biomedical Engineering. Berlin/Heidelberg: Springer-Verlag. arXiv: arXiv:1011.1669v3.

Flayhan, Ali et al. (Feb. 2018). “Saposin Lipid Nanoparticles: A Highly Versatile and Modular Tool for Membrane Protein Research”. In: *Structure* 26.2, 345–355.e5.

Franke, Daniel and Dmitri I. Svergun (Apr. 2009). “DAMMIF , a program for rapid ab-initio shape determination in small-angle scattering”. In: *J. Appl. Crystallogr.* 42.2, pp. 342–346.

Frauenfeld, Jens et al. (Apr. 2016). “A saposin-lipoprotein nanoparticle system for membrane proteins”. In: *Nat. Methods* 13.4, pp. 345–351. arXiv: 15334406.

Frey, Lukas et al. (Jan. 2017). “Micelles, Bicelles, and Nanodiscs: Comparing the Impact of Membrane Mimetics on Membrane Protein Backbone Dynamics”. In: *Angew. Chemie Int. Ed.* 56.1, pp. 380–383. arXiv: NIHMS150003.

Gabel, Frank (2015). “Small-Angle Neutron Scattering for Structural Biology of Protein–RNA Complexes”. In: *Methods Enzymol.* Vol. 558. 1. Elsevier Inc., pp. 391–415.

Gabel, Frank et al. (July 2014). “Probing the Conformation of FhaC with Small-Angle Neutron Scattering and Molecular Modeling”. In: *Biophys. J.* 107.1, pp. 185–196.

Garavito, R. Michael and Shelagh Ferguson-Miller (Aug. 2001). “Detergents as Tools in Membrane Biochemistry”. In: *J. Biol. Chem.* 276.35, pp. 32403–32406.

Garidel, P. et al. (June 2000). “Thermodynamic Characterization of Bile Salt Aggregation as a Function of Temperature and Ionic Strength Using Isothermal Titration Calorimetry”. In: *Langmuir* 16.12, pp. 5267–5275.

Garni, Martina et al. (Apr. 2017). “Biopores/membrane proteins in synthetic polymer membranes”. In: *Biochim. Biophys. Acta - Biomembr.* 1859.4, pp. 619–638.

Gasteiger, Elisabeth et al. (2005). “Protein Identification and Analysis Tools on the ExPASy Server”. In: *Proteomics Protoc. Handb.* Vol. 50. 2. Totowa, NJ: Humana Press, pp. 571–607. arXiv: arXiv:1011.1669v3.

Giusti, Fabrice et al. (Oct. 2014). “Synthesis, Characterization and Applications of a Perdeuterated Amphipol”. In: *J. Membr. Biol.* 247.9-10, pp. 909–924.

Glatter, O. (1977). “A new method for the evaluation of small-angle scattering data”. In: *J. Appl. Crystallogr.* 10.5, pp. 415–421. arXiv: NIHMS150003.

Glatter, Otto (2002). “The Inverse Scattering Problem in Small-Angle Scattering”. In: *Neutrons, X-rays Light Scatt. Methods Appl. to Soft Condens. Matter*. Ed. by P. Lindner and Th. Zemb. North-Holland. Chap. Chapter 4, pp. 73–102.

Gohon, Yann et al. (May 2008). “Bacteriorhodopsin/Amphipol Complexes: Structural and Functional Properties”. In: *Biophys. J.* 94.9, pp. 3523–3537.

Gohon, Yann et al. (Jan. 2006). “Well-Defined Nanoparticles Formed by Hydrophobic Assembly of a Short and Polydisperse Random Terpolymer, Amphipol A8-35”. In: *Langmuir* 22.3, pp. 1281–1290. arXiv: arXiv:1408.1149.

González-Pérez, Alfredo, Kenneth M. Persson, and Pablo Taboada (May 2018). “Structural stability of SoPIP2;1 aquaporin under reconstitution in polymersomes”. In: *J. Mol. Liq.* 257, pp. 26–31.

Graewert, Melissa A. et al. (Sept. 2015). “Automated Pipeline for Purification, Biophysical and X-Ray Analysis of Biomacromolecular Solutions”. In: *Sci. Rep.* 5.1, p. 10734.

Grinkova, Y. V., I. G. Denisov, and S. G. Sligar (Nov. 2010). “Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers”. In: *Protein Eng. Des. Sel.* 23.11, pp. 843–848.

Gripon, C. et al. (June 1997). “Lysozyme solubility in H2O and D2O solutions: a simple relationship”. In: *J. Cryst. Growth* 177.3-4, pp. 238–247.

Grouleff, Julie et al. (Sept. 2015). “The influence of cholesterol on membrane protein structure, function, and dynamics studied by molecular dynamics simulations”. In: *Biochim. Biophys. Acta - Biomembr.* 1848.9, pp. 1783–1795.

Guinier, A. and G. Fournet (1955). *Small Angle Scattering of X-Rays*. New York: Wiley.

Hagn, Franz et al. (Feb. 2013). “Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins”. In: *J. Am. Chem. Soc.* 135.5, pp. 1919–1925.

Hall, Stephen C. L. et al. (2018). “An acid-compatible co-polymer for the solubilization of membranes and proteins into lipid bilayer-containing nanoparticles”. In: *Nanoscale* 10.22, pp. 10609–10619.

Hansen, Sara K. et al. (Aug. 2015). “Solid-state NMR methods for oriented membrane proteins”. In: *Prog. Nucl. Magn. Reson. Spectrosc.* 88-89, pp. 48–85.

Harayama, Takeshi and Howard Riezman (Feb. 2018). “Understanding the diversity of membrane lipid composition”. In: *Nat. Rev. Mol. Cell Biol.* 19.5, pp. 281–296.

Heberle, Frederick A. and Georg Pabst (Aug. 2017). “Complex biomembrane mimetics on the sub-nanometer scale”. In: *Biophys. Rev.* 9.4, pp. 353–373.

Heerklotz, Heiko and Joachim Seelig (Sept. 2001). “Detergent-Like Action of the Antibiotic Peptide Surfactin on Lipid Membranes”. In: *Biophys. J.* 81.3, pp. 1547–1554.

Hemberg, O., M. Otendal, and H. M. Hertz (Aug. 2003). “Liquid-metal-jet anode electron-impact x-ray source”. In: *Appl. Phys. Lett.* 83.7, pp. 1483–1485.

Herranz-Trillo, Fátima et al. (Jan. 2017). “Structural Analysis of Multi-component Amyloid Systems by Chemometric SAXS Data Decomposition”. In: *Structure* 25.1, pp. 5–15.

Hjelmeland, Leonard M., Daniel W. Nebert, and James C. Osborne (Apr. 1983). “Sulfobetaine derivatives of bile acids: Nondenaturing surfactants for membrane biochemistry”. In: *Anal. Biochem.* 130.1, pp. 72–82.

Hopkins, Jesse B. and Robert E. Thorne (June 2016). “Quantifying radiation damage in biomolecular small-angle X-ray scattering”. In: *J. Appl. Crystallogr.* 49.3, pp. 880–890.

Huda, Pie (May 2015). “Studying membrane protein structure and function using nanodiscs”. In: *PhD Thesis*.

Huda, Pie et al. (July 2015). “PET/CT Based In Vivo Evaluation of 64Cu Labelled Nanodiscs in Tumor Bearing Mice”. In: *PLoS One* 10.7. Ed. by Christoph E Hagemeyer, e0129310.

Imura, Tomohiro et al. (Apr. 2014). “Surfactant-like Properties of an Amphiphilic α-Helical Peptide Leading to Lipid Nanodisc Formation”. In: *Langmuir* 30.16, pp. 4752–4759.

Islam, Rafique M. et al. (Dec. 2018). “Structural properties of apolipoprotein A-I mimetic peptides that promote ABCA1-dependent cholesterol efflux”. In: *Sci. Rep.* 8.1, p. 2956.

Israelachvili, Jacob (1991). *Intermolecular & Surface Forces*. London: Academic Press, p. 450. arXiv: arXiv:1011.1669v3.

Jackson, Andrew J. and Kalliopi Kanaki (2014). “ESS Construction Proposal LoKI - A broad-band SANS instrument”. In:

Jackson, Scott M. et al. (Apr. 2018). “Structural basis of small-molecule inhibition of human multidrug transporter ABCG2”. In: *Nat. Struct. Mol. Biol.* 25.4, pp. 333–340.

Jacrot, B. (Oct. 1976). “The study of biological structures by neutron scattering from solution”. In: *Reports Prog. Phys.* 39.10, pp. 911–953.

Jahnen-Dechent, Wilhelm and Markus Ketteler (Feb. 2012). “Magnesium basics”. In: *Clin. Kidney J.* 5.Suppl 1, pp. i3–i14.

Jamshad, Mohammed et al. (Mar. 2015). “Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins”. In: *Nano Res.* 8.3, pp. 774–789.

Jaradat, Da’san M. M. (Jan. 2018). “Thirteen decades of peptide synthesis: key developments in solid phase peptide synthesis and amide bond formation utilized in peptide ligation”. In: *Amino Acids* 50.1, pp. 39–68.

Jeffries, Cy M. et al. (Mar. 2015). “Limiting radiation damage for high-brilliance biological solution scattering: practical experience at the EMBL P12 beamline PETRAIII”. In: *J. Synchrotron Radiat.* 22.2, pp. 273–279.

Jeffries, Cy M. et al. (Nov. 2016). “Preparing monodisperse macromolecular samples for successful biological small-angle X-ray and neutron-scattering experiments”. In: *Nat. Protoc.* 11.11, pp. 2122–2153.

Jordan, Ashley et al. (Dec. 2016). “SEC-SANS: size exclusion chromatography combined in situ with small-angle neutron scattering”. In: *J. Appl. Crystallogr.* 49.6, pp. 2015–2020.

Jørgensen, Ida Louise, Gerdi Christine Kemmer, and Thomas Günther Pomorski (Mar. 2017). “Membrane protein reconstitution into giant unilamellar vesicles: a review on current techniques”. In: *Eur. Biophys. J.* 46.2, pp. 103–119.

Josts, Inokentijs et al. (Aug. 2018). “Conformational States of ABC Transporter MsbA in a Lipid Environment Investigated by Small-Angle Scattering Using Stealth Carrier Nanodiscs”. In: *Structure* 26.8, 1072–1079.e4.

Kariyazono, Hirokazu et al. (Feb. 2016). “Formation of stable nanodiscs by bihelical apolipoprotein A-I mimetic peptide”. In: *J. Pept. Sci.* 22.2, pp. 116–122.

Kiessling, Volker, Chen Wan, and Lukas K. Tamm (Jan. 2009). “Domain coupling in asymmetric lipid bilayers”. In: *Biochim. Biophys. Acta - Biomembr.* 1788.1, pp. 64–71.

Knowles, Timothy J. et al. (June 2009). “Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer”. In: *J. Am. Chem. Soc.* 131.22, pp. 7484–7485.

Kobilka, Brian (June 2013). “The Structural Basis of G-Protein-Coupled Receptor Signaling (Nobel Lecture)”. In: *Angew. Chemie Int. Ed.* 52.25, pp. 6380–6388. arXiv: NIHMS150003.

Konarev, Petr V. and Dmitri I. Svergun (July 2018). “Direct shape determination of intermediates in evolving macromolecular solutions from small-angle scattering data”. In: *IUCrJ* 5.4, pp. 402–409.

Kondo, Hiroaki, Keisuke Ikeda, and Minoru Nakano (Oct. 2016). “Formation of size-controlled, denaturation-resistant lipid nanodiscs by an amphiphilic self-polymerizing peptide”. In: *Colloids Surfaces B Biointerfaces* 146, pp. 423–430.

Koutsioubas, Alexandros (Dec. 2017). “Low-Resolution Structure of Detergent-Solubilized Membrane Proteins from Small-Angle Scattering Data”. In: *Biophys. J.* 113.11, pp. 2373–2382.

Kramer, Ryan M. et al. (Apr. 2012). “Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility”. In: *Biophys. J.* 102.8, pp. 1907–1915.

Krause, Lennard et al. (Feb. 2015). “Comparison of silver and molybdenum microfocus X-ray sources for single-crystal structure determination”. In: *J. Appl. Crystallogr.* 48.1, pp. 3–10. arXiv: /repositorio. ug.edu.ec/handle/redug/10118 [http:].

Kučerka, Norbert, Mu-Ping Nieh, and John Katsaras (Nov. 2011). “Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature”. In: *Biochim. Biophys. Acta - Biomembr.* 1808.11, pp. 2761–2771.

Kuwamoto, Shigeo, Shuji Akiyama, and Tetsuro Fujisawa (Nov. 2004). “Radiation damage to a protein solution, detected by synchrotron X-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants”. In: *J. Synchrotron Radiat.* 11.6, pp. 462–468.

Kwan, Ann H. et al. (2011). “Macromolecular NMR spectroscopy for the non-spectroscopist”. In: *The FEBS Journal* 278.5, pp. 687–703.

Kynde, Søren A. R. et al. (Feb. 2014). “Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment”. In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70.2, pp. 371–383.

Larsen, Andreas Haahr et al. (Nov. 2018). “Small-angle neutron scattering studies on the AMPA receptor GluA2 in the resting, AMPA-bound and GYKI-53655-bound states”. In: *IUCrJ* 5.6, pp. 780–793.

Larsen, Andreas N. et al. (2016). “Dimeric peptides with three different linkers self-assemble with phospholipids to form peptide nanodiscs that stabilize membrane proteins”. In: *Soft Matter* 12.27, pp. 5937–5949.

Lasic, Danilo D. (Nov. 1988). “The mechanism of vesicle formation”. In: *Biochem. J.* 256.1, pp. 1–11.

Le Bon, Christel et al. (Sept. 2018). “Folding and stabilizing membrane proteins in amphipol A8-35”. In: *Methods* 147.April, pp. 95–105.

Lee, Robin van der et al. (July 2014). “Classification of Intrinsically Disordered Regions and Proteins”. In: *Chem. Rev.* 114.13, pp. 6589–6631.

Lemmon, Mark A. (Feb. 2008). “Membrane recognition by phospholipid-binding domains”. In: *Nat. Rev. Mol. Cell Biol.* 9.2, pp. 99–111.

Lévy, Daniel et al. (June 1990). “A systematic study of liposome and proteoliposome reconstitution involving Bio-Bead-mediated Triton X-100 removal”. In: *Biochim. Biophys. Acta - Biomembr.* 1025.2, pp. 179–190.

Li, Jun et al. (Oct. 2016). “Characterizing the Size and Composition of Saposin A Lipoprotein Picodiscs”. In: *Anal. Chem.* 88.19, pp. 9524–9531.

Lichtenberg, Dov, Hasna Ahyayauch, and Félix M. Goñi (July 2013). “The Mechanism of Detergent Solubilization of Lipid Bilayers”. In: *Biophys. J.* 105.2, pp. 289–299. arXiv: arXiv:1011.1669v3.

Lorber, Bernard, John B. Bishop, and Lawrence J. DeLucas (Apr. 1990). “Purification of octyl β-d-glucopyranoside and re-estimation of its micellar size”. In: *Biochim. Biophys. Acta - Biomembr.* 1023.2, pp. 254–265.

Luchette, Paul A. et al. (Aug. 2001). “Morphology of fast-tumbling bicelles: a small angle neutron scattering and NMR study”. In: *Biochim. Biophys. Acta - Biomembr.* 1513.2, pp. 83–94.

Lydic, Todd A. and Young-Hwa Goo (Dec. 2018). “Lipidomics unveils the complexity of the lipidome in metabolic diseases”. In: *Clin. Transl. Med.* 7.1, p. 4.

Lyukmanova, E.N. et al. (Mar. 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”. In: *Biochim. Biophys. Acta - Biomembr.* 1818.3, pp. 349–358.

Maguire, Michael E. (Sept. 2006). “Magnesium transporters: properties, regulation and structure.” In: *Front. Biosci.* 11, pp. 3149–63.

Maric, Selma et al. (Jan. 2015). “Biosynthetic preparation of selectively deuterated phosphatidylcholine in genetically modified Escherichia coli”. In: *Appl. Microbiol. Biotechnol.* 99.1, pp. 241–254. arXiv: NIHMS150003.

Maric, Selma et al. (Feb. 2014). “Stealth carriers for low-resolution structure determination of membrane proteins in solution”. In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70.2, pp. 317–328.

Marty, Michael T. et al. (May 2013). “Nanodisc-solubilized membrane protein library reflects the membrane proteome.” In: *Anal. Bioanal. Chem.* 405.12, pp. 4009–16. arXiv: NIHMS150003.

Maslova, V. A. and M. A. Kiselev (May 2018). “Structure of Sodium Cholate Micelles”. In: *Crystallogr. Reports* 63.3, pp. 472–475.

Mathew, Elizabeth, Ahmed Mirza, and Nick Menhart (July 2004). “Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins”. In: *J. Synchrotron Radiat.* 11.4, pp. 314–318.

Matsui, Ryoichi et al. (Dec. 2016). “Magnetically Alignable Bicelles with Unprecedented Stability Using Tunable Surfactants Derived from Cholic Acid”. In: *ChemPhysChem* 17.23, pp. 3916–3922.

Matthies, Doreen et al. (Feb. 2016). “Cryo-EM Structures of the Magnesium Channel CorA Reveal Symmetry Break upon Gating”. In: *Cell* 164.4, pp. 747–756.

Matthies, Doreen et al. (Aug. 2018). “Single-particle cryo-EM structure of a voltage-activated potassium channel in lipid nanodiscs”. In: *Elife* 7, pp. 1–18.

MAXIV (2018). *https://www.maxiv.lu.se/accelerators-beamlines/beamlines/cosaxs/techniques/*.

McGregor, Clare-Louise et al. (Feb. 2003). “Lipopeptide detergents designed for the structural study of membrane proteins”. In: *Nat. Biotechnol.* 21.2, pp. 171–176.

Meer, Gerrit van, Dennis R. Voelker, and Gerald W. Feigenson (Feb. 2008). “Membrane lipids: where they are and how they behave”. In: *Nat. Rev. Mol. Cell Biol.* 9.2, pp. 112–124. arXiv: NIHMS150003.

Mei, Xiaohu and David Atkinson (Nov. 2011). “Crystal Structure of C-terminal Truncated Apolipoprotein A-I Reveals the Assembly of High Density Lipoprotein (HDL) by Dimerization”. In: *J. Biol. Chem.* 286.44, pp. 38570–38582.

Midtgaard, Søren Roi (2013). “Self-assembling peptide and protein nanodiscs for studies of membrane proteins”. In: *PhD Thesis*.

Midtgaard, Søren Roi, Martin Cramer Pedersen, and Lise Arleth (July 2015). “Small-Angle X-Ray Scattering of the Cholesterol Incorporation into Human ApoA1-POPC Discoidal Particles”. In: *Biophys. J.* 109.2, pp. 308–318.

Midtgaard, Søren Roi et al. (Jan. 2018). “Invisible detergents for structure determination of membrane proteins by small-angle neutron scattering”. In: *FEBS J.* 285.2, pp. 357–371. arXiv: 0608246v3 [arXiv:physics].

Midtgaard, Søren Roi et al. (2014). “Self-assembling peptides form nanodiscs that stabilize membrane proteins”. In: *Soft Matter* 10.5, pp. 738–752.

Miehling, Jonas, David Goricanec, and Franz Hagn (Sept. 2018). “A Split-Intein-Based Method for the Efficient Production of Circularized Nanodiscs for Structural Studies of Membrane Proteins”. In: *ChemBioChem* 19.18, pp. 1927–1933.

Mio, Kazuhiro and Chikara Sato (Apr. 2018). “Lipid environment of membrane proteins in cryo-EM based structural analysis”. In: *Biophys. Rev.* 10.2, pp. 307–316.

Mishra, Vinod K. et al. (Mar. 2006). “Association of a Model Class A (Apolipoprotein) Amphipathic α Helical Peptide with Lipid”. In: *J. Biol. Chem.* 281.10, pp. 6511–6519.

Miyazaki, Masakazu et al. (Sept. 2010). “Static and Dynamic Characterization of Nanodiscs with Apolipoprotein A-I and Its Model Peptide”. In: *J. Phys. Chem. B* 114.38, pp. 12376–12382.

Mörs, Karsten et al. (Apr. 2013). “Modified lipid and protein dynamics in nanodiscs”. In: *Biochim. Biophys. Acta - Biomembr.* 1828.4, pp. 1222–1229.

Mylonas, Efstratios and Dmitri I. Svergun (Feb. 2007). “Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering”. In: *J. Appl. Crystallogr.* 40.s1, s245–s249.

Naing, Swe-Htet et al. (Feb. 2018). “Solution Structure of an Intramembrane Aspartyl Protease via Small Angle Neutron Scattering”. In: *Biophys. J.* 114.3, pp. 602–608.

Nakano, Minoru et al. (June 2009). “Static and Dynamic Properties of Phospholipid Bilayer Nanodiscs”. In: *J. Am. Chem. Soc.* 131.23, pp. 8308–8312.

Nasr, Mahmoud L. et al. (Jan. 2017). “Covalently circularized nanodiscs for studying membrane proteins and viral entry”. In: *Nat. Methods* 14.1, pp. 49–52.

Nicolson, Garth L. (June 2014). “The Fluid—Mosaic Model of Membrane Structure: Still relevant to understanding the structure, function and dynamics of biological membranes after more than 40years”. In: *Biochim. Biophys. Acta - Biomembr.* 1838.6, pp. 1451–1466.

Nieh, Mu-Ping et al. (May 2002). “SANS Study on the Effect of Lanthanide Ions and Charged Lipids on the Morphology of Phospholipid Mixtures”. In: *Biophys. J.* 82.5, pp. 2487–2498.

Oliver, Ryan C., Sai Venkatesh Pingali, and Volker S. Urban (Oct. 2017). “Designing Mixed Detergent Micelles for Uniform Neutron Contrast”. In: *J. Phys. Chem. Lett.* 8.20, pp. 5041–5046.

Oliver, Ryan C. et al. (May 2013). “Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group”. In: *PLoS One* 8.5. Ed. by Laurent Kreplak, e62488.

Oluwole, Abraham Olusegun et al. (Feb. 2017). “Solubilization of Membrane Proteins into Functional Lipid-Bilayer Nanodiscs Using a Diisobutylene/Maleic Acid Copolymer”. In: *Angew. Chemie Int. Ed.* 56.7, pp. 1919–1924.

Orthaber, Doris, Alexander Bergmann, and Otto Glatter (Apr. 2000). “SAXS experiments on absolute scale with Kratky systems using water as a secondary standard”. In: *J. Appl. Crystallogr.* 33.2, pp. 218–225.

Ostanevich, Yurii M. (Mar. 1988). “Time-of-flight small-angle scattering spectrometers on pulsed neutron sources”. In: *Makromol. Chemie. Macromol. Symp.* 15.1, pp. 91–103.

Overington, John P., Bissan Al-Lazikani, and Andrew L. Hopkins (Dec. 2006). “How many drug targets are there?” In: *Nat. Rev. Drug Discov.* 5.12, pp. 993–996. arXiv: NIHMS150003.

Palmer, Liam C. et al. (June 2007). “Supramolecular self-assembly codes for functional structures”. In: *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 365.1855, pp. 1417–1433.

Panjkovich, Alejandro and Dmitri I. Svergun (June 2018). “CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data”. In: *Bioinformatics* 34.11. Ed. by Alfonso Valencia, pp. 1944–1946.

Park, Sang Ho et al. (Oct. 2011). “Nanodiscs versus Macrodiscs for NMR of Membrane Proteins”. In: *Biochemistry* 50.42, pp. 8983–8985. arXiv: NIHMS150003.

Park, Sang Ho et al. (Nov. 2012). “Structure of the chemokine receptor CXCR1 in phospholipid bilayers”. In: *Nature* 491.7426, pp. 779–783. arXiv: NIHMS150003.

Parmar, Mayuriben et al. (Feb. 2018). “Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure”. In: *Biochim. Biophys. Acta - Biomembr.* 1860.2, pp. 378–383.

Pedersen, Jan Skov (2002). “Modelling of Small-Angle Scattering Data”. In: *Neutrons, X-rays Light Scatt. Methods Appl. to Soft Condens. Matter*. Ed. by P. Lindner and Th. Zemb. North-Holland. Chap. Chapter 16, pp. 391–420.

Pedersen, Martin Cramer, Lise Arleth, and Kell Mortensen (Dec. 2013). “WillItFit : a framework for fitting of constrained models to small-angle scattering data”. In: *J. Appl. Crystallogr.* 46.6, pp. 1894–1898.

Pedersen, Martin Cramer et al. (Dec. 2014). “Quantification of the information in small-angle scattering data”. In: *J. Appl. Crystallogr.* 47.6, pp. 2000–2010.

Perlmutter, Jason D., Jean-Luc Popot, and Jonathan N. Sachs (Oct. 2014). “Molecular Dynamics Simulations of a Membrane Protein/Amphipol Complex”. In: *J. Membr. Biol.* 247.9-10, pp. 883–895.

Pernot, Petra et al. (July 2013). “Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution”. In: *J. Synchrotron Radiat.* 20.4, pp. 660–664.

Petoukhov, Maxim V. and Dmitri I. Svergun (Aug. 2005). “Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data”. In: *Biophys. J.* 89.2, pp. 1237–1250.

Planchard, Noelya et al. (Oct. 2014). “The Use of Amphipols for Solution NMR Studies of Membrane Proteins: Advantages and Constraints as Compared to Other Solubilizing Media”. In: *J. Membr. Biol.* 247.9-10, pp. 827–842.

Polenova, Tatyana, Rupal Gupta, and Amir Goldbourt (June 2015). “Magic Angle Spinning NMR Spectroscopy: A Versatile Technique for Structural and Dynamic Analysis of Solid-Phase Systems”. In: *Anal. Chem.* 87.11, pp. 5458–5469.

Popot, J.-L. et al. (June 2011). “Amphipols From A to Z*”. In: *Annu. Rev. Biophys.* 40.1, pp. 379–408.

Popot, Jean-Luc (June 2010). “Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions.” In: *Annu. Rev. Biochem.* 79.1, pp. 737–75.

Popot, Jean-Luc and Donald M. Engelman (Jan. 2016). “Membranes Do Not Tell Proteins How To Fold”. In: *Biochemistry* 55.1, pp. 5–18.

Popovic, K. et al. (Feb. 2012). “Structure of saposin A lipoprotein discs”. In: *Proc. Natl. Acad. Sci.* 109.8, pp. 2908–2912.

Postis, Vincent et al. (Feb. 2015). “The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy”. In: *Biochim. Biophys. Acta - Biomembr.* 1848.2, pp. 496–501.

Poveda, J.A. et al. (July 2008). “Protein-promoted membrane domains”. In: *Biochim. Biophys. Acta - Biomembr.* 1778.7-8, pp. 1583–1590.

Privé, Gilbert G. (Aug. 2009). “Lipopeptide detergents for membrane protein studies”. In: *Curr. Opin. Struct. Biol.* 19.4, pp. 379–385.

Puthenveetil, Robbins, Khiem Nguyen, and Olga Vinogradova (Jan. 2017). “Nanodiscs and solution NMR: preparation, application and challenges”. In: *Nanotechnol. Rev.* 6.1, pp. 111–126.

Putnam, Christopher D. et al. (Aug. 2007). “X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution”. In: *Q. Rev. Biophys.* 40.03, pp. 191–285.

Radoicic, Jasmina, Sang Ho Park, and Stanley J. Opella (July 2018). “Macrodiscs Comprising SMALPs for Oriented Sample Solid-State NMR Spectroscopy of Membrane Proteins”. In: *Biophys. J.* 115.1, pp. 22–25.

Rasmussen, Søren G. F. et al. (Nov. 2007). “Crystal structure of the human β2 adrenergic G-protein-coupled receptor”. In: *Nature* 450.7168, pp. 383–387.

Rigaud, J.-L. (July 2002). “Membrane proteins: functional and structural studies using reconstituted proteoliposomes and 2-D crystals”. In: *Brazilian J. Med. Biol. Res.* 35.7, pp. 753–766.

Rigaud, J.-L. et al. (June 1998). “Detergent removal by non-polar polystyrene beads”. In: *Eur. Biophys. J.* 27.4, pp. 305–319.

Ritchie, T.K. et al. (Jan. 2009). “Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs”. In: *Methods Enzymol.* Ed. by Nejat Düzgünes. Vol. 464. C. Academmic Press. Chap. Methods in, pp. 211–231. arXiv: NIHMS150003.

Robertson, Janice L (Nov. 2018). “The lipid bilayer membrane and its protein constituents”. In: *J. Gen. Physiol.* 150.11, pp. 1472–1483.

Roh, Soung-Hun et al. (Mar. 2018). “The 3.5-Å CryoEM Structure of Nanodisc-Reconstituted Yeast Vacuolar ATPase V o Proton Channel”. In: *Mol. Cell* 69.6, 993–1004.e3.

Rothblat, George H. and Michael C. Phillips (June 2010). “High-density lipoprotein heterogeneity and function in reverse cholesterol transport”. In: *Curr. Opin. Lipidol.* 21.3, pp. 229–238.

Rouck, John E. et al. (July 2017). “Recent advances in nanodisc technology for membrane protein studies (2012-2017)”. In: *FEBS Lett.* 591.14, pp. 2057–2088. arXiv: 15334406.

Ryan, Timothy M. et al. (Feb. 2018). “An optimized SEC-SAXS system enabling high X-ray dose for rapid SAXS assessment with correlated UV measurements for biomolecular structure analysis”. In: *J. Appl. Crystallogr.* 51.1, pp. 97–111.

Saito, Hiroyuki et al. (May 2004). “α-Helix Formation Is Required for High Affinity Binding of Human Apolipoprotein A-I to Lipids”. In: *J. Biol. Chem.* 279.20, pp. 20974–20981.

Sali, Andrej et al. (July 2015). “Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop”. In: *Structure* 23.7, pp. 1156–1167.

Sanders, Charles R. and R. Scott Prosser (Oct. 1998). “Bicelles: a model membrane system for all seasons?” In: *Structure* 6.10, pp. 1227–1234.

Saotome, Kei et al. (Dec. 2017). “Structure of the mechanically activated ion channel Piezo1”. In: *Nature* 554.7693, pp. 481–486. arXiv: 15334406.

Sauer, Ines et al. (Apr. 2006). “Dipalmitoylation of a cellular uptake-mediating apolipoprotein E-derived peptide as a promising modification for stable anchorage in liposomal drug carriers”. In: *Biochim. Biophys. Acta - Biomembr.* 1758.4, pp. 552–561.

Scheidelaar, Stefan et al. (Jan. 2015). “Molecular Model for the Solubilization of Membranes into Nanodisks by Styrene Maleic Acid Copolymers”. In: *Biophys. J.* 108.2, pp. 279–290. arXiv: arXiv:1408.1149.

Schlame, Michael (Aug. 2008). “Thematic Review Series: Glycerolipids. Cardiolipin synthesis for the assembly of bacterial and mitochondrial membranes”. In: *J. Lipid Res.* 49.8, pp. 1607–1620.

Schulz, Sarah et al. (Apr. 2017). “Molecular architecture of the N-type ATPase rotor ring from Burkholderia pseudomallei”. In: *EMBO Rep.* 18.4, pp. 526–535.

Sears, Varley F. (1992). “Neutron scattering lengths and cross sections”. In: *Neutron News* 3.3, pp. 26–37.

Seddon, Annela M., Paul Curnow, and Paula J. Booth (Nov. 2004). “Membrane proteins, lipids and detergents: not just a soap opera”. In: *Biochim. Biophys. Acta - Biomembr.* 1666.1-2, pp. 105–117.

Sezgin, Erdinc et al. (Mar. 2017). “The mystery of membrane organization: composition, regulation and roles of lipid rafts”. In: *Nat. Rev. Mol. Cell Biol.* 18.6, pp. 361–374. arXiv: NIHMS150003.

Shaw, Andrew W., Mark A. McLean, and Stephen G. Sligar (Jan. 2004). “Phospholipid phase transitions in homogeneous nanometer scale bilayer discs”. In: *FEBS Lett.* 556.1-3, pp. 260–264.

Shimizu, Kentaro et al. (May 2018). “Comparative analysis of membrane protein structure databases”. In: *Biochim. Biophys. Acta - Biomembr.* 1860.5, pp. 1077–1091.

Simon, Kailene S., Naomi L. Pollock, and Sarah C. Lee (Nov. 2018). “Membrane protein nanoparticles: the shape of things to come”. In: *Biochem. Soc. Trans.* 0.November, BST20180139.

Skar-Gislinge, Nicholas and Lise Arleth (2011). "Small-angle scattering from phospholipid nanodiscs: derivation and refinement of a molecular constrained analytical model form factor". In: *Phys. Chem. Chem. Phys.* 13.8, pp. 3161–3170.

Skar-Gislinge, Nicholas et al. (Oct. 2010). "Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein". In: *J. Am. Chem. Soc.* 132.39, pp. 13713–13722. arXiv: NIHMS150003.

Skou, Soren, Richard E. Gillilan, and Nozomi Ando (July 2014). "Synchrotron-based small-angle X-ray scattering of proteins in solution". In: *Nat. Protoc.* 9.7, pp. 1727–1739. arXiv: 15334406.

Skrypnychuk, Vasyl et al. (2016). "Small-angle scattering : a view on the properties , structures and structural changes of biological macromolecules in solution". In: *J. Mater. Chem. C* 4.19, pp. 4143–4149. arXiv: NIHMS150003.

Smith, Ross and Charles Tanford (June 1972). "The critical micelle concentration of l-α-dipalmitoylphosphatidylcholine in water and water/methanol solutions". In: *J. Mol. Biol.* 67.1, pp. 75–83.

Sohlenkamp, Christian and Otto Geiger (2016). "Bacterial membrane lipids: Diversity in structures and pathways". In: *FEMS Microbiol. Rev.* 40.1, pp. 133–159.

Son, Woo Sung et al. (Jan. 2012). "'q-Titration' of long-chain and short-chain lipids differentiates between structured and mobile residues of membrane proteins studied in bicelles by solution NMR spectroscopy". In: *J. Magn. Reson.* 214, pp. 111–118.

Stetsenko, Artem and Albert Guskov (July 2017). "An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization". In: *Crystals* 7.7, p. 197.

Stuart, Marc C.A. and Egbert J. Boekema (Nov. 2007). "Two distinct mechanisms of vesicle-to-micelle and micelle-to-vesicle transition are mediated by the packing parameter of phospholipid–detergent systems". In: *Biochim. Biophys. Acta - Biomembr.* 1768.11, pp. 2681–2689.

Sud, Manish et al. (Jan. 2007). "LMSD: LIPID MAPS structure database". In: *Nucleic Acids Res.* 35.Database, pp. D527–D532.

Svergun, D.I. (June 1999). "Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing". In: *Biophys. J.* 76.6, pp. 2879–2886.

Svergun, Dmitri I. and Michel H J Koch (Oct. 2003). "Small-angle scattering studies of biological macromolecules in solution". In: *Reports Prog. Phys.* 66.10, pp. 1735–1782.

Tanford, Charles (1980). "The hydrophobic effect: Formation of micelles and biological membranes, 2nd edition". In: pp. 1–232.

Tao, Houchao et al. (Aug. 2013). "Engineered nanostructured β-sheet peptides protect membrane proteins". In: *Nat. Methods* 10.8, pp. 759–761. arXiv: NIHMS150003.

Thonghin, Nopnithi et al. (Sept. 2018). “Cryo-electron microscopy of membrane proteins”. In: *Methods* 147, pp. 176–186. arXiv: arXiv:1011.1669v3.

Tonge, S.R and B.J Tighe (Dec. 2001). “Responsive hydrophobically associating polymers: a review of structure and properties”. In: *Adv. Drug Deliv. Rev.* 53.1, pp. 109–122.

Trevino, Saul R., J. Martin Scholtz, and C. Nick Pace (Feb. 2007). “Amino Acid Contribution to Protein Solubility: Asp, Glu, and Ser Contribute more Favorably than the other Hydrophilic Amino Acids in RNase Sa”. In: *J. Mol. Biol.* 366.2, pp. 449–460. arXiv: NIHMS150003.

Tribet, C., R. Audebert, and J.-L. Popot (Dec. 1996). “Amphipols: Polymers that keep membrane proteins soluble in aqueous solutions”. In: *Proc. Natl. Acad. Sci.* 93.26, pp. 15047–15050.

Tunuguntla, Ramya et al. (Sept. 2013). “Lipid Bilayer Composition Can Influence the Orientation of Proteorhodopsin in Artificial Membranes”. In: *Biophys. J.* 105.6, pp. 1388–1396.

VanAken, T. et al. (1986). “Alkyl glycoside detergents: synthesis and applications to the study of membrane proteins.” In: *Methods Enzymol.* 125, pp. 27–35.

Vargas, Carolyn et al. (2015). “Nanoparticle self-assembly in mixtures of phospholipids with styrene/maleic acid copolymers or fluorinated surfactants”. In: *Nanoscale* 7.48, pp. 20685–20696.

Wagner, Samuel et al. (Sept. 2007). “Consequences of Membrane Protein Overexpression in Escherichia coli”. In: *Mol. Cell. Proteomics* 6.9, pp. 1527–1550.

Wang, LiGuo and LiGe Tonggu (Jan. 2015). “Membrane protein reconstitution for functional and structural studies”. In: *Sci. China Life Sci.* 58.1, pp. 66–74. arXiv: 15334406.

Ward, Andrew B., Andrej Sali, and Ian A. Wilson (Feb. 2013). “Integrative Structural Biology”. In: *Science (80-.).* 339.6122, pp. 913–915. arXiv: arXiv:1011.1669v3.

Wennerström, Håkan (Apr. 1979). “Micelles. Physical chemistry of surfactant association”. In: *Phys. Rep.* 52.1, pp. 1–86.

White, Stephen H (July 2004). “The progress of membrane protein structure determination”. In: *Protein Sci.* 13.7, pp. 1948–1949.

White, Stephen H. (2018). *http://blanco.biomol.uci.edu/mpstruc/.*

Whitten, Andrew E. et al. (Apr. 2007). “The Structure of the KinA-Sda Complex Suggests an Allosteric Mechanism of Histidine Kinase Inhibition”. In: *J. Mol. Biol.* 368.2, pp. 407–420.

Wild, Rebekka et al. (Feb. 2018). “Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation”. In: *Science (80-.).* 359.6375, pp. 545–550. arXiv: NIHMS150003.

Winchester, Bryan G. (Jan. 2001). “Lysosomal membrane proteins”. In: *Eur. J. Paediatr. Neurol.* 5.SUPPL. A, pp. 11–19.

Wolfe, Aaron J. et al. (2017). “Quantification of Membrane Protein-Detergent Complex Interactions”. In: *J. Phys. Chem. B* 121.44, pp. 10228–10241.

Wyhe, John van (July 2016). “The evolution revolution”. In: *New Sci.* 231.3082, pp. 35–39.

Yamamoto, Kazutoshi, Ronald Soong, and Ayyalusamy Ramamoorthy (2009). “Comprehensive analysis of lipid dynamics variation with lipid composition and hydration of bicelles using nuclear magnetic resonance (NMR) spectroscopy”. In: *Langmuir* 25.12, pp. 7010–7018.

Yang, Kui and Xianlin Han (2016). “Lipidomics: Techniques, Applications, and Outcomes Related to Biomedical Sciences”. In: *Trends Biochem. Sci.* 41.11, pp. 954–969. arXiv: 15334406.

Yeh, Vivien et al. (2018). “Highly Efficient Transfer of 7TM Membrane Protein from Native Membrane to Covalently Circularized Nanodisc”. In: *Sci. Rep.* 8.1, pp. 1–11.

Yoshikawa, H et al. (2016). “High-resolution cryo-EM analysis of the yeast ATP synthase in a lipid membrane”. In: *Science (80-.).* 14.October 2012, pp. 39–57. arXiv: 15334406.

Yusuf, Yustina et al. (Feb. 2018). “Optimization of the Production of Covalently Circularized Nanodiscs and Their Characterization in Physiological Conditions”. In: *Langmuir* 34.11, pp. 3525–3532.

Zaccai, Nathan R. et al. (2016). “Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures and Releases Unfolded Outer Membrane Proteins”. In: *Methods Enzymol.* Ed. by Zvi Kelman. Vol. 566. Academic Press. Chap. Methods in, pp. 159–210. arXiv: 15334406.

Zhang, Qinghai et al. (2007). “Designing facial amphiphiles for the stabilization of integral membrane proteins”. In: *Angew. Chemie - Int. Ed.* 46.37, pp. 7023–7025.

Zhang, Sensen et al. (2017). “Cryo-EM structures of the mammalian endo-lysosomal TRPML1 channel elucidate the combined regulation mechanism”. In: *Protein Cell* 8.11, pp. 834–847.

Zoonens, Manuela and Jean-Luc Popot (2014). “Amphipols for Each Season”. In: *J. Membr. Biol.* 247.9-10, pp. 759–796. arXiv: NIHMS150003.

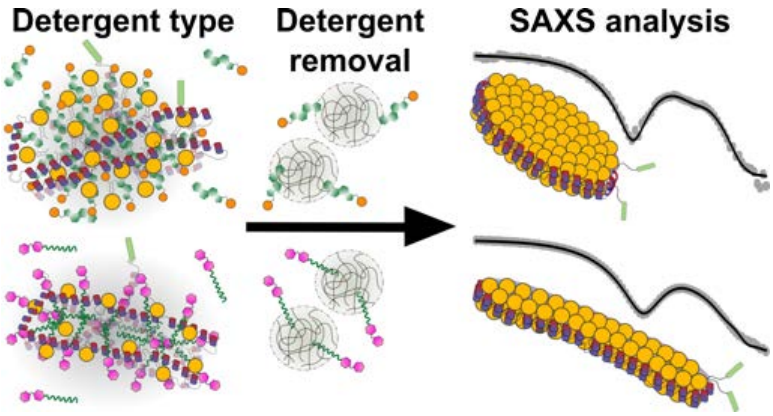
Part II

PAPERS

4

PAPER I

COMPREHENSIVE STUDY OF THE SELF-ASSEMBLY OF PHOSPHOLIPID NANODISCS: WHAT DETERMINES THEIR SHAPE AND STOICHIOMETRY?



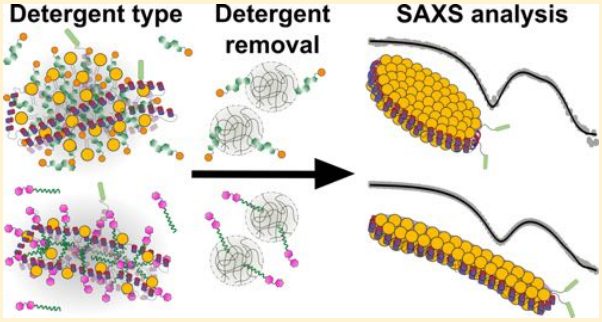
LANGMUIR
Article
Cite This: Langmuir 2018, 34, 12569−12582
pubs.acs.org/Langmuir

Comprehensive Study of the Self-Assembly of Phospholipid Nanodiscs: What Determines Their Shape and Stoichiometry?

Nicholas Skar-Gislinge,[†] Nicolai Tidemand Johansen, Rasmus Høiberg-Nielsen,[‡] and Lise Arleth*

Structural Biophysics, Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

ABSTRACT: Phospholipid nanodiscs have quickly become a widely used platform for studies of membrane proteins. However, the molecular self-assembly process that ultimately should place a membrane protein inside a nanodisc is not well understood. This poses a challenge for a successful high-yield reconstitution of general membrane proteins into nanodiscs. In the present work, the self-assembly process of POPC-MSP1D1 nanodiscs was carefully investigated by systematically modulating the reconstitution parameters and probing the effect with a small-angle X-ray scattering analysis of the resulting nanodiscs. First, it was established that nanodiscs prepared using the standard protocol followed a narrow but significant size distribution and that the formed nanodiscs were stable at room temperature over a time range of about a week. Systematic variation of the POPC/MSP1D1 stoichiometry of the reconstitution mixture showed that a ratio of less than 75:1 resulted in lipid-poor nanodiscs, whereas ratios of 75:1 and larger resulted in nanodiscs with constant POPC/MSP1D1 ratios of 60:1. A central step in the self-assembly process consists in adding detergent-absorbing resin beads to the reconstitution mixture to remove the reconstitution detergent. Surprisingly, it was found that this step did not play a significant role for the shape and stoichiometry of the formed nanodiscs. Finally, the effect of the choice of detergent used in the reconstitution process was investigated. It was found that detergent type is a central determining factor for the shape and stoichiometry of the formed nanodiscs. A significantly increasing POPC/MSP1D1 stoichiometry of the formed nanodiscs was observed as the reconstitution detergent type is changed in the order: Tween80, DDM, Triton X-100, OG, CHAPS, Tween20, and Cholate, but with no simple correlation to the characteristics of the detergent. This emphasizes that the detergents optimal for solution storage and crystallization of membrane proteins, in particular DDM, should not be used alone for nanodisc reconstitution. However, our data also show that when applying mixtures of the reconstitution detergent cholate and the storage detergents DDM or OG, cholate dominates the reconstitution process and nanodiscs are obtained, which resemble those formed without storage detergents.



INTRODUCTION

Phospholipid nanodiscs[1] prepared from phospholipid bilayer disks encircled by two membrane scaffold proteins (MSP1D1) are widely applied for the investigation of membrane proteins in solution[2] and have quickly found a broad range of applications within studies of membrane proteins. The field has been recently reviewed by Denisov and Sligar.[3] When studying this quickly expanding literature, it is striking that the majority of the studies focus on functional investigations of membrane proteins inside nanodiscs, while only a minor fraction of the literature reports from structural studies. This observation is even more astounding when taking into account the general lack of structural information about membrane proteins combined with their widely accepted biological and pharmaceutical importance. About 30% of the proteins coded for by the human genome are believed to be membrane proteins. However, at present, only 800[a] of the about 60 000[b] unique protein structures published in the Protein Data Bank correspond to membrane protein structures. The commonly accepted explanation for the low number of membrane protein crystal structures is that membrane proteins are notoriously difficult to handle and to achieve crystallization from. With the recent increased resolution of cryo-electron microscopy (cryo-EM), information about new important membrane protein structures are becoming available. However, still the sample preparation includes a membrane protein reconstitution step as one of the several bottlenecks.[4]

Nanodiscs are most commonly self-assembled by a detergent-driven reconstitution process (see Figure 1) closely resembling the approaches applied for detergent-based liposome formation[5−8] and for reconstitution of membrane proteins into liposomes.[9−11] The chosen phospholipids are mixed with a so-called reconstitution detergent having a high critical micellar concentration (CMC) compared to the phospholipid. In the original nanodisc reconstitution protocol from the Sligar lab[1,12] cholate, with a CMC of ∼10 mM, is chosen as the reconstitution detergent. Phospholipid−cholate mixed micelles are spontaneously formed by mixing them in

Received: May 14, 2018
Revised: August 14, 2018
Published: September 21, 2018

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

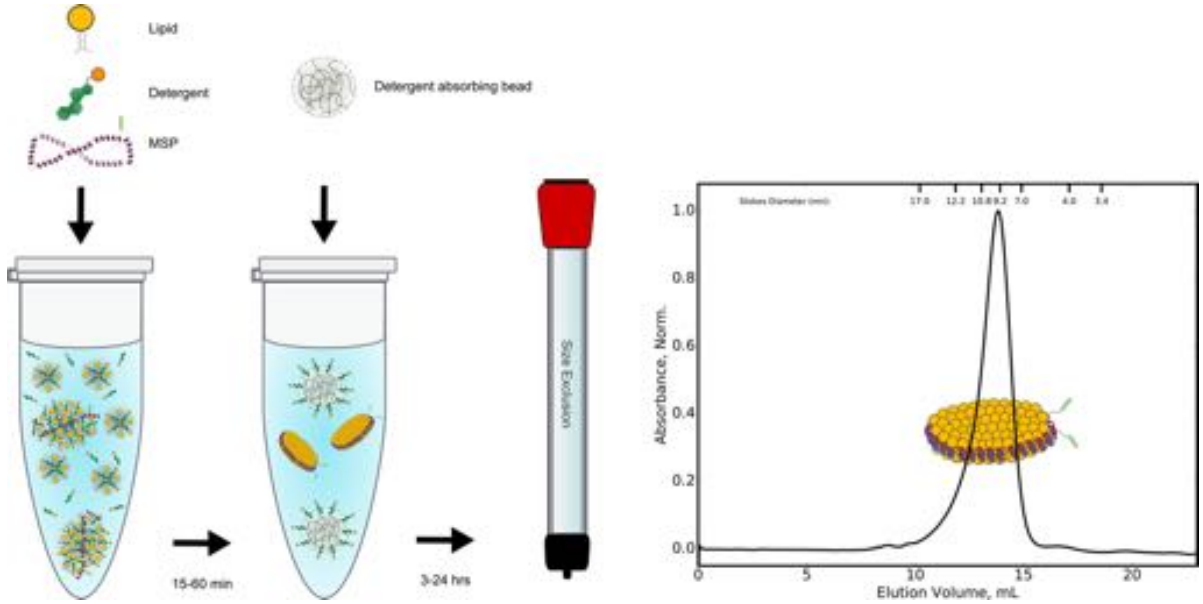


Figure 1. Nanodisc reconstitution process. Membrane scaffold protein (MSP), phospholipid, and reconstitution detergent are mixed in a buffer solution and left to equilibrate. Detergent-absorbing resin beads are added to remove the detergents. This leads to phospholipid−MSP nanodiscs and possibly other aggregates. The sample is then run over an SEC column to purify the formed nanodiscs.

aqueous buffer at a cholate concentration above its CMC.[5,11] To form nanodiscs, an amphipatic membrane scaffolding protein is introduced to stabilize the rims of the disks. In the Sligar protocol, the so-called MSP1D1 or its derivatives are introduced into the buffer and is believed to mix into the phosholipid−cholate mixed micelles. MSP1D1 is genetically engineered with close resemblance to the human ApoA1 protein and is an amphipatic protein with a secondary structure of 10 α-helices. Two MSP1D1s combine to encircle and stabilize the rims of the nanodiscs and thereby define the nanodisc circumference. MSP1D1-based nanodiscs have a diameter close to 10 nm, but by varying the length of the MSP through the number of inserted α-helices, smaller or larger disks may be stabilized with diameters ranging from 7 to 15 nm.[12,13] After the phospholipid−cholate−MSP sample has equilibrated, detergent-absorbing polymeric resin beads are added to the sample. These resin beads are colloidal particles with a highly porous and hydrophobic interior,[11,14] and monomers of the high-CMC cholate will relatively quickly adsorb to the resin interior[11] and thus be removed from the mixed phospholipid−MSP1D1−cholate particles. This process leads to the formation of monodisperse phospholipid−MSP1D1 nanodiscs, and in most cases, also to larger aggregates, which are removed in a subsequent purification step by size exclusion chromatography (SEC). If membrane proteins are to be incorporated into the nanodiscs, these are added before the detergent removal step.

The nanodisc platform for reconstituting membrane proteins has the advantage over detergent systems of providing a more nativelike phospholipid environment. Phospholipid liposomes are for obvious reasons also widely used for investigations of membrane proteins and have the intrinsic possibility of having a functionally relevant inside and outside environment and an associated membrane potential. However, phospholipid liposomes are generally less suited for structural studies than nanodiscs, among other reasons due to their low mass ratio of membrane protein to phospholipid.

Despite several attempts,[c] nanodiscs or membrane proteins inside nanodiscs have never been crystallized, and there is only a single report in the literature that makes use of the MSP−nanodisc system to obtain membrane protein crystals.[15] However, solution-based techniques providing intrinsically lower structural resolution, such as electron microscopy, including cryo-EM[16,17] and small-angle scattering, have been able to elucidate the overall structure of membrane proteins based on the nanodisc platform.[2,18,19] In addition, high-resolution nuclear magnetic resonance (NMR) techniques are able to provide crucial structural information about small membrane proteins reconstituted into nanodiscs.[13,20] A recent NMR study even provided a full molecular structure of a slightly truncated version of the traditional MSP1D1 nanodiscs.[21] The number of structural results based on these techniques are, however, also still relatively limited. We hypothesize that a central explanation for this is that it is challenging to obtain a structurally homogeneous reconstitution of a membrane protein in nanodiscs. In contrast, nanodisc-enabled functional studies, which are indeed very abundant by now, may be performed at very low concentrations and even on samples where only a fraction of membrane proteins is successfully reconstituted.

To establish the knowledge necessary to develop a better control of the nanodisc self-assembly, we have, in the present work, systematically investigated the self-assembly process by varying the reconstitution conditions and correlating this with the detailed structure of the resulting nanodiscs as determined by small-angle X-ray scattering (SAXS) and a structural modeling of the SAXS data. The investigation of the individual steps in the self-assembly process is strongly inspired by the very detailed analysis conducted in the 1980s and the 1990s by Rigaud and co-workers in relation to liposome formation.[9−11]

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

The study presented in this article starts from an investigation of the homogeneity and stability of the ensemble of POPC-MSP1D1 nanodiscs formed within a standard preparation through combined SEC-SAXS and static light scattering (SLS) (case A). Then, the effect of varying the POPC/MSP1D1 ratio in the reconstitution mixture is examined (case B), as well as the effect of the detergent removal speed (case C). Finally, the effect of the choice of the reconstitution detergent is investigated (case D). Overall, we find that the nanodiscs formed when using the standard Sligar protocol are homogeneous but with an underlying small variation of the POPC/MSP1D1 ratio in the ensemble of nanodiscs and that the formed nanodiscs are indeed very stable over time. In contrast to liposome reconstitution, the detergent removal speed plays no role in the structure of the formed disks; however, the choice of reconstitution detergent has a major effect on the structure of the formed disks. We therefore suggest that for optimization of protocols for membrane protein reconstitution in nanodiscs, given that the sample stoichiometries are reasonably chosen, the main impact is to be expected from the choice of reconstitution detergents and from the balance between reconstitution detergents and membrane protein storage detergents.

MATERIALS AND METHODS

Materials. 1-Palmitoyl-2-oleyl-*sn*-glycero-3-phosphocholine (POPC) was obtained from Avanti Polar Lipids (Alabaster, Alabama). n-dodecyl β-D-maltoside detergent (DDM) and n-octyl β-D-glucoside (OG) were obtained from Carbosynth (Compton, UK). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was obtained from Fisher Scientific (Hampton, New Hampshire). Trizma base, sodium chloride, sodium cholate, ammonium molybdate tetrahydrate, ethylenediaminetetraacetic acid (EDTA), guanidinium hydrochlorid (GuHCl), dithiothreitol (DTT), 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS), Tween20, Tween80, Triton X-100 (TX100), and Amberlite XAD-2 were obtained from Sigma-Aldrich (Darmstadt, Germany). All chemicals were used as received and without further purification. Amberlite XAD-2 was washed once in methanol and three times in Milli-Q water and used as the detergent-absorbing polymer resin beads in the present study.

MSP1D1 Protein Expression and Purification. pET28(+) encoding the His-tagged MSP1D1 gene was kindly provided by Steven Sligar (University of Illinois). The plasmid was transformed into *E. coli* BL21 star DE3 (Novagen) and plated on LB-agar containing kanamycin. A single colony was picked and inoculated in LB-containing kanamycin and grown overnight at 37 °C. The next day, the culture was diluted to an optical density of 600 nm (OD_{600}) = 0.05, grown to OD_{600} = 0.6 and induced with 1 mM IPTG for 3 h, all at 37 °C. The cells were pelleted by centrifugation and resuspended in 3.5 mL of buffer A (50 mM Tris-HCl, pH 8, 300 mM NaCl, 20 mM imidazole, 6 M GuHCl) per gram of cell wet weight. The cells were lysed by shaking and freeze/thawing, debris was pelleted by centrifugation, and the supernatant was loaded on a column loaded with NiNTA resin (Qiagen) equilibrated in buffer A. The column was washed in three column volumes of buffer A, three column volumes of wash buffer (50 mM Tris-HCl, pH 8, 300 mM NaCl, 40 mM imidazole) including 10 mM sodium cholate, three column volumes of wash buffer, and protein was eluted in buffer B (50 mM Tris-HCl, pH 8, 300 mM NaCl, 250 mM imidazole). Tobacco etch virus (TEV) protease was added to the protein sample, and the mix was dialyzed 100-fold against TEV buffer (50 mM Tris-HCl, pH 8, 100 mM NaCl, 1 mM EDTA, 1 mM DTT) overnight, producing N-terminally cleaved MSP1D1 lacking the His-tag. MSP1D1 was separated from the His-tagged TEV protease and residual His-tagged MSP1D1 on NiNTA resin. MSP1D1 was concentrated in vivaspin20 10k molecular weight cutoff Satorius concentration columns to 5 mg/mL, flash-frozen, and stored at −80 °C.

Nanodisc Preparation. Nanodiscs were prepared as described elsewhere[12] for all but case D. Briefly, POPC was dissolved in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 100 mM cholate to a concentration of 50 mM using vortexing and sonication, mixed with MSP1D1 in the appropriate POPC/MSP1D1 molar ratio, and then diluted with gel filtration buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl) to yield a final POPC concentration of 10 mM and a cholate concentration of 20 mM. The sample was thoroughly mixed and Amberlite XAD-2 beads were added in appropriate amounts and incubated at 5 °C overnight. The following day, the samples were separated from the beads and gel-filtrated on a SuperDex 200 10/300GL column (GE Healthcare) equilibrated in gel filtration buffer. The samples were kept cold until use.

Case A: Homogeneity and Stability of the Formed Nanodiscs. Two sets of samples were investigated by SAXS based on a nanodisc sample prepared using a POPC/MSP1D1 reconstitution ratio of 75:1. One experiment was conducted in on-line SEC-SAXS mode and the other in a serial mode with SEC, followed by SAXS of the individual fractions. In the latter case, the samples were gel-filtrated on site and measured immediately after elution.

Case B: Nanodiscs Produced at Different POPC/MSP1D1 Stochiometries. The POPC/MSP1D1 ratio of the reconstitution mixture was systematically varied by adaptation of the MSP1D1 concentration to a fixed 10 mM POPC concentration. Samples based on POPC/MSP1D1 ratios of 30:1, 45:1, 55:1, 65:1, 75:1, 85:1, and 110:1 were prepared.

Case C: Nanodiscs Produced at Different Detergent Removal Speeds. Excess water was removed from hydrated Amberlite XAD-2 beads, which were then weighed. Two series of samples were prepared from the standard protocol, both using a POPC/MSP1D1 ratio of 75:1. In the first series, the absorbing beads were added in appropriate volumes to yield samples containing bead concentrations of 0.11, 0.14, 0.18, 0.23, 0.30, 0.38, 0.50, 0.64, 0.83, and 1.06 g/mL. In the second series, beads were added to yield concentrations of 0.001, 0.01, 0.1, and 1.0 g/mL.

Case D: Nanodiscs Produced with Different Reconstitution Detergents. Both the CMC of a detergent and its ability to solubilize phospholipid bilayers have to be taken into account when comparing across detergent types.[22,23] For this reason, the detergent concentrations for the different reconstitutions were modified compared to the cholate reconstitution to take into account the relevant CMC as well as the detergent concentration required to solubilize a vesicle bilayer at 10 mM POPC (or in some cases, EYPC), which we denote C_{sol} (see Table 1). Detergent concentrations, $C_{reconst}$, slightly higher than C_{sol} were systematically used to assure that the lipids and

Table 1. Solution Behavior and Solubilization Characteristics of the Investigated Detergents: Critical Micellar Concentration (CMC) and Micellar Aggregation Number (N_{agg})[a]

reconstitution detergent	CMC (mM)	N_{agg}	C_{sol} (mM)	$C_{reconst}$ (mM)
cholate	10[25]	2−3	12.1[26]	20
CHAPS	8	10	15[27]	40
Tween20	0.059		60[24]	50
Tween80	0.012	60	250[24]	50
DDM	0.17	98[28]	20	30
TX100	0.23	100−155	20−40[29,30]	50
OG	18−20	84	42[31]	100

[a]When nothing else is noted, values for CMC and N_{agg} are taken from Sigma-Aldrich's detergent handbook.[32] C_{sol} denotes the detergent concentration required to solubilize a POPC or EYPC vesicle bilayer at 10 mM. References to the obtained valued inserted in the table. $C_{reconst}$ denotes the detergent concentration applied for the nanodisc reconstitution (based on 10 mM POPC) in the present work.

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

detergent were in a mixed micelle phase (see Table 1). For Tween80 and Tween20, we deviated slightly from this approach as we observed that the solutions became clear (indicating transition from vesicles to small mixed micelles) at much lower concentrations than proposed by the results reported in the literature.[24] POPC and the used amount of detergent were solubilized in chloroform and mixed in appropriate ratios to ensure complete mixing. DDM, Tween20, and Tween80 were not soluble in pure chloroform, but solubilized in 50/50% v/v chloroform/ethanol. The mixtures were thoroughly mixed by bath sonication, then dried under a stream of nitrogen and subsequently under high vacuum overnight. Next day, the POPC/detergent mix was resuspended in gel filtration buffer and mixed with MSP1D1 to yield a final concentration of 10 mM POPC and detergent concentrations of $C_{reconst}$ (see Table 1). Hereafter, the standard protocol was followed to produce the final nanodiscs. For nanodisc samples reconstituted from a mix of cholate and either DDM or OG, the standard protocol was followed replacing a small volume of gel filtration buffer with either 1% DDM or 10% OG in gel filtration buffer, respectively. For the (MSP1D1-free) mixed micelles also analyzed under case D, the same protocol was applied but without adding MSP1D1.

Phosphorus Analysis. The POPC/MSP1D1 stoichiometry of the formed nanodiscs was determined through a phosphorus analysis.[33] This analysis was carried out by chemical digestion of phospholipids in H_2SO_4 to yield free phosphate, which was reacted with molybdate to produce a blue product, whose absorption was measured at 812 nm. Phospholipids were quantified using a standard curve obtained from standards containing known amounts of disodium phosphate. From previous experiments, this method is found to be accurate, at least within ±10%.

Light Scattering. Dynamic light scattering (DLS) and static light scattering (SLS) were measured on a Brookhaven BI-200SM (Brookhaven Instruments Corporation, New York) with the detector placed at 90°. The temperature in the vat was controlled and kept at 20 ± 0.1 °C using an external water bath. For measuring the temporal stability of nanodiscs (Figure 4A), 30 s SLS and DLS frames were recorded over the course of 6 days. The DLS data were binned in 300 s frames, and representative frames for different time points were fitted with the DLS analysis tool available at bayesapp.org,[34] which yields an intensity-weighted distribution of the particle sizes as a function of the particle hydrodynamic radius, R_h. The DLS data for POPC−detergent-mixed premicelles (Figure 7) were analyzed with the second-order cumulant method[35] to yield the average particle size and the width of the particle size distribution.

SAXS Experiments. For cases A and D, SAXS experiments were performed partly at the ID13-4 BioSAXS beamline and partly at the BM29 BioSAXS beamline, both at ESRF, Grenoble, France.[36] Case B samples were measured at the I911 SAXS beamline at MAXLAB, Lund Sweden,[37] and samples of cases C and D were measured partly at the ID13-4 BioSAXS beamline[36] at ESRF, Grenoble, France, and partly at P12 at PETRAIII at DESY, Hamburg, Germany.[38] The data were reduced according to standard procedures using the beamline software at the relevant beamlines and scaled to absolute scattering intensity, $I(q)$, in units of 1/cm as a function of q, where $q = 4\pi \sin\theta/\lambda$, where θ is half the scattering angle and λ is the wavelength of the incoming beam. Absolute intensity calibration was conducted using a secondary protein standard[39] and/or Milli-Q water.[40] The indirect Fourier transforms[41] (IFT) to determine the pair-distance distribution functions, $p(r)$, were performed using a Bayesian version of the IFT implementation.[42]

THEORY AND SAXS DATA ANALYSIS

The model based analysis of the measured SAXS data from nanodisc samples are performed as described in our previous articles.[43,44] Generally, the small-angle X-ray scattering intensity of a dilute suspension of particles can be expressed by

$$I(q) = n\Delta b^2 P(q) \quad (1)$$

where n is the particle number density; Δb is the particle excess scattering length, which may also be expressed as $\Delta b = V\Delta\rho$, where V is the single particle volume and $\Delta\rho$ is the average excess scattering length density of the particle compared to the solvent; and $P(q)$ is the single-particle form factor intensity and contains the information about the particle shape. It is normalized such that $P(0) = 1$. The amplitude form factor, $A(q)$, is related to the intensity form factor by $P(q) = A(q) \times A^*(q)$, where $A^*(q)$ is the complex conjugate of $A(q)$.

For SAXS, the amplitude form factor, $A(q)$, is calculated from the Fourier transform of the excess scattering length density distribution, $\Delta\rho(r)$, which is directly proportional to the excess electron density distribution of the particle. Because of this, the form factor for various geometric objects can readily be calculated in a semianalytical manner[45] and combined to yield even more complex objects. This principle is used to construct a model of the small-angle scattering from nanodiscs, which are represented as an assembly of geometrical shapes: The central phospholipid bilayer is represented as a stack of five cylinders with an elliptical cross section. Each cylinder represents a different molecular part of the phospholipid bilayer. The top and bottom cylinder layers represent the phospholipid head groups and associated hydration water. The second and fourth layers represent the alkyl tails of the phospholipids, and the third/middle layer represents the methyl end groups of both the top and bottom phospholipids. The two MSP1D1s spanning the rim of the bilayer are represented by a hollow cylinder with elliptical cross section.

To minimize the number of free parameters in the model, constraints based on the molecular composition of the nanodisc have systematically been incorporated. For example, the volume of the cylindrical stack representing the phospholipid bilayer is constrained to the phospholipid aggregation number times the partial molecular volume of a phospholipid molecule. By fitting the model to the measured data, we can obtain the structural parameters describing the nanodisc.[43] A detailed description of the model can be found in Skar-Gislinge and Arleth.[44] For the analysis in the present work, the following model parameters were fitted: Axis ratio, lipid aggregation number, number of hydration water molecules per lipid head group, lipid molecular volume, MSP molecular volume, interface roughness parameter, and a constant background. The disks in case B have two histidine tags attached, which are modeled by two Gaussian random coils, adding the coil radius of gyration to the fitted parameters in this case. For these parameters, good initial guesses were available based on the molecular information about the system and, indeed, the fitted values landed close to the prior guesses. The final parameter in the model is the area per lipid molecule in the bilayer plane, which was locked to 70 Å^2 based on previous results, to minimize the number of free parameters.[43] As most of the central description of the nanodiscs is captured in the obtained values for the number of phospholipids per nanodisc and in the related value for the ellipticity in terms of the axis ratio, we have chosen only to report these values.

RESULTS

Case A: Homogeneity and Stability of the Formed Nanodiscs. Figure 2A shows the SEC elution profile and the obtained SEC-SAXS data of a POPC nanodisc preparation. The black line shows the absorption at 280 nm, proportional to the protein concentration. The dots indicate the forward

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

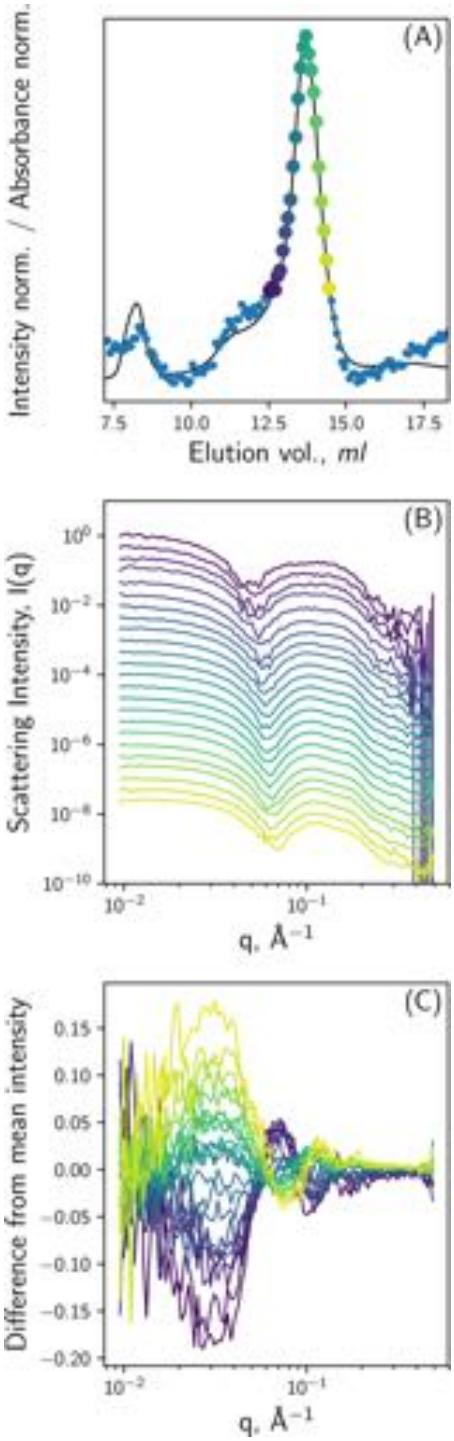


Figure 2. SEC elution profile of POPC nanodiscs. (A) Forward scattering intensity, $I(0)$, as dots scaled to the measured UV absorption at 280 nm (continuous black curve). The highlighted points correspond to the scattering curves shown in (B). (B) Scattering profiles corresponding to the points shown in (A) (data on arbitrary scale and rescaled for clarity, same color code in (A)–(C)). (C) Difference plots of the scattering profiles to the average scattering profile.

scattering, $I(0)$ (estimated as the average intensity of the first five points in each scattering curve). The two are scaled to one another. The SAXS data, $I(q)$, corresponding to the highlighted points in Figure 2A are plotted in Figure 2B. These data display a scattering profile characteristic of nanodiscs,[43] but with some systematic variation along the elution peak. It is seen that the position of the first minimum increases and that the shape of the broad bump around 0.1 Å^{-1} evolves as a function of elution volume. To emphasize this variation, Figure 2C shows the difference between each normalized curve and the average of the normalized curves. This plot removes the common behavior of the curves and clearly shows that the shape of the $I(q)$ changes gradually and systematically along the whole elution profile. The data show that the peak fraction of the nanodisc elution profile contains a slightly heterogeneous population of nanodiscs. This observed small heterogeneity is fully consistent with previous reports on similar systems.[46−48]

Inspection of the SEC elution profile of Figure 2A suggested that this observed heterogeneity might be due to the overlap between the small shoulder seen between 10 and 12.5 mL and the main peak at 13 mL. To investigate whether this was the case, a second data set was collected (see Figure 3). Here, the peak fraction from the SEC-run of a nanodisc preparation was first collected and then rerun on the same column. The obtained chromatogram (Figure 3A) clearly shows that the contribution from the shoulder fraction is minimal in the peak fraction. From this SEC-run, five peak fractions were collected for further analysis using SAXS. Due to the low concentration of the samples after two consecutive SEC runs, the five fractions were simply measured in a standard solution SAXS setup, which allowed for longer measurement times and better counting statistics. The obtained SAXS data of the five fractions are plotted in Figure 3B as dots along with the nanodisc model fits to the data (continuous black curves). Again, there is a small but clearly visible development of the shape of $I(q)$ implying a variation in the structure of the formed nanodiscs. Figure 3C shows the POPC/MSP1D1 stoichiometry of the nanodiscs as a function of elution volume, as determined via the nanodisc model fit to the SAXS data. A decrease from an initial value of about 65 POPC/MSP1D1 to slightly below 50 POPC/MSP1D1 is observed as a function of elution volume. Figure 3D shows the nanodisc axis ratio obtained from the nanodisc model fits to the SAXS data. This value increases from ∼1.35 to ∼1.55 as a function of elution volume.

The model fits including the obtained values for the POPC/MSP1D1 ratio (Figure 3C) show that the heterogeneity observed for the main peak can be understood as a distribution in the number of incorporated lipids in the nanodiscs, which modulate both their size and elliptical shape: The first fraction is found to have a smaller axis ratio and a higher number of lipids per nanodisc (given as twice the POPC/MSP1D1 ratio) than the nanodiscs in the later fractions, which have a lower number of lipids per nanodisc and become more elliptical (Figure 3C,D).[43]

To assess the nanodisc stability over time, a nanodisc sample from a standard preparation was investigated in a combined static and dynamic light scattering (SLS/DLS) experiment. The sample was kept at 20 °C and left in the instrument so that the count rate, i.e., the static part of the signal, of the light scattered at a 90° angle could be continuously measured over the time course of 6 days. The resulting SLS data, plotted in

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

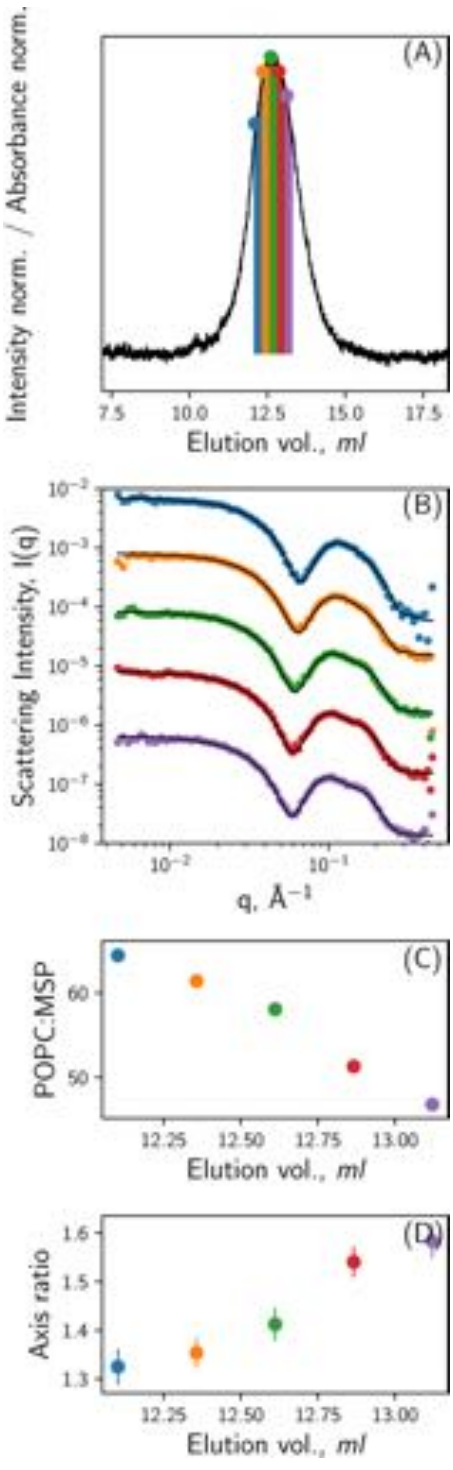


Figure 3. (A) SEC elution profile of the peak fraction of POPC nanodiscs. Forward scattering intensity, $I(0)$ (dots), of the collected fractions are scaled to the UV absorption at 280 nm (continuous black curve). (B) SAXS data (dots) from the five collected fractions, along with the corresponding nanodisc model fits (continuous black curves). (C) POPC/MSP1D1 ratios derived from the model fits and plotted as a function of elution volume. Errors on the determined values are smaller than the size of the symbols. (D) Axis ratios derived from the model fits plotted as a function of elution volume.

Figure 4A, show a very minor increase of the scattering signal during most of the experiment. Only after 140 h and toward

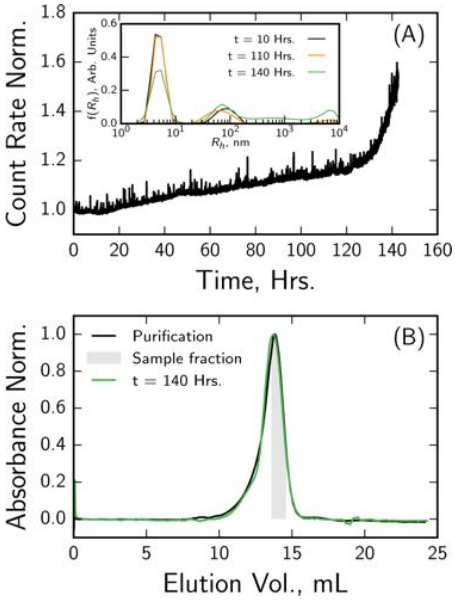


Figure 4. (A) SLS data from a nanodisc sample plotted as a function of time. The count rate is normalized by the count rate at the beginning of the experiment. The experiment was conducted at 20 °C. (B) SEC analysis of the nanodisc sample performed before and after the experiment in (A). The sample fraction for the light scattering experiment is indicated in the plot.

the end of the experiment, the intensity starts to increase more rapidly, indicating the formation of larger particles. The DLS part of the signal was analyzed by converting the obtained correlation curves into a distribution of hydrodynamic radii, R_h. The R_h distributions from the beginning of the experiment (10 h) and before (110 h) and after (140 h) the onset of the SLS signal increase of the SLS signal are plotted in the inset of Figure 4A. It is seen that at the late times, a significant fraction of the scattering comes from large particles that are not present in the first part of the experiment. Keeping in mind that light scattering is highly sensitive to the presence of a few large particles, an SEC measurement was performed on the nanodisc sample at the end of the experiment. The data (Figure 4B) show that the SEC chromatogram of the sample run after 140 h at 20 °C overlay almost completely with the t = 0 sample, indicating that the nanodisc sample was only negligibly deteriorated or otherwise modulated after 6 days at 20 °C in the light scattering instrument. In particular, no indications of the large particles visible in the SLS/DLS were visible in the subsequent SEC. On the basis of the present data, we cannot tell whether the explanation for this is that the large particles, despite scattering strongly due to their size, constitute only a very minor fraction of the sample that is not visible in the SEC or that the large particles are loosely bound agglomerates of nanodiscs that fall apart again in the SEC dilution. The data nevertheless clearly show that the nanodiscs are indeed very stable for several days, in particular when it is taken into account that nanodiscs are usually stored in cold conditions, i.e., at ∼5 °C.

Overall, we conclude from case A that despite that nanodiscs can be prepared to be very monodisperse, a nanodisc sample

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

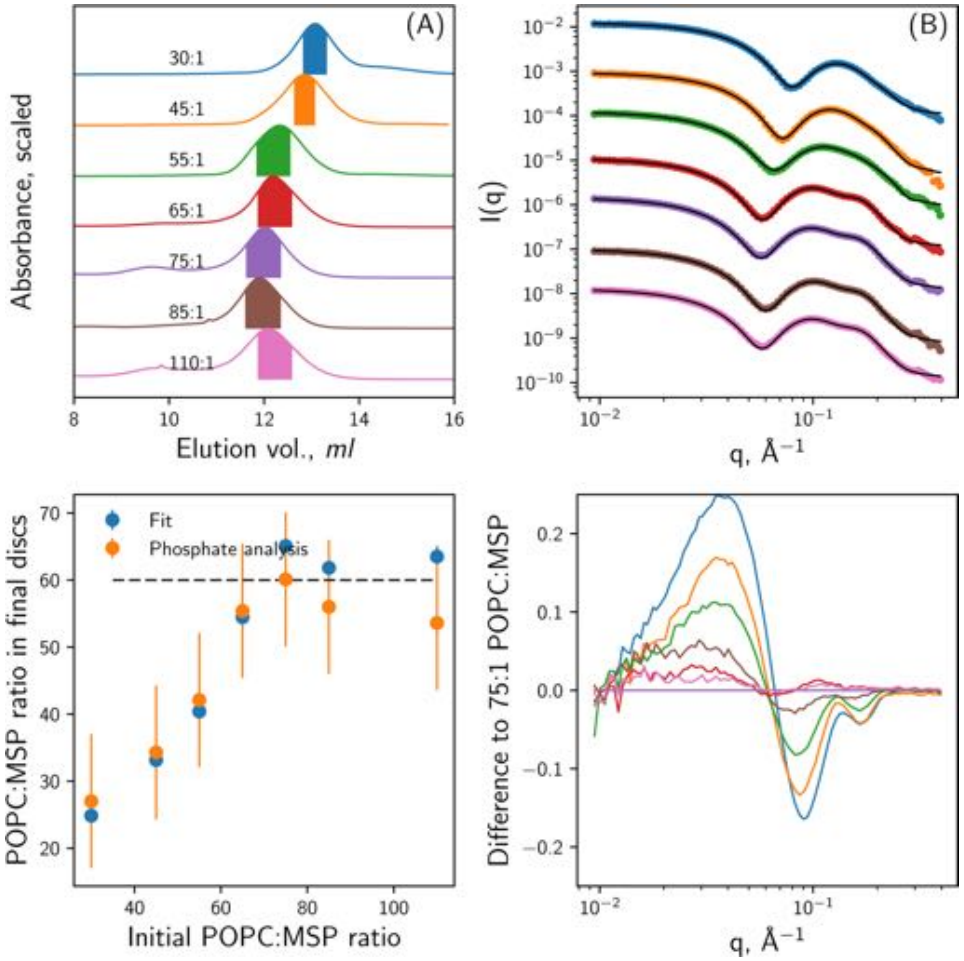


Figure 5. (A) SEC data of nanodiscs prepared at POPC/MSP1D1 ratios from 30:1 to 110:1. The colored areas mark the fractions collected for subsequent SAXS analysis. (B) SAXS data (points) along with structural model fits (full curves) for the different POPC/MSP1D1 preparations. (C) POPC/MSP1D1 ratio of the formed nanodiscs plotted as a function of POPC:MSP ratio of the reconstitution mixture. (D) Difference of the scattering data compared to nanodisc prepared with an initial POPC/MSP1D1 ratio of 75:1.

has a small but significant dispersity in the POPC/MSP1D1 ratios of the underlying disk and consequently in the axis ratios of the formed disks. The disks are structurally stable over several days at room temperature, and while we start to observe a small fraction of large particles after 140 h, the major part of the nanodiscs maintains their integrity.

Case B: Nanodiscs Produced at Different POPC/MSP1D1 Stochiometries. Figure 5A shows SEC data from nanodiscs prepared with different POPC/MSP1D1 reconstitution ratios and indicates the collected peak fractions. From a POPC/MSP1D1 ratio of 110:1 to 65:1, there is only a little change in the retention volume, which stays close to 12 mL. However, below a POPC/MSP1D1 ratio of 65:1, the retention volume increases to about 13 mL with decreasing POPC/MSP1D1 ratio, indicating that the hydrodynamic radius of the self-assembled particles decrease with a decreasing initial POPC/MSP1D1 ratio. Figure 5B shows the SAXS data from the collected peak fractions. The obtained $I(q)$ shows a displacement of the characteristic minimum from q = 0.08 to 0.058 Å^{-1} as the POPC/MSP1D1 increases. Also the shape of the maximum at around q = 0.15 Å^{-1} changes as the reconstitution stoichiometry increases until a reconstitution ratio of 65:1, above which the data are indistinguishable. Again, to emphasize the variation, Figure 5D shows the difference between each normalized curve and the data obtained with the 75:1 reconstitution ratio. This plot clearly shows that the main deviation is obtained upon "under loading" of the disks with a too low POPC/MSP1D1 ratio.

The SAXS data were again fitted using the nanodisc model,[43,44] and the resulting model fits are shown in black in Figure 5B. The fits provide values for the POPC/MSP1D1 ratio of the formed nanodiscs (Figure 5C). The values fall systematically below the POPC/MSP1D1 ratio of the reconstitution mixture. As an independent control, the POPC/MSP1D1 ratio was determined from a direct measurement of the ratio between the total phosphorus concentration and the protein concentration in each sample. These values are also plotted as a function of the initial POPC/MSP1D1 ratio in Figure 5C. The POPC/MSP1D1 ratios obtained from the two independent methods are seen to be in very good agreement within the precision of the measurements, and in both cases, we find that the disks can maximally be loaded to have a ratio close to 60:1 in the final disks.

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

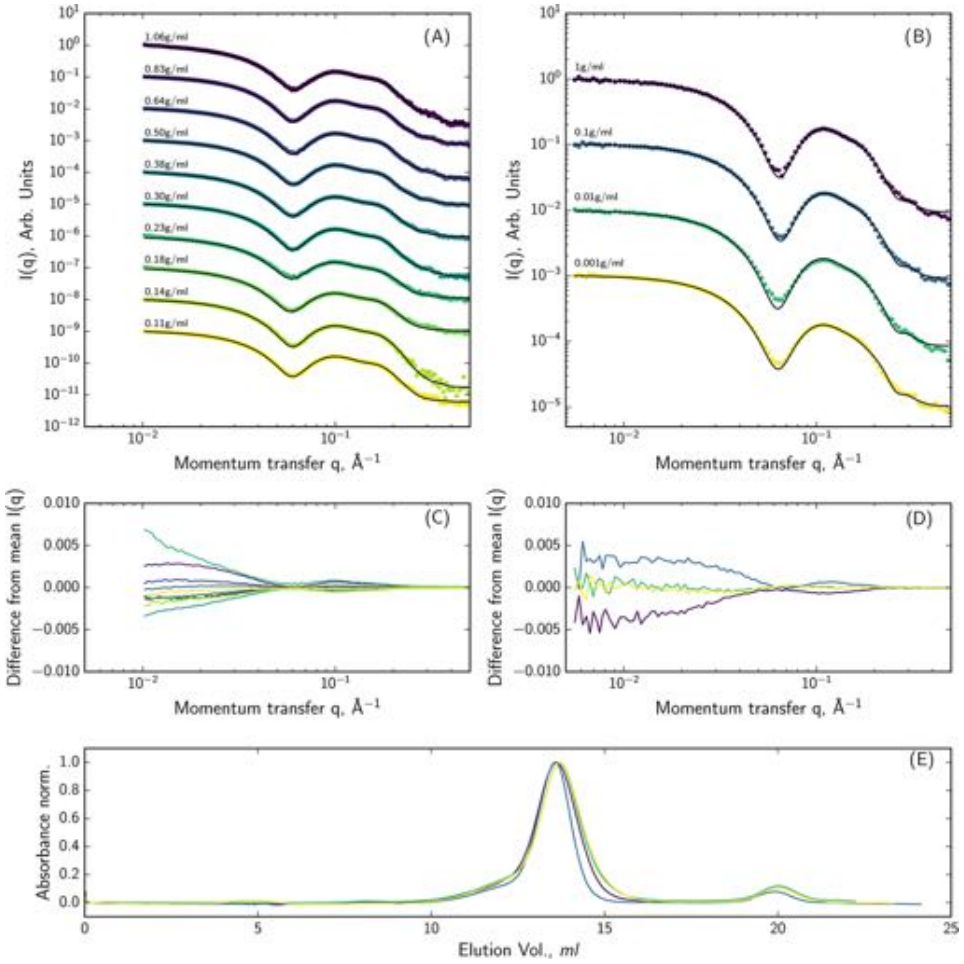


Figure 6. (A) SAXS data (points) and model fits (continuous lines) of nanodiscs prepared using cholate and an increasing amount of detergent-absorbing beads (indicated in figure) starting from the theoretical minimum amount needed to absorb all cholate, to simulate an increasing detergent removal speed. (B) Same as (A), but the amount of detergent-absorbing beads varied over 3 orders of magnitude on an exponential scale and starting from close to zero. (C, D) Difference plots for SAXS curves in (A) and (B), respectively. (E) SEC profiles for sample preparations leading to the SAXS curves in (B).

Overall, we conclude from the analysis in case B that the nanodiscs can be underloaded by not using a sufficiently high POPC/MSP1D1 ratio in the reconstitution mixture. But when attempting to overload the nanodiscs, the excess lipid is not incorporated in the disks but instead simply washed out in the subsequent SEC purification process.

Case C: Nanodiscs Produced at Different Detergent Removal Speeds. Inspired by previous work on detergent-mediated liposome reconstitution (see refs 8, 11 and the references therein), which reports a clearly visible dependence between the detergent removal speed and the average size and homogeneity of the formed vesicles, we hypothesized that such a dependence could also be present in the nanodisc reconstitution. The detergent may, in principle, be removed during either the step, where the detergent-absorbing polymer beads are added, or in the SEC purification step, where the associated sample dilution below the CMC of the detergent also serves to remove the detergent from the nanodiscs, provided that the CMC is sufficiently high.

As documented by Rigaud et al.,[11] the detergent removal speed is directly controllable in the first of the two steps through the amount of detergent-adsorbing polymer beads added to the solution as well as by the temperature where the detergent removal takes place. On the basis of estimates of the cholate absorption capacity of the applied amberlite XAD-2,[11] an initial series was prepared, where 10 different amounts of detergent-removing beads, ranging from 0.11 to 1.06 g/mL, were added to 10 vials of the nanodisc reconstitution mixtures. The samples were then left to equilibrate, after which the usual SEC purification procedure was performed. The obtained SEC profiles indicated no dependence of the amount of added detergent-absorbing beads (data not shown). The SAXS data of the finally formed nanodiscs, plotted in Figure 6A, are likewise almost indistinguishable, and subsequent analysis by inspection of the difference plot (Figure 6C) and by fitting the nanodisc model to the SAXS data did not indicate any systematic dependence between the amount of detergent-removing beads added during the detergent removal step and the final structure of the formed nanodiscs. Instead, it showed that the individual variation between the nanodiscs is indeed both very small and with no systematic dependence of the removal speed.

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

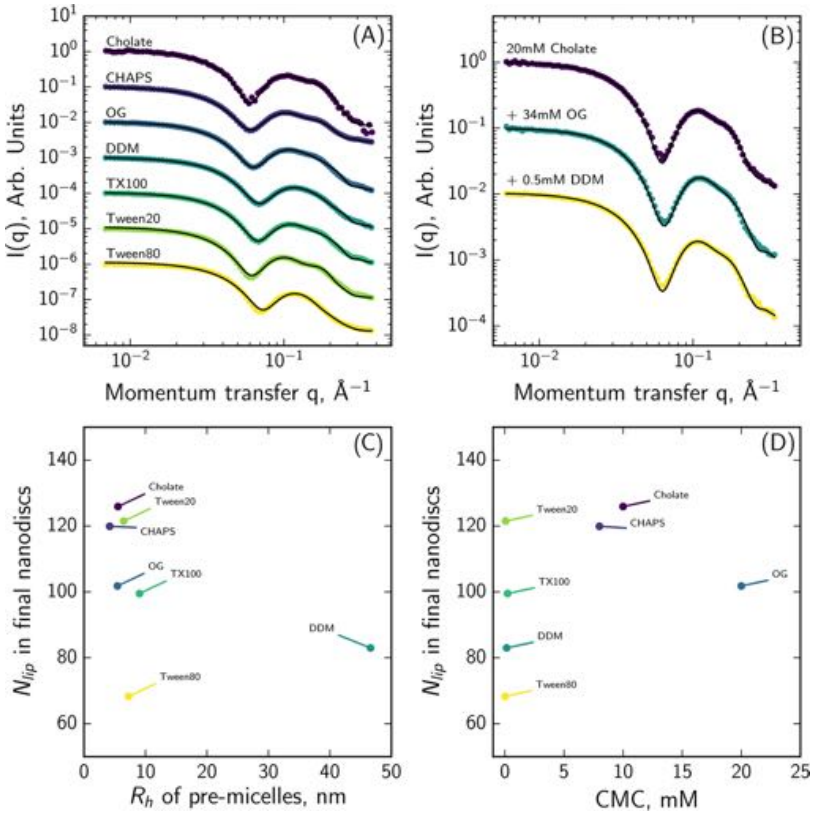


Figure 7. (A) SAXS data from nanodiscs prepared using different detergents and the corresponding model fits. (B) SAXS data from nanodiscs reconstituted from cholate mixed with typical membrane protein reconstitution detergents. (C) Number of lipids per nanodisc, as determined by model fits to the scattering data, plotted against the hydrodynamic radius of the POPC−detergent mixed micelles in the reconstitution mixture. (D) Number of lipids per nanodisc plotted against the CMC of the reconstitution detergent.

To investigate whether an effect would be visible for other amounts of detergent-removing beads, a second series was prepared, but the amount of beads were 1.0, 0.1, 0.01, and 0.001 g/mL (Figure 6B). In this case, the preparation with 1.0 g/mL should have a large excess absorbing capacity, while the three preparations with 0.1, 0.01, and 0.001 g/mL were expected to not have enough capacity for a full cholate removal of the studied samples.[11] Again, there was no visible effect on the initial detergent removal speed (see Figure 6B,D). As the amount of added beads in the 0.01 and 0.001 mg/mL samples were clearly insufficient to remove all of the detergent, the remaining detergent would have to be removed during the subsequent SEC process. However, a subsequent SEC analysis did also not indicate any significant difference between the particles (Figure 6E).

To summarize, this shows that within the studied time frame, the removal speed does not affect the nanodisc self-assembly, and actually whether or not detergent is removed quickly with the detergent-removing beads or through the dilution in a later SEC purification does also not affect the finally formed disks.

Case D: Nanodiscs Produced with Different Reconstitution Detergents. Figure 7A shows SAXS data from nanodiscs prepared with different detergents. This experiment was inspired by the previous observation[9,49] that in detergent-mediated liposome preparation, the choice of detergent is a central determining factor for the size and homogeneity of the formed liposomes. The obtained data show that a similar detergent effect is visible for nanodiscs as the structure of the formed nanodiscs is significantly affected by the choice of reconstitution detergent. This effect is quantified by fitting the nanodisc model to the experimental data, which, among other parameters, yields the dependence between the reconstitution detergent and the number of lipids per formed nanodisc. These data (see Table 2) show that the choice of reconstitution detergent play a major role for the number of POPC-molecules, N_{lipids}, that end up in the formed nanodiscs. Often the user wants to maximize this number. In this respect,

Table 2. Average Number of Lipids Per Nanodisc, N_{lipids}, Determined for Nanodisc Reconstitutions Made with Different Types of Detergent[a]

reconstitution detergent	N_{lipids}	axis ratio
Tween80	68 ± 3	1.46 ± 0.07
DDM	83 ± 1	1.64 ± 0.02
TX100	99 ± 2	1.45 ± 0.03
OG	102 ± 2	1.54 ± 0.02
CHAPS	120 ± 2	1.35 ± 0.02
Tween20	121 ± 1	1.22 ± 0.03
cholate	127 ± 1	1.35 ± 0.02
cholate + 34 mM OG	120 ± 4	1.35 ± 0.06
cholate + 0.5 mM DDM	122 ± 5	1.32 ± 0.08

[a]The listed CMC values for the different detergents were found in Sigma-Aldrich's detergent handbook.[32]

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

Tween80 and DDM clearly performed the least, Triton X-100 and octyl glucoside both showed medium performance and produced nanodiscs with $N_{lipids} \approx 100$, while CHAPS, Tween20, and cholate performed almost equally well with N_{lipids} around 120. These data clearly show that despite that the reconstitution detergent is removed from the final formed disks by means of the detergent-removing beads, as clearly demonstrated in previous works,[50,51] it still plays a central role for the stoichiometry of the formed disks and can not simply be interchanged with another detergent. We note that the cholate reconstitution for case D gave a slightly larger N_{lipids} of 127 ± 1 than the 120 reported under case A and, indeed, we observe a small batch-dependent variation between the different experiments. The effect is, however, minor compared to the effect of detergent choice.

In a typical membrane protein reconstitution into nanodiscs,[12] phospholipid/cholate mixed micelles are mixed with membrane protein stored and stabilized in DDM, OG, or other detergents. Mixed aggregates of all of the different constituents are then formed and the cholate and DDM (or OG) are removed by the addition of the detergent-absorbing beads. Figure 7B shows data from an investigation of how a reconstitution mixture of 20 mM cholate and 0.5 mM DDM, or 20 mM cholate and 34 mM OG would affect the formed disks compared to reconstitutions with only 20 mM cholate. Table 2 lists the obtained N_{lipids} numbers. It is seen that in these detergent mixtures, which are representative for what would be used in a typical membrane protein reconstitution, the nanodisc self-assembly is apparently dominated by the effect of the cholate such that comparable structures of the formed nanodiscs, including comparable N_{lipids} values, are obtained in the three cases. This shows that the effect of the added storage detergent on the cholate-based reconstitution is negligible.

Two final analyses were conducted to investigate whether a simple explanation for why certain reconstitution detergents yielded nanodiscs with certain N_{lipids} values could be identified.

In the first analysis, it was hypothesized that there would be a correlation between the size of the mixed POPC−detergent premicelles and the finally formed nanodiscs. Mixed micelles of POPC and the different detergents were prepared and their size, represented in terms of their hydrodynamic radius (R_h) were measured by DLS. Figure 7C plots the number of POPC in the formed nanodiscs against the obtained R_h. All premicelles except those of POPC−DDM were rather compact with R_h values in the range of 4−8 nm, suggesting small spherical or ellipsoidal micelles. The POPC−DDM sample was more viscous than the other samples and had an R_h of around 45 nm, consistent with a previously described underlying wormlike micelle structure.[28] It is clearly seen from Figure 7C that there is no simple correlation between the two sets of parameters.

In the second analysis, it was hypothesized that there would be a correlation between the CMC of the reconstitution detergent and the finally formed nanodiscs. This was based on the rationale that a lower CMC would give rise to a slower detergent depletion, which could then affect the stoichiometry of the formed particles. Figure 7D plots the N_{lipids} values as a function of the CMC values of the applied detergents. This plot shows that there is also no simple correlation between the N_{lipids} values and the CMC, which is in line with the observation that detergent removal speed is not important for the structure of the final nanodisc, as shown under case C and in Figure 6A,B.

DISCUSSION

In brief summary, we find the following for the self-assembly process:

1. The study in case A shows that the peak fractions normally collected when purifying nanodiscs using SEC contain a small but significant distribution of sizes. The analysis of the SAXS data provides a more detailed view and shows that the POPC/MSP1D1 ratio of the formed nanodiscs decreases as a function of elution volume of the SEC peak, giving rise to a larger axis ratio of the nanodiscs in the tailing fractions. The most loaded and hence less elliptical nanodiscs are observed in the leading fractions of the SEC peak. SLS and DLS data combined with SEC analysis show that the samples are very stable over a time span of up to 140 h at room temperature.
2. The study in case B shows that it is possible to underload the nanodiscs during the self-assembly, producing more elliptic and "lipid-poor" nanodiscs. However, within a reasonable range of POPC/MSP1D1 ratios around the normally recommended value, it is not possible to overload and produce extra lipid-enriched nanodiscs through high POPC/MSP1D1 ratios. This shows that the MSP defines an upper bound on size and POPC/MSP1D1 ratio of the formed nanodisc. We find that this upper limit lies close to 60:1, which is obtained at a reconstitution ratio of 75:1.
3. The study in case C shows that the detergent removal speed does not affect the structure of the nanodiscs prepared, with cholate as the reconstitution detergent and within the time scale accessible in this experiment. In continuation of this, we do not observe the addition of smaller or larger amounts of detergent-removing beads to have any effect on the finally formed disks, when a subsequent SEC purification is performed to remove excess cholate.
4. Given that the remaining reconstitution parameters are well chosen and that a sufficiently high POPC/MSP1D1 ratio is available, the choice of reconstitution detergent is the central factor in defining the shape and stoichiometry of the formed nanodiscs.

The observation 1 of a small but significant distribution of the stoichiometries of the formed nanodiscs is in line with previous observations.[46−48] In particular, circular nanodiscs were not observed for any stoichiometries. As previously suggested by some authors,[43] this may be explained by assuming that the two MSPs define a constant upper bound to the circumference of the disks, making a circular disk unlikely as there is only one way that the disk can be maximally loaded but many more that it can be less than fully loaded. Hence, given that there will be a distribution of POPC/MSP1D1 stoichiometries, the average disk will have to be less than maximally loaded. The assumption of the constant circumference is further supported by the recent NMR structure of DMPC-based nanodiscs based on a truncated form of apoA1,[21] which is very similar to MSP1D1. These data show that the two amphiphatic MSPs are bound together in a zipper-like manner, thus locking them into a preferred circumference.

Observation 2 shows that the free energy of the nanodiscs as a function of POPC/MSP1D1 ratio of the formed disks must

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

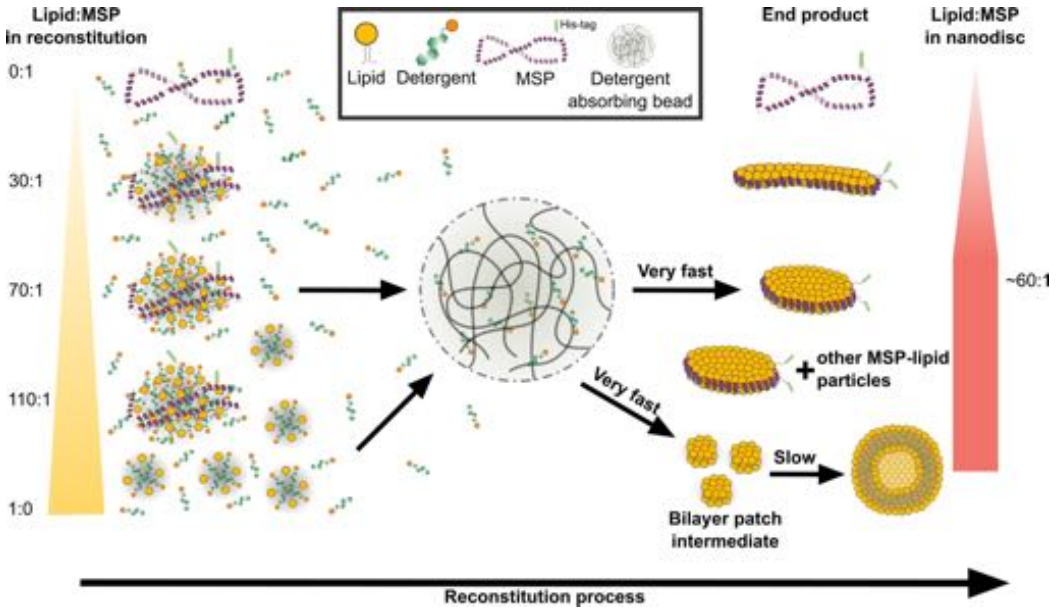


Figure 8. Graphical representation of the nanodisc self-assembly process compared to the commonly accepted model for vesicle formation (figure inspired by Egelhaaf et al.[54]). In the first step (to the left), lipids, detergent, and MSP coexist as preaggregates of lipid−detergent mixed micelles (bottom left), lipid−detergent−MSP mixed micelles (middle left), or as solubilized MSP (top left). In the second step, detergent-absorbing beads are added and detergent is removed from the solution without significant fusion or other exchange of lipids between the preaggregates. This results in a reorganization of the preaggregates into disks, either with MSP (middle right) or as bilayer patches without MSP (bottom right). At this stage, the MSP-containing system has already formed the final nanodiscs, while in the case of the MSP-free vesicle systems (bottom), an additional slow fusion and closure of the disks lead to the further formation of vesicles (bottom right). Very low lipid/MSP ratios result in nanodiscs with low and ultimately zero lipid/MSP ratios (top right).

be a broad flat function. We see that the nanodiscs prepared with a too low POPC-to-MSP ratio contain fewer lipids and become more elliptical. This gives rise to a shift in the elution peak position (Figure 5) rather than splitting up into populations with, respectively, maximally loaded and unloaded nanodiscs as it would have been the case with a $\Delta G_{\text{nanodisc}}$ that favored fully loaded disks and thereby had a strong dependence on the POPC/MSP1D1 ratio. This shows that the MSPs are relatively flexible with respect to how many lipids they can accommodate as long as they are not saturated. In a previous study,[43] we found that upon heating of nanodiscs made from MSP1D1 and, respectively, DLPC and POPC, the temperature-induced increase of the lipid areas per head groups in both cases lead to more circular nanodiscs. This observation is in good agreement with the description of the MSP1D1 belts as relatively flexible. When considering that the MSP is derived from apoA1, this flexibility is perhaps not that surprising: As a component in the reverse cholesterol transport in the human body, apoA1 undergoes significant structural transformations, from lipid-free to discoidal nascent high-density lipoprotein (HDL) particles and finally more spherical cholesterol-loaded mature HDL particles.[52] A large adaptability is crucial for the apoA1 to serve this purpose.

Observation 3 is in immediate contrast to similar experiments on the formation of phospholipid vesicles and it is relevant to compare our results to the similar self-assembly of vesicles from detergent micelles.[8,11,53] In these studies and in later studies by Egelhaaf and co-workers (see, e.g., ref 54), it is proposed that removal of detergents from the mixed lipid−detergent micelles drives the formation of first elongation to wormlike micelles, which is, for example, observed in the lecithin-bile salt system,[54] which then, upon additional detergent removal, reorganize in a very fast process into short-lived discoidal lipid particles. These again fuse with other disks to minimize the hydrophobic rim of the particles.[8,50,54] This disk-fusion process is in competition with a process where the individual lipid bilayers undulate, which may also lead them to form closed vesicle structures. The first fusion process is limited by the diffusion speed and concentration of the larger and slower prevesicle disks, while the second process is limited by the size of the prevesicle disks and the general rigidity of their lipid bilayer. The combination of the two processes provides an explanation for the observed dependence between the speed of the reconstitution process and the observed average size and polydispersity of the formed vesicles, where a fast detergent removal generally leads to smaller and more narrowly dispersed vesicles, whereas a slow detergent removal leads to, on average, larger and more polydisperse vesicles. In the case of the nanodisc self-assembly, we surprisingly observed no such correlation between the detergent removal speed and the homogeneity of the formed nanodiscs. This suggests that the slower disk-fusion process, which is highly relevant to the vesicle self-assembly, plays a minor or negligible role for the nanodisc self-assembly. Instead, the choice of detergent plays a major role. This suggests that the nanodisc self-assembly takes place via a more or less direct transition from the preaggregates, constituted by the MSP1D1, the POPC lipid, and the reconstitution detergent to the final formed disks.

We were surprised by observation 4, both that the effect of the detergent choice was so significant and by the lack of a clearly visible correlation to the detergent CMC or the size of the premicelles in terms of their hydrodynamic radius R_h. Some of the explanation to this may, however, lie in the fact

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

that R_h is a relatively superficial parameter not reporting on the underlying mixed micelle structure, which for some detergents could be, for instance, wormlike and which does also not reveal more subtle effects like local phase separations of the POPC and detergent within the formed micelles, as has been described for some lipid−detergent mixtures capable of forming, respectively, bicelles and wormlike micelles in a temperature-dependent fashion.[55] Such underlying structures would most probably affect the molecular self-assembly process ultimately driving to the optimally loaded nanodiscs. Also, it should be noted that nanodisc reconstitution has so far been primarily optimized with respect to cholate; hence, it can not be ruled out that it could be better optimized for other detergents as well.

Figure 8 summarizes our proposed description of the nanodisc self-assembly process and compares it to the commonly accepted explanation for the vesicle formation process. The figure builds directly on a similar figure proposed by Egelhaaf et al.[54] to describe vesicle formation. Three different situations are illustrated: (1) with only lipid and reconstitution detergent (bottom of the diagram), which would just lead to vesicles through the pathway illustrated in the original figure by Egelhaaf et al. and discussed in further detail above; (2) with MSPs and lipids in an optimal lipid/MSP reconstitution ratio, which for POPC/MSP1D1 would be close to 75:1 (middle of the diagram); and (3) with a very low lipid/MSP ratio, which yields very lipid-poor nanodiscs. The extreme case where no lipid is added and the lipid/MSP ratio is zero (top of the figure) was not studied in the present work, but the addition and subsequent detergent removal is not expected to significantly alter the MSP1D1 solution structure, which we would expect to resemble the model that has been recently proposed for ApoA1.[56]

In all cases, starting from an equilibrium mixture of lipid−detergent and/or lipid−detergent−MSP mixed micelles along with free detergent monomers, the addition of the detergent-absorbing beads and the associated detergent removal, or alternatively the dilution in the SEC purification process, will drive the self-assembly into bilayer-based structures (with the exception of the very lipid-poor systems). To minimize energy, exposed lipid tails must be matched either in a bilayer vesicle structure or by the hydrophobic side of two stacked MSPs in a nanodiscs (both illustrated in Figure 8) and probably also by other and less well-described structures (not indicated in the figure). The difference dependence between size, homogeneity, and detergent removal speed observed for the nanodiscs and the vesicles combined with the observation that the choice of reconstitution detergent determines the size of the finally formed particles suggests that preaggregates with more or less final POPC/MSP1D1 ratios form already from the beginning of the process, i.e., corresponding to the left panel of Figure 8 such that nanodiscs of these stoichiometries are formed already in the first "very fast" (see Figure 8) part of the self-assembly process. In this process, there is basically no time for exchange of the lipids from the preaggregates via fusion or other mechanisms, as these processes would be too limited by the relatively slow diffusion of the preaggregates. The assumption of these ready-packed POPC-MSP1D1-cholate preaggregates is in good accordance with a previous combined SAXS and MD simulation study on cholate-mediated disassembly of nanodiscs.[57]

The subsequent slow fusion between lipid disks, which is relevant to the vesicle formation,[54] is hence not relevant for the nanodisc formation. Instead, as opposed to vesicle formation, the presence of the MSPs halts the fusion of the lipid−detergent particles once they are associated to the MSP. As the MSPs are flexible, the resulting lipid aggregation number in the formed nanodiscs is determined by the availability of lipid, but an upper limit is defined by the size of the MSP. All in all, we propose a simple mechanism where the lipids and the MSP are to a large extent prepacked already in the preaggregates, where the lipid-to-MSP ratio is on the one hand determined by the overall stoichiometry, but on the other hand also determined by the packing properties of the applied reconstitution detergent.

CONCLUSIONS

We have presented a comprehensive study of the factors governing the self-assembly of phospholipid nanodiscs and observed that the self-assembly of nanodiscs, like preparation of phospholipid liposomes, is highly dependent on the preparation procedure. However, as opposed to liposomes, the nanodisc self-assembly process is not affected by the detergent removal speed, but is highly affected by the choice of detergent. On the basis of this, a description of the nanodisc self-assembly process is proposed, where the shape and POPC/MSP1D1 stoichiometry of the formed disks is basically defined by the stoichiometries of the preaggregates. An important finding is that the detergents often used for solubilizing and stabilizing membrane proteins are not the same detergents that result in an optimal nanodisc preparation. This information is crucial when optimizing incorporation of membrane proteins in phospholipid nanodiscs. Of practical importance, we however also observe that when the storage detergents, DDM or OG, are used in combination with the reconstitution detergent cholate, it is the behavior of the cholate that dominates the self-assembly process and the formed disks resemble those prepared with cholate alone.

AUTHOR INFORMATION

Corresponding Author

*E-mail: arleth@nbi.ku.dk.

ORCID

Lise Arleth: 0000-0001-6694-0626

Present Addresses

‡Novo Nordisk A/S, Novo Nordisk Park 1, 2760 Måløv, Denmark (R.H.-N.).

†Copenhagen Business School, Porcelænshaven 18B, 2000 Frederiksberg, Denmark (N.S.-G.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge SAXS beamtime allocated at respectively the European Synchrotron Radiation Facility (ESRF), Max-Lab Synchrotron (MAX-2) in Sweden and PETRAIII at DESY, Germany. In particular, the authors thank Dr. A. Round at the ID14-3 BioSAXS beamline and Martha Brennich at the BM29 BioSAXS beamline, both at ESRF, Thomas Plivelic at the I911 SAXS beamline at Max-Lab and Al Kikhney at the P12 SAXS at PETRAIII at DESY. Lasse Sander Dreyer's help with measuring some of the SAXS data is also acknowledged. Finally the authors acknowledge co-funding from the Lundbeck Foundation Brainstruc project, the Novo Nordisk Foundation Synergy project, the KU2016 BioSynergy

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

project and the Danish Government-funded UNIK Synthetic Biology program.

■ ADDITIONAL NOTES

[a]According to the database of membrane protein structures maintained by Stephen White as of April 2018 and counting only entries solved by diffraction methods: http://blanco.biomol.uci.edu/mpstruc/.

[b]According to a search in the Protein Data Bank April 2018. Counting only protein entries solved by diffraction methods and removing similar sequences at 100% identity: http://www.rcsb.org/pdb/search/advSearch.do.

[c]Personal communication with Prof. Sligar, University of Illinois.

■ REFERENCES

(1) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2002**, *2*, 853−856.

(2) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. Assembly of single bacteriorhodopsin trimers in bilayer nanodiscs. *Arch. Biochem. Biophys.* **2006**, *450*, 215−222.

(3) Denisov, I. G.; Sligar, S. G. Nanodiscs in Membrane Biochemistry and Biophysics. *Chem. Rev.* **2017**, *117*, 4669.

(4) Stark, H.; Chari, A. Sample preparation of biological macromolecular assemblies for the determination of high-resolution structures by cryo-electron microscopy. *Microscopy* **2016**, *65*, 23−34.

(5) Mazer, N. A.; Benedek, G. B.; Carey, M. C. Quasi-Elastic Light-Scattering-Studies of Aqueous Biliary Lipid Systems - Mixed Micelle Formation in Bile-Salt Lecithin Solutions. *Biochemistry* **1980**, *19*, 601−615.

(6) Lasic, D. D. A Molecular-Model for Vesicle Formation. *Biochim. Biophys. Acta* **1982**, *692*, 501−502.

(7) Lichtenberg, D.; Robson, R. J.; Dennis, E. A. Solubilization of Phospholipids by Detergents - Structural and Kinetic Aspects. *Biochim. Biophys. Acta* **1983**, *737*, 285−304.

(8) Lasic, D. D. The Mechanism of Vesicle Formation. *Biochem. J.* **1988**, *256*, 1−11.

(9) Paternostre, M. T.; Roux, M.; Rigaud, J. L. Mechanisms Of Membrane-Protein Insertion Into Liposomes During Reconstitution Procedures Involving the Use of Detergents .1. Solubilization of Large Unilamellar Liposomes (Prepared by Reverse-Phase Evaporation) by Triton X-100, Octyl Glucoside, and Sodium Cholate. *Biochemistry* **1988**, *27*, 2668−2677.

(10) Rigaud, J. L.; Paternostre, M. T.; Bluzat, A. Mechanisms of Membrane-Protein Insertion Into Liposomes During Reconstitution Procedures Involving the Use of Detergents .2. Incorporation of The Light-Driven Proton Pump Bacteriorhodopsin. *Biochemistry* **1988**, *27*, 2677−2688.

(11) Rigaud, J.-L.; Levy, D.; Mosser, G.; Lambert, O. Detergent removal by non-polar polystyrene beads. *Eur. Biophys. J.* **1998**, *27*, 305−319.

(12) Ritchie, T. K.; Grinkova, Y.; Bayburt, T.; Denisov, I.; Zolnerciks, J.; Atkins, W.; Sligar, S. *Methods Enzymol.* **2009**, *464*, 211−231.

(13) Hagn, F.; Etzkorn, M.; Raschle, T.; Wagner, G. Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. *J. Am. Chem. Soc.* **2013**, *135*, 1919−1925.

(14) Holloway, P. W. A simple procedure for removal of Triton X-100 from protein samples. *Anal. Biochem.* **1973**, *53*, 304−308.

(15) Nikolaev, M.; Round, E.; Gushchin, I.; Polovinkin, V.; Balandin, T.; Kuzmichev, P.; Shevchenko, V.; Borshchevskiy, V.; Kuklin, A.; Round, A.; Bernhard, F.; Willbold, D.; Büldt, G.; Gordeliy, V. Integral Membrane Proteins Can Be Crystallized Directly from Nanodiscs. *Cryst. Growth Des.* **2017**, *17*, 945−948.

(16) Frauenfeld, J.; Gumbart, J.; Sluis, E. O. vd; Funes, S.; Gartmann, M.; Beatrix, B.; Mielke, T.; Berninghausen, O.; Becker, T.; Schulten, K.; Beckmann, R. Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. *Nat. Struct. Mol. Biol.* **2011**, *18*, 614−621.

(17) Matthies, D.; et al. Cryo-EM Structures of the Magnesium Channel CorA Reveal Symmetry Break upon Gating. *Cell* **2016**, *164*, 747−756.

(18) Kynde, S. A.; Skar-Gislinge, N.; Pedersen, M. C.; Midtgaard, S. R.; Simonsen, J. B.; Schweins, R.; Mortensen, K.; Arleth, L. Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, 371−383.

(19) Skar-Gislinge, N.; Kynde, S. A. R.; Denisov, I. G.; Ye, X.; Lenov, I.; Sligar, S. G.; Arleth, L. Small-angle scattering determination of the shape and localization of human cytochrome P450 embedded in a phospholipid nanodisc environment. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *71*, 2412−2421.

(20) Nasr, M. L.; Baptista, D.; Strauss, M.; Sun, Z.-Y. J.; Grigoriu, S.; Huser, S.; Plückthun, A.; Hagn, F.; Walz, T.; Hogle, J. M.; Wagner, G. Covalently Circularized Nanodiscs for Studying Membrane Proteins and Viral Entry. *Nat. Methods* **2017**, *14*, 49−52.

(21) Bibow, S.; Polyhach, Y.; Eichmann, C.; Chi, C. N.; Kowal, J.; Albiez, S.; Mcleod, R. A.; Stahlberg, H.; Jeschke, G.; Güntert, P.; Riek, R. Solution structure of discoidal high-density lipoprotein particles with a shortened apolipoprotein A-I. *Nat. Struct. Mol. Biol.* **2017**, *24*, 187−193.

(22) Heerklotz, H. Interactions of surfactants with lipid membranes. *Q. Rev. Biophys.* **2008**, *41*, 205−264.

(23) Lichtenberg, D.; Hasna, A.; Félix, G. The Mechanism of Detergent Solubilization of Lipid Bilayers. *Biophys. J.* **2013**, *105*, 289−299.

(24) Lim, W. H.; Lawrence, M. J. Influence of surfactant and lipid chain length on the solubilisation of phosphatidylcholine vesicles by micelles comprised of polyoxyethylene sorbitan monoesters. *Colloids Surf., A* **2004**, *250*, 449−457.

(25) Garidel, P.; Hildebrand, A.; Neubert, R.; Blume, A. Thermodynamic Characterization of Bile Salt Aggregation as a Function of Temperature and Ionic Strength Using Isothermal Titration Calorimetry. *Langmuir* **2000**, *16*, 5267−5275.

(26) Hildebrand, A.; Beyer, K.; Neubert, R.; Garidel, P.; Blume, A. Temperature dependence of the interaction of cholate and deoxycholate with fluid model membranes and their solubilization into mixed micelles. *Colloids Surf., B* **2003**, *32*, 335−351.

(27) Cladera, J.; Ricaud, J.-l.; Villaverde, J.; Dunach, M. Liposome solubilization and membrane protein reconstitution using Chaps and Chapso. *Eur. J. Biochem.* **1997**, *243*, 798−804.

(28) Lambert, O.; Levy, D.; Ranck, J. L.; Leblanc, G.; Rigaud, J. L. A new "gel-like" phase in dodecyl maltoside-lipid mixtures: Implications in solubilization and reconstitution studies. *Biophys. J.* **1998**, *74*, 918−930.

(29) Urbaneja, M. A.; Alonso, A.; Gonzalez-Mañas, J. M.; Goñi, F. M.; Partearroyo, M. A.; Tribout, M.; Paredes, S. Detergent solubilization of phospholipid vesicle. Effect of electric charge. *Biochem. J.* **1990**, *270*, 305−308.

(30) Schnitzer, E.; Kozlov, M. M.; Lichtenberg, D. The effect of cholesterol on the solubilization of phosphatidylcholine bilayers by the non-ionic surfactant Triton X-100. *Chem. Phys. Lipids* **2005**, *135*, 69−82.

(31) Vinson, P. K.; Talmon, Y.; Walter, A. Vesicle-micelle transition of phosphatidylcholine and octyl glucoside elucidated by cryo-transmission electron microscopy. *Biophys. J.* **1989**, *56*, 669−681.

(32) Bhairi, S. M.; Mohan, C.; Ibryamova, S.; LaFavor, T. *Lab Mater. Supplies*; MilliporeSigma, 2017; pp 1−20.

(33) Rouser, G.; Siakotos, A. N.; Fleischer, S. Quantitative analysis of phospholipids by thin-layer chromatography and phosphorus analysis of spots. *Lipids* **1966**, *1*, 85−86.

(34) Hansen, S. DLSanalysis.org: a web interface for analysis of dynamic light scattering data. *Eur. Biophys. J.* **2018**, *47*, 179−184.

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

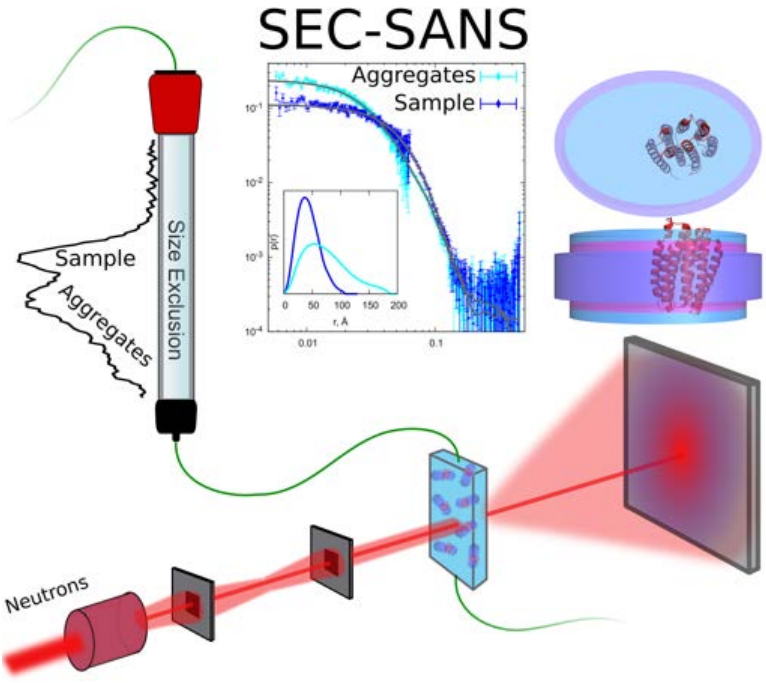
(35) Frisken, B. J. Revisiting the method of cumulants for the analysis of dynamic light-scattering data. *Appl. Opt.* **2001**, *40*, 4087−4091.
(36) Pernot, P.; Theveneau, P.; Giraud, T.; Fernandes, R. N.; Nurizzo, D.; Spruce, D.; Surr, J.; McSweeney, S.; Round, A.; Felisaz, F.; Foedinger, L.; Gobbo, A.; Huet, J.; Villard, C.; Cipriani, F. New beamline dedicated to solution scattering from biological macromolecules at the ESRF. *J. Phys.: Conf. Ser.* **2010**, *247*, No. 012009.
(37) Labrador, A.; Cerenius, Y.; Svensson, C.; Theodor, K.; Plivelic, T. In The Yellow Mini-Hutch for SAXS Experiments at MAX IV Laboratory. 11th International Conference on Synchrotron Radiation Instrumentation (SRI), Lyon, France, Jul 09−13, 2012.
(38) Blanchet, C. E.; Spilotros, A.; Schwemmer, F.; Graewert, M. A.; Kikhney, A.; Jeffries, C. M.; Franke, D.; Mark, D.; Zengerle, R.; Cipriani, F.; Fiedler, S.; Roessle, M.; Svergun, D. I. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J. Appl. Crystallogr.* **2015**, *48*, 431−443.
(39) Mylonas, E.; Svergun, D. I. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J. Appl. Crystallogr.* **2007**, *40*, s245−s249.
(40) Orthaber, D.; Bergmann, A.; Glatter, O. SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J. Appl. Crystallogr.* **2000**, *33*, 218−225.
(41) Glatter, O. A new method for the evaluation of small-angle scattering data. *J. Appl. Crystallogr.* **1977**, *10*, 415−421.
(42) Hansen, S. BayesApp: a web site for indirect transformation of small-angle scattering data. *J. Appl. Crystallogr.* **2012**, *45*, 566−567.
(43) Skar-Gislinge, N.; Simonsen, J. B.; Mortensen, K.; Feidenhans'l, R.; Sligar, S. G.; Lindberg Møller, B.; Bjørnholm, T.; Arleth, L. Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. *J. Am. Chem. Soc.* **2010**, *132*, 13713−13722.
(44) Skar-Gislinge, N.; Arleth, L. Small-angle scattering from phospholipid nanodiscs: derivation and refinement of a molecular constrained analytical model form factor. *Phys. Chem. Chem. Phys.* **2011**, *13*, 3161−3170.
(45) Pedersen, J. S. Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting. *Adv. Colloid Interface Sci.* **1997**, *70*, 171−210.
(46) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A.; Sligar, S. G. Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *J. Am. Chem. Soc.* **2004**, *126*, 3477−3487.
(47) Marty, M. T.; Zhang, H.; Cui, W.; Gross, M. L.; Sligar, S. G. Interpretation and Deconvolution of Nanodisc Native Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 269−277.
(48) Graziano, V.; Miller, L.; Yang, L. Interpretation of solution scattering data from lipid nanodiscs. *J. Appl. Crystallogr.* **2018**, *51*, 157−166.
(49) Rotenberg, M.; Lichtenberg, D. What Determines the Size of Phospholipid-Vesicles Made by Detergent-Removal Techniques. *J. Colloid Interface Sci.* **1991**, *144*, 591−594.
(50) Rigaud, J. L.; Lévy, D.; Mosser, G.; Lambert, O. Detergent removal by non-polar polystyrene beads. *Eur. Biophys. J.* **1998**, *27*, 305−319.
(51) Lévy, D.; Bluzat, A.; Seigneuret, M.; Rigaud, J.-L. A systematic study of liposome and proteoliposome reconstitution involving Bio-Bead-mediated Triton X-100 removal. *Biochim. Biophys. Acta, Biomembr.* **1990**, *1025*, 179−190.
(52) Wang, M.; Briggs, M. R. HDL: the metabolism, function, and therapeutic importance. *Chem. Rev.* **2004**, *104*, 119−137.
(53) Almog, S.; Kushnir, T.; Nir, S.; Lichtenberg, D. Kinetic And Structural Aspects Of Reconstitution of Phosphatidylcholine Vesicles by Dilution of Phosphatidylcholine Sodium Cholate Mixed Micelles. *Biochemistry* **1986**, *25*, 2597−2605.
(54) Egelhaaf, S.; Schurtenberger, P. Micelle-to-Vesicle Transition: A Time-Resolved Structural Study. *Phys. Rev. Lett.* **1999**, *82*, 2804−2807.
(55) Singh, D. Small Angle Scatting Studies of Self Assembly in Lipid Mixtures. Ph.D. Thesis, The Johns Hopkins University, 2009; Available from Proquest http://www.proquest.com/products-services/dissertations.
(56) Melchior, J. T.; et al. A consensus model of human apolipoprotein A-I in its monomeric and lipid-free state. *Nat. Struct. Mol. Biol.* **2017**, *24*, 1093−1099.
(57) Shih, A. Y.; Freddolino, P. L.; Sligar, S. G.; Schulten, K. Disassembly of nanodiscs with cholate. *Nano Lett.* **2007**, *7*, 1692−1696.

DOI: 10.1021/acs.langmuir.8b01503
Langmuir 2018, 34, 12569−12582

5

PAPER II

INTRODUCING SEC-SANS FOR STUDIES OF COMPLEX SELF-ORGANISED BIOLOGICAL SYSTEMS



ARTICLE IN PRESS – Acta Cryst. D

ISSN 2059-7983

Introducing SEC–SANS for studies of complex self-organized biological systems

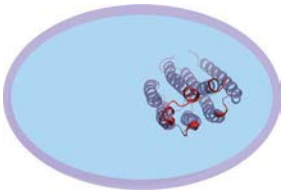
Proof instructions

Proof corrections should be returned by **9 November 2018**. After this period, the Editors reserve the right to publish your article with only the Managing Editor's corrections.

Please
(1) Read these proofs and assess whether any corrections are necessary.
(2) Check that any technical editing queries highlighted in **bold underlined** text have been answered.
(3) Send corrections by e-mail to **lj@iucr.org**. Please describe corrections using plain text, where possible, giving the line numbers indicated in the proof. Please do not make corrections to the pdf file electronically and please do not return the pdf file. If no corrections are required please let us know.

To arrange payment for **open access**, please visit **http://shop.iucr.org/iucrshop/viewitem/openaccess/?code=fs5159**. To purchase printed offprints, please complete the attached order form and return it by e-mail.

Please check the following details for your article



Thumbnail image for contents page

Synopsis: Small-angle neutron scattering (SANS) is coupled with online size-exclusion chromatography (SEC). The obtained SEC–SANS was combined with SEC–SAXS and utilized to investigate solution structures of phospholipid nanodiscs with and without incorporated membrane proteins.

Abbreviated author list: Johansen, N.T.; Pedersen, M.C.; Porcar, L.; Martel, A.; Arleth, L.

Keywords: size-exclusion chromatography; small-angle neutron scattering; small-angle X-ray scattering; membrane proteins; phospholipid nanodiscs

How to cite your article in press

Your article has not yet been assigned page numbers, but may be cited using the doi:

Johansen, N.T., Pedersen, M.C., Porcar, L., Martel, A. & Arleth, L. (2018). *Acta Cryst.* D**74**, https://doi.org/10.1107/S2059798318007180.

You will be sent the full citation when your article is published and also given instructions on how to download an electronic reprint of your article.

STRUCTURAL BIOLOGY

ISSN 2059-7983

Received 31 January 2018
Accepted 14 May 2018

Edited by V. T. Forsyth, Institut Laue-Langevin, France, and Keele University, UK

Keywords: size-exclusion chromatography; small-angle neutron scattering; small-angle X-ray scattering; membrane proteins; phospholipid nanodiscs.

Introducing SEC–SANS for studies of complex self-organized biological systems

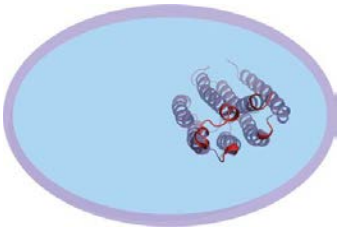
Nicolai Tidemand Johansen,[a] Martin Cramer Pedersen,[a] Lionel Porcar,[b] Anne Martel[b]* and Lise Arleth[a]*

[a]Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark, and [b]Institut Laue–Langevin, Grenoble, France. *Correspondence e-mail: martela@ill.fr, arleth@nbi.ku.dk

Small-angle neutron scattering (SANS) is maturing as a method for studying complex biological structures. Owing to the intrinsic ability of the technique to discern between ^{1}H- and ^{2}H-labelled particles, it is especially useful for contrast-variation studies of biological systems containing multiple components. SANS is complementary to small-angle X-ray scattering (SAXS), in which similar contrast variation is not easily performed but in which high signal-to-noise data are more easily obtained. Obtaining small-angle scattering (SAS) data on monodisperse complex biological structures is often challenging owing to sample degradation and/or aggregation. This problem is enhanced in the D_2O-based buffers that are typically used in SANS. In SAXS, such problems are solved using an online size-exclusion chromatography (SEC) setup. In the present work, the feasibility of SEC–SANS was investigated using a series of complex and difficult samples of membrane proteins embedded in nanodisc particles that consist of both phospholipid and protein components. It is demonstrated that SEC–SANS provides data of sufficient signal-to-noise ratio for these systems, while at the same time circumventing aggregation. By combining SEC–SANS and SEC–SAXS data, an optimized basis for refining structural models of the investigated structures is obtained.

1. Introduction

In structural biology, X-ray crystallography is still the dominating source of information. However, many proteins, for example membrane proteins, are notoriously difficult to crystallize, and intrinsically disordered proteins or proteins with large intrinsically disordered regions may not be crystallizable or accessible by high-resolution cryo-electron microscopy (cryo-EM). Structural information on these types of systems is sparse, although the potential impact from new knowledge is high: approximately one quarter of the genome in eukaryotes encodes membrane proteins (Fagerberg *et al.*, 2010), which are the target of an estimated 60% of all drugs (Overington *et al.*, 2006). Similarly, approximately one third of the genome in eukaryotes codes for intrinsically disordered proteins (Oldfield *et al.*, 2005), which mediate important regulatory interactions (Oldfield & Dunker, 2014). The term 'integrative structural biology' (Ward *et al.*, 2013) was recently coined to describe an approach in which a high-resolution structural model is refined against information from several complementary techniques, both computational and experimental, with the aim of obtaining a more complete description of the structure and dynamics of a system. For this purpose, high-resolution techniques such as crystallography, nuclear



magnetic resonance spectroscopy (NMR) and cryo-EM are optimally combined with more easily accessible information from more low-resolution techniques such as negative-stain EM and small-angle scattering (SAS), as well as various computational approaches for describing the relevant structure or ensemble of structures.

With no inherent size limitation, as is the case for NMR spectroscopy and cryo-EM, SAS is a widely applied method for obtaining low-resolution structural information on biological systems in solution (Svergun & Koch, 2003). Small-angle X-ray scattering (SAXS) has gained popularity in biology over the past few decades, facilitated by excellent synchrotron sources and the development of easily accessible software (Schneidman-Duhovny *et al.*, 2010; Förster *et al.*, 2010; Graewert & Svergun, 2013; Skou *et al.*, 2014; Hopkins *et al.*, 2017; *SasView*, http://www.sasview.org/). As part of this development, SEC–SAXS has been introduced and is now a standard method to circumvent problems of aggregating samples (Mathew *et al.*, 2004; David & Pérez, 2009; Pernot *et al.*, 2013; Graewert *et al.*, 2015; Jeffries *et al.*, 2016; Ryan *et al.*, 2018). SEC–SAXS combines size-exclusion chromatography (SEC) and SAXS in such a way that the SEC purification is coupled directly to the capillary in the X-ray beam, such that sample eluting from the column is illuminated immediately.

During this development, small-angle neutron scattering (SANS) on biological systems has remained a more exclusive technique, mainly owing to the more limited access to neutrons. However, SANS is highly complementary to SAXS owing to the possibility for contrast variation through $^{1}H/^{2}H$ substitution. This has facilitated structural studies of multi-component complexes containing one or more components of protein, lipid and nucleic acids, such as DNA–protein complexes, membrane proteins in detergent micelles and quarternary protein complexes (Whitten *et al.*, 2007; Niemann *et al.*, 2008; Appolaire *et al.*, 2014; Gabel, 2015; Zaccai *et al.*, 2016; Midtgaard *et al.*, 2018). An additional advantage is that radiation damage, which is often an issue in synchrotron SAXS, is not a problem in SANS. Despite these advantages, SANS is more challenging than SAXS. The required sample concentration and volume are higher, the exposure times are longer and, to obtain the best signal-to-noise ratio, samples are usually measured in 100% D_2O. Although chemically very similar to D_2O, this slightly different solvent often promotes protein aggregation (Larsson, 1988; Gripon *et al.*, 1997).

A membrane protein in solution is inherently a multi-component system. To remain soluble, the membrane protein must be stabilized by an amphiphilic entity such as a detergent micelle or a lipid bilayer that shields the hydrophobic transmembrane region of the membrane protein. Detergents provide the most simple reconstitution system in terms of sample preparation, but the polydispersity of the resulting micelle around the membrane protein and the existence of excess micelles decrease the theoretical structural resolution that is obtainable with this approach, although several methods for clever sample preparation and data treatment have been developed (Gabel *et al.*, 2014; Oliver *et al.*, 2017; Koutsioubas, 2017; Midtgaard *et al.*, 2018).

Nanodiscs can be utilized to obtain more monodisperse samples containing membrane proteins embedded in native-like environments (Bayburt *et al.*, 2002). A traditional nanodisc consists of two membrane-scaffold proteins (MSPs) encircling a disc-shaped patch of phospholipid bilayer that can serve as a carrier for a membrane protein. Attempts at crystallizing nanodiscs or membrane proteins in nanodiscs have so far been unsuccessful, but nanodiscs have recently been used to facilitate membrane-protein crystallization (Nikolaev *et al.*, 2017). The diameter of the nanodisc is directly correlated to the length of the MSP (Denisov *et al.*, 2004; Hagn *et al.*, 2013), and NMR and spin-labelling techniques have recently been combined to describe the first high-resolution structure of an 8 nm nanodisc based on a short MSP (Bibow *et al.*, 2017).

SAXS and SANS have previously been combined to study bacteriorhodopsin inside a nanodisc (Kynde *et al.*, 2014), where a combined analysis, *i.e.* global fitting, of SAXS and SANS data was desired because of the high complementarity of the obtained information. The protein component of the nanodisc had higher excess scattering-length density in the X-ray contrast, whereas the lipid component had higher excess scattering-length density in the SANS contrast in 100% D_2O. The simultaneous fitting of the model to the two contrasts allowed a more detailed low-resolution structure of the system to be refined and gave information about both the membrane protein and the surrounding lipid membrane structure.

Because of the structural complexity of the nanodisc system, the analysis of SAS data is not straightforward and generally relies on data from structurally homogeneous/monodisperse samples. However, sample preparation of such homogeneous nanodisc samples is challenging, and until now it has been our experience that the success rate of obtaining good SANS data from these systems is relatively low owing to combined issues with membrane-protein reconstitution and the abovementioned increased aggregation in D_2O-based buffers. Despite the higher sample requirements than for SAXS, the feasibility of performing online SEC–SANS with relative short exposure times has recently been demonstrated on a set of water-soluble proteins (Jordan *et al.*, 2016). Overall, the expected outcome for SEC–SANS (and SEC–SAXS) is higher sample homogeneity, which leads to a scattering signal that is not smeared because of structural dispersity. In turn, structural models can be refined to better resolution.

In the present study, we aimed to test and benchmark the new SEC–SANS setup in combination with complex self-assembled systems: nanodiscs without membrane proteins (empty nanodiscs) and membrane-protein-loaded nanodiscs (loaded nanodiscs). Using SANS data from monodisperse samples with good signal-to-noise ratios in an investigated q-range up to $q \simeq 0.15$ Å^{-1}, we show that the SEC–SANS method is applicable to and relevant for studying these types of samples, which are usually difficult to produce in a sufficiently pure and homogenous form in the required amounts and concentrations. The SEC–SANS data are complemented by SEC–SAXS data, and we model empty nanodiscs and the membrane protein proteorhodopsin (PR) inside a nanodisc, respectively, with global fits to the SANS and SAXS data sets.

2. Sample preparation

2.1. Materials

All chemicals were obtained from Sigma–Aldrich. The salts used in the SEC–SAS buffers were obtained in the anhydrous forms. 99.9% D_2O was obtained from Cambridge Isotope Laboratories.

2.2. Protein expression and purification

The plasmids encoding the genes for MSPs were obtained from Professor Steven Sligar (University of Illinois, USA), the plasmid encoding PR was obtained from Professor Richard Neutze (University of Gothenburg, Sweden) and the plasmid encoding *Thermotoga maritima* CorA (TmCorA) was obtained from Professor Mikaela Rapp (University of Stockholm, Sweden). MSPs were expressed and purified essentially as described elsewhere (Ritchie *et al.*, 2009), but under denaturing conditions. TmCorA was expressed and purified using a slightly modified protocol to that reported elsewhere (Palombo *et al.*, 2012), and PR was expressed and purified essentially as described elsewhere (Gourdon *et al.*, 2008). A detailed description of the preparation of all proteins can be found in the supporting information. The MSP used for loaded nanodiscs had the N-terminal His_6 tag cleaved off with *Tobacco etch virus* (TEV) protease before the reconstitution procedure described below. The N-terminal His_6 tag was not cleaved off from the MSPs used for empty nanodiscs.

2.3. Preparation of empty and loaded nanodiscs

Empty nanodisc samples were prepared off-site one week in advance, although without the final SEC step. Briefly, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC; Avanti Polar Lipids) lipids were dissolved to a final concentration of 50 m*M* in gel-filtration buffer (20 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl) containing 100 m*M* sodium cholate. The lipid mixture was mixed with MSP and diluted to 10 m*M* lipid concentration with gel-filtration buffer and incubated briefly at room temperature. Damp Amberlite XAD-2 detergent-absorbing beads (Sigma–Aldrich) were added to approximately 15%(*w*/*v*) and the sample was incubated at 5°C for 4 h under fast agitation. The detergent-absorbing beads were removed and the sample was purified by SEC on a Superdex 200 10/300 GL column (GE). A 1 ml peak fraction was collected, flash-frozen and stored at −80°C.

For loaded nanodiscs, the membrane protein of interest was added to the reconstitution mixture, displacing some volume of gel-filtration buffer, and allowed to equilibrate for 1 h before the addition of detergent-absorbing beads. To avoid excess membrane protein and reduce the amount of oligomeric membrane protein, PR was mixed with MSP1D1 in a 1:16 ratio with a POPC:MSP1D1 ratio of 70:1, whereas TmCorA was mixed with MSP1E3D1 in a 1:20 ratio with a POPC:MSP1E3D1 ratio of 130:1. Empty nanodiscs were separated from the loaded nanodiscs on a 1 ml HisTrap column (GE), utilizing the His_6 tag on the membrane protein. Empty nanodiscs flowing through the column were discarded, and the bound loaded nanodiscs could be eluted from the column with gel-filtration buffer containing 250 m*M* imidazole. The samples were diluted twofold in gel-filtration buffer and concentrated to approximately 1 ml in a 10 kDa molecular-weight cutoff (MWCO) spin filter. Each sample was split into two for SANS and SAXS measurements, flash-frozen and stored at −80°C. Samples were transported to the Institut Laue–Langevin (ILL) or the ESRF on dry ice and were stored at −80°C until use. Before measurement, the samples were thawed and centrifuged for 10 min at 14 000*g*. The final volumes and concentrations of the injected samples for SEC–SANS and SEC–SAXS are shown in Table 1.

Table 1
Sample volumes and associated absorptions measured at 280 nm (A_{280}) for the samples investigated in this study.

Sample	A_{280}	Volume (µl)	
		SANS	SAXS
MSP1D1–POPC nanodiscs	1.30	200	200
MSP1D1–POPC–PR nanodiscs	3.34	250	160
MSP1E3D1–POPC–TmCorA nanodiscs	3.64	160	140

3. Methods and instrumentation

3.1. Dedicated SEC–SANS setup on D22

SANS data were recorded on the D22 small-angle scattering diffractometer at the ILL, Grenoble, France. The online SEC (Fig. 1) was performed using a modular HPLC system (Serlabo), including two BlueShadow pumps (Knauer), an active mixing valve, an automated sample loader (ALIAS), a Foxy R1 fraction collector (Teledyne ISCO) and a SmartLine 2600 diode-array spectrophotometer (Knauer). The spectrophotometer was connected *via* optic fibres either to an optic cell (3 mm path length, used in measurements of empty nanodiscs) placed at the outlet of the chromatography column or directly to the middle of the SANS measurement cell, perpendicularly to the neutron beam (10 mm path length, used in measurements of loaded nanodiscs), enabling the simultaneous recording of chromatograms at four different wavelengths, in this case 220, 280, 535 and 600 nm.

The setup was equipped with a Superdex 200 10/300 GL column (GE) with a void volume of approximately 7.5 ml. The data were obtained during two separate beam times. Empty nanodisc data were obtained in the first session and loaded nanodisc samples in the second session. For the empty nanodiscs a flow rate of 0.3 ml min^{-1} was used, whereas loaded nanodisc samples were pumped at 0.5 ml min^{-1} in most of the column void volume; this was reduced to 0.2 ml min^{-1} during sample elution. Furthermore, all components of the HPLC setup, including the buffers and the column, were placed in a closed cabinet connected to an air-cooling system set to 10°C to control the temperature. For empty nanodiscs, the buffer consisted of 20 m*M* Tris–DCl pH 7.5, 100 m*M* NaCl and for PR-loaded nanodiscs the buffer consisted of 20 m*M* Tris–DCl pH 7.5, 300 m*M* NaCl. For TmCorA-loaded nanodiscs, 100 m*M* $MgCl_2$ was added to the latter buffer.

The experiments were carried out with a nominal neutron wavelength (λ) of 6.0 Å, a wavelength distribution ($\Delta\lambda/\lambda$) of 10% full-width at half-maximum, a rectangular collimation of 40 × 55 mm and a rectangular sample aperture of 7 × 10 mm. Three sample-to-detector distances (11.2, 5.6 and 1.4 m, with collimation lengths of 11.2, 5.6 and 2.8 m, respectively) were used for characterization of the empty nanodiscs, covering a momentum-transfer range q of 0.0044–0.61 Å^{-1}; $q = 4\pi\sin(\theta)/\lambda$, where θ is half of the angle between the incoming and the scattered neutrons. For the measurement of loaded nanodiscs, sample availability limited the number of possible sample-to-detector distances to two: 11.2 and 2.0 m (with collimation lengths of 11.2 and 2.8 m, respectively), giving a q-range of 0.0044–0.46 Å^{-1}. The measured intensities were binned into 30 s frames.

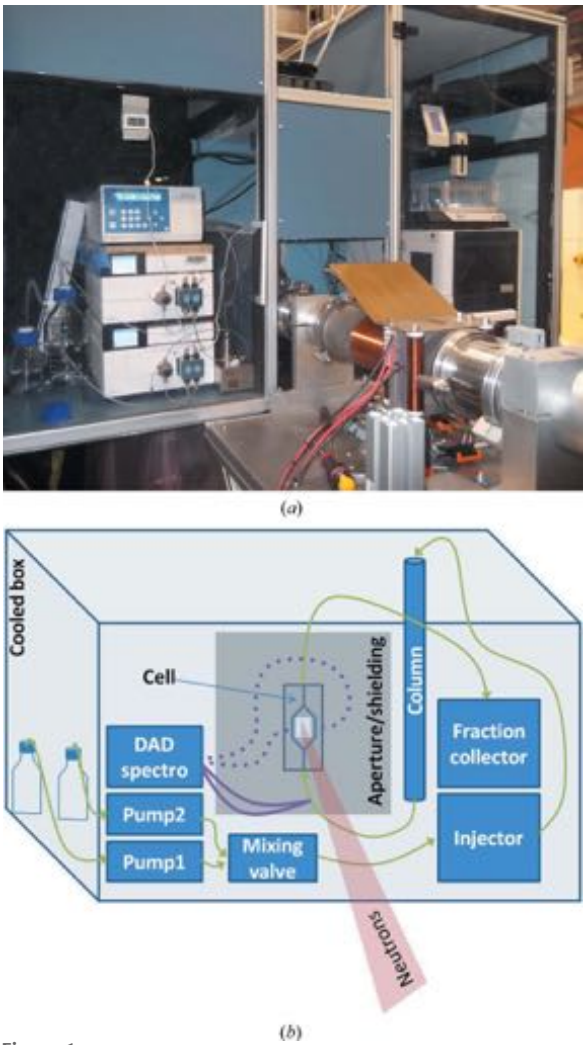


Figure 1
(*a*) Photograph of the setup implemented at D22. (*b*) Schematic describing the SEC–SANS setup, with the different components placed identically to those in the photograph in (*a*). The scheme highlights the two possible options for UV–visible light absorbance measurement: either through a specific cell (optic fibres drawn as continuous purple lines) or directly across the SANS measurement cell (optic fibres drawn as purple dotted lines). The green lines represent the solution flow.

Sample transmission was approximated by the buffer measured at a sample-to-detector distance of 11.2 m. The measured intensity, $I(q)$, was brought to an absolute scale in units of scattering cross-section per unit volume (cm^{-1}) using the direct beam flux measured for each collimation length prior to the experiment. In this way, problems with mismatch between intensities measured at different detector settings were avoided. Data reduction was performed using the *GRASP* software (Dewhurst, 2017).

3.2. SEC–SAXS data acquisition

SEC–SAXS was performed using the BioSAXS instrument on BM29 at ESRF, Grenoble, France (Pernot *et al.*, 2013). This setup includes an HPLC controlled separately from the SAXS measurement, coupled to a UV–Vis array spectrophotometer measuring absorption from 190 to 800 nm. Data were collected with an X-ray wavelength of 0.9919 Å at momentum transfers q ranging from 0.003 to 0.49 Å^{-1}. The capillary was cooled to 10°C, but with the HPLC including the SEC column placed at ambient temperature. Experiments were run on the same column and in H_2O-based solvents containing the same components as in SEC–SANS at flow rates of 0.5 ml min^{-1} for empty nanodiscs and 0.4 ml min^{-1} for loaded nanodisc. Data reduction was carried out using the in-house software at the beamline, and subsequent conversion to absolute units was performed with water as a calibration standard (Orthaber *et al.*, 2000). The 1 s frames recorded were averaged in 10 s bins; this is the typical measurement time for static SAXS samples on BM29.

3.3. SEC–SAS data processing

3.3.1. Selection of sample frames. Scattergrams, *i.e.* scattering intensity as a function of elution volume, were obtained by normalizing the total intensities of the individual frames. Chromatograms of absorption at 280 nm were normalized, and for the setups with the UV cell located before the sample cell the chromatograms were shifted to align with the scattergrams. Where measured, absorptions at wavelengths other than 280 nm were normalized to the absorption at 280 nm. The radius of gyration (R_g) of background-subtracted data was calculated for individual frames with *AutoRg* (Petoukhov *et al.*, 2007). The interval used for extracting the optimal scattering data was chosen to have a converged R_g within error bars in a window covering approximately the right-side half of the desired SEC peak, corresponding to a typical fraction collected for other types of experiments to avoid the overlap of larger particles. The selected frames were averaged to produce a single scattering curve. For SAXS, the averaged data were logarithmically rebinned.

Pair distance [$p(r)$] distributions were calculated by inverse Fourier transformations (IFTs; Glatter, 1977) using the *BayesApp* software available online (Hansen, 2014). For empty nanodiscs, $p(r)$ distributions were calculated for the

individual frames of the selected interval, whereas they were only calculated for averaged data for loaded nanodiscs. The maximum pair distances (D_{max}) and R_g values presented in Figs. 2(*g*), 3(*e*) and 4(*e*) were obtained from this IFT analysis.

Sample concentrations were estimated from the online absorption measurements at 280 nm divided by sample-specific molar extinction coefficients calculated using *ProtParam* on the ExPASy server (Gasteiger *et al.*, 2005). For the empty nanodiscs, however, better estimates of the concentration were obtained from the values of the fractional area of the normalized absorption measured at 280 nm, the volume measured and the total injected amount of sample. This estimate was only possible for the empty nanodiscs, which were already monodisperse before applying them to the column, and thus not for the loaded nanodiscs.

3.3.2. Modelling of empty nanodiscs. The recorded data were analyzed by refining the analytical model for the nanodisc presented in the literature (Skar-Gislinge *et al.*, 2010, 2011) using the data (Fig. 5*a*). The model was simultaneously refined using the SAXS and SANS data by minimizing the combined χ^2 of both data sets. The source code for the models and descriptions thereof, as well as the optimization routines, can be found in the literature (Pedersen *et al.*, 2013).

In this model, the nanodisc is described by a combination of analytical form factors (Pedersen, 1997), each of which are endowed with the scattering properties of the molecules they are meant to represent. Mathematically, the model has the form

$$I(q) = nP(q), \tag{1}$$

where $P(q)$ is the form-factor intensity for a single nanodisc and n is the number density. $P(q)$ describes the scattering properties of an orientationally averaged (over the two angles α and β) nanodisc described by the complex form-factor amplitude $F(\mathbf{q})$,

$$P(q) = \int_0^{\pi} \sin\alpha \, \mathrm{d}\alpha \int_0^{2\pi} F(\mathbf{q}) F^*(\mathbf{q}) \, \mathrm{d}\beta, \tag{2}$$

where the vector $\mathbf{q}$ can be expressed in spherical coordinates (q, α, β), with q being the norm. The asterisk denotes complex conjungation.

The form-factor amplitude of the (oriented) nanodisc is described by a sum of contributions accounting for the different components in the nanodisc,

$$\begin{aligned} F(\mathbf{q}) = {} & \Delta\rho_{\text{protein belt}} V_{\text{protein belt}} A(\mathbf{q})_{\text{protein belt}} \\ & + \Delta\rho_{\text{histidine tag}} V_{\text{histidine tag}} A(\mathbf{q})_{\text{histidine tag}} \\ & + \Delta\rho_{\text{lipid headgroups}} V_{\text{lipid headgroups}} A(\mathbf{q})_{\text{lipid headgroups}} \\ & + \Delta\rho_{\text{lipid alkyl groups}} V_{\text{lipid alkyl groups}} A(\mathbf{q})_{\text{lipid alkyl groups}} \\ & + \Delta\rho_{\text{lipid methyl groups}} V_{\text{lipid methyl groups}} A(\mathbf{q})_{\text{lipid methyl groups}}, \end{aligned} \tag{3}$$

where the $\Delta\rho$ terms are the excess scattering-length densities of the given component, the V terms are the (partial specific) volumes of the components and lastly the $A(\mathbf{q})$ terms are the (normalized) form factors describing the shape of the given component. As an example, the form factor for the lipid methyl groups, $A(\mathbf{q})_{\text{lipid methyl groups}}$, is that of a flat cylinder with elliptical cross-section (see, for example, Pedersen, 1997). The parameters in the model describe these shapes, their volumes and their scattering properties. The dimensions of the aforementioned cylinder depends on the area per lipid headgroup, the number of lipids in the disc, the volume of a POPC molecule and the axis ratio of the disc, while the relevant excess scattering-length density, $\Delta\rho$, may be calculated from the chemical composition and molecular volume of the constituents (see, for example, Skar-Gislinge *et al.*, 2010). In this way, by systematically using molecular constraints, the number of fit parameters to describe the model could be kept at a minimum (see Fig. 4*e* and Section 4).

In the refinement process, the height of the MSP belt was fixed to 25.78 Å to match the recently published high-resolution structure of a nanodisc (Bibow *et al.*, 2017). Similarly, the R_g of the protruding His_6 tags was fixed to 12.70 Å in line with previous studies of disordered proteins (Kohn *et al.*, 2004). It was assumed that 50% of the exchangeable H atoms in the MSP1D1 proteins had exchanged with the solvent (100% D_2O) in the SANS experiments, in line with the literature (Morgan *et al.*, 2011). The effects of instrumental smearing was accounted for in the fits in accordance with the approaches outlined in the literature (Pedersen *et al.*, 1990).

The minimization was performed using the Levenberg–Marquardt (Levenberg, 1944; Marquardt, 1963) algorithm in *WillItFit* (Pedersen *et al.*, 2013), which was also used to compute the estimated refinement errors.

3.3.3. Modelling of PR-loaded nanodiscs. A model analogous to the model of bacteriorhodopsin in a nanodisc presented in the literature (Kynde *et al.*, 2014) was refined from the data obtained from monomeric PR-loaded nanodiscs (Fig. 5*b*). Apart from the analytical description of the nanodisc described above, the model was based on the published solution NMR structure of PR (Reckel *et al.*, 2011; PDB entry 2l6x) in the Protein Data Bank (Berman *et al.*, 2000), acquired from the Orientations of Proteins in Membranes (OPM) database (Lomize *et al.*, 2006). For reference, the OPM database lists the hydrophobic height of the protein as 27.8 ± 1.3 Å.

H atoms were added to the structure using the built-in algorithms in the *PHENIX* software suite (Adams *et al.*, 2010) and solvent exchange was accounted for using the same tool. Based on some of the few available data (Hebling *et al.*, 2010; Mehmood *et al.*, 2012), it was assumed that no solvent exchange had occurred in the transmembrane part of the protein, whereas it was assumed that 100% of the exchangeable H atoms in the flexible intracellular and extracellular loops of the protein had exchanged. In the calculations of scattering intensities, the PR structure was coarse-grained to amino-acid residue level for computational efficiency. As explained in the previous section, the height of the MSP belt was fixed during the refinement process.

The model of PR in a nanodisc is similar to that presented in (1)–(3). We simply add another contribution from the PR to (3) as described previously (Kynde *et al.*, 2014).

4. Results

4.1. Empty nanodiscs: feasibility of the SEC–SANS setup

To test and benchmark the new dedicated SEC–SANS setup at D22 (Fig. 1), empty nanodiscs were first measured. For all three detector settings, the scattergrams in terms of the normalized total scattering intensity are plotted as a function of retention volume along with the normalized chromatogram measured at 280 nm and the calculated R_g values (Fig. 2*a*). The three scattergrams and the chromatogram all feature a well defined single peak, indicating a monodisperse sample, as well as a flat region corresponding to the buffer, which was used for background subtraction. Fig. 2(*b*) shows a complementary SEC–SAXS data set measured on the same sample as used for the SEC–SANS data, although in an H_2O-based buffer. The SAXS data in Supplementary Fig. S1 show that the structure of the empty nanodisc is preserved after exchange from H_2O to D_2O.

The normalized and scaled scattering data of the individual frames for SANS and SAXS, respectively, are shown in Fig. 2(*d*), with the top data set corresponding to the first frame of the selected interval and the rest in the order of elution. As they were very noisy at $q > 0.1$ Å^{-1}, the data were plotted without error bars to make the individual data sets easily separable for the eye. All scattering data exhibit a well defined Guinier region (a flat low-q region on a double-logarithmic scale) down to the minimum q value, indicating no aggregation or presence of larger particles. Guinier plots for the averaged SANS and SAXS data (Fig. 5*a*), respectively, are shown in Supplementary Figs. S2(*a*) and S2(*b*) and confirm this observation.

$p(r)$ distributions for the individual frames are shown in Fig. 2(*f*). Overall, the $p(r)$ distributions for the different frames are rather similar and terminate at approximately the same D_{max} in SANS and SAXS, respectively. This shows that the samples are rather monodisperse within the resolution of the SEC and that the SANS and SAXS samples indeed have the same underlying structure, forming an optimal basis for a subsequent simultaneous analysis. The values for the concentrations calculated from the absorption at 280 nm, as well as the D_{max} and R_g values derived from the $p(r)$ distributions, are summarized in Fig. 2(*g*).

A more careful inspection of the data reveals that while the individual SANS data sets appear to be highly similar, the SAXS data sets have a small but clear systematic variation as a function of retention volume (Fig. 2*d*). Inspection of the $p(r)$ distributions (Fig. 2*f*) and the characteristic sizes (Fig. 2*g*) confirms this apparent change in both SANS and SAXS as seen by the R_g decreasing systematically from 31.2 to 27.8 Å for the SANS data and from 45.4 to 40.1 Å for the SAXS data. Hence, the SEC–SANS and SEC–SAXS analyses provide additional information by revealing a small but significant structural dispersion for the empty nanodiscs. Despite these small variations of the particle shape over the SEC peaks, averaged data sets were used for modelling as described in Section 4.4 in order to improve the statistics.

4.2. PR-loaded nanodiscs: separation of different species

Next, PR-loaded nanodiscs were measured by SEC–SAS. As for the empty nanodiscs, scattergrams and chromatograms (280 and 535 nm) were plotted for both SEC–SANS and SEC–SAXS, yielding two dominating but overlapping peaks eluting after the column void volume and with approximately 3 ml separation between their peaks (Fig. 3*b*). PR has a light red colour owing to the absorption of green light, and the

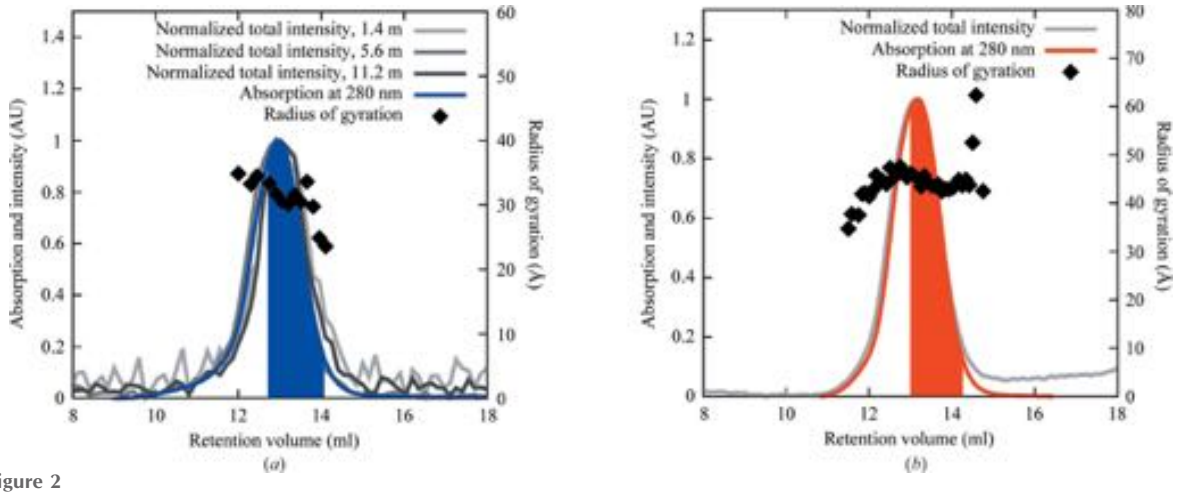
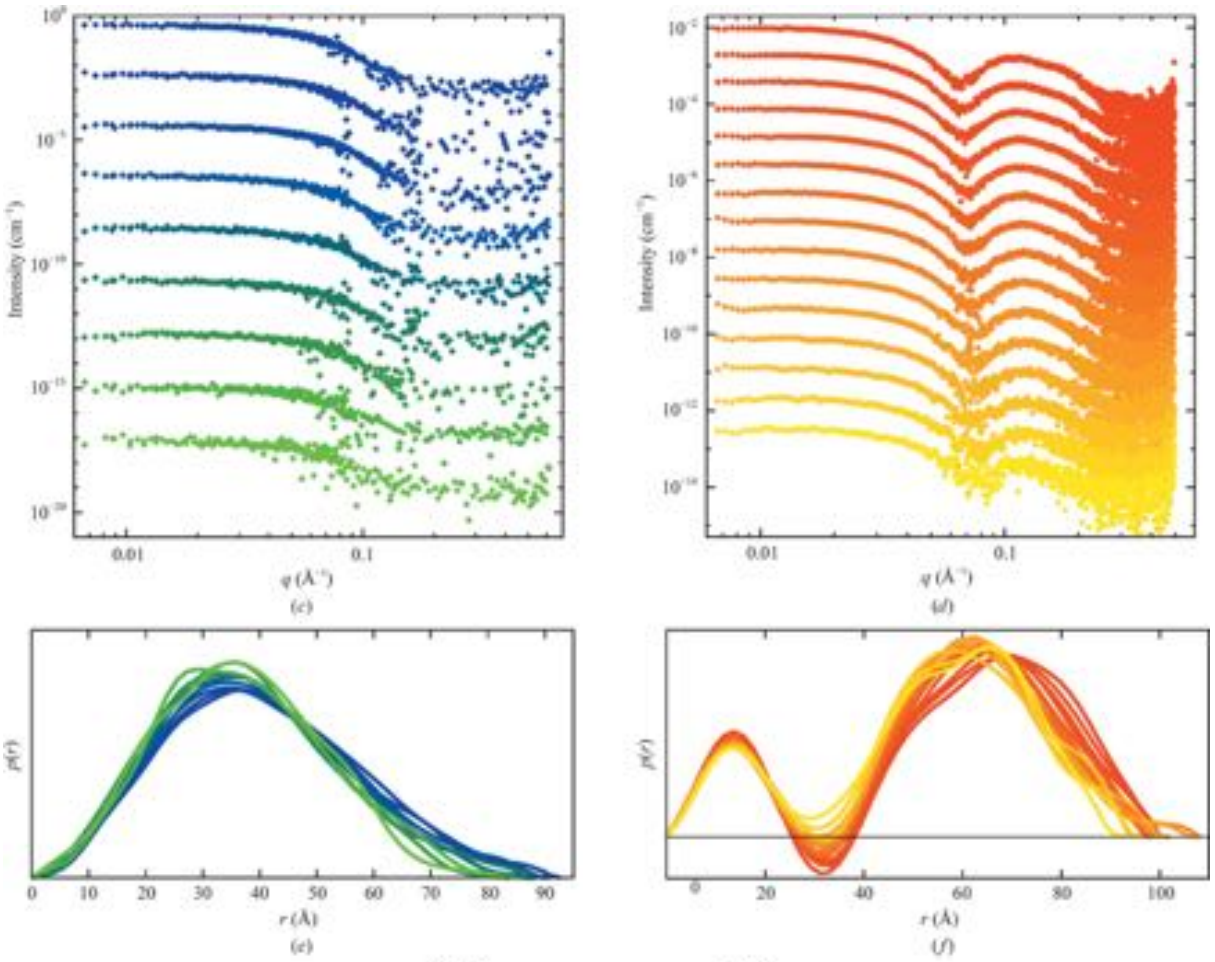


Figure 2
(*a*, *b*) SEC–SAS data for the empty nanodiscs, including scattergrams and the chromatograms measured at 280 nm as well as calculated R_g values. The coloured areas indicate the intervals from which individual frames were taken. (*c*, *d*) Reduced and background-subtracted SAS data for the individual frames taken from the indicated intervals in (*a*) and (*b*) shown without the associated error bars. The topmost data sets are presented on an absolute scale, whereas the others have been rescaled by a factor of 100 (SANS data) and a factor of five (SAXS data) relative to the data set above. (*e*, *f*) $p(r)$ distributions refined from each of the data sets in (*c*) and (*d*), respectively. The averaged data are shown in Fig. 5(*a*). (*g*) Parameters from *BayesApp* refined from the IFT of the data.

chromatogram measured at 535 nm confirmed the presence of PR in both peaks. Clearly, the two peaks observed in the scattergrams and chromatograms were not separated sufficiently to extract a monodisperse sample from the relatively broad first peak (peak 1). The selected frames were picked from the middle of the peak and processed to obtain an idea of



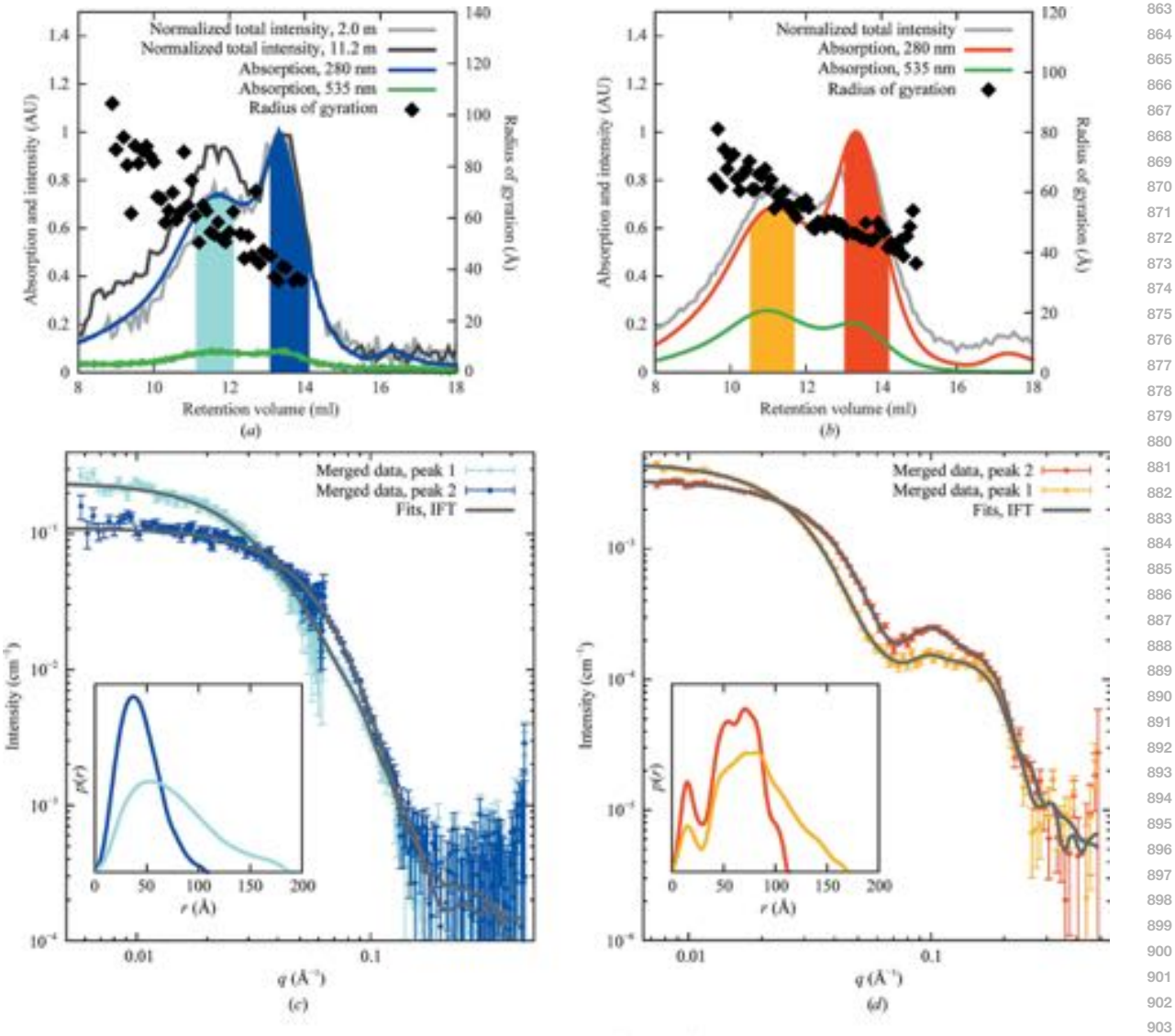
(g)

SANS				SAXS			
Frame	c (μ*M*)	D_{max} (Å)	R_g (Å)	Frame	c (μ*M*)	D_{max} (Å)	R_g (Å)
1	4.2	85.0	31.2	1	4.4	97.4	45.4
2	4.6	90.0	31.0	2	4.3	98.3	44.9
3	4.4	86.2	30.4	3	4.5	97.8	44.3
4	4.2	87.9	30.1	4	4.4	97.4	43.8
5	3.5	85.1	29.2	5	4.0	95.3	43.2
6	2.8	86.2	29.2	6	4.0	96.2	42.8
7	2.0	85.1	27.8	7	3.4	95.3	42.4
8	1.3	85.1	28.5	8	3.2	97.0	42.2
9	0.9	85.1	27.8	9	2.7	105.9	42.3
				10	2.1	97.2	41.8
				11	1.8	103.9	42.0
				12	1.5	98.1	41.7
				13	1.1	92.3	41.3
				14	0.8	93.3	40.6
				15	0.6	88.5	40.1
Merged	3.1	90.1	30.2	Merged	2.85	98.9	43.3

Figure 2 (continued)

the overall size of the particles that were present, but without further interpretation. To minimize interference from the larger particles present in peak 1, the frames from the second peak (peak 2) were chosen from the right side of this peak.

The obtained averaged data sets for both peaks are plotted along with the fits of the IFT analysis in Fig. 3(*d*), with the derived $p(r)$ distributions shown in the insets. The parameters refined from the IFT analysis are summarized in Fig. 3(*e*). Owing to the low sample concentration, the SANS data are slightly noisy in both the low-q and high-q regions, but nevertheless feature well determined Guinier regions along with significant structural differences between the SANS data



Data set	D_{max} (Å)	R_g (Å)
Peak 1, SANS	186.1	59.1
Peak 1, SAXS	168.8	60.5
Peak 2, SANS	105.2	33.3
Peak 2, SAXS	110.2	45.0

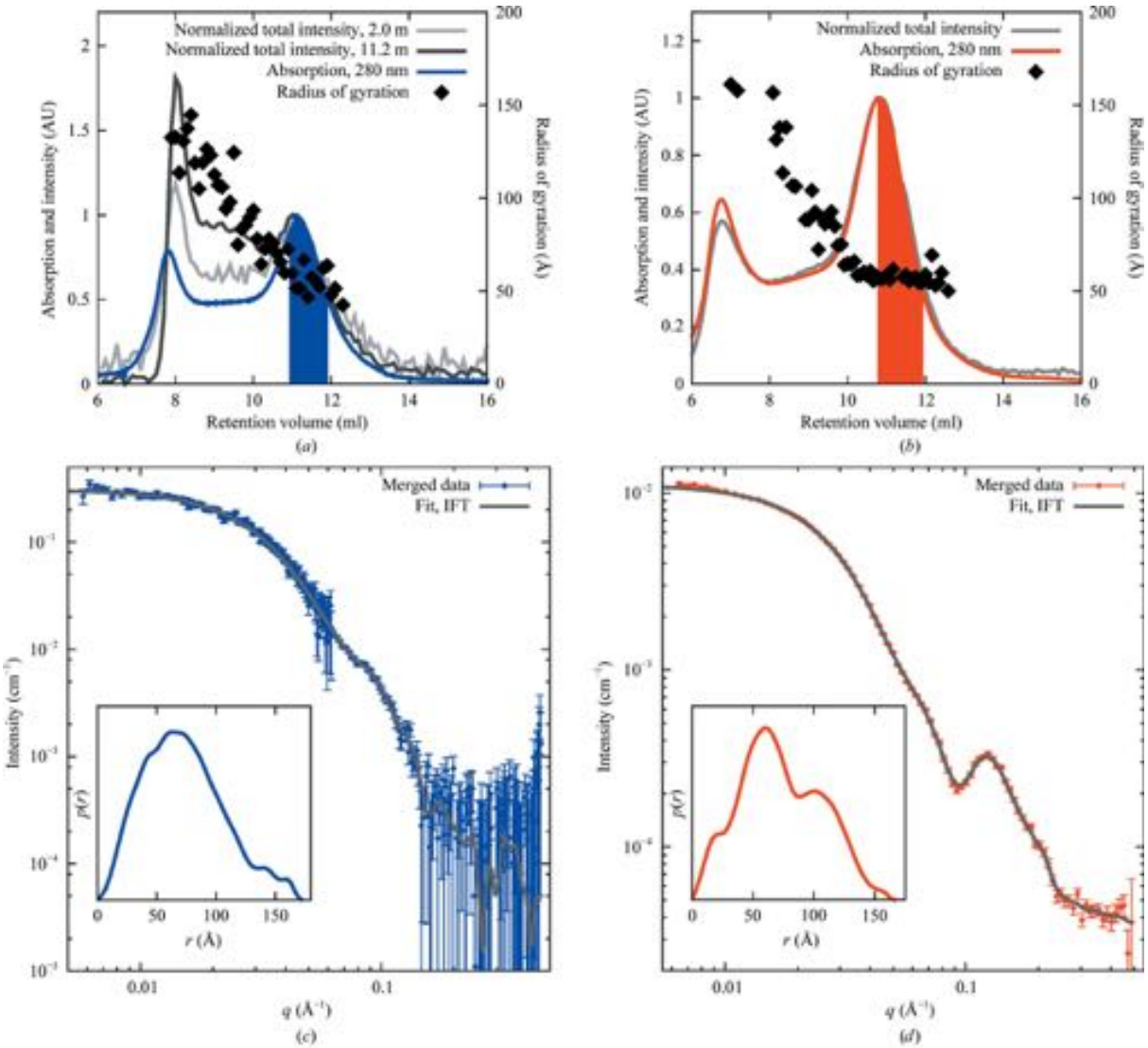
(*e*)

Figure 3
(*a*, *b*) SEC–SAS data from the experiment on PR-loaded nanodiscs as presented for empty nanodiscs in Figs. 2(*a*) and 2(*b*). (*c*, *d*) Averaged data for the frames recorded across the highlighted peaks in (*a*) and (*b*), respectively. Insets: $p(r)$ distributions refined from the data. The data sets collected from peak 2 were further analyzed as explained in the text and shown in Fig. 5(*b*). (*e*) Parameters from *BayesApp* refined from the IFT of the data.

obtained from the two peak populations (Fig. 3*c* and Guinier plots in Fig. 2*b*). The SAXS data also feature well defined Guinier regions for both samples as well as oscillating features at around $q \simeq 0.1$ Å^{-1} (Fig. 3*d*). For the sample in peak 2, the D_{max} and R_g values are only slightly larger than for the empty nanodisc (Fig. 2*g*), whereas for the sample in peak 1 they are much larger. These sizes indicate that the sample from peak 2 is monodisperse, whereas the sample from peak 1 is likely to be a polydisperse and larger aggregate structure. Further modelling and data interpretation was only carried out for the sample in peak 2, as described in Section 4.4.

4.3. TmCorA-loaded nanodiscs: removal of aggregates

Finally, to test the SEC–SANS method on a larger membrane protein with large intracellular domains, TmCorA-loaded nanodiscs were measured. Scattergrams and chromatograms are plotted for both SEC–SANS and SEC–SAXS



Data set	D_{max} (Å)	R_g (Å)
SANS ■	166.1	56.7
SAXS ■	160.3	56.8

(e)

Figure 4
(*a*, *b*) SEC–SAS data from the experiments on TmCorA-loaded nanodiscs are shown in the same manner as in Figs. 2 and 3. (*c*, *d*) Averaged scattering data. Insets: $p(r)$ distributions refined from the averaged data. (*e*) Parameters from *BayesApp* refined from the IFT of the data.

together with the calculated R_g values as shown in Fig. 4(*b*), yielding a dominating peak at 11.5 ml with a trailing signal into a peak in the void volume of the column.

Averaged data sets calculated from the frames in the selected interval are plotted together with the fit of the IFT analysis in Fig. 4(*d*) along the refined $p(r)$ distributions shown in the insets. The SANS data have reasonable statistics in the low-q region and display a well defined Guinier region (Guinier plots are shown in Supplementary Figs. S2*a* and S2*b*). Furthermore, a small distinct feature is observed at $q \simeq 0.08$ Å^{-1}. The SAXS data are also of good quality, with a well defined Guinier region and several features in the q-region above 0.04 Å^{-1}.

TmCorA has been proposed to undergo large conformational changes upon the binding of magnesium ions (Matthies *et al.*, 2016), making it an ideal system to probe with SANS and SAXS. In this work, we stay with the observation that the samples are suited for SEC–SAS and leave out further modelling. This will form part of future work focusing on TmCorA alone.

4.4. Modelling: joint refinement from SEC–SANS and SEC–SAXS data

The nanodisc model was refined from the averaged data of the empty nanodisc (Fig. 5*a*). The refined model is sketched as an inset in Fig. 5(*a*) and the refined parameters are shown in Table 2. The model resulted in a combined χ^2 of 8.76 and the presented confidence intervals are based on this value (Pedersen *et al.*, 2014). Apart from the listed parameters, a constant background (close to zero) and a term accounting for interface roughness (Als-Nielsen & McMorrow, 2011; typically up to 7 Å) were refined for each of the data sets, giving a total of four fit parameters in addition to those listed in Table 2.

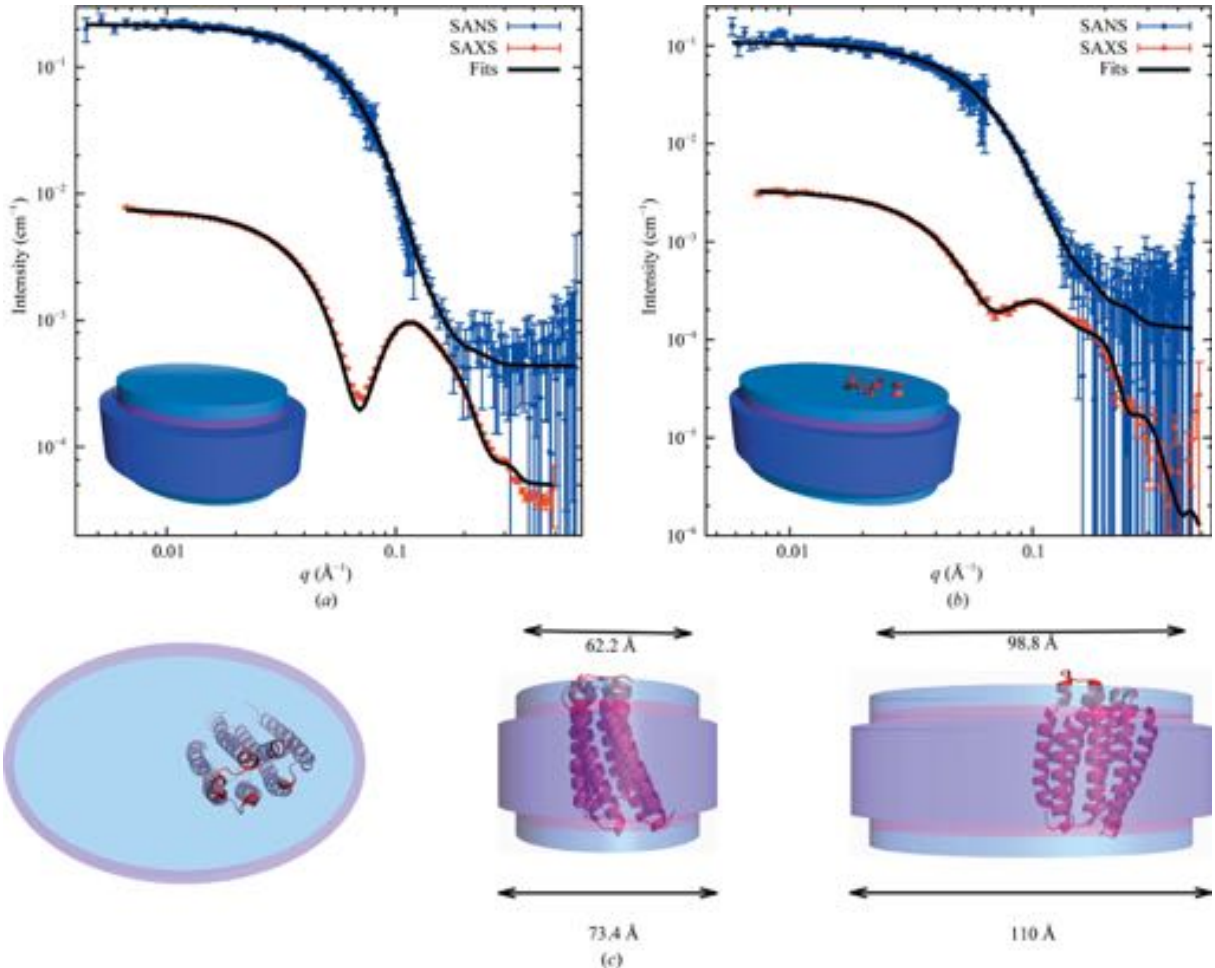


Figure 5
(*a*) Averaged data from empty nanodiscs (Fig. 2*d*) along with the fits of the nanodisc model shown in the inset. The SAXS data have been logarithmically rebinned. The His_6 tags on the MSP1D1 proteins, modelled as random coils attached to the protein belt, are not shown in the rendering. (*b*) Data from PR-loaded nanodiscs (Fig. 3*d*) with the fits of a model of PR embedded in a nanodisc as shown in the inset. (*c*) The refined model of PR in the nanodisc shown from different perspectives.

Table 2
Parameters describing the model refined from the data shown in Fig. 5.

We note that the empty nanodisc sample has not had the His_6 tag cleaved off and that the refined volume of MSP1D1 reflects this. Also note that the model for a membrane protein in a nanodisc is parametrized differently to the model presented in the previous section.

Parameter	Empty nanodiscs	PR-loaded nanodiscs
Axis ratio of bilayer patch	1.37 ± 0.11	1.59 ± 0.51
Average area per lipid headgroup (Å^2)	57.1 ± 2.1	60.0†
Hydrophobic height of bilayer (Å)	33.0†	33.8 ± 4.0
No. of lipids per nanodisc	117 ± 3.1	141 ± 11
In-plane displacement of PR from centre (Å)	—	18.8 ± 9.6
Volume of MSP1D1 (Å^3)	28200 ± 420	19500 ± 6300
Volume of POPC (Å^3)	1270 ± 7.4	1315 ± 86
Thickness of protein belt (Å)	8.40†	5.59†
Minor radius of bilayer patch (Å)	33.0†	31.1†
Major radius of bilayer patch (Å)	44.3†	49.4

† Parameters derived from the refined parameters.

The PR nanodisc model was refined from the data for the PR-loaded nanodisc eluting in peak 2 (Fig. 3*b*). The data with the associated fits, together with a rendering of the refined model, is presented in Fig. 5(*c*), and the associated parameters are shown in Table 2. The combined χ^2 of the presented model is 1.56 and, as before, the presented parameter confidences are based on this value. As for the empty nanodisc, the models include four fit parameters that are not listed in Table 2.

By requiring our models to reproduce the trends of the SANS as well as the SAXS data, we constrain the number of feasible models in our search by effectively enabling the rejection of models with greater confidence owing to the increased amount of data. This effect is amplified by the different contrasts in SAXS and SANS data, models refined from each of which predict some parameters better than others. By combining sets of data, we harvest the predictive power of both.

As outlined in the literature (Pedersen *et al.*, 2014), the effect of refining models simultaneously from several sets of data is improved confidence intervals of the refined parameters. As an example, in our case, owing to the different contrasts of the constituents of the nanodisc, adding SANS data to the refinement improves our confidence on the number of lipids in the nanodisc in particular: refining the nanodisc model from only the presented SAXS data results in an estimate of 106 ± 8.0 lipids per nanodisc, whereas refining including the SANS data in the model refinement yields an estimate of 117 ± 3.1 for the same parameter.

5. Discussion

5.1. Benchmarking the SEC–SANS setup on empty nanodiscs

The nanodisc is a popular platform for reconstituting membrane proteins in a native-like environment and studying their structure and function (Denisov & Sligar, 2016). However, samples of membrane proteins inside nanodiscs are demanding to produce in large amounts, and monodisperse samples are difficult to achieve, especially in D_2O-based solvents, because of issues of increased aggregation. This makes them difficult to study by SANS. Using a newly developed SEC–SANS setup on D22 at the ILL (Jordan *et al.*, 2016), we obtained SANS data with a sufficiently high signal-to-noise ratio from monodisperse samples of nanodiscs, both empty (Fig. 2*a*) and loaded with the membrane proteins PR (Fig. 3) and TmCorA (Fig. 4), respectively, in 100% D_2O. This shows that the method is practically applicable to such complex biological systems.

Compared with static SANS, sample handling is easier in SEC–SANS as the final purification step is carried out online. The success rate of the SEC–SANS experiments presented here, in terms of the proportion of good data sets, was markedly higher than for standard solution SANS experiments performed previously by our group. The method has the additional advantage that both buffer and sample are passed through the same flow cell, making background subtraction more accurate than in standard cuvette-based static SANS experiments. Non-optimal background subtraction from SANS data might result in incorrect values for calculated sizes (Rubinson *et al.*, 2008).

SEC–SAS is especially advantageous for the study of aggregation-prone systems or systems in an equilibrium between structural or oligomeric states, whereas highly stable and monodisperse systems would preferably be measured with longer static measurements to obtain better counting statistics. The nanodisc is usually described as a very stable and monodisperse system, but its scattering pattern varied slightly but significantly as a function of retention volume as observed by SEC–SAS (Figs. 2*c* and 2*d*). This variation is ascribed to a small structural dispersity of the nanodiscs, with a variation in the lipid aggregation number as a function of retention volume, in line with other results from our group (manuscript in preparation) and consistent with recently reported findings for MSP1E3D1 nanodiscs (Graziano *et al.*, 2018). This structural effect would not have been visible without SEC–SAS.

The empty nanodiscs were found to have a slightly smaller D_{max} in SANS compared with SAXS (Fig. 2*e*), which is in agreement with previous studies (Kynde *et al.*, 2014; Skar-Gislinge *et al.*, 2010) and is most likely to be a result of the nanodiscs having a higher peripheral excess scattering-length density in the SAXS contrast than in the SANS contrast. Interestingly, however, the D_{max} values determined in this study from both SAXS and SANS are approximately 10–20 Å smaller than those reported in these studies. For a static SAXS or SANS measurement the nanodiscs are SEC-purified in advance, which may allow sufficient time for a monodisperse population of nanodiscs to equilibrate to a broader distribution of nanodiscs with different lipid:MSP ratios or even for the sample to start forming small populations of slightly larger particles. Our data indicate that in the SEC–SAS setups such issues are under better control as the measured samples of the individual frames are more monodisperse.

5.2. Feasibility for demanding complex biological samples

The SEC–SAS experiments were carried out with sample volumes in the range of a few hundred microlitres and

concentrations of the reconstituted membrane protein of a few mg ml^{-1} (Table 1) and yielded surprisingly good-quality SANS and SAXS data in terms of signal-to-noise ratio.

Firstly, PR-loaded nanodiscs were measured. A typical purification of PR yields several species, including monomers, trimers and hexamers (Gourdon *et al.*, 2008; Stone *et al.*, 2013), and two major species were indeed observed in the PR-loaded nanodisc samples (Fig. 3): a polydisperse population with charateristic sizes that were too large for a monomeric species and a population with the expected sizes for a monomeric PR-loaded nanodisc. The D_{max} values for the monomeric PR-loaded nanodisc sample from SANS and SAXS (Fig. 3*e*), respectively, were close to those refined for the empty nanodisc (Fig. 2*g*), which is expected given that PR has almost all of its structure embedded in the lipid bilayer of the nanodisc. Despite the low concentration of the sample in the frames extracted from peak 1, the resulting data were of sufficient quality and allowed refinement of the model of PR embedded in a nanodisc (Fig. 5*c*). The D_{max} was approximately 10 Å smaller than that reported for the similar membrane protein bacteriorhodopsin in MSP1D1-based nanodiscs as measured by standard static SANS and SAXS (Kynde *et al.*, 2014). This could indicate that a more monodisperse sample, and hence data of better quality, could be obtained using the combined SEC–SANS and SEC–SAXS setup.

The concentration of the TmCorA-loaded nanodisc sample was lower than that of the PR-loaded nanodisc, but the larger size of the system led to an increased scattering intensity and the obtained SANS and SAXS data (Fig. 4*d*) had a slightly better signal-to-noise ratio than the data from the PR-loaded nanodisc sample (Fig. 4). Our results might serve as a guideline for the required amount of sample for similar types of membrane proteins in nanodiscs, including the highly interesting class of G-protein coupled receptors, which possess a similar overall structure in the membrane to that of PR.

5.3. Discussion of model fit results

A good model fit was obtained to the data from the empty nanodiscs (Fig. 5*a*). However, the model did not quite manage to reproduce the trend of the SAXS data around the minimum at $q \simeq 0.08$ Å^{-1}, a property that is shared by several published fits on similar structures (Midtgaard *et al.*, 2015). This small deviation is most likely to be the result of the underlying structural dynamics and small structural dispersity of the nanodiscs (Fig. 2), as the model, for simplicity, assumes a static monodisperse structure. Still, the refined model of the empty nanodisc is in line with our expectations in terms of the refined values for the various parameters of the model. In particular, the partial specific molecular volume is refined to a value of 28200 Å^3, corresponding to a specific mass density of 1.46 g ml^{-1}, which is in good agreement with the reported value of 1.35 g ml^{-1} for proteins (Mylonas & Svergun, 2007). It is noteworthy, though, that the refined value of the average area per POPC headgroup was found to be 57.1 Å^2. This is somewhat lower than the value of 60 Å^2 reported for POPC in large bilayer vesicles (Kučerka *et al.*, 2011) and is also lower than the value of 66 Å^2 determined for MSP1D1–POPC nanodiscs in a previous study (Skar-Gislinge *et al.*, 2010).

The fit of the data acquired from the PR-loaded nanodiscs managed to reproduce the trends of the data consistently (Fig. 5*b*). The POPC area per headgroup was derived as 60 Å^2 from the refined model, which is in agreement with the value reported for POPC in unperturbed bilayers (Kučerka *et al.*, 2011). However, the value for the specific mass density of MSP1D1 refined from the data on PR-loaded nanodiscs is 1.88 g ml^{-1} and thus differs considerably from the expected value. Similarly, the refined value for the number of lipids in the nanodisc is slightly higher than anticipated. As a consequence, the MSP belt surrounding the loaded nanodisc adopts somewhat unphysical dimensions in the model. However, the molecular volume of MSP1D1 is not well determined, as seen by the large error bars on the value of this parameter. This is of some concern given the good quality of the SAS data. We suspect that it could be the result of a weakly determined concentration in the SEC–SAXS experiment, which is an issue that has been addressed elsewhere (Graewert *et al.*, 2015; Ryan *et al.*, 2018). An accurate estimate of sample concentration is paramount for fitting a complex many-parameter model.

For both the empty nanodiscs and the PR-loaded nanodiscs, the obtained SEC–SANS data showed no indications of aggregates at low q as judged by the flat Guinier regions. Furthermore, the data had a good signal-to-noise ratio up to $q \simeq 0.15$ Å^{-1} (Figs. 5*a* and 5*b*). As seen from the figure, the obtained SEC–SAXS data for empty nanodiscs had a very good signal-to-noise ratio up to the maximum q value, whereas data from PR-loaded nanodiscs had significant error bars from $q \simeq 0.25$ Å^{-1} owing to the lower sample concentration. In the present analysis we found that good (aggregate-free) SANS data up to $q \simeq 0.1$ Å^{-1} were sufficient to have a major impact on the model fit when they were combined with high-quality SAXS data, in particular regarding the parameters describing the lipid component. All in all, the complementary information obtained in the different contrast situations in SANS and SAXS is essential when refining complex models.

6. Conclusions and future perspectives

SEC–SANS was demonstrated to be feasible for studying the structures of empty nanodiscs as well as of membrane-protein-loaded nanodiscs. Data of surprisingly good quality and without signs of aggregation were obtained for all presented samples. The quality of the combined SEC–SANS and SEC–SAXS data obtained was sufficient to refine structural models of empty and PR-loaded nanodiscs.

Complex biological samples are demanding to produce in sufficient amounts and homogeneity for structural measurements, including SAXS and especially SANS. The high flux and a state-of-the-art detector at D22 at the ILL lowers the limit for the required amount of sample, and the new SEC–SANS setup has made it possible to study truly monodisperse samples. In the presented sets of experiments, the flow rates were limited to 0.3 and 0.2 ml min^{-1}, respectively. With the

continued development of the setup, stable flow rates down to 0.01 ml min^{-1} are now available, which allow better tuning of exposure times, limited at one extent by the amount of sample and at the other by the stability of the sample, *i.e.* the resistance to aggregation over time. This opens the possibility of obtaining better statistics than for the data presented here, and/or of the loading of less material onto the column.

For the present article, the SEC–SANS experiments had to be repeated two or three times with different sample-to-detector distances to cover the desired q-range. The D22++ project within the Endurance phase II upgrade (proposal for funding submitted) includes a fixed detector placed at a short sample-to-detector distance, making it possible to cover a large q-range in a single SEC run. This upgrade will thus reduce both the time and the amount of sample needed for the experiment. Furthermore, a setup is being implemented with the possibility of running two SEC columns in parallel, enabling the overlap of sample elution from one column with the void volume from the other column, in turn making it possible to continuously measure relevant sample frames instead of wasting time and neutrons on excess buffer frames. Furthermore, this setup also has potential for automating measurements of contrast-variation series by programming the mixing of H_2O-based and D_2O-based buffers.

The anticipated high neutron flux at the upcoming European Spallation Source has potential for much improved SANS instrumentation and for either faster measurements and/or smaller sample volumes than at present sources. This, in combination with the general interest in complex biological structures, drives an increase in the level of ambition with respect to how complex biological samples may be investigated using SANS. As a consequence, dedicated BioSANS instruments or BioSANS instrument activities have become increasingly consolidated in several places in the world during the last decade. The high flux at D22 combined with online SEC is the optimal setup at present for studying complex and potentially unstable biological systems using SANS. Thus, for future studies of extremely precious samples, such as match-out deuterated nanodiscs (Maric *et al.*, 2014), this setup will be preferred.

Acknowledgements

The authors acknowledge the ILL and ESRF for developments in SEC–SAS and allocated beam time. We further thank the D22 technician Mark Jacques and the beamline scientists Marta Brennich and Petra Pernot at BM29 for their support during the beam times. Furthermore, we thank Søren Roi Midtgaard for help during the planning and execution of the first SEC–SANS experiment, as well as Frederik Grønbæk Tidemand for his help with MSP1E3D1 production and his assistance during the second SEC–SAXS beamtime. We thank Professor Steven Sligar for kindly providing the plasmids encoding MSP constructs, Professor Richard Neutze for the plasmid encoding PR and Professor Mikaela Rapp, Heena Sandhu and Jens Berndtsson (University of Stockholm, Sweden) for the plasmid encoding TmCorA and for helpful advice on protein purification.

Funding information

Funding for this research was provided by the Synergy programme of the Novo Nordisk Foundation, the BRAINSTRUC programme of the Lundbeck Foundation and BioSynergy KU2016.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213–221.
Als-Nielsen, J. & McMorrow, D. (2011). *Elements of Modern X-ray Physics*, 2nd ed. Chichester: Wiley.
Appolaire, A., Girard, E., Colombo, M., Durá, M. A., Moulin, M., Härtlein, M., Franzetti, B. & Gabel, F. (2014). *Acta Cryst.* D**70**, 2983–2993.
Bayburt, T. H., Grinkova, Y. V. & Sligar, S. G. (2002). *Nano Lett.* **2**, 853–856.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bibow, S., Polyhach, Y., Eichmann, C., Chi, C. N., Kowal, J., Albiez, S., McLeod, R. A., Stahlberg, H., Jeschke, G., Güntert, P. & Riek, R. (2017). *Nature Struct. Mol. Biol.* **24**, 187–193.
David, G. & Pérez, J. (2009). *J. Appl. Cryst.* **42**, 892–900.
Denisov, I. G., Grinkova, Y. V., Lazarides, A. A. & Sligar, S. G. (2004). *J. Am. Chem. Soc.* **126**, 3477–3487.
Denisov, I. G. & Sligar, S. G. (2016). *Nature Struct. Mol. Biol.* **23**, 481–486.
Dewhurst, C. (2017). *GRASP* v.7.15. https://www.ill.eu/fr/users-en/scientific-groups/large-scale-structures/grasp/.
Fagerberg, L., Jonasson, K., von Heijne, G., Uhlén, M. & Berglund, L. (2010). *Proteomics*, **10**, 1141–1149.
Förster, S., Apostol, L. & Bras, W. (2010). *J. Appl. Cryst.* **43**, 639–646.
Gabel, F. (2015). *Methods Enzymol.* **558**, 391–415.
Gabel, F., Lensink, M. F., Clantin, B., Jacob-Dubuisson, F., Villeret, V. & Ebel, C. (2014). *Biophys. J.* **107**, 185–196.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.
Glatter, O. (1977). *J. Appl. Cryst.* **10**, 415–421.
Gourdon, P., Alfredsson, A., Pedersen, A., Malmerberg, E., Nyblom, M., Widell, M., Berntsson, R., Pinhassi, J., Braiman, M., Hansson, Ö., Bonander, N., Karlsson, G. & Neutze, R. (2008). *Protein Expr. Purif.* **58**, 103–113.
Graewert, M. A., Franke, D., Jeffries, C. M., Blanchet, C. E., Ruskule, D., Kuhle, K., Flieger, A., Schäfer, B., Tartsch, B., Meijers, R. & Svergun, D. I. (2015). *Sci. Rep.* **5**, 10734.
Graewert, M. A. & Svergun, D. I. (2013). *Curr. Opin. Struct. Biol.* **23**, 748–754.
Graziano, V., Miller, L. & Yang, L. (2018). *J. Appl. Cryst.* **51**, 157–166.
Gripon, C., Legrand, L., Risenman, I., Vidal, O., Robert, M. C. & Boué, F. (1997). *J. Cryst. Growth*, **177**, 238–247.
Hagn, F., Etzkorn, M., Raschle, T. & Wagner, G. (2013). *J. Am. Chem. Soc.* **135**, 1919–1925.
Hansen, S. (2014). *J. Appl. Cryst.* **47**, 1469–1471.
Hebling, C. M., Morgan, C. R., Stafford, D. W., Jorgenson, J. W., Rand, K. D. & Engen, J. R. (2010). *Anal. Chem.* **82**, 5415–5419.
Hopkins, J. B., Gillilan, R. E. & Skou, S. (2017). *J. Appl. Cryst.* **50**, 1545–1553.

Jeffries, C. M., Graewert, M. A., Blanchet, C. E., Langley, D. B., Whitten, A. E. & Svergun, D. I. (2016). *Nature Protoc.* **11**, 2122–2153.

Jordan, A., Jacques, M., Merrick, C., Devos, J., Forsyth, V. T., Porcar, L. & Martel, A. (2016). *J. Appl. Cryst.* **49**, 2015–2020.

Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N., Dothager, R. S., Seifert, S., Thiyagarajan, P., Sosnick, T. R., Hasan, M. Z., Pande, V. S., Ruczinski, I., Doniach, S. & Plaxco, K. W. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 12491–12496.

Koutsioubas, A. (2017). *Biophys. J.* **113**, 2373–2382.

Kučerka, N., Nieh, M.-P. & Katsaras, J. (2011). *Biochim. Biophys. Acta*, **1808**, 2761–2771.

Kynde, S. A. R., Skar-Gislinge, N., Pedersen, M. C., Midtgaard, S. R., Simonsen, J. B., Schweins, R., Mortensen, K. & Arleth, L. (2014). *Acta Cryst.* D**70**, 371–383.

Larsson, U. (1988). *Eur. J. Biochem.* **174**, 139–144.

Levenberg, K. (1944). *Q. Appl. Math.* **2**, 164–168.

Lomize, M. A., Lomize, A. L., Pogozheva, I. D. & Mosberg, H. I. (2006). *Bioinformatics*, **22**, 623–625.

Maric, S., Skar-Gislinge, N., Midtgaard, S., Thygesen, M. B., Schiller, J., Frielinghaus, H., Moulin, M., Haertlein, M., Forsyth, V. T., Pomorski, T. G. & Arleth, L. (2014). *Acta Cryst.* D**70**, 317–328.

Marquardt, D. W. (1963). *J. Soc. Ind. Appl. Math.* **11**, 431–441.

Mathew, E., Mirza, A. & Menhart, N. (2004). *J. Synchrotron Rad.* **11**, 314–318.

Matthies, D., Dalmas, O., Borgnia, M. J., Dominik, P. K., Merk, A., Rao, P., Reddy, B. G., Islam, S., Bartesaghi, A., Perozo, E. & Subramaniam, S. (2016). *Cell*, **164**, 747–756.

Mehmood, S., Domene, C., Forest, E. & Jault, J.-M. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 10832–10836.

Midtgaard, S. R., Darwish, T. A., Pedersen, M. C., Huda, P., Larsen, A. H., Jensen, G. V., Kynde, S. A. R., Skar-Gislinge, N., Nielsen, A. J. Z., Olesen, C., Blaise, M., Dorosz, J. J., Thorsen, T. S., Venskutonytė, R., Krintel, C., Møller, J. V., Frielinghaus, H., Gilbert, E. P., Martel, A., Kastrup, J. S., Jensen, P. E., Nissen, P. & Arleth, L. (2018). *FEBS J.* **285**, 357–371.

Midtgaard, S. R., Pedersen, M. C. & Arleth, L. (2015). *Biophys. J.* **109**, 308–318.

Morgan, C. R., Hebling, C. M., Rand, K. D., Stafford, D. W., Jorgenson, J. W. & Engen, J. R. (2011). *Mol. Cell. Proteomics*, **10**, M111.010876.

Mylonas, E. & Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s245–s249.

Niemann, H. H., Petoukhov, M. V., Härtlein, M., Moulin, M., Gherardi, E., Timmins, P., Heinz, D. W. & Svergun, D. I. (2008). *J. Mol. Biol.* **377**, 489–500.

Nikolaev, M., Round, E., Gushchin, I., Polovinkin, V., Balandin, T., Kuzmichev, P., Shevchenko, V., Borshchevskiy, V., Kuklin, A., Round, A., Bernhard, F., Willbold, D., Büldt, G. & Gordeliy, V. (2017). *Cryst. Growth Des.* **17**, 945–948.

Oldfield, C. J., Cheng, Y., Cortese, M. S., Brown, C. J., Uversky, V. N. & Dunker, A. K. (2005). *Biochemistry*, **44**, 1989–2000.

Oldfield, C. J. & Dunker, A. K. (2014). *Annu. Rev. Biochem.* **83**, 553–584.

Oliver, R. C., Pingali, S. V. & Urban, V. S. (2017). *J. Phys. Chem. Lett.* **8**, 5041–5046.

Orthaber, D., Bergmann, A. & Glatter, O. (2000). *J. Appl. Cryst.* **33**, 218–225.

Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. (2006). *Nature Rev. Drug Discov.* **5**, 993–996.

Palombo, I., Daley, D. O. & Rapp, M. (2012). *J. Biol. Chem.* **287**, 27547–27555.

Pedersen, J. S. (1997). *Adv. Colloid Interface Sci.* **70**, 171–210.

Pedersen, J. S., Posselt, D. & Mortensen, K. (1990). *J. Appl. Cryst.* **23**, 321–333.

Pedersen, M. C., Arleth, L. & Mortensen, K. (2013). *J. Appl. Cryst.* **46**, 1894–1898.

Pedersen, M. C., Hansen, S. L., Markussen, B., Arleth, L. & Mortensen, K. (2014). *J. Appl. Cryst.* **47**, 2000–2010.

Pernot, P., Round, A., Barrett, R., De Maria Antolinos, A., 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). *J. Synchrotron Rad.* **20**, 660–664.

Petoukhov, M. V., Konarev, P. V., Kikhney, A. G. & Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s223–s228.

Reckel, S., Gottstein, D., Stehle, J., Löhr, F., Verhoefen, M. K., Takeda, M., Silvers, R., Kainosho, M., Glaubitz, C., Wachtveitl, J., Bernhard, F., Schwalbe, H., Güntert, P. & Dötsch, V. (2011). *Angew. Chem. Int. Ed.* **50**, 11942–11946.

Ritchie, T. K., Grinkova, Y. V., Bayburt, T. H., Denisov, I. G., Zolnerciks, J. K., Atkins, W. M. & Sligar, S. G. (2009). *Methods Enzymol.* **464**, 211–231.

Rubinson, K. A., Stanley, C. & Krueger, S. (2008). *J. Appl. Cryst.* **41**, 456–465.

Ryan, T. M., Trewhella, J., Murphy, J. M., Keown, J. R., Casey, L., Pearce, F. G., Goldstone, D. C., Chen, K., Luo, Z., Kobe, B., McDevitt, C. A., Watkin, S. A., Hawley, A. M., Mudie, S. T., Samardzic-Boban, V. & Kirby, N. (2018). *J. Appl. Cryst.* **51**, 97–111.

Schneidman-Duhovny, D., Hammel, M. & Sali, A. (2010). *Nucleic Acids Res.* **38**, W540–W544.

Skar-Gislinge, N. & Arleth, L. (2011). *Phys. Chem. Chem. Phys.* **13**, 3161–3170.

Skar-Gislinge, N., Simonsen, J. B., Mortensen, K., Feidenhans'l, R., Sligar, S. G., Lindberg Møller, B., Bjørnholm, T. & Arleth, L. (2010). *J. Am. Chem. Soc.* **132**, 13713–13722.

Skou, S., Gillilan, R. E. & Ando, N. (2014). *Nature Protoc.* **9**, 1727–1739.

Stone, K. M., Voska, J., Kinnebrew, M., Pavlova, A., Junk, M. J. N. & Han, S. (2013). *Biophys. J.* **104**, 472–481.

Svergun, D. I. & Koch, M. H. J. (2003). *Rep. Prog. Phys.* **66**, 1735–1782.

Ward, A. B., Sali, A. & Wilson, I. A. (2013). *Science*, **339**, 913–915.

Whitten, A. E., Jacques, D. A., Hammouda, B., Hanley, T., King, G. F., Guss, J. M., Trewhella, J. & Langley, D. B. (2007). *J. Mol. Biol.* **368**, 407–420.

Zaccai, N. R., Sandlin, C. W., Hoopes, J. T., Curtis, J. E., Fleming, P. J., Fleming, K. G. & Krueger, S. (2016). *Methods Enzymol.* **566**, 159–210.

Acta Cryst D STRUCTURAL BIOLOGY

ISSN: 2059-7983

ORDER FORM

YOU WILL AUTOMATICALLY BE SENT DETAILS OF HOW TO DOWNLOAD
AN ELECTRONIC REPRINT OF YOUR PAPER, FREE OF CHARGE.
PRINTED REPRINTS MAY BE PURCHASED USING THIS FORM.

Please scan your order and send to lj@iucr.org

INTERNATIONAL UNION OF CRYSTALLOGRAPHY
5 Abbey Square
Chester CH1 2HU, England.

VAT No. GB 161 9034 76

Article No.: D180718-FS5159

Title of article	Introducing SEC–SANS for studies of complex self-organized biological systems
Name	Lise Arleth
Address	Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
E-mail address (for electronic reprints)	arleth@nbi.ku.dk

OPEN ACCESS

IUCr journals offer authors the chance to make their articles open access on **Crystallography Journals Online**. For full details of our open-access policy, see http://journals.iucr.org/services/openaccess.html. For authors in European Union countries, VAT will be added to the open-access charge.

If you wish to make your article open access please go to http://shop.iucr.org/iucrshop/viewitem/openaccess/?code=FS5159 The charge for making an article open access is **1300 United States dollars**.

DIGITAL PRINTED REPRINTS

I wish to order . . . . . paid reprints

These reprints will be sent to the address given above. If the above address or e-mail address is not correct, please indicate an alternative:

PAYMENT (REPRINTS ONLY)

Charge for reprints . . . . . USD

☐ An official purchase order made out to **INTERNATIONAL UNION OF CRYSTALLOGRAPHY** ☐ is enclosed ☐ will follow

Purchase order No.

☐ Please invoice me

☐ I wish to pay by credit card

EU authors only: VAT No:

Date	Signature

OPEN ACCESS

The charge for making an article open access is **1300 United States dollars**. For authors in European Union countries, VAT will be added to the open-access charge.

A paper may be made open access at any time after the proof stage on receipt of the appropriate payment. This includes all back articles on **Crystallography Journals Online**. For further details, please contact support@iucr.org. Likewise, organizations wishing to sponsor open-access publication of a series of articles or complete journal issues should contact support@iucr.org.

DIGITAL PRINTED REPRINTS

An electronic reprint is supplied free of charge.

Printed reprints without limit of number may be purchased at the prices given in the table below. The requirements of all joint authors, if any, and of their laboratories should be included in a single order, specifically ordered on the form overleaf. All orders for reprints must be submitted promptly.

Prices for reprints are given below in **United States dollars** and include postage.

	Size of paper (in printed pages)				
Number of reprints required	1–2	3–4	5–8	9–16	Additional 8's
50	184	268	372	560	246
100	278	402	556	842	370
150	368	534	740	1122	490
200	456	664	920	1400	610
Additional 50's	86	128	178	276	116

PAYMENT AND ORDERING

Open-access fees should be paid at http://shop.iucr.org/iucrshop/viewitem/openaccess/?code=FS5159

Official purchase orders should be made out to **INTERNATIONAL UNION OF CRYSTALLOGRAPHY**.

Orders should be returned by email to lj@iucr.org

ENQUIRIES

Enquiries concerning reprints should be sent to support@iucr.org.

Combining SEC-SANS and SEC-SAXS Introducing SEC-SANS for studies of complex self-organised biological systems

Nicolai Tidemand Johansen,[a] Martin Cramer Pedersen,[a]
Lionel Porcar,[b] Anne Martel[b]* and Lise Arleth [a]*

[a]*Niels Bohr Institute, University of Copenhagen, Denmark, and* [b]*Institut Laue-Langevin, Grenoble, France. E-mail: martela@ill.fr, arleth@nbi.ku.dk*

size-exclusion chromatography; small-angle neutron scattering; small-angle X-ray scattering; membrane proteins; phospholipid nanodiscs

1. Supporting information

1.1. Protein expression and purification

Plasmids were transformed into *Escherichia coli* (*E. coli*) BL21(DE3) star (Novagen) under appropriate antibiotic selection. A starter culture in Lysogeny Broth medium was grown from a single bacterial colony and used to inoculate a larger culture in Terrific Broth medium. For MSPs, protein expression was induced with 1 mM IPTG at an optical density at 600 nm = 0.8. For membrane proteins, cultures were grown to an optical density at 600 nm = 0.8 and cooled to 15 °C before adding 0.4 mM IPTG or 1.25 g/L arabinose and 1.5 mg/L all trans-retinal for TmCorA and PR, respectively. MSP cultures were harvested after 3 h, whereas the membrane protein cultures were harvested after 24 h.

For purification of MSPs, cells were lysed in lysis buffer (50 mM TrisHCl pH 8, 300 mM NaCl, 20 mM imidazole) containing 6 M guanidinium hydrochloride using vigorous shaking. Insoluble material was pellet at 25000g for 30 min and the supernatant was loaded onto equilibrated NiNTA resin (Qiagen) in a disposable column. The column was washed in 1 column volume (CV) of lysis buffer containing 6 M guanidinium hydrochloride, then in 3 CV of wash buffer (50 mM TrisHCl pH 8, 300 mM NaCl, 40 mM imidazole) containing 10 mM sodium cholate, and finally in 3 CV of wash buffer. Protein was eluted in elution buffer (50 mM TrisHCl pH 8, 300 mM imidazole) and dialysed 100 fold against gel filtration buffer (20 mM TrisHCl pH 7.5, 100 mM NaCl). MSPs used for loaded nanodiscs had the N-terminal His_6-tag cleaved by TEV protease (1 : 100) overnight at room temperature in gel filtration buffer with 1 mM DTT and 0.5 mM EDTA added. TEV protease and uncleaved MSP was removed by passing the sample over a NiNTA column. Samples were concentrated to A280 $\geq$ 5 using 10 kDa MWCO spin filters and frozen and stored at -80 °C.

For membrane isolation of TmCorA and PR, cells were resuspended in lysis buffer or PR-lysis buffer (50 mM MOPS pH 7.5, 1 mM EDTA), respectively, with 1 mM PMSF, protease inhibitor tablets (Roche) and 100 μg/ml lysozyme. After incubation on ice for 30 min, 2 mM $MgCl_2$ and benzonase (Merck) was added and the cells were incubated with mild stirring until the suspension had homogenized. The cells were then lysed with a cell disruptor (Constant Systems Ltd) at 34 kPSI. Insoluble material was spun down at 14000g for 30 min and the supernatant was transferred to ultracentrifugation tubes (Beckman). After centrifugation at 150000g for 1 h, the supernatant was discarded and the pelleted membranes were resuspended in lysis buffer or PR-lysis buffer, respectively, using a tissue grinder. For solubilizing membranes, the TmCorA sample had n-dodecyl-β-D-maltoside (DDM) added to a final concentration of 19.6 mM, and the PR sample had n-octyl-β-D-glucoside (OG) added to a final concentration of 205

mM. Solubilization was carried out overnight at 4 °C with slow magnetic stirring. The following day, insoluble material was removed with centrifugation at 150000g for 30 min.

For purification of TmCorA, the supernatant was mixed with equillibrated NiNTA resin and incubated at 4 °C for 1 h before loaded onto a disposable column. The column was washed with 1 CV of lysis buffer, then 3 CV of TmCorA-wash buffer (50 mM TrisHCl pH 7.9, 150 mM NaCl, 40 mM imidazole, 0.59 mM DDM) and finally, protein was eluted in 0.5 CV fractions in TmCorA elution buffer (50 mM TrisHCl pH 7.9, 150 mM NaCl, 300 mM imidazole, 0.59 mM DDM). Fractions containing TmCorA were pooled and diluted 10 fold in lysis buffer containing 0.59 mM DDM and concentrated in a 30 kDa MWCO spin filter (Merck). The sample was aliquoted, flash frozen in liquid nitrogen and stored at -80°C.

For purification of PR, the supernatant was diluted 1 : 1 in DIL buffer (50 mM MOPS pH 7.5, 34.2 mM OG) and mixed with equillibrated NiNTA resin. The column was washed in 3 CV IMAC1 buffer (50 mM MOPS pH 7.5, 30 mM imidazole, 34.2 mM OG), and protein was eluted in IMAC2 buffer (50 mM MOPS pH 7.5, 250 mM imidazole, 34.2 mM OG), which could visibly be followed by the strong red color of the protein. The eluate was concentrated to < 1 mL in a 30 kDa MWCO spin filter and subsequently diluted 50 fold in ION1buffer (50 mM sodium phosphate pH 6.0, 34.2 mM OG). Aggregated protein was spun down at 10000g for 30 min, and the supernatant was loaded onto a SP sepharose cation exchange column (Pharmacia) using an Äkta Pure system (GE). The protein was eluted with a 25 CV gradient from 0 to 50% ION2 buffer (50 mM sodium phosphate pH 8.0, 34.2 mM OG). The eluted protein was pooled, concentrated and added glycerol to a final concentration of 10% v/v glycerol, then flash frozen and stored at -80°C.

1.2. SAXS measurements of empty nanodiscs in H_2O and D_2O

A sample of empty nanodiscs was prepared in H_2O-based buffer and split in two equal samples. One sample was buffer exchanged into D_2O-based buffer on a NAP-5 desalting column (GE). Complete buffer exchange in the sample was verified by densitometry. Standard static SAXS was measured at BM29, and background subtracted data was taken directly from the output of the in-house software and logarithmically rebinned. The data was normalized for easy comparison of the two samples and plotted in Figure **??** along with $p(r)$-functions shown in the inserts. The data show that the empty nanodiscs maintain their overall structure after exchange into D_2O.

1.3. Guinier plots

Guinier plots were generated for all averaged data presented in the main article (Figure **??**). The data were fitted by the Guinier approximation at low q:

$$I(q) = I_0 e^{-{R_g}^2 q^2/3} \tag{1}$$

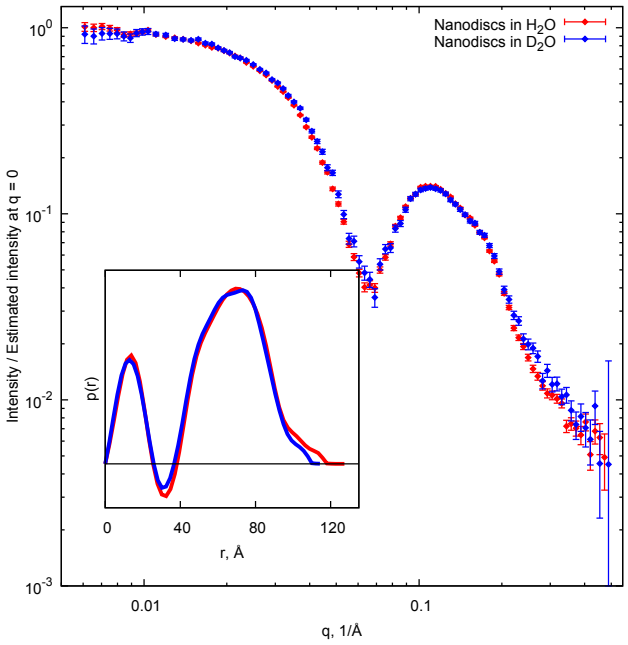


Fig. S1. (a): Empty nanodiscs measured with standard static SAXS at BM29, before (blue) and after solvent exchange into D_2O (red) (b): $p(r)$-distributions refined from the data in (a).

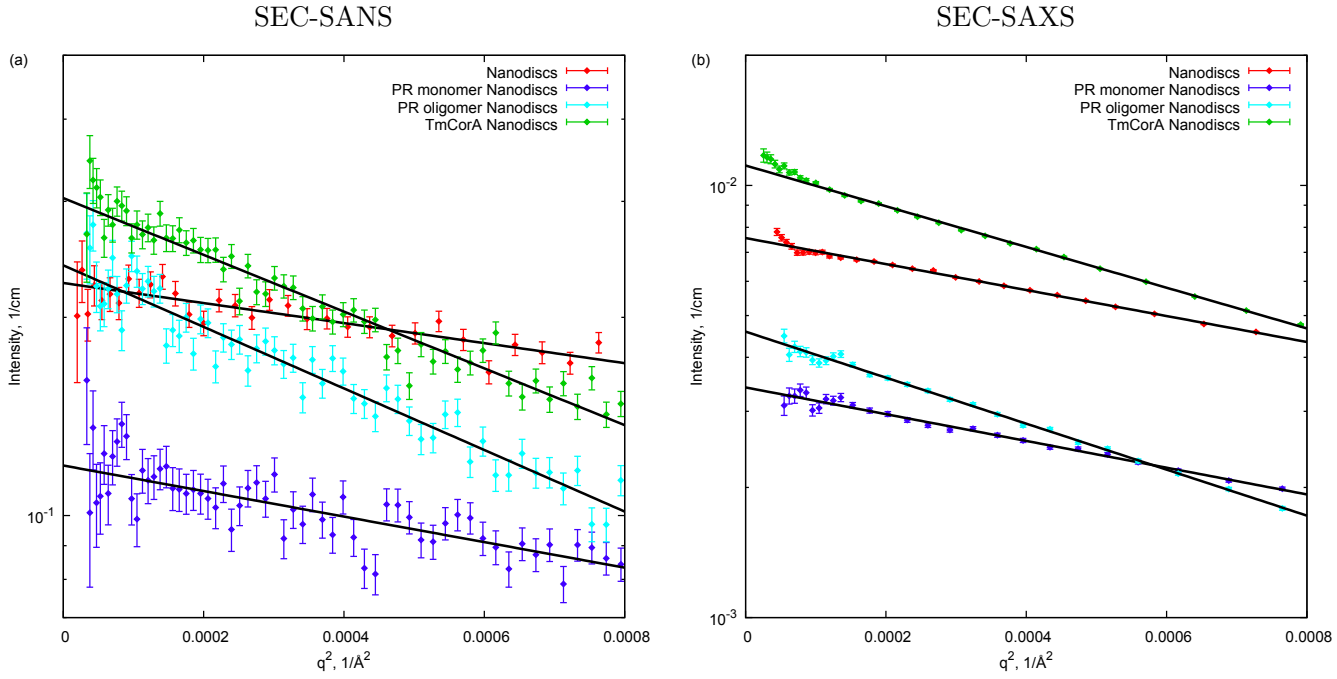
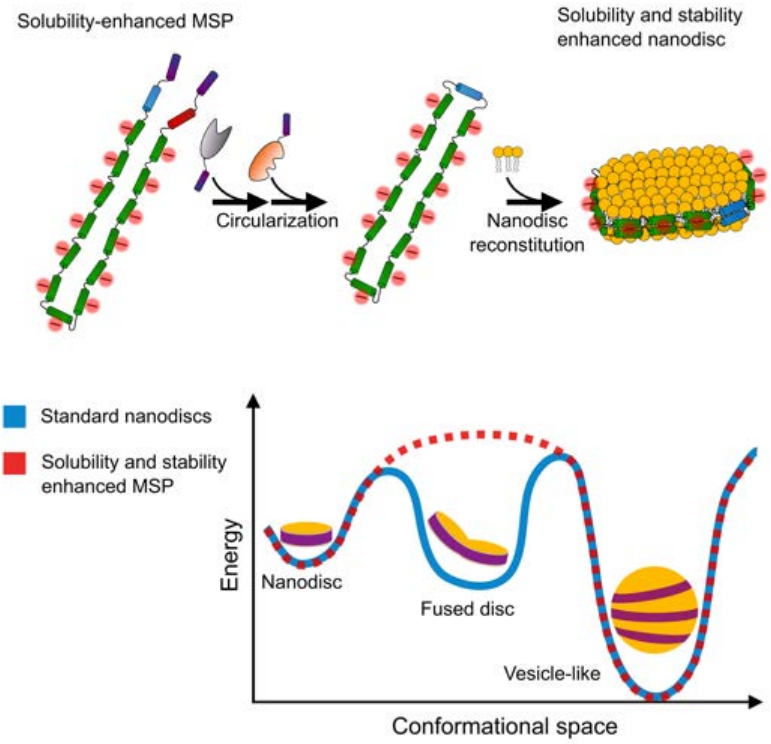


Fig. S2. (a) and (b): Guinier plots including fits for all averaged data sets presented in the main article. The datasets are shown on absolute scale. The plots confirm that data from aggregation-free samples were obtained. A small upturn is noted in the innermost SAXS data points from empty nanodiscs and TmCorA-loaded nanodiscs. We ascribe this to small systematic errors in the experimental setup.

6

PAPER III

CIRCULARIZED AND SOLUBILITY-ENHANCED MSPS FACILITATE SIMPLE AND HIGH YIELD PRODUCTION OF STABLE NANODISCS FOR STUDIES OF MEMBRANE PROTEINS IN SOLUTION



ORIGINAL ARTICLE

Circularized and solubility-enhanced MSPs facilitate simple and high yield production of stable nanodiscs for studies of membrane proteins in solution

Nicolai Tidemand Johansen[1]* | Frederik Grønbæk Tidemand[1]* | Tam T. T. N. Nguyen[2] | Kasper Dyrberg Rand[2] | Martin Cramer Pedersen[1] | Lise Arleth[1]

[1]Structural Biophysics, Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

[2]Protein Analysis Group, Department of Pharmacy, University of Copenhagen, Universitetsparken 2, 2100 Copenhagen, Denmark

Correspondence
Structural Biophysics, Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
Email: arleth@nbi.ku.dk

Funding information
University of Copenhagen, bioSYNergy; Novo Nordisk Foundation, Synergy; Lundbeck Foundation, Brainstruc

The nanodisc technology has diverse applications in studies of membrane protein structure and function in solution. Recently, circularized nanodiscs were developed, utilizing an enzymatic reaction to covalently link the N- and C-termini of the membrane scaffold protein. These nanodiscs were shown to be more homogeneous and stable than standard nanodiscs. We continue this development and aim for obtaining high yields of stable and monodisperse nanodiscs for structural studies of membrane proteins by solution small-angle scattering techniques. Based on the template MSP1E3D1, an optimized membrane scaffold protein (lsMSP1E3D1) is designed with a sortase recognition-motif and high abundance of solubility-enhancing negative charges. We find that with these modifications, high expression is maintained and the circularization reaction becomes highly efficient such that a high yield of the circularized membrane scaffold protein (csMSP1E3D1) and downstream circularized nanodiscs is obtained. The circularized protein and corresponding nan-

Abbreviations: 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine, POPC; 1,2-dimyristoyl-sn-glycero-3-phosphocholine, DMPC

* Equally contributing authors.

odiscs are characterized biophysically by circular dichroism spectroscopy, light scattering and small-angle X-ray scattering and compared to non-circularized samples. We find that lsMSP1E3D1 nanodiscs become stabilized as compared to the template MSP1E3D1 nanodiscs as an effect of the extra negative charges. Upon circularization, further improvement of csMSP1E3D1 nanodisc stability is achieved, while maintaining the structure. Finally, we show that a membrane protein can be efficiently incorporated in csMSP1E3D1 nanodiscs. Large scale production methods for circularized nanodiscs with improved thermal and temporal stability will facilitate easier access to the nanodisc technology and enable applications at physiologically relevant temperatures.

KEYWORDS
nanodisc, membrane protein, sortase, protein circularization, self-assembly, small-angle X-ray scattering, thermal and temporal stability

1 | INTRODUCTION

Nanodiscs are discoidal particles composed of a lipid bilayer surrounded by two amphiphatic and helical membrane scaffold proteins (MSP) [1, 2, 3]. Membrane proteins can be incorporated in nanodiscs, allowing for structural and functional studies [3] and their size may be modulated [4] to accommodate membrane proteins of different sizes. Compared to other popular membrane protein carrier systems, such as detergent micelles [5], vesicles [6] or bicelles[7], nanodiscs provide a lipid bilayer environment in a relatively small and monodisperse particle, which are desired properties for structural studies of membrane proteins in solution.

To our knowledge, all attempts at obtaining high resolution crystal structures of nanodiscs or membrane proteins inside nanodiscs have so far been unsuccessful. This reflects a dynamic nature of nanodiscs in general, making the system better suited for solution based structural studies using e.g. nuclear magnetic resonance (NMR) [2, 8], small-angle X-ray and neutron scattering (SAXS and SANS) [9, 10], or electron microscopy (EM) techniques including cryo-EM [11, 12]. Significant optimization of the membrane protein reconstitution is typically required prior to such experiments to obtain structurally homogenous and functionally relevant samples of the membrane proteins inside the nanodiscs. This implies that often tens to hundreds of milligrams of MSP are required as starting material for a project, making it essential to be able to produce MSP in large amounts with relative ease. Fortunately, the common uncircularized MSP constructs (MSP1D1, MSP1E3D1, etc.) express more than 100 mg per L of bacterial culture [1] and purification is straight-forward by immobilized metal-affinity chromatography (IMAC) [13].

The original MSP1D1 and MSP1D1E3 nanodiscs [1, 4] have been developed since the early 2000s and are widely used for functional and structural studies [3]. The development of smaller nanodiscs based on helix-truncated versions of the MSP1D1 later made the technology compatible with NMR [2]. More recently, an additional design improvement was introduced in terms of covalently circularized nanodiscs (cND's) [14], where an evolved variant of the

enzyme sortase A (eSrt) from *S. aureus* [15] was utilized to covalently link the N- and C-termini of the MSP. The main improvements reported on cNDs were a more narrow size distribution, *i.e.* reduced polydispersity, increased stability and better control of the number of membrane proteins incorporated per nanodisc compared to non-circularized nanodiscs [14]. Shortly after these pioneering studies, an improved protocol for circularization of a number of different MSPs was published, showing the need for large amounts of the detergent Triton X-100 present in the reaction to obtain a higher yield [16]. However, the final yield was still significantly lower than for uncircularized MSPs with an estimated 12.5 mg of cMSP1E3D1 per L culture [16]. Furthermore, the detergent used for the optimization, Triton X-100 is almost impossible to remove and can interfere with downstream applications such as UV absorption measurements or in extreme cases alter membrane protein structure and function in the nanodisc. An alternative approach was recently demonstrated, where split-intein mediated circularization of MSP was readily achieved inside the producing *E. coli* host [17]. While this study increased the speed of the purification process and eliminated the need for detergents, the yield was still modest.

The previously described circularized MSPs [14] are all based on the original MSPs [4] and do not contain specific mutations other than the introduction of the relevant eSrt-recognition motif. Given the fact that the original MSPs are simply truncations of the human apolipoprotein A-I (apoA-I) [1], these constructs are genetically optimized for specific interactions in the body, but not for the purpose of being highly expressing homogeneous membrane protein carriers. This leaves room for sequence optimization towards that goal, which we have exploited in the present work through the introduction of extra negatively charged and polar amino acids. This approach is inspired by the negative correlation found between aggregation propensity of proteins in solution and the introduction of negative charges on the protein surface [18, 19]. The obvious substitutions are Asn to Asp and Gln to Glu, where the length of the side chains are not altered but charges are introduced. Furthermore, substituting Thr to Ser also can improve solubility by increasing the polarity of the side chain. These substitutions were implemented in the work described in this article.

For standard nanodiscs, an above-optimum lipid:MSP ratio in the reconstitution mixture results in the formation of nanodiscs together with larger and not so well-defined lipid-MSP particles [4]. Furthermore, elevated temperatures facilitate the formation of other types of lipid-MSP particles than the nanodisc [20], for instance by fusion of nanodiscs. These observartions indicate that the energy landscape as a function of lipid and MSP content has several local minima. As such, the thermal stability of nanodiscs is given both by its resistance to structural unfolding, aggregation propensity and access to other structural states. For many applications, it is desirable to keep the nanodisc structure intact and it is expected that improved solubility and stability of nanodiscs will facilitate new interesting studies of membrane proteins. In functional studies, it is of interest to conduct experiments at elevated temperatures, commonly 37 °C, in order to be able to compare the results to physiological conditions and hence emphasize the biological relevance of the results. Furthermore, NMR experiments are also preferably conducted at high temperatures to improve the resolution of the data. As an example, the recently published high resolution structure of the MSP in a nanodisc was determined by NMR measured at 41 °C for several days [21].

All in all, MSP designs that can be produced in high yield and that provide homogenous nanodiscs that are highly stable over time (ideally several days) at elevated temperatures are desirable. The previously introduced circularization was a major step towards this goal as it improved the nanodisc stability [14, 16]. In the present work we provide a protocol for large scale preparation (75 mg per L of bacterial culture) of a novel MSP construct based on MSP1E3D1 [4], which is sequence-optimised with solubility-enhancing mutations and added a C-terminal eSrt-recognition motif for circularization by eSrt. We show that nanodiscs made from this new MSP construct (csNDs) are more stable at physiologically relevant temperatures than non-circularized nanodiscs (lsNDs), which are again more stable than standard MSP1E3D1 NDs, while the solution structures of csNDs and lsNDs are virtually indifferent as determined by SAXS. Furthermore, we show that a membrane protein can be effectively incorporated in csNDs, yielding a monodisperse

sample from which high quality SAXS data can be obtained.

2 | RESULTS AND DISCUSSION

2.1 | Production of circularized MSP

To develop more stable and soluble nanodiscs, a new construct was designed on the basis of MSP1E3D1, which has been shown to form nanodiscs of approximately 12 nm in diameter [4, 22]. This template was added a sortase recognition motif, LPETG, in the C-terminus to make it compatible with circularization along with an additional His_6-tag for purification following the work of Nasr and colleagues [14]. Furthermore, solubility-enhancing mutations were made at all possible positions, *i.e.* Asn to Asp, Gln to Glu and Thr to Ser, respectively [18, 19]. We name this construct His-tagged-linear-soluble-MSP1E3D1, in short His-lsMSP1E3D1. A sequence alignment of MSP1E3D1 and His-lsMSP1E3D1 is shown in Figure 1A with a graphical representation of important sequence elements described in the legend of Figure 1B. Figure 1B shows the different steps of the circularization of His-lsMSP1E3D1, *i.e.* digestion by tobacco etch virus (TEV) protease to remove the N-terminal His_6-tag and yield a free N-terminal Gly followed by circularization by eSrt and reverse IMAC purification.

MSPs have proven to be highly over-expressed in *E. coli*, with MSP1D1 as an example expressing at 250 mg protein per L culture in a fermentor [1], and in our experience around 150 mg per L culture in shaker flasks, which usually yields around 100 mg of pure protein. For His-lsMSP1E3D1, the expression level was similarly high, as indicated by the appearance of an intense band at the 35 kDa mark in SDS-PAGE analysis of the total *E. coli* lysate after 3 h induction of protein expression by isopropyl β-D-1-thiogalactopyranoside (IPTG) (Figure 1C). IMAC purification under denaturing conditions was highly effective, leaving only few impurities and yielding an estimated 100 mg His-lsMSP1E3D1 (Figure 1B, "(1)" and Figure 1D, lane 1) per L culture. Subsequent proteolytical digestion of the N-terminal with TEV protease yielded lsMSP1E3D1 (Figure 1B (2) and Figure 1D lane 2), the substrate for eSrt in the circularization reaction, with undetectable sample loss.

2.1.1 | Circularization efficiency

After IMAC purification, the concentration of lsMSP1E3D1 was approximately 166 μM, a factor of 16.6 higher than the concentration used by others for the circularization reaction in order to avoid oligomers by intermolecular circularization [14]. Circularized MSP can be visualized by SDS-PAGE, where it migrates slightly longer than the non-circularized MSP [14]. A series of small reactions with varying concentrations of lsMSP1E3D1 confirmed that a concentration of 10 μM was also optimal for our construct with a yield of 95% monomeric circular-soluble-MSP1E3D1 (csMSP1E3D1) (Figure 2A and Figure 2C). Whereas Nasr *et al.* used a MSP:eSrt ratio of 2:1 and quenched the reaction by addition of a covalent inhibitor, we tested the effect of the lsMSP1E3D1:eSrt ratio on the yield of the reaction, and found that a much lower ratio could be used. In fact, the reaction was only incomplete at a ratio of 100:1 at the time scale of the experiment, whereas a ratio of 10:1 yielded an equally complete reaction as using a ratio of 2:1 (Figure 2B and Figure 2D). By using a high lsMSP1E3D1:eSrt ratio, only a minor part of the sample can be complexed with eSrt and thus, the covalent inhibitor can be excluded.

For the large scale circularization reaction presented here (Figure 1D), eSrt was added to a ratio of 20:1 lsMSP1E3D1:eSrt at a lsMSP1E3D1 concentration of 10 μM. The formation of csMSP1E3D1 (Figure 1B "(3)" and Figure 1D, lane 3) was verified by SDS-PAGE, observing a small but significant gel shift compared to lsMSP1E3D1 (Figure 1B, "(2)" and Figure 1D, lane 2). The significance of the shift was confirmed by loading both lsMSP1E3D1 and csMSP1E3D1 in the

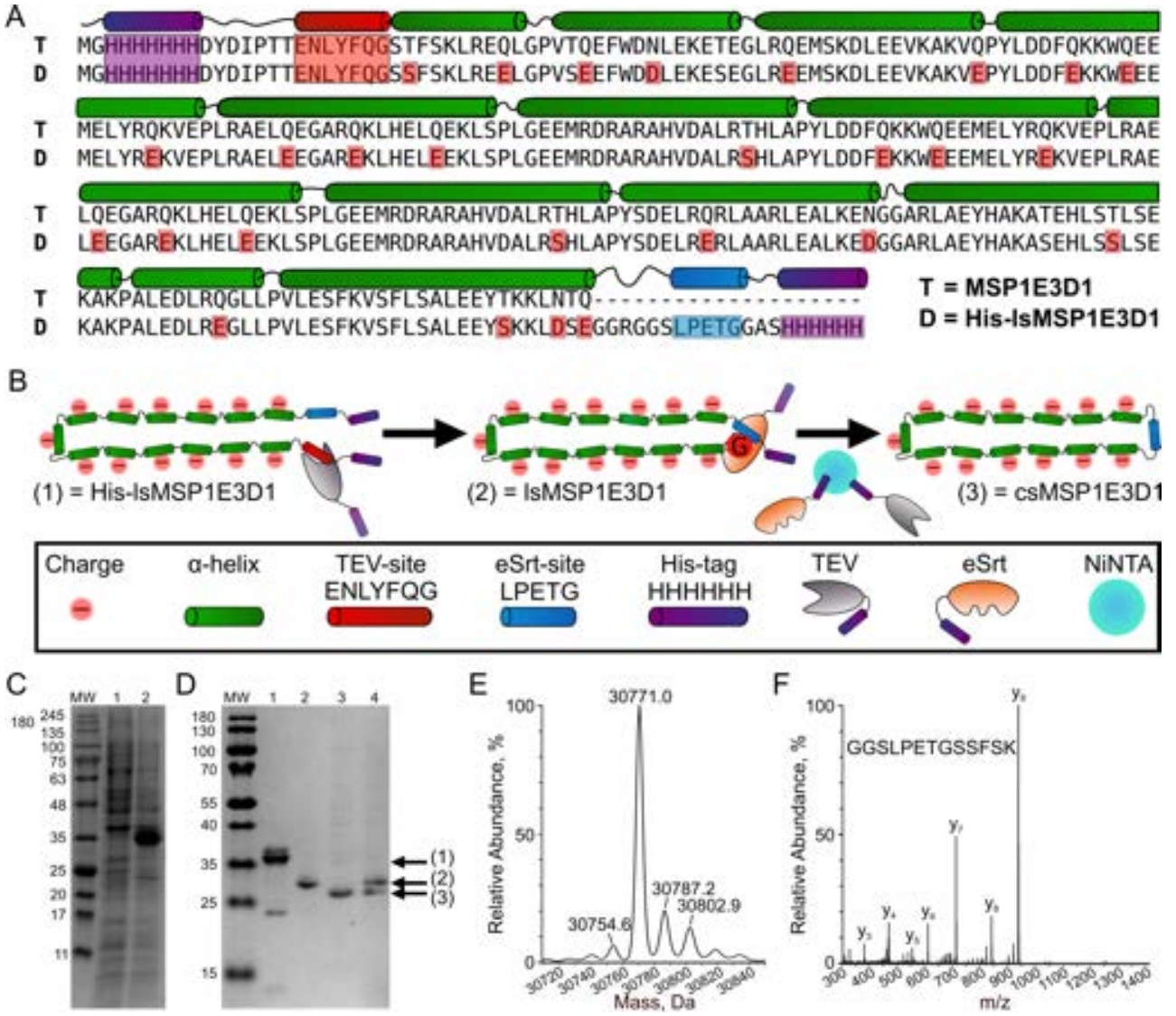


FIGURE 1 Overview of sequence optimized MSP and its purification. A: Primary sequence alignment of the template MSP1E3D1 (**T**) and the designed His-lsMSP1E3D1 (**D**), including graphical and highlighted sequence elements. B: Schematic illustrating the sequence elements of His-lsMSP1E3D1 and the circularization process. C: SDS-PAGE of the expression of His-lsMSP1E3D1. MW = molecular weight marker, 1 = whole cell extract before induction and 2 = whole cell extract after induction with IPTG. D: SDS-PAGE of different steps in the circularization process, indicated by arrows labeled according to the constructs shown in B. 1 = IMAC purified His-lsMSP1E3D1, 2 = lsMSP1E3D1, 3 = csMSP1E3D1, 4 = lsMSP1E3D1 and csMSP1E3D1 together. E: Intact mass analysis of csMSP1E3D1 by ESI-MS. The major peak corresponds to intact csMSP1E3D1 (2.2 ppm accuracy), while the two peaks with higher masses indicated correspond to oxidized products (See Figure S2 and S3). F: CID MS/MS fragmentation spectrum of a tryptic peptide corresponding to the peptide segment linking the N-and C-terminal of csMSP1E3D1. The precursor ion was $[M+2H]^{2+}$ (m/z 627.32).

same well, which yielded two clearly distinguishable bands (Figure 1D, lane 4). We note that the samples in Figure 1D lanes 2-4 were loaded in small amounts to better visualize the mobility shift of lsMSP1E3D1 compared to csMSP1E3D1, which made TEV protease, eSrt, oligomeric csMSP1E3D1 and eventual impurities undetectable.

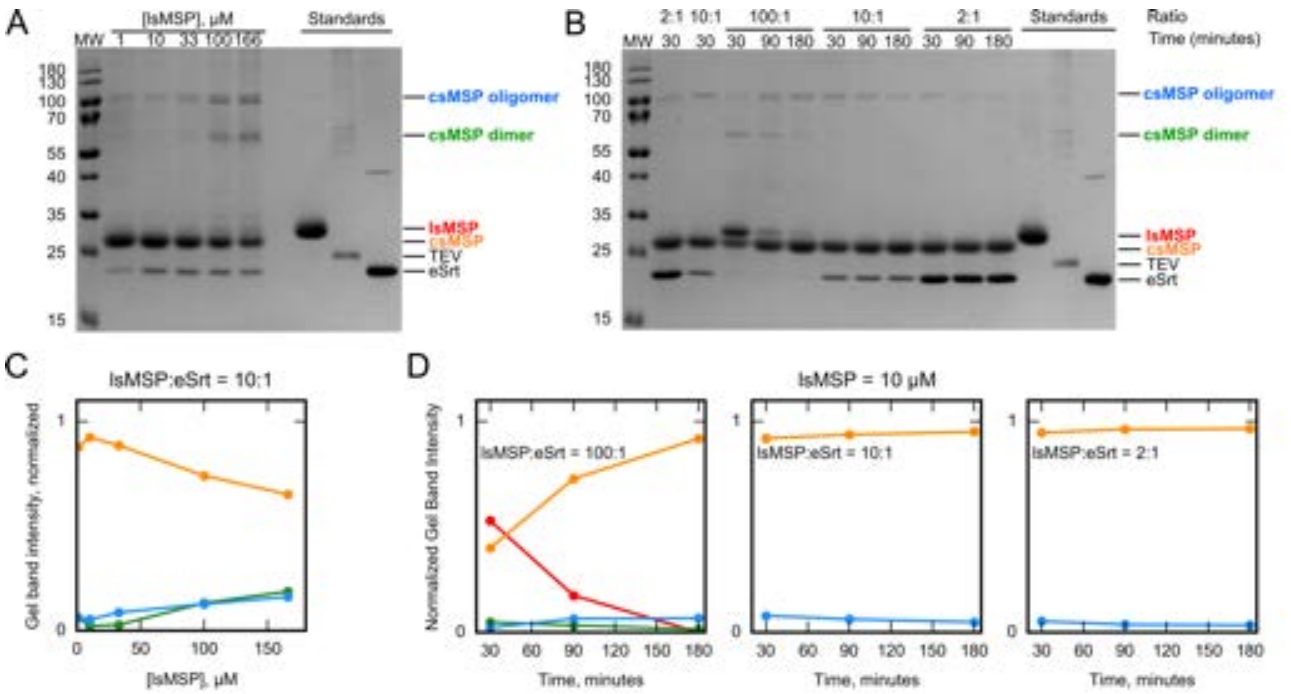


FIGURE 2 Circularization conditions. A: SDS-PAGE analysis of a 3 h long circularization reaction with varying concentration of lsMSP1E3D1 but with a fixed lsMSP1E3D1:eSrt ratio of 10:1. B: SDS-PAGE analysis of circularizartion reactions with three fixed lsMSP1E3D1:eSrt ratios of 100:1, 10:1, and 2:1 at a fixed lsMSP1E3D1 concentration of 10 μM. C: Quantification of gel band intensities in panel A. D: Quantification of gel band intensities in panel B.

2.1.2 | Purity and MS analysis

After circularization and reverse IMAC purification, csMSP1E3D1 was diluted 2.5 fold in milliQ water to a conductivity of approximately 8 mS/cm and purified by ion-exchange chromatography (IEC). The IEC purification was optimized by initial tests with linear and many-step gradients, respectively, which showed that two major populations could be separated (Figure S1). However, it was also evident that monomeric csMSP1E3D1 was present in both populations, and thus the best yield and highest concentration were achieved in a step-gradient eluting both populations in a single fraction (Figure 3A). The two fractions of highest purity and concentration (Figure 3B, lane 3 and lane 4) were pooled. The final monomeric csMSP1E3D1 was estimated to be 95% pure from SDS-PAGE, which is in general better than or equal to that obtained for other MSPs (data not shown). The sample had a concentration of 191 μM, which was in the desired range for downstream nanodisc preparation. The total yield was 75 mg csMSP1E3D1 per L of bacterial culture. The apparent loss of sample since the IMAC purification originates mainly from the choice of only the most concentrated fractions from the IEC purification.

To fully asses the integrity of the final product, a sample of pure csMSP1E3D1 was analyzed by mass spectrometry (MS). The experimentally determined molecular mass of the csMSP1E3D1 from LC-MS is 30771.00 (Figure 1E), which corresponds well to the theoretical mass of the circularized protein (30770.93 Da), with a mass deviation of 2.2 ppm. Furthermore, two modifications corresponding to oxidations of the protein were detected at approx. 20% and 10% relative abundance. Most importantly, however, we observed no hydrolyzed product, *i.e.* non-circularized MSP. To further verify the primary structure of the protein, peptide mapping analysis was performed by trypsin digestion followed by LC-MS/MS analysis. Tryptic peptides covering 92.4% of the sequence were identified (Figure S2). Furthermore,

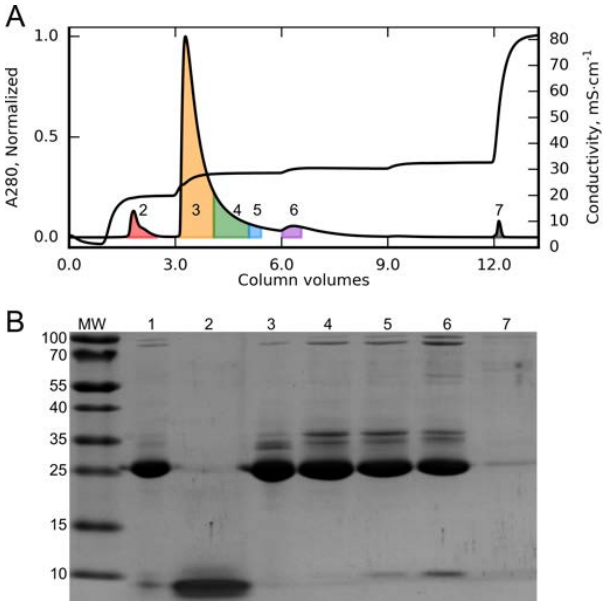


FIGURE 3 Purification and concentration on IEC. A: Chromatogram of the final purification on IEC. A Step gradient was applied from 0-0.4 M NaCl. Peaks are labeled with the lane numbers on the gel in B. B: SDS-PAGE gel of the corresponding fractions in A. MW is a ladder of protein standards and lane 1 is in the input sample for IEC purification.

a tryptic peptide corresponding to the peptide segment linking the N-and C-terminal of csMSP1E3D1 (GGSLPET-GSSFSK) was also identified (Figure 1F), confirming the successful ligation of the N-terminal residues (GSSFSK) to the C-terminal residues (GGSLPET). Finally, based on LC-MS/MS of the tryptic digest of csMSP1E3D1, the observed two major oxidative species of intact csMSP1E3D1 were conclusively assigned to oxidation of Met-59 and Met-125, which are placed in two repeats in the sequence of csMSP1E3D1 (Figure S3). Minor additional oxidation of the remaining Met residues was also observed, however, at very low abundance. Complete oxidation of two Met residues has been reported to slightly destabilize apoA-I lipoproteins [23]. However, given the relatively low abundance of Met oxidation in csMSP1E3D1, we are not concerned about such a destabilizing effect in downstream nanodiscs.

Summing up the purification, we devise a large scale production and purification that yields 75 mg per L bacterial culture of pure csMSP1E3D1 in a straight-forward manner. All intermediates of the constructs in the different purification steps are stable and highly soluble, making them compatible with enzymatic reactions at or above room temperature without the need for addition of detergent or other stabilizing agents. Furthermore, the solubility-enhancing mutations, *i.e.* extra negative charges, of the construct aid in the IEC step, where dilution with milliQ water to a relatively high conductivity of 8 mS/cm is enough to make the protein bind the column, thus excluding the need for tedious dialysis of very large volumes.

2.2 | Structural verification of nanodiscs by SAXS

2.2.1 | Model-free analysis

Solution SAXS is carried out either at in-house instruments or at international x-ray large-scale facilities, synchrotrons, as in the case of this study. Samples in solution are irradiated by an intense x-ray beam and the scattered intensity is collected as a function of scattering angle (or rather scattering momentum transfer, $q = 4\pi \sin(\theta)/\lambda$, where θ is the scattering angle, and λ is the wavelength of the incoming radiation). The method is a direct structural technique, which

provides structural information about the irradiated sample in length scales from 10 Å to 1000 Å. Specifically, SAXS probes the internal pairwise distances in the sample, and is thus capable at resolving changes in the overall structure of a sample or differences between structures. SAXS has been widely used for characterization and determination of nanodisc structures [9, 24, 22, 25] and related biophysical samples. In this context, we have recently shown that that the technique has resolution to distinguish between more subtle details of the nanodiscs [26].

Here, nanodiscs made from POPC and csMSP1E3D1 (csND) or lsMSP1E3D1 (lsND), respectively, were measured with SAXS to investigate their structure. The chromatograms obtained from the size-exclusion chromatography (SEC) purification prior to SAXS featured a single major peak for both samples (Figure 4A), giving a first indication that nanodisc reconstitution with these constructs was readily done by the usual protocol for other MSPs [13]. The identical retention volume and peak widths of the two samples furthermore indicate that they are overall similar in size. SAXS samples were picked from the middle and right-hand side of the main peak to eliminate contribution from larger particles present in the shoulder fraction on the left-hand side, as indicated by the colored areas under the chromatograms (Figure 4A).

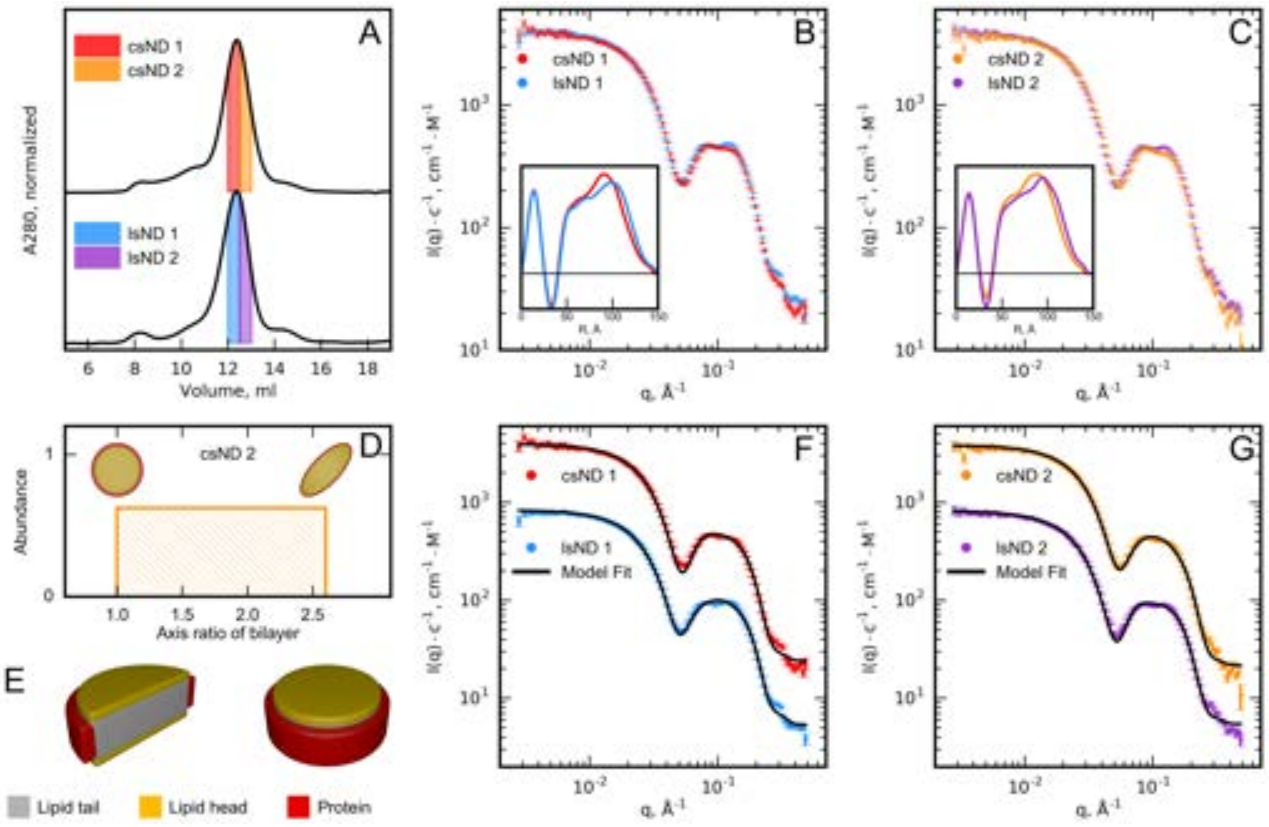


FIGURE 4 Structural verification and modeling by SAXS. A: SEC profiles of nanodiscs made from csMSP1E3D1 (csND) or lsMSP1E3D1 (lsND), respectively. The fractions collected for SAXS samples are indicated by the colored areas. B: SAXS data recorded on the fractions collected from the left-hand side of the peaks as indicated in A. The insert shows $p(r)$-distributions determined from the SAXS data. C: Same as B, but for fractions collected on the right-hand side of the peaks. D: A representative model distribution is shown for the csND 2 SAXS data set in C, together with top views of the models with the smallest (1) and largest (2.6) bilayer axis ratios. E: Graphical representation of the fitted csND model with an axis ratio of 1. The model is shown both in a full and in a cross-section representation to visualize the internal organization of the lipid bilayer in the nanodisc. F and G: Models fits to the data shown in B and C, respectively. The lsND data are rescaled by a factor 1/5 for clarity.

Scattering data from the csND and lsND samples from the two fractions collected from the SEC were normalized to sample concentration, c, and plotted on absolute scale (Figure 4B and Figure 4C). At first glance, the scattering data from the presented samples of csND and lsND differed only slightly. The same general trends and features are visible in both datasets and for both fractions indicating very similar overall structure of the samples. The $p(r)$-distributions refined from the data also showed profiles with only small differences (Figure 4B insert). The parameters inferred from this $p(r)$-distribution refinement are given in Table 1. The characteristic sizes of the samples are similar to those reported for MSP1E3D1 nanodiscs [22] and in agreement with the overall understanding and expectation of the nanodisc structure. Interestingly, we found that the R_g D_{max} decrease as a function of retention volume. This phenomenon has been described for both MSP1D1 nanodiscs [26] and MSP1E3D1 nanodiscs [25]. Although symmetric peaks are obtained from SEC purifications, nanodisc samples collected at low retention volumes appear to have higher lipid:MSP stoichiometry [26]. Consequently, due to the finite resolution of the SEC coulumn, one collected fraction of nanodisc sample contains a small distribution of lipid stoichiometries.

When comparing the $p(r)$-distributions for the csND to the $p(r)$-distributions for the lsND (Figure 4B and Figure 4C), we note that, despite the great similarity, there are for both fractions slightly more pair-distances at higher R-values for the lsNDs. Furthermore, D_{max} and R_g values of the lsNDs (see Table 1) are systematically slightly larger than those of csNDs. This shows that the small differences between the circularised and uncircularised nanodiscs go in the direction of slightly larger non-circularised discs.

TABLE 1 Parameters determined from the IFT-analysis and corresponding to the $p(r)$-distributions shown in Figure 4B insert, Figure 4C insert and Figure 6B insert.

Parameter	csND 1	csND 2	lsND 1	lsND 2	CorA csND
Largest distance in sample, D_{max} / Å	144	140	148	143	150
Radius of gyration, R_g / Å	58.2	56.2	60.5	58.9	55.6

In a recent study, cNDs were reported to be 1-2 nm wider in diameter than corresponding non-circularized nanodiscs as evaluated by SEC and EM [17]. In contrast, we find in the present study that the retention volumes of the SEC-peaks are identical, and that both R_g and D_{max} values for csNDs are in fact slightly smaller than for lsNDs (Table 1). In another EM study, circularization did not appear to influence the mean size of the formed nanodiscs [14], and the discrepancy between all of these data could be due to differences between the specific MSP constructs.

2.2.2 | Modeling

Structural models of the empty nanodiscs were refined from the SAXS data of csND and lsND, respectively. As described in the Methods section, each model includes a uniform distribution of nanodiscs with different axial ratios of the bilayer (Figure 4D) and hence a distribution of different lipid to MSP ratios in the individual nanodiscs, but otherwise the same description of the MSP, lipid core and lipid head groups, as illustrated for csND 2 with an axial ratio of 1 (Figure 4E). The model fits to the csND and lsND data are plotted on top of the respective data in Figure 4F and Figure 4G. The models reproduce the overall behavior of the data, but with minor, systematic discrepancies present. In particular, the models do not entirely reproduce the structure of the two shoulders around $q = 0.08$ Å^{-1} and $q = 0.15$ Å^{-1}. We suspect that these shortcomings are due to the simple description of the bilayer structure.

The parameters describing the models refined from the individual data sets are listed in Table 2. In general, the values are in line with expectations and agree with the current understanding of the nanodisc systems. The area per

headgroup for POPC is refined to values between 65.2 Å^2 and 60.9 Å^2, which is in good accordance with the 62.7 Å^2 reported for POPC in vesicles at 20 °C [27]. The number of lipids per nanodisc is comparable to values obtained in a previous study of nanodiscs made from POPC and MSP1E3D1 [22], while the values for the thicknesses of the MSP belt agree well with the published NMR structure of a similar MSP [21], and our previous results [9]. Interestingly, the number of lipids is lower in csND compared to lsND. This observation is in qualitative agreement with the recent work by Yusuf *et al.* [16], who found MSP1E3D1 nanodiscs to contain 241 DMPC lipids, and circularized MSP1E3D1 nanodiscs to contain only 157 DMPC lipids. It is also well in line with the smaller sizes found in the model free-analysis (Table 1). These results indicate that circularization causes a tighter constraint on the nanodisc, allowing fewer lipids inside. In a recent study, we found that the degree of lipid-loading of the final nanodiscs is mainly controlled by how many lipids that can be packed into the MSP-lipid-detergent pre-aggregates in the early phase of the nanodisc self-assembly process [26]. Consequently, it is not so surprising that the pre-aggregates with the covalently circularized MSPs that are presumably less flexible can generally incorporate fewer lipids than those of the more flexible non-circularised MSPs. Finally, the specific densities of the MSPs were refined to 1.46 g/cm^3 for both samples. This value is higher than the 1.35 g/cm^3 reported for globular proteins [28] and implies that the MSPs are slightly more dense than average globular proteins.

TABLE 2 Structural parameters refined from the SAXS data presented in Figure 4F and Figure 4G. The top-most parameters are refined directly from the data, whereas the bottom-most are calculated from these.

Refined parameters	**csND 1**	**csND 2**	**lsND 1**	**lsND 2**
Area per PC headgroup / Å^2	62.5 ± 1.0	65.2 ± 1.3	60.9 ± 0.8	62.4 ± 0.9
Largest axis ratio of bilayer in distribution	2.2 ± 0.2	2.6 ± 0.3	2.5 ± 0.3	2.1 ± 0.2
Thickness of protein belt / Å	8.0 ± 0.2	8.6 ± 0.2	7.5 ± 0.2	8.3 ± 0.2
Roughness / Å	5.0 ± 0.2	5.2 ± 0.2	4.6 ± 0.2	5.0 ± 0.2
Volume of POPC / Å^3	1270 ± 5.1	1280 ± 6.1	1280 ± 4.5	1270 ± 4.4
Volume of belt protein / Å^3	34200 ± 490	35100 ± 500	33900 ± 500	36200 ± 470
Derived parameters				
χ^2	11.0	10.7	18.8	12.7
Thickness of bilayer / Å	40.8	39.3	41.9	40.7
Thickness of hydrophobic bilayer / Å	30.3	29.3	31.1	30.3
Density of belt protein / g/cm^3	1.49	1.46	1.56	1.46
Number of lipids in biggest nanodisc	237	206	278	248
Number of lipids in smallest nanodisc	203	166	226	215

Overall, we find both csND and lsND to be structurally highly similar to each other and to nanodiscs made from the template MSP1E3D1 [22]. This high similarity between csNDs and lsNDs is in contrast to the previous finding that circularization of the MSP should enhance the monodispersity of nanodiscs as evaluated by EM [14]. In fact, we find that the formed nanodiscs are already rather monodisperse before circularization. A sample with a high degree of polydispersity would have given rise to scattering intensity profiles with fewer features than the ones shown in Figure 4B and Figure 4C.

Particles are measured under standard solution conditions in SAXS. This is in contrast to EM where samples are measured upon staining, fixation on grids and cryo-cooling, which can give rise to artifacts. In addition, SAXS probes a large ensemble (on the order of 10^{13} particles) compared to the few hundreds to thousands in EM. From our studies of the stability- and solubility enhanced MSP, we conclude that an increased structural homogeneity of the resulting nanodiscs does not appear to be a main advantage of circularization

2.3 | Stability enhancement by negative charges and circularization

2.3.1 | Secondary structure quantification

Association of lipids to MSP or similar types of peptides have been shown to induce changes in secondary structure, in most cases explained by stabilization of helical structure [1, 29, 14]. To probe secondary structure stability, csMSP1E3D1, lsMSP1E3D1, and nanodiscs based on these MSPs and POPC lipids were measured with circular dichroism spectroscopy (CD) and compared to the template MSP1E3D1 and nanodiscs based on MSP1E3D1 and POPC.

The obtained CD spectra confirm that all samples are helical in solution as given by the presence of the two distinct minima at 208 and 222 nm, respectively (Figure 5A). Comparing csMSP1E3D1 to lsMSP1E3D1 and csND to lsND, respectively, does not reveal any significant changes to the secondary structure upon circularization. However, in all instances, it is clear that MSP1E3D1-based samples have a smaller CD signal, *i.e.* lower helical content, than the corresponding csMSP1E3D1 and lsMSP1E3D1-based samples.

When comparing the lipid-free MSP samples (csMSP1E3D1, lsMSP1E3D1 and MSP1E3D1) to the corresponding nanodisc samples, we observe an increased magnitude of the CD signals at 222 nm for the nanodiscs (Figure 5A). This indicates a general stabilization of the helical structure by the lipids. For csND and lsND, the ratio between the signals at 208 and 222 nm is different from those of the corresponding lipid-free MSPs, suggesting a conformational change in the secondary structure upon lipid association. The 222nm/208nm ratio can indicate whether α-helices exist as free (222nm/208nm $\approx$ 0.9) or in a coiled-coil conformation (222nm/208nm $\approx$ 1.1) [30]. Thus, the observed ratio close to 0.9 for csMSP1E3D1 and lsMSP1E3D1 indicates that these proteins contain mainly free α-helices, which is not the case for the MSP1E3D1 samples, where the CD signals for both lipid-free and nanodisc samples are similar in shape to those of csND and lsND (Figure 5A). The negative charges on the solubility-enhanced MSPs are likely repelling each other, which is supported by SEC data that shows that lsMSP1E3D1 and csMSP1E3D1 have less tendency to form dimers compared to MSP1E3D1 (data not shown).

In summary, the CD analysis reveals that the secondary structures of the new MSPs and NDs are perturbed by the introduction of negative charges, rendering the samples slightly more helical in solution compared to the template and with the free MSPs having a slightly altered conformation in solution compared to MSP1E3D1. On the other hand, circularization does not perturb secondary structure.

2.3.2 | Thermal denaturation

Next, thermal unfolding was monitored by CD spectroscopy at 220 nm. The obtained curves show that all lipid-free MSPs unfold with a long transition spanning approximately 30 °C (Figure 5B), whereas the nanodisc samples unfold only partially towards the maximum temperature of the experiment (Figure 5C). In the refolding of the lipid-free MSPs, the curves almost overlay those of the unfolding (Figure 5B), whereas for the nanodisc samples, the refolding curves are very different from the unfolding curves (Figure 5C). In fact, they resemble the unfolding curves of the lipid-free samples better. This behavior suggests that a fraction of the nanodiscs irreversibly denatures upon heating to very

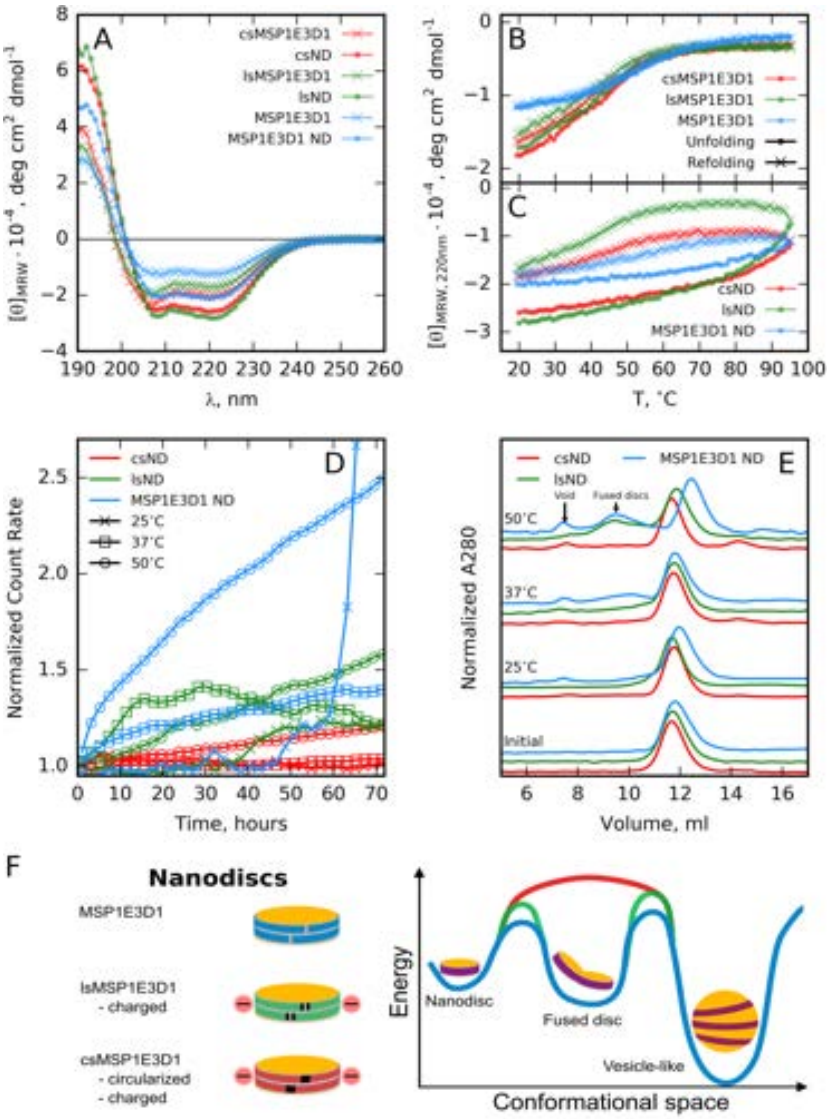


FIGURE 5 MSP and nanodisc stability. A: Near-UV CD spectroscopy of lipid-free and nanodisc samples. B: Thermal unfolding and refolding of the lipid-free samples in A monitored by the CD signal at 220 nm. C: Same as in B, but for nanodiscs samples. D: SLS measurements at different temperatures and as a function of time. E: SEC chromatograms (280 nm) measured before and after the SLS measurements shown in D. Arrows indicate the retention volume for aggregates (void) and intermediary particles (fused discs), respectively. F: Proposed energy landscape of nanodiscs made from the three investigated MSPs. Energy barriers between different energy minima are color coded according to the nanodiscs sketches shown on the left.

high temperatures, and that some of the MSP is separated from lipids and thus refolds as lipid-free MSP.

Given the incomplete unfolding curves for the nanodisc samples, it was not possible to extract thermodynamic parameters and make a quantitative evaluation of the effect of circularization. However, the magnitude of the signal as a function of temperature is markedly lower, and with a more calm slope, for csND than for lsND (Figure 5C). Qualitatively, this observation reflects a slightly higher thermal stability of csND. The slope and the magnitude of the signal change are even more subtle for the MSP1E3D1 ND than for csND, and after refolding, the signal even returns to a value close to the initial value before unfolding (Figure 5C). Given the the large differences in primary and secondary structure of csND/lsND and MSP1E3D1 ND, a quantitative comparison of the CD unfolding curves from

these samples is not feasible.

The melting temperatures of the lipid-free MSPs can be estimated from the inflection points of the unfolding curves. For csMSP1E3D1 and lsMSP1E3D1, the melting temperatures are determined to be 41.6 °C and 44.4 °C, respectively, and 52.5 °C for MSP1E3D1. Thus, it appears that the extra helical structure gained from the introduction of negative charges is less stable than the original helical structure in the template MSP1E3D1. Furthermore, the circularization of lsMSP1E3D1 to obtain csMSP1E3D1 appear to slightly destabilize the secondary structure. Comparing the obtained values to those reported for MSP1D1 and circularized MSP1D1 (53.8 °C and 65.4 °C, respectively) [14], it is evident that the melting temperature of the template MSP1E3D1 is very similar to that reported for MSP1D1. However, csMSP1E3D1 and lsMSP1E3D1 behave very differently from circularized MSP1D1, likely because the secondary is perturbed by the extra negative charges.

2.3.3 | Temporal stability and aggregation propensity

Although not observed from the CD measurements on csMSP1E3D1 (Figure 5B), improved thermal stability is a common property of circularized proteins or peptides [31]. Coupled effects, such as aggregation propensity, could also be lowered due to the smaller likelihood of the protein unfolding locally.

With the CD data difficult to interpret, the stability of csND and lsND were also evaluated as a function of time at different temperatures by SLS and compared to MSP1E3D1 nanodiscs (Figure 5D). All nanodiscs were measured at three different temperatures to probe the effect on stability arising from circularization and addition of negative charges, corresponding to the temperature of a typical room temperature experiment (25°C), physiological temperature (37 °C), and at high temperature where accelerated aggregation was expected (50 °C). SEC chromatograms were measured for the samples immediately after purification as well as after the SLS experiments to investigate the integrity of the measured samples (Figure 5E).

In contrast to lsND, the SLS signals show that csND remains fully intact at both 25 °C and 37 °C throughout the 72 h experiment (Figure 5D). The corresponding SEC chromatograms confirm that the csND samples measured at the two different temperatures are structurally similar and elute at the same retention volume (Figure 5E). At 50 °C, the SLS signal for csND increases linearly, but with a significantly smaller slope than that of lsND, indicating increased thermal tolerance of the circularized nanodisc. Interestingly, the corresponding SEC analysis of csND shows only a small peak in the column void volume beside the main peak at 12 mL, whereas a distinct peak at 9.5 ml is observed for lsND (Figure 5E). In fact, a tiny shoulder at 9.5 ml was also observed next to the main peak for the lsND after the experiment at 37 °C. This suggests that the circularized MSP indeed stabilizes the nanodisc structure and prevents formation of other larger intermediary types of lipid-MSP particles. A similar effect is described by Yusuf *et al.* [16] for nanodiscs made from non-circularized and circularized MSP1D1, respectively, and DMPC lipids.

MSP1E3D1 nanodiscs were measured for probing the stabilizing effect of the solubility-enhancing mutations in csND and lsND. In the SLS experiment, the MSP1E3D1 appears to behave similarly to lsND at 25 °C and 37 °C for around 60 h, whereas at 50 °C, the signal increases much faster and to a higher end-value (Figure 5D). After the 60 h at 25 °C, some large particles or aggregates are likely formed in the MSP1E3D1 ND sample, causing the signal to accelerate. The SEC analysis confirms this formation of aggregates at 25 °C, while it shows the formation of both aggregates (void) and other larger intermediary particles (fused discs) at 37 °C and 50 °C (Figure 5E).

After the 50 °C measurement of MSP1E3D1 ND, the main peak in the SEC chromatogram is also seen to move from the initial 12 ml to around 13 ml (Figure 5E). This observation could be explained by lipid depletion of the initial nanodisc population by transfer of lipids into larger particles having a higher lipid-MSP ratio. Interestingly, the peak appearing at approximately 9.5 mL, similar to the one observed for lsND at 50 °C, indicates the formation of

intermediary larger lipid-MSP particles already at 37 °C for MSP1E3D1 ND. This suggests that the solubility enhancing mutations in csND and lsND are favorable in terms of physical stability in the sense that the formed nanodiscs appear to be less prone to collide and form other types of particles, for instance fused nanodiscs with four MSPs surrounding a larger lipid bilayer patch. This structure would be impossible to obtain for csND, since the termini of csMSP1E3D1 are covalently linked. However, the fairly broad peak observed at 9.5 ml may also represent other types of less well-defined particles.

2.3.4 | The nanodisc energy landscape

Based on the observations described above, we propose a simple model for the energy landscape of the nanodiscs made from the different MSPs (Figure 5F). Here, the intermediary larger MSP-lipid particles are represented as fused discs, and even larger particles as vesicle-like structures. We note that these are only labels as the structures have not been characterized experimentally. As the initial reconstitution of nanodiscs by detergent depletion yields mainly monodisperse nanodiscs, it implies that the nanodisc structure is the kinetically most stable structure. Only upon heating the samples, larger particles are formed, which suggest that the larger lipid-MSP particles are thermodynamically more stable than the nanodisc. The thermodynamic stability is then governed by favorable lipid-lipid interactions, gaining higher stability the closer the formed structures become to resemble the vesicle-like structure.

The proposed energy landscape (Figure 5F) illustrates that the energy barrier for disc-fusion is higher for the solubility-enhanced lsND compared to MSP1E3D1 ND, whereas the fused disc intermediate can not be visited by the csND. The csND must take another, more direct, path to the aggregated vesicle-like structure. This direct path has a higher activation energy than that of lsND as inferred from the higher resistance to aggregation observed in the SLS experiment. The energy barrier for fused lsNDs to the vesicle-like structure has the same relative height to that of MSP1E3D1 ND as the barrier from nanodisc to fused discs.

All in all, the combined CD, SLS and SEC results show that the csND is a stable particle suited for experiments at physiological temperatures. Furthermore, the reduced aggregation propensity is a highly desired property for future SAXS (and SANS) studies of membrane proteins incorporated in nanodiscs.

2.4 | Incorporation of the membrane protein CorA in circularized nanodiscs

In the reconstitution of membrane proteins inside nanodiscs, the membrane protein is usually incubated with cholate and lipids (mixed micelles) before addition of MSP. The MSP will thus have to wrap around the membrane protein and form a mixed micelle aggregate. This process is not well-understood and depends on a number of parameters such as protein size, hydrophobicity and presence of domains of the membrane protein protruding from the membrane. Thus, it was not clear whether or not csMSP1E3D1 was compatible with effective incorporation of a large membrane protein due to the more restricted MSP after circularization, although previous work with circular nanodiscs had shown promising results for a similar system [14].

Here, the 200 kDa pentameric membrane protein CorA from *Thermogata maritima* was used as model for testing membrane protein incorporation in csND. The SEC purification of CorA-loaded csND reveals a sample containing several sub-populations, with a minor peak in the void volume trailing into the major peak at approximately 11 mL (Figure 6A). This chromatogram is very similar to chromatograms obtained after incorporation of CorA in MSP1E3D1 nanodiscs [32], suggesting that nanodisc reconstitution of CorA is not particularly affected by the more restricted csMSP1E3D1.

The sample was collected from the right-hand side of the major peak to avoid large particles from the shoulder

of the peak and measured with SAXS (Figure 6B). As described for the unloaded nanodiscs, the $p(r)$-distribution was determined from the data (Figure 6B insert), yielding the parameters listed in Table 1. The scattering data show no signs of aggregation or presence of contaminating large particles, as seen from the flat Guinier-region in the SAXS data and the smooth descent to $D_{\max}$ in the $p(r)$-distribution.

The presence of CorA and csMSP1E3D1 in the sample was verified by SDS-PAGE (Figure 6C). Based on gel quantification of CorA- and csMSP1E3D1-standards as well as the sample, the CorA-monomer:csMSP1E3D1 ratio was found to be close to 2.5:1. This is in agreement with one CorA pentamer per csND. That CorA exists as a single pentamer is further supported by the apparent $D_{\max}$ of 150 Å, as a nanodisc particle containing *e.g.* a dimeric CorA pentamer would have a $D_{\max}$ of approximately 200 Å independent of the orientation of the CorA dimer (Figure 6D). In all, these results indicates that CorA is effectively incorporated as a single pentamer into the csND.

We observe that the shape of the $p(r)$-distribution for the CorA sample (Figure 6B insert) has smaller oscillations than for the empty nanodiscs (Figure 4B insert and Figure 4C insert). Furthermore, the SAXS data from the CorA sample (Figure 6B) have inferior counting statistics compared to the csND sample (Figure 4B and Figure 4C). In a model refinement situation, one would likely have to supplement the presented SAXS data with complementary SANS, ideally contrast-variation SANS utilizing e.g. match-out deuterated nanodiscs [33], in order to confidently refine a solution structure of the CorA samples. This emphasizes the need for reliable production of large amounts of sample, including deuterated MSP, due to the relatively larger sample requirements of SANS.

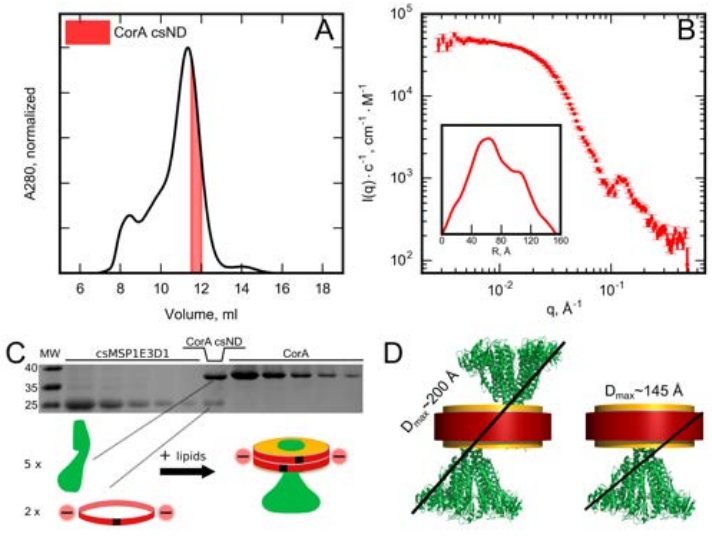


FIGURE 6 Incorporation of CorA in csND. A: SEC purification of CorA-loaded csND. The colored area indicates the fraction collected for structural analysis by SAXS. B: SAXS data on the sample purified in A with the refined $p(r)$-distribution shown in the insert. C: SDS-PAGE analysis of CorA- and csMSP1E3D1-standards along with the CorA-loaded csND sample. The illustration under the gel shows the stoichiometry of the CorA-loaded csND. D: Structural representations of csNDs loaded with dimeric or monomeric CorA and with the approximate $D_{\max}$ values indicated. The CorA structures are based on a crystal structure (Protein Data Bank ID: **4eeb**), and the csND structure on the refined model shown in Figure 4E.

3 | CONCLUSION

Membrane protein carrier systems that are compatible with maintaining the relevant structure and function of the studied membrane proteins while at the same time being sufficiently abundant to allow for sample optimization are highly desirable. We have devised an efficient method for large scale production of sortase-mediated circularized MSP yielding 75 mg of pure final product per L of culture. A central step in the protocol was the design of a solubility-enhanced MSP through incorporation of extra negatively charged amino acids. This effectively eliminated the need for detergents in any of the applied purification and circularization steps.

Although the introduced charges slightly destabilized the secondary structure of the free MSP at elevated temperatures, we found that nanodiscs made with these cs- and lsMSP1E3D1 constructs were much more stable over time at physiological temperatures than nanodiscs formed with the original template MSP1E3D1 construct. This shows that the solubility-enhancing mutations also improved the physical stability of the nanodiscs. Not only is this an obvious desired property for structural and functional studies of membrane protein at physiological temperatures, it also indicates that nanodisc stability, or propensity for aggregation, is governed not only by the flexibility of the MSP, but also from its propensity of collision with other nanodiscs. Circularization further enhanced nanodisc stability. This was most clearly seen at 50 °C, where the circularized nanodiscs did not produce the larger intermediary lipid-MSP particles over time that we observed for the non-circularized nanodiscs.

Other studies based on the original MSP1E3D1 template report on a gain of structural homogeneity for circularized nanodiscs [14]. As evaluated from solution SAXS data, such a change in homogeneity was not observed for nanodiscs produced with our constructs. Consequently, modeling of the SAXS data yielded very similar structural models for circularized and non-circularized nanodiscs, respectively. However, the number of lipids was found to be systematically lower in the circularized nanodisc. Finally, we showed that the membrane protein CorA could be effectively incorporated into the circularized nanodiscs and that high quality SAXS data could be obtained. Such a membrane protein in nanodisc system is highly dynamic and contains multiple components, which means that more elaborate structural studies are needed in order to dissect the complex in detail, for instance by contrast-variation SANS, EM and NMR. The solubility- and stability-enhanced csMSP1E3D1 will aid in obtaining sufficiently large amounts of good samples for such purposes.

4 | METHODS

4.1 | Materials

The plasmid harboring the gene for MSP1E3D1 was kindly provided by Professor Steven Sligar (Uni. Illinois). For His-lsMSP1E3D1 and eSrt, codon-optimized genes were custom synthesized and inserted into the pET28a(+) vector by Genscript (Piscataway, NJ, USA). The protein sequences were:

His-lsMSP1E3D1

MGHHHHHHHD YDIPTTENLY FQGSSFSKLR EELGPVSEEF WDDLEKESEG LREEMSKDLE
EVKAKVEPYL DDFEKKWEEE MELYREKVEP LRAELEEGAR EKLHELEEKL SPLGEEMRDR
ARAHVDALRS HLAPYLDDFE KKWEEEMELY REKVEPLRAE LEEGAREKLH ELEEKLSPLG
EEMRDRARAH VDALRSHLAP YSDELRERLA ARLEALKEDG GARLAEYHAK ASEHLSSLSE
KAKPALEDLR EGLLPVLESF KVSFLSALEE YSKKLDSEGG RGGSLPETGG ASHHHHHH

eSrt

MASGGSGGSG GSGGSGGSQA KPQIPKDKSK VAGYIEIPDA DIKEPVYPGP ATREQLNRGV
SFAEENESLD DQNISIAGHT FIDRPNYQFT NLKAAKKGSM VYFKVGNETR KYKMTSIRNV
KPTAVEVLDE QKGKDKQLTL ITCDDYNEET GVWETRKIFV ATEVKLEHHH HHH

Unless specifically stated, buffer components were from Sigma-Aldrich (Darmstadt, Germany). All buffers were prepared with Milli-Q water (Merck/Millipore, Darmstadt, Germany). TB medium components, *i.e.* yeast extract and peptone were from VWR (Radnor, PA, USA) and glycerol from Sigma-Aldrich (Darmstadt, Germany).

4.2 | Expression and purification of MSPs

Plasmids were transformed into *E. coli* BL21 Star (Merck/Novagen, Darmstadt, Germany). For protein expression, the cells were grown at 37 °C in TB medium and induced with 1 mM IPTG (Fisher Scientific, Loughborough, UK) at an optical density at 600 nm (OD600) of 0.8. After 3 h, cells were pelleted by centrifugation and subsequently resuspended and shaken vigorously in binding buffer (50 mM TrisHCl pH 8, 300 mM NaCl, 20 mM imidazole) containing 6 M guanidinium HCl. Insoluble material was pelleted by centrifugation at 30000g for 30 min, and the supernatant was added equilibrated NiNTA resin (10 ml resin for an expected amount of 100 mg MSP). After incubation for 1 h, the resin was loaded onto a disposable column and washed with 1 column volume (CV) of denaturing binding buffer, 5 CV of wash buffer (50 mM TrisHCl pH 8, 300 mM NaCl, 40 mM imidazole) and finally eluted with elution buffer (50 mM TrisHCl pH 8, 300 mM imidazole). Digestion with TEV protease was carried out (1:100 TEV protease:MSP) at room temperature (RT) for 3 h in a 10 kDa molecular weight cut off dialysis bag submerged in a 100 fold volume of TEV buffer (20 mM TrisHCl pH 8, 100 mM NaCl, 1 mM EDTA, 1 mM DTT). At this point, some of the lsMSP1E3D1 was concentrated and stored at -80 °C for making non-circularized reference samples. For the circularization reaction, the remaining sample was transferred to an Erlenmeyer flask and diluted buffer, salts, and milliQ water to a final concentration of 10 μM lsMSP1E3D1 in circularization buffer (30 mM TrisHCl, pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 1 mM DTT). eSrt (see purification below) was added to obtain a desired ratio of lsMSP1E3D1:eSrt (20:1 used in the final production) and the mixture was incubated at 37 °C. After 3 h, NiNTA resin (Qiagen, Hilden, Germany) equilibrated in circularization buffer was added to the flask, and the reaction was carried out for another 30 minutes. The mixture was loaded onto a disposable column, and the flow through was collected. 1 CV of circularization buffer was added to wash out the remaining sample from the column. The flow through was diluted with milliQ water to a conductivity of 8 mS/cm (approximately 2.5 fold dilution) and loaded onto a 6 ml ReSourceQ column (GE Healthcare, Chicago, IL, USA) connected to an Äkta Pure system (GE Healthcare, Chicago, IL, USA). Impurities and csMSP1E3D1 were eluted separately in a gradient from buffer A (20 mM TrisHCl pH 8) to buffer B (20 mM TrisHCl pH 8, 1 M NaCl). The sample was dialysed against gel filtration buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl) over night (ON) at 4°C, then flash frozen and stored at -80°C.

4.3 | Expression and purification of eSrt

A codon-optimized version of eSrt based on the sequence presented in the work by Chen *et al.* [15] was synthesized and cloned into the pET28a(+) vector by Genscript. The plasmid was transformed into *E. coli* BL21 Star cells and grown at 37 °C in TB media until expression was induced with 1 mM IPTG at OD600 of 0.6 after which cells were grown at 20 °C ON. All subsequent steps were performed at 4 °C. Cells were harvested by centrifugation and the pellet was dissolved in binding buffer and passed twice through a cell disrupter at 30 kPSI. Cellular debris was pelleted by

centrifugation at 20000g for 20 min before the supernatant was incubated with equilibrated NiNTA resin for 45 min. The preparation was loaded onto a disposable column and the resin was washed with 3 CV of wash buffer and eluted with elution buffer. Fractions containing eSrt were pooled and dialyzed against gel filtration buffer (containing 1 mM DTT) to remove imidazole.

4.4 | Expression and purification of CorA

The plasmid encoding *Thermogata Maritima* CorA was kindly provided by professor Mikaela Rapp (University of Stockholm, Sweden). CorA was expressed and purified as described elsewhere [34]. The N-terminal His_6 was left intact to allow IMAC purification of CorA-loaded nanodiscs. The purified protein was flash frozen in liquid nitrogen and stored at -80 °C until use.

4.5 | Nanodisc preparation

Nanodiscs were prepared as previously described. [1]. POPC (Avanti Polar Lipids, Alabaster, AL, USA) was dissolved in gel filtration buffer containing 100 mM sodium cholate. For all nanodiscs, POPC and MSP were mixed in a ratio of 130:1 and diluted to a final concentration of 10 mM POPC. For CorA-loaded nanodiscs, CorA (pentamer) was added to produce a ratio of 1300:10:1 for POPC:csMSP1E3D1:CorA, i.e. one CorA pentamer for every five nanodiscs, and the sample was diluted with gel filtration buffer to a final concentration of 10 mM POPC. Samples were thoroughly mixed and subsequently incubated for 15 min at RT before addition of detergent absorbing beads, Amberlite XAD-2 (Sigma-Aldrich, Darmstadt, Germany), at 15% w/v for empty nanodiscs and 30% w/v for CorA-loaded nanodiscs. The samples were incubated at 5°C while shaking at 800 rpm ON to remove cholate. Next day, the samples were separated from the detergent absorbing beads, and CorA-loaded nanodiscs were applied to a 1 ml HisTrap column (GE Healthcare, Chicago, IL, USA) equilibrated in gel filtration buffer containing 20 mM imidazole to bind only loaded nanodiscs, which were subsequently eluted with elution buffer. Appropriate fractions were diluted two-fold in gel filtration buffer, concentrated to < 1 mL and purified by SEC on a Superdex 200 10/300 GL (GE Healthcare, Chicago, IL, USA) equilibrated in an appropriate buffer. Unloaded nanodiscs were directly applied to the column after the removal of detergent absorbing beads. For the SLS/SEC experiments, the buffer was PBS, whereas for SAXS, the buffer was gel filtration buffer.

4.6 | Mass Spectrometry

A ProteinWorks^{TM} eXpress Digest Kit (Waters, Miliford, USA) was used to digest the protein samples. Fibrinopeptide (Glu-Fib), formic acid (FA) and acetonitrile (ACN) were purchased from Sigma-Aldrich (St. Louis, MO, USA). All other chemicals and reagents were of at least analytical grade.

A 3-step protocol was used for tryptic digestion as described by the manufacturer. Samples containing 500 pmol of csMSP1E3D1 were denatured by incubation with digestion buffer for 10 min at 80 °C. The samples were diluted with digestion buffer to a final volume of 175 μL followed by addition of 25 μL trypsin solution. The samples were subsequently incubated for 2 h at 45 °C. To quench trypsin activity, 5 μL of digestion inactivation reagent were added to the samples and incubated for 15 min at 45 °C. The sample supernatants were collected after centrifugation at 1800g for 15 min at 10 °C.

LC-MS was performed using a NanoAQUITY UPLC setup coupled to a ESI hybrid Q-TOF mass spectrometer (Synapt G2-SI, Waters, Wilmslow, UK). Intact mass analysis was performed by loading 50 pmol of intact protein in

gel filtration buffer for trapping on a reversed-phase C18 trap-column (ACQUITY UPLC BEH C18 1.7 μm VanGuard column, Waters, Miliford, USA) and eluted directly onto the mass spectrometer by a chromatographic gradient ranging 8-95% of mobile phase B (ACN, 0.23% FA) over 9 min at 40 μL/min. Mass spectrometry was in positive ion mode with a scan time of 1 second in a range of 300-2000 m/z. Deconvoluted mass spectra were produced by MaxEnt 1 algorithm in the MassLynx software, after being processed with baseline subtraction parameters set to (20,5) and Savitzky-Golay smoothing parameters set to (1,15). The experiment was performed in three replicates.

Peptide mapping analysis was performed by loading 20 pmol of the digested protein onto the UPLC system and trapped and desalted on a C18 trap column (ACQUITY UPLC BEH C18 1.7 μm VanGuard column, Waters, Miliford, USA) for 3 min at 150 μL/min of mobile phase A (0.23% FA). The tryptic peptides were eluted across a C18 analytical column (ACQUITY UPLC BEH C18 1.7 μm 1.0 x 100 mm column, Waters, Miliford, USA) and into the mass spectrometer using a gradient from 1% to 50% of mobile phase B for 20 min at a flow-rate of 40 μL/min. Mass spectrometry was performed in positive ion mode, with internal mass-calibration using a reference lock-spray signal of Glu-Fib. Peptide identification was performed by DDA (Data-Dependent Acquisition) CID MS/MS. Mass spectra were acquired over an m/z-range of 300-2000 with a scan time of 0.5 s for MS survey and 0.3 s for each subsequent MS/MS scan. UPLC-MS/MS data of the digested samples were processed against the csMSP1E3D1 sequence via Protein Lynx Global Server (PLGS, Waters, Milford, MA). Criteria for peptide identification were a mass tolerance for precursors ions < 15 ppm and a ladder score of product ions > 40. One missed trypsin cleavage was tolerated and methionine oxidation was chosen as variable modification.

The "linking" peptide (GGSLPETGSSFSK) was identified through MS/MS data with a mass accuracy of 20 ppm and 26 ppm for the singly charged and doubly charged precursor ion, respectively, and an MS/MS ladder score > 70 with a mass accuracy of assigned fragment ions from 5 to 20 ppm.

4.7 | Gel quantification of MSP circularization

For the quantification of the influence of substrate concentration, lsMSP1E3D1 was mixed with eSrt a ratio of 10:1 but in varying concentrations under the buffer conditions described for the large scale circularization. The samples were incubated for 3 h at 37 °C and quenched by the addition of 100% trichloracetic acid to a final concentration of 12 %. Precipitated protein was pelleted by centrifugation at 14,500g for 10 min at 4°C and the pellet was subsequently washed in ice cold acetone. This step was repeated before solubilizing the pellet in SDS loading buffer (50mM TrisHCl pH 6.8, 10% glycerol, 2% w/v SDS, 0.05% w/v bromphenol blue) to a final concentration of 8 μM MSP. For the quantification of the circularization reaction with varying eSrt:lsMSP1E3D1 ratio, reactions were setup with different amounts of eSrt in samples containing 10 μM lsMSP1E3D1. At indicated time points, 20 μL samples were taken out and quenched by the addition of 5 μL 5x SDS loading buffer. Samples were run on 12 % acryl amide gels. Gels were stained in Coomassie brilliant blue G250 (Bio-Rad, Hercules, CA, USA) and images of the gels were obtained on a Gel Doc (Bio-Rad, Hercules, CA, USA). The images were quantified in ImageLab (Bio-Rad, Hercules, CA, USA) using standard settings. Relative gel band intensities were calculated to report the relative quantities of species.

4.8 | Gel quantification of CorA-loaded nanodiscs

Standards with an absorption at 280nm were prepared of CorA and csMSP1E3D1 and loaded in different amounts on the same gel as the sample. Gels were stained with Coomassie G-250 (Bio-Rad, Hercules, CA, USA) and subsequently destained in 10% acetic acid, 20% ethanol. Gels were imaged using a GelDoc (Bio-Rad) and band intensities were quantified in the ImageLab software (Bio-Rad, Hercules, CA, USA) using standard settings. The intensities of the

standards were fitted with linear regression in order to quantify the amount of CorA and csMSP1E3D1 in the nanodisc sample.

4.9 | Static light scattering (SLS)

Samples of 5 μM nanodiscs were measured on a BI-200SM light scattering instrument (Brookhaven Instruments, Holtsville, NY, USA) equipped with at diode laser emitting light at a wavelength of 637 nm and with the detector placed at 90° to the incoming light. The temperature of the instrument was kept constant by an external water bath, and the temperature inside the VAT was measured by a built-in thermometer. Prior to measurement, samples were filtered four times into the sample cuvette through a 0.22 μm filter (Merck/Millipore, Darmstadt, Germany) to eliminate dust particles. Hereafter, samples were degassed for 15 minutes to prevent the formation of bubbles. Frames of 30 s were continuously measured over the course of 72 h, and subsequently, bad frames with abnormally high intensity, probably due to traces of dust particles, were removed. Frames were averaged in 2 h bins and normalized to the intensity of the first frame, which maintained the overall trends of the data. The recorded SLS intensity is approximately proportional to the average molecular weight of the species in solution as described by Høiberg-Nielsen *et al.* [35].

4.10 | Circular dichroism spectroscopy (CD)

Samples were exchanged into 20 mM sodium phosphate pH 7.0 on a NAP5 column (GE Healthcare, Chicago, IL, USA) and subsequently diluted to 1.5 μM in that buffer and degassed for 15 min. Sample concentrations were determined from absorption at 280 nm measured on a UV-1600PC spectrophotometer (VWR, Radnor, PA, USA). Circular dichroism (CD) spectroscopy was measured on a J-815 spectrometer (Jasco, Easton, MD, USA) equipped with a variable temperature sample holder controlled by a Peltier element. CD spectra were collected in three replicas with a bandwidth of 2 nm, scan speed of 20 nm/min, and data pitch of 1 nm at 20°C. The CD signal was recorded in millidegrees ($m\theta$) and subsequently converted to mean residue ellipticity ($[\theta]_{MRW}$) with the relation

$$[\theta]_{\mathrm{MRW}} = \frac{m\theta}{10cN_{pb}l}$$

where c is the protein concentration in molar, N_{pb} is the number of peptide bonds, i.e. the number of amino acids minus one, and l is the path length of the sample cuvette in cm, in this case 0.1 cm. For thermal scans, the CD signal at 220 nm was monitored at a heating/cooling rate of 1°C/min. To estimate the melting temperature, the data was noise-filtered using the Savitsky-Golay algorithm [36] from which derivatives of the melting curves were computed.

4.11 | Small angle X-ray scattering (SAXS)

Samples were prepared in the lab in Copenhagen and kept on ice until the time of measurement. Before exposure to X-rays, the absorptions at 280 nm of the samples were measured in triplicates using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) to estimate the sample concentration. The sample concentrations were: csND 1 = 10.7 μM, csND 2 = 10.3 μM, lsND 1 = 12.3 μM, lsND 2 = 11.1 μM and CorA csND = 0.9 μM. SAXS measurements were done at the P12 BioSAXS beamline at PETRAIII at Deutsches Elektronen-Synchrotron (DESY) in Hamburg, Germany at 10°C [37]. The intensity was measured as function of the magnitude of the scattering vector q, where $q = 4\pi \sin(\theta)/\lambda$ with θ being the scattering angle and λ the wavelength, in this case 1.24 Å. Data reduction was done automatically with the software available at the beam line and the 1D data was brought to absolute scale

in units of cm^{-1} using water as a calibration standard [38, 28]. Pair distance, $p(r)$, distributions were obtained from indirect Fourier transformations (IFT) calculated at the BayesApp-server [39].

Modeling was carried out in WillItFit [40]. The structural models (see futher description below) refined from the data are similar to the ones introduced by Huda *et al.* [22], which in turn are generalizations of similar models [9, 41]. The refinement was done using the same software (WillItFit?) [40], with the source code of the model available in the associated repository. Apart from the parameters listed in Table 2, the model also includes a constant background added to the model. The term accounting for interface roughness is described in the literature [42]. In addition, the height of the protein belt surrounding the bilayer was fixed to 25.78 Å to match the properties of a recently published structure of a similar protein [21].

The model describes a uniform distribution of nanodiscs with differing ellipticity parameterized by the axis ratio of the bilayer patch. Consequently, nanodiscs with large axis ratio will contain fewer lipids than nanodiscs with axis ratios closer to 1. The individual nanodiscs are described by analytical form factors [43] as explained by Skar-Gislinge *et al.* [9, 41] The model is fitted on absolute scale and utilizes information from the molecular composition of the MSP and POPC to calculate the scattering lengths of these compounds. The resulting values of χ^2 can be found listed in table 2 and the error bars associated to the refined parameters are based on these values as described in the literature. [44, 45]

Mathematically, our model has the form for the measured intensity, $I(q)$:

$$I(q) = \frac{n}{\kappa_{\max} - 1} \int_1^{\kappa_{\max}} d\kappa \; P(q, \vec{p})$$

where $\vec{p}$ is the list of parameters in our model for a nanodisc with a particular shape, κ is the aforementioned axis ratio of the bilayer, and n is the number density of nanodiscs estimated from the absorption at 220 nm. In particular, note that $\kappa \in \vec{p}$, whereas $\kappa_{\max}$, the largest axis ratio in the distribution, is a fitting parameter. $P(q, \vec{p})$ is the form factor intensity for an orientationally averaged model of the nanodisc described by the particular set of parameters, $\vec{p}$. [9, 41]

acknowledgements

The authors thank DESY in Hamburg for allocated beam time at the bioSAXS beam line at P12. In this context, we thank Alexey Kikhney and Melissa Gräwert for support during the beam time. Furthermore, we thank professor Mikaela Rapp (University of Stockholm, Sweden) for the plasmid harboring the gene for CorA, and professor Steven Sligar for the plasmid harboring the gene for MSP1E3D1. We thank Søren Midtgaard and Lasse Dreyer for measuring the SAXS data presented in this article. We thank Peter W. Thulstrup for providing access to the CD measurements.

author contributions

NTJ, FGT and LA conceived the project. NTJ and FGT produced proteins and nanodiscs and performed experiments. MCP performed SAXS data analysis. TTTNN and KDR performed MS experiments and data analysis. All authors wrote the paper.

conflict of interest

The authors declare no conflicts of interest.

supporting information

IEC purification tests (Figure S1); MS sequence coverage (Figure S2); Identification of oxidized species (Figure S3).

references

[1] Bayburt TH, Grinkova YV, Sligar SG. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. Nano Lett 2002 aug;2(8):853–856.

[2] Hagn F, Etzkorn M, Raschle T, Wagner G. Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. J Am Chem Soc 2013 feb;135(5):1919–1925.

[3] Denisov IG, Sligar SG. Nanodiscs in Membrane Biochemistry and Biophysics. Chem Rev 2017 mar;117(6):4669–4713.

[4] Denisov IG, Grinkova YV, Lazarides AA, Sligar SG. Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. J Am Chem Soc 2004 mar;126(11):3477–3487.

[5] Seddon AM, Curnow P, Booth PJ. Membrane proteins, lipids and detergents: not just a soap opera. Biochim Biophys Acta - Biomembr 2004 nov;1666(1-2):105–117.

[6] Jørgensen IL, Kemmer GC, Pomorski TG. Membrane protein reconstitution into giant unilamellar vesicles: a review on current techniques. Eur Biophys J 2017 mar;46(2):103–119.

[7] Dürr UHN, Gildenberg M, Ramamoorthy A. The Magic of Bicelles Lights Up Membrane Protein Structure. Chem Rev 2012 nov;112(11):6054–6074.

[8] Hagn F, Nasr ML, Wagner G. Assembly of phospholipid nanodiscs of controlled size for structural studies of membrane proteins by NMR. Nat Protoc 2018;13(1):79–98.

[9] Skar-Gislinge N, Simonsen JB, Mortensen K, Feidenhans'l R, Sligar SG, Lindberg Møller B, et al. Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. J Am Chem Soc 2010 oct;132(39):13713–13722.

[10] Kynde SAR, Skar-Gislinge N, Pedersen MC, Midtgaard SR, Simonsen JB, Schweins R, et al. Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment. Acta Crystallogr Sect D Biol Crystallogr 2014 feb;70(2):371–383.

[11] Shen PS, Yang X, DeCaen PG, Liu X, Bulkley D, Clapham DE, et al. The Structure of the Polycystic Kidney Disease Channel PKD2 in Lipid Nanodiscs. Cell 2016 mar;167(3):763–773.

[12] Gao Y, Cao E, Julius D, Cheng Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. Nature 2016 may;534(7607):347–351.

[13] Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, et al. Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs. In: Düzgünes N, editor. Methods Enzymol., vol. 464 Academmic Press; 2009.p. 211–231.

[14] Nasr ML, Baptista D, Strauss M, Sun ZYJ, Grigoriu S, Huser S, et al. Covalently circularized nanodiscs for studying membrane proteins and viral entry. Nat Methods 2017 jan;14(1):49–52.

[15] Chen I, Dorr BM, Liu DR. A general strategy for the evolution of bond-forming enzymes using yeast display. Proc Natl Acad Sci 2011 jul;108(28):11399–11404.

[16] Yusuf Y, Massiot J, Chang YT, Wu PH, Yeh V, Kuo PC, et al. Optimization of the Production of Covalently Circularized Nanodiscs and Their Characterization in Physiological Conditions. Langmuir 2018 feb;34(11):3525–3532.

[17] Miehling J, Goricanec D, Hagn F. A Split-Intein-Based Method for the Efficient Production of Circularized Nanodiscs for Structural Studies of Membrane Proteins. ChemBioChem 2018 sep;19(18):1927–1933.

[18] Trevino SR, Scholtz JM, Pace CN. Amino Acid Contribution to Protein Solubility: Asp, Glu, and Ser Contribute more Favorably than the other Hydrophilic Amino Acids in RNase Sa. J Mol Biol 2007 feb;366(2):449–460.

[19] Kramer RM, Shende VR, Motl N, Pace CN, Scholtz JM. Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility. Biophys J 2012 apr;102(8):1907–1915.

[20] Denisov IG, McLean MA, Shaw AW, Grinkova YV, Sligar SG. Thermotropic Phase Transition in Soluble Nanoscale Lipid Bilayers. J Phys Chem B 2005 aug;109(32):15580–15588.

[21] Bibow S, Polyhach Y, Eichmann C, Chi CN, Kowal J, Albiez S, et al. Solution structure of discoidal high-density lipoprotein particles with a shortened apolipoprotein A-I. Nat Struct Mol Biol 2017 feb;24(2):187–193.

[22] Huda P, Binderup T, Pedersen MC, Midtgaard SR, Elema DR, Kjær A, et al. PET/CT Based In Vivo Evaluation of 64Cu Labelled Nanodiscs in Tumor Bearing Mice. PLoS One 2015 jul;10(7):e0129310.

[23] Sigalov AB, Stern LJ. Oxidation of methionine residues affects the structure and stability of apolipoprotein A-I in reconstituted high density lipoprotein particles. Chem Phys Lipids 2001 nov;113(1-2):133–146.

[24] Midtgaard SR, Pedersen MC, Arleth L. Small-Angle X-Ray Scattering of the Cholesterol Incorporation into Human ApoA1-POPC Discoidal Particles. Biophys J 2015 jul;109(2):308–318.

[25] Graziano V, Miller L, Yang L. Interpretation of solution scattering data from lipid nanodiscs. J Appl Crystallogr 2018 feb;51(1):157–166.

[26] Skar-Gislinge N, Johansen NT, Høiberg-Nielsen R, Arleth L. Comprehensive Study of the Self-Assembly of Phospholipid Nanodiscs: What Determines Their Shape and Stoichiometry? Langmuir 2018 oct;34(42):12569–12582.

[27] Kučerka N, Nieh MP, Katsaras J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. Biochim Biophys Acta - Biomembr 2011 nov;1808(11):2761–2771.

[28] Mylonas E, Svergun DI. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. J Appl Crystallogr 2007 feb;40(s1):s245–s249.

[29] Larsen AN, Sørensen KK, Johansen NT, Martel A, Kirkensgaard JJK, Jensen KJ, et al. Dimeric peptides with three different linkers self-assemble with phospholipids to form peptide nanodiscs that stabilize membrane proteins. Soft Matter 2016;12(27):5937–5949.

[30] Greenfield NJ, Hitchcock-Degregori SE. Conformational intermediates in the folding of a coiled-coil model peptide of the N-terminus of tropomyosin and $\alpha\alpha$-tropomyosin. Protein Sci 1993 aug;2(8):1263–1273.

[31] Antos JM, Popp MWL, Ernst R, Chew GL, Spooner E, Ploegh HL. A Straight Path to Circular Proteins. J Biol Chem 2009 jun;284(23):16028–16036.

[32] Johansen NT, Pedersen MC, Porcar L, Martel A, Arleth L. Introducing SEC–SANS for studies of complex self-organized biological systems. Acta Crystallogr Sect D Struct Biol 2018 dec;74(12).

[33] Maric S, Skar-Gislinge N, Midtgaard S, Thygesen MB, Schiller J, Frielinghaus H, et al. Stealth carriers for low-resolution structure determination of membrane proteins in solution. Acta Crystallogr Sect D Biol Crystallogr 2014 feb;70(2):317–328.

[34] Palombo I, Daley DO, Rapp M. The Periplasmic Loop Provides Stability to the Open State of the CorA Magnesium Channel. J Biol Chem 2012 aug;287(33):27547–27555.

[35] Høiberg-Nielsen R, Fuglsang CC, Arleth L, Westh P. Interrelationships of Glycosylation and Aggregation Kinetics for Peniophora lycii Phytase. Biochemistry 2006 apr;45(15):5057–5066.

[36] Savitzky A, Golay MJE. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal Chem 1964 jul;36(8):1627–1639.

[37] Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, et al. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). J Appl Crystallogr 2015 apr;48(2):431–443.

[38] Orthaber D, Bergmann A, Glatter O. SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. J Appl Crystallogr 2000 apr;33(2):218–225.

[39] Hansen S. Update for BayesApp : a web site for analysis of small-angle scattering data. J Appl Crystallogr 2014 aug;47(4):1469–1471.

[40] Pedersen MC, Arleth L, Mortensen K. WillItFit : a framework for fitting of constrained models to small-angle scattering data. J Appl Crystallogr 2013 dec;46(6):1894–1898.

[41] Skar-Gislinge N, Arleth L. Small-angle scattering from phospholipid nanodiscs: derivation and refinement of a molecular constrained analytical model form factor. Phys Chem Chem Phys 2011;13(8):3161–3170.

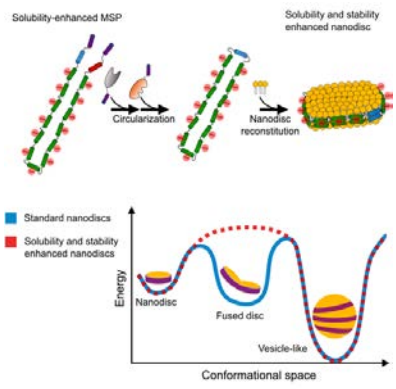
[42] Als-Nielsen J, McMorrow D. Elements of Modern X-ray Physics, 2nd Edition. Wiley, West Sussex; 2011.

[43] Pedersen JS. Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting. Adv Colloid Interface Sci 1997 jul;70(1-3):171–210.

[44] Pawitan Y. In All Likelihood. Clarendon Press, Cork, Ireland; 2001.

[45] Pedersen MC, Hansen SL, Markussen B, Arleth L, Mortensen K. Quantification of the information in small-angle scattering data. J Appl Crystallogr 2014 dec;47(6):2000–2010.

GRAPHICAL ABSTRACT



A solubility-enhanced membrane scaffold protein was genetically designed by introducing high abundance of negatively charged amino acids. Its enhanced solubility facilitate straight-forward and high yield expression, purification, and circularization. Nanodiscs made with this protein are highly conformationally stable and thus suitable for membrane protein incorporation and studies at physiologically relevant temperatures.

SUPPORTING INFORMATION

Circularized and solubility-enhanced MSPs facilitate simple and high yield production of stable nanodiscs for studies of membrane proteins in solution

Nicolai Tidemand Johansen[1] | Frederik Grønbæk Tidemand[1] | Tam T. T. N. Nguyen[2] | Kasper Dyrberg Rand[2] | Martin Cramer Pedersen[1] | Lise Arleth[1]

Circularization: IEC

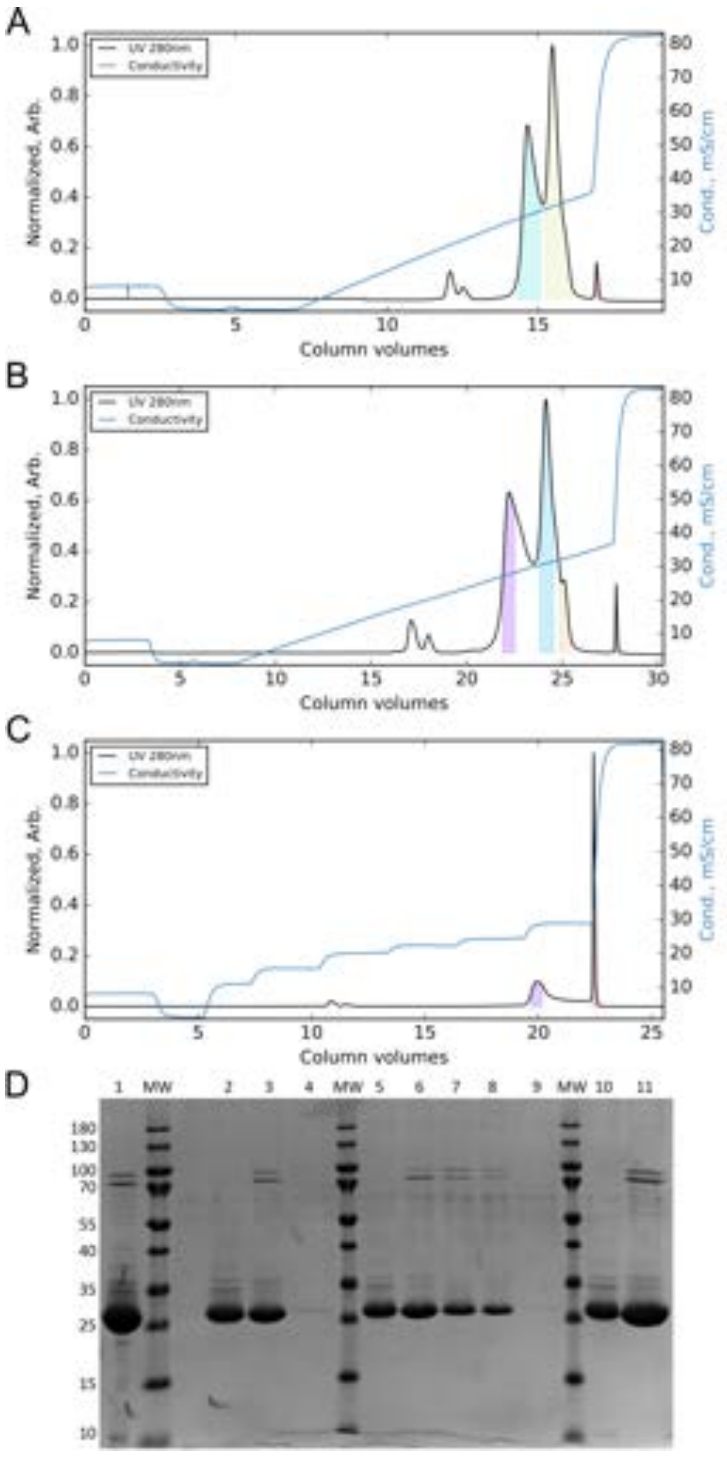


FIGURE S1 IEC purification tests. A: csMSP1E3D1 elution with 10 CV gradient from 0 to 0.5 M NaCl. B: 20 CV gradient from 0 to 0.5 M NaCl. C: Multi-step gradient from 0 to 0.32 M NaCl. D: SDS-PAGE analysis of the samples collected in A, B and C (colored areas). Lane 1 is the sample before the IEC purification, lanes 2 – 4 are the samples from A, lanes 5 – 9 are the samples from B, and lanes 10 – 11 are the samples from C. MW are molecular wight standards.

Mass spectrometry: Sequence coverage

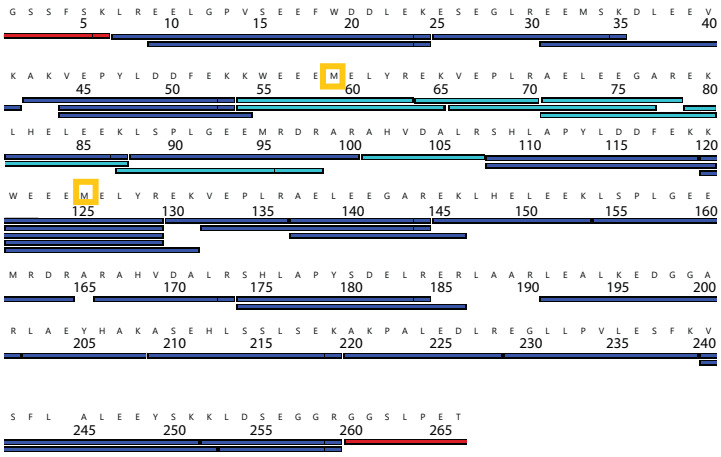


FIGURE S2 MS sequence coverage. Peptide map of csMSP1E3D1 after digestion with trypsin. The blue bars indicate the identified peptides by LC-MS/MS. The cyan bars indicate some identical peptides detected in the segment 46-113 relative to the segment 112-179 as these two segments are identical in sequence. The red bars indicate a tryptic peptide corresponding to the peptide segment linking the N-and C-terminal of csMSP1E3D1 (GGSLPETGSSFSK). Methionines marked with the yellow squares are the two major sites of oxidation.

Mass spectrometry: Methionine oxidation

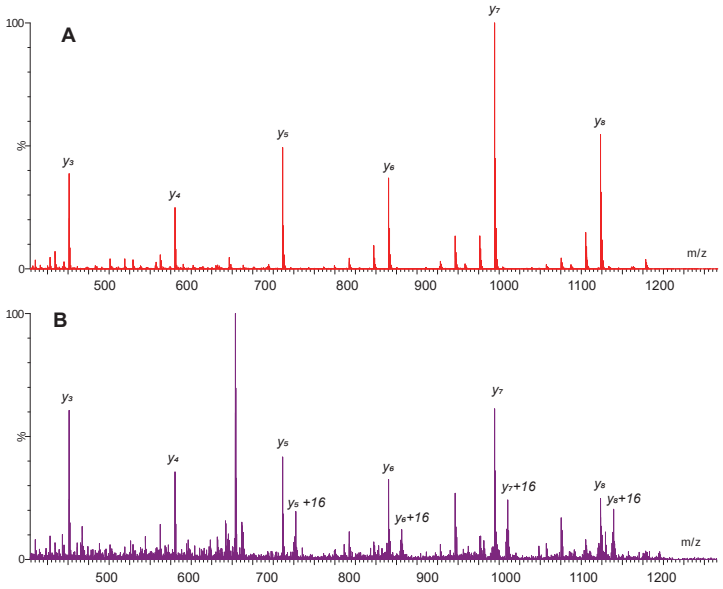


FIGURE S3 Identification of oxidized methionines. CID MS/MS fragmentation spectra of (A) non-oxidized and (B) oxidized form of a tryptic peptide 121-129 containing methionine (WEEEMELYR). The precursor ions were $[M+2H]^{2+}$ (m/z 642.79) and $[M+2H+16]^{2+}$ (m/z 650.79), respectively .

7

PAPER IV

CORA SOLUTION STRUCTURES

Small-angle neutron scattering show that the solution structures of the bacterial Mg^{2+}-channel CorA are overall similar with and without Mg^{2+} bound

Nicolai T. Johansen,[†] Andreas H. Larsen,[†] Martin C. Pedersen,[†] Pie Huda,[†,‖] Frederik G. Tidemand,[†] Jens Berndtsson,[‡] Thomas Günther Pomorski,[¶,§] Mikaela Rapp,[‡] and Lise Arleth[*,†]

†*Structural Biophysics, Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark*

‡*Department of Biochemistry and Biophysics, Stockholm University, Svante Arrhenius Väg 16C, SE-106 Stockholm, Sweden*

¶*Section for Transport Biology, Department of Plant and Environmental Science, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Denmark*

§*Molecular Biochemistry, Department of Biochemistry II, Ruhr University Bochum, Universitätsstraße 150, 44780 Bochum, Germany*

‖*Current address: Australian Institute for Bioengineering and Nanotechnology, University of Queensland, Cnr College Rd and Cooper Rd, Brisbane City QLD 4072, Australia*

E-mail: arleth@nbi.ku.dk

ND: The author list is incomplete at this time. More authors will be added when the manuscript is to be finalized.

Abstract

CorA is a homopentameric ion-channel involved in Mg^{2+}-regulation in prokaryotes. In several crystal structures, CorA adapts a highly symmetric structure in presence of Mg^{2+}. Based on recent cryo electron microscopy structures, it has been suggested that symmetry break is the main mechanism for regulation of Mg^{2+}. Here we use SANS and contrast-matched micelles and nanodiscs to interrogate the solution structure of CorA in presence and absence of Mg^{2+}. Surprisingly, we find no difference in scattering signal between these conditions, and furthermore, no current models fit to the obtained data. The SANS data is complemented by negative stain electron microscopy and activity assays to show that the samples are indeed intact under the D_2O condition used in SANS. Overall, this study shows that the structural fluctuations in CorA between Mg^{2+}-free and bound states are likely to be small.

Introduction

Magnesium is the most abundant divalent cation (Mg^{2+}) inside the cell, where it is mainly associated with the biological energy source adenosine triphosphate and other negatively charged molecules.[1] Mg^{2+} serves several biological functions, *e.g.* as co-factor for many enzymes, and Mg^{2+} deficiency is linked to severe diseases including cardiac syndromes, muscular dysfunction and bone wasting[2] The important task of upholding Mg^{2+} homeostasis is carried out by Mg^{2+} transporting proteins residing in the cell membrane.[3] The best described of these systems are prokaryotic Mg^{2+} transport membrane proteins (MPs), in particular MgtE and CorA.[4,5]

For the latter, several studies have provided insight to the mechanism of Mg^{2+} transport based on both cellular and biophysical studies.[6–14] Notably, Mg^{2+} is the most hydrated divalent cation; it hosts two water layers and the hydrated ion is approximately 400 times larger than the naked ion.[1] Stripping all water layers of Mg^{2+} would be energetically costly, and the ion selectivity mechanism of CorA has consequently evolved to an interaction with

the first-shell water layer by the conserved GMN-motif .[12,15–20]

A number of structures of CorA from the hyperthermophilic bacteria *Thermogata Maritima* have been determined by X-ray crystallography in the presence of divalent cations, Mg^{2+} [20–22] or calcium (Ca^{2+}).[23] In all of these, CorA is a symmetric pentamer (exemplified by 4I0U in Figure 1A) with a narrow 40 Å long hydrophobic pore thought to represent a closed state that does not allow Mg^{2+} to pass. The structure of the monomer is characterized by a transmembrane domain (TMD) of two helices that are linked to an intracellular domain (ICD) via a long stalk-helix.[21] Medium to low resolution structures of CorA, both in dodecyl-β-D-maltopyranoside (DDM) micelles and in phospholipid nanodiscs, recently emerged from cryo electron microscopy (cryoEM).[24] These structures revealed more open conformations of Mg^{2+}-free CorA, seen as large scale displacements of the ICDs (Figure 1B).[24] A conformational change has been supported by several structural and biochemical studies.[9,13,14,23]

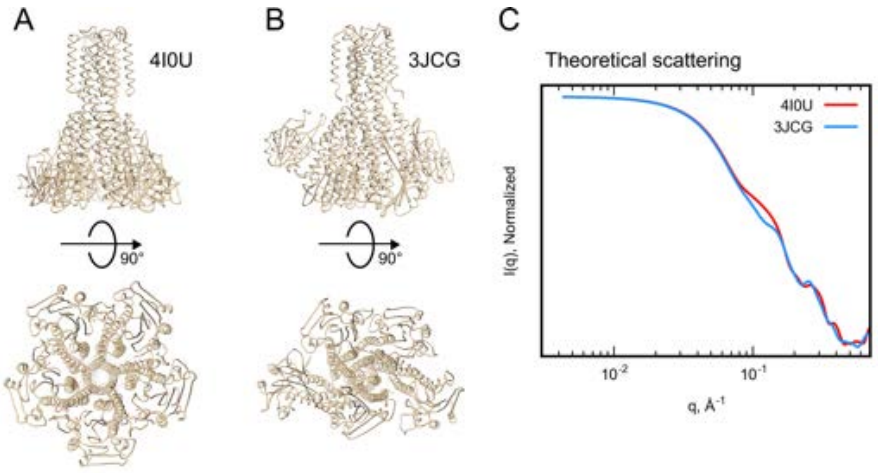


Figure 1: Symmetric and asymmetric CorA structures. **A**: CorA in the Mg^{2+}-bound symmetric structure obtained by crystallography (PDB ID: *4I0U*). **B**: CorA in the Mg^{2+}-free asymmetric structure obtained by cryoEM (PDB ID: *3JCG*). Both models have been added missing residues and/or side chains to match the full sequence of CorA. **C**: Theoretical scattering of a Mg^{2+}-free, open CorA structure (PDB ID: *3JCG*) and the Mg^{2+}-bound, closed crystal structure (PDB ID: *4I0U*).

It is well-acknowledged that crystallization can lead to crystal packing effects.[25] In cryoEM, particles are selected and aligned in a iterative process, which often involves subjective

decisions on the number of structures to be generated to explain different conformations.[26] Thus, these methods are ideally combined with methods that can probe structure under solution conditions, such as nuclear magnetic resonance spectroscopy (NMR) or small-angle scattering (SAS).[27] To our knowledge, only a single study reports solution scattering data of Mg^{2+}-free CorA.[28] Notably, though, this study was conducted on CorA from the thermophilic archea *Methanocaldococcus jannaschii* (MjCorA), which has low sequence identity (24.3%) to *Thermogata Maritima* CorA. Structurally, MjCorA has smaller ICDs with largely open monomer-monomer interfaces, but the same overall architecture, compared to *Thermogata Maritima* CorA.[17] In combination with low resolution cryoEM and X-ray crystallography, this study presented open-state conformations of CorA to some degree resembling the medium resolution cryoEM structures described above.[24] However, MjCorA was measured in DDM micelles,[28] which contributes a lot to the scattering and makes data analysis by standard methods, such as bead modelling,[29] non-trivial.

Recent developments in carrier system for membrane proteins has made small-angle neutron scattering (SANS) the preferred method for performing optimal scattering experiments on MPs in solution, including invisible (stealth) detergents[30] and stealth nanodiscs (sND).[31,32] These carrier systems are designed to match the scattering length density (SLD) of 100% D_2O, thereby being invisible to neutrons in that solvent and allowing for accurate determination of structural parameters for incorporated MPs. A further advantage is that at 100% D_2O, the incoherent background from hydrogen is at a minimum while the contrast of a standard membrane protein is at a maximum, providing optimal signal-to-noise ratio.[33]

Based on current structural models, it is expected that the proposed large domain displacement in the ICDs of CorA can be observed with solution scattering in absence of signal from surrounding micelles or nanodiscs (Figure 1C). Thus, in the present study, we applied SANS and novel invisible carrier systems to probe the structural states of CorA in solution. SANS data obtained in stealth DDM (sDDM) micelles show that CorA forms dimers of pentamers at the relatively high concentrations needed for SANS, which implicates downstream

data analysis. Using a novel size-exclusion chromatography (SEC)-SANS setup at the D22 instrument at the ILL,[34] data of monomeric CorA could be obtained, both in sDDM micelles and sNDs, respectively. The calculated model-independent structural parameters are in good agreement with the crystal structure of CorA, but surprisingly, no difference in scattering pattern is found between Mg^{2+}-free and bound states. Also strikingly, neither the symmetric crystal structure nor the asymmetric cryoEM structures explain the obtained data, both having systematic deviations at q$\sim$0.1 $Å^{-1}$. This prompted us to question whether the D_2O conditions used in SANS had a negative effect on CorA given the importance of Mg^{2+} hydration. However, we confirm that CorA is active in this slightly altered solvent. An overview of experiments and system are shown in figure 2. Overall, this study shows that the solution structures of CorA in Mg^{2+}-free and bound states are similar on average and thereby unlikely to undergo large structural rearrangements.

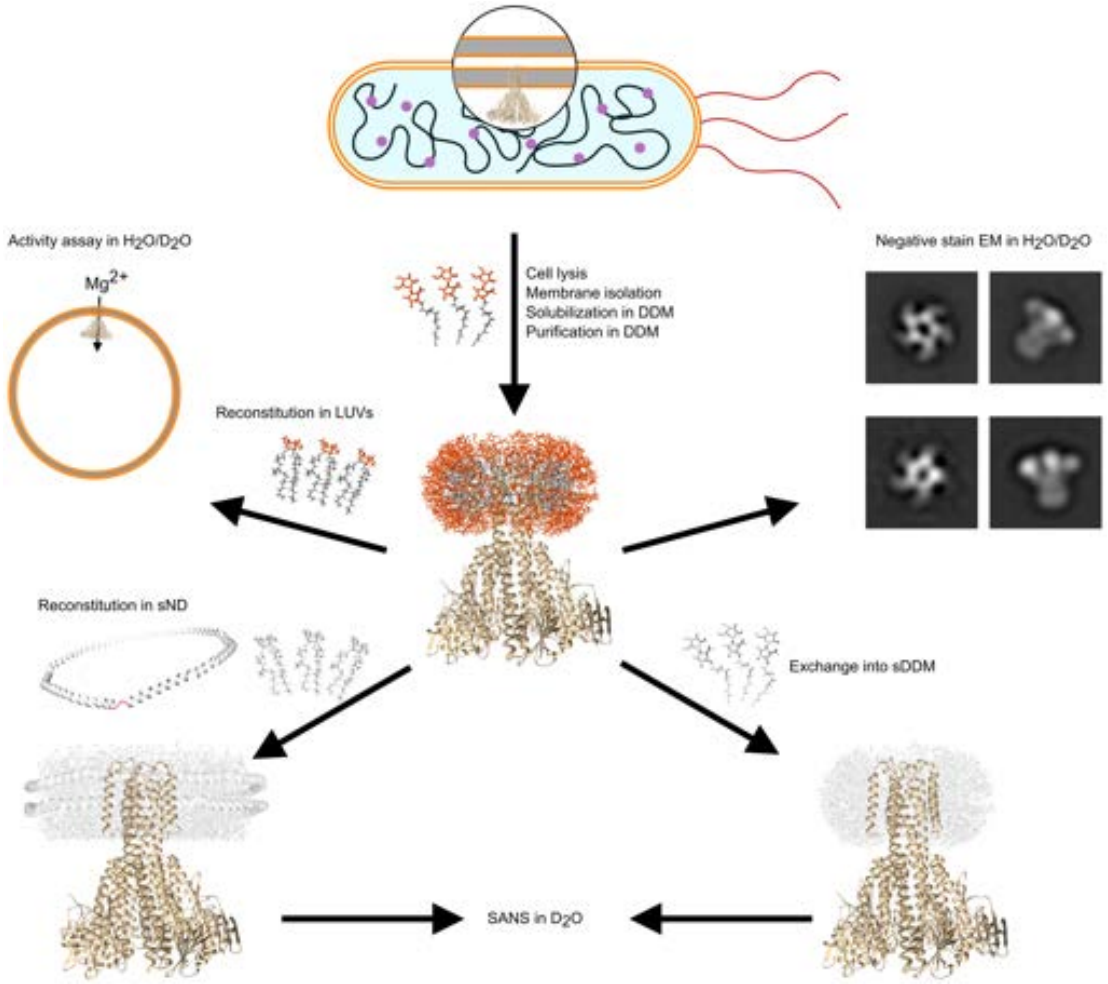


Figure 2: Overview of sample preparation and experiments. CorA was overexpressed in *E. coli* and extracted and purified in DDM. For activity assays, CorA was reconstituted in large uni-lamellar vesicles (LUVs) of 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) lipids. For negative stain EM studies, CorA was gel filtrated in H_2O or D_2O buffer containing DDM. For SANS studies, CorA was reconstituted in either stealth nanodiscs (sNDs) or stealth DDM (sDDM) micelles and gel filtrated in D_2O buffers.

Results

Generation of novel stealth nanodiscs (sNDs) for SANS

sNDs with almost zero forward scattering have been developed using phosphocholine (PC) lipids produced in *E. coli*[31] and shown to be applicaple for studying MP structural rear-

rangements.[32] Here, we develop and use sNDs made with novel chemically synthesized and pure d-POPC[35] with 94% deuteration in the tail-groups and 71% deuteration in the head-group. Furthermore, we use a deuterated version[36] of the recently developed csMSP1E3D1 belt protein[37] to generate slightly larger nanodiscs to accommodate CorA and at the same time providing more stable and soluble nanodiscs for SANS.

Because the deuterated d-csMSP1E3D1 and especially d-POPC were expensive material, nanodisc reconstitution was first optimized with hydrogenated csMSP1E3D1 and POPC. For empty NDs, the size of the formed NDs was observed to converge at a POPC:csMSP1E3D1 ratio of 125:1 (Figure S1A). The size of CorA-loaded NDs converged at a POPC:csMSP1E3D1 ratio of 100:1 using a fixed csMSP1E3D1:CorA ratio of 8:1 (Figure S1B). With a fixed POPC:csMSP1E3D1 ratio of 100:1, the csMSP1E3D1:CorA ratio could be lowered to 4:1 without compromising sample homogeneity (Figure S1C). Thus, empty sNDs were prepared with an initial d-POPC:d-csMSP1E3D1 ratio of 125:1, and CorA-loaded sNDs with a d-POPC:d-csMSP1E3D1:CorA ratio of 400:4:1.

The scattering of the empty sNDs were measured in a contrast series from 60% to 100% D_2O (Figure 3A upper panel). The scattering from the sND at 100% D_2O closely match that of buffer (Figure 3A, lower panel), giving a first indication of successful contrast matching. Importantly, the scattering signal has zero forward scattering and is flat in the entire q-range. For the samples in 60% to 90% D_2O, the forward scattering was found by Guinier-analysis (Figure 3B inset). To validate the contrast-match point (CMP) of the sND, the square root of the concentration normalized forward scattering, which is proportional to contrast, is plotted as a function of D_2O content (Figure 3B). It is seen that the data follow a linear trend with an intercept at 100.1% D_2O (Figure 3B). We note that the purple point for the sND at 100% was estimated from a linear fit of the scattering curve and was not included in the fit of the contrast series.

The structure of the sND was validated by SAXS (Figure 3C). The scattering data is indicative of a nanodisc with a characteristic minimum at $q = 0.06$ Å^{-1} and a double-bump

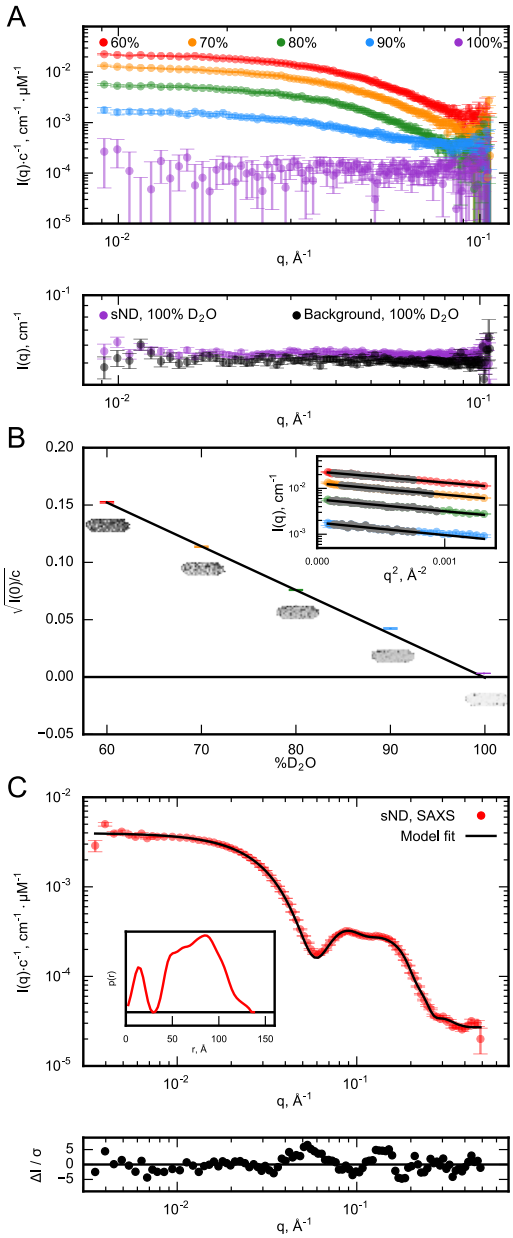


Figure 3: Validation of the sND. A: Contrast variation series of sNDs measured in 60-100% D_2O. The bottom plot shows the raw SANS data of the sND in 100% D_2O and the corresponding buffer. **B:** Plot of $\sqrt{I(0) \cdot c^{-1}}$, which is proportional to the contrast. $I(0)$ was estimated from the Guinier-approximation to which fits are shown in the insert. Grey nanodisc models are shown to indicate the progressively decreased contrast as the D_2O content in the buffer is increased. **C:** SAXS data from a sND sample with its $p(r)$-distribution shown in the inset. The solid black line is the fit of the nanodisc model.

in the range q = 0.1 - 0.2 Å^{-1}.[37] Furthermore, the $p(r)$-distribution refined from the SAXS data has the expected shape with a minimum around 30 Å due to the negative contrast of

the lipids and the expected D_{max} of approximately 130 Å(Figure 3C, inset), as shown for similar nanodiscs made with hydrogenated components.[37] Furthermore, a nanodisc model[38] could be fitted to the data (Figure 3C, solid black line). The overall fit is good with a χ^2 of 7.1, but with some discrepancy as illustrated by the variance normalized difference plot (Figure 3C). Such discrepancies are commonly encountered owing to the simple model of the nanodisc.[37,39] The refined parameters (Table 1) are overall in agreement with the previously determined values for csMSP1E3D1-POPC nanodiscs.[37]

Table 1: Structural parameters refined from the SAXS data presented in Figure 3C as well as derived parameters.
*Fixed parameter.
†Calculated based on an estimated 70% perdeuteration.

Refined parameters					
Axis ratio	A_{Head}, Å^2	h_{Belt}, Å	N_{Lipid}	V_{POPC}, Å^3	V_{MSP}, Å^3
1.45±0.06	59.9±4.1	25.78*	219±23	1244±5	32717±3540
Derived parameters					
$h_{Bilayer}$, Å	r_{major}, Å	r_{minor}, Å	w_{belt}, Å	ρ_{belt}, g cm^{-3}	χ^2
41.6	54.9	38.0	8.0	1.62†	7.1

SANS data on CorA in stealth carriers

SEC-SANS vs. static SANS

SANS data were obtained using standard static SANS or the novel SEC-SANS setup at D22.[34] In SEC-SANS, the sample is passed directly from SEC purification into the SANS flow-cell, which allow measurements of truly monodisperse species.[40] The chromatograms for both offline SEC purification for static SANS measurements as well as the chromatograms obtained during SEC-SANS are shown in Figure S2. In the offline purification, the chromatograms featured well-defined peaks (Figure S2A and B). Still, fractions were collected from the right-hand side of the peaks to avoid eventual larger particles present in the beginning of the peaks. In SEC-SANS, where different flow rates were used prior to sample elution and during sample elution, the elution time was not readily converted to elution volume,

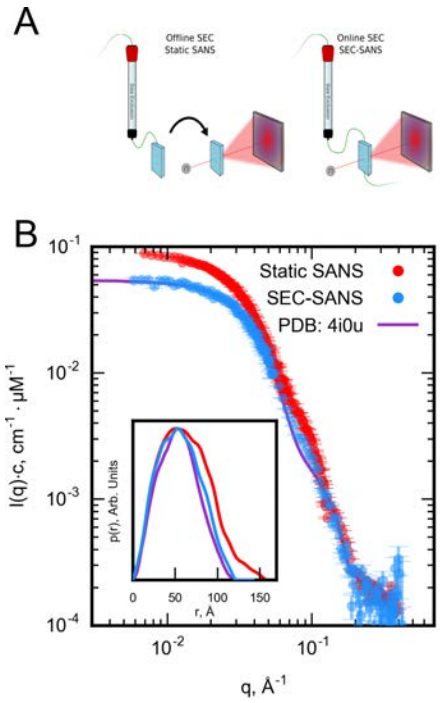


Figure 4: Comparison of standard static SANS vs SEC-SANS. A: Schematic illustrations of the difference between static SANS with offline SEC purification of samples and SEC-SANS with online SEC purification. **B:** Comparison of two data sets obtained with static SANS and SEC-SANS respectively, and compared to the theoretical scattefig from the CorA crystal structure (PDB ID: *4I0U*).

because the flow rate was not logged adequately (Figure S2C and D). Furthermore, due to the long path-length of the flow cell, the absorption at 280 nm saturated during sample elution. Thus, the absorption at 300 nm, which was used to estimate sample concentration, is shown as well for those data (Figure S2C and D). Again, frames were chosen from the right-hand side of part of the peaks, corresponding well to the fractions selected in the offline purification for static SANS.

The data obtained using static SANS show that CorA forms some fraction of dimers of pentamers in sDDM micelles (Figure 4B), although SEC purification was applied just prior to the measurement. With SEC-SANS, however, the data match the predicted scattering from the CorA crystal structure well (Figure 4B), suggesting that the single pentamer is measured with this technique. With SEC-SANS, the measurement time is limited by the flow rate of the SEC, which makes it difficult to obtain good signal-to-noise ratio at high-q

values.[37] Dimerization is unlikely to affect the local structure, *i.e.* short molecular distances, of CorA, and thus, the data obtained with static SANS is still valuable to evaluate differential scattering at high q-values between samples with and without Mg^{2+}.

SANS data indicate no large structural changes upon Mg^{2+} binding/release

All together, SANS data was collected on CorA in sDDM and sNDs using both static SANS and SEC-SANS (Figure 5). Samples were measured with or without Mg^{2+} present, and the first observation is that the scattering is invariant between these conditions for each respective carrier system and technique. The crystal structure of CorA was fitted to each data set, but did not produce good fits (Figure 5, solid black lines). In the data collected with static SANS, slight polydispersity is observed both in sDDM (Figure 5A) and sNDs (Figure 5C) as most clearly visualized from the refined $p(r)$-distributions (Figure 5, insets). This was effectively eliminated with the use of SEC-SANS both for sDDM (Figure 5B) and sNDs (Figure 5D). Thus, these data sets were chosen for further comparison in the following section. The R_g and D_{max} values obtained from $p(r)$-distributions (Figure 5) are shown in Table 2. For reference, the R_g and D_{max} values for the crystal structure (*4I0U*) are estimated to 41.3 Å and 135 Å, respectively, using Cryson.[41]

Table 2: Parameters obtained from IFT analysis. N_p is the estimated number of good parameters.[42]

	sDDM				sND			
	Static SANS		SEC-SANS		static SANS		SEC-SANS	
	w/ Mg	w/o Mg	w/ Mg	w/o Mg	w/ Mg	w/o Mg	w/ Mg	w/o Mg
R_g, Å	48.8	50.3	42.8	43.4	45.8	45.1	44.3	44.5
D_{max}, Å	154.1	165.7	118.6	126	140.9	139.4	118.0	117.4
N_p	12.9	13.4	8.3	8.9	11.7	10.9	6.9	6.7

No current structural models fit to the SANS data

To make a closer comparison of the SEC-SANS data obtained with or without Mg^{2+}, data obtained in sNDs were plotted together with a difference plot (Figure 6A). The difference

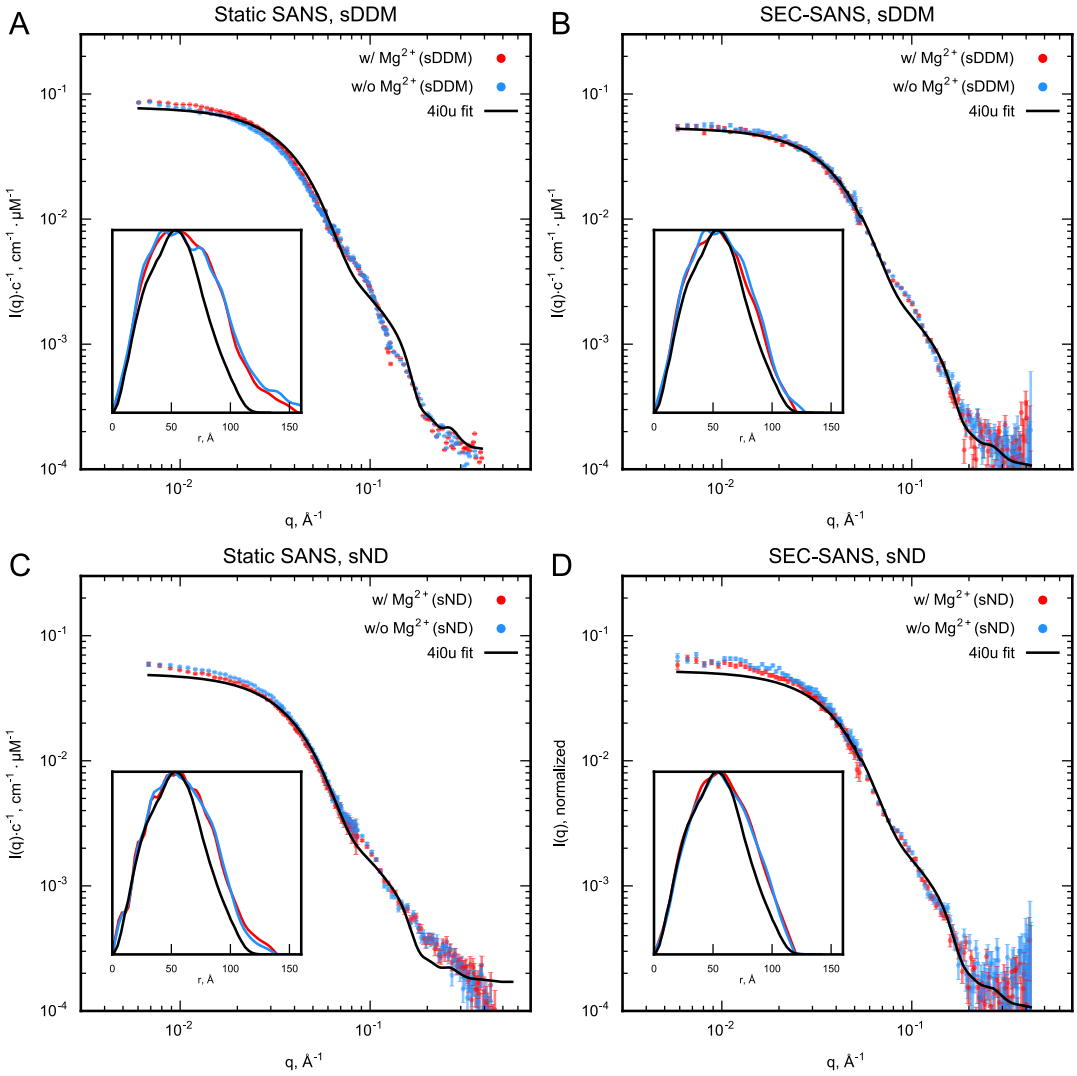


Figure 5: SANS data collected on CorA in sDDM and sNDs. SANS was collected by standard static SANS or with SEC-SANS with 40 mM Mg^{2+} or without Mg^{2+} (1 mM EDTA). **A:** Static SANS in sDDM. **B:** SEC-SANS in sDDM. **C:** Static SANS in sND. **D:** SEC-SANS in sND. Solid black lines indicate fits of the CorA crystal structure (PDB ID: *4I0U*) to the data.

plot shows the differential scattering intensity between the sample without Mg^{2+} relative to the sample with Mg^{2+} normalized to the error estimate on the intensity. Overall, there is no significant difference, which is also the case in the q-range where the theoretical scattering was expected to differ (0.1 $Å^{-1}$ to 0.15 $Å^{-1}$, see Figure 1C). The scattering intensity from CorA in sNDs varies significantly from CorA in sDDM micelles in the low-q region, but

insignificantly in the q-range 0.1 Å^{-1} to 0.15 Å^{-1} (Figure 6B).

Finally, different structural models, *i.e.* the symmetric crystal structure and the asymmetric cryo-EM structures were fitted to the data (Figure 6C). Common for all of these models is that they do not fit the data well. In particular they do not reproduce the feature in the data at $q \sim 0.08$ Å^{-1}. In the interesting q-range of 0.1 Å^{-1} to 0.15 Å^{-1}, the symmetric model fit the data slightly better than the asymmetric models. These observations were confirmed for all SEC-SANS data sets (data not shown).

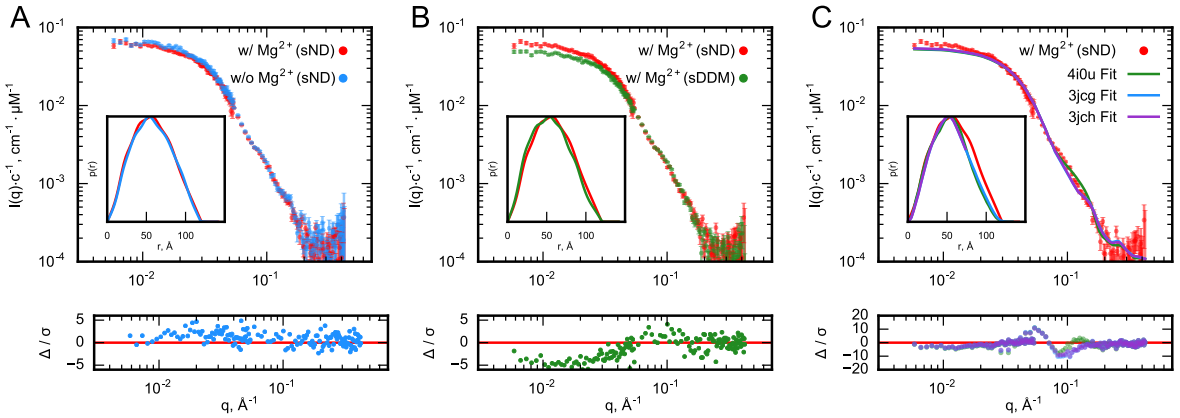


Figure 6: Comparison of SEC-SANS data. A: CorA measured on sNDs in presence (40 mM) or absence (1 mM EDTA) of Mg^{2+}. The inset shows $p(r)$-distributions refined from the SANS data, and the residual plot shows the error-normalized difference between the data sets as indicated by the colors. **B:** Same plots as in A, but for CorA measured in sNDs and sDDM micelles in presence of Mg^{2+}. **C:** Cryson fits to the indicated SANS data set, using the models shown in Figure 1 along with a second similar open model obtained by cryo-EM.[24] The relevant PDB IDs are given in the legend.

Ab initio modeling is unfeasible for the obtained SANS data

Due to the isolated signal of CorA in the stealth carriers, it is feasible to use common modeling approaches, such as *ab initio* bead-modeling.[30] Using the SEC-SANS data (Figure 5B and C) program DAMMIF from the ATSAS package was used to model average envelope densities.[29] Without imposing symmetry, the density envelopes for all data sets were elongated and relatively flat compared to the expected symmetric structure of CorA (Figure 7A). Imposing P5 symmetry, the obtained density envelopes resembled more the

expected symmetric CorA structure (Figure 7B). Notably, the normalized spatial discrepancy (NSD) values for the different average models were all between 1.2 and 1.6. In general, NSD values above 1 is a sign that individual models that are used to generate the averaged models are not similar.[43] Thus, *ab initio* bead-modeling is not feasible for CorA, at least not without further constraints applied.

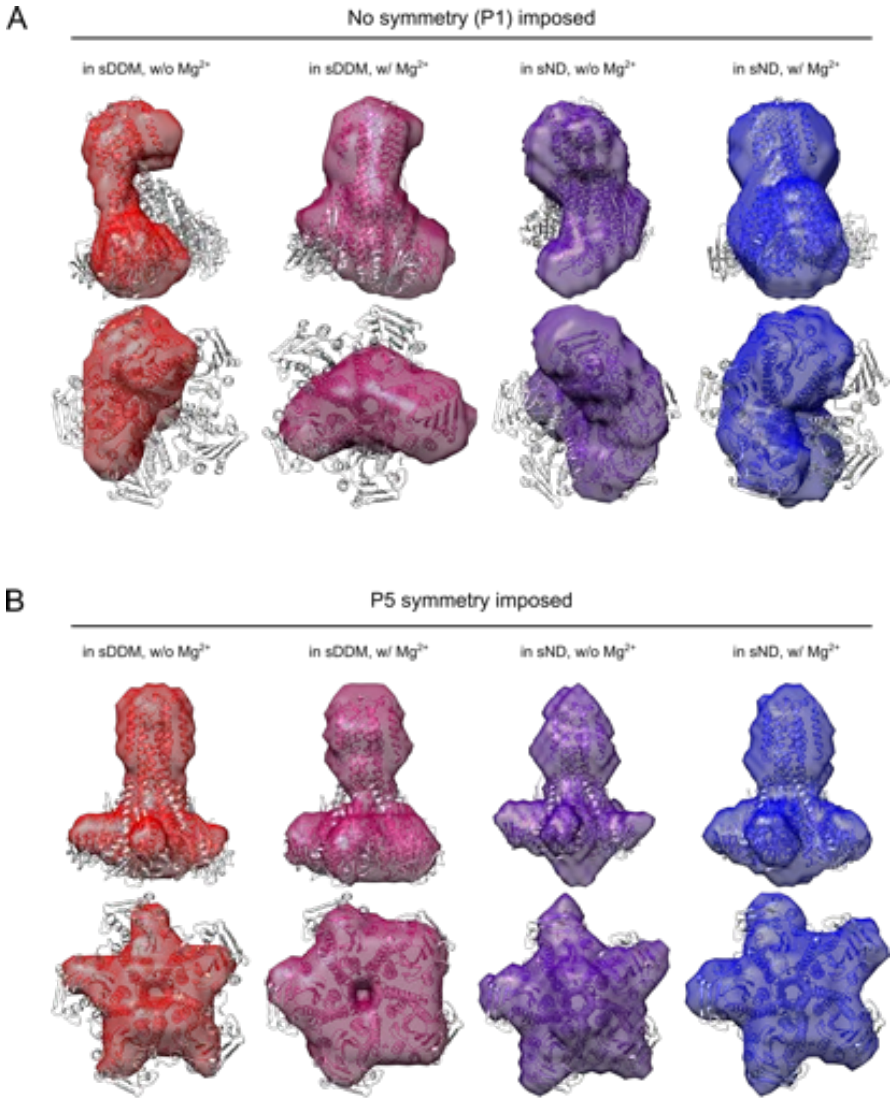


Figure 7: Bead-modeling of CorA from SEC-SANS data. A: Envelopes obtained from bead-modeling in DAMMIF are aligned with the CorA crystal structure (*4I0U*. The models are shown in a side- and top-view. **B:** Same as in **A**, but with P5 symmetry imposed in the modeling.

Negative stain EM confirms that CorA is a pentamer in D_2O

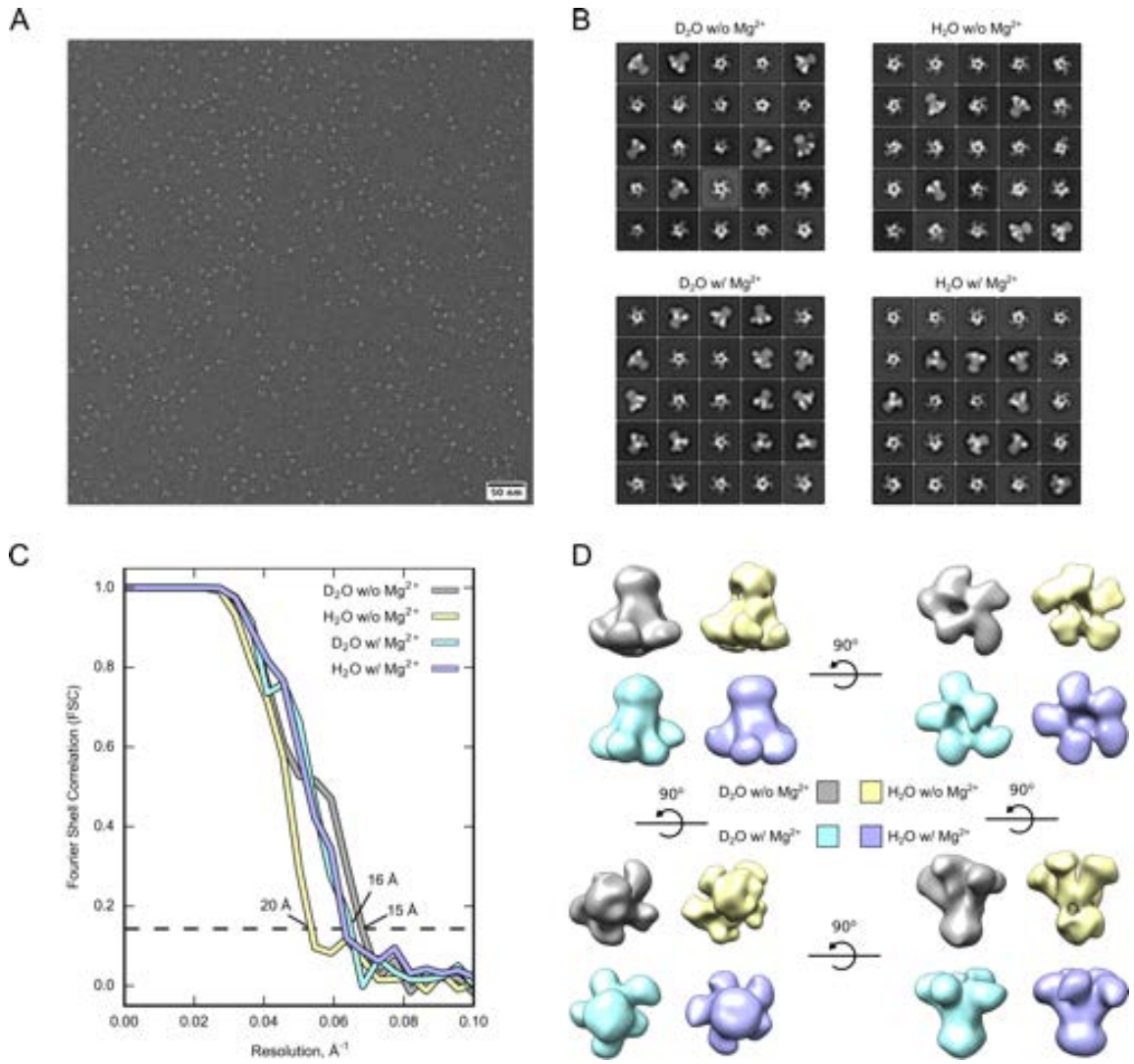


Figure 8: CorA structures from negative stain EM. A: A representative raw micrograph. **B:** A subset of 2D class averages selected for producing 3D models. **C:** Gold standard FSC plot for estimating resolution of the 3D models. Resolution estimates are given on the plot. D: 3D models refined from the 2D class averages in B.

The lack of structural difference in scattering intensities prompted us to compliment the SANS data with structural analysis by negative stain EM. In particular, we wished to verify the structural integrity of CorA in D_2O. Thus, samples of CorA in DDM micelles were measured in both H_2O and D_2O buffers. Good grids with a large number of well-dispersed particles could readily be obtained for all samples, as exemplified for the sample measured

in D_2O without Mg^{2+} (Figure 8A). The 2D class averages clearly show the pentameric architecture of CorA with a number of side, top and bottom views represented (Figure 8B). 3D models were refined from the 2D class averages without applying symmetry to resolutions ranging from approximately 15 Å to 20 Å, as determined by the gold standard Fourier Shell Correlation[44] at a cut-off of 0.143 (Figure 8C). The refined 3D models are overall quite similar, showing an overall pentameric architecture with slight asymmetry (Figure 8D). Notably, the ICDs appear better resolved in the samples containing Mg^{2+}, both in H_2O and D_2O (Figure 8D, turquoise and purple). Finally, we note that the TMD is to some degree shielded by the DDM micelle, which made alignment of the different 3D models difficult. Overall, these data show that the CorA pentamer is intact in D_2O.

Fluorometric assays confirm that CorA is active in D_2O

While the EM data showed that the CorA pentamer was structurally intact in D_2O, we speculated whether CorA activity was also preserved at the conditions used for SANS, *i.e.* D_2O and 10°C. To elaborate on this, we were particularly concerned about a "D_2O" effect, due to the highly hydrated nature of Mg^{2+}. To test activity, CorA was reconstituted in POPC LUVs containing an encapsulated Mg^{2+}-sensitive fluorophore, similar to a previous study on CorA.[9] In H_2O at 10°C, CorA is observed to conduct both Mg^{2+} and Ca^{2+} (Figure 9A). We note that the fluorescent probe has much higher affinity for Ca^{2+} (223 fold) causing a steeper slope of the signal with this ion, although CorA is selective for Mg^{2+}. No increase in fluorescence signal is observed in absence of Mg^{2+}, and activity is also inhibited in presence $Co[NH_3]^{3+}$, which resembles Mg^{2+} with a single hydration shell. The same observations are made in D_2O at 10°C, albeit at slower rates (Figure 9B). Comparison of the initial slopes of the data show that CorA activity in D_2O is 1.5 - 2 times slower than in H_2O (Figure 9C). Overall, these data show that CorA is active in D_2O and specifically inhibited by $Co[NH_3]^{3+}$ as is the case in H_2O.

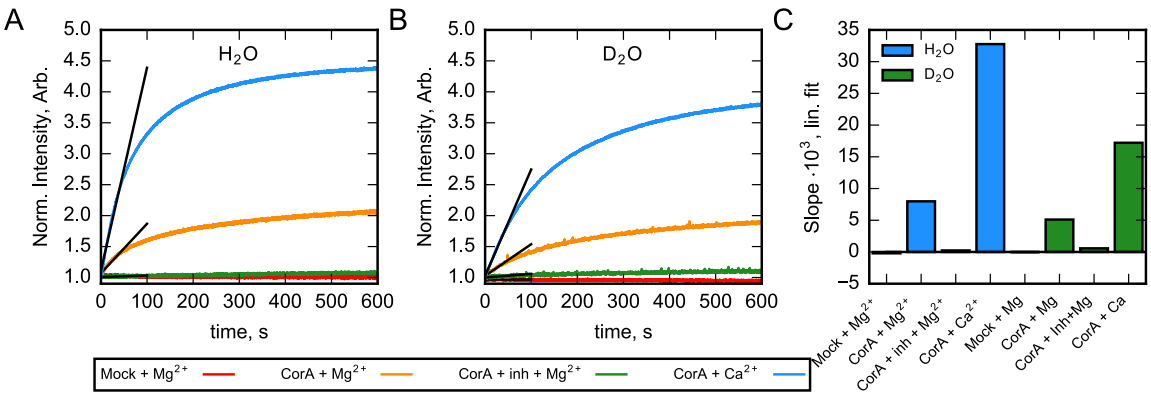


Figure 9: CorA channel activity in D_2O. A: CorA activity in H_2O-based buffer measured for empty liposomes (mock, red), and CorA-loaded liposomes added either 10 mM Mg^{2+} (orange), 1 mM $Co[NH_3]^{3+}$ (inhibitor) + 10 mM Mg^{2+} (green), or 10 mM $CaCl_2$ (blue). Solid lines are linear fits to the initial 100 s of data. **B:** Activity measured in D_2O-based buffer. **C:** Comparison of slopes obtained from linear fits in A and B.

Discussion

We have studied the solution structures of *Thermogata maritima* CorA using SANS and stealth carriers. As part of this study, novel sNDs were developed with chemically synthesized dPOPC and a deuterated version of an optimized nanodisc belt protein (d-csMSP1E3D1).[37] The sNDs were effectively contrast matched at 100% D_2O (Figure 3), facilitated by high-precision deuteration and chemically pure deuterated lipids. Notably, the high purity of these chemically synthesized lipids, *i.e.* identical tail lengths, allowed better control of the nanodisc reconstitution process compared to biologically deuterated PC lipids, which are a mix of lipids with different tail-lengths.[31] Furthermore, because of covalently circularized nanodisc, the MP:MSP stoichiometry could be better controlled than for standard MSPs,[45] which allowed us to use a minimal amount of the expensive dPOPC in sample preparation and obtain more homogeneous samples.

In sDDM micelles, CorA formed dimers of pentamers after SEC purification, which made analysis of the SANS data complicated. However, utilizing SEC-SANS,[34,40] a monomeric CorA-pentamer could be measured (Figure 4). Because the sample is flowed continuously in SEC-SANS, the minimum flow rate of the pumps effectively determine the available mea-

surement time. In our case, we were limited by a flow rate of 0.05 ml/min, which allowed us to extract frames corresponding to approximately 11 minutes. It is evident that the SEC-SANS data have lower signal-to-noise ratio than data obtained with standard static SANS (Figure 5), but good counting statistics were obtained in the interesting q-region (0.1 $Å^{-1}$ - 0.15 $Å^{-1}$), where the theoretic scattering intensities of symmetric and asymmetric CorA structures were expected to deviate (Figure 1). In this region, data obtained from sND samples were strikingly similar to the data obtained from sDDM samples (Figure 6B).

CorA has been proposed to undergo a large conformational change between a Mg^{2+}-bound symmetric non-conducting state to a Mg^{2+}-free asymmetric conducting state.[13,24] The most notable observation from the SANS studies presented here is the lack of a structural difference between samples with and without Mg^{2+} present for respective carrier systems and SANS setups (Figure 5). Comparing samples in different carrier systems, we note that the sND samples had some excess scattering intensity at low-q compared to samples measured in sDDM (Figure 6B). The lipid bilayer in nanodiscs is expected to provide a more native-like environment compared to detergent micelles.[46] Thus, one explanation for this discrepancy could be that CorA has enhanced dynamics in sNDs. On the other hand, the scattering intensities at higher q were highly similar, indicating that the finer structure of CorA was preserved in DDM. Furthermore, the reported cryoEM structures of CorA were obtained from samples in DDM.[24] By measuring CorA in both nanodiscs and DDM, we obtain an optimal argument that the solution structures of CorA in presence and absence of Mg^{2+} are overall similar in the resolution range, where they were expected to differ.

As the SANS data did not indicate the expected conformational change, negative stain EM was performed to verify the structural integrity of CorA in the conditions used for SANS, *i.e.* D_2O and 10°C. The negative stain EM data showed that the pentameric architecture of CorA was intact in D_2O and did not look compromised compared to in H_2O (Figure 8). Notably, the cryoEM structure of CorA in the closed state was refined with P5 symmetry imposed, whereas the open cryoEM structures of CorA were refined without symmetry.[24]

Here, we did all refinement without imposing symmetry to rule out eventual bias of this constraint, resulting in overall similar and slightly asymmetric models for all conditions (Figure 8D). The most striking difference between the models was that the individual ICDs appeared more well-defined in the Mg^{2+}-samples (Figure 8D). This is in agreement with a less dynamic structure in presence of Mg^{2+}.[23,24]

Furthermore, we speculated whether CorA was active in 100% D_2O. The heavier mass of D affects the energy of the O-D bond and consequently also hydrogen bonds between water molecules and thereby water structure.[47] Given that Mg^{2+} is the most hydrated metal-ion,[1] CorA activity could potentially be impacted in D_2O. Using a fluorimetric assay in POPC LUVs, we found that CorA was active in D_2O (Figure 9). The effective inhibition by $Co[NH_3]^{3+}$ veryfied that CorA was also specifically inhibited in this solvent. The rate of Mg^{2+} conduction in D_2O was reduced almost 2-fold compared to H_2O, which is, however, in line with observations for other MPs.[48]

All of these data combined indicate that the solution structures of CorA are overall similar in presence and absence of D_2O. Recently, a cryoEM structure was solved of the Zn^{2+} ZntB channel,[49] which is a homolog to CorA with same overall structure. A symmetric structure, with high resemblance to the symmetric structure of CorA, was obtained from a sample prepared without Zn^{2+}. The authors commented that, perhaps, the asymmetric CorA structures obtained without Mg^{2+} represented an artificial structural state because Mg^{2+} is always available in the cell, which is not the case for Zn^{2+}.[49] Thereby, they suggest that CorA has evolved to a symmetric state only in presence of Mg^{2+}. While our data do not strictly counter this argument, they show that the Mg^{2+}-free state of CorA is overall quite similar to the Mg^{2+} bound state. While we did not yet succeed in proposing a model to describe the SANS data, it is likely that CorA is overall symmetric on average, or in other words dynamically symmetric both with and without Mg^{2+}. At least, a large conformational change is not observed on average. To gain a more detailed view of the dynamics, other experiments are needed, such as solid-state NMR, together with simulations.

Conclusion

We have investigated the structure of the bacterial magnesium transport protein CorA under solution conditions using SANS and contrast matched deuterated micelles and nanodiscs in D_2O. Based on cryoEM structures[24] and electron paramagnetic resonance data,[13] symmetry-break has been suggested as the main mechanism for gating in CorA and Mg^{2+} uptake regulation. Using SANS, we were not able to observe such structural differences. This suggests that the solution structure of CorA is overall similar in presence and absence of Mg^{2+} and that eventual structural fluctuations are relatively small.

Experimental

Materials

All buffer components were from Sigma-Aldrich (Darmstadt, Germany), unless otherwise stated. Yeast extract and tryptone for bacterial growth media were from VWR (Radnor, PE, USA). 99.9% D_2O was from Cambridge Isotope Laboratories (Tewksbury, MA, USA) or from Sigma-Aldrich (Darmstadt, Germany). Nickel nitrilo-acetic acid (NiNTA) coupled agarose resin was from Thermo Fisher Scientific (Waltham, MA, USA). Amberlite XAD-2 was from Sigma-Aldrich (Darmstadt, Germany) and Biobeads SM-2 were from Biorad (Hercules, CA, USA). Gel filtration columns were from GE Healthcare (Chicago, IL, USA). Commercially available 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) was from Avanti Polar Lipids (Alabaster, AL, USA). Centrifugal filters for spin concentration (spin filters) were from Satorius (Göttingen, Germany). Dodecyl β-D-maltopyranoside (DDM) and octyl β-D-glucopyranoside (OG) were from Carbosynth (Compton, UK). Cell impermeable mag-Fluo-4 was from Thermo Fisher Scientific (Waltham, MA, USA).

Sample preparation

CorA expression and purification

CorA was expressed and purified as previously described, but leaving the N-terminal His-tag on for downstream purification of CorA-loaded nanodiscs.[18] In the final SEC step, the protein was purified in CorA-SEC buffer (20 mM Tris HCl pH 8, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.5 mM DDM). For making downstream nanodisc samples, the His-tag was left intact. The sample was concentrated to 100 μM in a 30,000 kDa MWCO spin filter, aliquoted and flash frozen in liquid nitrogen and stored at -80°C until further use. For making downstream samples in sDDM micelles, some of the purified CorA had the His-tag cleaved off by TEV protease: TEV protease was added to the CorA sample in a 1:10 ratio and incubated for 5 hours at room temperature followed by incubated overnight at 4°C. Uncleaved CorA was removed by Reverse IMAC purification on NiNTA resin. The sample was concentrated as above and purified by SEC in CorA-SEC buffer to remove the large excess of DDM micelles from the concentration step. A peak fraction of 30 μM CorA was extracted and used for SEC-SANS.

Deuterated csMSP1E3D1

Match-out deuterated csMSP1E3D1 (d-csMSP1E3D1) was expressed at the D-lab at ILL in Grenoble France[36] and purified as described previously.[37] The circularization protocol was modified slightly and was carried out at ambient temperature (28 °C) at pH 7.1 for 4 hours in eSrt buffer (20 mM Tris HCl pH 8, 150 mM NaCl, 10 mM $CaCl_2$, 1 mM DTT). Furthermore, 10 mM EDTA was added after reverse IMAC purification to suppress eventual trace sortase activity.[50] Finally, the protein was purified by SEC in MSP-buffer (20 mM TrisHCl, 150 mM NaCl, 1 mM EDTA). Unlike standard MSPs, csMSP1E3D1 is mainly monomeric and a well-resolved peak fraction could be extracted and subsequently concentrated in a 10,000 MWCO spin filter to a workable concentration for making nanodiscs (A280 $\sim$ 6).

Optimization of ND reconstitution

Because the deuterated d-csMSP1E3D1 and especially d-POPC were expensive material, nanodisc reconstitution was optimized with hydrogenated csMSP1E3D1 and POPC following the standard nanodisc protocol.[51] Briefly, csMSP1E3D1 and POPC solubilized by cholate were mixed in varying ratios of POPC:csMSP1E3D1 between 50:1 and 150:1, incubated for 15 min at room temperature, and had detergent removed by amberlite XAD-2 (15% w/v) overnight at 5°C under 800 rpm shaking. To evaluate homogeneity, samples were subjected to SEC on a Superdex 200 Increase 10/300 GL column in h-SEC buffer (20 mM TrisHCl pH 7.5, 150 mM NaCl in H_2O).

Empty sNDs for SANS

An empty sND was prepared as above, but with d-csMSP1E3D1 and d-POPC and using the optimal ratio of 125:1. The sample was purified by SEC using a Superdex 200 Increase 10/300 GL column equilibrated in d-SEC buffer (20 mM TrisDCl pH 7.5, 150 mM NaCl in 100% D_2O). A 1 ml fraction from the peak was collected and split in two samples. One sample was measured as the 100% D_2O sample, the other was diluted to 6:4 in h-SEC buffer to a final concentration of 60% D_2O. The remaining samples in the contrast series were prepared from dilution of the 100% D_2O sample with h-SEC buffer. Before each dilution, 50 to 100 μL were kept for UV absorption measurements.

Optimization of CorA reconstitution in NDs

CorA reconstitution was also first optimized with hydrogenated material. Again, the standard protocol for nanodisc reconstitution was followed,[51] *i.e.* solubilization of POPC with sodium cholate followed by mixing with CorA and csMSP1E3D1 at a final concentration of 10 mM POPC and 20 mM cholate. First, the POPC:csMSP1E3D1 ratio was varied between 50:1 and 125:1 at a fixed CorA:csMSP1E3D1 ratio of 8:1. Second, the csMSP1E3D1:CorA ratio was varied between 2:1 and 8:1 at a fixed POPC:csMSP1E3D1 ratio of 100:1. Samples

were incubated for 1 h at 5°C with 300 rpm shaking and amberlite XAD-2 were added to 50% w/v and incubated overnight at 5°C with 800 rpm shaking. Samples were added NiNTA resin and the flow through containing empty NDs was discarded. CorA-loaded NDs were eluted in h-SEC buffer containing 250 mM imidazole and immediately subjected to SEC on a Superdex 200 Increase 10/300 GL column in h-SEC buffer (20 mM TrisHCl pH 7.5, 150 mM NaCl in H_2O).

CorA reconstitution in sNDs

CorA-loaded sNDs were prepared as above but with d-csMSP1E3D1 and d-POPC at an optimized d-POPC:d-csMSP1E3D1:CorA ratio of 400:4:1. After detergent removal by amberlite XAD-2 and IMAC purification, the sample was added TEV protease and dialyzed against TEV buffer (20 mM TrisHCl pH 8, 100 mM NaCl, 0.5 mM EDTA, 1 mM DTT) for 3 h at room temperature and continued overnight at 4°C. TEV protease and uncleaved CorA-loaded sNDs were removed by reverse NiNTA purification. The sample was concentrated using a 30,000 MWCO spin filter to approximately 30 μM.

Exchange into sDDM for cuvette SANS

For standard static SANS with offline SEC purification, samples were purified by SEC on site on a Superdex 200 Increase 10/300 column equilibrated in d-DDM buffer (20 mM TrisDCl pH 7.5, 150 mM NaCl, 0.5 mM sDDM in 100% D_2O), using a flow rate of 0.3 ml/min to allow efficient exchange.[30] The collected peak fraction was split in two: One sample was added 1 mM EDTA from a 50 mM stock in d-DDM buffer, and the other sample was added 40 mM Mg^{2+} from a 2 M stock in d-DDM buffer.

UV absorption measurements

Sample concentrations were estimated from UV absorption at 280 nm measured on a Nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) in triplicates.

Molar extinction coefficients were calculated from the primary sequences online.

Reconstitution in large unilamellar vesicles (LUVs)

10 mg POPC in chloroform was dried under a stream of N_2. LUVs were prepared by dissolving the lipid film in reconstitution buffer (10 mM MOPS-KOH pH 7.2, 150 mM KCl, 100 μM EGTA) including 10 μM Mag-Fluo-4 to a POPC concentration of 15 mg/ml. The sample was estruded 41 times through a 0.2 μm filter, yielded a homogeneous sample. To reconstitute CorA in the LUVs, a solution was mixed to contain 10.5 mg/ml POPC, 2 μM CorA, 10 μM Mag-Fluo-4, and 50 mM OG and vortexed briefly. For empty LUVs (mock), CorA-SEC buffer was added instead. Biobeads SM-2 were added to 45% w/v and incubated with rotation for 30 min at room temperature. The sample solution was separated from Biobeads by centrifugation and applied to a Sephadex G50 (Sigma-Aldrich, Germany) column equilibrated in gel buffer (10 mM MOPS-KOH pH 7.2, 150 N-methyl-D-glucamine-HCl (NMDG-Cl), 100 μM EGTA). The sample was allowed to slowly enter the gel before centrifuged for 7.5 min at 180g. CorA-loaded LUVs were collected in the flow through.

Small-angle scattering

For all of the presented SAS data, the scattering vector has the relation $q = 4\pi \cdot sin(2\theta)/\lambda$, where θ is the scattering angle. Scattering intensity, $I(q)$, on absolute scale is given by the differential scattering cross section per unit volume $\frac{d\Sigma}{d\Omega}(q)$ in units of cm^{-1}.

SANS

The nanodisc contrast series was measured at the QUOKKA instrument at ANSTO, Lucas Heights, Australia[52] with a sample to detector distance of 8 m giving a q-range of 0.009 $Å^{-1}$ - 0.1 $Å^{-1}$ using a neutron wavelength of 5 Å±10% FWHM. Samples were measured at 10°C. Scattering intensity was brought to absolute scale using the direct beam intensity measured with an attenuator in place.

Standard static SANS on CorA in sDDM micelles were measured at the KWS-1 instrument at the Forschungs-Neutronenquelle Heinz Maier-Leibnitz (FRM II) in Munich, Germany, using three sample-to-detector distances, of 2 m, 4 m, and 8 m giving a q-range of 0.006 $Å^{-1}$ - 0.27 $Å^{-1}$ using a neutron wavelength of 5 Å±10% FWHM. Samples were measured at 10°C. Scattering intensity was brought to absolute scale using plexiglass as secondary calibration standard.

Samples of CorA in sDDM micelles and sNDs were measured at the D22 instrument at the ILL, Grenoble, France using the newly developed SEC-SANS setup.[34] sND samples were run on a superose6 Increase 10/300 GL column (GE) in reducing d-SEC buffer (20 mM TrisDCl pH 7.5, 150 mM NaCl, 1 mM DTT in 100% D_2O), whereas sDDM samples were run on a Superdex 200 Increase 10/300 GL column in reducing d-DDM buffer (20 mM TrisDCl pH 7.5, 150 mM NaCl, 1 mM DTT, 0.5 mM sDDM in 100% D_2O). For sNDs, an initial flow rate of 0.5 ml/min was reduced to 0.05 ml/min, when the sample peak eluted to optimize the measurement time. For sDDM, an initial flow rate of 0.3 ml/min was applied to allow efficient exchange into sDDM and reduced to 0.05 ml/min upon sample elution. For each sample, two SEC-SANS runs were performed to cover two sample to detector distances of 11.2 m and 2 m, respectively, giving a q-range of 0.0044 $Å^{-1}$ - 0.42 $Å^{-1}$, using a neutron wavelength of 5 Å±10% FWHM. The SEC-SANS setup was cooled to approximately 11°C. Scattering intensity was brought to absolute scale using the direct beam intensity measured with an attenuator in place.

SEC-SANS concentration estimation

In SEC-SANS, the absorption signal at 280 nm saturated during sample elution. Thus, to estimate concentration, the absorption at 300 nm was used to estimate concentration in stead. This was done by assuming that the A_{280}/A_{300} ratio was constant throughout the sample peak.

SAXS

SEC-SAXS was done at BM29 at the ESRF in Grenoble, France. An empty sND sample was run on a superose 6 10/300 GL column in reducing h-SEC buffer, using a flow rate of 0.7 ml/min. 1 s SAXS frames were recorded and the beam intensity was attenuated to 41% to avoid radiation damage. The scattering signal was brought to absolute scale using water as a secondary calibration standard.[53]

Data analysis

Bayesian Inverse Fourier Transform (IFT) was carried out online at the bayesapp server available at www.bayesapp.org to obtain $p(r)$-distributions.[42] For SAXS data, the transformation/regularization parameter was set to negative. For calculating the theoretical $p(r)$-distributions of PDB structures, the program CaPP[54] was used with the water layer density setting set to 6% and only applied for the regions of the protein in contact with water as determined by the OPM database.[55]

PDB structures were fitted to SANS data using CRYSON from the ATSAS package[41] or using WillItFit?.[56] For analysis by WillItFit?, PDB files were added Hs and Ds in Readyset in the PHENIX package.[57]

The SAXS data were fitted with the nanodisc model in WillItFit?.[38] The height of belt parameter was fixed to 25.78 Å as determined from the high resolution structure of a small nanodisc.[58] A concentration scale factor was also fitted given the weak concentration estimation from the UV cell in the HPLC. Besides the parameters presented, a surface roughness was was fitted. The nanodisc model is described in detail elsewhere.[38] Furthermore, a slightly modified version has been used to model the hydrogenated versions of the sNDs used in the present study study.[37]

Negative stain EM

CorA was purified by SEC in h-SEC and d-SEC buffer, respectively. The collected peak fractions were split in two and added EDTA to 1 mM or Mg^{2+} to 40 mM from stock solutions prepared in h-SEC and d-SEC buffer, respectively. Samples were diluted to 0.1 μM in appropriate buffers containing 1 mM EDTA or 40 mM Mg^{2+}.

Cupper grids were neutralized with a Easiglow glow discharger (Agar Scientific, Standsted, UK). 3 μL of sample was applied to the grid and incubated for 30 s. The grid was blotted onto a filter paper from the edge, and 3 μL 2% uranyal formate was added immediately and incubated for 30 s. The staining procedure was repeated two more times. After the final staining, the grid was left to dry for ten minutes. EM data were acquired on a Tecnai TEM (FEI, Thermo Fischer scientific) at Aarhus University, Denmark.

For data processing, particles were picked in xmipp.[59] 2D class averages and 3D models were refined in Relion.[60]

Activity assays

For activity assays, 20 μL LUV samples were diluted in 1000 μL - x μL of measurement buffer (gel buffer containing 10 μM valinomycin) and incubated for 1 h at 10°C. For measurements in presence of the inhibitor $Co[NH_3]_6^{3+}$, the LUVs were diluted in measurement buffer containing 1 mM of the inhibitor. For the experiments in D_2O, measurement buffer and ligand stocks were prepared in D_2O.

CorA activity was monitored by Mag-Fluo-4 fluorescence at 515 nm (excitation at 488 nm) on a FluoroMax fluorometer (Horiba, Kisshoin, Japan) upon addition of x μL metal ion stock. The fluorescence data was normalized to the flat part of the signal recorded before adding metal ion and the initial 100 s of data collect after addition was fitted by linear regression to estimate initial activity rates.

Acknowledgement

We thank the staff at BM29 at the ESRF for allocated SAXS beam time and support. Here, special thanks goes to Marta Brennich and Mark Tully. We also thank the staff at KWS-1 at FRM2, D22 at ILL and QUOKKA at ANSTO for allocation of SANS beam time and support. Special thanks goes to Henrich Frielinghaus at KWS-1, Anne Martel and Lionel Porcar at D22, and Elliot Gilbert at QUOKKA.

The work was funded by the Lundbeck Foundation Brainstruc programme, the Novo Nordisk Foundation Synergy project, and BioSynergy KU2016, and the authors gratefully acknowledge the support.

Supporting Information Available

Supporting info available: Optimization of nanodisc reconstitution (Figure S1).

References

(1) Jahnen-Dechent, W.; Ketteler, M. Magnesium basics. *Clin. Kidney J.* **2012**, *5*, i3–i14.

(2) Rude, R. K. Magnesium Deficiency: A Cause of Heterogenous Disease in Humans. *J. Bone Miner. Res.* **1998**, *13*, 749–758.

(3) Romani, A. M. Cellular magnesium homeostasis. *Arch. Biochem. Biophys.* **2011**, *512*, 1–23.

(4) Kehres, D. G.; Maguire, M. E. Structure, properties and regulation of magnesium transport proteins. *Biometals* **2002**, *15*, 261–70.

(5) Maguire, M. E. Magnesium transporters: properties, regulation and structure. *Front. Biosci.* **2006**, *11*, 3149–63.

(6) Smith, R. L.; Maguire, M. E. Distribution of the CorA Mg2+ transport system in gram-negative bacteria. *J. Bacteriol.* **1995**, *177*, 1638–40.

(7) Smith, R. L.; Gottlieb, E.; Kucharski, L. M.; Maguire, M. E. Functional similarity between archaeal and bacterial CorA magnesium transporters. *J. Bacteriol.* **1998**, *180*, 2788–91.

(8) Kucharski, L. M.; Lubbe, W. J.; Maguire, M. E. Cation Hexaammines Are Selective and Potent Inhibitors of the CorA Magnesium Transport System. *J. Biol. Chem.* **2000**, *275*, 16767–16773.

(9) Payandeh, J.; Li, C.; Ramjeesingh, M.; Poduch, E.; Bear, C. E.; Pai, E. F. Probing Structure-Function Relationships and Gating Mechanisms in the CorA Mg 2+ Transport System. *J. Biol. Chem.* **2008**, *283*, 11721–11733.

(10) Dalmas, O.; Cuello, L. G.; Jogini, V.; Cortes, D. M.; Roux, B.; Perozo, E. Structural Dynamics of the Magnesium-Bound Conformation of CorA in a Lipid Bilayer. *Structure* **2010**, *18*, 868–878.

(11) Svidová, S.; Sponder, G.; Schweyen, R. J.; Djinović-Carugo, K. Functional analysis of the conserved hydrophobic gate region of the magnesium transporter CorA. *Biochim. Biophys. Acta - Biomembr.* **2011**, *1808*, 1587–1591.

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

(13) Dalmas, O.; Sompornpisut, P.; Bezanilla, F.; Perozo, E. Molecular mechanism of Mg2+-dependent gating in CorA. *Nat. Commun.* **2014**, *5*, 3590.

(14) Dalmas, O.; Sandtner, W.; Medovoy, D.; Frezza, L.; Bezanilla, F.; Perozo, E. A repulsion mechanism explains magnesium permeation and selectivity in CorA. *Proc. Natl. Acad. Sci.* **2014**, *111*, 3002–3007.

(15) Knoop, V.; Groth-Malonek, M.; Gebert, M.; Eifler, K.; Weyand, K. Transport of magnesium and other divalent cations: evolution of the 2-TM-GxN proteins in the MIT superfamily. *Mol. Genet. Genomics* **2005**, *274*, 205–216.

(16) Moomaw, A. S.; Maguire, M. E. Cation Selectivity by the CorA Mg 2+ Channel Requires a Fully Hydrated Cation. *Biochemistry* **2010**, *49*, 5998–6008.

(17) Guskov, A.; Nordin, N.; Reynaud, A.; Engman, H.; Lundback, A.-K.; Jong, A. J. O.; Cornvik, T.; Phua, T.; Eshaghi, S. Structural insights into the mechanisms of Mg2+ uptake, transport, and gating by CorA. *Proc. Natl. Acad. Sci.* **2012**, *109*, 18459–18464.

(18) Palombo, I.; Daley, D. O.; Rapp, M. Why Is the GMN Motif Conserved in the CorA/Mrs2/Alr1 Superfamily of Magnesium Transport Proteins? *Biochemistry* **2013**, *52*, 4842–4847.

(19) Dudev, T.; Lim, C. Importance of Metal Hydration on the Selectivity of Mg 2+ versus Ca 2+ in Magnesium Ion Channels. *J. Am. Chem. Soc.* **2013**, *135*, 17200–17208.

(20) Nordin, N.; Guskov, A.; Phua, T.; Sahaf, N.; Xia, Y.; Lu, S.; Eshaghi, H.; Eshaghi, S. Exploring the structure and function of Thermotoga maritima CorA reveals the mechanism of gating and ion selectivity in Co 2+ /Mg 2+ transport. *Biochem. J.* **2013**, *451*, 365–374.

(21) Lunin, V. V.; Dobrovetsky, E.; Khutoreskaya, G.; Zhang, R.; Joachimiak, A.; Doyle, D. A.; Bochkarev, A.; Maguire, M. E.; Edwards, A. M.; Koth, C. M. Crystal structure of the CorA Mg2+ transporter. *Nature* **2006**, *440*, 833–837.

(22) Eshaghi, S. Crystal Structure of a Divalent Metal Ion Transporter CorA at 2.9 Angstrom Resolution. *Science (80-.).* **2006**, *313*, 354–357.

(23) Payandeh, J.; Pai, E. F. A structural basis for Mg2+ homeostasis and the CorA translocation cycle. *EMBO J.* **2006**, *25*, 3762–3773.

(24) Matthies, D.; Dalmas, O.; Borgnia, M. J.; Dominik, P. K.; Merk, A.; Rao, P.; Reddy, B. G.; Islam, S.; Bartesaghi, A.; Perozo, E.; Subramaniam, S. Cryo-EM Structures of the Magnesium Channel CorA Reveal Symmetry Break upon Gating. *Cell* **2016**, *164*, 747–756.

(25) Tsuchiya, Y.; Nakamura, H.; Kinoshita, K. Discrimination between biological interfaces and crystal-packing contacts. *Adv. Appl. Bioinform. Chem.* **2008**, *1*, 99–113.

(26) Egelman, E. H. The Current Revolution in Cryo-EM. *Biophys. J.* **2016**, *110*, 1008–1012.

(27) Sali, A. et al. Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* **2015**, *23*, 1156–1167.

(28) Cleverley, R. M.; Kean, J.; Shintre, C. A.; Baldock, C.; Derrick, J. P.; Ford, R. C.; Prince, S. M. The Cryo-EM structure of the CorA channel from Methanocaldococcus jannaschii in low magnesium conditions. *Biochim. Biophys. Acta - Biomembr.* **2015**, *1848*, 2206–2215.

(29) Franke, D.; Svergun, D. I. DAMMIF , a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **2009**, *42*, 342–346.

(30) Midtgaard, S. R. et al. Invisible detergents for structure determination of membrane proteins by small-angle neutron scattering. *FEBS J.* **2018**, *285*, 357–371.

(31) Maric, S.; Skar-Gislinge, N.; Midtgaard, S.; Thygesen, M. B.; Schiller, J.; Frielinghaus, H.; Moulin, M.; Haertlein, M.; Forsyth, V. T.; Pomorski, T. G.; Arleth, L. Stealth carriers for low-resolution structure determination of membrane proteins in solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2014**, *70*, 317–328.

(32) Josts, I.; Nitsche, J.; Maric, S.; Mertens, H. D.; Moulin, M.; Haertlein, M.; Prevost, S.; Svergun, D. I.; Busch, S.; Forsyth, V. T.; Tidow, H. Conformational States of ABC

Transporter MsbA in a Lipid Environment Investigated by Small-Angle Scattering Using Stealth Carrier Nanodiscs. *Structure* **2018**, *26*, 1072–1079.e4.

(33) Zaccai, N. R.; Sandlin, C. W.; Hoopes, J. T.; Curtis, J. E.; Fleming, P. J.; Fleming, K. G.; Krueger, S. In *Methods Enzymol.*; Kelman, Z., Ed.; Academic Press, 2016; Vol. 566; Chapter Methods in, pp 159–210.

(34) Jordan, A.; Jacques, M.; Merrick, C.; Devos, J.; Forsyth, V. T.; Porcar, L.; Martel, A. SEC-SANS: size exclusion chromatography combined in situ with small-angle neutron scattering. *J. Appl. Crystallogr.* **2016**, *49*, 2015–2020.

(35) Yepuri, N. R.; Darwish, T. A.; Krause-Heuer, A. M.; Leung, A. E.; Delhom, R.; Wacklin, H. P.; Holden, P. J. Synthesis of Perdeuterated 1-Palmitoyl-2-oleoyl- sn -glycero-3-phosphocholine ([D 82]POPC) and Characterisation of Its Lipid Bilayer Membrane Structure by Neutron Reflectometry. *Chempluschem* **2016**, *81*, 315–321.

(36) Dunne, O.; Weidenhaupt, M.; Callow, P.; Martel, A.; Moulin, M.; Perkins, S. J.; Haertlein, M.; Forsyth, V. T. Matchout deuterium labelling of proteins for small-angle neutron scattering studies using prokaryotic and eukaryotic expression systems and high cell-density cultures. *Eur. Biophys. J.* **2017**, *46*, 425–432.

(37) Johansen, N. T.; Tidemand, F. G.; Nguyen, T.; Rand, K. D.; Pedersen, M. C.; Arleth, L. Circularized and solubility-enhanced MSPs facilitate simple and high yield production of stable nanodiscs for studies of membrane proteins in solution. *Manuscript* **2018**,

(38) Skar-Gislinge, N.; Simonsen, J. B.; Mortensen, K.; Feidenhans'l, R.; Sligar, S. G.; Lindberg Møller, B.; Bjørnholm, T.; Arleth, L. Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. *J. Am. Chem. Soc.* **2010**, *132*, 13713–13722.

(39) Huda, P.; Binderup, T.; Pedersen, M. C.; Midtgaard, S. R.; Elema, D. R.; Kjær, A.;

Jensen, M.; Arleth, L. PET/CT Based In Vivo Evaluation of 64Cu Labelled Nanodiscs in Tumor Bearing Mice. *PLoS One* **2015**, *10*, e0129310.

(40) Johansen, N. T.; Pedersen, M. C.; Porcar, L.; Martel, A.; Arleth, L. Introducing SEC–SANS for studies of complex self-organized biological systems. *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74*.

(41) Svergun, D. I.; Richard, S.; Koch, M. H.; Sayers, Z.; Kuprin, S.; Zaccai, G. Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 2267–72.

(42) Hansen, S. Update for BayesApp : a web site for analysis of small-angle scattering data. *J. Appl. Crystallogr.* **2014**, *47*, 1469–1471.

(43) Volkov, V. V.; Svergun, D. I. Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **2003**, *36*, 860–864.

(44) Scheres, S. H. W.; Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **2012**, *9*, 853.

(45) Nasr, M. L.; Baptista, D.; Strauss, M.; Sun, Z.-Y. J.; Grigoriu, S.; Huser, S.; Plückthun, A.; Hagn, F.; Walz, T.; Hogle, J. M.; Wagner, G. Covalently circularized nanodiscs for studying membrane proteins and viral entry. *Nat. Methods* **2017**, *14*, 49–52.

(46) Frey, L.; Lakomek, N.-A.; Riek, R.; Bibow, S. Micelles, Bicelles, and Nanodiscs: Comparing the Impact of Membrane Mimetics on Membrane Protein Backbone Dynamics. *Angew. Chemie Int. Ed.* **2017**, *56*, 380–383.

(47) Steiner, T. The Hydrogen Bond in the Solid State. *Angew. Chemie Int. Ed.* **2002**, *41*, 48–76.

(48) Sugiyama, T.; Yoshiok, T. In *Protein Struct.*; Faraggi, E., Ed.; InTech, 2012; pp 291–308.

(49) Gati, C.; Stetsenko, A.; Slotboom, D. J.; Scheres, S. H. W.; Guskov, A. The structural basis of proton driven zinc transport by ZntB. *Nat. Commun.* **2017**, *8*, 1313.

(50) Nasr, M. L.; Wagner, G. Covalently circularized nanodiscs; challenges and applications. *Curr. Opin. Struct. Biol.* **2018**, *51*, 129–134.

(51) Ritchie, T.; Grinkova, Y.; Bayburt, T.; Denisov, I.; Zolnerciks, J.; Atkins, W.; Sligar, S. In *Methods Enzymol.*; Düzgünes, N., Ed.; Academmic Press, 2009; Vol. 464; Chapter Methods in, pp 211–231.

(52) Wood, K. et al. QUOKKA, the pinhole small-angle neutron scattering instrument at the OPAL Research Reactor, Australia: design, performance, operation and scientific highlights. *J. Appl. Crystallogr.* **2018**, *51*, 294–314.

(53) Orthaber, D.; Bergmann, A.; Glatter, O. SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J. Appl. Crystallogr.* **2000**, *33*, 218–225.

(54) Larsen, A. github.com/Niels-Bohr-Institute-XNS-StructBiophys/CaPP. 2018.

(55) Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. OPM: orientations of proteins in membranes database. *Bioinformatics* **2006**, *22*, 623–5.

(56) Pedersen, M. C.; Arleth, L.; Mortensen, K. WillItFit : a framework for fitting of constrained models to small-angle scattering data. *J. Appl. Crystallogr.* **2013**, *46*, 1894–1898.

(57) Adams, P. D. et al. PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213–221.

(58) Bibow, S.; Polyhach, Y.; Eichmann, C.; Chi, C. N.; Kowal, J.; Albiez, S.; McLeod, R. A.; Stahlberg, H.; Jeschke, G.; Güntert, P.; Riek, R. Solution structure of discoidal high-density lipoprotein particles with a shortened apolipoprotein A-I. *Nat. Struct. Mol. Biol.* **2017**, *24*, 187–193.

(59) de la Rosa-Trevín, J.; Otón, J.; Marabini, R.; Zaldívar, A.; Vargas, J.; Carazo, J.; Sorzano, C. Xmipp 3.0: An improved software suite for image processing in electron microscopy. *J. Struct. Biol.* **2013**, *184*, 321–328.

(60) Scheres, S. H. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **2012**, *180*, 519–530.

Supporting Information

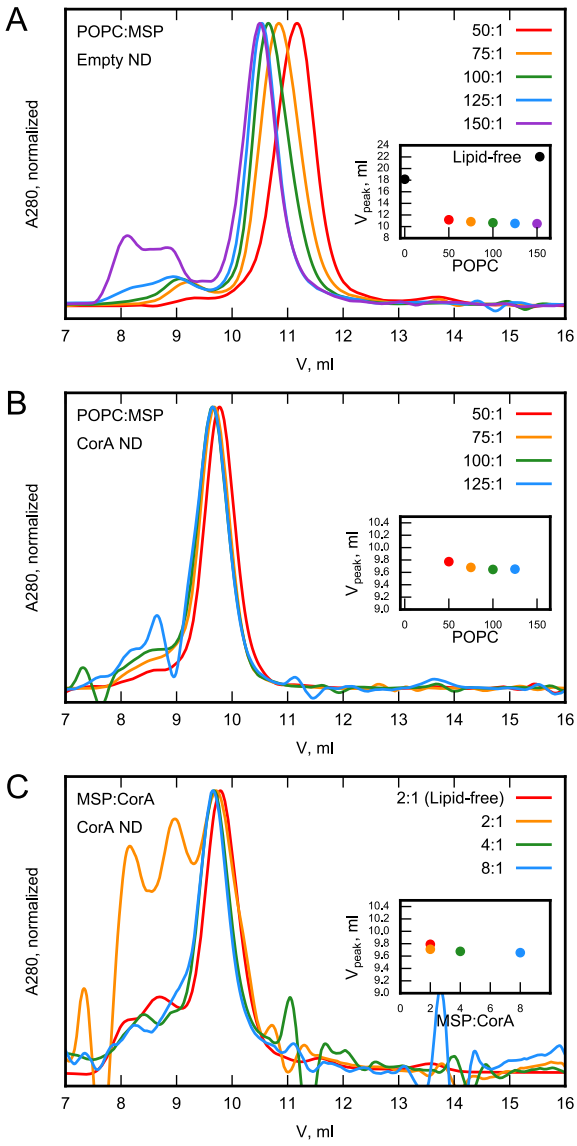


Figure S1: Optimization of nanodisc reconstitution. A: Nanodiscs prepared with varying POPC:csMSP1E3D1 ratios. Free csMSP1E3D1 eluted at 18 ml, which is only showed in the inset. **B:** CorA-loaded nanodiscs prepared with varying POPC:csMSP1E3D1 ratios and with a constant csMSP1E3D1:CorA ratio of 8:1. **C:** CorA-loaded nanodiscs prepared with varying csMSP1E3D1:CorA ratios and with a constant POPC:csMSP1E3D1 ratio of 100:1. One sample was prepared without POPC (lipid-free). In all panels, the insets show the peak elution volumes as a function of reconstitution ratios.

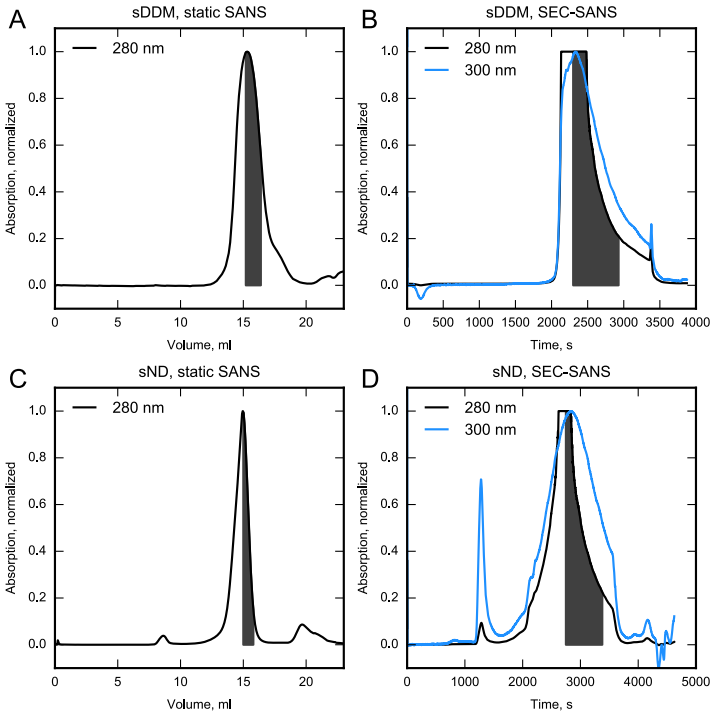


Figure S2 CorA SEC profiles. A: Chromatogram (280 nm) of CorA in sDDM purified with offline SEC for static SANS measurement. **B:** Chromatograms (280 nm and 300 nm) of CorA in sDDM in SEC-SANS measurement. **C:** Chromatogram (280 nm) of CorA in sNDs purified with offline SEC for static SANS measurement. **D:** Chromatograms (280 nm and 300 nm) of CorA in sNDs in in SEC-SANS measurement. The collected/measured fractions are indicated by the colored areas under the curves.

8

REPORT I

HEAVY WATER EFFECT ON PROTEINS AND NANODISCS?

Heavy water effect on proteins and nanodiscs?

Nicolai Tidemand Johansen[1]

[1]Niels Bohr Institute, University of Copenhagen, Copenhagen, 2100, Denmark

ABSTRACT

Heavy water, D_2O, is used as solvent or co-solvent for many biophysical and structural methods such as mass spectrometry, nuclear magnetic resonance spectroscopy, and small-angle neutron scattering (SANS). Although the D atom is equal in size to the H atom, its mass is double due to an extra neutron in the nuclei. This has an effect on the bonding energy between D and other atoms and can *e.g.* affect hydrogen-bonds in water. Increased aggregation propensity for biological samples has been observed in heavy water, which is particularly problematic in SANS, where aggregates have a large impact on the obtained scattering signal. This has been a bottleneck for preparation of monodisperse nanodisc samples for SANS. It has previously been suggested that decreased stability in heavy water was the main driver of this aggregation. In this present study, the heavy water effect is investigated on a set of small soluble proteins using differential scanning calorimetry and circular dichroism spectroscopy. Furthermore, the stability of nanodisc samples in heavy water is probed using light scattering techniques and size-exclusion chromatography. This study shows that protein stability is increased in heavy water and that as long as samples are buffer exchanged using desalting or gel filtration columns, monodisperse and stable nanodiscs can be obtained in heavy water. This implies that decreased stability is not the driver of aggregation in this solvent. Instead, low solubility is proposed as the main driver of aggregation. To address this, it is shown that negatively charged nanodiscs can be prepared, which can serve as a general strategy to improve solubility.

Introduction

Several biophysical and structural methods make us of hydrogen/deuterium (H/D) substitution or exchange (H/D-X) to report on dynamics or mask signal. These include mass spectrometry (MS), nuclear magnetic resonance (NMR) spectroscopy, and small-angle neutron scattering (SANS). In SANS, changing the D content of the buffer or the molecules under investigation is routinely used for contrast-variation studies to mask specific molecules in multi-components complexes to obtain a more detailed view of the organization of individual parts.[1] The nanodisc is a discoidal particle containing a central lipid bilayer core surrounded by two copies of so-called membrane scaffold proteins (MSP)[2,3] that can be used to stabilize membrane proteins in solution[4]. Thus, the nanodisc is an example of a multi-component system for which it is desired to be able to perform contrast-variation. Furthermore, using D_2O as solvent has the advantage that it provides little incoherent background in SANS compared to H_2O.[5]

However, in previous SANS and small-angle X-ray scattering (SAXS) studies, it has been observed that D_2O can be detrimental to the quality of nanodisc samples.[6,7] Using a standard preparation method of serial concentration and dilution in D_2O buffers in all cases yielded samples containing aggregates. Notably, this preparation method did not yield aggregated samples in H_2O buffers, which suggests a "heavy water effect" (HWE) on nanodisc samples. This is problematic for SANS studies because the complex models needed for describing biological structures are highly dependent on good quality and aggregation-free data together with contrast variation to yield meaningful parameters. Thus the increased aggregation propensity of some samples in D_2O is an important challenge to overcome.

Two observations make it clear that D_2O is different from H_2O. First, D_2O has a freezing point of 3.8°C,[8] and second, high levels of D_2O are toxic to living organisms,[9] although simple organisms can be adapted to survive.[10] Several studies have investigated the HWE for a number of bio-molecules, such as proteins[11–20], peptides[21], carbohydrates[22], and lipids.[23] Whereas perdeuteration decreases protein stability in D_2O[12,24], the general trend for standard hydrogenated proteins is that their stability is actually significantly increased in this D_2O.[20] Protein solubility, on the other hand, has in all cases been found to correlate inversely[13,17] and has been ascribed to increased attractive intermolecular interactions in D_2O.[16] The Ubbelohde effect[25] is the phenomena that the hydrogen bond length increases upon deuteration. Apart from changing the properties of water, this is also expected to affect protein structures, particularly secondary structure, which is stabilized by hydrogen bonds. The hydrogen bond forming chemical groups in proteins will eventually exchange D for H in D_2O solvents.

Here, the HWE is probed on a set of small soluble proteins as well as a number of nanodisc samples. For the small proteins, the HWE is quantified by determination of transition melting temperatures (T_M) as determined by differential scanning calorimetry (DSC) or circular dichroism spectroscopy (CD) in buffers of varying D_2O content. These studies are then followed by longer, quantitative studies reporting on the ability of nanodiscs to remain homogenous over time in D_2O as compared to H_2O, using static and dynamic light scattering (SLS and DLS, respectively) combined with size-exclusion chromatography

(SEC). Overall, these studies show that protein stability is enhanced by D_2O and that nanodiscs aggregation is not a problem if the samples are prepared by mild buffer exchange into D_2O. This suggest that issues with aggregation in D_2O could rather be due to decreased solubility. Based on this observation, a more negatively charged MSPs is designed and shown to make nanodiscs, which presumably are more soluble than standard nanodiscs.

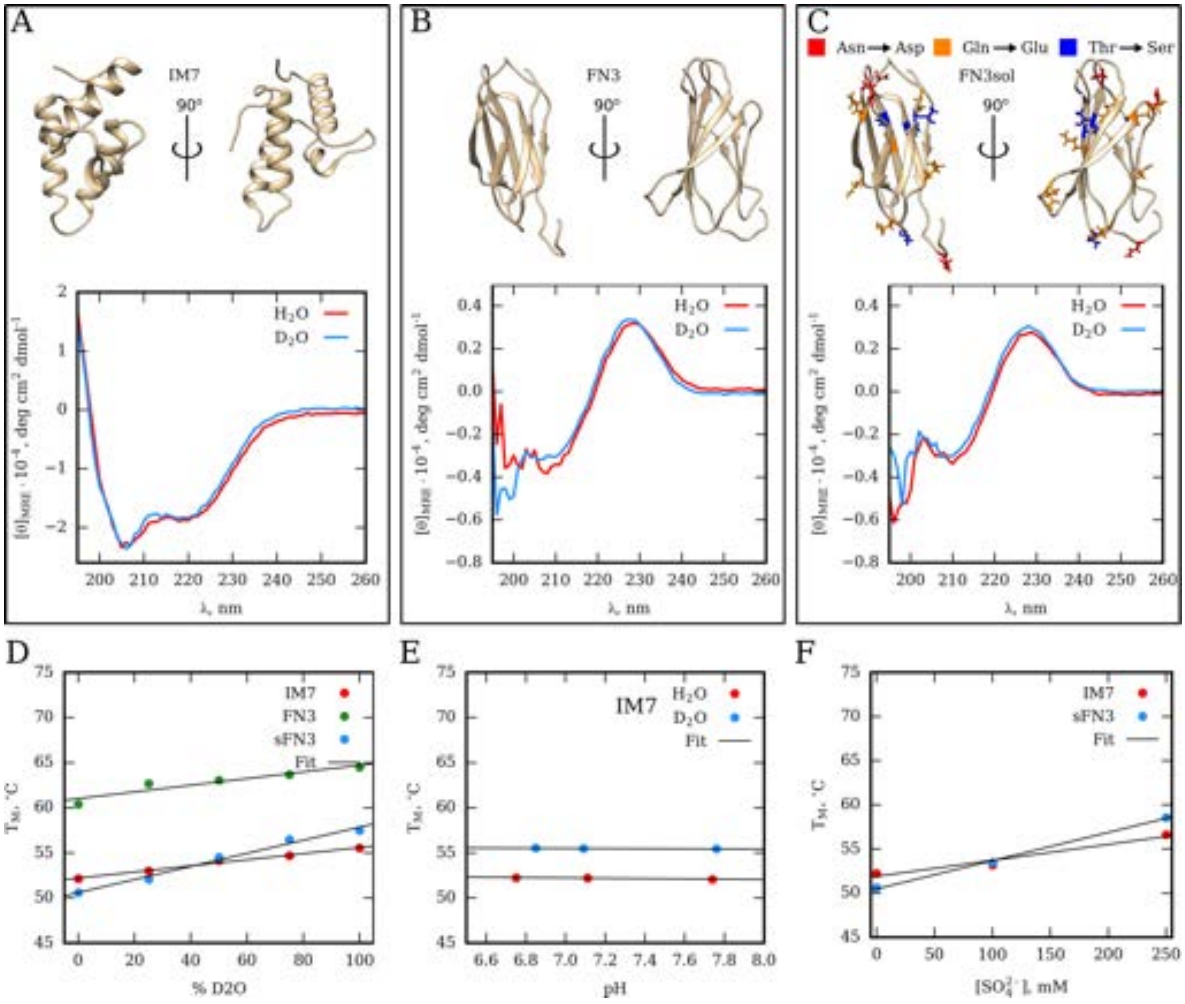


Figure 1 Heavy water effect on proteins. A: Crystal structure of IM7 (PDB ID: 1AYI) shown in ribbon model along with far-UV CD spectra obtained for samples in H_2O and D_2O solvents, respectively. **B:** Solution structure of FN3 obtained by NMR (PDB ID: 2KBG) as well as far-UV CD spectra as in A. **C:** Same as B, where mutated residues in FN3 have been colored according to the legend, along with far-UV CD spectra of the resulting mutant, sFN3. The CD spectra were measured in 20 mM sodium phosphate pH 7, 100 mM NaCl buffer. **D:** Melting temperatures (T_M) of the three proteins as a function of D_2O content in the same buffer as for CD. The data follow a linear trend (Fit, solid line). **E:** IM7 T_M measured for different pH, all in 20 mM sodium phosphate, 100 mM NaCl. Linear fits are shown to emphasize the almost identical T_M values. **F:** T_M measured for IM7 and sFN3 in 20 mM sodium phosphate pH 7, 100 mM NaCl with varying concentrations of sodium sulfate added.

Results

Stability of small, soluble proteins in D_2O

To probe the HWE on relatively simple samples, a set of small and water-solible proteins were investigated in D_2O. Two major scaffolds, one purely α-helix, the other purely β-sheet, were chosen to probe the HWE on secondary structure. The α-helical protein immunity collicin 7 (IM7) is shown in Figure 1A, and the β-sheet protein fibronectin-like domain III (FN3) is shown in Figure 1B. The third protein was a mutated FN3 (sFN3), with all Asn, Gln, and Thr mutated to Asp, Glu, and Ser, respectively, as shown in Figure 1C.

These mutations are expected to enhance protein solubility[26,27] and lower the approximate pI to 5.39 from 7.14 in the wild type, which makes it feasible to conduct experiments at neutral pH.

For measurements in D_2O, protein samples were buffer exchanged 24 h in advance to ensure complete H/D-X with the solvent. Initially, protein structures were verified by far-UV CD. The data show characteristics of α-helix for IM7 (Figure 1A) and characteristic spectrums for FN3 (Figure 1B) and sFN3 (Figure 1C) due to ordered aromatic amino acids.[28] For all three proteins, the spectra obtained in H_2O and D_2O are highly similar, indicating that the respective protein structures are retained in D_2O. Furthermore, the severe level of mutation did not appear to have affected the overall structure of sFN3 compared to the wild type FN3 as judged by their highly similar CD spectra.

The stabilities of IM7 and sFN3 were probed by DSC, whereas the stability of FN3 was probed by CD because this protein aggregated in the calorimeter before a sufficient enthalpogram could be obtained (data not shown). T_M, as given by the transition midpoint between the folded and unfolded state, were extracted as a proxy for protein stability. For all proteins, the T_M increases linearly with the D_2O content (Figure 1D) and has an overall change of 3°C to 7°C between D_2O and H_2O, respectively. The slopes of the T_M plots vary between the three proteins, with sFN3 having the steepest response to D_2O, and IM7 having the flattest response. Notably, FN3 has a markedly higher T_M than sFN3 throughout the D_2O series but with a flatter response to he D_2O content (Figure 1D). The thermodynamic parameters extracted from the DSC and CD data are given in Table 1.

The correction pD = pH + 0.4 is often suggested for D_2O, but recent research show that the correction is much smaller and buffer dependent.[29] Such correction is also expected to be difficult for protein samples in D_2O, where similar effects on the ionization state of protein side chains could potentially impact the net effect of the buffer. Here, the pH is simply reported as the value measured on a pH-meter without applying corrections. To make sure that the increase in T_M was not an effect of pH/pD, IM7 was measured in a pH-range in steps of roughly 0.4 units (Figure 1E). Identical T_M values are obtained over a range of 1 pH unit both in H_2O and D_2O . The increase of stability stability in D_2O is thus not ascribed to altered protonation states of protein side chains, at least for IM7.

Two-state folding of globular proteins is governed by the hydrophobic effect[30] and is a highly cooperative process. The stability is related to the accessible surface area of the unfolded protein that is buried upon folding, and it is thus clear that solvation is a major determinant for protein stability.[31] It has been shown that some ions, such as sulfate (SO_4^{2-}), can increase the surface tension of water, which leads to protein stabilization.[32] Furthermore, sulfate is at the same time known to decrease protein solubility.[33] These effects are qualitatively similar to the HWE on small proteins,[13,20] and thus, IM7 and sFN3 were investigated in presence of sulfate. Addition of sulfate indeed increased the stabilities of IM7 and sFN3 (Figure 1F). Notably, the increases in stability had linear trends similar to the ones for stabilization in D_2O (Figure 1A). It was also attempted to measure the solubility of the three proteins in H_2O and D_2O by sulfate-induced precipitation. However, the proteins were for all practical purposes too soluble (>50 mg/ml) to achieve precipitation with up to 2M Na_2SO_4 added (data not shown).

	T_M, °C					T_M, °C	
% D_2O	IM7	FN3	sFN3	MSP1D1	$[SO_4^{2-}]$, mM	IM7	sFN3
0	52.2	60.4	50.6	56.3	0	52.2	50.6
25	53.0	62.7	52.0	-	100	53.2	53.4
50	54.1	63.1	54.5	-	250	56.6	58.6
75	54.7	63.7	56.7	-			
100	55.5	64.5	57.5	61.2			
ΔT_M	3.3	4.1	6.9	4.9		4.4	8.0

Table 1. T_M values corresponding to the plots in Figure 1 and Figure 4.

Stability of nanodiscs in D_2O

The HWE was of special concern regarding nanodisc samples, which have proven extremely difficult to obtain homogeneous in D_2O. It was previously suggested that the method of buffer exchange could be the key to obtain better samples.[7] The low D_2O-consuming method of sequential concentration and dilution using spin concentration columns was found to be the main driver of aggregation, presumably because nanodiscs in the process are in a very high local concentration in the vicinity of the membrane.[6] Thus, only mild methods for buffer exchange, such as desalting or SEC, were applied in this study to prepare samples of nanodiscs made from the MSP1D1 belt protein[2] and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipids.

Structural integrity over time monitored by SEC

To test structural homogeneity of nanodiscs over time, SEC purified nanodisc samples were buffer-exchanged on small desalting columns into different H_2O or D_2O buffers and incubated at three different temperatures (Figure 2A). At specific time points, 50 μL of sample was gel filtrated in a standard SEC buffer (Figure 2A, right panel).

The initial SEC purification shows that the main peak in the chromatogram is slightly skewed and with a small shoulder on the left (Figure 2A left panel). Peaks with this appearance will be referred to as homogenous nanodiscs in the following. Samples incubated in SEC buffer at 4°C remain almost completely homogenous over the course of 12 days, whereas at 20°C, degradation sets in somewhere between one and 12 days (Figure 2B, left and middle panels). At 40°C, severe degradation is visible already at the first time points of 6 hours (Figure 2B, right panel). This overall picture is identical for samples that were incubated in SEC buffer with 5% w/v glycerol added (Figure 2C) or in a 100 mM phosphate buffer (Figure 2D). For some samples, the aggregation appears more aggressive in D_2O after 12 days (Figure 2B, 40°C and Figure 2D, 20°C). However, after that long storage, other effects than simple

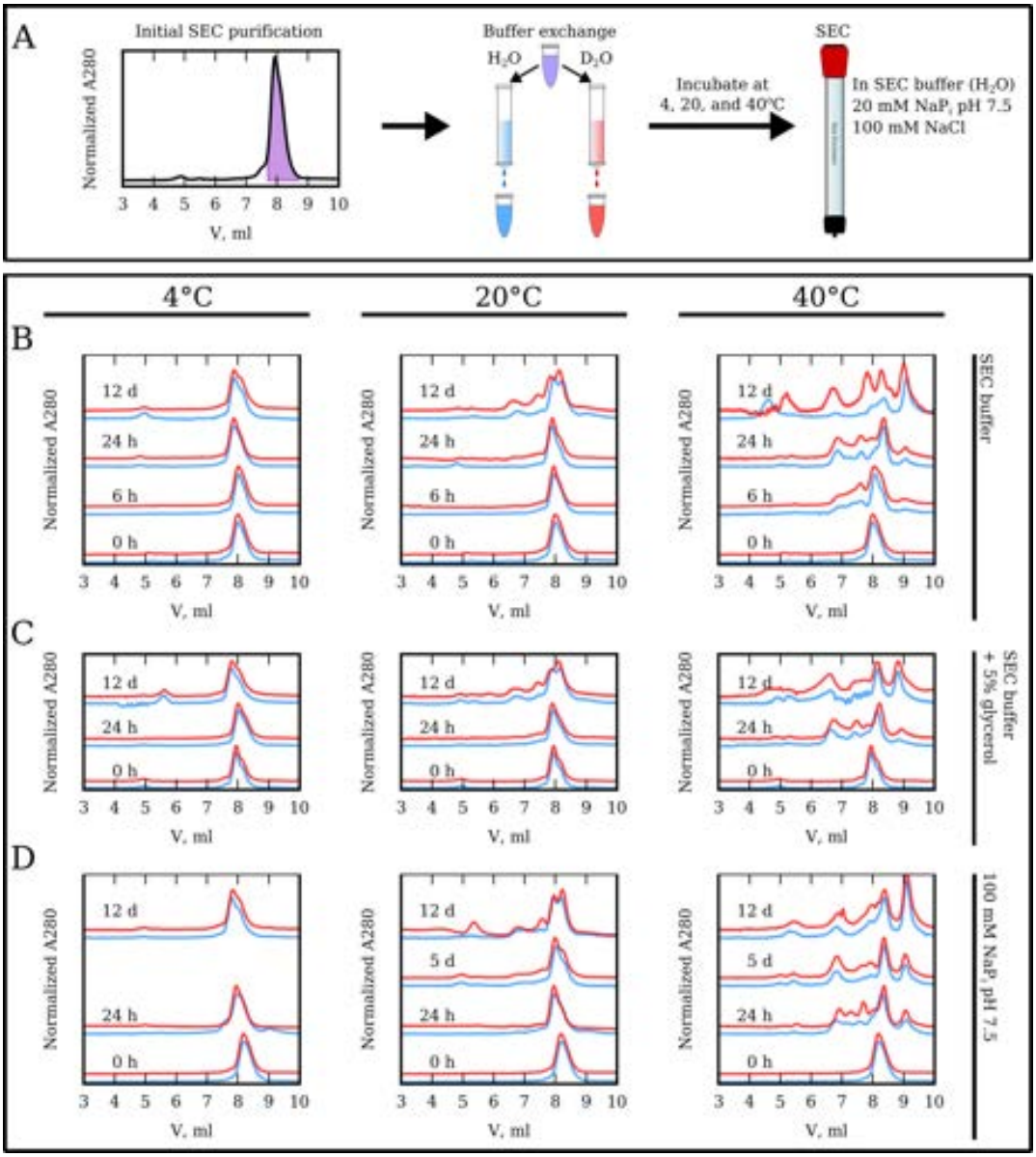


Figure 2 Nanodisc stability monitored by SEC. A: Overview of the experimental procedure. Initially, a large nanodisc sample was purified. This was exchanged into various buffers in H_2O (blue) and D_2O (red) using NAP-5 desalting columns. Note the color of the H_2O and D_2O samples, which serve as legend for the obtained SEC data. These samples were incubated at various temperatures and at given time points evaluated by SEC, using a Bio SEC-3 column with a void volume of 5 ml. **B**, **C**, and **D:** SEC chromatograms for samples incubated in SEC buffer, SEC buffer with 5% v/v glycerol added, or 100 mM sodium phosphate buffer, respectively, as indicated.

nanodisc aggregation, such as hydrolytic degradation of the lipids in the nanodisc, are expected to be prevalent and make assessment of the data complicated.

For all samples incubated at 4°C for 12 d, small peaks are observed in the column void volume at approximately 5 ml. However, no major intermediary shoulder peaks to the left of the main nanodisc peak are observed as is the case for the degraded samples at 20°C and especially at 40°C. For these degraded nanodiscs, the intermediary peaks are accompanied by a peak to the right of the nanodisc peak, which is

likely to be free MSP. Notably, all of these observations are similar for samples incubated in H_2O and D_2O buffers, respectively.

Combining SEC with SLS

From the data presented above, nanodisc stability does not seem to be compromised in D_2O. To gain a more detailed insight to the temporal disintegration of the nanodisc, samples were measured continuously with SLS. To probe an effect of increased solubility, resembling the introduction of negatively charged amino acids in sFN3, samples with varying amount of negatively charged 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG) lipids were measured. Furthermore, 1% mol/mol of fluorescent 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl) (RhodPE) lipids were added to nanodiscs made from POPC to clarify whether the larger species monitored by SEC after incubation at 40°C in Figure 2 contained lipid or were purely aggregates of the MSP1D1 belt protein.

Samples were prepared as above (Figure 2A) in SEC buffer in H_2O and D_2O, filtered into the sample cuvette to eliminate dust particles and degassed to avoid formation of air bubbles. SLS was then continuously measured over the course of 40 hours. The first observation is that the sample behavior is overall similar in H_2O (Figure 3A) and D_2O (Figure 3B). In both solvents at 20°C, the normalized SLS signals lie at around 1 throughout the measurement, whereas at 40°C, the signal has a significant and systematic slope, indicating the appearance of larger particles. The large signal deviations of the POPC:RhodPE nanodiscs around 10 hours to 20 hours are attributed to experimental artifacts as the signals return to 1 at the end of the experiment.

As introduction of POPG lipids had little observable effect at the two temperatures, further analysis was only done for the POPC:RhodPE samples. Size-distributions were calculated from the DLS data that were recorded simultaneously with the SLS data (Figure 3C). At the beginning of the experiment at 20°C, the nanodisc sample contained a major fraction of particles with the expected hydrodynamic radius of a nanodisc of approximately 8 nm as well as a broad population of species with much higher hydrodynamic radii around 80 nm (Figure 3C, black lines). Notably, DLS is very sensitive to large particles as the intensity is proportional to the particle volume squared,[34] and thus, these large particles only constitutes a minor part of the sample. This measurement serves as reference for comparison to size-distributions analyzed for different time points of the measurements at 40°C. An apparent systematic feature is seen in the gradual increase in size of the main peak at around 8 nm, both in H_2O and D_2O (Figure 3C), whereas the large population at around 80 nm had no such systematic variation. Overall, the data are similar for H_2O and D_2O.

The resolution of DLS is relatively low, and in general it is not expected that a particle less than three fold larger than another particle can be distinguished. SEC chromatograms obtained after the 40°C experiments for POPC:RhodPE samples show a set of overlapping peaks, which differ less than three fold in size (Figure 3D). Overall, the chromatograms are similar to those reported in Figure 2A (note that a

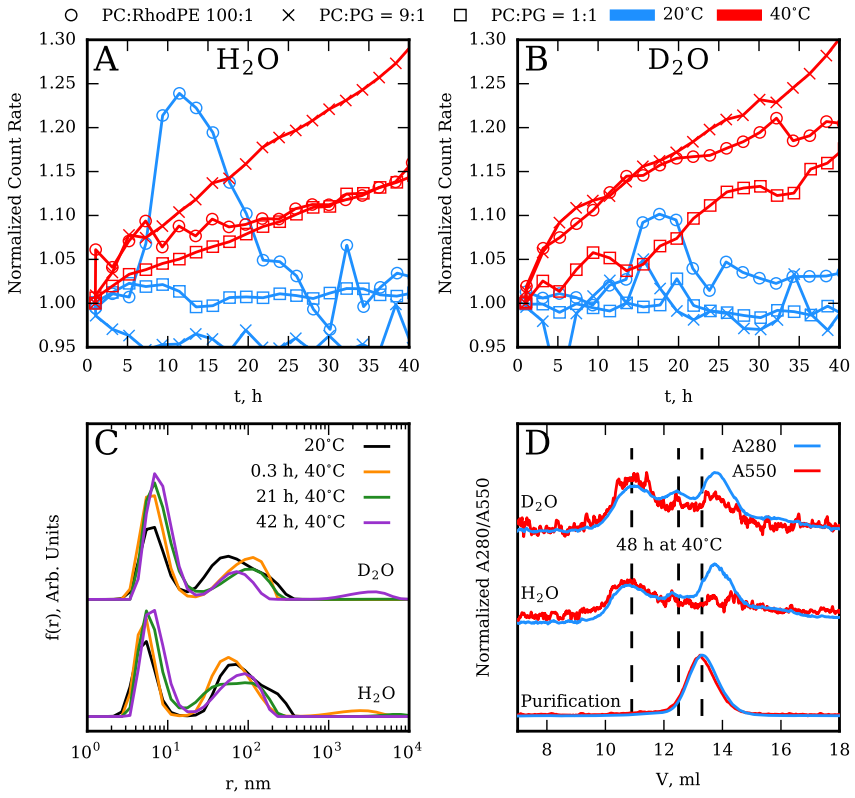


Figure 3 Nanodisc stability monitored by SLS and SEC. A and **B**: SLS measurements of nanodiscs with three different lipid compositions (PC = POPC, PG = POPG) at 20°C and 40°C. The signal was normalized to the starting value to follow the relative increase in signal. **C:** Size-distributions obtained from DLS frames obtained during the SLS measurements at the given time points. **D:** SEC chromatograms of the initial sample as well as the POPC:RhodPE (100:1) samples after the SLS measurements at 40°C. Dashed lines indicate the location of the three major peaks. The SEC column was a Superdex200 10/300 GL with a void volume of 8 ml.

different column was used). The fluorescent RhodPE lipid's absorption was monitored at 550 nm and showed that lipid was present in all three major peaks. This shows that other MSP1D1-lipid particles than the nanodisc are indeed formed at elevated temperatures. Again, the chromatograms were close to identical between the H_2O and D_2O samples.

Thermal stability of MSP1D1 and nanodiscs

MSP1D1 and nanodisc stability was measured by CD spectroscopy. The lipid-free MSP1D1 is α-helical in solution and can be reversibly unfolded by thermal denaturation (Figure 4A). At 95°C, the maximum temperature in the unfolding experiment, MSP1D1 appears to be largely disordered but with some residual α-helix (Figure 4A, crossed lines). The unfolding of MSP1D1 shows a broad transition, indicating low degree of cooperativity (Figure 4B). T_M is increased around 5°C in D_2O compared to H_2O as given by the transition midpoint (Table 1). Refolding of MSP1D1 was fully reversible with the CD signal closely overlaying that of the unfolding (data not shown).

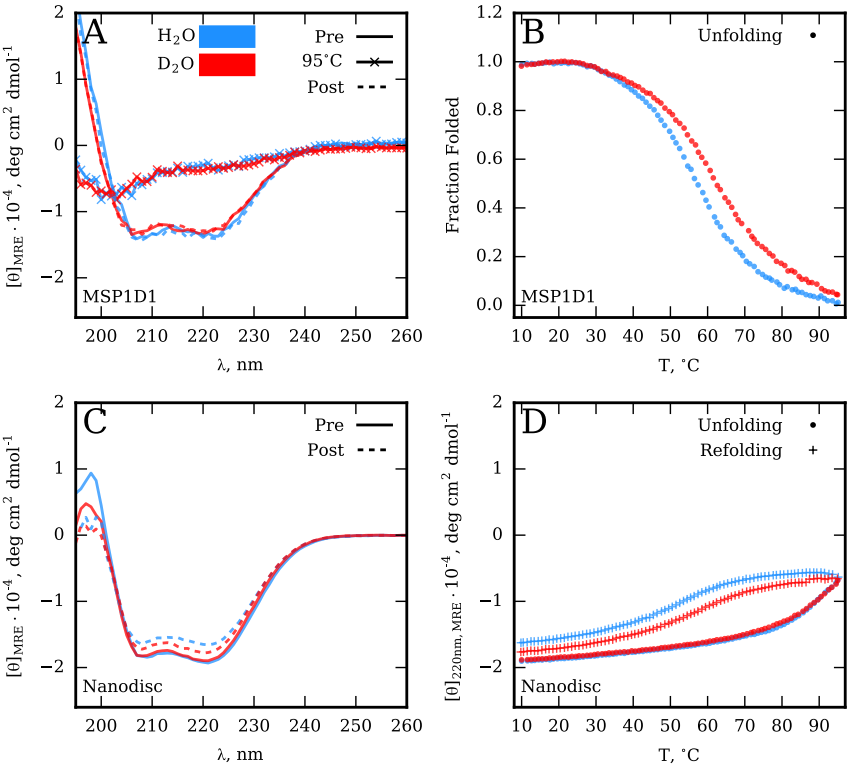


Figure 4 MSP and nanodisc stability monitored by CD. A: Far-UV CD spectra of MSP1D1 before, at, and after heating from 10°C to 95°C. **B:** Thermal denaturation followed by the CD signal at 220 nm and converted to fraction folded. **C:** Far-UV CD spectra of MSP1D1-POPC nanodiscs before and after heating from 10°C to 95°C. **D:** Thermal denaturation followed by the CD signal at 220 nm. The signal is not converted to fraction folded due to the lack of baselines at high temperature.

Nanodisc samples were studied by the same means as MSP1D1. First, it is noted that lipid-bound MSP1D1 in nanodiscs is more α-helical (Figure 4C, solid lines) than lipid-free MSP1D1 (Figure 4A, solid lines) as judged by the magnitude of the CD signal at 220 nm. Furthermore and in contrast to the lipid-free MSP1D1, the α-helix content of the nanodisc is reduced after thermal denaturation (Figure 4C, dotted lines). Nanodisc unfolding do not reach a plateau at high temperature, indicating incomplete denaturation (Figure 4D). Furthermore, the refolding do not overlay the unfolding, suggesting that the nanodisc structure is altered upon heating. This is in line with the nanodisc disintegration observed by SEC (Figure 2) and SLS (Figure 3). No T_M values were extracted due to the complex nanodisc disintegration scheme and the lack of a defined baseline at high temperature. Qualitatively, however, it is seen that the change in the magnitude of the CD signal is lower in D_2O compared to H_2O, possibly reflecting a slightly higher stability in D_2O.

Introducing solubility-enhancing mutations in MSP1D1

The above presented results all indicate that protein and nanodisc stability is higher in D_2O. Thus, aggregation induced by protein unfolding is unlikely to be the main problem for these samples in D_2O. Instead, the main problem might be related to decreased solubility. Therefore, a MSP was generated with

high abundance of negative charges (sMSP1D1), using the same amino acid substitutions as for sFN3 (Figure 1C), *i.e.* all Asn, Gln, Thr to Asp, Glu and Ser, respectively. Protein sequences are provided in the Methods section.

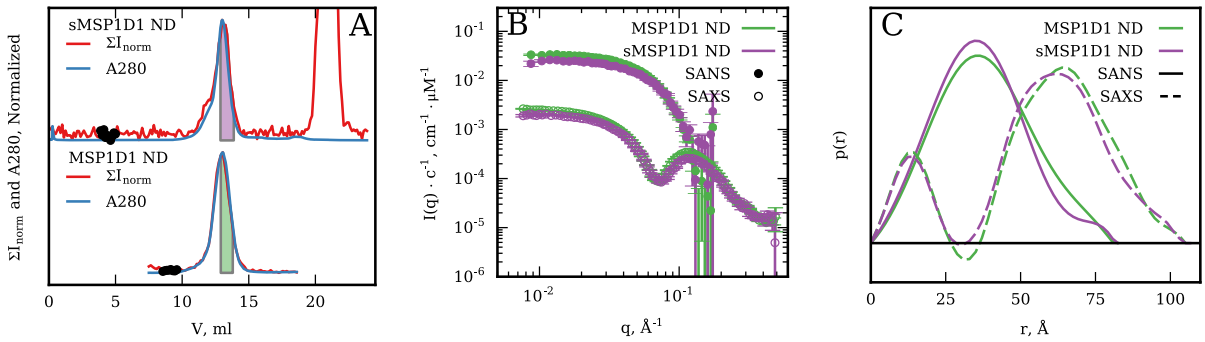


Figure 5. Nanodisc structure. A: Normalized chromatograms and scattergrams obtained by SEC-SANS (ND = nanodiscs). The column was a Superdex200 10/300 GL with a void volume of 7.5 m*l*. The colored area indicate the interval of frames selected for further analysis. **B:** Reduced and background subtracted SAXS and SANS data. For sMSP1D1 nanodiscs, the SAXS data set was collected in conventional static mode, whereas the remaining data sets were from SEC-SANS and SEC-SAXS. **C:** Pair-distance distributions ($p(r)$-distributions) refined from the SAXS and SANS data.

Nanodiscs could readily be prepared from sMSP1D1 and POPC lipids using the same protocol as for MSP1D1 and POPC: The SEC chromatograms of sMSP1D1 nanodiscs feature a single peak with a small shoulder on the left side and elutes at the same retention volume as MSP1D1 nanodiscs (Figure 5A). The structure of the sMSP1D1 nanodisc was probed by SAXS and SANS and compared to MSP1D1 nanodiscs (Figure 5B). Notably, the SANS data were for both samples obtained through online SEC coupled to SANS (SEC-SANS),[35] whereas online SEC coupled to SAXS (SEC-SAXS) was only available for the MSP1D1 nanodisc. At first glance, the scattering pattern of sMSP1D1 nanodiscs are highly similar to those of MSP1D1 nanodiscs (Figure 5B). Pair-distance distribution functions, $p(r)$-distributions, are consequently also quite similar (Figure 5C). The characteristic sizes of the nanodisc samples refined from the $p(r)$-distributions are shown in Table 2 and confirm that the samples are overall identical.

	MSP1D1 nanodisc		sMSP1D1 nanodisc	
	SAXS	SANS	SAXS	SANS
R_g, Å	43.6	29.2	42.7	27.8
D_{max}, Å	102.1	79.5	101.1	80.1

Table 2. Nanodisc sizes. Characteristic sizes obtained from the $p(r)$-distriubtions in Figure 5C. R_g is the radius of gyration, D_{max} is the maximum pair-distance.

The solubility of sMSP1D1 was not probed quantitatively, but during purification, there was no visible precipitation in the dialysis after elution from the affinity column, which is normally the case for MSP1D1. This serves a qualitative indication that higher solubility has been achieved for sMSP1D1.

Discussion

Soluble proteins of mainly α-helix (IM7 and MSP1D1) or β-sheet (FN3 and sFN3) were shown to retain their structures when exchanged into D_2O but showed an increase in T_M (Figure 1 and 4). Going from H_2O to D_2O, there was a shift in T_M in the range of 3.3°C (IM7) to 6.9°C (sFN3). These increases are in line with the literature.[20] Enthalpies were also extracted from the fits to the data, but very similar values were obtained for the different conditions for each protein, respectively (data not shown). Taking T_M as a proxy for stability, D_2O was concluded to increase protein conformational stability.

In an attempt to mimic the increase in stability and presumed decrease in solubility[13,17], T_M was measured for the two negatively charged proteins IM7 and sFN3 in presence of sulfate (Figure 1F). A similar slope of T_M was found with sulfate compared to D_2O (Figure 1D). However, IM7 and sFN3 were too soluble to readily carry out protein precipitation experiments within practical limits. Other types of experiments could provide insight to the degree of molecular attraction, such as obtaining second virial coefficients from SAXS. These data will be pursued in the near future to make valuable correlations between gain of stability and (assumed) loss of solubility.

Obtaining monodisperse nanodisc samples in D_2O was the initial goal of this study. Using gentle buffer exchange methods, such as gel filtration, this is readily achieved. The data obtained in this report suggest that stability is not an issue. Nanodiscs show identical behavior when incubated in H_2O and D_2O buffers over long periods of time (Figure 2 B, C and D). On short time scales, there are also no observable differences (Figure 3). At 40°C, nanodiscs disintegrate into larger particles. As evaluated by SLS and SEC, this process is also highly similar in H_2O and D_2O (Figure 3 A and B). Finally, the CD data indicate a very modest stabilization of nanodiscs in D_2O as monitored by the relative absolute signal change (Figure 4 C and D). Overall, stability and homogeneity does not seem compromised in D_2O.

The reason that concentration in spin filters produce aggregates must be related to solubility. As for the small proteins, it was attempted to measure nanodisc solubility, but with little luck. Furthermore, generating a concentrated stock in D_2O would not be feasible given that spin filters ruin the samples. Instead, a more immediate solution is proposed by generating more negatively charged MSPs for increasing solubility. The absence of protein aggregation during purification suggest that this property was achieved, although it was not quantified experimentally. Nanodiscs can be readily formed with sMSP1D1 and have overall similar structure and size to nanodiscs made with MSP1D1 (Figure 5).

In conclusion, the problems associated with obtaining monodisperse nanodisc samples in D_2O can be readily overcome by gentle buffer exchange, such as gel filtration. Compared to buffer exchange in spin filters, gel filtration requires high consumption of D_2O, which is expensive. However, so is *e.g.* SANS beam time. It is therefore recommended to utilize gel filtration in D_2O for sample preparation. If the concentration from a gel filtration peak fraction is not adequate, it might be worth going for solubility-enhanced nanodiscs prepared with sMSP1D1. However, it remains to be tested whether these can tolerate harsh buffer exchange methods, such as serial concentration with spin filters and dilution in D_2O.

Methods

Sample preparation

All salts were obtained from Sigma-Aldrich (Merck), Darmstadt, Germany. 99% D_2O was obtained from Cambridge analytical laboratories, Tewksbury, MA, USA.

Protein sequences

IM7

MELKNSISDYTEAEFVQLLKEIEKENVAATDDVLDVLLEHFVKITEHPDGTDLIYYPSDNRDDSPE
GIVKEIKEWRAANGKPGFKQGHHHHHH

FN3

MGREPSPPSIHGQPSSGKSFKLSITKQDDGGAPILEYIVKYRSKDKEDQWLEKKVQGNKDHIILEH
LQWTMGYEVQITAANRLGYSEPTVYEFSMPPKPNIIKDHHHHHH

sFN3

MGREPSPPSIHGEPSSGKSFKLSITKEDDGGAPILEYIVKYRSKDKEDEWLEKKVEGDKDHIILEH
LEWTMGYEVEITAADRLGYSEPTVYEFSMPPKPDIIKDHHHHHH

sMSP1D1

MGHHHHHHHHDYDIPTTENLYFQGSSFSKLREELGPVSEEFWDDLEKESEGLREEMSKDLEEVKAK
VEPYLDDFEKKWEEEMELYREKVEPLRAELEEGAREKLHELEEKLSPLGEEMRDRARAHVDALR
SHLAPYSDELRERLAARLEALKEDGGARLAEYHAKASEHLSSLSEKAKPALEDLREGLLPVLESFK
VSFLSALEEYSKKLDSE

MSP1D1

MGHHHHHHHHDYDIPTTENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK
VQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALR
THLAPYSDELRQRLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFK
VSFLSALEEYTKKLNTQ

Genes encoding for IM7, FN3, sFN3 and sMSP1D1, respectively, were custom synthesized and cloned into pET28a vectors by Genscript (Genscript, Piscataway, NJ, USA). The pET28a plasmid encoding MSP1D1 was obtained from professor Stephen Sligar (University of Illions, Urbana-Champaign IL, USA).

Protein expression

Plasmids were transformed into BL21 DE3 star (Novagen (Merck), Darmstadt, Germany) and plated on LB agar containing kanamycin. Single colonies were used to inoculate 1 l cultures of TB medium in 5 L baffled shaker flasks and grown at 37°C until an optical density (OD) at 600 nm of 0.8 had been reached. Protein expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Fisher Scientific, Hampton, NH, USA) and cultures were incubated at 37°C for 3 h (IM7, FN3, MSP1D1) or at 20°C for 16 h (sFN3). Cells were harvested by centrifugation (8 min at 8000g at 10°C).

For IM7, FN3 and sFN3, the cell pellets were resuspended in binding buffer (50 mM TrisHCl pH 8,

300 mM NaCl, 20 mM imidazole) and lysed using a cell disruptor at 30 kPSI (Constant Systems, Low March, Daventry, UK). Insoluble debris was pelleted by centrifugation (30 min at 25000*g* at 10°C) and the supernatant was mixed with NiNTA resin (Qiagen, Hilden, Germany) equilibrated in binding buffer and incubated for 45 min at 4°C. The mixture was added to a disposable column (BioRad, Hercules, CA, USA), and the resin was washed in three column volumes (CV) of wash buffer (50 mM TrisHCl pH 8, 300 mM NaCl, 40 mM imidazole). Protein was eluted from the resin in elution buffer (50 mM TrisHCl pH 8, 250 mM imidazole) and immediately transferred to a dialysis bag and submerged in at least a 100 fold larger volume of buffer A (20 mM TrisHCl pH 8). Dialysis was continued over night. In a final step, the proteins were purified by ion-exchange chromatography (IEC) on a 6 ml ressource Q column (GE Healthcare, Pittsburgh, PA, USA). Protein was bound in buffer A and later eluted using a step gradient with buffer B (20 mM TrisHCl pH 8, 1 M NaCl). The highly negative IM7 and sFN3 eluted around 30% buffer B, whereas FN3 was eluted with 10% buffer B. The purified protein was stored at -20°C.

For MSP1D1 and sMSP1D1, the cell pellet was resuspended in binding buffer containing 6 M guanidine hydrochlorid. The cells were lysed by magnetic stirring followed by vigorous shaking. Insoluble debris was pelleted by centrifugation (60 min at 25000*g* at 10°C) and the supernatant was mixed with NiNTA resin equilibrated in the denaturing binding buffer and incubated for 30 min at room temperature. The mixture was added to a disposable column and the resin was washed in three CV of detergent wash buffer (50 mM TrisHCl pH 8, 300 mM NaCl, 40 mM imidazole, 10 mM cholate) and subsequently in three CV of wash buffer. Protein was eluted from the resin in elution buffer and dialysed 50 fold against tris-buffer (20 mM TrisHCl pH 7.5, 100 mM NaCl) over night. The protein was flash frozen in liquid nitrogen and stored at -80°C.

Nanodisc preparation

When using lipid mixtures, the respective lipids were dissolved in chloroform and mixed. Chloroform vas evaporated under a stream of nitrogen followed by hard vacuum over night. Nanodiscs were prepared according to the standard protocol.[36] Lipids were dissolved in tris-buffer containing 100 mM cholate to a final lipid concentration of 50 mM. Lipid solution was mixed with purified MSP and tris-buffer to yield a final lipid:MSP ratio of 70:1. Samples were briefly incubated at room temperature before added amberlite XAD-2 (Sigma-Aldrich, St. Louis, MO, USA) and incubated for 4 h at 10°C. Monodisperse nanodiscs were purified by SEC using either a Bio SEC-3 7.8/300 column (Agilent, Santa Clara, CA, USA) or a Superdex200 10/300 GL column (GE Healthcare, Pittsburgh, PA, USA), using a common running buffer of 20 mM sodium phosphate pH 7.5, 100 mM NaCl, and flow rates of 1 ml/min and 0.5 1 ml/min for the two columns, respectively.

Buffer exchange

Buffer exchange was done on Illustra NAP-5 columns (GE Healthcare, Pittsburgh, PA, USA) following the manufacturer's instructions. However, in the elution from these columns, a smaller volume than specified was used to avoid overlap of the old buffer into the sample.

Experimental

DSC

DSC scans were obtained on a Microcal VP-DSC (Malvern, Malvern, UK). All buffers and samples were degassed for 15 min under mild vacuum. Buffer-buffer scans were measured until a stable signal had been obtained and the sample was loaded carefully to avoid bubble formation. Enthalpograms were recorded from 20°C to 95°C with a scan rate of 1°C/min and the feed back sensitivity set to low. DSC data were background subtracted, normalized with protein concentration and fitted using a 2-state model in the Origin software provided by the manufacturer. IM7 samples were measured in concentrations ranging from 260 μM to 290 μM, whereas sFN3 samples were measured in concentrations ranging from 80 μM to 96 μM.

SLS

After buffer exchange, samples were filtered three times through a 0.22 μm filter and into the sample cuvette, a round 1 cm wide quartz cuvette (Hellma, Müllheim, Germany). Before measurements, samples were degassed under mild vacuum for 15 min. SLS and DLS data were recorded on a BI-200SM light scattering instrument (Brookhaven Instruments, Holtsville, NY). Frames of 30 s were continuously recorded over indicated periods of time.

In SLS, the signal is proportional to the weight average molecular weight of the sample, as described elsewhere.[37,38] The DLS data were fitted using the *bayesapp* server for light scattering data as described in the literature.[39]

CD

CD spectra and thermal scans were recorded on a Jasco-815 (Jasco, Easton, MD, USA). Far-UV spectra were obtained in a wavelength range from 260 nm to 190 nm using a scan rate of 20 nm/min. Three spectra were recorded and subsequently averaged and background subtracted. Thermal scans were recorded with a scan rate of 1 °C/min and by monitoring the CD signal at 220 nm. The CD signal in millidegrees was converted to mean residue ellipticity by the relation

$$[\theta]_{MRE} = \frac{[\theta]_{mdeg}}{10 \cdot c \cdot N_{pb} \cdot l} \tag{1}$$

where c is the molar concentration of the sample, N_{pb} is the number of peptide bonds and l is the pathlength in cm.

The far-UV CD spectra were measured with the following concentrations: IM7 = 6 μM, FN3 = 16 μM, sFN3 = 16 μM, and MSP1D1 4 μM. The nanodisc samples were measured at 11.5 μM (nanodisc concentration), which was the same used for monitoring thermal denaturation. For FN3, the concentration was also identical in the thermal and spectral scans.

Thermal scans were fitted with a two-state model as described elsewhere.[40]

SAXS and SANS

SAXS data were recorded at BM29 at the ESRF, Grenoble, France. MSP1D1 nanodiscs were measured using the inline SEC-SAXS, whereas sMSP1D1 nanodiscs were measured in static mode using the robotic sample changer.[41] Scattering intensity, $I(q)$ was measured as function of scattering vector, $q = 4\pi\sin(\theta)\lambda^{-1}$, where θ is half the scattering angle and λ is the X-ray wavelength. With $\lambda = 0.992$ Å, a q-range of 0.003 Å^{-1} to 0.5 Å^{-1} was obtained. $I(q)$ was brought to absolute scale in units of cm^{-1} with water as a calibration standard.[42] Samples were measured at 10°C.

SEC-SANS was measured at D22 at ILL, Grenoble, France.[35] The specifications were identical those those described elsewhere.[43] With $\lambda = 5$ Å$\pm$10% FWHM and a sample-to-detector distance of 5.6 m, a q-range of 0.0087 Å^{-1} to 0.1 Å^{-1} was obtained. $I(q)$ was absolute calibrated with the direct beam flux.

$p(r)$-distributions were calculated using the *bayesapp* server for small-angle scattering data.[44]

References

1. Zaccai, N. R. *et al.* Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures and Releases Unfolded Outer Membrane Proteins. In Kelman, Z. (ed.) *Methods Enzymol.*, vol. 566, chap. Methods in, 159–210, DOI: 10.1016/bs.mie.2015.06.041 (Academic Press, 2016). 15334406.

2. Bayburt, T. H., Grinkova, Y. V. & Sligar, S. G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2**, 853–856, DOI: 10.1021/nl025623k (2002).

3. Skar-Gislinge, N. *et al.* Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. *J. Am. Chem. Soc.* **132**, 13713–13722, DOI: 10.1021/ja1030613 (2010). NIHMS150003.

4. Denisov, I. G. & Sligar, S. G. Nanodiscs in Membrane Biochemistry and Biophysics. *Chem. Rev.* **117**, 4669–4713, DOI: 10.1021/acs.chemrev.6b00690 (2017). 15334406.

5. Jacrot, B. The study of biological structures by neutron scattering from solution. *Reports Prog. Phys.* **39**, 911–953, DOI: 10.1088/0034-4885/39/10/001 (1976).

6. Midtgaard, S. R. Self-assembling peptide and protein nanodiscs for studies of membrane proteins. *PhD Thesis* (2013).

7. Huda, P. Studying membrane protein structure and function using nanodiscs. *PhD Thesis* (2015).

8. la Mer, V. K. & Baker, W. N. The Freezing Point of Mixtures of H 2 O and D 2 O. The Latent Heat of Fusion of D 2 O. *J. Am. Chem. Soc.* **56**, 2641–2643, DOI: 10.1021/ja01327a030 (1934).

9. Kushner, D. J., Baker, A. & Dunstall, T. G. Pharmacological uses and perspectives of heavy water and deuterated compounds. *Can. J. Physiol. Pharmacol.* **77**, 79–88, DOI: 10.1139/y99-005 (1999).

10. Dunne, O. *et al.* Matchout deuterium labelling of proteins for small-angle neutron scattering studies using prokaryotic and eukaryotic expression systems and high cell-density cultures. *Eur. Biophys. J.* **46**, 425–432, DOI: 10.1007/s00249-016-1186-2 (2017).

11. Antonino, L. C., Kautz, R. A., Nakano, T., Fox, R. O. & Fink, A. L. Cold denaturation and 2H2O stabilization of a staphylococcal nuclease mutant. *Proc. Natl. Acad. Sci.* **88**, 7715–7718, DOI: 10.1073/pnas.88.17.7715 (1991).

12. Makhatadze, G. I., Clore, G. M. & Gronenborn, A. M. Solvent isotope effect and protein stability. *Nat. Struct. Biol.* **2**, 852–855, DOI: 10.1038/nsb1095-852 (1995).

13. Gripon, C. *et al.* Lysozyme solubility in H2O and D2O solutions: a simple relationship. *J. Cryst. Growth* **177**, 238–247, DOI: 10.1016/S0022-0248(96)01077-9 (1997).

14. Kuhlman, B. & Raleigh, D. P. Global analysis of the thermal and chemical denaturation of the N-terminal domain of the ribosomal protein L9 in H 2 O and D 2 O. Determination of the thermodynamic parameters, Δ H, Δ S, and Δ C p , and evaluation of solvent isotope effects. *Protein Sci.* **7**, 2405–2412, DOI: 10.1002/pro.5560071118 (1998).

15. Chakrabarti, G., Kim, S., Gupta, M. L., Barton, J. S. & Himes, R. H. Stabilization of Tubulin by Deuterium Oxide. *Biochemistry* **38**, 3067–3072, DOI: 10.1021/bi982461r (1999).

16. Budayova-Spano, M., Lafont, S., Astier, J.-P., Ebel, C. & Veesler, S. Comparison of solubility and interactions of aprotinin (BPTI) solutions in H2O and D2O. *J. Cryst. Growth* **217**, 311–319, DOI: 10.1016/S0022-0248(00)00495-4 (2000).

17. Legrand, L., Rosenman, I., Boué, F. & Robert, M.-C. Effect of the substitution of light by heavy water on lysozyme KCl and NaNO3 solubility. *J. Cryst. Growth* **232**, 244–249, DOI: 10.1016/S0022-0248(01)01157-5 (2001).

18. Cioni, P. & Strambini, G. B. Effect of Heavy Water on Protein Flexibility. *Biophys. J.* **82**, 3246–3253, DOI: 10.1016/S0006-3495(02)75666-X (2002).

19. Shi, Z., Olson, C. A., Kallenbach, N. R. & Sosnick, T. R. D/H Amide Isotope Effect in Model α-Helical Peptides. *J. Am. Chem. Soc.* **124**, 13994–13995, DOI: 10.1021/ja027740n (2002).

20. Efimova, Y. M., Haemers, S., Wierczinski, B., Norde, W. & van Well, A. A. Stability of globular proteins in H2O and D2O. *Biopolymers* **85**, 264–273, DOI: 10.1002/bip.20645 (2007).

21. Cho, Y. *et al.* Hydrogen Bonding of β-Turn Structure Is Stabilized in D 2 O. *J. Am. Chem. Soc.* **131**, 15188–15193, DOI: 10.1021/ja9040785 (2009).

22. Cardoso, M. V., Carvalho, L. V. & Sabadini, E. Solubility of carbohydrates in heavy water. *Carbohydr. Res.* **353**, 57–61, DOI: 10.1016/j.carres.2012.03.005 (2012).

23. Beranová, L. *et al.* Effect of heavy water on phospholipid membranes: experimental confirmation of molecular dynamics simulations. *Phys. Chem. Chem. Phys.* **14**, 14516, DOI: 10.1039/c2cp41275f (2012).

24. Meilleur, F., Contzen, J., Myles, D. A. A. & Jung, C. Structural stability and dynamics of hydrogenated and perdeuterated cytochrome P450cam (CYP101). *Biochemistry* **43**, 8744–8753, DOI: 10.1021/bi049418q (2004).

25. Steiner, T. The Hydrogen Bond in the Solid State. *Angew. Chemie Int. Ed.* **41**, 48–76, DOI: 10.1002/1521-3773(20020104)41:1<48::AID-ANIE48>3.0.CO;2-U (2002).

26. Trevino, S. R., Scholtz, J. M. & Pace, C. N. Amino Acid Contribution to Protein Solubility: Asp, Glu, and Ser Contribute more Favorably than the other Hydrophilic Amino Acids in RNase Sa. *J. Mol. Biol.* **366**, 449–460, DOI: 10.1016/j.jmb.2006.10.026 (2007). NIHMS150003.

27. Kramer, R. M., Shende, V. R., Motl, N., Pace, C. N. & Scholtz, J. M. Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility. *Biophys. J.* **102**, 1907–1915, DOI: 10.1016/j.bpj.2012.01.060 (2012).

28. Yang, L. *et al.* Expression, refolding and spectroscopic characterization of fibronectin type III (FnIII)-homology domains derived from human fibronectin leucine rich transmembrane protein (FLRT)-1, -2, and -3. *PeerJ* **5**, e3550, DOI: 10.7717/peerj.3550 (2017).

29. Rubinson, K. A. Practical corrections for p(H,D) measurements in mixed H 2 O/D 2 O biological buffers. *Anal. Methods* **9**, 2744–2750, DOI: 10.1039/C7AY00669A (2017).

30. Baldwin, R. L. & Rose, G. D. How the hydrophobic factor drives protein folding. *Proc. Natl. Acad. Sci.* **113**, 12462–12466, DOI: 10.1073/pnas.1610541113 (2016).

31. Privalov, P. L. & Crane-Robinson, C. Role of water in the formation of macromolecular structures. *Eur. Biophys. J.* **46**, 203–224, DOI: 10.1007/s00249-016-1161-y (2017).

32. Tadeo, X., López-Méndez, B., Castaño, D., Trigueros, T. & Millet, O. Protein Stabilization and the Hofmeister Effect: The Role of Hydrophobic Solvation. *Biophys. J.* **97**, 2595–2603, DOI: 10.1016/j.bpj.2009.08.029 (2009).

33. Wingfield, P. Protein Precipitation Using Ammonium Sulfate. In *Curr. Protoc. Protein Sci.*, vol. 421, A.3F.1–A.3F.8, DOI: 10.1002/0471140864.psa03fs13 (John Wiley & Sons, Inc., Hoboken, NJ, USA, 1998). 15334406.

34. Øgendal, L. *Light Scattering: A brief introduction*. May (2015).

35. Jordan, A. *et al.* SEC-SANS: size exclusion chromatography combined in situ with small-angle neutron scattering. *J. Appl. Crystallogr.* **49**, 2015–2020, DOI: 10.1107/S1600576716016514 (2016).

36. Ritchie, T. *et al.* Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs. In Düzgünes, N. (ed.) *Methods Enzymol.*, vol. 464, chap. Methods in, 211–231, DOI: 10.1016/S0076-6879(09)64011-8 (Academmic Press, 2009). NIHMS150003.

37. Høiberg-Nielsen, R., Fuglsang, C. C., Arleth, L. & Westh, P. Interrelationships of Glycosylation and Aggregation Kinetics for Peniophora lycii Phytase. *Biochemistry* **45**, 5057–5066, DOI: 10.1021/bi0522955 (2006).

38. Midtgaard, S. R. *et al.* Self-assembling peptides form nanodiscs that stabilize membrane proteins. *Soft Matter* **10**, 738–752, DOI: 10.1039/C3SM51727F (2014).

39. Hansen, S. DLSanalysis.org: a web interface for analysis of dynamic light scattering data. *Eur. Biophys. J.* **47**, 179–184, DOI: 10.1007/s00249-017-1232-8 (2018).

40. Greenfield, N. J. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* **1**, 2527–2535, DOI: 10.1038/nprot.2006.204 (2007).

41. Pernot, P. *et al.* Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664, DOI: 10.1107/S0909049513010431 (2013).

42. Orthaber, D., Bergmann, A. & Glatter, O. SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J. Appl. Crystallogr.* **33**, 218–225, DOI: 10.1107/S0021889899015216 (2000).

43. Johansen, N. T., Pedersen, M. C., Porcar, L., Martel, A. & Arleth, L. Introducing SEC-SANS for studies of complex self-organised biological systems. *Accepted* 1–38 (2018).

44. Hansen, S. Update for BayesApp : a web site for analysis of small-angle scattering data. *J. Appl. Crystallogr.* **47**, 1469–1471, DOI: 10.1107/S1600576714013156 (2014).