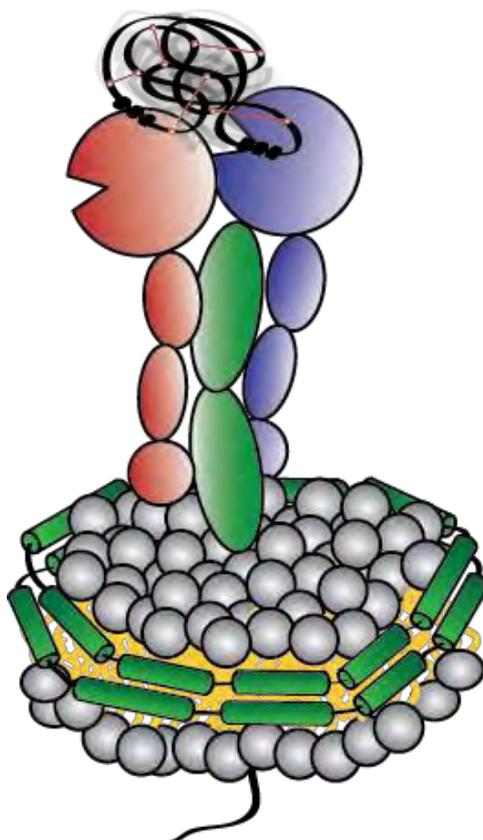




PhD Thesis  
Frederik Grønbæk Tidemand

# Structural investigation of the tissue factor:Factor VIIa:Factor Xa complex

Optimizing carrier systems to provide experimental setups allowing for the investigation of lipid interactions and the structure of a multicomponent system



Principal supervisor: Professor Lise Arleth

Co-supervisor: Professor Birthe B. Kragelund

13<sup>th</sup> of September 2019



# Project description

This project was conducted in the Structural Biophysics group, at the Niels Bohr Institute, University of Copenhagen. The project was supervised by Lise Arleth and co-supervised by Birthe B. Kragelund in the period between the 15<sup>th</sup> of September 2016 to the 14<sup>th</sup> of September 2019. Exactly eight days after the official initiation of the project, I joined my first beamtime in Grenoble, France, and throughout the entire course of the project I have participated in a total of 16 beamtimes in France, Germany, England, and Australia. The project was conducted and financed as a part of the Novo Synergy consortium headed by Birthe B Kragelund. A large part of the three years has been spent in different labs, namely in the lab at the Niels Bohr Institute, the Institute of Biology and a major fraction of the time at Anja Thoe Fuglsangs lab at the Department of Plant and Environmental Sciences at the Faculty of Science. The thesis has been submitted for the degree of Doctor of Philosophy at the University of Copenhagen. A more detailed description of the focus of the work and the content of this thesis will be described in the following sections.

## Motivation and objectives

The interplay between lipids and proteins is to an increasing degree, recognized as an essential parameter in many biological functions, and this interplay was also one of the main motivations to engage in this project. Most structural experimental methods currently available, are mainly directed towards investigating proteins or lipids separately, creating a gap in our knowledge on the structural interplay between the two. During the last three years, I have investigated the blood coagulation system, which contains one membrane protein (MP): Tissue Factor (TF), and several lipid-binding proteins. As for many other similar systems, high-resolution structures of many of the soluble proteins are available, but the interaction with the lipids is poorly understood. The pathway is initiated by the complex formation between TF, and the soluble Factor VIIa (FVIIa). The TF:FVIIa complex activates Factor X (FX) to FXa, in a catalytic mechanism that is markedly increased by the presence of a lipid bilayer. Intriguingly, the lipid composition of the membrane is also essential for the activity, as the Phosphatidylserine (PS) lipids have been shown to increase the activity of the complex even further. The structural nature of this effect is not entirely understood, but it is well known that the affinity between FX and a membrane increases with the amount of PS present in the membrane. The lipid interaction recruits FX to the membrane, where the TF:FVIIa complex is situated, thus increasing the probability of activation. It is also likely PS affects the activation in various other ways, but

this is still unclear.

The structures of the soluble proteins in the complex have been solved, using either X-ray crystallography or nuclear magnetic resonance, but these methods are not useful for the full complex due to the presence of the lipid membrane (which hampers crystallization), and the large size of the system, respectively.

Therefore, other techniques, such as small angle scattering (SAS) or cryo electron microscopy has to be employed. SAS is a technique that provides structural information in the medium to low-resolution range. Measurements are performed in solution, and are compatible with the presence of lipids and therefore ideal for bridging the gap between the high-resolution pieces and the unknown overall structure of the complex. However, in order to perform such experiments, TF needs to be incorporated in a carrier system, providing a hydrophobic environment for the transmembrane domain, but also a lipid bilayer for the interactions with FVIIa and FXa. A fraction of the available carrier systems contains lipids, while others keep the MP in solution by attaching to it directly, such as detergents and amphipols. One of the lipid-containing systems is the nanodisc (ND), which has been used throughout this study. The ND is a molecule composed of a lipid bilayer surrounded by two amphipathic membrane scaffolding proteins (MSPs). This provides a particle which is interesting as a membrane mimic on its own, but also very useful to facilitate studies on MPs incorporated in the ND. In relation to the blood coagulation system, it is a very advantageous feature that the lipid composition of the ND easily can be varied. This is especially important as recent studies have suggested that a broader range of lipids, besides just PS, influences the activity of the TF:FVIIa complex, but this is still to be explored.

In summary, the primary motivation to perform this project was to gain a deeper understanding of the lipid-induced increase in activity as well as the structural features of the lipid:TF:FVIIa:FXa interplay. This was done by utilizing the ND technology which allowed for obtaining data on the tertiary TF:FVIIa:FXa complex on a lipid bilayer.

## Objectives

Three main objectives were defined for this project, and the main part of this thesis will revolve around these. The first objective serves to lay the foundation for the experiments performed to investigate objective 2 and 3.

1. **Establish protocols which enable structural studies on TF in a membrane environment.**

Before this project was initiated, initial attempts at producing and reconstituting TF in NDs had been performed in our group. But as these previous studies were hampered by low production yields and consequently, low sample concentrations for structural studies, we decided to pursue a new production strategy. Also, optimization of carrier systems for SAS were pursued as well as a sample deposition method for neutron reflectometry (NR) experiments.

2. **Investigate the mechanism behind the lipid-induced increase in activity of TF:FVIIa.**

Using the protocols established in objective 1, the aim was to obtain a combination of

structural and functional data using different lipid compositions that could shed light on the still elusive mechanism behind the lipid-induced activity increase of TF:FVIIa.

### 3. Obtain experimental data on the structure of the TF:FVIIa:FXa complex.

The overall objective of this project was to provide experimental data on the tertiary complex, of which the structure has been pursued for decades. The protocols developed in objective 1 should pave the road towards SAS data, which in combination with molecular dynamics simulations could provide an experimentally supported model of the complex.

## Cronology

The main goal of this project was to investigate the structure of the tertiary complex TF:FVIIa:FXa, but before it was possible to prepare such a sample, many roadblocks were faced. The first obstacle was the production of TF in sufficient amounts and purity, which proved to be much more complicated than first expected, and thus, an entire optimization study was conducted. In parallel, a study on optimizing the NDs currently available was conducted, to ensure the availability of a stable carrier system once the TF production was established. The two studies converged at roughly the same time, allowing production of samples with TF and FVIIa on a solubility enhanced and circularized ND. However, to stabilize the tertiary complex, we needed to utilize an inhibitor of the complex, as the FXa binding to the TF:FVIIa complex is transitory. In humans, the primary inhibitor of the complex is tissue factor pathway inhibitor (TFPI), but this is reported to be difficult to produce. Therefore, for the present study, NAPc2 was chosen as it possible to produce in *E. Coli*, had been structurally characterized and shown to stabilize the quaternary complex. Though, as I started the investigation of NAPc2, it became clear that the currently available structure did not represent the data I obtained, and thus what was supposed to be a small part of the story ended up being a study on its own. At the same time, an opportunity to perform NR emerged, and this seemed to be the perfect way to investigate the protein-lipid interplay in the TF:FVIIa complex. However, to perform such studies, a new deposition technique had to be developed, which ensured incorporation of oriented membrane proteins in a bilayer on a surface. I spent quite some time optimizing samples for these measurements before ending up with a protocol that yielded the desired amount of TF in the membrane. In this process we also discovered an interesting but rather annoying tendency of TF preferring incorporation into discs containing PS lipids. Unfortunately, we also found that samples with pure PC membrane were challenging to prepare in reasonable amounts. The mechanism behind the PS preferring behaviour was not explored further due to time limitations, but this is definitely the plan. However, the new deposition technique allowed for measurements on TF, FVIIa, and FXa on a PS containing bilayer, which in combination with a reference bilayer (e.g. PC or PC/PE) might shed light on parts of the structural mechanism behind the PS induced increase in activity. Lastly, SAS measurements were performed on the developed NDs with TF and the complex assembled. Initial attempts at Cryo-EM measurements were also performed, but due to technical issues with the microscope, the first attempts did not provide useful data.

## Thesis content

This thesis is divided into three parts, which serve to introduce the biology and experimental methods, explain the main results, and show the seven resulting manuscripts, respectively. The following gives a brief introduction to the different parts.

### Part 1: Introduction

The introduction consists of two sections; *i*) introducing the biological background and the available information about the structural features of the coagulation pathway. *ii*) introducing the theory behind the experimental methods utilized in this project. The hope is that this introduction will prepare the reader for the following results and manuscripts, and provide the overview needed to put the results presented in this thesis into the right context.

### Part 2: Results and conclusion

This part functions as a summation of the most important findings in this project. I have both included results that are also described in the manuscripts, but also some more preliminary results that are not presented elsewhere. All results will be discussed ongoing and will be wrapped up by a conclusion.

### Part 3: Manuscripts

I have included six manuscripts and one report in which the main results obtained are presented. One of the manuscripts is accepted in a peer-reviewed journal, while two are ready for submission. The rest of the manuscripts and the report are currently lacking some final work before they are ready for submission. A brief description of the status, the content, and my contributions to all manuscripts can be found below.

#### Paper I

*'Circularized and solubility-enhanced MSPs facilitate simple and high yield production of stable nanodiscs for studies of membrane proteins in solution'*

Authors: Nicolai T. Johansen\*, Frederik G. Tidemand\*, Tam T. T. N. Nguyen, Kasper D. Rand, Martin C. Pedersen, and Lise Arleth

\*Authors contributed equally to the work.

This paper initiated the work on sequence optimized MSPs. By introducing a substantial amount of solubility enhancing negatively charged residues, the yield of the circularization reaction increased, and the final ND was in addition more stable at physiologically relevant temperatures. An MP was incorporated, and SAXS studies showed that the structural integrity of the ND was conserved despite all the mutations.

Status: Published in *The FEBS journal*, January 2019.

Author contributions: NTJ, FGT, and LA: Founded and planned the project. NTJ and FGT: Designed the new constructs and conducted expression and purification of the proteins and did the reconstitution of NDs. Also, SAXS, CD, SLS, and SEC measurement were performed and

analyzed, except SAXS modeling, which was performed by MCP. TTTNN and KDR performed MS and analyzed the data. NTJ, FGT, LA, and MCP co-wrote the manuscript.

## Paper II

*'Circularizing reconstituted nanodiscs using a fusion construct'*

Authors: Frederik Grøn**b**æk Tidemand, Xiaoxuan Chen, Nicolai Tidemand Johansen, Tam T. T. N. Nguyen, Kasper Dyrberg Rand, Birthe B. Kragelund, Lise Arleth

Continuation of Paper I, where the possibility to reconstitute an ND and perform the circularization reaction post reconstitution was explored. In addition, a novel fusion construct containing both the MSP and the Sortase was introduced, showing the possibility to rethink the circularization process. An MP was also embedded, and the structure and activity were shown to be unaffected by the post-circularization.

Status: *In preparation*. A fraction of the SAXS data presented is affected by radiation damage and is preferably remeasured. Also, one of the purifications would benefit from further optimization.

Author contributions: FGT, NTJ, and LA: Founded and planned the project. NTJ and FGT: Designed the fusion constructs. FGT and XC: All experimental work except MS measurements which were performed by TTTNN and KDR. FGT, XC, NTJ, BBK, and LA: Discussion of results throughout the project and inputs to the direction of the project. FGT: Analysis of data and writing of the manuscript.

## Paper III

*'Peptide disc mediated control of membrane protein orientation in supported bilayers for surface-sensitive investigations'*

Authors: Alessandra Luchini, Frederik Grøn**b**æk Tidemand, Nicolai Tidemand Johansen, Mario Campana, Javier Sotres, Michael Ploug, Marité Cárdenas, Lise Arleth

To be able to investigate MPs in a lipid bilayer using NR, a method ensuring a unidirectional orientation of the MP of interest was developed. Using peptidediscs as carriers of a patch of membrane including the MP of interest, a complete membrane could be formed simply by removing the peptide after deposition. This left a bilayer with all MPs in the same direction ready for experiments such as NR, SPR, and AFM.

Status: *In submission*.

Author contributions: AL and LA: Founded and planned the project. AL and FGT: Optimization of samples and sample preparation for experiments. AL and MCar conducted NR experiments. AL, JS, and MCar conducted AFM measurements and did data analysis. MP conducted SPR measurements and data analysis. AL conducted QCMD experiments and analyzed NR data. AL, FGT, NTJ, and LA discussed the data and results throughout the project. AL wrote the manuscript with inputs from all authors.

## Paper IV

*'Efficient refolding and reconstitution of Tissue Factor into nanodiscs facilitates structural investigation of a multi-component system on a lipid bilayer'*

Authors: Frederik Grønbæk Tidemand, Henrik Østergaard, Michael Ploug, Birthe B. Kragelund and Lise Arleth

This work emerged as obtaining the quantities of TF needed for structural studies proved more difficult than first expected. TF was expressed in inclusion bodies, and the subsequent refolding was optimized as well as the reconstitution into NDs. The activity of the obtained TF was probed to ensure a correct fold, and the first structural data on TF in NDs was obtained by SAXS measurements, laying the foundation for some of the later work in this thesis.

Status: *In submission.*

Contributions: All authors contributed to the experiment design, FGT conducted all experiments and data treatment with input from HØ, BBK, and LA, except SPR measurements which were conducted by MP. FGT wrote the manuscript with contributions and input from all co-authors.

## **Paper V**

*'Probing the orientation of the tissue factor:Factor VIIa complex in a lipid bilayer'*

Authors: Frederik Grønbæk Tidemand, Alessandra Luchini, Raul Araya-Secchi, Katrine Bugge, Michael Ploug, Birthe B. Kragelund, Lise Arleth

The mechanism behind the increased activity of the TF:FVIIa complex upon the presence of PS lipids is still not elucidated. One of the proposed mechanisms is interactions between the ECD of TF and the membrane, and that this potentially could alter the activity. We explore this by using the developed protocols from Paper III and IV and combining NR, NMR, SAXS, SPR, QCMD, and molecular dynamics simulations to probe the importance of the bilayer in the orientation of the TF:FVIIa complex.

Status: *In preparation.* Unfortunately it has proven difficult to produce the samples with TF in a pure POPC bilayer, and hence a useful dataset under those conditions is still needed before this manuscript can be finalized for submission.

Contributions: FGT, AL, and LA founded the project. FGT produced proteins and samples. AL performed NR and QCMD experiments and analyzed the data. RAS performed MD simulations. FGT performed NMR and SAXS experiments. All results were discussed throughout the project with all authors. FGT wrote the manuscript.

## **Paper VI**

*'Probing the flexibility of the blood coagulation inhibitor, NAPc2'*

Authors: Frederik Grønbæk Tidemand, Katrine Bugge, Raul Araya-Secchi, Birthe B. Kragelund, Lise Arleth

This project was initiated as NAPc2 was needed to stabilize the quaternary complex assembled around TF. Initial studies of the structure of NAPc2, however, revealed that there was more to the story than previously published, and a more thorough investigation was initiated. A

combination of SAXS measurements and MD simulations revealed a large degree of flexibility in NAPc2, and the initial NMR results also support this. Mapping the interface between NAPc2 and FVIIa was unfortunately not finalized within the project time-frame, but will become important for modeling the entire complex.

Status: *In preparation.* Due to time limitations, NMR experiments using  $^{13}\text{C}^{15}\text{N}$ -NAPc2 were not done, and hence, assignments of the residues were not possible. This leaves the NMR part of the manuscript unfinished, and these experiments have to be conducted before the manuscript can be finalized.

Contributions: FGT, BBK, and LA founded the project. FGT optimized expression and purification, produced all proteins, performed all activity assays and SAXS measurements. FGT and KB conducted NMR experiments and data analysis. RAS performed all simulations. RAS and FGT analyzed SAXS data. Results were discussed throughout the project with all authors. FGT wrote the manuscript.

## Report I

*'Structural investigation of the tissue factor:Factor VIIa:Factor Xa:NAPc2 complex on a lipid bilayer'*

Authors: Frederik Grønbæk Tidemand

This report contains SAS data on tissue factor in an ND, and with relevant interaction partners bound. Data on the quaternary complex in an 'invisible' ND is also presented.

Status: *In preparation.* Most data is collected, but modeling of the data is lacking. This involves modeling of the individual parts, but also the assembly of the complex and subsequent refinement using SAS data.

Contributions: FGT conducted all the work presented in this report.

## Other manuscripts

During this project, I also contributed to the following paper, but as the contributions are rather small, it is not included in the thesis.

*'PSX: Protein-Solvent Exchange - Software for calculation of deuterium-exchange effects in SANS measurements from protein coordinates'*

Authors: Martin Cramer Pedersen, Yong Wang, Frederik Grønbæk Tidemand, Anne Martel, Kresten Lindorff-Larsen and Lise Arleth

Status: Accepted in *Journal of Applied Crystallography*, April 2019.

FGT contributions: Sample preparation and SANS measurements of BSA at different pH-values and  $D_2O$  exposure times. Writing of methods section related to the SANS experiments.

## Other work

As a part of the last three years, I have also conducted work on another study which is not included in this thesis and has not resulted in publications. An extensive investigation of the Sodium dodecyl sulfate mediated unfolding of various proteins probed by isothermal calorimetry,

SAXS, and circular dichroism. This has also been the topic of three Bachelors's projects of which I was the daily supervisor. Plenty of data is available from these studies, and the plan is to finalize a manuscript based on this data after handing in this thesis.

Also, I was the daily supervisor of a Bachelors's and a Master's project on the work presented in Paper II.

# Abstract

Structural investigation of proteins in a lipid context remains challenging due to the unavailability of suitable experimental techniques and sample environments. Two of the most common techniques, for probing protein structures, are X-ray crystallography and nuclear magnetic resonance (NMR). However, for systems involving lipids, these two techniques are not commonly applied, as the presence of a lipid membrane makes crystallization infeasible, and the relatively large size of most lipid-containing systems makes the tumbling too slow for NMR measurements. Therefore, other techniques, such as small angle scattering (SAS) or cryo electron microscopy, have to be used. SAS measurements are performed in physiologically relevant conditions, and measurements on systems containing both proteins and lipids are possible, albeit information is obtained with a relatively low resolution.

In this thesis, the membrane protein tissue factor (TF), and its soluble interaction partner Factor VIIa, has been investigated. Upon the rupture of a blood vessel, the formation of the TF:FVIIa complex occurs and initiates the extrinsic blood coagulation pathway. TF:FVIIa recognizes the macromolecular substrate, Factor X (FX) and activates it to FXa, which subsequently participates in the common coagulation pathway, ultimately resulting in blood clotting. The activity of the TF:FVIIa complex, is markedly increased upon the presence of a lipid membrane, and in general, lipid interactions are essential in the blood coagulation system. High-resolution structures of the individual soluble parts of the tertiary TF:FVIIa:FXa complex have been solved, but as the tertiary structure can only be formed in the presence of a lipid bilayer, it is challenging to approach experimentally. In the presented work, the nanodisc (ND), which is a small lipid bilayer encircled by two amphipathic proteins, was utilized to facilitate the formation of the tertiary complex.

Firstly, the production of TF was established in order to obtain the relatively large amounts required for structural studies. In the same study, the incorporation of TF into the developed circularized and solubility enhanced NDs (csND) was optimized. The csNDs were shown to possess higher temporal stability, and to be easy to produce in large amounts. Having TF in a csND enabled functional studies, showing the expected activity of TF. Also SAS measurements were performed both on TF in a csND, but also with FVIIa added, demonstrating the anticipated enlargement of the particle. However, the interaction between the TF:FVIIa complex and FXa is only transitory, and thus we had to deploy an inhibitor, NAPc2, which stabilize the complex, allowing for SAS measurements. In parallel, neutron reflectometry was used to probe the heights of TF and FVIIa on a membrane. These measurements provided the opportunity to evaluate the orientation of both TF and the TF:FVIIa complex on a membrane. Measurements were also

performed on FXa, revealing that the FVIIa active site and the cleavage site on FX, are located at the same height, which highlights an essential function of the membrane, and potentially explains part of the membrane induced increase in activity.

In summary, this project provided new insights into the lipid-protein interplay in the TF:FVIIa:FXa system. In addition, the first structural data on the quaternary TF:FVIIa:FXa:NAPc2 complex on a lipid bilayer was obtained, laying the foundation for modeling of the TF:FVIIa:FXa complex.

# Resumé

Strukturelle undersøgelser af proteiner i samspil med lipider forbliver en udfordring på grund af de eksperimentelle metoder der er til rådighed. To af de mest anvendte metoder til protein struktur bestemmes er røntgen krystallografi og kerne magnetisk resonans (NMR). Ingen af disse to metoder er dog særlig anvendelige når der er lipider tilstede, da fleksibiliteten gør krystallisering svær, mens den store størrelse som protein-lipid komplekser ofte har, gør NMR umulig. Derfor må man ty til andre metoder ved undersøgelse af disse prøver, og her kommer småvinkel spredning (SAS) ind i billedet. Modsat krystallografi og NMR, kan man med SAS kun få information med lav opløsning, men prøverne kan måles under fysiologiske betingelser, og man kan måle på prøver der både indeholder proteiner og lipider.

In denne afhandling er membranproteinet tissue factor (TF) of dets opløselige binding partner Factor VIIa (FVIIa) blevet undersøgt. Ved skade på en blodåre kan TF og FVIIa danne komplekset der starter den indre blod koagulerings kaskade. TF:FVIIa komplekset genkender så Factor X (FX) og aktiverer det til FXa som efterfølgende indgår i den fælles koagulerings kaskade der ender med stoppe blødning.

Aktiviteten af TF:FVIIa komplekset bliver markant forøget ved tilstedeværelsen af en lipid-membran, og generelt set er lipid interaktioner fundamentale i blodkoaguleringskaskaden. Høj opløsnings strukturer af de individuelle opløselige dele af det tertiære TF:FVIIa:FXa kompleks er blevet løst, men da det tertiære kompleks kun kan dannes i tilstedeværelsen af en membran er det svært at undersøge det eksperimentelt. I dette projekt er nanodiscen (ND), der er et lille stykke membran omkredset af to amfipatiske proteiner, blevet brugt som membran til bindingen af det tertiære kompleks. Først skulle produktionen af TF dog optimeres, da den var vigtig for at kunne få de relativt store mængder rent protein der er nødvendige for at lave strukturelle studier. I det samme delprojekt blev inkorporeringen af TF i de udviklede cirkulære og opløseligheds optimerede NDs (csND) også optimeret. Vi viste også at csND'erne have en højere stabilitet over tid, og at de var relativt lette at lave i store mængder. Med TF i csND'ere kunne vi udføre funktionelle studier der viste at TF havde den forventede aktivitet, men vi kunne også lave strukturelle studier. SAS-eksperimenter blev lavet på både TF i ND'ere men også med FVIIa tilsat, hvilket viste den forventede større størrelse af partiklen. Vi forsøgte derefter at tilsætte FXa, men da interaktionen kun er kortvarig var vi nødt til at bruge inhibatoren NAPc2 til at låse komplekset. Da vi tilsatte NAPc2 kunne vi samle det nu kvaternære kompleks på en ND og derefter udføre SAS målinger.

Samtidig lavede vi neutron reflectometri målinger der tillod os at måle højderne og orienteringerne af både TF, FVIIa og FXa på en membran. Vi så at højen af FVIIAs aktive site

var den samme som kløvningssitet på FX, hvilket understreger at lokalisering af disse er en vigtig membran funktion. Dette kan muligvis også være del af forklaringen på den lipid inducerede stigning i aktiviteten af TF:FVIIa komplekset.

Alt I alt har dette projekt givet ny viden om lipid-protein sammenspillet i TF:FVIIa:FXa komplekset og lagt fundamentet for modelring af komplekset på en lipidmembran.

# Acknowledgement

First and foremost, I would like to thank my supervisor Lise Arleth, for excellent supervision and continuous support during the project. I would also like to thank you for encouraging me to pursue the research ideas I found interesting, even though they were quite different from most of the projects conducted in the group. Also, a big thanks to my co-supervisor Birthe B. Kragelund for great supervision throughout the three years. I would also like to thank Anja Thoe Fulgsang for letting me work in the labs at Frederiksberg campus. It is probably the place I have spent the most time in the last three years.

I would also like to thank all the members of the Structural Biophysics group for creating a very pleasant working environment. In this context, a special thanks to Nicolai Tidemand Johansen for numerous good collaborations, beamtimes, and idea-development sessions, where many of the projects presented in this thesis were founded. Also, a big thanks to Alessandra Luchini, Raul Araya-Secchi and Martin Cramer Pedersen for good collaborations.

Finally, I would like to thank my girlfriend, Josephine, for proofreading, general support, and a genuine attempt at appreciating the world of protein structures!

# List of abbreviations

Acronym	Description
Acronym	Description
ABC	Anything but choline
AFM	Atomic force microscopy
CMP	Contrast match point
cMSP	Circularized membrane scaffolding protein
cND	Circularized nanodisc
csdH5	Circularized and solubility enhanced MSP1\delta H5D1
csE3	Circularized and solubility enhanced MSP1E3D1
d18A	Dark 18A
dcsND	Deuterated circularized and solubility enhanced nanodisc
DG-FVIIa	Des-gla Factor VIIa
DG-FXa	Des-gla Factor Xa
dND	Deuterated nanodisc
ECD	Extracellular domain
EGF	Epidermal growth factor
EM	Electron microscopy
EPOr	Erythropoietin receptor
eSrt	Evolved Sortase
F	Factor
fE3	Fusion construct of csE3 and eSrt
FL	Full length
FRET	Förster resonance energy transfer
FVII	Factor VII
FVIIa	Factor VIIa
FX	Factor X
FXa	Factor Xa
GHr	Growth hormone receptor
Gla	$\gamma$ -carboxyglutamate
GuHCl	Guanidine hydrochloride
hND	Hydrogenated ND
ICD	Intracellular domain
IEC	Ion exchange chromatography
K <sub>d</sub>	Dissociation constant

$K_m$	Michaelis–Menten constant
lsE3	Linear and solubility enhanced MSP1E3D1
MD	Molecular dynamics
MO	Match-out
MP	Membrane protein
MSP	Membrane scaffolding protein
MTL	Mass transport limitation
NAPc2	Nematode Anticoagulant protein c2
ND	Nanodisc
NiNTA	Nikkel Nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NR	Neutron reflectometry
NTA	Nitrilotriacetic acid
PC	Phosphocholine
PDB	Protein data bank
PE	Phosphatidylethanolamine
POPC	1-palmitoyl-2-oleoyl-glycero-3-phosphocholine
POPS	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine
PS	Phosphoserine
QCMD	Quartz crystal microbalance with dissipation monitoring
RU	Resonance units
SANS	Small angle neutron scattering
SAS	Small angle scattering
SAXS	Small angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SLB	Supported lipid bilayers
SLD	Scattering length density
SM	Sphingomyelin
SPR	Surface plasmon resonance
sTF	Soluble tissue factor
TEV	Tobacco etch virus
TF	Tissue factor
TFPI	Tissue factor pathway inhibitor
TMD	Transmembrane domain

# Contents

<b>Project description</b>	<b>i</b>
<b>Abstract</b>	<b>ix</b>
<b>Resumé</b>	<b>xi</b>
<b>Acknowledgement</b>	<b>xiii</b>
<b>List of abbreviations</b>	<b>xiv</b>
<b>I Introduction</b>	<b>1</b>
<b>1 The human blood coagulation system</b>	<b>3</b>
1.1 Coagulation pathways . . . . .	3
1.2 Molecular features of the TF complex . . . . .	5
1.3 Related diseases . . . . .	21
1.4 Outlook . . . . .	23
<b>2 Experimental methods</b>	<b>24</b>
2.1 Sample production . . . . .	24
2.2 Small angle scattering . . . . .	30
2.3 Experiments on a surface . . . . .	44
2.4 Neutron reflectometry . . . . .	47
2.5 Surface plasmon resonance . . . . .	54
<b>II Main results and conclusions</b>	<b>59</b>
<b>3 Results and discussion</b>	<b>61</b>
3.1 Development of carrier systems and purification protocols . . . . .	61
3.2 Tissue factor and Factor VIIa . . . . .	73
3.3 Investigation of the quaternary complex . . . . .	78
<b>4 Conclusions and outlooks</b>	<b>83</b>
<b>References</b>	<b>104</b>

<b>III Manuscripts</b>	<b>105</b>
<b>5 Paper I</b>	<b>107</b>
<b>6 Paper II</b>	<b>131</b>
<b>7 Paper III</b>	<b>147</b>
<b>8 Paper IV</b>	<b>173</b>
<b>9 Paper V</b>	<b>187</b>
<b>10 Paper VI</b>	<b>204</b>
<b>11 Report I</b>	<b>215</b>



## Part I

# Introduction



# 1 | The human blood coagulation system

The main aim of this PhD-project was to obtain structural information on the tissue factor (TF):Factor VIIa (FVIIa):Factor Xa (FXa) complex, which plays an essential role in the human blood coagulation pathway. This chapter presents an overview of the coagulation system in humans and an in-depth description of the mechanisms and structural features of the TF:FVIIa:FXa complex.

## 1.1 Coagulation pathways

The blood coagulation pathway can be divided into three subpathways: The intrinsic, the extrinsic, and the common pathway, as illustrated in figure 1.1A. These pathways consists of many different proteins, all having their specific place in these cascade mechanisms. All the pathway proteins share the common name: Factor (F) followed by and a Roman numeral unique for each protein, with some exceptions listed in table 1.1. All the factors listed as Protease zymogens in table 1.1 can be activated by limited proteolysis. Upon activation a lowercase 'a' is added to the name [1]. In some activation mechanisms the protein mass is kept intact, while in others the so-called activation peptide is released. Following activation, the now separated domains are held together by disulfide bonds, conserving the same overall shape.

The *in vivo* triggering mechanism of the intrinsic pathway is still not known, and it is likely that multiple activators can initiate the pathway. Currently, activators ranging from specific proteins to extracellular nucleic acids and glycosaminoglycans have been proposed [1]. *In vitro* studies show that activation of FXII to FXIIa is initiated when blood comes in contact with an artificial surface [2], as shown as the first step in figure 1.1. The next steps are well understood: FXIIa activates FXI and FXIa activates FIX. FIXa and FVIIIa form the intrinsic Tenase complex that activates FX, which then goes into the common pathway [3].

The extrinsic pathway contains fewer steps, and the triggering mechanism is well documented [1, 4]. While all the other coagulation factors are present in the blood stream, TF is located on the surface of the cells surrounding the blood vessels [5, 6]. Upon rupture of a blood vessel, FVII is able to interact with TF by forming the TF:FVII complex which initiates the extrinsic pathway. The complex formation increases the activity of FVIIa by more than 100,000 times, making TF an essential regulator. Low amounts of the active FVIIa ( $\sim 1\%$ ) is present in the blood, while the major fraction is zymogen FVII constituting 99% of FVII [7, 8]. Activation of FVII can be performed by FVIIa, FXa, or FIXa and mainly occurs after complex formation with TF [9]. However, the precise mechanism by which FVII activation occurs is still not fully

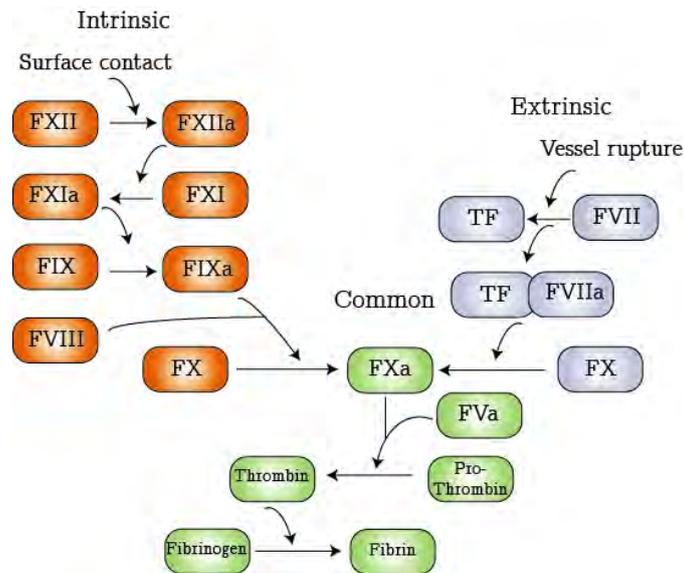


Figure 1.1: Simplified overview of the human blood coagulation pathway.

understood and will be discussed in greater detail in a later section. The activated TF:FVIIa complex is also called the extrinsic tenase complex as it activates FX to FXa. FXa interacts with FVa in a complex on the membrane surface, called the Prothrombinase complex, which activates Prothrombin to Thrombin. Thrombin cleaves Fibrinogen leading to formation of fibrin which participates as structural elements in clotting the blood vessel leak.

As seen in table 1.1, the plasma levels of the different coagulation factors spans four orders of magnitude, which is really one of the strengths of the pathway. Generally the concentration of the specific protein involved is increased with each step of the pathway and in combination with the consecutive activation this results in a really efficient way to amplify the signal. Especially Fibrinogen is present in high amounts ( $\sim 3 \text{ mg/ml}$ ), which is needed as it is the structural foundation of clotting, but also both Prothrombin and FX are present in relatively high amounts as they constitute the two final steps before clotting.

No	Common name	Function	MW (kDa)	Plasma level (nM)
I	Fibrinogen	Structural	340	8824
II	Prothrombin	Protease zymogen	72	1389
III	Tissue Factor	Cofactor	37	0
V	Factor V	Cofactor	330	30
VII	Factor VII	Protease zymogen	50	10
VIII	Factor VIII	Cofactor	330	0.3
IX	Factor IX	Protease zymogen	56	89
X	Factor X	Protease zymogen	56	179
XI	Factor XI	Protease zymogen	160	31
XII	Factor XII	Protease zymogen	80	375

Table 1.1: Overview of relevant proteins involved in coagulation. Information about molecular weight and plasma levels are from [6]. Plasma levels can differ from person to person.

In addition to the mentioned coagulation factors, platelets are important constituents of a well-functioning coagulation system. Platelets are cellular fragments of about 2-3  $\mu\text{m}$  in diameter, which do not contain a nucleus, but still retain other cellular components such as mitochondria, mitochondrial DNA and endoplasmic reticulum [10]. Platelets circulate in the plasma, and are upon vascular injury recruited to the site of injury, forming an initial plug. Subsequently, the platelets provide the phospholipid membrane facilitating the activation of many of the mentioned coagulation factors and also participates as a part of the plug itself [11, 10].

## 1.2 Molecular features of the TF complex

### 1.2.1 The TF:FVIIa complex

As previously mentioned, the formation of the TF:FVIIa complex initiates the extrinsic pathway. TF is a 37 kDa integral MP with a single transmembrane helix connected to an N-terminal globular extracellular domain (ECD) and a, likely disordered, short intracellular domain (ICD) in the C-terminal. The ICD is generally accepted to be insignificant for the clotting activity but possesses a possible regulatory mechanism in tumor growth and angiogenesis, both regulated by phosphorylations in the ICD [12, 13]. However, as this introduction focuses on the mechanisms behind blood clotting, the regulatory functions of the TF ICD is out of scope. The TF ECD spans the first 219 residues, and is composed of two fibronectin type III domains, which is a beta sandwich structure. Both fibronectin domains contains a disulfide bridge and are joined at an angle of 125 degrees (figure 1.2). C-terminally of the ECD, a 10 residue long linker is present before the TMD starts. No structural information is published about the linker, but it is likely flexible, as no electron density was present in the crystal structure, even though the crystallized protein contained part of the linker. Also the TMD is structurally uncharacterised, and the exact start end end points are predicted by MP topology predictors.

FVIIa is a soluble 50 kDa protein consisting of a Gla-domain, two epidermal growth factor (EGF)-like domains, and a C-terminal serine protease domain. A signal peptides ensures localization to the Golgi apparatus and later secretion to the blood. The Gla-domain is named after the ten  $\gamma$ -carboxyglutamate (Gla) residues situated in the membrane binding region and is only completely folded upon the binding of multiple calcium ions. The Gla residues are formed by a carboxylase in the Golgi and has been shown to be essential for clotting [14]. The binding of the calcium induces a conformational change in the domain which serves to direct three hydrophobic residues outwards enabling the lipid interaction of the domain. The active site of FVIIa is located in the protease domain, is called the catalytic triad and consists of a His<sup>193</sup>, Asp<sup>242</sup> and a Ser<sup>344</sup> (red residues, figure 1.2A) [15]. This triad enables a charge transfer which prepares the serine for a nucleophilic attack [16]. Therefore, it is essential that these three residues are located precisely in the correct spatial configuration, and this configuration is highly conserved among serine proteases. Zymogen FVII is a single polypeptide string, but upon activation to FVIIa, cleavage of the scissile bond (between Arg<sup>152</sup> and Ile<sup>153</sup>) divides the protein into the heavy and light chain. The two chains are kept together by a disulfide bridge, and the protein is now active. Upon activation, the newly formed N-terminal (from the cleavage) inserts into

the active site of FVIIa and stabilize the catalytic triad in the correct position [17]. The region in FVII containing the scissile bond is unstructured, which hampers structural determination of FVII, and thus the precise structural rearrangement is not fully characterized.

The binding between TF and FVIIa is very strong with a  $K_d$  in the pM range [18, 19] (figure 1.3F) which is below the reported plasma levels of FVII (around 10 nM, see table 1.1). Many studies have been conducted using a truncated version of TF only containing the soluble ECD of TF called sTF. sTF has a lower affinity for FVIIa with a  $K_d$  three orders of magnitude higher than WT TF in a lipid bilayer [18] (figure 1.3E). The crystal structure of the sTF:FVIIa complex was solved in 1996 [20] (figure 1.2A) showing an extended binding interface providing many contact points between the two proteins. FVII and FVIIa bind TF through approximately the same interface, suggesting that TF confers a dual role at the membrane surface: *i*) recruiting FVII to the surface for activation and *ii*) increasing the activity of FVIIa. It is well established that TF binding induces a stabilization of the FVIIa catalytic triad and also loops which play a vital role in substrate recognition [21, 22, 23]. Biochemical studies have shown that especially Met<sup>164</sup> on FVIIa plays a critical role in the TF induced activity of FVIIa [24].

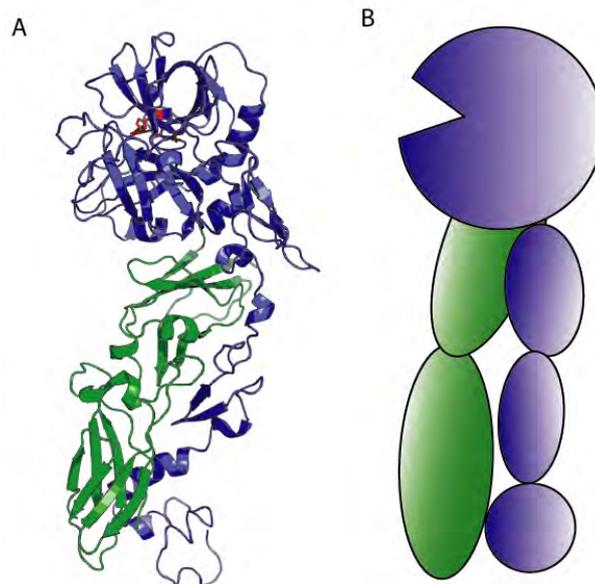


Figure 1.2: sTF:FVIIa complex. Green: TF. Blue: FVIIa. PDB ID: 1DAN [25].

The membrane is also important for FVIIa binding: Firstly, the FVIIa Gla-domain is known to have affinity for the membrane (figure 1.3B) [26] (which will be discussed in details later) and secondly, interacting with the membrane will restrict FVIIa in the correct orientation for TF binding. To elucidate the importance of the FVIIa Gla-domain, Neuenschwander and colleagues used a variant of FVIIa lacking the Gla-domain (des-gla, DG-FVIIa) [27], which did not have any affinity for the membrane (figure 1.3A). Previous studies had shown that the affinity between FVIIa and a membrane containing 30% PS had a  $K_d$  in the low  $\mu$ M range [26, 28] (figure 1.3B). The  $K_d$  of DG-FVIIa and TF in the membrane was found to be 50 nM which is a 1000 times higher than that of FVIIa [27, 18] and is almost similar to that of sTF and FVIIa (figure 1.3D, F and E).

In conclusion, it seems that the strength of the TF:FVIIa interaction mainly arises from the

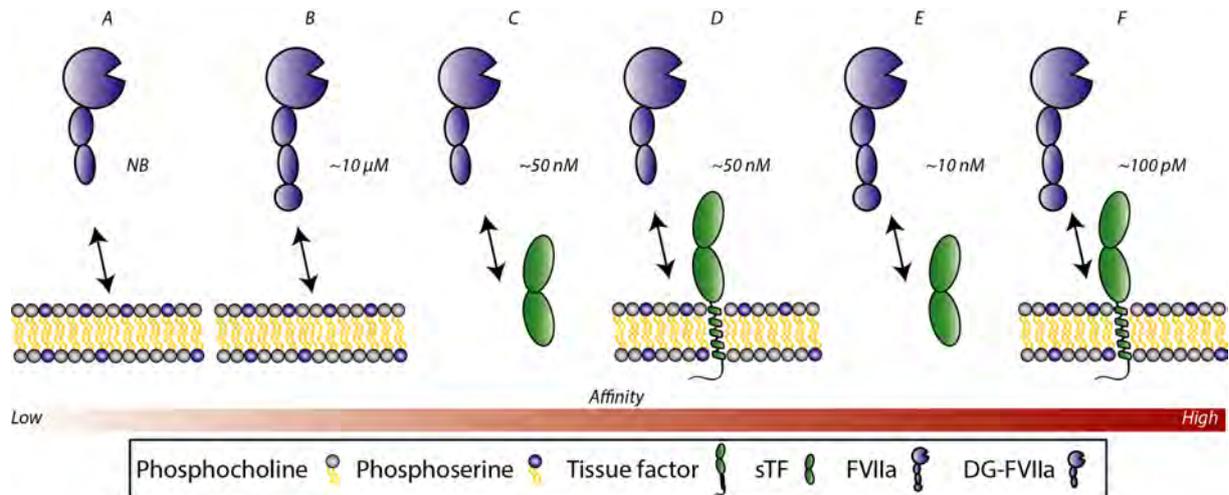


Figure 1.3: Affinity between FVIIa and TF in different cases with the associated  $K_d$  values stated in the figure. All values are from [27, 18]. NB: No binding.

protein-protein interactions, but it is also evident that in order to obtain the maximal  $K_d$ , the membrane has to be present.

### 1.2.2 Factor X

Structurally FX is very similar to FVIIa, with the same overall domains: The Gla-domain, two EGF-like domains and the serine protease domain containing the active site (figure 1.4). As for FVIIa, the Gla-domain binds a number of calcium ions, and mediates the lipid binding of FX(a). The affinity for the lipids is critical for the function of FXa, as the complex with FVa (the Prothrombinase complex) is formed on a lipid bilayer, only mediated by the Gla-domains of FXa and FVa. The FXa catalytic triad is very similar to that of FVIIa and consists of His<sup>236</sup>, Asp<sup>282</sup> and Ser<sup>379</sup> [29].

The structure of full length (FL) FXa is not solved, and the model presented in figure 1.4A is based on existing structures of parts of FXa and completed by homology modeling [30]. Contrary to FVIIa, FX is divided into the heavy and light chain by specific proteolysis in the Golgi apparatus [31]. This means that upon activation of FX by FVIIa, the cleavage of the Arg<sup>194</sup>-Ile<sup>195</sup> bond results in release of the activation peptide. The structure of the protease domain of FXa is also similar to that of Thrombin, however it is much more specific. The residue Asp<sup>373</sup> has proven important in substrate recognition and is probable the explanation for the relatively high specificity of FXa compared to e.g. thrombin [29].

In the activation by the TF:FVIIa complex, FX docks on the complex and is after activation released again. The precise binding interface is not known, as the tertiary structure is not determined (will be discussed in greater detail in a later section) but fluorescence studies have highlighted some important residues [32]. All these residues are located on the lower part of TF on the opposite side of the FVIIa interface. As the formation of the tertiary complex requires the presence of lipids, high resolution structural information is highly challenging to obtain.

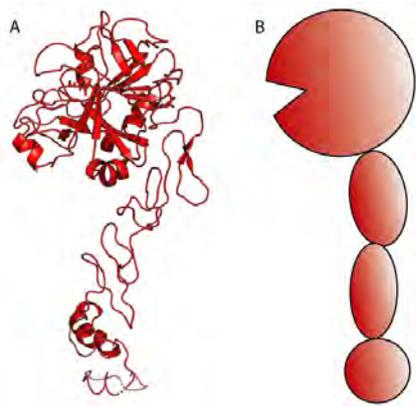


Figure 1.4: Overview of the structure of FXa. A: FXa from computational model presented in [30]. B: Illustration of FXa.

### 1.2.3 Glycosylations

Both TF, FVII and FX are glycosylated in their natural form. Recombinant TF expressed in mammalian cells has three N-glycosylations in the ECD, namely Asn<sup>11</sup>, Asn<sup>124</sup> and Asn<sup>137</sup> [33]. Contradicting results have been published on the importance of these glycosylations with regards to clotting activity. The very first study on this matter demonstrated that adding the carbohydrate-binding protein Con A, markedly reduced the TF activity indicating glycosylations are important for the function of TF [34]. However, later studies using recombinant purified TF from mammalian cells (containing glycosylations) and *E. Coli* (without glycosylations) suggested that glycosylations were without implications for the activity, as the two proteins exhibited similar activity [35, 36]. Other studies suggests that there is a decrease in activity of TF from human placenta compared to TF produced in insect cells and TF produced in *E. Coli*. In support of this, deglycosylation of placenta TF has been shown to result in a decreased activity [37]. It is still unknown what gives rise to these discrepancies in the literature. One explanation could be the difference in experimental setups, as the TF dependency on glycans might be different when probed *in vivo* or *in vitro*, but this is still to be determined [38].

FVII contains four glycosylation sites, two O-glycosylations (Ser<sup>52</sup> and Ser<sup>60</sup>) and two N-glycosylations (Asn<sup>145</sup> and Asn<sup>322</sup>) [39]. It has been suggested that the N-glycosylations are mainly involved during folding of FVII, while O-glycosylations are involved in the function of FVII [40]. Studies by Iino and colleagues showed that mutations of Ser<sup>52</sup> and Ser<sup>60</sup> to Ala reduced the clotting activity of FVII to only 14% of the WT activity [41]. This correlates with a small increase in  $K_d$  of the double mutant towards TF compared to WT FVII [41] suggesting that the O-glycosylations are involved in TF binding. The O-glycosylations have been found to be relatively homogenous as only one species is found on Ser<sup>60</sup> while three different glycans were found on Ser<sup>52</sup>. On the other hand the glycans on Asn<sup>322</sup> and on Asn<sup>145</sup> are rather heterogeneous [40].

Similarly to FVII, FX contains two O-linked (Thr<sup>159</sup> and Thr<sup>171</sup>) and two N-linked (Asn<sup>181</sup> and Asn<sup>191</sup>) glycosylation sites. Very interestingly all four sites are located in the activation peptide [42, 43]. A study by Wang *et al.* showed that the glycans on FX were essential in restricting the interaction between FX and FVIIa if TF were not present. Removing all four glycosylations

resulted in a 20-fold increase in  $K_m$  in the activation by the TF:FVIIa complex, which suggests that the glycosylations provide important interactions mediating FX binding [43]. Lastly, it was found that the glycosylations protected FX from unspecific activation.

### 1.2.4 Encryption of TF

In the 1970s it was suggested that only a small fraction of the TF present on cell surfaces is actually in an active state [44]. Later studies confirmed this finding and estimated that around 80% of TF on cell surfaces is in an inactive state, which is typically referred to as cryptic TF [45]. In addition, a study by *P. Giesen et al.* found that contrary to previously proposed theories, low levels of TF are present in the blood [46]. This also points towards the fact that TF must be regulated by other means than just being separated physically from FVIIa, as unwanted clotting would otherwise occur. The origin of this cryptic state has been discussed and investigated ever since, but to date no definite answer has been found. This is probably because there is no single mechanism that can explain this phenomenon, rather it is the sum of different contributions [47]. In the following sections the two main encryption mechanisms proposed will be described as the encryption is a key question in the mechanism of TF regulation, and has been so for the last decades.

#### 1.2.4.1 Structural encryption

The ICD of TF is, as previously described, not important for clotting activity, but it was hypothesised that it could influence the cryptic state of TF [12, 38]. TF is a member of the cytokine receptor superfamily which also contains the growth hormone receptor (GHR) and erythropoietin receptor (EPOr) [48]. Both the GHR and the EPOr are crystalized as dimers [49, 50] and TF could therefore potentially also be regulated by a dimer:monomer equilibrium. Chemical cross-linking experiments revealed the presence of dimeric TF on the cell surface and laid the foundation for the theory illustrated in figure 1.5A and B [51, 52]. As seen in the figure, TF contains a free cysteine in the ICD which was expected to mediate the dimerization, but studies showed that both the ICD and the TMD were important for the dimerization [52]. When treating cells with an ionophore that binds calcium irreversibly and triggers the TF activity it was found that this ionophore also induced monomerization of TF as probed by cross-linking [52]. It was proposed that the mechanism behind this finding was blocking of the FX binding interface due to dimerization, resulting in deactivation of the complex. Upon calcium influx monomerization was thus proposed to increase hence increasing the TF activity. However, other studies have shown that TF is mainly expressed as a monomer, which should provide unwanted clotting according to the hypothesis stated above. In addition, a study using an sTF construct forced to form a dimer in the C-terminal did not show any decrease in activity [53]. Furthermore, two studies published shortly after, showed that the ICD does not interfere with the encryption of TF, again suggesting that the explanation should be found elsewhere [54, 55]. Even though the current data demonstrate conflicting results, the most common opinion is that dimerization is not a part of the encryption mechanism of TF.

The other suggested structural encryption mechanism for TF, is the formation of the second

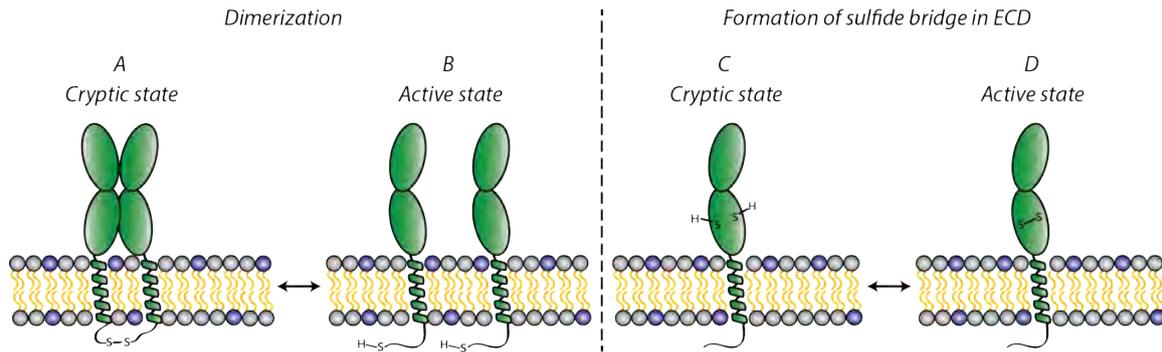


Figure 1.5: Structural encryption of TF: A + B: Dimerization of TF. A: Cryptic state when TF is a dimer. B: Active state upon monomerization of TF. C + D: Formation of a disulfide bridge in the TF ECD. C: Cryptic state when disulfide bridge is not formed. D: Active state upon formation of disulfide bridge.

disulfide bridge (between Cys<sup>186</sup> and Cys<sup>209</sup>) located in the C-terminal part of the ECD of TF (figure 1.5C and D). A mutational study showed that mutating both cysteines (Cys<sup>186</sup> and Cys<sup>209</sup>) to serines decreased affinity to FVIIa 20-fold and the activity of the complex 30-fold, while mutating the two other cysteines in the ECD did not affect binding nor activity [56]. A recent molecular dynamics study showed that sTF with all disulfide bridges formed, introduced a rigid conformation of FVIIa upon complex formation [57]. When performing the same simulation with the mutant unable to form the second disulfide bridge, a much more flexible conformation of FVIIa was observed. The introduced stability in FVIIa was suggested to mediate the correct orientation, needed to reach a high activity. The dramatic loss in affinity of the mutant sTF towards FVIIa is contrary to what has been described for the cryptic state of TF, as the affinity of cryptic TF is unchanged, and thus the mutant do not directly confer the cryptic TF. Thus in conclusion with the data currently available, it is impossible to draw any conclusions on this encryption mechanism.

#### 1.2.4.2 Lipid mediated encryption

It is well established that the human lipid membrane has an asymmetrical distribution of lipids, and that this asymmetry aids and regulates many cellular processes [58, 59, 60]. The membrane contains a variety of phospholipids, sterols and MPs and the general comprehension moves towards the hypothesis that many biological processes involving either of the three is regulated by an interplay between all three. When disregarding the MPs, the human membrane consists of about 50% Cholesterol and 50% phospholipids [58]. The distribution of the most common phospholipids can be seen in figure 1.6A, with mainly Sphingomyelin (SM) and Phosphocholine (PC) in the outer leaflet and mainly Phosphoserine (PS) and Phosphatidylethanolamine (PE) in the inner leaflet. As seen from the phospholipid structures in top of figure 1.6B, this also gives an asymmetrical distribution of charges, as PS is the only one of the four carrying net charged headgroup, while both SM, PC and PE are zwitterionic. The structures in figure 1.6B are the ones primarily used *in vitro* studies, but in the human membrane a variety of different tail lengths can be found.

It has been evident for many decades that the blood clotting cascade is heavily dependent on

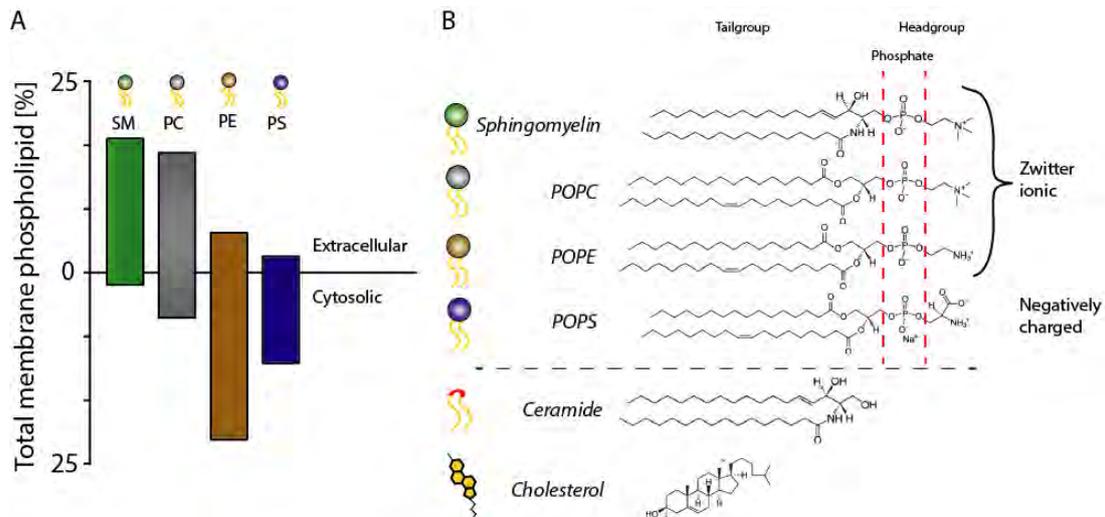


Figure 1.6: A: Distribution of the most common phospholipids (based on headgroup) as found in the human membrane. Numbers from [61, 62]. B: Chemical structures of relevant lipids.

the presence of a lipid membrane, and especially the fact that the PS enhances the activity of the TF:FVIIa complex [63, 64, 65]. It is widely accepted that PS is needed to obtain decrypted TF, but whether exposing TF to PS can cause full TF activity is still discussed [66]. In addition, the mechanism by which PS increases the activity of the TF:FVIIa complex is not understood and is still investigated. As mentioned earlier FVIIa has a  $K_d$  in the low  $\mu\text{M}$  for a membrane containing a mixture of PS (30%) and PC(70%), while no interaction occurs in with a pure PC membrane [26, 28]. Therefore, it is quite curious that the  $K_d$  of the interaction between FVIIa and TF in a membrane with an increasing amount of PS is not affected [18].

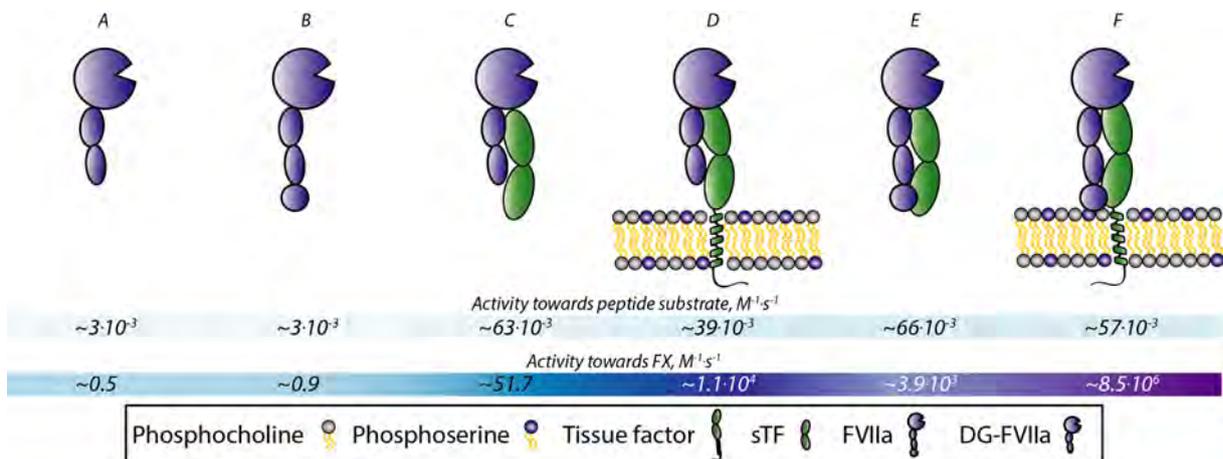


Figure 1.7: Activity of different FVIIa variants with or without different TF variants. Upper panel: Activity of the given FVIIa towards a peptide substrate. Lower panel: Activity of the given FVIIa towards FX. All values are from [26, 28, 27, 18].

A decrease in  $K_d$  is also observed for the FX:membrane interaction when increasing the PS content and no binding when the membrane only consists of PC lipids [19]. This probably correlates with the decrease in  $K_m$  of the FX activation by TF:FVIIa when increasing the PS content in the membrane, but other factors might also participate in the increased activity. An explanation for the increased activity upon PS exposure was proposed to be a PS induced allosteric

conformational change of FVIIa, which would alter the active site. This was investigated by comparing membrane-anchored TF bound FVIIa and DG-FVIIa cleavage of a peptide substrate mimicking FX [27]. From this study *Neuenschwander et al.* found that the activity towards the peptide substrate was identical for the two proteins (figure 1.7D and F), and concluded that the interaction between FVIIa and PS did not change the active site of FVIIa. On the other hand the activation of FX decreased 740-fold by removing the FVIIa Gla-domain (figure 1.7D and F) suggesting that the Gla-domain plays an essential part in the activation of FX [27]. It is also very interesting that activity towards the peptide substrate is identical for FVIIa and DG-FVIIa bound to sTF, but that the activity towards FX is three orders of magnitude higher for the FVIIa complex (figure 1.7C and E). This underlines the importance of the Gla-domain, not only for membrane binding, but also for substrate (FX) recognition.

To further asses these different scenarios, *McCallum et al.* investigated the height of the active site of FVIIa over the membrane [67, 68]. This was done by fluorescence resonance energy transfer (FRET) measurements between a probe attached to the active site of FVIIa and one attached to the lipid head groups. The height of free FVIIa on the membrane was determined to be  $83 \pm 3.3 \text{ \AA}$  (figure 1.8A), while when TF was in the membrane the height decreased to  $75 \pm 1.8 \text{ \AA}$  (figure 1.8B). This suggests that TF induces a change in the orientation of FVIIa towards the membrane, to a more tilted angle. The same height of the active site was found by an MD study performed on the TF:FVIIa complex on a membrane, supporting this theory [69]. Performing the FRET experiment with DG-FVIIa provided a distance of  $78 \pm 1.8 \text{ \AA}$  (figure 1.8C), which is within the error range of TF bound FVIIa, and thus these data suggests that the Gla-domain does not influence the FVIIa orientation when bound to TF [67, 68]. Potentially, the tilted orientation of FVIIa upon TF binding provides the correct height for accessing the scissile bond in FX, but as the structure of the tertiary complex is not solved, this is yet to be shown (figure 1.8D).

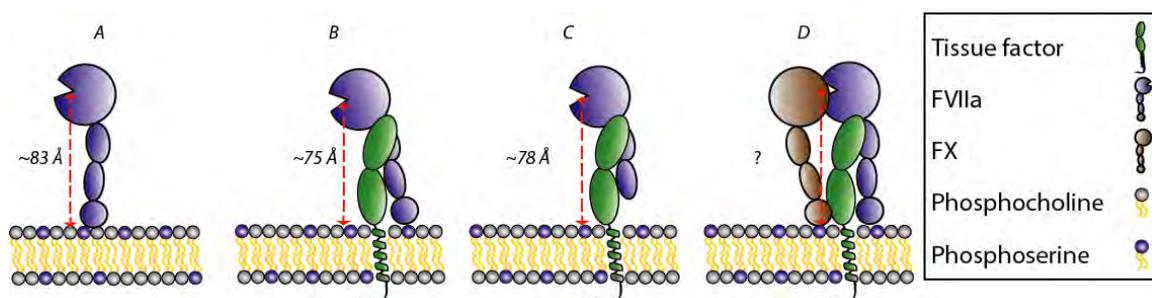


Figure 1.8: Distance between the active site of FVIIa and the membrane as measured by FRET [67, 68]. A: FVIIa on a membrane. B: TF:FVIIa on a membrane. C: TF:DG-FVIIa on a membrane. D: TF:FVIIa:FX on a membrane. Distance to the FVIIa cleavage site on FX is unknown.

Assuming that the TF:FVIIa and TF:DG-FVIIa complexes have the same orientation in relation to the membrane (figure 1.8B and C), the only thing guiding this orientation must be TF. As the only structural data available on the ECD and TMD of TF in a lipid bilayer are the FRET studies just presented, the mechanism by which TF guides the orientation is not known. Also the relationship between the orientation and the activity of the TF:FVIIa complex is not known yet. It could be hypothesized that the linker region between the ECD and the TMD of TF could be involved in controlling this orientation, but studies using an artificial linker attaching TF to the

headgroup of lipids, showed a similar activity to that of TF incorporated in a bilayer [70]. This suggests that neither the specific sequence of the TMD or the linker is important for maintaining the activity, but solely their ability to anchor TF to the membrane. It was speculated by *Ke et al.* that the similar activity was correlated with a similar orientation of TF with respect to the membrane, which then must then be controlled purely by the ECD. Therefore, they proposed that the ECD of TF could interact with the membrane, particularly with PS lipids which could potentially explain the PS induced increase in activity, by changing the orientation of TF. *K. Ke et al.* performed an Alanine scan on 20 residues located in the C-terminal region of sTF which they, based on the available structure of sTF, believed to interact with the membrane [71]. They found that FX activation was significantly decreased for eight of these mutants and confirmed that for all eight the amidolytic activity of FVIIa was still intact. This suggests that these eight residues are either involved in the orientation of the TF:FVIIa complex, in the recognition of FX binding, or both. To verify this hypothesis, molecular dynamics (MD) simulations were performed on sTF in the presence of a membrane containing PS lipids. Indeed, *K. Ke et al.* did find TF residues that were in contact with PS lipids more than 30% of the time and even a few residues which were in contact with PS more than 70% of the time of the simulation [72, 71]. Especially the loop containing the two residues, Lys<sup>165</sup> and Lys<sup>166</sup> which are expected to interact with FX, [73, 32, 74] showed many interactions with PS in the MD simulation. This suggests that this loop both mediates the interaction with lipids, and also functions as a docking point for FX.

After the discovery of the PS influence on TF activity, other membrane constituents have been investigated, with a main focus on phospholipids. PC lipids are usually used as balance when performing experiments with a fraction of PS present, as PC is by far the most common phospholipid in the outer leaflet of the cell membrane (figure 1.6). *Neuenschwander et al.* investigated the effect of PE, which are also present in the outer leaflet in small amounts [75]. They found that PE and PS synergize with the increasing of TF activity, and that only 3% PS was necessary to obtain full TF activity in the presence of PE, while 30% was needed without PE. Later studies showed that the Gla-domain of FVIIa contains a single specific PS binding site and multiple other phosphate binding sites [76], which lead *Tavoosi et al.* to propose the so-called ABC hypothesis - anything but choline [76, 77]. As the name indicates, the theory concludes that it is actually not specifically PS that is needed (at least in high amounts) to obtain full TF activity, but any phospholipid, without the bulky choline head group would be able to support the activity.

While the ABC theory was mainly directed at the discovery of the synergetic effect between PS and PE, it might be stretched even further as more recent studies suggests a decreasing effect of SM on TF activity [78]. While not being a classical phospholipid, SM still contains a headgroup which is identical to that of PC (figure 1.6B), and the inhibitory effect is thus in agreement with the ABC hypothesis. In a study by *Wang et al.* it was discovered that vesicles containing 50% SM only exhibited 20% TF activity compared to vesicles without SM, while samples with 20% SM showed similar activity to vesicles without SM [78]. As seen in figure 1.6A, the SM content on the outside of cells is roughly 20%, which should suggest that the decreased activity of TF induced by SM is a neglectable amount in normal cells. However, other studies have shown that TF is present in caveolae, which are a specific type of lipid rafts rich in proteins, cholesterol

and SM [79, 80, 81]. This finding asks the question of whether the local concentration of SM around TF on the cellular surface is higher than the average SM content, and in extension if it is possible that the SM induced decrease in activity actually occurs. To suggest a mechanism for the decryption of TF, *Wang et al.* investigated the effect of ceramide on TF activity. Ceramide is the hydrolysis product of SM, with the same tail group, but without the phosphocholine head (figure 1.6B). *Wang et al.* found that ceramide did not have a significant effect on the TF activity, suggesting that hydrolysis of SM to ceramide might be an efficient decryption of TF as illustrated in figure 1.9.

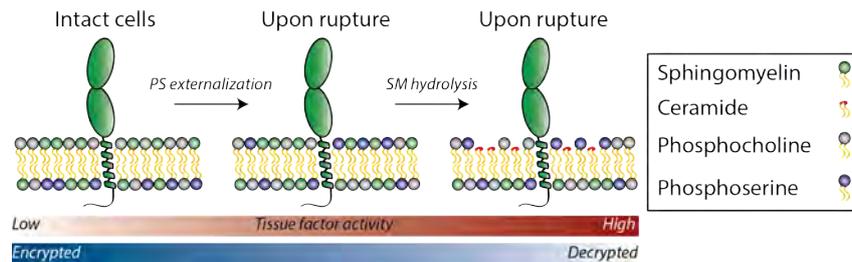


Figure 1.9: Schematic overview of the correlation between lipid content and encryption level of TF.

Lastly, cholesterol has also been suggested to influence TF activity, as it is also present in caveolae [81, 82]. Studies by *Dietzen et al.* showed that removing cholesterol caused disruption of the caveolae which lead to a 3-fold increase of the TF activity [79]. However, once again there are contradictory studies in the literature, as studies by *Awasthi et al.* showed that disrupting caveolae (albeit in another cell line) had no impact on TF activity [80]. A possible link between cholesterol and TF activity was proposed to be that cholesterol affects the cells ability to externalize PS, which would otherwise naturally decrease TF activity [83]. Yet again other studies suggest that the effect of cholesterol on TF seems to be independent of PS, and it is proposed that cholesterol modulates the interaction between TF and FVIIa [82]. As opposed to all the other lipids presented in this section, no biochemical studies investigating cholesterol effects on TF:FVIIa affinity or TF activity in a reconstituted systems have been published yet, also leaving this topic with a lot of controversies.

In conclusion, studies on lipid encryption has demonstrated conflicting results obscuring the overall picture. The only thing that is certain is that the activity of TF is highly dependent on the lipid environment and that the exact dependencies and mechanisms are far from well understood.

### 1.2.5 FVII activation

While the above sections have mainly focused on studies concerning FVIIa, the first step in the extrinsic coagulation pathway is actually the association between TF and FVII (figure 1.1). As mentioned, FVII is the zymogen of the activated FVIIa, and is the main FVII constituent in the blood as 99% is in the zymogen form while only around 1% is active FVIIa [84]. Therefore, the conversion of FVII to FVIIa is an essential step in coagulation, but compared to many of the later events, this step is more diverse. It is well established that many factors in the coagulation pathway are capable of activating FVII. Both IXa [85], FXIIa [86], Thrombin [87]

and FXa [87, 88, 89] (figure 1.10A) have been shown to be able to activate FVII and lastly, FVIIa can activate FVII in a mechanism called autoactivation [90]. It is well established that TF promotes both the activation and autoactivation of FVII [91] but whether the autoactivation is mediated by two FVII(a) molecules which dock onto the same TF (figure 1.10B) or two TF molecules each bound to one FVII(a) that are in close proximity (figure 1.10C) is currently not completely understood [90]. *Neuenschwander et al.* presented evidence that the main mechanism of autoactivation could be the one depicted in figure 1.10C. The most distinct difference between these two mechanisms is the presentation of the substrate, which in the case of mechanism B is dependent on the affinity between FVII and the substrate on TF, and as well as the concentration of FVII. In the case of mechanism C, presentation of substrate is dependent on lateral diffusion of the TF molecules and the density of TF molecules on the surface [9]. As mentioned *Neuenschwander et al.* found that enzyme kinetics correspond to the latter, however, it is very likely that the TF region which interacts with the Gla-domain of FXa [92] is also able to interact with other Gla-domains, such as that of FVIIa, and thereby also be able to mediate the autoactivation depicted in figure 1.10B which is also still suggested [93]. To distinguish between these two mechanisms, a system with TF in a state or system where lateral diffusion is impossible, is needed.

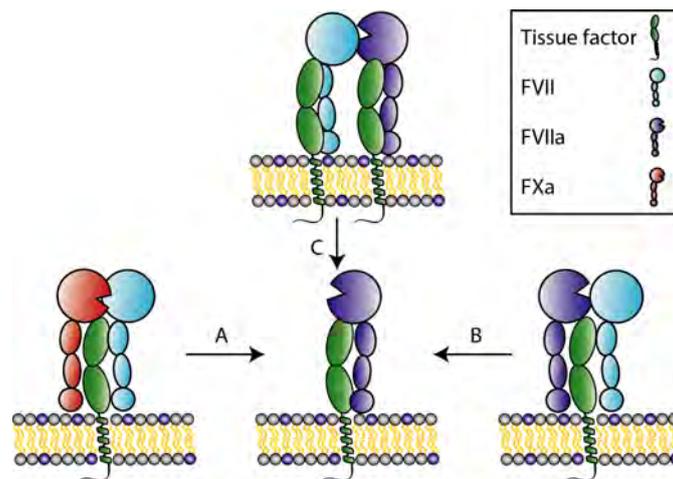


Figure 1.10: Schematic overview of different mechanisms of FVII activation. A: Activation by FXa. B: Activation by FVIIa docked onto same TF as FVII. C: Activation by FVIIa docked onto another TF molecule than FVII.

### 1.2.6 Inhibitors of clotting

As important as it is to be able to clot unwanted bleeding, it is equally important to be able to control and eventually stop the clotting. The primary inhibitor of the blood coagulation pathway is the tissue factor pathway inhibitor (TFPI), which is a 42 kDa protein containing three Kunitz-like domains each containing three disulfide-bonds [94]. TFPI is able to simultaneously inhibit the TF:FVIIa complex and FXa by forming a quaternary complex which inhibits FXa from participating in the prothrombinase complex [95]. The first Kunitz domain inhibits the active site of FVIIa, while the second Kunitz domain inhibits the active site of FXa [96]. The plasma concentration of TFPI in healthy adults has been reported to be approximately 2 nM [97].

As the number of people suffering from thrombosis is increasing, an increase in the research of

new anticoagulants has emerged in the last couple of decades [98]. *Stassen et al.* investigated the anticoagulant properties of the nematode *Ancylostoma caninum*, and found that three proteins (NAP5, NAP6 and NAPc2) were the main contributors to anticoagulation [99]. All three proteins were found to inhibit in much the same way as TFPI which is by forming a quaternary complex with TF, FVIIa and FXa. NAPc2 was found to form a high-affinity complex with FXa with a  $K_d$  of 1 nM [100] but when the crystal structure of the FXa:NAPc2 complex was solved, it revealed that the binding site on FXa was not in the active site, but in an exosite site [101]. This was also supported by biochemical studies showing that FXa in the quaternary complex TF:FVIIa:FXa:NAPc2 was still active on peptide substrates [100]. Other studies showed that NAPc2 bound equally well to FX and FXa, which both functioned as a scaffold for the inhibition of the TF:FVIIa complex [102]. This is in contrast to TFPI which only binds FXa, as the binding site is in the active site of FXa. The resulting FX:NAPc2 complex has a very strong inhibitory effect on the TF:FVIIa complex with an  $K_i$  in the low pM range [102].

The mechanism by which these binding events occur is still debated, but two mechanisms are the most commonly presented. As the interaction between TF and FVIIa and the interaction between NAPc2 and FX(a) are both very strong, it is broadly accepted that these two complexes are formed initially (figure 1.11) [102]. The FXa:NAPc2 complex can then either bind onto the membrane (figure 1.11A) and by translation encounters a TF:FVIIa complex and form the quaternary complex, or interact with a TF:FVIIa complex directly from solution (B).

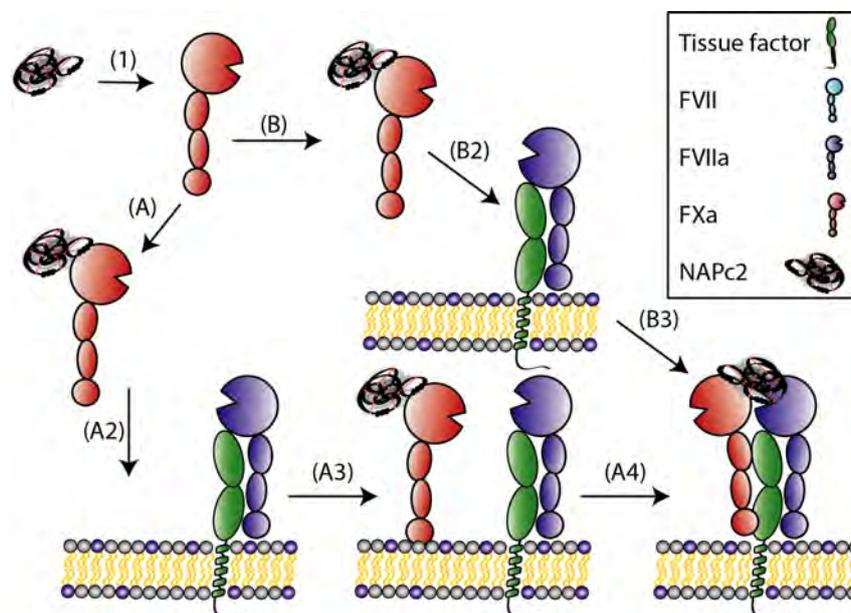


Figure 1.11: Schematic overview of NAPc2 mediated inhibition of FVIIa and FXa. A: NAPc2:FXa complex binds membrane surface and translate to find a TF:FVIIa complex to interact with. B: NAPc2:FXa complex binds directly to a TF:FVIIa complex from solution.

*Bergum et al.* investigated the importance of the FXa Gla-domain in the formation of the quaternary complex, and found that it was essential for the inhibitory effect on a membrane [102]. They found the affinity between DG-FXa and NAPc2 to be similar to that of FXa and NAPc2, but that the DG-FX(a) was not able to mediate the scaffold function. This suggests that the main function of FXa in NAPc2 mediated inhibition of FVIIa, is to stabilize NAPc2

at the correct position enabling it to interact with FVIIa. In addition it was shown that the inhibition was more effective in the presence of PS lipids, as these increase the affinity of FXa towards the membrane, leading to the mechanism illustrated in figure 1.11A.

The structure of NAPc2 was solved by NMR by *Duggan et al.* in 1999, and generally showed an unstructured conformation with a few small secondary structure elements [103]. The main reasons of the protein size was conferred by five disulfide bridges and a number of salt bridges formed by the high number of charged residues (40% of the sequences is charged). The low amount of hydrophobic residues lead to a lack of a hydrophobic core, and thus the protein is not folded with a compact core, which results in a certain flexibility. This flexibility is likely essential to the function of NAPc2, as it is to function in other organisms than the nematode itself. Therefore, a promiscuous behaviour with regards to binding partners is important, and it has been shown that NAPc2 also has a high affinity for bovine FXa [101]. Human and Bovine FXa are structurally very similar, and have a 70 % sequence similarity, which results in a  $K_d$  between Bovine FXa and NAPc2 of 266 nM. Currently, there are not determined affinities for other FX homologues, but FX among many animals have sequence similarities ranging between 60 and 80 % suggesting that these proteins might also confer a tight binding with NAPc2.

The binding interface between NAPc2 and FVIIa has not been determined by a high-resolution structure, but biochemical assays have suggested that the loop region of NAPc2 spanning Leu<sup>42</sup> to Val<sup>45</sup> binds to the active site of FVIIa [102]. Taken together the NMR structure and the information about the binding interface with FXa as obtained by the crystal structure (figure 1.12A), it seems that two flexible regions interact with FVIIa and FXa (figure 1.12B). However, it is also evident that a folding-upon-binding event occurs for NAPc2 upon interaction with FXa, and it is very likely that a similar event is happening upon FVIIa binding.

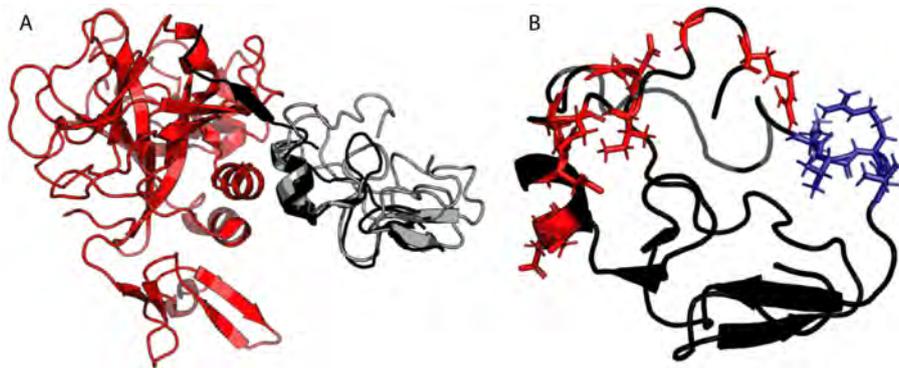


Figure 1.12: NAPc2 bindingsites. A: Crystal structure of NAPc2 (black) and FXa (red) (2H9E) overlaid with NMR structure of NAPc2 (grey) (1COU). B: NMR structure of NAPc2 (1COU) with binding site of FXa (red) and FVIIa (blue) marked.

### 1.2.7 Structure of the TF:FVIIa:FX:NAPc2 complex

The overall aim of this project was to obtain structural information about the TF:FVIIa:FXa complex. As mentioned, this complex is only formed transiently and only on a lipid bilayer which hampers the experimental possibilities substantially. The first issue was addressed using NAPc2 which locks the complex by the high affinity interactions, leaving a stable quaternary

complex. The second issue points towards a more general consideration when investigating an MP: Choosing a suitable carrier system. All integral MPs have a hydrophobic domain which is normally situated in the membrane, but when performing *in vitro* studies, a carrier system is needed to keep the MP in solution. Detergents are usually employed during purification, but as TF is highly dependent on the lipid environment, a carrier system containing lipids is necessary to investigate the TF:FVIIa:FXa complex. The nanodisc (ND) was the carrier system of choice in this project, and the following section will introduce this technology.

### 1.2.7.1 Nanodiscs

The ND is a soluble particle consisting of two amphipathic membrane scaffolding proteins (MSPs) surrounding a lipid bilayer patch [104, 105]. Besides functioning as a model-system for a lipid bilayer itself, it can also carry an MP inserted into the bilayer patch. The reconstitution of an ND is a self-assembly process in which the MSP, lipids (dissolved in detergent, usually cholate) and the MP of choice is mixed. After an equilibration time, the detergent molecules are removed, usually using detergent absorbing beads, and the NDs self-assemble (figure 1.13A). After reconstitution, the beads are removed and in order to separate empty (MP free) NDs from MP loaded NDs, a NiNTA purification is usually performed (if the MP contains a His-tag) followed by a size exclusion chromatography run, ensuring monodispersity of the final sample (figure 1.13B).

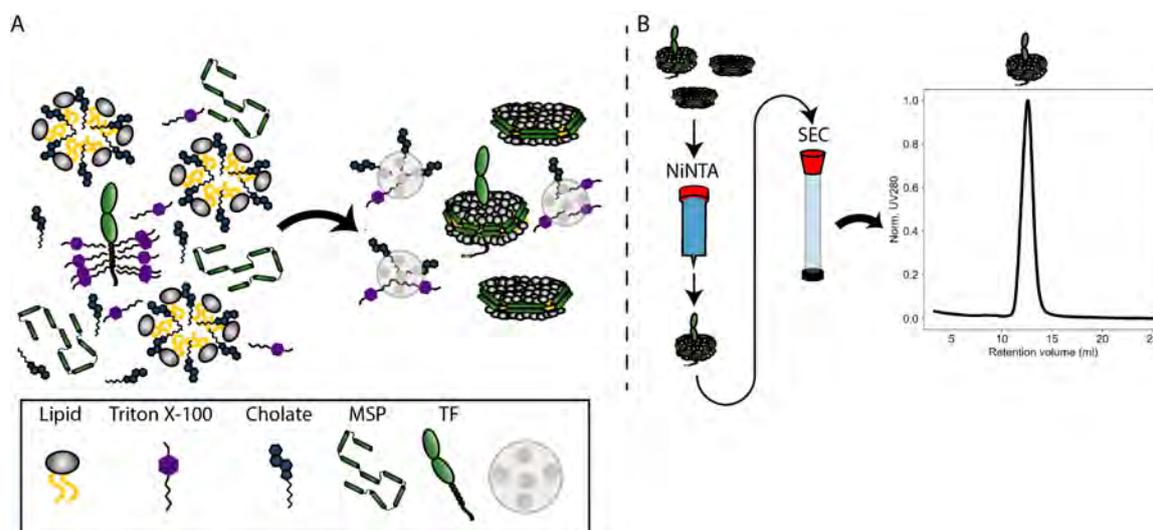


Figure 1.13: Schematic overview of the production of a TF loaded ND. A: Reconstitution of TF into an ND. B: Final purification steps of the production.

The circumference of the ND is to a large degree controlled by the MSP encompassing the lipids, and thus there have been attempts at adjusting the length of the MSP, as this would allow tuning of the final ND size [105, 106]. One of the first developed MSPs, MSP1D1, was derived from the protein Apolipoprotein A1, which mediates lipid transport in the human body [104]. The sequences consists of amphipathic helices termed helix 1 to helix 10, which are punctured by either proline or glycine linkers (figure 1.14). Most attempts at modifying the length of the MSP has been either elongating MSP1D1 with helix segments or deleting helix segments [105, 106].

Two of the most successful new sequences are MSP1E3D1 which is elongated by the three helices H4, H5 and H6 [105], and MSP1 $\Delta$ H5D1 in which H5 is deleted [106] (figure 1.14).

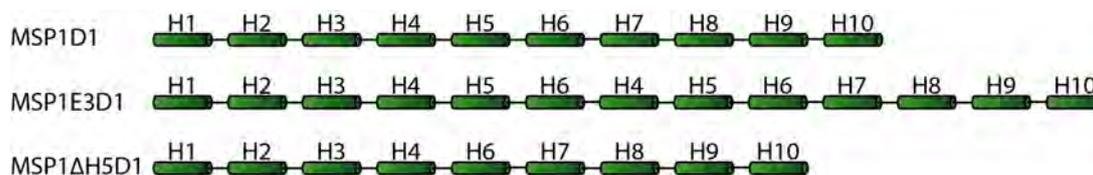


Figure 1.14: ND sizes: Overview of the three most common MSPs originating from MSP1D1.

Tuning the size of the ND is relevant for mainly two purposes; 1) the MP of interest needs to fit inside the ND, and especially if lipid interactions of either the MP or binding partners are of interest, it is desirable to have an excess lipid bilayer to mediate those interactions. 2) Some experimental techniques are sensitive to the size of the particle of interest. Eg. NMR is hampered by slow tumbling time of a large particle, and by decreasing the MSP length from MSP1D1 to MSP1 $\Delta$ H5D1 studies have shown a substantial increase in the peak sharpness of the embedded MP [106]. On the other hand electron microscopy usually provides the best data on large particles, and larger NDs are usually employed for this technique [107, 108]. For small angle scattering (SAS) experiments, the aim is usually to have an ND which is just large enough to fit the given MP, but does not have a large excess of bilayer, as the lipids will contribute considerably to the overall scattering signal, thus blurring out the contribution from the MP, hampering the final modelling of the MP.

During the last couple of years another major improvement has been developed by *Nasr et al.* who used the enzyme Sortase to circularize the MSP [109]. Sortase recognizes a five residues motif and a free N-terminal glycine, and links the two by forming a covalent bond (figure 1.15).

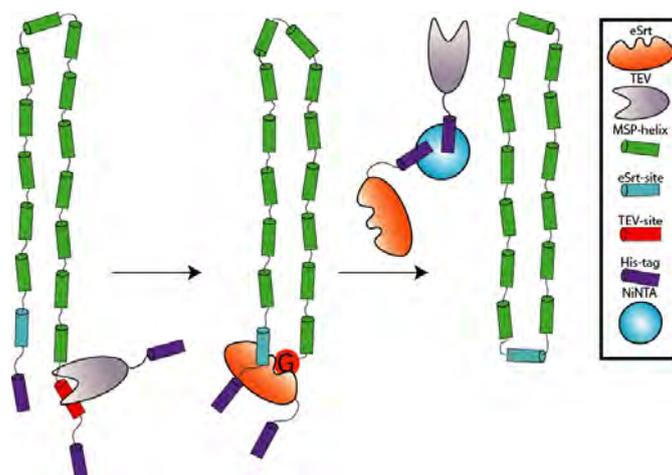


Figure 1.15: Circularization process of an MSP. First the MSP is cleaved by TEV protease, leaving a free Glycine which together with the eSrt-motif is recognized by the eSRT which forms a covalent bond. Lastly the two enzymes are removed using the His-tag. Adapted from Paper I.

After circularization of the MSP, the reconstitution of an ND can be performed using the standard protocol. The circularized nature of the ND improved the stability and induced monodispersity which are both desired properties, especially for structural studies [109, 110], but producing them in large amounts is challenging [111]. Later studies by our group, further developed

these circularized NDs, by introducing a substantial amount of negatively charged residues to enhance the solubility of the final ND (called circularized and solubility enhanced MSP1E3D1, csE3) (presented in Paper I) [110]. Besides improving the temporal stability of the NDs, the circularizing reaction was also enhanced by the extra charges, and yielded 75 mg of pure csE3 per liter of culture which was much more than reported in previous studies [109, 111].

NDs have been used for structure investigation of MPs using both NMR [112] Cryo-EM [108, 107] and SAXS [113, 114] and as the csE3 provides high stability and a relatively large bilayer it is ideal for investigation of the quaternary TF:FVIIa:FXa:NAPc2 complex.

### 1.2.7.2 Available structures

As mention throughout the last sections, parts of TF:FVIIa:FXa complex has been solved by X-ray crystallography (overview in table 1.2). There are many more deposited structure of the proteins in the Protein data bank (PDB), but only the first structures are presented in the table.

Protein(s)	PDB ID	Exp. method	Year	Comment	Ref
sTF	1BOY	X-ray	1994	Soluble TF	[115]
sTF:FVIIa	1DAN	X-ray	1996	Structure of the complex	[25]
DG-FXa	1HCG	X-ray	1993	Only DG-FXa	[116]
NAPc2	1COU	NMR	1999	The experiments are performed at a pH close to pI of the protein	[103]
FXa:NAPc2	2H9E	X-ray	2007	Only the binding interface of NAPc2 is determined	[101]
sTF:FVIIa:FXa	-	Computational model	2003	Based on docking of FXa on sTF:FVIIa structure (1DAN)	[30]

Table 1.2: Overview of a number of published structures of proteins in the TF:FVIIa:FXa:NAPc2 complex.

The 1BOY and 1DAN structures almost contain all residues of the given proteins, and show that the conformation of sTF is almost identical in the free and FVIIa bound state. As mentioned earlier there is no high-resolution data on the FL TF, but there is currently 43 structures of sTF deposited with different binding partners or different experimental conditions. More than 200 different structures have been deposited for FVIIa, but interestingly not a single FL structure of free FVIIa is available. It seems that the flexible nature of free FVIIa [72, 117, 57] is not suitable for crystallization, which also means that the possible structural rearrangement of FVIIa upon TF binding is still not described on an atomistic level. A similar issue exists for FXa, as a FL high-resolution structure is missing, yet 140 structures containing parts of FXa is deposited in the PDB. Most of the structures are of DG-FXa, and all of them are complexed with different inhibitors to stabilize the structure. One of the inhibitors is NAPc2, and the DG-FXa:NAPc2 crystal structure was published in 2007 [101]. The structure only contains electron density of NAPc2 in the binding interface, which shows more secondary structure than what is present in the structure of free NAPc2. In a study in 2003 *Norledge et al.* tried to assemble

the TF:FVIIa:FXa puzzle using all the structural pieces available at the time [30]. The main piece was the structure of sTF:FVIIa and a structure of FXa was modelled and docked onto the sTF:FVIIa complex (figure 1.16A). However, when overlaying the structure of FXa:NAPc2 which was published four years later (figure 1.16B) it was clearly visible that NAPc2 was not able to bind in the active site of FVIIa in the proposed conformation. Therefore, it is very likely that the proposed model in figure 1.16A is not representative of the actual complex, but structural data on the entire complex is needed to verify the model or propose a new one.

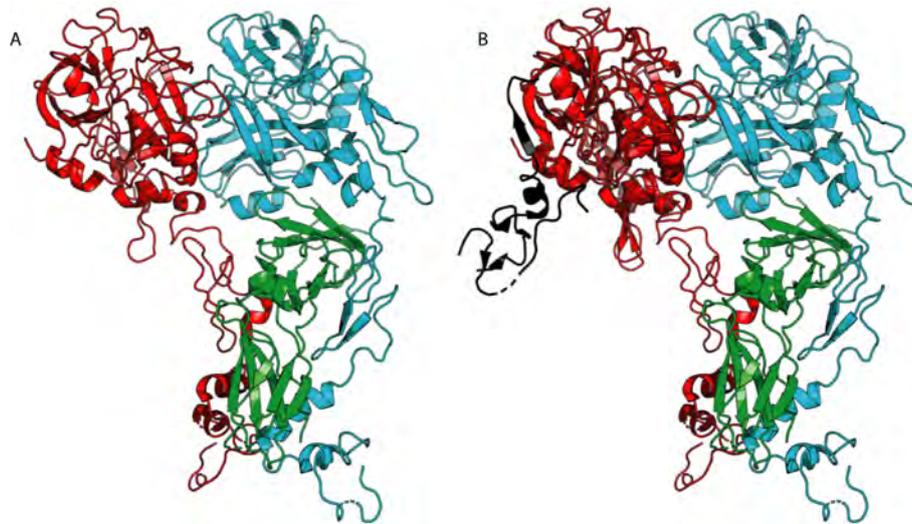


Figure 1.16: For both figures: sTF: Green. FVIIa: Teal. FXa: Red. NAPc2: Black. A: Model of the sTF:FVIIa:FXa complex [30]. B: Same model as in A, but with crystal structure of the FXa:NAPc2 complex overlaid [101].

### 1.3 Related diseases

While all the sections above retain to how the healthy body functions, there are also cases where parts of the cascade do not work, or work suboptimally. Table 1.3 shows four bleeding disorders which arise from lack of the proteins involved in coagulation. Haemophilia A, B and C arise due to the lack of FVIII, FIX and FXI respectively. The lack of any one of these factors effectively stops the coagulation cascade, and they will be discussed in greater detail later. In contrast, the Scott syndrome highlights a mechanism which is not yet well understood: The externalization of PS lipids. This disease was reported for the first time in 1979 [118] as a patient had frequent bleeding events, but normal levels of coagulation factors and platelets. However, the calcium dependent PS exposure was substantially decreased leading to a very ineffective clotting. About 30 years later it was discovered that the MP, TMEM16F, was essential for the PS exposure and subsequent genomic analysis revealed that Scott syndrome patients have a loss-of-function mutation in both alleles encoding TMEM16F [119].

Subsequently, it was determined that TMEM16F is a calcium-dependent scramblase, which makes it essential for regulating blood clotting activity [123, 124, 125]. This information provides another step in the mechanism of enhancing TF activity shown in figure 1.17. The finding of a 'molecular-switch' such as this scramblase which has this big an effect on blood clotting is

Disease	Defect in	Ref
Haemophilia A	FVIII	[120]
Haemophilia B	FIX	[121]
Haemophilia C	FXI	[122]
Scott syndrome	TMEM16F	[123]

Table 1.3: An overview of diseases related to missing proteins involved in blood coagulation.

potentially also very important in the development of anticoagulant drugs, as repressing the activity of TMEM16F would effectively lead to less clotting.

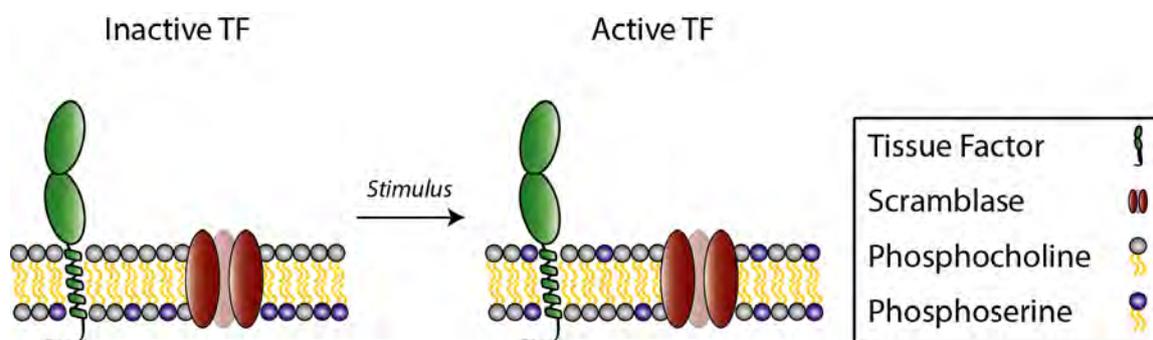


Figure 1.17: Schematic overview of the activity of the TMEM16F scramblase.

Haemophilia A, B and C are the most common bleeding disorders and those who have been studied the most. Hemophilia A is an X-linked disease and the most common bleeding disorder estimated to affect 1 out of 5000 males [126, 127]. The disease results in recurrent bleeding events usually in the joints, and is normally treated with intravenous FVIII upon bleeding [120]. Both Hemophilia B and C are also usually treated with injections of the missing coagulation factor, but patients carrying either of the three mutations are expected to have shorter lives and with a much higher number of hospital stays than the average person [121]. One problem is that patients can develop antigens for the injected coagulation factor, and in that case a backup treatment has to be used. In many cases this is injection of recombinant FVIIa, which does not solve the problem with the missing coagulation factor, but instead boosts the FX activation to a degree that will cause clotting [128]. The problem with this treatment is that it is only applicable upon bleeding and not to prevent bleeding, and the resulting amount of clotting can be difficult to control.

On the other hand excessive clotting (thrombosis) is also a major problem and in the United States more than 300,000 people die from thrombosis annually [129, 130]. When the coagulation system is active to a degree undesirable for the body, it is called hypercoagulation. The exact mechanism leading to this phenomenon is not fully understood, but it is well known that the environment, unhealthy habits and age influences the risk [131]. The unhealthy habits usually highlighted are smoking and alcoholism [131].

Many different drugs against thrombosis are available, normally called blood-thinners. The most common drugs are in the class of Heparins which are glycosaminoglycans mainly functioning by inhibiting Thrombin and FXa [132]. Also proteins have been investigated for their potential use

as anticoagulants, and in one study where NAPc2 was used to inhibit the unwanted bleeding of patients during and just after a surgery, but to date there is no approved NAPc2 drug on the market [133].

## 1.4 Outlook

Even though the coagulation system and the TF:FVIIa:FXa complex have been investigated for many decades, and the amount of literature on the topic is extensive, there are still some important mechanisms and structural features that are not yet understood. One of these is the increase of TF:FVIIa activity upon the presence of a membrane and especially PS exposure. The increase in activity was established many years ago, but the structural features and the mechanism explaining the increase, is still not fully elucidated. Similarly the large increase in FVIIa activity upon TF binding, is still not understood on a structural level. Two key pieces in the puzzle are missing: 1) Structural information about the complex in a lipid environment. As mentioned two FRET studies are available [67, 68], but more information in a lipid context is needed to fully elucidate the role of the lipids. 2) A high-resolution structure of the free form of FVIIa. As mentioned, obtaining a high-resolution structure of the free FVIIa is hampered by the flexibility of the free protein as shown by MD simulations. The structure of FVIIa was determined in complex with sTF and also with other binding partners which all introduces a certain rigidity in the protein favouring crystallization.

One of the reasons why it is difficult to clarify the role of the lipids in the increased activity, is that it is multi faced. The lipids participates in the binding of the substrate (FX) but likely also affects the TF:FVIIa complex (possibly the orientation). In addition, FX also binds the TF:FVIIa complex in a not fully understood way, making it difficult to pinpoint the individual contributions. It is likely that the heights of FX and FVIIa are important for the activity, as they most likely has to be correct before an optimal enzymatic reaction can take place. This height might be guided by the lipid bilayer, but more studies have to be performed to shed light in this. In addition, the components of the lipid bilayers used for the studies might have to be extended, as recent studies have suggests an influence of both SM and cholesterol.

Lastly, the structure of the TF:FVIIa:FXa complex still remains the missing cornerstone, although multiple mutational studies provides information on the binding interface of TF:FXa. However, a full structure of the complex on a lipid bilayer will provide invaluable insights to the mechanism of the lipids, and also to provide a basis for development of new drugs targeting this complex. As a lipid bilayer is needed to form the complex, a carrier system containing lipids and TF have to be utilized. This could potentially be an ND which also would provide an excess lipid bilayer for FVIIa and FXa binding. In addition, an inhibitor such as NAPc2 is needed to stabilize the complex. However, the size and flexibility of the system makes it difficult to investigate, as an ND has never been crystallized, and the complex will exceed the size limitation of NMR, other techniques such as SAS and cryo-EM has to be applied.

## 2 | Experimental methods

### 2.1 Sample production

Even though this project has been conducted at the physics department, a major part of the projects have had a more protein chemistry-oriented focus, which has been aided by the use of many biophysical experiments. In this section, some of the considerations done before starting up this project as well as some of the adjustments made along the way will be described. This will include protein purification, the considerations laying the foundation for the protein design as well as sample preparation for structural analysis.

#### 2.1.1 Designing constructs

Generally speaking, this project has revolved around two different kinds of proteins: *i*) proteins with a biologically relevant sequence, which has been investigated to obtain information about their structure and function in the human body and *ii*) proteins which solely purpose is to function as MP carriers. The use of proteins in the *i* class is the main focus of Paper IV, V, VI, and Report I, while the *ii* class is the focus of Paper I, II, and III. The considerations behind the construct designs are critical in both cases but usually more difficult in the *ii* case. In the *i* case, the first consideration should be if working with the full length (FL) protein or if a truncated version is preferable. The obvious reason to work with the FL protein is that it represents the native situation the best, while working with a truncated version can be selected to make either the expression and purification or subsequent experiments easier (or feasible at all). In this project, a truncated version of TF lacking the ICD has been used for all experiments. As mentioned earlier, it has been shown that the ICD is not important for the coagulant properties of TF, which was the main focus of this project. In addition, the ICD contains a free Cys which can potentially cause problems with undesired dimerization during purification. The next consideration should be whether a tag is needed for purification. In the purification of most proteins produced recombinantly, a tag is very convenient. In this project, the His-tag has been used for almost all proteins, and has proved very valuable, as it has provided a relatively high purity. In addition, it is functional under denaturing conditions, which is used in a fraction of the used protocols. Many other tags are available, and there is generally proportionality between the provided purity and the cost of the resin (high cost = high purity). In addition, a linker region should be considered as the tag needs to be available to the resin used for purification. A short sequence of flexible residues is usually selected as a linker, ensuring the desired reach

and flexibility of the tag. Commonly used residues are Gly and Ser, which are often seen in the combination called  $G_4S$  (Gly-Gly-Gly-Gly-Ser) [134].

### 2.1.1.1 Introduction of charges in the MSP sequence and joining eSrt

In case *ii*, it is difficult to outline a general strategy, as the purpose of the protein and the desired properties ultimately dictates the design. Therefore, this paragraph is a walkthrough of the thoughts behind the constructs designed during this project.

The development of the csE3 MSP was based on the idea that it should be possible to optimize the sequence of the MSP, as the MSP sequences are simply derived from the human apolipoprotein A1 (ApoA1), which posses other functions in the body than acting as MP carrier [104]. Previous studies on optimized MSPs were conducted using truncated or elongated versions of MSPs, but still using the original MSP1D1 sequence [105, 106]. As even a small amount of aggregation in SAS experiments will have a large impact on the final scattering signal, our major aim was to produce more stable and if possible also more monodisperse NDs. As an article on circularized NDs was just published [109], demonstrating production of more monodisperse NDs, we decided to continue this development using Sortase for circularization. In addition, we decided to introduce a considerable amount of negatively charged residues, as previous studies (on completely unrelated proteins) have shown these residues to induce solubility [135, 136]. We used the original MSP1E3D1 sequence [105] as a template and substituted all *Thr*  $\rightarrow$  *Ser*, all *Asn*  $\rightarrow$  *Asp* and all *Gln*  $\rightarrow$  *Glu*. With that, the number of negatively charged residues increased from 53 to 73, changing the pI of the protein from 5.8 to 5.0. Of course more charges could be introduced, but by only including these specific mutations, we tried to conserve the overall size of the residues in order to 'disturb' the MSP structure the least possible. As described in more detail in Paper I, the extra charges increased the temporal stability of the final ND, and also aided in providing a high yield from the circularization process.

To further develop on the circularized and solubility enhanced MSP (cs-MSP), a fusion construct was designed containing both the MSP and the eSrt in one construct, in order to ease the production of samples (figure 2.1A). Two main objectives were addressed using this strategy; *i*) circumventing the need for producing the sortase in parallel and *ii*) decreasing the amount of dimers and oligomers created during circularization.

A problem with the csMSP is that, besides circularizing a MSP, eSrt can also open an already circularized MSP and form a dimer by attaching it to a neighboring MSP. This is most likely part of the explanation of the more than 35% of dimers and oligomers formed when performing the reaction at high concentrations (see Paper I). In Paper I, the issue was solved by decreasing the concentration, which decreased the amount of dimers and oligomers. However, this leads to a very large volume increase, which was a tedious step in the production, and the hope was that we could avoid the need for large volumes by using the fusion construct for an on-column strategy.

The idea was to perform on-column circularization and purification in one step (see figure 2.1) after the TEV cleavage was performed in dialysis (as performed for csE3 production, see Paper I).

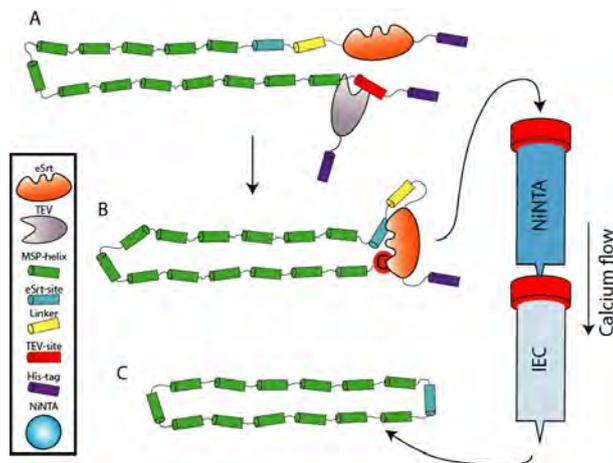


Figure 2.1: Fusion construct design and proposed production protocol. A: Overview of the construct. B: After TEV cleavage, a free Gly is recognized by eSrt which performs the circularization. C: Circularized product.

As the eSrt is reported to be calcium-dependent, the idea was to load the TEV-cleaved fusion construct (figure 2.1B) on the NiNTA column and then run a buffer with low salt content ( $\sim 80$  mM) and 10 mM calcium over the column. This should circularize the MSP, which would then be released from the NiNTA column and bind the IEC column (the eSrt would still be attached to the NiNTA by the His-tag). After end circularization, the NiNTA column would simply be removed, and the circularized product eluted from the IEC column using a salt gradient. The hope was that the MSP would be released from the column as soon as the reaction was performed as this would circumvent the possibility of an additional unwanted reaction. In addition, we believed that by joining the MSP and eSrt a proximity effect would be present, meaning that the chances of the two termini of the MSP meeting would be higher than two different MSPs meeting, thus resulting in a majority of monomeric circularized MSPs.

However, as will be described in a later section (and Paper II), we faced some unforeseen problems with the fusion construct, which resulted in a change of plans and production strategy. This is often the case when designing new constructs from scratch, and usually much optimization (and updated constructs) are needed before a construct comprising all the desired properties is obtained. Therefore, the on column circularization has not yet been performed, but it was possible to produce csE3 from the fusion construct without producing Sortase in parallel.

### 2.1.1.2 Tuning the MSP length

Besides the csE3 and the fusion construct, a few other attempts at optimizing the MSPs were conducted, which did not yet result in functional MSPs. However, the ideas and considerations will be explained in the following section. All the new MSPs contained the mutations to negatively charged residues and the eSrt motif in the same manner as csE3.

A primary goal in the ND field has been to develop discs of different sizes, which are useful for different purposes. The most commonly used MSPs are shown in figure 2.2A, which are all based on the same sequence (MSP1D1) but either elongated with three helices (MSP1E3D1) or decreased in length by one helix (MSP1 $\Delta$ H5D1). The helices in MSP1D1 are relatively well

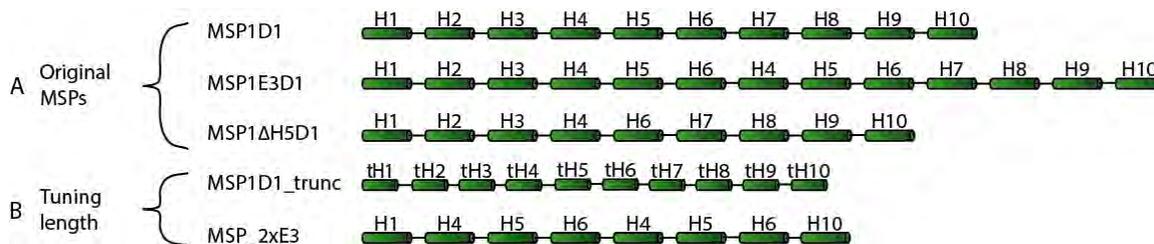


Figure 2.2: Construct design. The 'Original MSPs' are included for comparison. New attempts of tuning the length were made (Tuning length). tH1: truncated Helix 1.

determined, as they are separated by Pro or Gly residues. The helices vary in length but are all amphipatic in nature. Previous studies using a number peptides corresponding to truncated versions of one of the helices from ApoA1 showed that even peptides consisting of only 14 residues (compared to the initial 22 residues) were able to form lipid containing discs [137]. Peptides shorter than 14 residues were not able to solubilize the lipids, while all peptides longer than 14 residues were able to form discs. Whether this is also true for MSPs is not certain as linking the helices provides different properties, but it is worth considering when designing new constructs. Therefore, we decided to try shortening each helix by one turn (e.g. four residues), thus maintaining the amphipatic nature, as well as the number of helices (which were all 14 residues or longer) and punctuations (MSP\_trunc, figure 2.2B).

One of the most succesful MSP designs is the MSP1E3D1 which seems to be stabilized by the repeat of H4, H5, and H6. Thus we decided to design a construct with these three helices as the major part of the sequence and only conserve the terminal helices. We started by designing the MPS\_2xE3 (figure 2.2B), and the idea was that it could potentially be elongated to MSP\_3xE3 and so on. Both of these new MSPs were expressed and purified using the standard MSP protocols, but they yielded poor amounts of the final product (around 5 mg per liter of culture, data not shown). In addition, an initial experiment showed that the MSPs did not form NDs with POPC lipids, as a larger particle, which ended up in the void of the SEC column, emerged instead of a peak at the expected size (performed by a Masters student, data not shown). For now, no further experiments have been conducted, as changing the constructs is necessary. Maybe a less dramatic change would be a better starting point. Although the final goal is to make significant changes in the sequence, smaller changes might be a better starting point, as these might not affect the expression levels as significantly.

### 2.1.2 Purification of proteins

Throughout this project, establishing solid protein production strategies has proven to be the most critical factor in obtaining good results, although the process it is rarely described in much detail in the final papers. Paper IV and VI describe two new protocols for TDF and NAPc2, respectively. In this section, a brief overview of the different considerations, advantages, and disadvantages related to different strategies will be given (all strategies are outlined in table 2.1). All proteins produced in this project have been made recombinantly in *E. Coli*, and thus only strategies using *E. coli* expression will be discussed.

The protocols for producing the two enzymes TEV and eSrt are probably the most straight forward methods utilized in this project. They rely on expression of the protein in a soluble conformation, followed by cell lysis and an affinity purification using the His-tag. For both TEV and eSrt this provides an ample amount of protein with a purity acceptable for use in enzymatic reactions. A similar protocol was used for NAPc2 in Paper V apart from the use of a special cell line *BL21 Origami* which provides an oxidizing environment to support the formation of the five disulfide bridges in NAPc2. Before NAPc2 was ready for experiments, the His-tag was removed using TEV, and a SEC run was performed to remove minor impurities as well as the fraction of the protein present as a dimer. The important parameters for the expression of TEV, eSrt and NAPc2 have been time and temperature, as lowering the temperature to 20 °C and inducing expression for around 20 hours increased the yields of all three considerably, compared to expression at 37 °C. TEV and eSrt have been used for almost all projects presented in this thesis and have been produced throughout the three years from solid protocols available beforehand [109, 138]. Most earlier experiments on NAPc2 have been performed using a *Pichia pastoris* produced protein [100, 101], while a single example of an *E. Coli* produced protein is present [139]. In this project, we further developed the *E. Coli* based strategy and simplified the protocol (see details in Paper V). When introducing a new production protocol it is always important to prove that the final product possesses the same properties as the same protein produced using previous protocols. Therefore, the inhibitory efficiency of NAPc2 was probed in two different assays confirming native activity (Paper V).

Similar to the protocol for TEV and eSrt, the protocol to produce MSPs (which is also similar to the protocol used for the Fusion construct) has been used extensively in this project. The protocol has been described elsewhere and has only been adjusted slightly (see details in Paper I). The protocol is based on expression at 37 °C, which provides a very high yield but also a fraction of the MSPs produced as inclusion bodies (IBs). However, by lysing cells in 6M GuHCl, all proteins are solubilized and unfolded. Affinity chromatography is performed under these denaturing conditions, and the protein is refolded on the column. As all proteases are unfolded until purification is done, the risk of undesired degradation is minimized. Refolding on a column is probably only feasible for the MSPs as they contain no tertiary structure, and only helices have to fold.

Developing the protocol for the production of TF was the single most time-consuming project presented in this thesis. The aim was to establish a protocol providing ample amounts of protein for structural studies, which was not possible with the protocol available at the time [140]. The original protocols are all based on expression in a membrane-anchored state, with a subsequent solubilization of membranes in detergent [140]. We pursued a strategy involving expression in inclusion bodies, and a subsequent refolding. To begin with, the expression levels were relatively high, but the refolding yields were tiny. This is a commonly encountered problem [141], and after much optimization, the refolding yields increased, enough to yield protein in sufficient amounts for functional and structural studies (see details in paper IV).

In conclusion, the production strategy ultimately depends on the nature of the given protein, and if a protocol is not established previously, it is likely that different attempts are needed before successful production is obtained. In this project, a variety of different strategies have been used, and much time has been spent on optimizing some of these.

Expression	Purification	Proteins	Used in
20 °C BL21 Star	His-tag	TEV and eSRT	Paper I, II, IV, V, VI and Report I
20 °C BL21 Origami	His-tag	NAPc2	Paper VI and Report I
37 °C BL21 Star	Refolding on column using His-tag	All MSPs and Fusion construct	Paper I, II, IV, V, VI and Report I
	Refolding in reservoir and His-tag	TF	Paper IV

Table 2.1: Overview of protein production strategies used in this project. Details of the different protocols can be found in the relevant papers.

### 2.1.3 Production of samples

In many research fields, protein purification would be the last step before conducting the final experiment, but for this project, additional sample preparation was usually required. As working with MPs often involves incorporation of lipids into the sample of interest, the final sample preparation was therefore usually a self-assembly process. This either involved the incorporation of the MP of choice into a carrier system or just self-assembly of the MP-free carrier system. The work presented in this thesis mainly revolved around NDs and peptidediscs, which are produced using two different self-assembly protocols (see overview in figure 2.3) [104, 142]. One of the reasons for the two different protocols is that the peptides were able to solubilize the lipids on their own, while the MSPs, needed to be assisted by a detergent initially.

Both processes start with the purified MP in detergent. Subsequently, for the ND, MP, MSP and detergent-solubilized lipids were mixed and incubated before detergent absorbing beads were added (described in detail in Paper IV). After incubation, the beads were removed, and both preparations were applied to a NiNTA column to capture only the MP containing discs. For the peptide discs, MP was added to a lipid film and subsequently incubated, before purifying the MP containing discs using a NiNTA column (described in detail in Paper III). The NiNTA elutions were then analyzed by SEC to verify the existence of a single species and remove potential aggregates/impurities before the sample was ready for experiments.

Many parameters can be optimized in the self-assembly process including the ratio between lipid and MSP, the ratio between MSP and MP, temperature, and incubation time. Usually the effects are evaluated by SEC, and the success criteria is usually a single symmetrical peak eluting at the expected volume. Both optimization of TF reconstitution into NDs and peptidediscs have been performed in this project, which was necessary to obtain useful data in many of the subsequent structural studies.

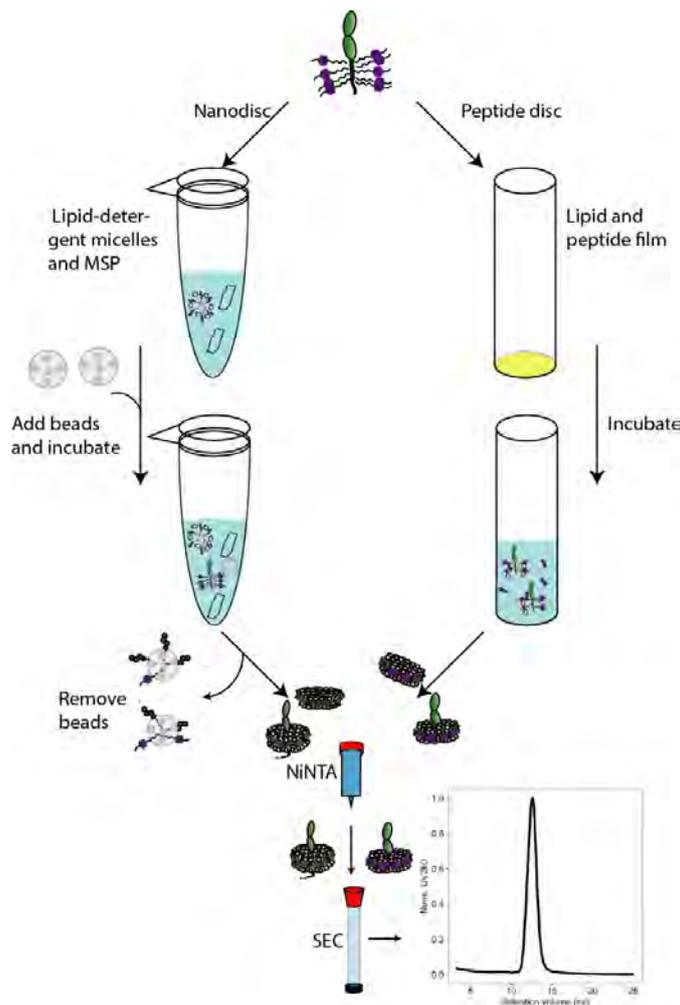


Figure 2.3: Incorporation of MPs in carrier systems. Left: Incorporation of TF into nanodiscs. Right: Incorporation of TF into peptidediscs. See details in main text.

## 2.2 Small angle scattering

Small angle scattering (SAS) has, for many decades, been used to probe the size and shape of molecules. The method has developed extensively since the early experiments in the 1940s [143] to the large scale facilities performing thousands of measurements each year today. Due to the relatively large size span this technique covers, a wide variety of samples ranging from superconducting material, over polymers, and purified proteins can be measured, the latter being the focus of this project. In contrast to methods such as X-ray crystallography and electron microscopy, SAS experiments on proteins are performed in solution. This is an advantage as well and a disadvantage. The fact that the experiments are performed in solution and under physiological temperatures makes the comparison to a native situation more reliable. On the other hand, the free tumbling of the molecules lowers the information content of the final data, leading to problems of resolving the 3D structure of the molecule of interest, and necessitates the combination with extensive structural modeling.

The scattering experiments performed in this project have either been conducted using X-rays

or neutrons, hence called either small angle X-ray scattering (SAXS) or small angle neutron scattering (SANS). Most of the theory presented in this section is the same for X-rays and neutrons and only differ slightly. In this project, many samples have been investigated with both techniques as they provide different information on the same sample due to the fact that X-rays are scattered by electrons and neutrons are scattered by the nuclei of an atom. This information is highly complementary, and it is an advantage to obtain both, especially in the modeling of more advanced samples.

In this section, the general theory behind a SAS experiment will be explained, as well as the initial data treatment and methods for a more in-depth analysis. This theory is well explained in the following reviews and books which have been used for the entire section [144, 145, 146, 147].

### 2.2.1 The scattering vector

A monochromatic incident beam of X-rays or neutrons can be described as a wave  $\mathbf{k}$  with the wavelength,  $\lambda$ , and the length of  $\mathbf{k}$  given by  $|\mathbf{k}| = 2\pi/\lambda$ . When the wave interacts with a scatterer (red point, figure 2.4) an elastic scattering event can occur, which gives the scattered wave vector  $\mathbf{k}'$  with an angle of  $2\theta$  relative to the incident direction. For all elastic scattering events it can be assumed that  $|\mathbf{k}| = |\mathbf{k}'|$ . From this, the scattering vector  $\mathbf{q}$  is defined as  $\mathbf{q} = (\mathbf{k} - \mathbf{k}')$  and from this it follows that

$$q = |\mathbf{q}| = 2|\mathbf{k}|\sin(\theta) = \frac{4\pi\sin(\theta)}{\lambda} \quad (2.1)$$

If scattering occurs by two scatterers separated by the distance  $\mathbf{r}$ , then the phase difference is  $\mathbf{k} \cdot \mathbf{r} - \mathbf{k}' \cdot \mathbf{r} = \mathbf{r} \cdot \mathbf{q}$  which explains the angle dependent modulation of the scattering.

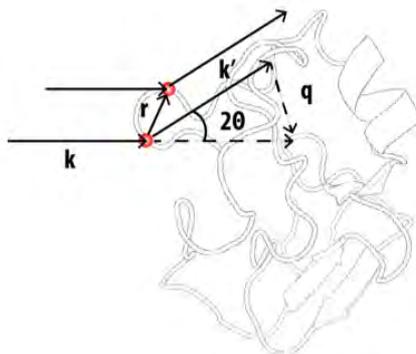


Figure 2.4: Scattering geometry. See main text for discussion.

### 2.2.2 Scattering lengths

As mentioned above, the scattering from X-rays arises from the interaction between the electric field of the X-ray photon and the electric field of the electron in the particle. On the other hand, the neutron scattering occurs from interaction with the nuclei of the atom. This means that different isotopes of an atom will scatter X-rays in a similar fashion, while the scattering of neutrons can be very different.

The amplitude of a given scattering event,  $E_{scat}$  is dependent on the scattering length,  $b$ , as given by equation 2.2:

$$E_{scat}(\theta) = \frac{E_{in} \cdot b}{R} \sqrt{\frac{1 - \cos^2(2\theta)}{2}} \quad (2.2)$$

Here  $R$  is the distance of the scattered photons or neutrons and  $E_{in}$  is the amplitude of the incoming beam. The last term in equation 2.2 is only valid for X-rays and is not included for neutrons. From this equation, it is seen that the unit of  $b$  is a length, usually provided in *cm*. For small angles of  $\theta$ , the last term in equation 2.2 approaches unity, and  $E_{scat}$  can be assumed to be independent of  $\theta$ . For larger values of  $\theta$ , the  $E_{scat}$  will be affected, and the so-called atomic form factor will be included in equation 2.2.

The scattering lengths of atoms commonly found in the samples used in this project can be seen in table 2.2. The values for X-ray scattering lengths are simply dependent on the number of electrons, and increase linearly, while the neutron scattering lengths are isotope dependent and empirically determined [148]. For neutrons, both coherent and incoherent values are provided. The coherent scattering contains the structural information, which is of interest, while the incoherent simply gives a background signal for the experiment. The incoherent values are only provided for H and D, as the values for the rest of the elements are negligible.

Atom	$b \cdot 10^{-13}, cm$		
	X-ray	SANS <sub>coh</sub>	SANS <sub>incoh</sub>
H	2.82	-3.740	25.27
D	2.82	6.671	4.04
<sup>12</sup> C	16.92	6.646	
<sup>14</sup> N	19.74	9.360	
<sup>16</sup> O	22.56	5.803	
<sup>31</sup> P	42.30	5.130	
<sup>32</sup> S	45.12	2.847	

Table 2.2: Scattering lengths of atoms commonly found in samples used in this project. Values for neutrons are found in *Sears, 1993* [148]

When performing a SAS experiment it is not the amplitude of the scattered X-rays or neutrons,  $E_{scat}$  that is measured, but the intensity of the scattered X-rays or neutrons,  $I_{scat}$ , hitting the detector. As  $I_{scat} = |E_{scat}|^2 \cdot A_{detector}$ , where  $A_{detector}$  is the area of the detector hit, and the incoming amplitude,  $E_0 = |E_{in}|^2$ , then the ratio between the scattered and the incoming X-rays or neutrons is

$$\frac{E_{scat}}{E_0} = \frac{|E_{scat}|^2 \cdot A_{detector}}{|E_{in}|^2} = \frac{|b|^2 \cdot A_{detector}}{R^2} = |b|^2 \Delta\Omega \quad (2.3)$$

where  $\Delta\Omega$  is the solid angle spanned by the detector at a distance  $R$ . This can also be expressed as in equation 2.4 which gives the ratio between scattered and incoming X-rays or neutrons in the detector direction.

$$\frac{E_{scat}}{E_0 \Delta\Omega} = |b|^2 \quad (2.4)$$

To obtain the full scattering from a sample in all directions, all angles will have to be measured:

$$\int_{\Omega} |b|^2 d\Omega = 4\pi |b|^2 = \sigma \quad (2.5)$$

Which implies that:

$$\frac{d\sigma}{d\Omega} = |b|^2 \quad (2.6)$$

where  $\sigma$  is the scattering cross-section, while  $\frac{d\sigma}{d\Omega}$  is called the differential scattering cross-section. This is the value that we measure in the scattering experiments, and it is typically reported in  $cm^2$ , or more commonly in scattering cross-section per unit volume ( $cm^{-1}$ ).

### 2.2.3 The form factor

In order to calculate the scattering intensity, or differential scattering cross section from a particle we have to have a good model for its structure. The scattering amplitude of a particle changes with the angle of view and the distance between the scattering atoms in the particle. All atoms that scatter will contribute with their scattering length,  $b$  and a phase factor  $e^{-i\mathbf{q}\cdot\mathbf{r}}$ . This can be summed up to give the expression for the scattering amplitude of a particle consisting of  $n$  scattering atoms:

$$A(\mathbf{q}) = b_1 e^{-i\mathbf{q}\cdot\mathbf{r}_1} + b_2 e^{-i\mathbf{q}\cdot\mathbf{r}_2} + \dots + b_n e^{-i\mathbf{q}\cdot\mathbf{r}_n} = \sum_{j=1}^n b_j e^{-i\mathbf{q}\cdot\mathbf{r}_j} \quad (2.7)$$

Knowing the scattering lengths of all the atoms in the particle, allows one to calculate the scattering length density (SLD),  $\rho$ , which is simply the sum of the scattering lengths divided by the volume of the particle. However, this is more difficult if the particle contains multiple contrasts (e.g., the ND), which will be discussed in greater detail in a later section. All SAS experiments conducted in this project are on particles in solution. In this case, the solution will also contribute to the overall scattering, which has to be taken into account by subtracting the SLD of the solvent, to provide the excess SLD:

$$\Delta\rho = \rho_{particle} - \rho_{solvent} \quad (2.8)$$

The excess SLD is often called the contrast of the system and is a critical parameter to have in mind when designing an experiment. Assuming that the contrast function  $\Delta\rho(\mathbf{r})$  is homogeneous, equation 2.7 can be rewritten as an integral over the volume,  $V$ , of the particle of interest:

$$A(\mathbf{q})_{particle} = \int_v \Delta\rho(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \quad (2.9)$$

It is important to note that the amplitude can not be measured and instead the form factor intensity of particle in a sample can be described as follows:

$$I(\mathbf{q})_{particle} = A(\mathbf{q})_{particle}A(\mathbf{q})_{particle}^* = A^2(\mathbf{q}) = \frac{d\sigma(q)_{particle}}{d\Omega} \quad (2.10)$$

where  $A(\mathbf{q})$  is the form factor amplitude (as calculated in 2.9) and  $A(\mathbf{q})^*$  is its complex conjugated. This relationship will become important later on.

When particles are in solution, they are free to rotate, which means that upon a measurement, we can assume a random orientation of all particles. In practice, this means that the signal measured upon a SAS experiment is orientationally averaged to yield  $I(q)$  instead of  $I(\mathbf{q})$ , which lowers the information content of the data. In addition, if the solution is sufficiently diluted, the scattering intensity of the sample can be described as:

$$I(q) = n\Delta\rho^2V^2P(q) \quad (2.11)$$

Here,  $n$  is the number of particles in the sample,  $\Delta\rho$  is the excess SLD, as calculated in equation 2.8,  $V$  is the volume of each particle, and  $P(q)$  is the normalized form factor intensity. The  $P(q)$  term contains all the structural information about the particle of interest and can be described by the following:

$$P(q) = \langle I(q)_{particle} \rangle = \frac{\langle (\int_V \Delta\rho(\mathbf{r})e^{i\mathbf{q}\cdot\mathbf{r}}d\mathbf{r})^2 \rangle}{M} \quad (2.12)$$

where  $M$  is the total scattering mass,  $M = \int_V \Delta\rho(\mathbf{r})d\mathbf{r}$ , and the brackets mean that it is orientationally averaged. Following equation 2.10,  $I(q)$  is the square of  $A(q)$  and thus all information about phases is lost. This, together with the orientational averaging, means that the data cannot be transformed into a full 3D SLD function,  $p(\mathbf{r})$ , which confers one of the major struggles in analyzing SAS data.

For all samples measured during this project, an attempt was made to use a concentration that did not introduce interparticle interaction, but at high sample concentrations, these can occur. The interactions can either be repulsive or attractive, and they will affect the obtained data in different ways, which are both undesired. For samples in this concentration range, an additional term has to be included in equation 2.11, that is the effective structure factor,  $S(q)$ , and the scattering intensity can then be described as

$$I(q) = n\Delta\rho^2V^2P(q)S(q) \quad (2.13)$$

Except for a few ideal cases, the  $S(q)$  term is almost impossible to calculate as it is dependent on both the shape, size, and orientation of the particles. This makes it highly desirable to avoid this contribution. This is done by lowering the concentration, and thus obtaining useful data is always a balance between avoiding interparticle interactions and having a high enough concentration to obtain good statistics.

### 2.2.4 Contrasts

As mentioned in equation 2.8, the contrast in a sample is simply the difference in the SLD between the particle and the solvent. Due to the inherent difference in the scattering lengths of the same atom between SAXS and SANS, the same particle will have different contrast situations in the two different experiments. This is illustrated for a number of molecules used in this project in table 2.3. This will, on its own, provide complementary information on the same sample, which will aid the subsequent data handling. However, especially for particles consisting of multiple components with different scattering lengths, performing contrast variation experiments can provide much stronger information, compared to just a single contrast. It is difficult to dramatically change the contrast in a SAXS experiment as it requires adding or removing a substantial amount of electrons to either the molecule or the solvent. Changing the molecule this much does not make sense, as it would also change the properties of the molecule. Changing the SLD of the solvent can be done by adding electron-rich molecules such as specific salts or sugars, but this will in many cases also affect the particle of interest and is not often used. On the other hand, changing the contrast in SANS experiments is one of the strongest and most used features of this technique. The main reason why this is possible in SANS is the large difference in scattering length between H and D, as seen in table 2.2.

Molecule	SAXS in H <sub>2</sub> O	SANS in D <sub>2</sub> O
	$\Delta\rho \cdot 10^{-14}, cm \cdot \text{\AA}^{-3}$	$\Delta\rho \cdot 10^{-14}, cm \cdot \text{\AA}^{-3}$
POPC <sub>head</sub>	5.10	-4.50
POPC <sub>tail</sub>	-1.65	-6.17
Protein	2.80	-3.30

Table 2.3: Excess scattering lengths of molecules used in this project.

Changing the solvent from H<sub>2</sub>O to D<sub>2</sub>O is relatively easy and will slightly increase the contrast between, e.g., protein and solvent (as seen in figure 2.5), but most importantly this will omit the high incoherent scattering from H, yielding an improved signal-to-noise ratio. As visible from figure 2.5, all the molecules will at some D<sub>2</sub>O content have precisely the same  $\rho$  as the solvent (water), which means that the contrast will be zero, which is also called the contrast match-point (CMP). In practice, this means that a molecule measured at a D<sub>2</sub>O content corresponding to its CMP will be 'invisible' to the neutrons (as  $\Delta\rho$  will be zero in equation 2.13), and in this situation, the molecules are so-called 'matched-out'. It might seem silly to make your particle of interest invisible, but when investigating a multicomponent system, it is often beneficial only to be able to only see one of the components by changing the D<sub>2</sub>O content in the solvent. The exact CMP of a molecule will usually have to be determined by doing a so-called contrast series, which is done by conducting SANS experiments at varying D<sub>2</sub>O contents, but for most molecules, the values in figure 2.5 are approximately correct. From figure 2.5, it can be seen that the CMP is found at approximately 65% D<sub>2</sub>O for DNA, and around 42% D<sub>2</sub>O for a protein. This can be used to investigate protein-DNA complexes, both at 42%, at 65% and 100% D<sub>2</sub>O and thus gain information about the DNA, the protein and the entire complex, respectively [149, 150, 151].

This method has also been used to investigate the ND, exploiting that at 42% D<sub>2</sub>O only the lipids are visible [152].

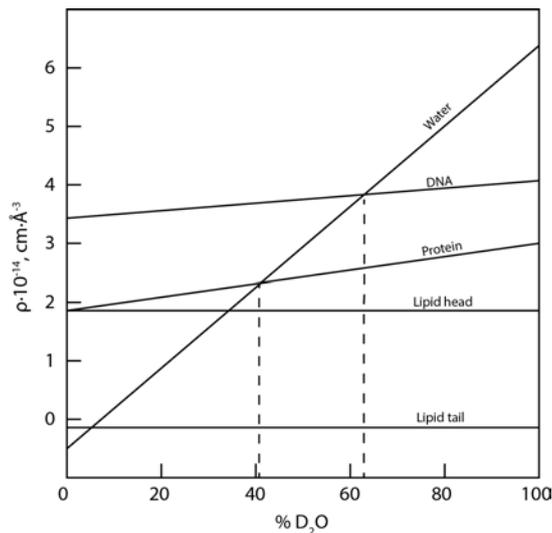


Figure 2.5: SANS scattering length densities for different molecules at different D<sub>2</sub>O contents.

Instead of changing the scattering length of the solvent, another approach is to modify the particle of interest, by changing the particle contrast. For biomolecules, this is usually done by deuteration, either chemically, or by producing them recombinantly in deuterated media. An example of the first is the chemically deuterated detergents used to study MPs without a contribution from the scattering of the detergent at 100% D<sub>2</sub>O [153]. A similar example was seen with a matched-out deuterated ND, where the MSPs and the lipids were produced in *E. Coli* in deuterated media [154, 155]. The neat thing about these methods is that the MP of interest can be produced as for any other experiments because it is only the carrier system that is modified. At present, the major drawback of the method is the cost of producing the material for the matchout deuterated carrier systems. Synthesizing deuterated detergent and lipids is expensive and time-consuming, and producing lipids and purifying from *E. Coli*, as performed in [154], is difficult to do in the high amounts typically needed for SANS experiments. However, establishing a solid synthesis protocol will potentially enable higher yields, and thus easier access to the deuterated material. An example of the use of a match-out deuterated ND is presented in Report I.

### 2.2.5 Experimental setup

Even though there are many differences in practice, the overall principles of the experimental setups at a SAXS and SANS beamline are quite similar. For all experiments in this project, X-rays were produced by a synchrotron, and neutrons by either a spallation source or a reactor, after which they were guided to the actual beamline. The first objective is to obtain a monochromatic beam which, for X-rays, can be done by the use of double crystals. For neutrons, it is usually done by a velocity selector, which can select neutrons with a specific wavelength band based on their velocity. After having obtained a monochromatic beam, the beam is collimated using slits, to have a narrow beam focused on the sample. The scattering intensity is recorded by a detector placed behind the sample. For SAXS experiments, the distance between the sample and the detector is usually fixed, but for all the SANS experiments performed in this project, two

sample-to-detector distances have been used for each sample. This ensures that a sufficiently broad  $q$ -range is obtained. The data from the different settings are then merged into one final dataset after reduction.

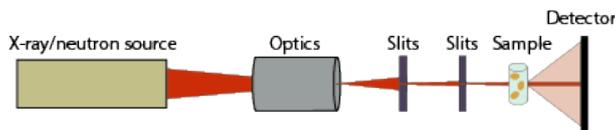


Figure 2.6: Schematic illustration of the setup of a SAS experiment. See main text for detailed description.

In a SAXS setup, a robot is typically used to load the sample, which will be flowed through a capillary where it is irradiated while flowing. The flow ensures that the X-ray dose is spread out on a larger volume, hence minimizing the X-ray induced destruction of the sample, called radiation damage [156]. For SANS the most typical setup is a rack containing about 10 slots for flat Hellma Quartz cuvettes. The cuvettes can have different shapes and thicknesses, depending on the beamline and the sample content, respectively.

For both SAXS and SANS it is now possible to perform a SEC run directly into the beam [157, 158, 159]. This is especially useful to avoid contributions from aggregates in the sample preparation. As scattering patterns are otherwise heavily influenced by the presence of only a minor fraction of aggregates, this method has quickly become a standard method for SAXS experiments and is also gaining ground for SANS. The main drawback of this method is the prolonged measurement time, which for SAXS is determined by the time it takes to run an entire column volume (typically between 35 and 50 min). For SANS the flow usually has to be decreased in the region of interest, giving measurement times of up to 2 hours. In addition, if measuring two settings for SANS, the experiment has to be performed twice, using the double amount of sample and time.

Most SAS data presented in this thesis is obtained using SEC-SAS. In addition, to avoiding aggregates, it is also an advantageous way to separate excess binding partners when probing a multicomponent system. The sample can be prepared with excess binding partners, ensuring elution of the main peak with the complex of interest, and subsequently excess binding partners. In a regular SAXS or SANS measurement, the presence of excess binding partners will contribute to the overall scattering. Examples of this can be found in Paper III and Paper VII.

### 2.2.6 Initial data treatment

As mentioned, SAS data is recorded on a detector in two dimensions, but as the molecules are free to rotate, the data can be radially averaged as illustrated in figure 2.7 to provide the one-dimensional curve of the intensity,  $I$ , as a function of  $q$ .

In all scattering experiments conducted in this project, a background containing the buffer was also measured. For most SANS data, the background is well described by a flat line, but in all SAXS data, it is essential that the buffer is identical to the one of the sample. In figure 2.7, an example of a buffer subtraction from a protein sample is seen (black curve is scaled for visualization). Small deviations in the buffer measurement are usually seen by a lack of

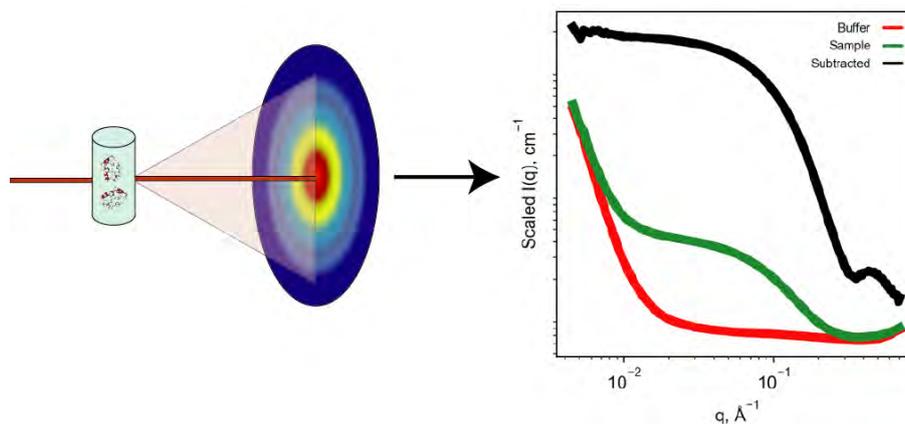


Figure 2.7: A: Illustration of a detector and the radially averaged data. B: Subtraction of a background from a sample measurement. Black: Final data, scaled for visualization.

convergence of sample and background in the high- $q$  region. For SEC-SAS measurements, frames are measured during the entire course of the SEC run. This means that frames corresponding to the particle of interest as well as a buffer region have to be selected afterward. Many frames will usually be averaged, and buffer subtraction will then be performed as usual. Buffer frames from SEC-SAS are typically chosen from the region before the void volume, as it is certain that no particles are eluting at this point.

For further analysis of the scattering data, the data has to be calibrated to an absolute scale in units of  $\text{cm}^{-1}$ . Obtaining data on absolute scale can be performed in many ways, but in this project, the SAXS data was always converted using  $\text{H}_2\text{O}$ . The isotropic scattering signal of  $\text{H}_2\text{O}$  at  $20^\circ\text{C}$  has been determined to be  $1.632 \cdot 10^{-2} \text{cm}^{-1}$  and can be used to calibrate the data [160]. For the SANS data presented in this work, the absolute calibration was performed using a measurement of the incoming beam with the use of mechanical attenuation (at D22, ILL, Grenoble). However, this can also be done by using a secondary standard (such as  $\text{H}_2\text{O}$  or a polymer) as is the routine at FRMII, Munich.

## 2.2.7 Data analysis

With background subtracted and absolute calibrated SAS data at hand, the data analysis can begin. There are a few relatively easy methods that can give an initial estimate of some parameters describing the sample, but more advanced modeling methods are commonly applied. In this section, some of the methods used in this project by me or others will be described to give an overview of the available SAS data modeling possibilities.

### 2.2.7.1 Guiner analysis

The first parameter usually estimated is the radius of gyration,  $R_g$ , which is the root mean square distance from excess scatters to the center of mass in the particle. The  $R_g$  is only meaningful if the particle has a homogeneous SLD. This makes it very useful for protein samples, but less

easy to interpret for, e.g., protein-DNA complexes and NDs. The  $R_g$  is related to scattering intensity at small  $q$ -values, and can be described by the following equation:

$$I(q \sim 0) \simeq I(0)e^{-q^2 R_g^2/3} \quad (2.14)$$

A sample only containing particles giving rise to scattering at higher  $q$ -values than the  $q$ -range covered (meaning that they are sufficiently small), will have a linear relationship between  $\log(I(q))$  and  $q^2$  in the low  $q$ -region, which is usually called the Guinier-region. Obtaining a flat Guinier region is often the first criterion of success for a useful dataset.

If the scattering data is on absolute scale, and the particle of interest has a homogeneous SLD, the molecular weight,  $M_w$  can be estimated using the  $I(0)$  and the concentration,  $c$ , in weight per volume:

$$I(0) = \frac{M_w \Delta \rho_m^2 c}{N_A} \quad (2.15)$$

Where  $N_A$  is Avogadro's number and  $\Delta \rho_m = (\rho_m - \rho_{solvent})\nu$ .  $\rho_m$  is the scattering length of the particle,  $\rho_{solvent}$  is the scattering length of the solvent and  $\nu$  is the specific molecular volume of the particle. For globular proteins in a SAXS contrast  $\Delta \rho_m$  has an average value of  $2 \cdot 10^{10} \text{ cm}^2/\text{g}$ [161]. This can be rewritten into a relatively simple relationship between the molecular weight of the sample and the forward scattering:

$$M_w = \frac{I(0)N_A}{\Delta \rho_m^2 c} \quad (2.16)$$

Where  $I(0)$  may be approximated from the Guinier fit.

### 2.2.7.2 Pair distance distributions

Another commonly used parameter is the maximal distance within the particle,  $D_{max}$ , which is closely related to the so-called pair distance distribution,  $p(r)$ , function proposed by *O. Glatter* [146]. The idea of the  $p(r)$ -function is to represent the SAXS data in real space instead of reciprocal space, to give a more intuitive understanding of the data. This could in principle be done by doing a Fourier transformation of the data, except that this would require a dataset with a  $q$ -range from 0 to  $\infty$ , which is not possible to obtain experimentally. The low- $q$  region would be affected by the beamstop, and the high- $q$  would drown in the background of the instrument at some point, as well as the detectors can not be infinitely large. The principle introduced by *O. Glatter* was to propose a  $p(r)$  function in a finite  $r$ -interval, Fourier transform the  $p(r)$  and fit it to the data. This is why this method is also called the indirect Fourier transformation (IFT). However, two constraints have to be imposed during this transformation; one is that the curve should be smooth, and the second that the data has a finite  $D_{max}$ . Using these assumptions, the scattering intensity,  $I(q)$ , is related to the  $p(r)$  by

$$I(q) = 4\pi \int_0^\infty p(r) \frac{\sin(qr)}{qr} dr \quad (2.17)$$

in which  $r$  is the distance between two scatters.

Figure 2.8 shows four different samples with the buffer-subtracted data in A and the corresponding normalized  $p(r)$ -functions in B. The globular and the flexible protein show the same trend in the  $p(r)$ -functions which roughly descends toward the same  $D_{max}$ . On the other hand, the aggregated protein has a definite upturn in the low- $q$  region of the raw data but also has a long tail in the  $p(r)$  function, which is indicative of aggregation. In the last sample, the nanodisc has a  $p(r)$ -function that is clearly different from the other three with the two peaks. This feature arises from the negative contrast of the lipid tails as seen in table 2.3. The small oscillation in the ND curve around  $130\text{\AA}$  probably arises from imperfections in the data, and the 'true'  $D_{max}$  is most likely around  $130\text{\AA}$ .

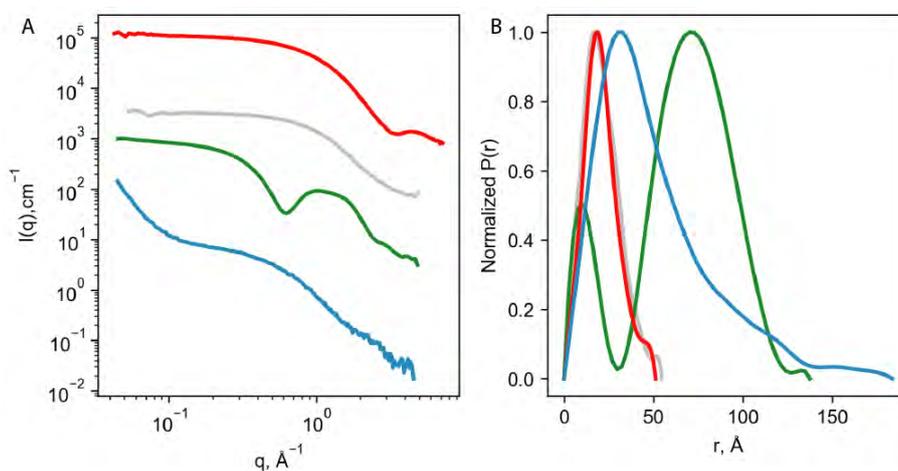


Figure 2.8: A: SAXS data on four different samples. Red: Globular protein. Grey: Globular flexible protein. Green: Nanodisc. Blue: Aggregated protein. Scaled for visualization. B: Normalized  $p(r)$  functions of the same samples.

### 2.2.7.3 Kratky plot

Even though the most common way to represent scattering data is  $I(q)$  vs.  $q$ , many other representations have also gained ground, each emphasizing different features in the data. The Kratky plot is simply  $q^2 \cdot I(q)$  vs  $q$  plotted as seen in figure 2.9B.

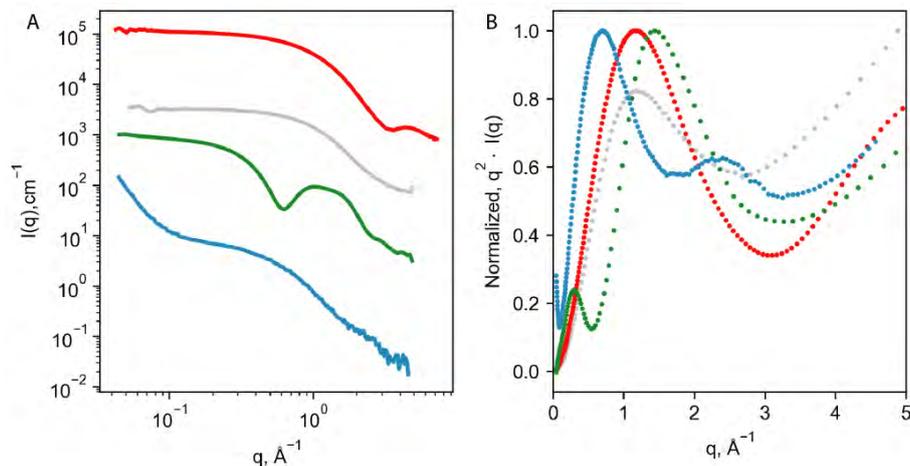


Figure 2.9: A: SAXS data on four different samples. Red: Globular protein. Grey: Globular flexible protein. Green: Nanodisc. Blue: Aggregated protein. Scaled for visualization. B: Normalized Kratky plot of the same samples.

The Kratky plot is primarily used for probing differences in the flexibilities of samples. Globular proteins will have a bell-shaped curve (similar to the red curve in figure 2.9B) while more flexible samples will have a more upwards trend (as seen for the grey curve in figure 2.9B). This is a rapid way to evaluate the flexibility of the system being investigated, but the results should be evaluated with care, and especially multi-contrast samples are challenging to evaluate using this method.

### 2.2.7.4 Modeling of proteins

To obtain more information than given by the parameters described in the above sections, different options are available depending on the scientific case. Generally, when modeling a protein, two scenarios are possible: *i*) a high-resolution 3D structure of the protein (or parts of it) is known (e.g., from crystallography, NMR or Cryo-EM) or *ii*) the 3D structure is entirely unknown. For case *i*), it may seem silly to investigate a protein with a known structure using a low-resolution technique such as SAXS, but in many cases, it is relevant to investigate the solution structure of the protein as this might differ from the crystal structure. Also, investigation of the protein with a ligand bound, or under different buffer conditions are very common experiments. Software such as Crysol, Cryson and SASREF [162, 163, 164] performs rigid-body modeling of known protein structures to SAS data. As the computational possibilities increase with the continuing development of better molecular dynamic (MD) simulation, a combination of MD and rigid body modeling is commonly applied, as the crystal structure seldom represents the solution structure. Another common problem is that the protein contains domains that are not included in the crystal structure. In this case, this part of the model would preferably be

modeled and then fitted to the SAS data. In general, MD and SAS fitting fields are merging more and more, as they complement each other to a high degree [165, 166].

In case *ii*) *ab initio* modeling can be performed, e.g., modeling without any prior knowledge about the structure. In practice it is rare that no prior knowledge is available, as information about the protein sequence, molecular weight (e.g., from mass spectrometry and SDS-PAGE), the overall size (e.g., from SEC), the secondary structure (e.g., from circular dichroism) ect., would be available in most cases, but the 3D structure is often unknown. A typical approach is bead-modeling which can be performed using the DAMMIF software or similar software [167]. The idea is that the protein is represented by small dummy beads of a given size. In the starting conformation, the beads are placed in a sphere where the radius is the  $D_{max}$  of the protein. Following this initial step, one bead at a time is moved in a random orientation, and the theoretical scattering pattern is calculated and fitted to the data. This process is repeated until the fit is satisfactory. A continuum penalty is introduced so that all the beads will always be interconnected [167]. The same idea has been developed for multi-contrast systems in the software called MONSA [147]. The bead modeling is usually performed multiple times to evaluate the consistency of the end-model.

### 2.2.7.5 Geometrical models

Another approach to modeling is using geometrical models build from simple geometrical shapes. As an example the ND can be build from different geometrical components [168, 152] as seen in figure 2.10.

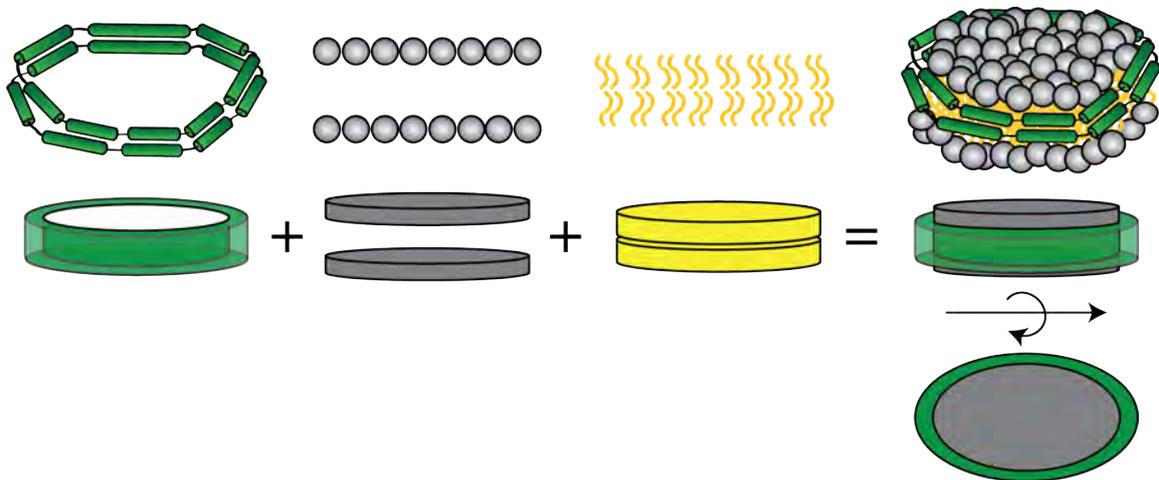


Figure 2.10: Geometrical model of a ND. Simplified model from [168, 152].

The form factor-amplitudes from the different components can then be summed to the total form factor amplitude of the ND:

$$A(\mathbf{q})_{\text{MSP}} + A(\mathbf{q})_{\text{lipidheadgroup}} + A(\mathbf{q})_{\text{lipidtails}} = A(\mathbf{q})_{\text{nanodisc}} \quad (2.18)$$

This results in a model with many parameters to fit, but most of them can be constrained by physically relevant bounds. As a part of this, the molecular composition is well-known along

with the molecular specifics such as volume and scattering length of all the components. In this system, it is valuable to have data from the same sample in different contrast situations (e.g., SAXS and SANS, or SANS at 100% and 42%  $D_2O$ ) as simultaneous fitting to more contrasts helps to restrain the fit. This model and other similar models can be found in the WillItFit software [169]. Usually, parameters such as the number of lipids and the axis ratio of the disc can be determined using this method.

Lastly, the two mentioned methods can be combined. That is the structure of an MP incorporated in an ND is to be determined. This can be done by fitting an empty reference ND, and then using these parameters for the ND with the MP, but fitting the MP either using rigid body or bead modeling [114, 113]. From these strategies, the position of the MP in the disc (e.g., in the rim or the center) and potentially also the orientation of the MP can be determined. In principle, if the MP is of an unknown structure, it is also possible to make a bead-based model of the structure in the ND[114].

## 2.3 Experiments on a surface

In contrast to the SAS experiments, which are performed on particles freely rotating in solution, other techniques require the sample to be deposited on a surface. This comes with advantages and disadvantages, all depending on the technique of choice. In this project, two surface techniques have been used: Neutron reflectometry (NR), for probing structures on surfaces and surface plasmon resonance (SPR), for measuring binding kinetics. The basic concepts of these methods will be discussed in later sections. But before applying these methods, the sample of interest has to be deposited or immobilized on a suitable surface, which is a research field of its own. My personal experience from the last three years has been that preparing a good surface is the key to a successful experiment, whether it is for NR or SPR. Therefore, the deposition technique of choice and the general sample preparation must be performed carefully. In the following section, the pros and cons of different deposition techniques will be discussed.

### 2.3.1 Forming a supported lipid bilayer

In this project, NR has been used to investigate lipid bilayers on a surface, and the prerequisite for obtaining useful data using NR on such a sample is the formation of a homogeneous layer with high surface coverage. Supported lipid bilayers (SLB) are routinely created via vesicle fusion in the proximity of a hydrophilic solid support surface (figure 2.11) [170, 171, 172]. Formation of SLBs via vesicle fusion allows the lipid composition of the bilayer to be varied and the produced bilayers have been characterized with a range of different techniques [173, 174]. In addition to NR and SPR which will be covered in this section, SLBs can also be characterized by other surface sensitive techniques such as Quartz Crystal Microbalance with Dissipation monitoring (QCMD), Atomic Force Microscopy (AFM) and fluorescence microscopy [175, 176, 177].

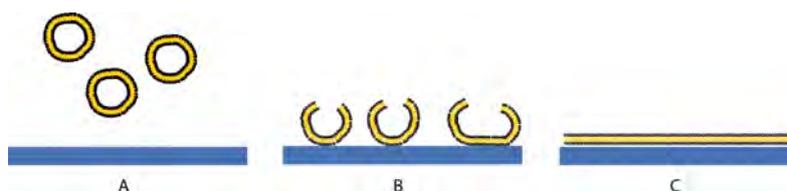


Figure 2.11: Deposition of a lipid bilayer on a solid support using vesicle fusion. A: Vesicles injected. B: Vesicle fuse on the support. C: The SLB is formed.

After the injection of vesicles into a sample cell containing the solid support with a hydrophilic surface (figure 2.11A), the vesicles will adsorb on the support surface and subsequently fuse once a critical concentration of adsorbed vesicles is reached (figure 2.11B). The vesicle fusion can be optimized by temperature and/or ionic strength of the buffer but either way, it is a spontaneous process, leaving a lipid bilayer on top of the support surface (figure 2.11C).

Creating SLBs with different lipid compositions can shed light on the behavior of cellular membranes and the different micro-domains forming in the membrane [178]. Furthermore, these membranes can be used to investigate the binding of soluble proteins and peptides to lipids [179, 180]. These interactions can be probed with regards to the height of the soluble protein over the surface, and the penetration depth in the lipid bilayer.

### 2.3.2 Forming a bilayer containing a membrane protein

Forming an SLB containing MPs is much more challenging than forming a pure lipid bilayer. At present, two methods have mainly been used, but in paper III, a novel method is presented. All three methods will be discussed in this section.

The first method is very similar to the one presented in figure 2.11, except that the vesicles now contain the MP of interest as seen in figure 2.12 [181]. However, the random orientation of the MPs in the vesicle might compromise an effective formation of the supported membrane on the support surface. For MPs with an extramembrane domain (such as TF), it is very likely that this domain will have repulsive interaction with the support. This means that in most cases no stable supported bilayer can be formed, which will result in low surface coverage and MPs in both orientations (see schematics in 2.12C). For an MP without any extramembrane domain, the main problem will be that the MPs end up with a random orientation in the bilayer, which hampers any subsequent binding experiments or similar studies.

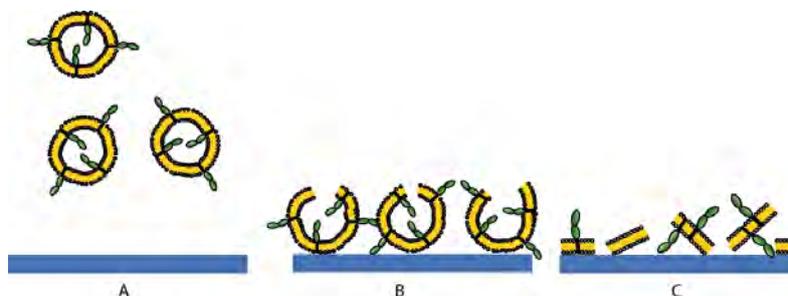


Figure 2.12: Deposition of a lipid bilayer containing TF on a solid support using vesicle fusion. A: Vesicles containing TF injected. B: Vesicle collapse on the support. C: Lipid bilayer containing TF molecules pointing in both directions is formed.

For any binding study, the random orientation means that half of the MPs are unavailable to the binding partner, making data interpretation difficult. For structural studies, the added heterogeneity and low surface coverage of the bilayer decrease the information content of the obtained data. In conclusion, this method has several drawbacks, and other methods have started to emerge.

To overcome the problem of random orientation and to produce supported membranes with high surface coverage, the second approach uses chemical anchoring of the detergent-solubilized MPs to the surface, ensuring a unidirectional orientation of all MPs (figure 2.13A and B) [182]. Afterward, the lipids are added to the system by vesicle fusion, which also ensures detergent removal (figure 2.13C).

This method requires chemical modification of the surface, usually involving gold coating, which limits the application of the method to a few surfaces sensitive techniques. After coating the substrate, the anchoring molecules have to be introduced (usually thiols with an NTA group) before the sample can actually be deposited. One of the critical parameters to optimize in this method is the number of anchoring molecules grafted on the support surface in order to load a suitable amount of MPs in the membrane, as too few will provide to low signal while overcrowding the membrane will provide undesired MP-MP interactions. Using the NTA anchoring requires

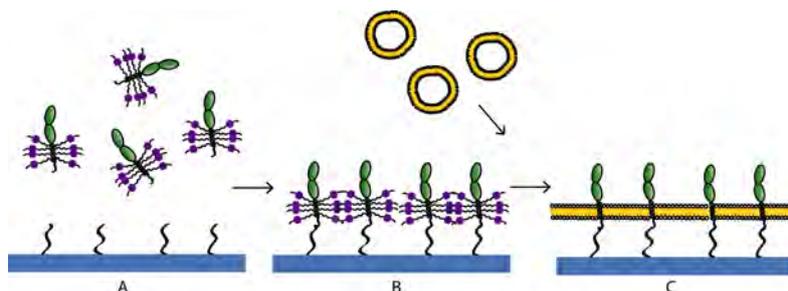


Figure 2.13: Forming a lipid bilayer containing TF on a solid support using chemical anchoring. A: Tissue factor in detergent is injected on a chemically modified surface. B: Tissue factor is anchored to the surface, and vesicles are injected. C: Vesicles collapse and fill the gaps between the TF molecules to form the bilayer.

the MP to have a His-tag in the correct position to reconstitute the MP in the membrane with its native conformation. Besides the advantage of the unidirectional orientation, the ability to choose the length of the surface anchor can be convenient as it determines the height of the bilayer and allows for MPs with large extramembrane domains to be included in the lipid bilayer also with such domains facing the support surface. However, the MP anchoring omits lateral diffusion of the MP, which is an important asset when aiming to replicate a native membrane. In conclusion, the many different preparation steps make this method time consuming, and also very sensitive to the success of each preparation step, setting the need for more convenient deposition methods.

To overcome the problem of modifying the support surface and the protein sequence while concomitantly controlling the orientation of the MP, a third approach was developed as described in Paper III. The method is based on the formation of a so-called peptidedisc, which is very similar to the ND except that the MSP is replaced by the amphipathic peptide, 18A. 18A is well studied and has been shown to be able to form discs with different lipids as well as the incorporation of an MP [142]. However, these discs are much more unstable than the original ND, as the peptides are in a constant equilibrium between the disc and the free peptide in solution. As unwanted a feature as this is for many structural studies in solution, it is also precisely this feature that is utilized for the production of SLB with oriented MPs. The schematics of the formation of the bilayer on the surface using these peptidediscs is demonstrated in figure 2.14.

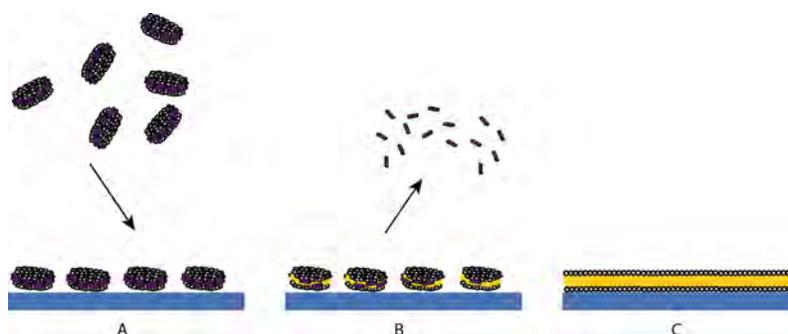


Figure 2.14: Forming a lipid bilayer using peptidediscs. A: Injection of peptidediscs. B: Washing with buffer which removes the peptide. C: Final lipid bilayer without peptide.

After injection of peptidediscs (figure 2.14A) absorption to the surface occurs, and after a certain incubation time, the peptide can simply be washed away with buffer (figure 2.14B) leaving an

intact lipid bilayer without peptides (figure 2.14C). The surface coverage of the SLB formed via the peptidediscs is as good as for the ones formed by vesicles, and no peptide is left in the bilayer after washing with buffer according to the characterization performed by NR (Paper III). The sample preparation of the peptide discs is straightforward and only requires buying 18A (which is commercially available).

However, the real strength of the method becomes apparent when depositing a lipid bilayer with MPs exhibiting one large extramembrane domain. An MP can be incorporated in a peptidedisc, by simply adding the purified MP in the reconstitution step. When injecting the MP-loaded peptidediscs on the surface, the discs will all be oriented with the large extramembrane domain pointing away from the surface (figure 2.15A), as this is the most energetically favorable orientation since the lipid headgroups can interact with the support surface. The rest of the procedure is identical to that of the empty discs, apart from now leaving a surface with a lipid bilayer containing oriented MPs incorporated (figure 2.15C).

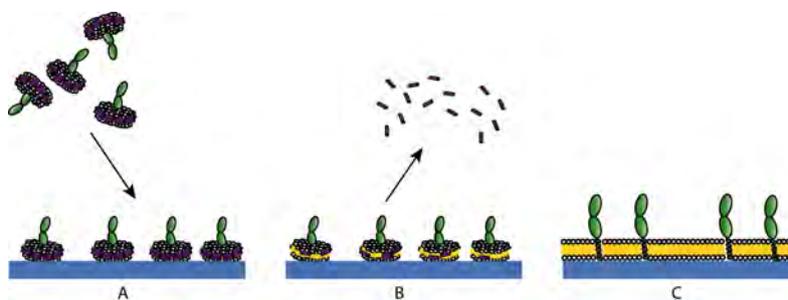


Figure 2.15: Forming a lipid bilayer containing TF using peptidediscs. A: Injection of TF loaded peptidediscs. B: Washing with buffer which removes the peptide. C: Final lipid bilayer containing TF.

The clear drawback of this method is that it requires an MP with an extramembrane domain on one side only, as this is what guides the orientation. Furthermore, it is not possible to control the orientation of an MP, which is only situated inside the membrane. With that said, this method was shown to orient two different MPs (one of them being TF, as shown in figure 2.15) without the need of chemical anchoring. This leaves an ideal sample for NR experiments, but the sample preparation method was also shown to work for SPR, AFM, and QCMD as described in Paper III. This method has been used for all NR experiments conducted during this project.

For most of the SPR experiments performed in this thesis, samples were immobilized using an anti-His antibody, as this is also the most used technique in the SPR field. This would not be preferable in an NR experiment, as the large antibody would contribute significantly to the signal, but in an SPR setup, this is not a problem. This issue will be addressed in greater detail in the following two sections.

## 2.4 Neutron reflectometry

In this project, NR has been used to investigate the structure and orientation of MPs with respect to the lipid bilayer as well as of protein complexes, and thus, a basic introduction will be given in the following section. At present, there are many good reviews describing the theory and

the instrumental setups available for the different beamlines, and these have laid the foundation for the following section [183, 184, 185, 186].

### 2.4.1 Theoretical background

When going through the theory of NR, it is often convenient to treat neutrons as particles which make the Schroedinger equation come in handy:

$$-\frac{\hbar^2}{8\pi^2m}\nabla^2\Psi + V\Psi = E\Psi \quad (2.19)$$

Here  $\hbar$  is Planck's constant,  $m$  is the mass of the neutron,  $V$  is the potential to which the neutron is subject and  $E$  is the energy of the neutron. In the case of NR,  $V$  is the interaction potential describing the effect of the interaction between the scatters in the sample, i.e., the nuclei, and the neutron. A good approximation of  $V$  is the Fermi pseudopotential:

$$V = \frac{\hbar^2}{2\pi m}\rho \quad (2.20)$$

where  $\rho$  is the SLD as we know it from earlier. However, in this case, we need a new way to calculate  $\rho$  which is given by

$$\rho = \sum_j b_j n_j \quad (2.21)$$

where  $b_j$  is the scattering length of nucleus  $j$  and  $n_j$  is the number of nuclei per volume. The scattering lengths are, of course, the same as provided for SANS in table 2.2.

Lets first have a look at the simple setup in which a neutron beam is approaching the surface of an infinite layer as depicted in figure 2.16A.

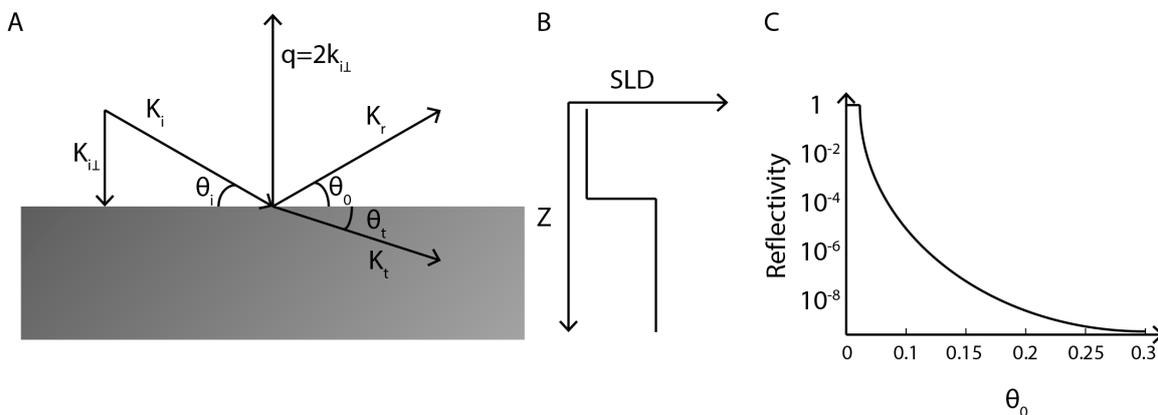


Figure 2.16: A: Schematics of a reflectivity event from an incoming beam. B: Example of a scattering length density profile. C: Example of reflectivity data.

As there is no structure on the surface, the only force present is perpendicular to the surface. This means that the incoming vector  $K_i$  is only affected by the barrier potential, and furthermore

that it is the kinetic energy and the composition of the medium that determines if the neutron is totally reflected. The kinetic energy can be described as

$$E_{i\perp} = \frac{(hK_i \sin \theta_i)^2}{8\pi^2 m} \quad (2.22)$$

where  $\theta_i$  is the angle of  $K_i$  as seen in figure 2.16A. Total reflection will occur in the case where  $E_{i\perp}$  is larger than  $V$ . The critical angle,  $\theta_c$ , which is exactly when  $E_{i\perp} = V$ , can thus be determined. Looking at the following expression for  $q$  at the critical angle:

$$q_c = \frac{4\pi}{\lambda} \left(1 - \cos^2 \theta_c\right)^{1/2} = (16\pi\Delta\rho)^{1/2} \quad (2.23)$$

it becomes clear that the final expression is not dependent on the angle but solely the difference in SLD between the two media conferring the interface.

From a setup as the one seen in figure 2.16A the SLD profile as seen in figure 2.16B would be obtainable, from the raw reflectivity data (figure 2.16C). The SLD profile clearly shows the transition between the two layers, while the reflectivity curve on its own can be difficult to interpret.

In NR two different types of reflection can occur; *i*) nonspecular reflection and *ii*) specular reflection. *i*) nonspecular reflection is defined by the angle of the reflected beam,  $\theta_0$  being different from the angle of the incoming beam,  $\theta_i$ . This phenomenon occurs when the sample is characterized by an inhomogeneous composition in-plane (along the direction parallel to the surface). In principle, from nonspecular reflectometry measurements, structural information about the lateral direction of the substrate can be obtained. This was not used in this project and will not be discussed any further. In contrast *ii*) specular reflection is when  $\theta_0 = \theta_i$ , which provides information on the structure and composition of the sample along the direction perpendicular to the surface and is the technique actually used in this thesis.

Generally speaking, when the neutron beam is not totally reflected because  $E_{i\perp} > V$ , the beam can either be reflected (figure 2.16A,  $K_r$ ) or transmitted (figure 2.16A,  $K_t$ ). In terms of NR, it is very convenient to introduce the momentum transfer vector perpendicular to the surface,  $q$ .  $q$  is defined as follows

$$q = 2k_{i\perp} = \frac{4\pi}{\lambda} \sin \theta_i \quad (2.24)$$

where  $\lambda$  is the wavelength of the neutrons. In the case of NR, the Schroedinger equation (equation 2.19) with the Fermi pseudo potential leads to the same results as treating the neutron as waves and it is possible to apply the laws from optics. Therefore, we can use the Fresnel coefficient:

$$r = \frac{k_{i\perp} - k_{t\perp}}{k_{i\perp} + k_{t\perp}} \quad (2.25)$$

and it becomes clear that by substituting in  $q$  we get

$$R = r^2 = \left[ \frac{q - (q^2 - q_c^2)^{1/2}}{q + (q^2 - q_c^2)^{1/2}} \right]^2 \quad (2.26)$$

where  $q_c$  is the  $q$ -vector of the critical angle,  $\theta_c$ , and  $R$  is the reflectivity. In practice, the  $q$ -value for most measurements is much larger than  $q_c$ . This means that a high  $q$ -values, the following approximation can be used (called the Born approximation [187]):

$$R \approx \frac{16\pi^2}{q^4} \Delta\rho^2 \quad (2.27)$$

We now have an expression for the reflectivity only dependent on  $q$  and  $\Delta\rho$ .

However, for all data obtained in this project, it was necessary to use equation 2.26 to describe the data. This equation only describes a single interface, and in practice, multiple interfaces will typically be present in a biological sample. The equations to fit such data are made by combining equation 2.26 for each interface, while also considering interferences between the neutron wave reflected at each interface. The data analysis will be discussed in more detail in a later section.

### 2.4.2 The instrumental setup

The exact instrumental setup of the reflectometers around the world differs a bit, but in general, they all contain some fundamental components, as shown in figure 2.17. This is a schematic of the SURF instrument at ISIS, where all the data presented in this thesis are collected [188]. As evident from equation 2.24,  $q$  is dependent on  $\lambda$  and  $\theta_i$ . This means that in order to cover the desired  $q$ -range, either of the two parameters can be changed. At some beamlines, a  $q$ -scan is performed by changing  $\theta_i$ , usually by moving the sample, but at other beamlines (such as SURF), the time-of-flight (TOF) method is used. The TOF method relies on the relationship between wavelength and velocity of the neutron, which means that by using a polychromatic beam containing neutrons with many different wavelengths, the desired  $q$ -range can be obtained in a few settings. The time of arrival of the neutron at the detector can then be used to calculate the wavelength of that exact neutron, and hence, it can be correlated to its  $q$ -value.

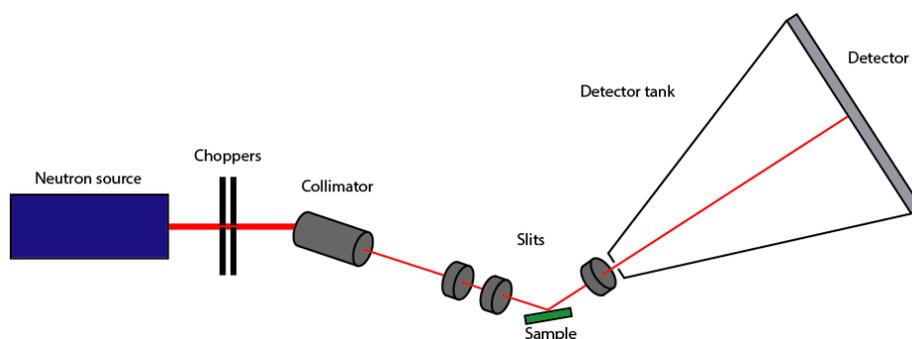


Figure 2.17: Schematic illustration of the SURF instrument, based on [189, 188].

At ISIS, the neutrons are produced by a spallation source, which means that the neutrons arrive in pulses, whereas, e.g., in Grenoble a reactor is used as the neutron source to produce a stable

beam of neutrons. At SURF, the fast neutrons are removed from the beginning of each pulse by the use of two choppers, which ensures that only neutrons with the desired wavelength (between  $0.5\text{\AA}$  and  $13\text{\AA}$ , depending on chopper settings) enters the rest of the instrument. Subsequently, the slow neutrons are also removed using overlap mirrors, which makes the final wavelength range  $0.5\text{\AA}$  to  $7\text{\AA}$ . Afterward, the beam is narrowed down to approximately  $60 \times 10$  mm using a collimator (including a set of mirrors) and a set of slits. Then the sample is irradiated in an upside-down position, meaning that the SLB is facing downwards. The neutrons corresponding to the transmitted and reflected neutrons are first separated by a third slit, which is used to select only the reflected neutrons to reduce the background and finally, the reflected neutrons go into the detector tank and hits the detector. In the case of NR experiments at solid-liquid interfaces, it is usually possible to mount a number of suitably designed solid-liquid cells that can be moved in or out of the incoming beam with a sample changer. Lastly, it is possible to control the buffer flow in the cells using external pumps.

### 2.4.3 Performing an experiment

As with many other methods, obtaining useful data is really about sample preparation. NR is no exception, and the cleaning of the cell and the solid support, as well as the deposition of the sample of interest, are critical parameters in obtaining a good data set. Depending on the desired surface, different cleaning procedures are performed but obtaining a clean surface before the deposition is in any case essential. In this project, we used monocrystalline silicon supports in the dimensions:  $5 \times 8 \times 1.4$  cm. The substrate was cleaned adequately before depositing the samples, in order to obtain a hydrophilic silicon oxide surface, which spontaneously grows on the silicon support surface when exposed to air. The two layers can relatively easy be separated as their SLDs are quite different. When depositing a lipid bilayer (with or without MP) the hydrophilic  $SiO_2$  surface is usually selected. If any impurities or hydrophobic areas are present on the surface, this will most likely affect the deposition of the lipids.

A wide range of commonly used buffers can be used for NR experiments, and in most cases, the same sample should be measured under three or four different  $D_2O$  contents. This procedure enables a robust data analysis by simultaneously fitting the same structural model to the data collected in the different contrast situations. Samples should always be measured in 100% and 0%  $D_2O$  but also in 38%  $D_2O$ , which is called silicon matched water (SMW). Sometimes a fourth contrast is also applied, which is measured at 68% and is called 4MW. The different contrasts are essential to be able to constrain the fitting of the data later on.

### 2.4.4 Data handling

The data analysis is performed by comparing a theoretical model of the sample to the collected data. The data analysis performed as part of this project was performed using the Rascal software and based on the optical matrix method. By this approach the sample is described as a stack of layers characterized by different molecular compositions and each layer is assigned with a thickness,  $t$ , a scattering length,  $\rho$ , a roughness,  $\sigma$ , and a solvent fraction,  $\phi$ . The  $\rho$  is simply the average scattering length of the molecules in the layer, while  $\sigma$  takes into account

that the interface of each layer is not perfectly flat.  $\phi$  is the solvent fraction which both accounts for imperfections in the coverage of the surface but also naturally occurring solvent in the layer (e.g., between lipid headgroups).

After fitting the defined model to the data, the obtained parameters can be used to calculate an SLD profile which contains information on the distribution of the molecules in the direction perpendicular to the substrate.

The lipid bilayer can be separated into four different layers as visualized in figure 2.18A. The silicon oxide is, of course, not of interest but has to be included in the model. In addition, the values for the inner and outer lipid head groups can, in many cases, be constrained to have the same structural parameters in order to reduce the number of free parameters in the fitting model. As all experiments in this project have been conducted using lipids of known composition, it is possible to calculate the SLD for both headgroups and tails, which further aids the fitting.

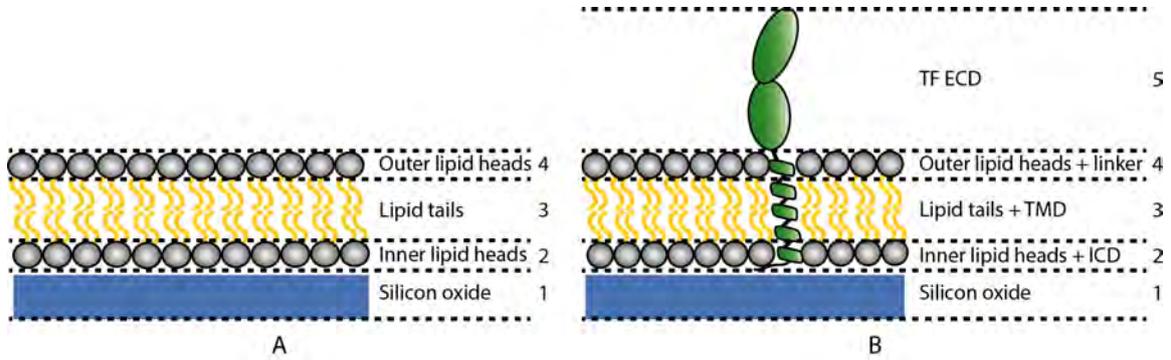


Figure 2.18: A: Separation of a lipid bilayer on a silicon surface into four different layers. B: Tissue factor in a lipid bilayer. Separation into five different layers for analysis of NR data.

In the case where an MP is incorporated in the bilayer (in this case TF), more layers have to be defined and the existing layers have to be redefined. As TF only has a single extramembrane domain (the ECD), a model with five layers was applied in Paper III. In addition to the fifth layer with the ECD of TF, the other parts of TF were included in the other layers, as seen in figure 2.18B. Creating these layers requires some prior knowledge about the structure of the MP, and how the different domains behave in a membrane.

In the case of TF, the lipid tail layer can be described by the following equation:

$$\rho_{Lipidtail+TMD} = \phi_{TMD}\rho_{TMD} + (1 - \phi_{TMD})\rho_{Lipidtail} \quad (2.28)$$

where  $\rho_{TMD}$  is calculated using the sequence of the TMD.  $\rho_{Lipidtail}$  is known from the data previously collected on the pure bilayer. To reduce the number of free parameters, the volume fraction of TF in the acyl chain layer can be described by:

$$\phi_{TMD} = \phi_{ECD} \frac{t_5 V_{TMD}}{t_3 V_{ECD}} \quad (2.29)$$

which means that the number of TMDs and ECDs of TF are the same.  $\phi$  and  $V$  are the volume fraction and the molecular volume, respectively.

A dataset from a measurement on a POPC/POPS membrane on a silicon oxide surface can be seen in figure 2.19A (example from Paper III).

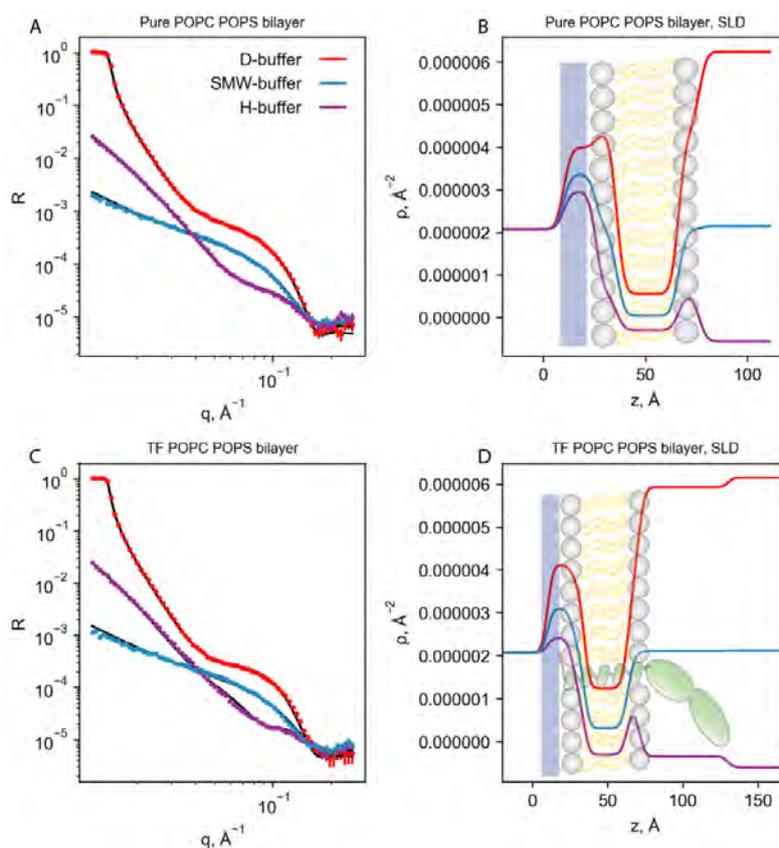


Figure 2.19: A: Reflectivity data on a POPC/POPS bilayer under three different buffer conditions, with fits. Data is scaled for visualization. B: SLD profiles of the data from A. The illustration of the model used for the fitting is overlaid. C: Reflectivity data obtained on a POPC/POPS bilayer with TF embedded under three different buffer conditions, with fits. Data is scaled for visualization. D: SLD profiles of the data from C. The illustration of the model used for the fitting is overlaid.

The model from figure 2.18A was applied to the data, and the obtained SLD profile can be seen in figure 2.19B. As seen, the different layers are visible in the SLD profiles of the different buffer, and it is also evident that the different buffers provide a different contrast (difference in SLD between the sample and the buffer) with the different parts of the membrane. The same measurements were performed on a bilayer with TF embedded (figure 2.19C), and the model presented in figure 2.18B was fitted to the data. In this case, the height of the ECD layer is what's the most interesting, as it indicates the orientation of the domain relative to the membrane. As demonstrated in figure 2.19D, the ECD layer can also be seen in the SLD profile, albeit not as distinct as the other layers, probably due to a higher solvent fraction in this layer.

## 2.4.5 Outlook

Even though NR has been used for decades to investigate surfaces of different substances, biomembranes and especially the interplay between MPs and membranes is still a relatively new subject in this field. This is, as it is for many techniques, in part due to problems with

sample preparation. However, as more and more techniques for depositing membranes with incorporated MPs emerge, this technique will certainly become more relevant for the general protein structure community [190, 191, 192]. In particular, lipid-protein interplay is becoming increasingly important. Many physiological processes are dependent on the presence of a lipid bilayer, and in some cases on specific lipids to maintain the activity (as seen for the coagulation pathway).

Interactions between membranes and soluble proteins (or peptides) are also studied to a greater extent, as the sample preparation is easier, and it provides interesting results about the penetration depth of the protein, and also the height of the protein over the membrane [193, 194]. In addition, this technique provides a very good constraint for supplementary MD simulations, as the height and penetration depth can be used. On the other hand, the MD results can provide an atomistic picture of the events on the membrane [195, 196, 197].

## 2.5 Surface plasmon resonance

SPR was used to determine binding kinetics and binding constants in this project and will be discussed in the following section. Many reviews are available on the topic, and a selection of these has been used for this section [198, 199, 200].

### 2.5.1 General theory

As for NR, SPR is all about a beam hitting a surface; in this case, the beam is of light instead of neutrons. When a beam of light hits a surface, the light is slightly bent, if it passes from a medium with a high density to another medium with a lower density (as illustrated in figure 2.21A, dashed line). By changing the angle of the incoming beam,  $\theta$ , it will at some point reach the critical angle, where total reflection occurs (figure 2.21A, solid line).

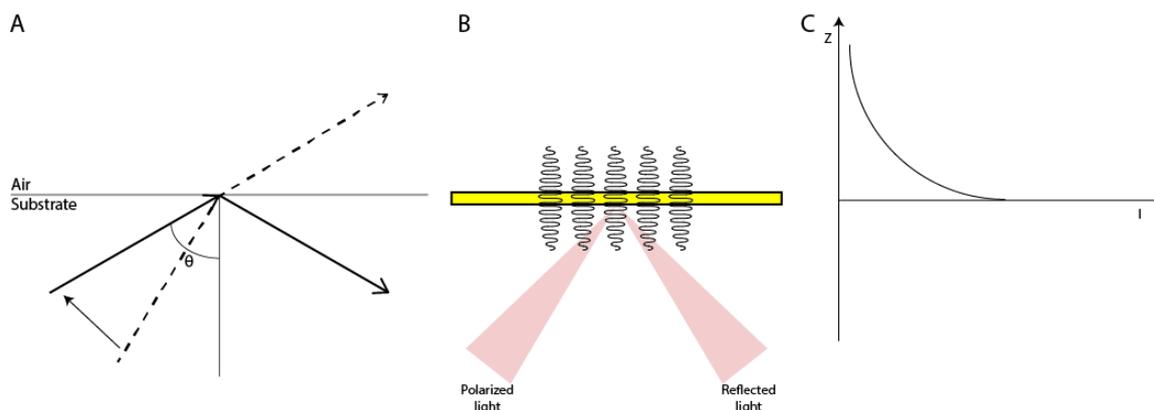


Figure 2.20: A: Illustration of a beam of light hitting a surface. B: Illustration of evanescent waves from a gold film. C: Intensity of the evanescent wave as a function of the distance perpendicular to the surface ( $Z$ ). See main text for detailed explanation.

In an SPR setup, the chip (the surface) is usually covered with a thin gold film, as it possesses some beneficial properties for the experiment, but other metals can also be used. At a specific

electrical field of the incoming photons, it can interact with the free electrons in the gold film. The energy of the photon beam is then absorbed by the gold film and transferred to the free electrons which are converted to plasmons. A plasmon is defined as the electron density wave, and it propagates parallel to the gold film. The oscillations of the plasmons create the so-called evanescent wave which is perpendicular to the surface and has a range of about  $300\text{ nm}$  (see illustration in figure 2.21B). The intensity profile of the evanescent wave in the  $Z$  direction is demonstrated in figure 2.21C.

In summary, this means that when the energy of the light, at the critical angle, is precisely right, the conversion of photons to plasmons will change the intensity of the reflected light. The exact angle (the SPR angle) at which this resonance happens, is dependent on the refractive index of the material on the other side of the surface. This means that if the refractive index on the other side of the gold film changes, the SPR angle will also change, and this change in angle is exactly what that is measured in the experiment.

## 2.5.2 Performing an experiment

To explain the course of an SPR experiment, we will look at an example from Paper IV where TF is reconstituted in an ND and immobilized using an anti-his antibody on the surface. The aim of the experiment is to investigate the binding kinetics between TF and FVIIa. The experiment starts with the immobilization of TF on the antibodies. At this point,  $\theta_1$ , is the angle giving rise to the largest decrease in intensity (figure 2.21 (1)). Upon injection of FVIIa and subsequent binding to TF, the refractive index changes, and hence the angle giving rise to the largest decrease in intensity also changes (now  $\theta_2$ ).

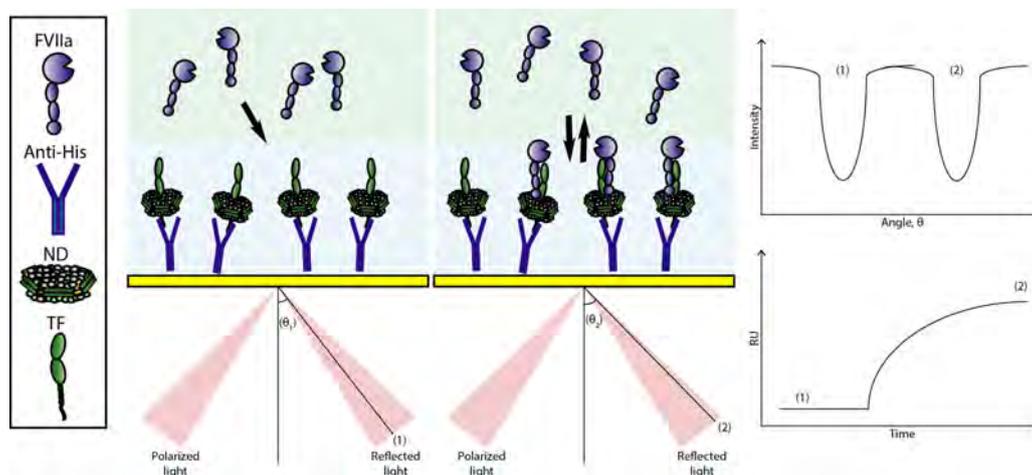


Figure 2.21: Schematic overview of a simplified SPR experiment indicating the change in the SPR angle upon binding of FVIIa. See main text for detailed explanation.

This change in angle can be converted to resonance units ( $RU$ ), adhering to the definition:  $1RU = 10^{-4}$  degree, and thus the development of  $RU$  as a function of time can be followed (figure 2.21, bottom right).

When performing an actual experiment, both the association and the dissociation will be followed as they contain information about the  $K_{on}$  and  $K_{off}$  rates, respectively. Continuing with the

same example as before, the experimental output would look similar to the one presented in figure 2.22.

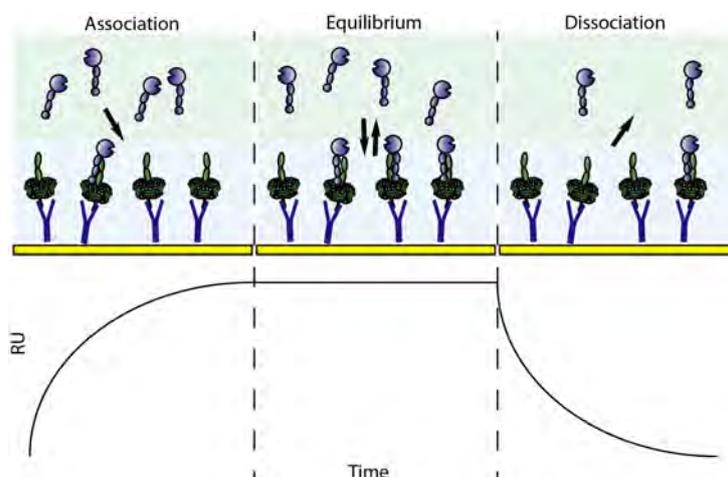


Figure 2.22: Schematic overview of an SPR experiment with an example of the obtained  $RU$  as a function of time in the lower panel.

After FVIIa injection, an increase in  $RU$  is observed as above. After a certain period, the system will reach equilibrium, and the  $RU$  will be flat, indicating that no more FVIIa will bind at this concentration. Lastly, buffer is flushed through the system, and the dissociation of FVIIa is followed until reaching a new flat  $RU$  signal, which should be the same as the initial signal. This cycle is repeated using different FVIIa concentrations ranging around the  $K_D$  of the interaction. In practice, a reference experiment is always conducted simultaneously and is subtracted from the actual experiment to remove contributions from, eg. FVIIa binding to the surface, ND or similar events. Having a good reference experiment is essential for obtaining useful data, and can, in some cases, be difficult.

When fitting the kinetic constants, multiple different concentrations should be measured, ranging around the  $K_D$  of the interaction. This was done for the TF:FVIIa interaction in Paper IV, as seen in figure 2.23.

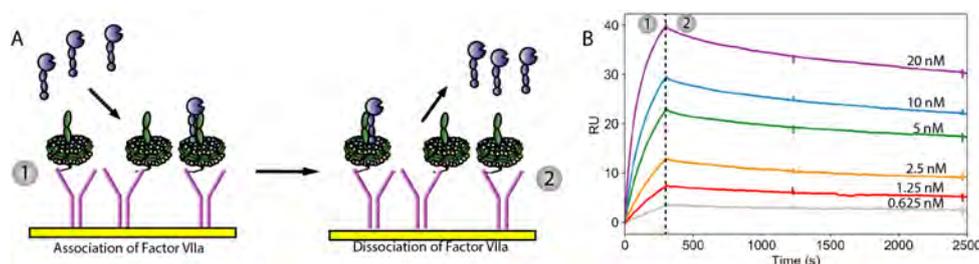


Figure 2.23: A: Experimental setup of FVIIa binding to immobilized TF. B: SPR data obtained on the interaction between TF in a nanodisc and FVIIa (as presented in Paper IV)

Assuming pseudo first order kinetics, the association phase can be described by the following equation:

$$\frac{d[TF : FVIIa]}{dt} = k_{on} \cdot [TF] \cdot [FVIIa] - k_{off} \cdot [TF : FVIIa] \quad (2.30)$$

where  $K_{off}$  is the off rate (in units of  $s^{-1}$ ) and  $K_{on}$  is the on-rate (in units of  $M^{-1}s^{-1}$ ). In other words,  $K_{off}$  is the fraction of the complexes that dissociate per second and  $K_{on}$  is the number of moles that associate per second. These values are usually fitted using this relationship:

$$R = R_{eq} \left( 1 - e^{-(k_{on} \cdot c + k_{off})(t-t_0)} \right) \quad (2.31)$$

where  $R$  is the resonance in  $RU$ ,  $R_{eq}$  is the resonance at equilibrium, and  $c$  is the concentration of FVIIa.

The dissociation phase of the data can be described by the following equation:

$$\frac{dR}{dt} = -k_{off} \cdot R \quad (2.32)$$

and is usually fitted using:

$$R = R_0 \cdot e^{-k_{off}(t-t_0)} \quad (2.33)$$

A global fit to all measured curves is usually applied to obtain common values for  $K_{on}$  and  $K_{off}$ . These values can be used to determine the affinity of the interaction by using the relation to the dissociation constant,  $K_D$ :

$$K_D = \frac{K_{off}}{K_{on}} \quad (2.34)$$

### 2.5.3 Experimental challenges

Multiple experimental challenges can be faced when performing an SPR measurement, which is often related to the immobilization of one of the interaction partners. This immobilization has to facilitate the correct orientation and be of sufficient strength to ensure that the interaction partner is situated at the surface during the entire experiment. In many cases, this is done by attaching an antibody that recognizes the protein of interest, but also covalent methods can be used. As such, in order for this immobilization to be successful, the protein of interest should contain a tag or a site for covalent interaction in order for this immobilization to be successful.

When performing the experiment, one of the things to be aware of is the so-called mass transport limitation (MTL). This phenomenon usually occurs if the binding partner is large and thus diffuses slowly. Upon association, this will result in a depletion zone (figure 2.24, left) where the concentration of the ligand is lower than in the injection, simply because it diffuses slowly. This gives rise to a  $RU$  signal as the solid line, while the binding without MTL is depicted as the dashed line. However, eventually, the same equilibrium  $RU$  as without MTL will be reached. The same issue is present upon the dissociation, now just with a retention zone, in which binding and unbinding events will continue to happen because the binding partner does not diffuse away (see figure 2.24, right).

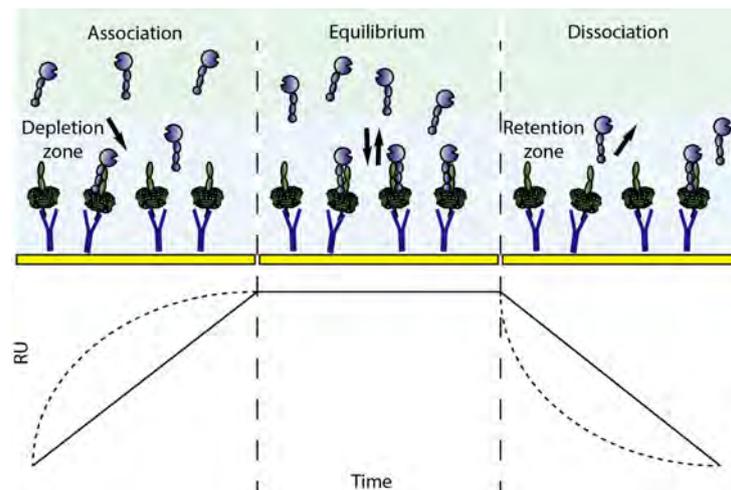


Figure 2.24: Illustration of mass transport limitation. The solid line is the obtained data while the dashed line represents the data if mass transport limitation had not been an issue.

If MTL is observed in a dataset, the first thing to try is to reverse the experiment and immobilize the other binding partner. If this is not possible, or if MTL still occurs, only equilibrium binding information can be obtained using SPR.

## **Part II**

# **Main results and conclusions**



## 3 | Results and discussion

This section is divided into three subsections, each describing the main results obtained during this project. Most of the results are also included in the manuscripts in Part III, however, a few results of a more preliminary nature are also presented here. The first subsection describes the process of method and protocol optimization performed during this project, which indeed laid the experimental foundation for the results in the two last subsections. The first section describes the findings on the orientation of TF in a bilayer, while the last concerns the quaternary TF:FVIIa:FXa:NAPc2 complex on an ND.

### 3.1 Development of carrier systems and purification protocols

Having a suitable carrier system for the membrane protein (MP) of interest, is key to a successful experiment, and hence development of new and improved carrier systems is an essential part of the MP research field. For both structural and functional experiments, three parameters are highly valued, when selecting a carrier system; stability of the MP, sustaining a native-like activity and structure, and lastly the ease of the method. In addition, the SAS methods that have been used extensively throughout this project are also dependent on a relatively narrow size distribution of the carrier system. In conclusion, the demands for an ideal carrier system are very high, and in practice, compromises on some of the parameters have to be made. In this project, most of the experiments have been conducted using NDs, as they provide the necessary lipid environment for the TF complex to be assembled. An optimized ND with a higher temporal stability and easier production was developed during this project, to facilitate subsequent SAS measurements. For NR experiments, the available methods of sample deposition all have drawbacks inconsistent with our aim, and in consequence we also developed a new method producing samples with the desired composition in a straightforward manner.

Lastly, TF had to be produced in the relatively large amounts needed for structural studies. The available protocols all relied on expression of a membrane-anchored TF, which resulted in low yields, not suitable for SAS experiments. We developed a protocol based on expression in inclusion bodies (IB) and subsequent refolding, which yielded more than 10-times as much protein, and amble for structural studies.

### 3.1.1 Circularized nanodiscs

Based on the work on the circularized ND performed by *M. Nasr. et al.*, we designed a new construct called circularized and solubility enhanced MSP1E3D1 (csE3) [109]. *M. Nasr et al.* utilized an evolved version of the enzyme Sortase (eSrt) to form a covalent bond between the two termini of the MSP, resulting in a circularized protein. We adapted this strategy, but also introduced a substantial number of negatively charged residues, which had already been shown to enhance the solubility of other proteins (figure 3.1A for an overview of the construct) [135, 136]. The general idea was that an initial TEV cleavage would provide a N-terminal free Glycine which together with the C-terminal eSrt-site would be linked by the eSrt, leaving a circularized MSP without any His-tags (figure 3.1B).

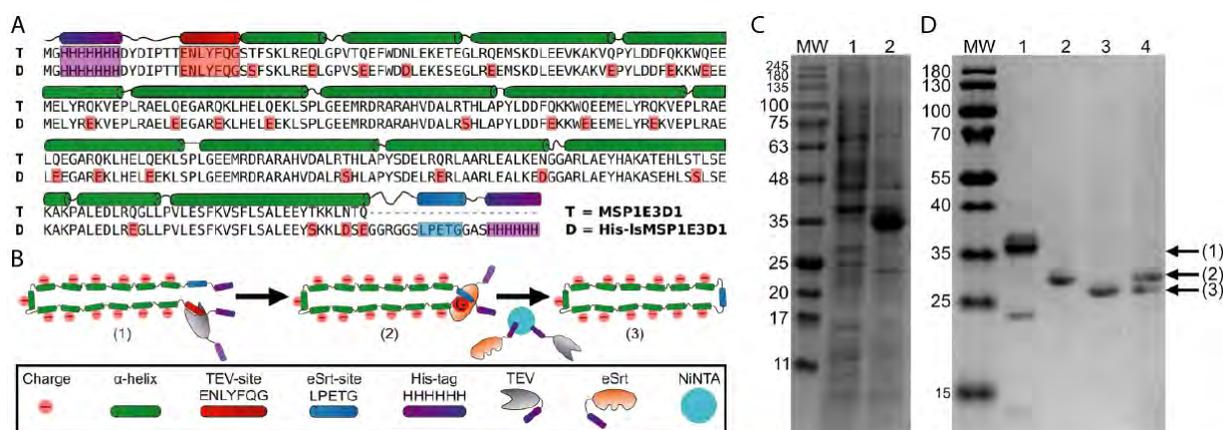


Figure 3.1: Production of csE3. A: Overview of the sequence of csE3 and the respective enzymatic sites. B: Schematic overview of the enzymatic reactions. C: Expression of csE3, lane 1: Before induction, lane 2: After induction with IPTG. D: Visualization of enzymatic reactions. Lane 1: After purification by NiNTA. Lane 2: After TEV cleavage. Lane 3: After circularization by eSrt. Lane 4: Both non-circularized and circularized loaded to visualize the difference in migration.

The expression of the new csE3 construct was very high (figure 3.1C) (similar to many other MSPs), and also the subsequent steps were performed without any issues, resulting in 75 mg of pure csE3 per liter of bacterial culture (figure 3.1D).

A vital optimization step in this protocol was the concentration of the substrate in the sortase reaction, linear soluble E3 (lsE3). We found an almost linear relationship between the fraction of csE3 oligomers and dimers, and the lsE3 concentration used in the reaction (figure 3.2). To maximize the fraction of monomeric csE3, a concentration of 10  $\mu\text{M}$  lsE3 was used in all further reactions, as it provided a yield of  $\sim 95\%$ . This was contrary to the previously reported circularization reactions, which had shown a much lower circularization yield as well as a lower final yield (around 20 mg per liter of culture for similar-sized MSPs) [109, 201, 202]. Another previous study found that the reaction yield could be optimized by adding detergent to the preparation [201], and it is very likely that the high abundance of charges in our new construct introduced the same repulsive properties, circumventing the need for detergent.

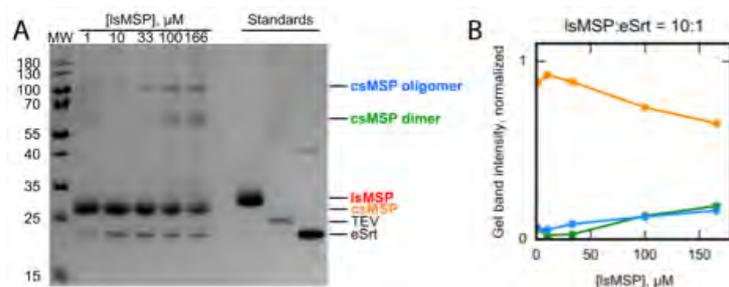


Figure 3.2: Yields of csE3 by different lsE3 concentrations. A: Gel of the samples at different lsE3 concentrations and relevant standards. B: Band intensities from A estimated and plotted as a function of lsE3 concentration.

To investigate the ND formation using the new construct, NDs of csE3 and lsE3 were reconstituted and analyzed by SEC. The two samples looked virtually identical (figure 3.3A), and fractions were selected for SAXS analysis. Both the  $I(q)$  and the  $p(r)$  representations of the data revealed highly identical structures (figure 3.3B) (see details in Paper I).

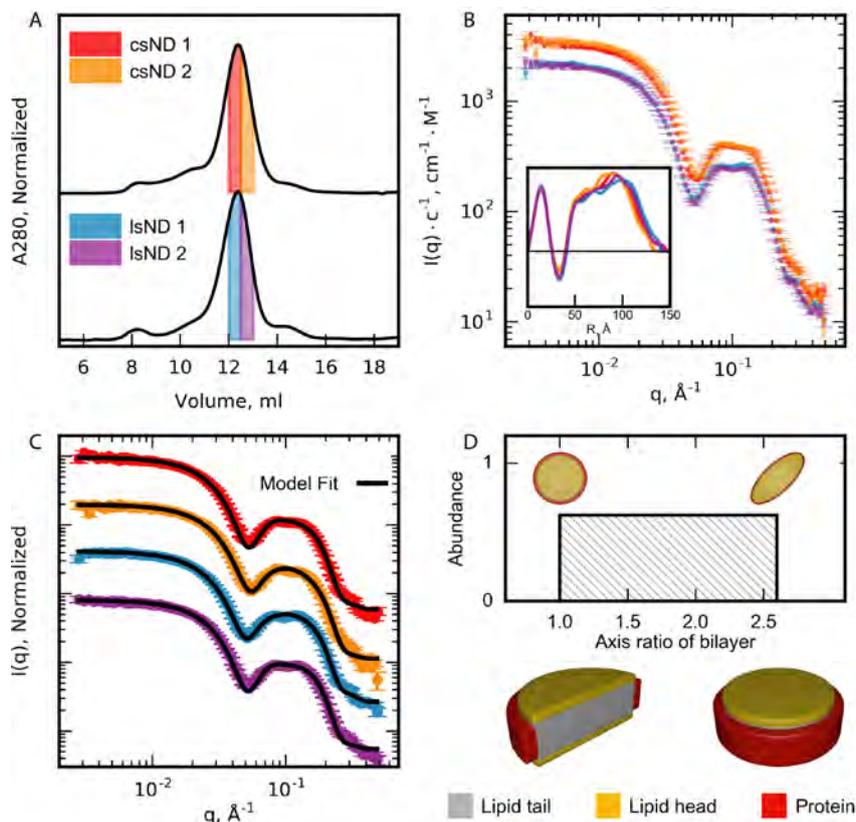


Figure 3.3: Structural investigation of NDs formed by lsE3 and csE3. A: SEC analysis of the reconstituted NDs. Marked fractions were used for SAXS measurements. B: SAXS data of the respective fractions. Insert:  $p(r)$  functions. C: Nanodisc model fitted to the SAXS data (see fitting parameters in Paper I). D: Representation of the axis ratios applied in the ND model to account for polydispersity.

An ND model [168, 152] was fitted to all four data sets, to evaluate the properties of the new MSPs as ND scaffolds. When a range of axis ratios were included in the fit, to account for the slight polydispersity, the model fitted the data (figure 3.3C+D). This range of axis ratios is commonly applied when investigating NDs made by the MSP1E1D1 belt, as they provide a relatively large bilayer which confers some polydispersity.

With an easy production at hand and ensured that the belt still formed NDs, we investigated the temporal stability at different temperatures. Besides csE3 and lsE3, we also investigated the template for the new constructs, MSP1E3D1, with the aim of evaluating the effects of the extra charges. As seen by following the static light scattering signal over time, at three different temperatures, both the charges and the circularization improved the stability (3.4A). This was especially evident at 37 °C where the csE3 signal was completely flat, while the two other curves increased in count rate, suggesting formation of larger particles. The same samples were analyzed by SEC subsequent to SLS measurements (figure 3.4B).

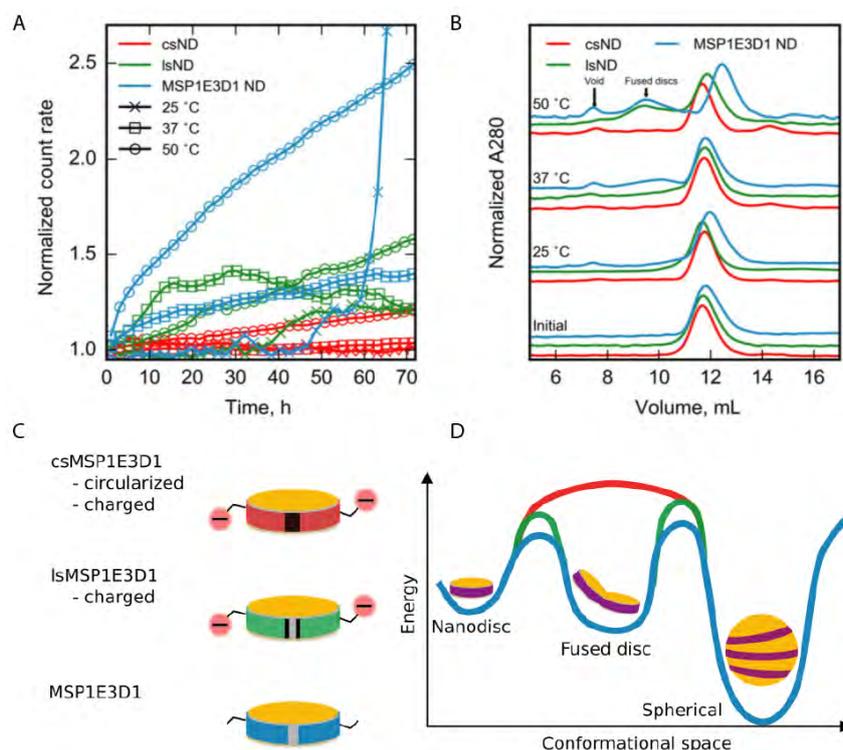


Figure 3.4: Temporal stability of NDs. See sample overview in C. A: SLS measurements as a function of time, performed at three different temperatures. B: SEC analysis of the same samples as in A. C: Overview of the three different samples. D: Proposed energy diagram of the three different samples.

From SEC analysis, it was evident that at 50 °C both the lsE3 and MSP1E3D1 NDs contained material eluting at an intermediate size, but also in the void volume of the column, the latter indicative of aggregation. On the other hand, the csE3 ND only contained material eluting in the void and in the ND peak. This made us propose the energy diagram seen in figure 3.3D. We speculated that circularization restricts csE3 NDs from forming what we call a 'fused disc', which in the SEC profiles eluted at an intermediate size (it should be noted that we have no structural data suggesting a merger of discs, but the size approximately corresponds to two discs). This 'fused-disc' state was visited by lsE3 and MSP1E3D1, but the barrier was higher for lsE3, as suggested by SLS and SEC. Thus we concluded that both the charges and the circularization enhanced the temporal stability of the NDs and that this made them ideal for experiments performed at physiologically relevant temperatures. We also showed that an MP could be incorporated using existing protocols, meaning that neither extra charges nor circularization of the MSP affected reconstitution in this case.

As a next step we wanted to further optimize the production of a circularized MSP, as well as the reconstitution of a circularized ND. Therefore, we designed a new construct which joins csE3 and eSrt (figure 3.5) resulting in a fusion protein (called fE3), with a flexible linker between the two proteins. The idea was to circumvent the need for production of Sortase in parallel, hoping that it could also minimize the production time, as the enzymatic capabilities will be inherent in the protein. In addition, we hoped that this construct would be useful for the class of MPs with hydrophilic domains on both sides of the membrane. These might struggle with the classical incorporation into a cND, as one of the hydrophilic domains would have to pass through the hydrophobic interior of the MSP, and circularization after reconstitution will circumvent this.

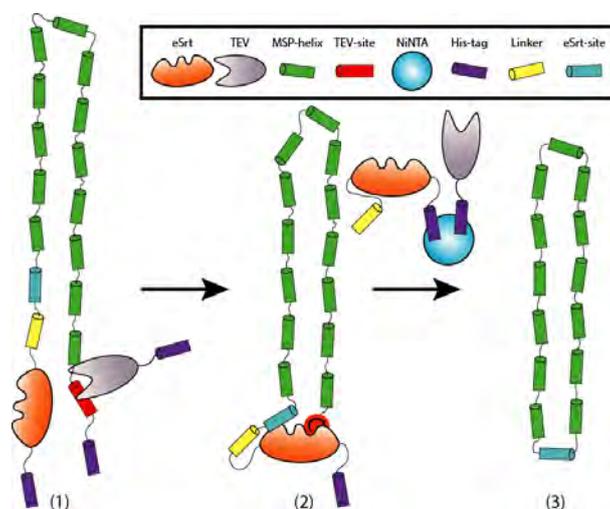


Figure 3.5: Overview of the fusion construct and the proposed reaction mechanism.

The proposed reaction scheme for fE3 was very similar to the one for csE3, except that the eSrt would be present all the time, and did not have to be added. The expression and purification of fE3 proved to be more challenging than first expected, as a fraction of fE3 was degraded during expression at both 20 °C and 37 °C (figure 3.6). We tested four different constructs with different linker lengths and different versions of Sortase (see detailed construct design in Paper II), but all four showed a propensity towards degradation. This could either be due to proteolysis of the linker region, or to Sortase being enzymatically active and attaching the MSP to a free Glycine in the bacterial cells. However, neither shortening the linker nor choosing a less active Sortase solved the problem. The same problem was encountered in studies using a similar construct, but the problem was not further addressed in these [203, 204, 205].

Thus, we continued using expression at 37 °C as it provided the highest amount of fE3. Purification was performed using NiNTA resin followed by an ion-exchange chromatography column providing relatively pure fE3 (details in Paper II). Sortase is reported to be calcium-dependent [206]; thus we tested the circularization of fE3 in the presence and absence of calcium (figure 3.7A).

fE3 was circularized even in the absence of calcium, albeit, at a slower rate than when calcium was present, at which the reaction was almost immediate (figure 3.8A). This was in contrast to the circularization of lsE3 (figure 3.7B), which was much slower in the absence of calcium, while the reaction with calcium was again almost immediate. This was probably due to a proximity

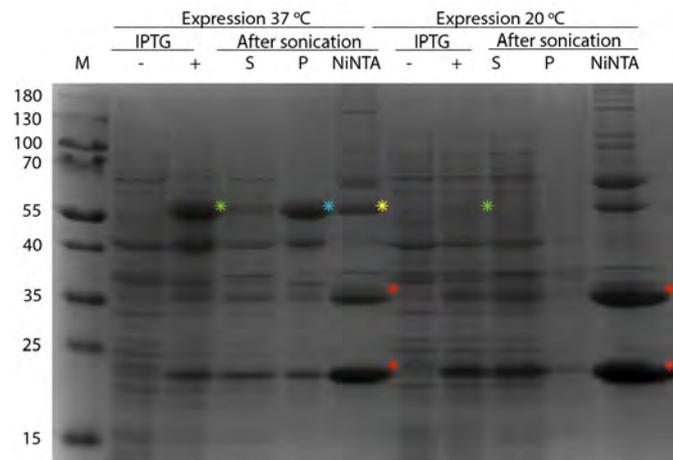


Figure 3.6: Expression of fE3. Green asterix: Soluble fE3. Blue asterix: fE3 in inclusion bodies. Red asterix: Degradation/enzymatic products of fE3.

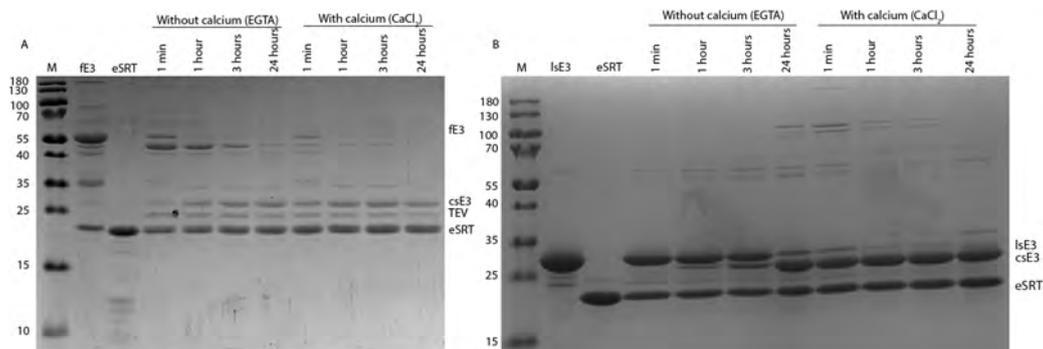


Figure 3.7: A: Circularization of fE3 in the presence or absence of calcium. B: Circularization of lsE3 in the presence or absence of calcium.

effect, as the local concentration of Sortase substrate was much higher for fE3. The csE3 from fE3 was also analyzed by mass spectrometry, which confirmed the overall mass and identified the unique peptide arising as a result of the circularization (Paper II). After confirming the circularization properties, we investigated the possibility to post-circularize a reconstituted ND. The schematics of the process is illustrated in figure 3.8A and is very similar to that in figure 3.5 except that the ND was reconstituted with a non-circularized MSP, and thus lipids are present throughout the enzymatic reactions.

The circularization of a reconstituted ND was indeed possible, using both fE3 (figure 3.8B) and lsE3 (figure 3.8C). In the case of lsE3 the process was yet again, dependent on calcium, but showed similar turnover times. However, in the case of fE3, the reaction was slower compared to the reaction without lipids. This could probably be explained by the sterical hindrance posed by the lipids restricting the reach of the Sortase. To evaluate the size of the NDs formed using this new reconstitution process, four samples were analyzed by SEC (figure 3.8D). The free fE3 eluted in a relatively broad peak around 14 ml, which is commonly seen for MSPs, which tend to migrate as oligomers. Forming an ND with fE3 resulted in a peak in the void, and a peak eluting slightly earlier than the normal csE3 ND. In contrast to the two post-circularized NDs (from either fE3 or lsE3) eluted at the expected (and identical) size. It is currently not clear why

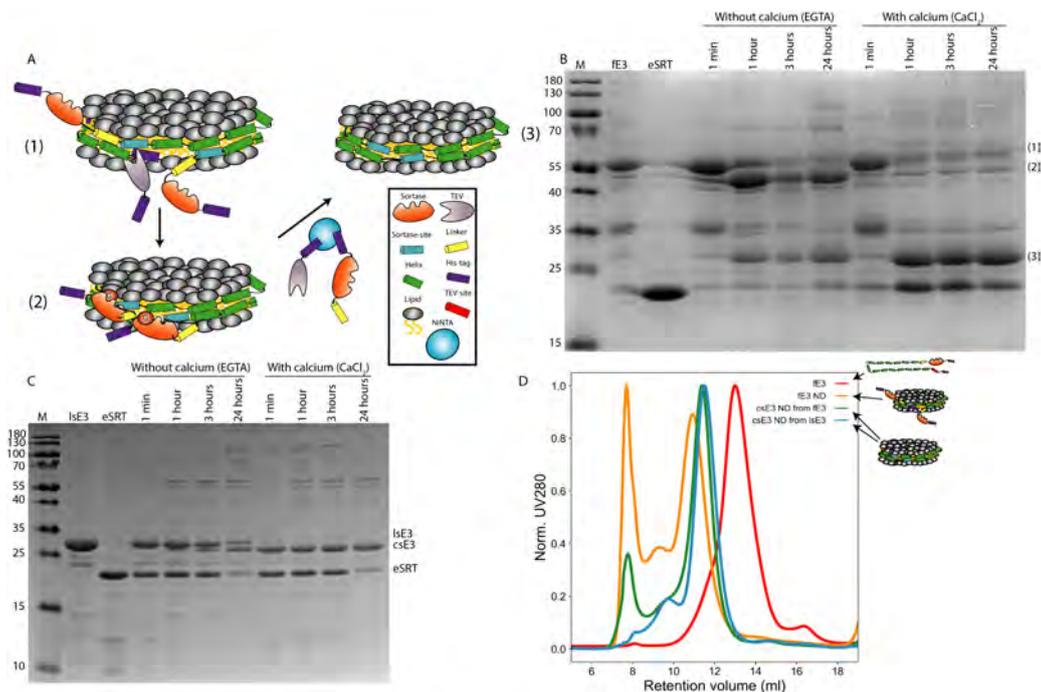


Figure 3.8: A: Schematic overview of the circularization of a reconstituted ND using fE3. B: Circularization of a fE3 ND. C: circularization of a lsE3 ND. D: SEC analysis of NDs and fE3 (see legend).

a fraction of the fE3 NDs ended in the void, and due to time limitations, this was not further explored.

To gain a deeper understanding of the fE3 structural features and the post-circularized NDs, five samples were prepared for SAXS analysis (top panel, figure 3.9).

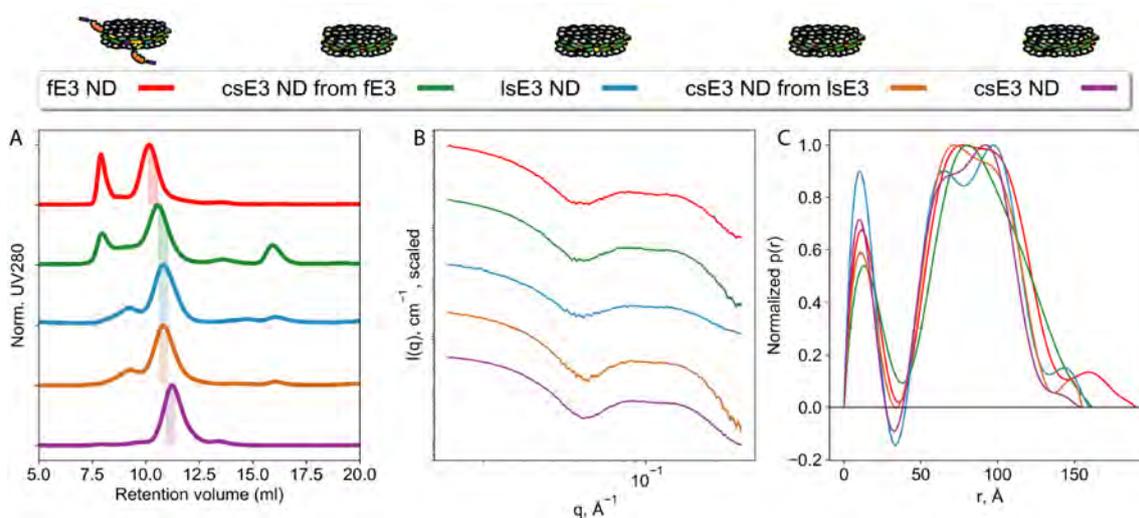


Figure 3.9: SAXS measurements of different NDs. A: SEC analysis of the five samples, with the fraction selected for SAXS marked. B: Buffer subtracted SAXS data on the five samples. Scaled for visualization. C:  $p(r)$  distributions of the five samples. Normalized  $p(r)$ -functions calculated using Bayesapp.org.

As evident from the SEC peaks, the fraction eluting in the void volume was still present in the fE3 ND sample (red), but seemed to be smaller when the post-circularization reaction was allowed (csE3 ND from fE3) (figure 3.9A, green). All five samples contained an oscillation in

the SAXS data (figure 3.9B) around  $0.1\text{\AA}^{-1}$  which is characteristic for an ND, and the  $p(r)$  distributions are also quite similar. The only distinct difference was the extra bump in the  $p(r)$  of the fE3 ND, which could correspond to the two sortases attached to the periphery of the ND. Unfortunately, issues with radiation damage occurred during the beamtime, which was probably the reason why the buffer subtractions looked unreliable, making it challenging to extract conclusive information from the data. However, this data strongly suggested that NDs were formed, and the structure was conserved throughout the post-circularization reaction of a reconstituted ND.

Lastly, we embedded TF in NDs, and performed the circularization after reconstitution. This worked well using both fE3 and lsE3, yielding circularized NDs containing TF with similar activity to probing it in a cND produced using the original protocol (Paper II and IV).

These findings showed the potential for improving the currently available NDs by optimizing the sequence. Furthermore, the available production protocol can be challenged by different purification steps, but also by designing new constructs as the fusion construct. Some optimization is still needed for the fE3 to function optimally, but developing new constructs and protocols easing the use of the ND technology will increase the applicability of the technology and aid the MP community in general.

### 3.1.2 Deuterated nanodiscs

The possibility to match-out (MO) deuterate bio-molecules, has previously been used to make a deuterated ND (dND) that was 'invisible' to neutrons. This was done by using the original MSP1D1 belt and PC-lipids both from bacteria grown in deuterated medium [154, 155]. By performing a contrast variation series, this ND was shown to almost match-out in 100%  $\text{D}_2\text{O}$ . Insertion of an MP into the ND and subsequent SANS data allowed probing different conformational states of the MP, without contributions to the signal from the ND [155]. After developing the csE3 belts, and finding them to confer high stability, we believed they would provide a good scaffold for dNDs, as the  $\text{D}_2\text{O}$  content in the buffer had previously proven to be a problem for less stable NDs. In addition, chemically deuterated POPC lipids had been developed by a collaborator, and the combination of the two resulted in the deuterated, circularized and solubility enhanced NDs (dcsND) presented here. We conducted SANS experiments in buffer with a varying amount of  $\text{D}_2\text{O}$  to estimate the precise MO point (figure 3.10A+B). As seen in figure 3.10A, the intensity decreased when the  $\text{D}_2\text{O}$  content increased, and when visualized as in figure 3.10B, the intercept of the fit provided the match point, which was 101%  $\text{D}_2\text{O}$ . Naturally, it is not possible to perform an experiment in more than 100%  $\text{D}_2\text{O}$ , but the overlay of the raw dcsND data with background measurement in 100%  $\text{D}_2\text{O}$  (figure 3.10C) revealed that the signal from the ND, over background, was neglectable.

In conclusion, we have developed different NDs for different SAS experiments. The high stability of the csE3 NDs was ideal for both SAXS and SANS, and the use of fE3 might make circularized NDs applicable to a broader range of MPs. On the other hand, the dcsND was perfect for SANS measurements in 100%  $\text{D}_2\text{O}$ , as it resulted in data with a signal only from the MP of interest. All the NDs have an approximate diameter of 13 nm, leaving enough lipid surface to facilitate

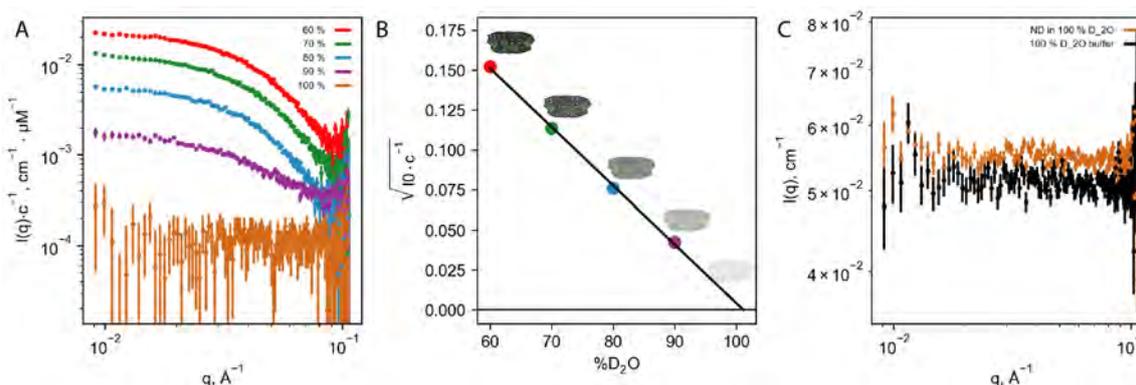


Figure 3.10: Matchout deuterated ND. A: Contrast variation series. SANS measurements at varying  $\text{D}_2\text{O}$  contents. B: Match point determination.  $I(0)$  values determined from A. C: Unsubtracted data for the ND in 100%  $\text{D}_2\text{O}$  and the corresponding background.

the formation of a complex on top. In later sections, all of these features will prove valuable in the investigation of the quaternary complex structure.

### 3.1.3 Deposition of a lipid bilayer

For the NR experiments conducted in this project, the preparation of a lipid bilayer containing TF in a unidirectional manner was essential. Therefore, we set out to develop a deposition method using the so-called peptidediscs. The peptidedisc is very similar to the ND, except that the MSPs are substituted with the amphipathic peptide 18A. Peptidediscs have previously been used to study MPs [142, 207], but as the peptides are in constant equilibrium with the free monomer state, it promotes a rather unstable nature of the molecule, which makes them infeasible for longer measurement times. However, it is this equilibrium that makes them ideal for deposition of a lipid bilayer, and it is this exact feature that was used in this new protocol. The first step was to prepare peptidediscs with the desired lipid content. Subsequently, the peptidediscs were injected into the sample cell and after a short incubation time, buffer was flushed over the surface, which removed the peptides in solution (figure 3.11A). As the peptides are in constant equilibrium, eventually all the peptides would be washed away, leaving a pure lipid bilayer. We followed this process by QCMD (figure 3.11B) and found that after washing, a frequency shift ( $\Delta F$ ) of about 26 Hz was observed, which was exactly what we expected for a POPC bilayer, confirming the utility of the method. Furthermore, NR measurements confirmed that no peptides were left in the membrane (Paper III).

The next step was to use the peptidediscs for incorporation of TF in the membrane. Peptidediscs were produced with TF incorporated, and the same procedure (as above) was followed (figure 3.12A). To confirm the presence of TF in the (POPC) membrane FVIIa was injected, and the decrease in  $\Delta F$  proved the presence of accessible TF (figure 3.12B). As evident from the QCMD data, a small fraction of FVIIa was washed away when flushed with buffer, but this was probably FVIIa bound to the membrane and not to TF. As a control experiment, FVIIa was injected onto a pure POPC membrane, and no binding was observed (Paper III).

To confirm that all TF molecules were in the same direction, NR measurements were conducted,

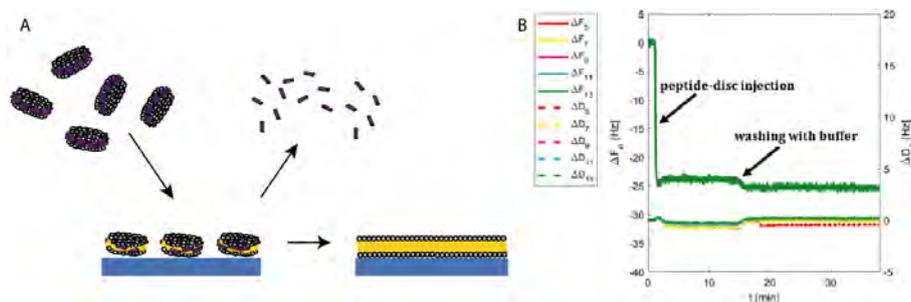


Figure 3.11: Depositing a lipid bilayer using peptidediscs. A: Schematic overview of the procedure. After injection of peptidediscs and a brief incubation time, buffer was flushed over the surface, removing the peptides and leaving a pure lipid membrane. B: QCMD data on the procedure presented in A.

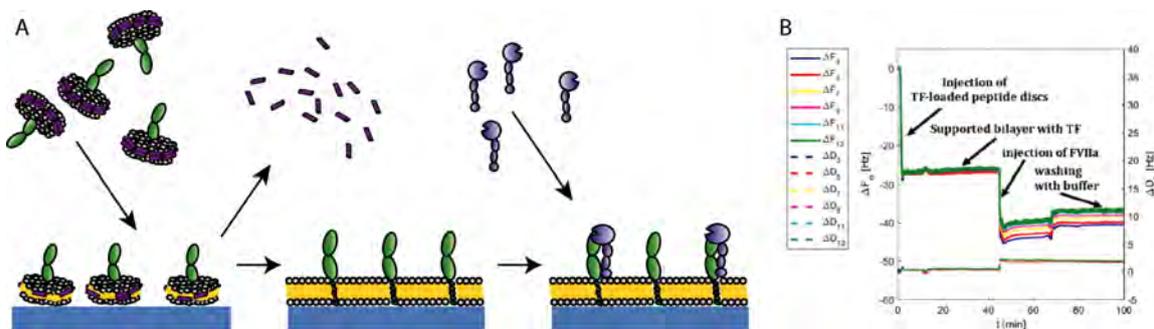


Figure 3.12: Depositing a lipid bilayer containing TF and subsequent binding of FVIIa. A: After injection of peptidediscs with TF incorporated the surface is flushed with buffer which removes the peptide and leaves a bilayer with TF incorporated in a unidirectional manner. Afterwards, FVIIa is injected and can bind to TF. B: QCMD data on the events presented in A.

demonstrating that the TF ECD was only present on the side of the membrane pointing away from the substrate (see Paper III).

A drawback of this new method, is the difficulty in estimating the concentration of the MP in the sample, as the absorption at 280 nm was a combination of the contributions from the MP and an unknown number of peptides. Therefore, we explored the helical structure of 18A (figure 3.13A) to see if we could mutate the Tryptophane and the Tyrosine, which contributes to the UV280 absorption. As evident from the peptide wheel, the Tyrosine is located on the hydrophobic side of the peptide, while the Tryptophane is in between two charged Lysines but close to the border of the hydrophobic part. Therefore, we decided to substitute the Tyrosine to a hydrophobic Leucine, while the Tryptophane was substituted to a Glutamine (figure 3.13B). We believed that these mutations would preserve the amphipathic properties of the peptide while removing the UV280 absorption, and hence called the mutant dark-18A (d18A).

The first test we did was to determine if d18A formed lipid-particles of a similar size to 18A. To be able to monitor the d18A peptidediscs, UV absorption at 214 nm was also monitored, as contributions from peptide bonds would be visible here. An 18A peptidedisc elutes around 14 ml on an SEC (figure 3.13C), and as seen in figure 3.13D, the d18A peptidedisc eluted at the same volume, but without any UV280 signal. The d18A chromatogram also showed a minor fraction eluting around 19 ml, which was probably free peptide. This could either be due to a slightly skewed lipid:peptide ratio (which is known to provide some free peptide) or to a slight

change in the properties of d18A compared to 18A. Therefore, a thorough optimization of the lipid:peptide ratio is needed, as it is essential for this kind of sample.

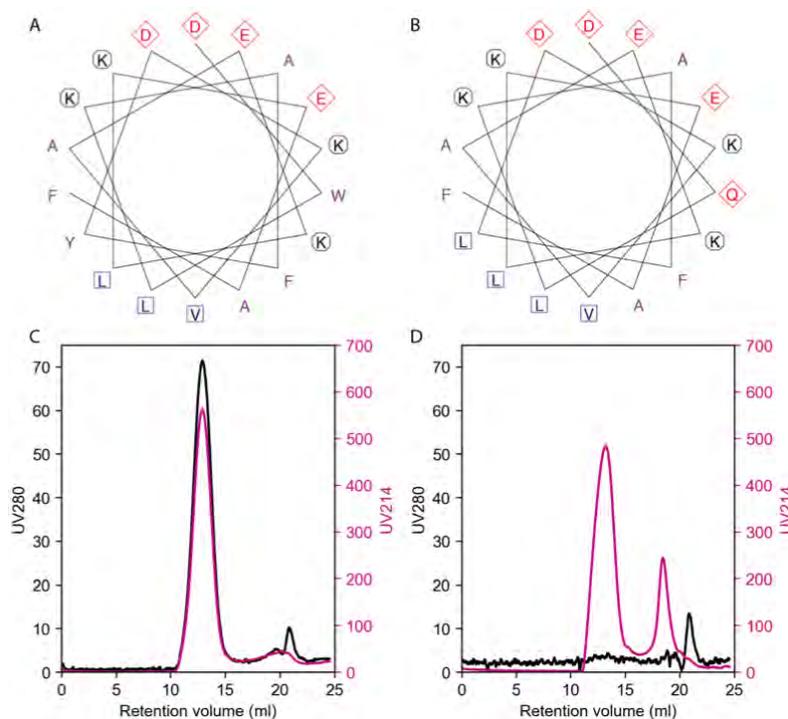


Figure 3.13: d18A design. A: Peptide wheel of 18A. B: Peptide wheel of d18A. Peptide wheels created using the EMBOSS server. C: SEC chromatogram monitoring UV absorption at both 280 nm and 214 nm for an 18A peptide disc. D: same as C, but with a d18A peptidedisc.

The incorporation of TF in 18A (figure 3.14A) and d18A (figure 3.14B) was monitored by SEC. As evident from the 18A chromatogram, it was impossible to see if TF was incorporated, but the UV280 peak in the d18A confirmed the presence of TF. The fractions from the chromatograms were analyzed by SDS-PAGE showing a TF band (figure 3.14C). The gel also demonstrated a band corresponding to the MP CorA, which was also incorporated with success (SEC chromatograms not shown). The fraction of UV280 in the void of d18A was probably due to pre-aggregated TF, as the same feature was present in the sample with 18A.

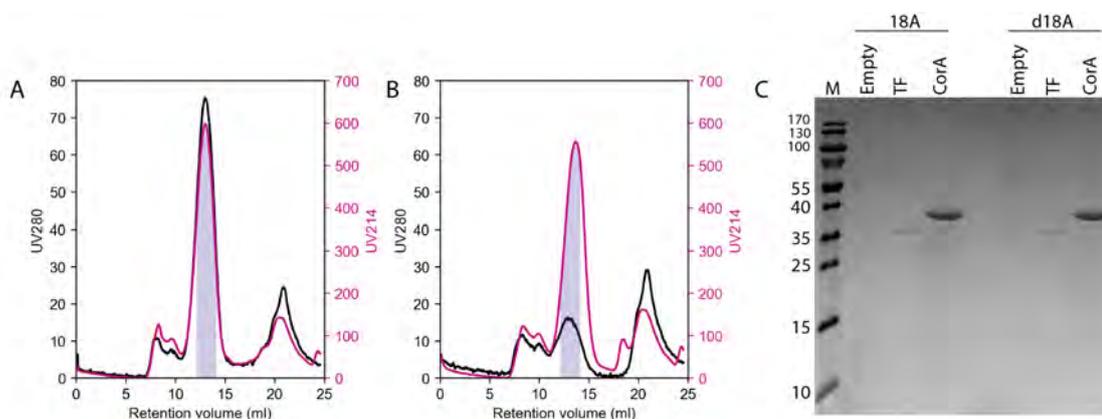


Figure 3.14: A: 18A peptidedisc with TF incorporated. B: d18A peptidedisc with TF incorporated. C: Gel of fractions from empty, TF (from A and B) and CorA incorporated peptidediscs made using either 18A or d18A.

To evaluate d18As properties as a deposition tool, the same QCMD tests as demonstrated for 18A were performed with the same outcome (data not shown). This means that a peptide with the same MP carrying properties is now available, and that this enables monitoring only the MP signal on the SEC, and also estimation of the final sample concentration using the UV280 absorption. This is very useful as knowing the concentration is important when performing most experiments, or when optimizing experimental conditions.

In conclusion, this novel technique enables a relatively straight forward sample deposition method, allowing for experiments such as NR and SPR to be performed on unidirectional MPs in a lipid bilayer.

### 3.1.4 Refolding of tissue factor

As the overall goal of this project was to obtain structural information on TF and its interaction partners, solid production of TF had to be established. Previously reported recombinant TF production relied on expression of a membrane-anchored TF [33], but as this provided relatively low yields, we pursued a strategy involving expression in inclusion bodies (IBs). This led to the overall production scheme illustrated in figure 3.15A.

The expression of TF was clearly visible in the cellular lysate (figure 3.15B+C, lane 3) and after cell lysis, TF was present in IBs, as the lysis supernatant almost contained no TF. After washing the IBs, they were dissolved in a 6 M GuHCl buffer, which unfolded everything, making the preparation ready for refolding. The refolding was performed by dripping the preparation into a buffer reservoir containing different volumes of refolding buffer, and extracting samples over a timespan of 24 hours. After NiNTA purification, the refolding yields were estimated by gel quantification, and the solubility of the sample by SEC (see Paper IV for details). The refolding yields, as a function of dilution factor and time, can be found in figure 3.16D. As figure 3.16D clearly demonstrates, the yield is dependent on time and dilution factor, with an optimal yield of 80% obtained using a dilution factor of 60 and an incubation time of 24 hours. This provided a total of 5 mg TF per liter of bacterial culture, which could then be used in future experiments.

To provide the required lipid bilayer for TF, reconstitution into the developed cse3 NDs was optimized with different lipid compositions (Paper IV). In order to probe the activity of the refolded TF, an activity assay based on FX activation by the TF:FVIIa complex was used (see the experimental setup in figure 3.16A). This assay relies on the substantial difference between the activity of free FVIIa and TF bound FVIIa, as described earlier. The activity of free FVIIa is neglectable, and thus only signal from TF bound FVIIa was measured, which allowed for an indirect determination of the  $K_d$  of the interaction.

The measured signal arose from FXa cleavage of a peptide substrate. Subsequent to cleavage, the substrate absorbed light at 405 nm, and the slope of the absorbance was evaluated as a function of the added FVIIa concentration (figure 3.16B). The  $K_d$ s of the interactions were determined to be  $48 \pm 8 \mu\text{M}$  and  $10 \pm 1 \mu\text{M}$  for the sample containing 15% POPS and 30% POPS, respectively, which is in agreement with values reported previously from similar experiments [19]. Interestingly, the SPR measurements conducted (see schematics in figure 3.16C), did not show the same dependency on PS presence. For these measurements, the  $K_d$ s were determined

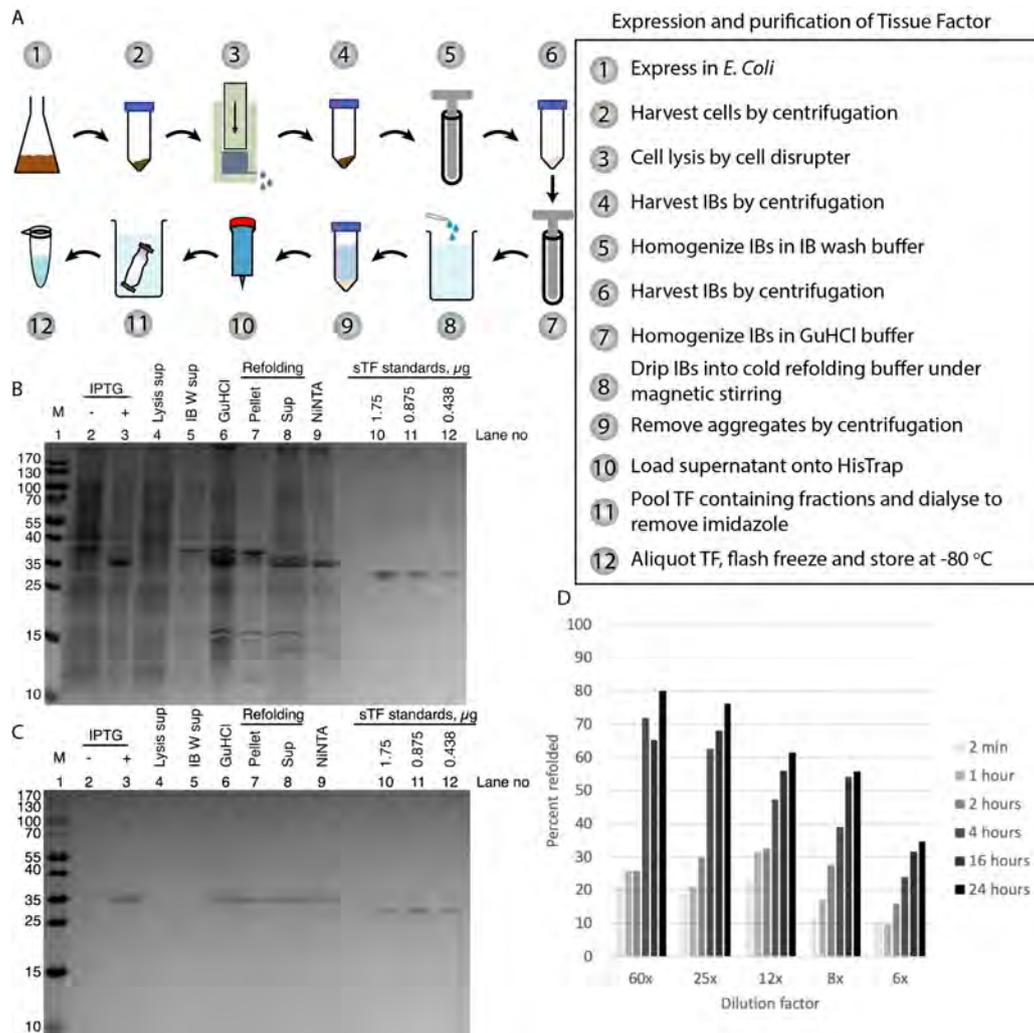


Figure 3.15: A: Schematic overview of the production of TF. B: SDS-PAGE gel, Coomassie-stained, of the different purification steps and relevant standards. C: SDS-PAGE gel, western blot with primary TF antibody, of the different purification steps and relevant standards. D: Refolding yields as a function of dilution factor and time.

to 126  $pM$ , 127  $pM$  and 125  $pM$  for 0% POPS, 15% POPS and 30% POPS, respectively. This was also in agreement with similar experiments [18], but highlights that affinity is not affected by the lipid composition, as discussed in the introduction. The reason for this dependency in the activity assay is currently unknown, but it could be due to the interaction between the Gla-domain of FX and PS [19].

### 3.2 Tissue factor and Factor VIIa

Prior to the work presented here, the only available structural data on TF in a membrane was two FRET studies indicating a change in the membranal orientation of FVIIa when it interacted with TF [67, 68]. This change in orientation was explained by lipid interactions of the TF ECD, which was supported by mutational studies and MD studies [71]. Therefore, we set out to investigate TF lipid interactions and the orientation of TF and FVIIa on a membrane, using the

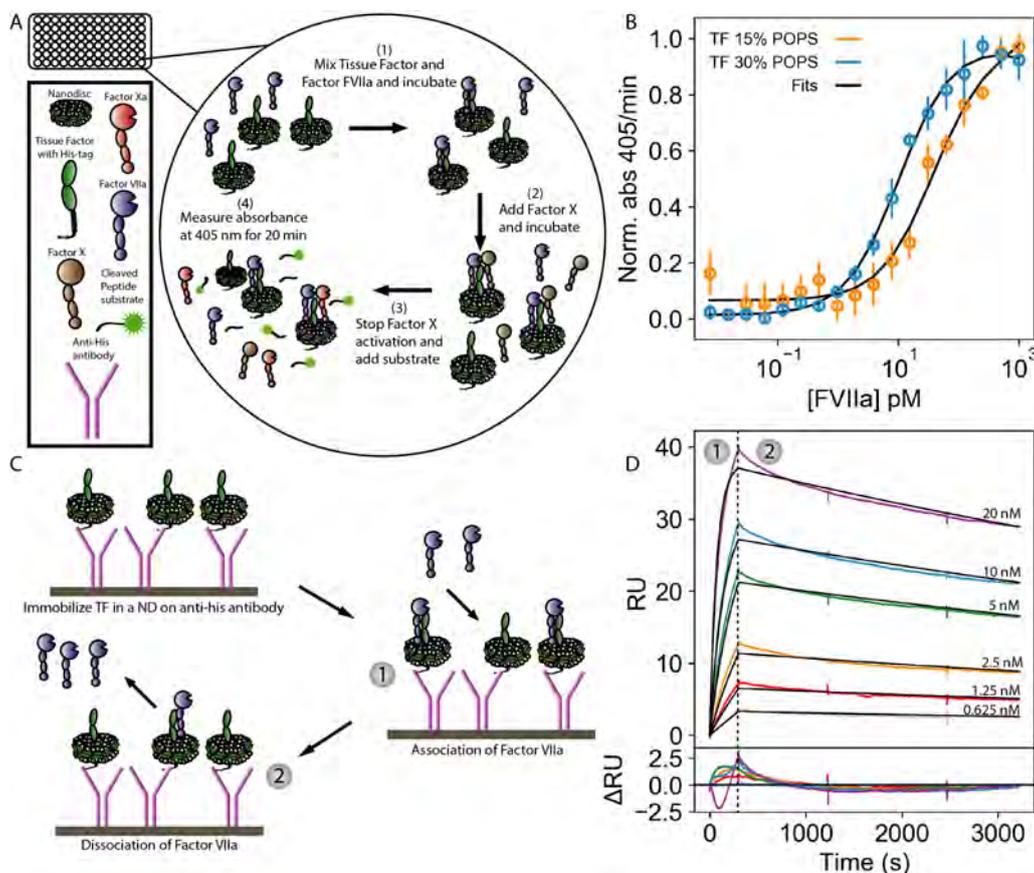


Figure 3.16: A: Experimental setup of the activity assay. B: Data obtained from the activity assay. The slope of the absorbance at 405 nm as a function of the FVIIa concentration added. The data were fitted by a quadratic binding curve to obtain the  $K_d$ . C: Setup of the SPR experiments. D: SPR data fitted using a non-linear regression model assuming pseudo-first-order reaction conditions.

tools developed in the previous section.

### 3.2.1 The orientation of Tissue factor

The first aim was to investigate the lipid binding properties of sTF and if possible, to map the lipid binding residues for a comparison with the MD study. In the first experiment conducted, we used QCMD and a pure lipid membrane containing 30% PS and  $10\mu M$  sTF, to see if the mass on the surface increased (figure 3.17A). However, as seen in the resulting data (figure 3.17B) no absorption on the bilayer was observed, as the  $\Delta F$  was completely flat throughout the experiment.

To further investigate the apparent lack of interaction, a labeled version of sTF ( $^2D^{15}N$ ) was mixed with NDs containing either pure POPC or a mix of 4:1 POPC:POPS (figure 3.17C). The NDs used in this experiment were the circularized and solubility enhanced MSP1 $\Delta$ H5D1 (csdH5), which were identical to csE3, except that the template was MSP1 $\Delta$ H5D1, and hence the final ND was smaller, with an approximate diameter of 9 nm [208]. The reason for choosing the smaller ND, was that a potential sTF:ND complex could tumble fast enough to provide a useful NMR signal, while the larger NDs would definitely tumble too slow. The labeled sTF provided

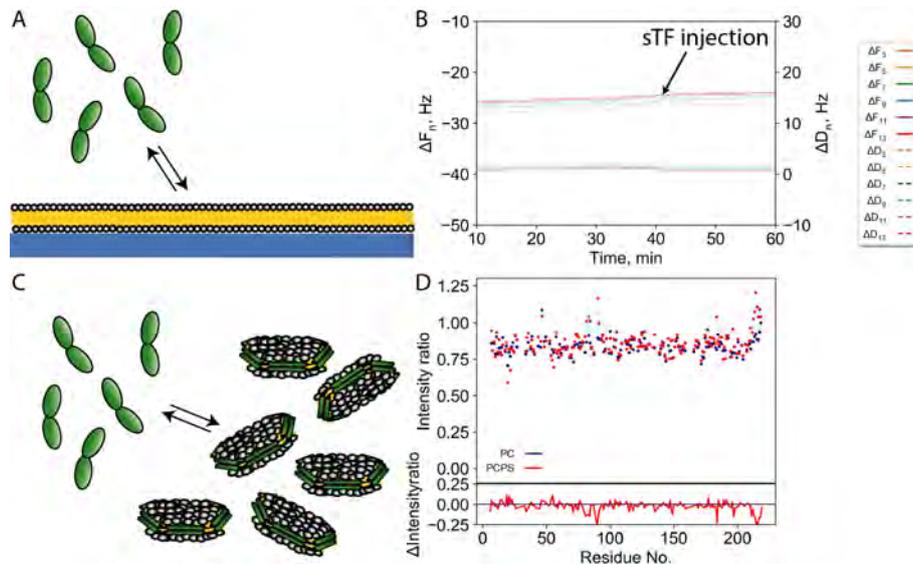


Figure 3.17: Lipid interactions of sTF. A: Setup of QCMD experiment. B: QCMD data. C: Setup of NMR experiment. D: NMR data. The intensity ratio between free sTF and sTF+ND. Experiments performed with pure POPC NDs and POPC/POPS NDs. Lower panel is the difference in ratios between POPC and POPC:POPS.

$^1H^{15}N$ -HSQC with nicely separated peaks (Paper V), and as all assignments were available for sTF, only the  $^1H^{15}N$ -HSQC spectra had to be collected at different sTF:ND ratios, to be able to see effects on a single residue level. This was obtained for the pure POPC ND as well as and the mixed ND, but there were no observable chemical shift changes (details in Paper V). Consequently, we investigated the change in intensity compared to free sTF, but observed no difference between the two different NDs (figure 3.17D), once again suggesting no lipid binding. As evident from the figure, the signal decreased for both NDs compared to the pure sTF sample, but this change was assigned to a change in viscosity, as a substantial amount of material was added (the NDs), and hence the tumbling of sTF was potentially slower. At least this change was very similar for the two different NDs, indicating the same effect, and no apparent effect of PS. In conclusion, we observed no interactions between sTF and lipids. However, this does not necessarily mean that no interactions takes place when FL TF is incorporated in a membrane, as the local concentration of the ECD is much higher as it is anchored in the proximity of the lipids.

The next step was, therefore, to investigate the structure of TF incorporated into a lipid membrane and to determine the orientation of TF. We used the deposition method, presented earlier, (and in Paper III) to make a PCPS bilayer containing TF. The obtained NR data suggested a height of  $61 \pm 3 \text{ \AA}$  of the TF ECD layer, which conferred a slightly tilted orientation of TF (figure 3.18). The parameters of the lipid bilayer were virtually identical to those obtained on the pure lipid membrane (Paper III).

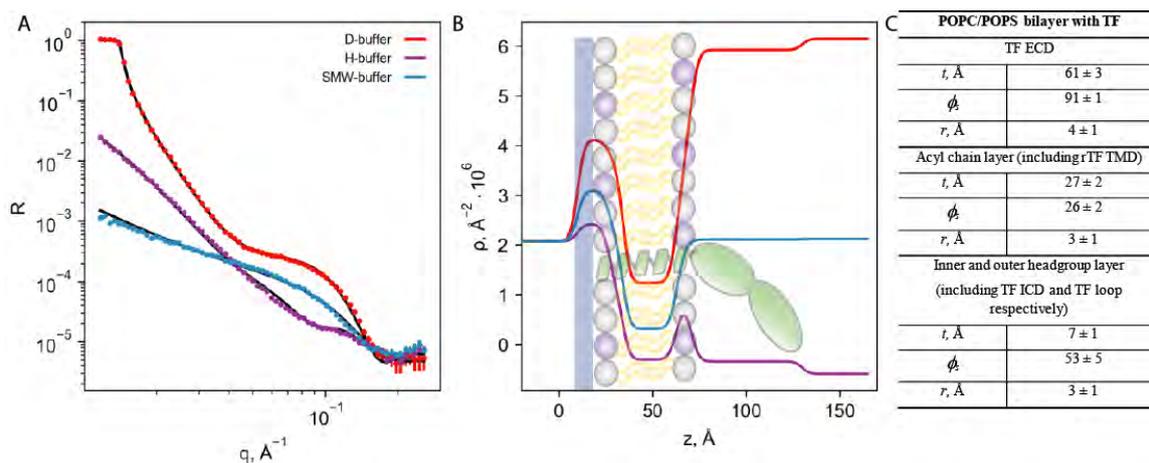


Figure 3.18: Orientation of TF in a lipid bilayer. A: NR data with corresponding fits. B: Calculated SLD functions with an overlay of the illustrated TF orientation. C: Fitting parameters from the fit in A.

### 3.2.2 Binding of FVIIa

The next step was to investigate the FVIIa interaction. As shown in figure 3.12, FVIIa was able to bind TF on a surface, meaning that NR experiments could be performed. The obtained data (figure 3.19A) could be fitted using the same strategy applied when analyzing TF in the membrane alone, only now factoring in FVIIa.

Once again, similar parameters were again obtained for the lipid bilayer, but this time the height of the protein layer (now TF and FVIIa) was determined to be  $99 \pm 3 \text{\AA}$ . This coincides with a greater height of FVIIa (compared to TF), and still confers a tilted angle of the complex.

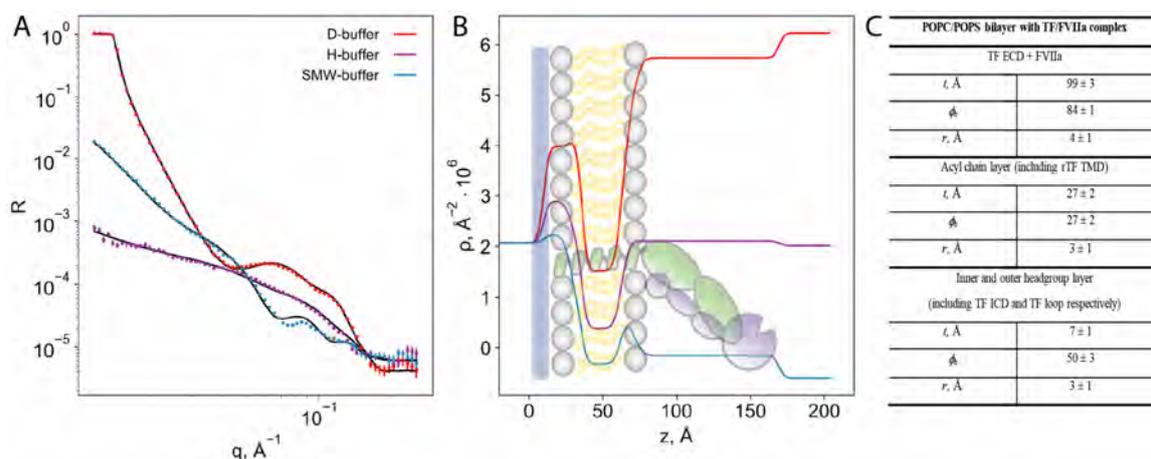


Figure 3.19: Orientation of the TF:FVIIa complex in a lipid bilayer. A: NR data with corresponding fits. B: Calculated SLD functions with an overlay of the illustrated complex. C: Fitting parameters from the fit in A.

In addition to these two measurements, we also used NR to measure FXa on a pure bilayer revealing a FXa height of  $92 \pm 3 \text{\AA}$ . The collected data from all the measured heights, and the available structural information on the proteins, allowed us to use simple trigonometry to calculate the orientations of the proteins (figure 3.20). From the available crystal structure [25], the longest distance in the ECD of TF was measured to be  $80 \text{\AA}$  (figure 3.20A). Measurements

on FVIIa (figure 3.20B) and FXa (figure 3.20C) were based on the model presented in [30]. The angle of TF in relation to the membrane was calculated to  $41^\circ$  (figure 3.20D) which probably reflects a broader range of conformations, with the average angle of  $41^\circ$ . This was also demonstrated by MD simulations, which showed that TF interacted with the lipid headgroups in a minor fraction of the time, while the majority of the time was spent occupying conformations corresponding to the NR values (Paper VI).

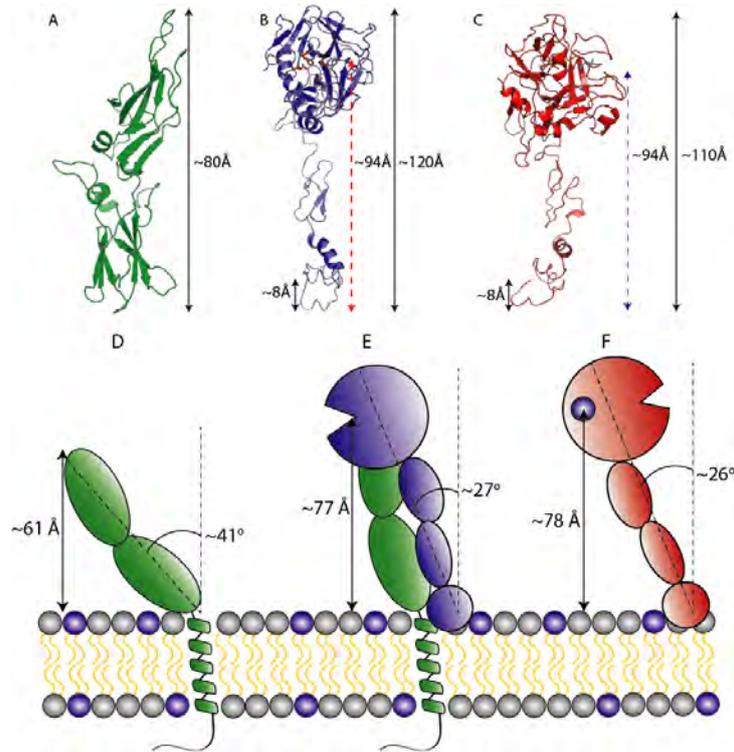


Figure 3.20: Distances and calculated angles in the TF, FVIIa and FXa lipid interplay. A: sTF, pdb-ID: 1BOY. B: FVIIa: Model from [30]. Active site residues in red. C: FXa: Model from [30]. New terminal residues after cleavage by FVIIa in teal. D: TF in a membrane with the height from NR and calculated angle. E: The TF:FVIIa complex in a membrane with the calculated angle and height of active site over the membrane. F: FXa on a membrane with calculated angle and height of cleavage site over the membrane.

The FVIIa and FXa Gla-domains have been reported to penetrate a PS containing membrane [209], and thus a penetration depth ( $8\text{\AA}$ ) was subtracted from the height when performing the calculations. Using the calculated angle of the TF:FVIIa complex and the position of the active site on FVIIa (red residues, figure 3.20B), the height of the active site over the membrane was calculated to be  $77\text{\AA}$  (figure 3.20E). This coincides well with the  $75 \pm 3\text{\AA}$  determined by FRET [67, 68] and the  $79\text{\AA}$  found in an MD study [69]. Using the NR measurements on FXa, a similar calculation was performed using the remaining residue left after cleavage of the scissile bond in FX (teal residue, figure 3.20C, blue circle, figure 3.20E). The angles of TF:FVIIa and FXa were almost identical and intriguingly, so were the heights of the FVIIa active site and the cleavage site on FXa. This suggests that an important feature of the membrane is positioning the two components for the catalytic reaction in the same height. It should be noted that the height measurements have to be confirmed using FX, but as there is no considerable difference between FX and FXa in the Gla-domain, the height is expected to be similar. In addition, it is important to underline that the NR data contains no orientational information, meaning

that the measurements of the active site and the cleavage sites are connected with uncertainties related to protein orientation.

Whether this particular localization of the active site is the underlying mechanism of the activity increase upon PS exposure, is still unclear. As the binding between the Gla-domain and the membrane is expected to change considerably when no PS is present, this will most likely also change the orientation of FVIIa, thus removing it from the FX cleavage site. However, NR measurements on the same samples in a pure PC membrane is needed to confirm this hypothesis.

### 3.3 Investigation of the quaternary complex

The overall aim of this project was to investigate the tertiary complex of TF, FVIIa, and FXa. A couple of computational models have been published on the soluble part of the complex [30, 210], but experimental evidence to support these are yet to be published. It is well established that a lipid bilayer is needed for the assembly of the complex, and thus the csE3 NDs developed here were utilized. In this section, results on the assembly of the complex is shown together with initial structural data describing the complex on an ND.

#### 3.3.1 Assembly of a multicomponent complex on a nanodisc

The aim was to investigate the complex on an ND using SEC-SAS, and the first step was to investigate the behavior of the samples using SEC. Therefore, a sample of TF in a POPC:POPS ND was prepared (details in Paper III), and split into four samples to which a different number of binding partners were added (figure 3.21, orange, green, blue and purple). All samples were analyzed using SEC (including an empty reference disc and a FXa sample) to evaluate the elution volumes (figure 3.21). The TF ND was enlarged upon addition of FVIIa (from orange to green), but was not further enlarged upon addition of FXa (blue). Therefore, the inhibitor of the complex, NAPc2, was added which caused an earlier elution (purple), suggesting that the quaternary complex was stable throughout an SEC run. This was very important for the subsequent SEC-SAS measurements, but the need of NAPc2 to be present further challenges following data modeling, as yet another protein has to be taken into account.

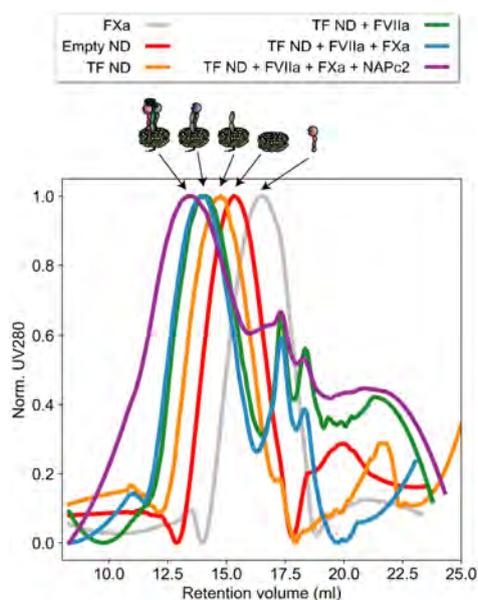


Figure 3.21: SEC analysis of the assembly of the quaternary TF:FVIIa:FXa:NAPc2 complex on an ND.

### 3.3.2 Solution structure of NAPc2

As NAPc2 proved necessary to assemble a stable complex, in depth information on the NAPc2 solution structure had to be obtained. A new production protocol was established, and the inhibition of FVIIa (both in the presence and absence of FX) was measured, and found to be similar to previously reported values (Paper V). Subsequently, SAXS measurements were performed on NAPc2, and the resulting data was compared to the available NMR structure (red, figure 3.22B). As evident from the fit, discrepancies were found between the NMR structure and the measured SAXS curves. This was probably a result of the markedly different buffer condition used for the two experiments, and thus we had to come up with a new model representing the structure in the buffer conditions used for the SAS measurements. In particular, it became evident that the NMR structure was more compact than the solution structure, as seen by the smaller  $D_{max}$  (figure 3.22C). To further explore this, an MD simulation was performed, which showed the ensemble to have a high degree of flexibility, as demonstrated by the overlaid structures (figure 3.22A) (details in Paper V). The generated ensemble also had a lower average  $R_g$ -value as well as a lower  $D_{max}$  than the one found for the experimentally determined (figure 3.22C). Therefore, a Bayesian/Maximum Entropy approach [211] was applied to reweight the ensemble to better represent the experimental  $R_g$ . Both the simulation and the reweighting improved the fits (figure 3.22B) and the best frame from the ensemble fitted the data decently. However, discrepancies were still present in the high- $q$  region, which was probably due to the fact that NAPc2 is mainly stabilized by disulfide bridges and the salt bridges. The salt bridges are often over stabilized in MD simulations, and the model might therefore not be an entirely correct representation of the actual structure. However, this new model represents the solution structure much better than the NMR model and will become important in the modeling of the full complex.

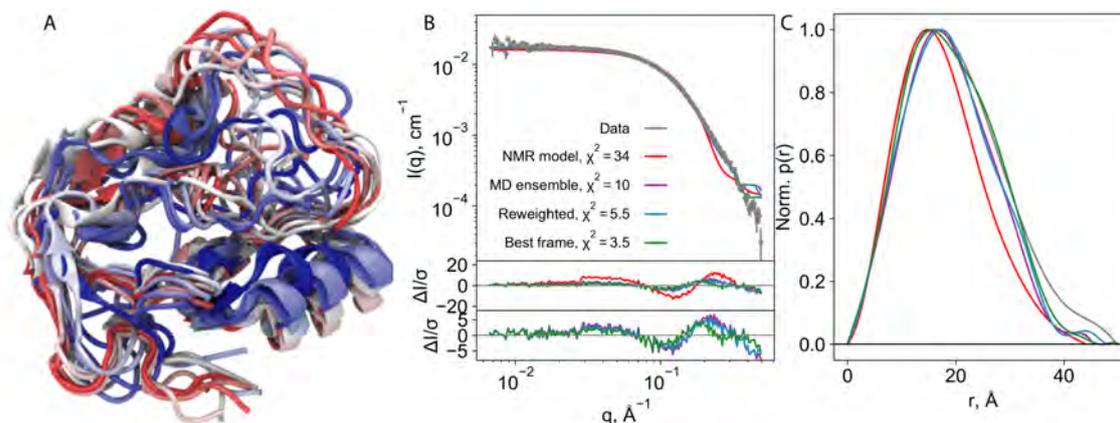


Figure 3.22: A: Overlaid structures taken every 40 ns of the simulation. B: SAXS data including fitted models. C:  $p(r)$ -functions of the same samples as in A. D: Kratky plot of the same samples as in A.

### 3.3.3 Overview of small angle scattering data

All SAS measurements on TF in NDs were performed in a SEC-SAS setup. This setup ensured removal of potential aggregates and excess binding partners, which would both have affected the resulting SAS data. The samples were all made from the same stock of TF in ND which was split into nine subsamples (three for SAXS and two times three for SANS), after which relevant binding partners were added. The SANS measurements were run twice to cover the desired  $q$ -range. Relevant frames were selected for further work (see details in Report I), and after a background subtraction, the SAXS data (figure 3.23A) exhibited an increasing  $I(0)$  as more binding partners were added. It should be mentioned that the concentration determinations from the SEC-SAS measurements are not particularly precise, which probably influences the obtained  $I(0)$ s. The increase in  $I(0)$  was also reflected in the  $p(r)$ -functions (figure 3.23B), as the  $D_{max}$  increased from  $\sim 130\text{\AA}$  (empty ND) to  $\sim 150\text{\AA}$  (TF ND) to  $\sim 220\text{\AA}$  (TF ND + FVIIa) and to  $\sim 290\text{\AA}$  (TF ND + FVIIa + FXa + NAPc2). The same trend was seen in the  $p(r)$ -functions from the SANS measurements (figure 3.23D), although all the  $D_{max}$ -values were smaller than for SAXS. This difference probably relies in the way the  $p(r)$ -functions are calculated. The contrast situations of the samples are very different in the SAXS and SANS measurements. The lipid tails have the largest contrast in SANS, while the MSPs have the largest contrast in SAXS, meaning that the interior of the ND scatters the most in SANS and the periphery in SAXS. When calculating  $p(r)$ -functions, the curve is ‘pushed’ towards 0 by the algorithm. This means a low signal at large distances will be forced to zero and hence a low  $D_{max}$  will be obtained, as in this case for SANS. In the SAXS contrast, the large distances will scatter more (due to the high contrast of the MSPs) and hence a larger  $D_{max}$  is obtained.

The SAXS data for the empty ND (figure 3.23A, red) demonstrated the expected double bump in the  $0.1\text{ \AA}^{-1}$  region, which is characteristic for an ND [152]. The same feature was present when TF was incorporated in the ND (green), but disappeared when FVIIa was added. The SANS data did not contain particular features, but still serves as an important restraint for the fitting when modeling will be performed.

In conclusion, the SAS data demonstrated the expected overall features (figure 3.23), and indeed

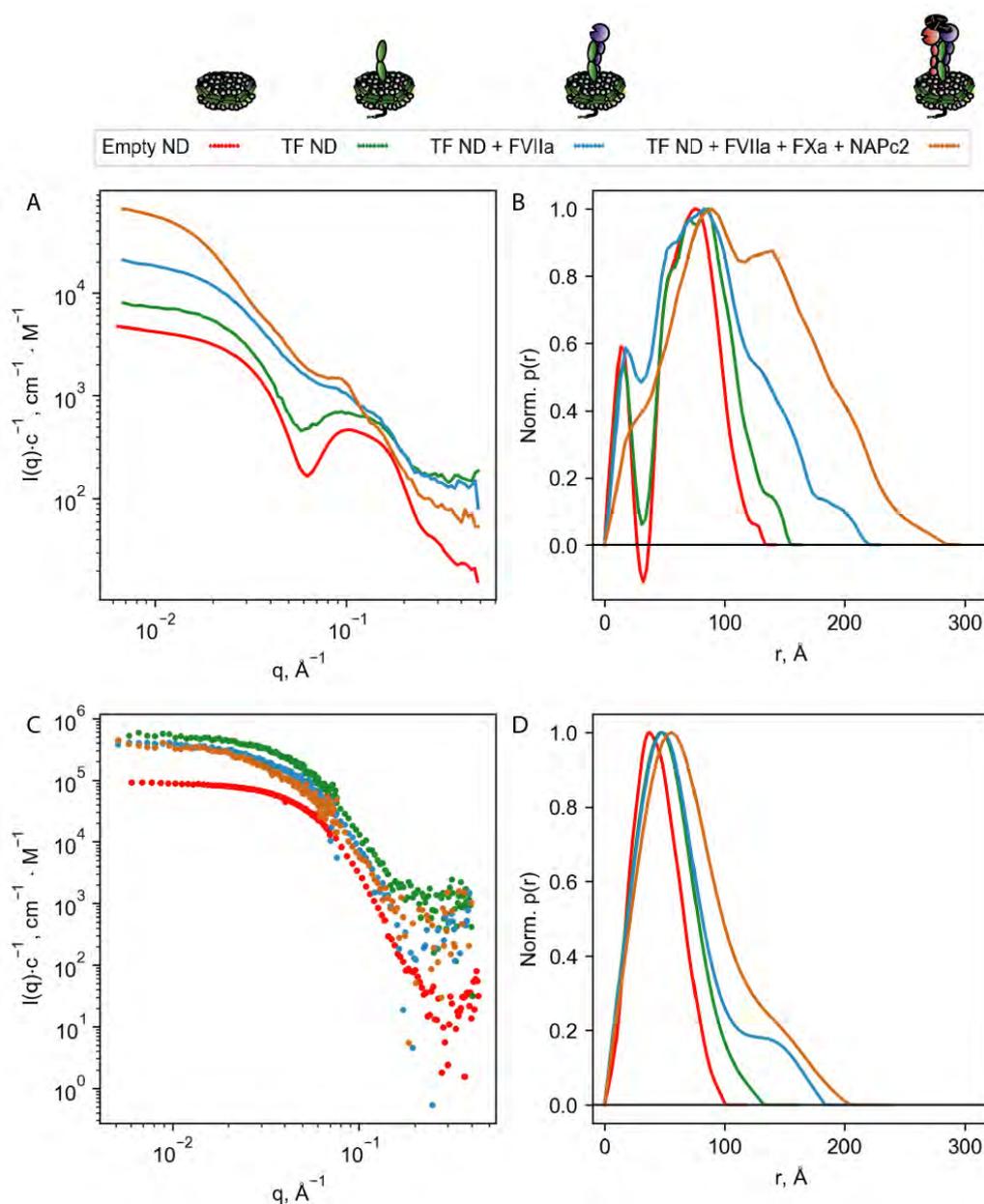


Figure 3.23: SAS data on four different samples: Empty ND (red), TF containing ND (green), TF containing ND + FVIIa (blue) and TF containing ND + FVIIa + FXa + NAPc2. A: SAXS data on absolute scale and normalized by concentration. B:  $p(r)$ -functions of the data presented in A. C: SANS data on absolute scale and concentration normalized. D:  $p(r)$ -functions of the data presented in C.

showed that it was possible to obtain data on the quaternary complex on an ND. Due to time limitations modeling of the data has not been performed yet, but the foundation for an experimentally supported model of the complex, is definitely present.

To further support the modeling of the complex, the previously described dcsNDs were utilized. TF was incorporated, and after purification, as usual, FVIIa, FXa, and NAPc2 were added in excess. As the ND was assembled with only dPOPC, the binding of the interaction partners was unclear. However, when analyzed by SEC, the elution volume was identical to the normal NDs with the complex assembled (figure 3.24A, red). A reference hND with the complex was

also made (figure 3.24A, blue) and fractions on the left-hand side of the peak were selected for SANS measurements. The SANS data (figure 3.24B) showed a flat Guinier region for the hND, but an ascending region for the dND. The reason was unclear, but the scattering from the dND was relatively weak due to a low concentration (same concentration as the hND, but much less scattering material) and the low signal could be an explanation.

The features in the SANS data were also reflected in the  $p(r)$ -functions, showing the expected  $D_{max}$  for the hND. On the other hand, the dND demonstrated an elongated  $D_{max}$ , which was unreliable, as the  $D_{max}$  was expected to be lower than for the hND.

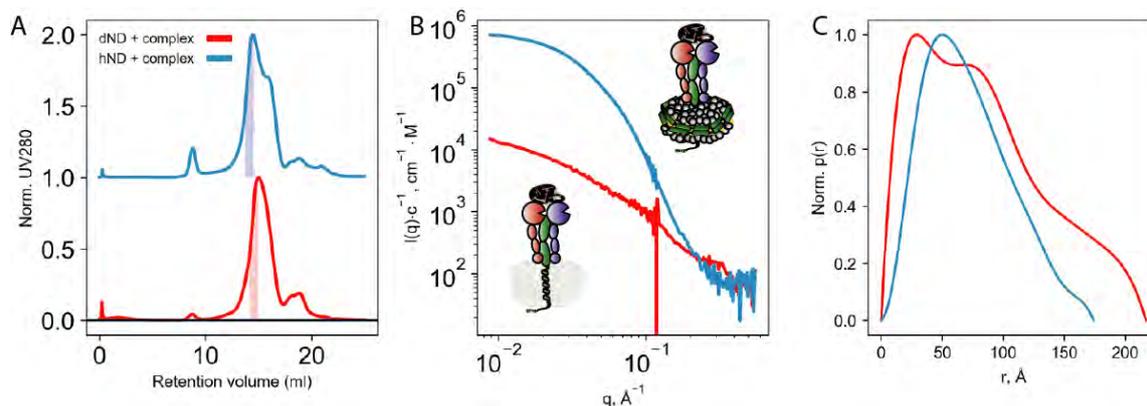


Figure 3.24: Comparison of the quaternary complex TF:FVIIa:FXa:NAPc2 assembled on a dND and an hND. A: SEC-chromatograms of the two different samples. Marked areas were the fractions selected for SANS measurements. B: SANS data on absolute scale and concentration normalized. C:  $p(r)$ -functions obtained using the Bayesapp server.

In summary, the dNDs provides the unique opportunity to obtain a scattering signal from the complex exclusively, and not the carrier system. The initial SEC run showed a promising peak at the expected elution volume, however, the scattering data was not as expected, and optimization can be done by testing a higher sample concentration.

## 4 | Conclusions and outlooks

To summarize the results presented in this thesis, the three overall objectives will be used, as these frame the main findings and the outlooks of the project.

### **1. Establish protocols which enable structural studies on TF in a membrane environment.**

Objective 1 has been the most time consuming, as it involved much optimization before the protocols yielded satisfying results. This has included optimization of the refolding time and dilution factor of TF, which proved to increase the refolding yield from 10% to 80%. This high refolding yield laid the foundation for the final yield of 5 mg per liter of culture, which provided sufficient amounts of TF for subsequent studies. Lastly, the activity and interaction with FVIIa were probed using an activity assay and SPR, respectively. Both provided similar results compared to earlier studies, confirming a successful refolding (Paper IV). Another part of the optimization process has been on NDs, as we devised the new csE3 NDs with additional negatively charged residues. These NDs proved to possess higher temporal stability than the template MSP1E3D1 at physiologically relevant temperatures, and at the same time they maintain the structural integrity as demonstrated by SAXS (Paper I). This study was followed up by the introduction of the fusion construct, containing both the MSP and the Sortase, which proved able to perform the circularization reaction without the need to produce Sortase in parallel. Using this construct, post-reconstitution circularization was performed. Using TF as a model MP, both the structural and functional integrity was shown to be conserved throughout the reconstitution and subsequent circularization. This novel circularization protocol has the potential to open up the cND technology to MPs with large hydrophilic domains on either sides of the membrane, such as GPCRs (Paper II).

The last method development study, was conducted to enable formation of a lipid bilayer on a surface with oriented MPs. Currently available strategies all possessed drawbacks, and thus a strategy involving peptidediscs was employed. This enabled an easy preparation of a surface with the desired lipid composition and unidirectional MPs (Paper III).

### **2. Investigate the mechanism behind the lipid-induced increase in activity of TF:FVIIa.**

Using NR on samples prepared using the new peptidedisc deposition, the heights of TF, the TF:FVIIa complex and FXa on a membrane were determined (Paper V). Using these heights together with the maximal distances from the crystal structures enabled determination of the membrane orientation of the proteins. Using this information, the height of the FVIIa active site and the FX cleavage site were determined to be 77Å and 78Å, respectively. Localizing

these two sites at the same height is probably one of the important features of the membrane, and likely a part of the mechanism behind the lipid-induced activity increase. To elucidate the mechanism behind the difference in PC and PS activity, data in a pure POPC membrane has to be obtained. This data, in combination with simulations, will hopefully shed more light on the presently elusive mechanism.

### **3. Obtain experimental data on the structure of the TF:FVIIa:FXa complex.**

The high yield of TF and the stable NDs achieved as a part of objective 1 proved critical for the success in this third objective. The inhibitor of the complex, NAPc2, proved to be essential in forming a stable quaternary complex on an ND, and thus NAPc2 also had to be structurally characterized. The SAXS data indicated an more extended structure of NAPc2 compared to the relatively rigid NMR structure available. Using MD simulations we refined the model of the NAPc2 solution structure, to match the obtained SAXS data. This was a flexible structure, which was mainly stabilized by the five disulfide and salt bridges, and interestingly almost no hydrophobic core nor secondary structure (Paper VI).

To investigate the entire complex on an ND, SEC-SAS measurements were performed. This provided scattering data without contributions from aggregates or excess interaction partners. The data shows the expected overall features, as the size of the particle increases upon the addition of the interaction partners, but further modeling has not been performed yet. However, this is the first experimental data on this complex, and also the first data in a lipid environment (Report I). Initial data on the complex in a dND was also obtained, but unfortunately, a low concentration hampered the data quality. However, the SEC data proved that it was indeed possible to assemble a TF containing dND and assemble the quaternary complex on the disc, which is very promising for future studies.

## References

- [1] Stephanie A Smith, Richard J Travers, and James H Morrissey. How it all starts: initiation of the clotting cascade. *Critical reviews in biochemistry and molecular biology*, 50(4):326–336, 2016. doi: 10.3109/10409238.2015.1050550.How.
- [2] Thomas Renné, Alvin H. Schmaier, Katrin F. Nickel, Margareta Blombäck, and Coen Maas. In vivo roles of factor XII. *Blood*, 120(22):4296–4303, 2012. ISSN 00064971. doi: 10.1182/blood-2012-07-292094.
- [3] L Autin, M. A. Miteva, W. H. Lee, K Mertens, K. P. Radtke, and B. O. Villoutreix. Molecular models of the procoagulant Factor VIIIa – Factor IXa complex. *Journal of Thrombosis and Haemostasis*, 3(May):2044–2056, 2005.
- [4] Nigel Mackman, Rachel E Tilley, and Nigel S Key. Role of the extrinsic pathway of blood coagulation in hemostasis and thrombosis. *Arteriosclerosis, thrombosis, and vascular biology*, 27(8):1687–93, 2007. ISSN 1524-4636. doi: 10.1161/ATVBAHA.107.141911.
- [5] Rebecca A. Fleck, L. Vijaya Mohan Rao, Samuel I. Rapaport, and Nissi Varki. Localization of human tissue factor antigen by immunostaining with monospecific, polyclonal anti-human tissue factor antibody. *Thrombosis Research*, 59(2):421–437, 1990. ISSN 00493848. doi: 10.1016/0049-3848(90)90148-6.
- [6] Bruce Furiel and Barbara C. Furie. The Molecular Basis of Blood Diseases. *Cell*, 53: 505–518, 1988. doi: 10.1136/jmg.25.11.789-a.
- [7] J H Morrissey, B G Macik, P F Neuenschwander, and P C Comp. Quantitation of activated factor VII levels in plasma using a tissue factor mutant selectively deficient in promoting factor VII activation. *Blood*, 81(3):734–44, 1993. ISSN 0006-4971.
- [8] Muriel Giansily-Blaziot and Jean-François Schved. Recombinant human factor VIIa (rFVIIa) in hemophilia: mode of action and evidence to date. *Therapeutic Advances in Hematology*, 8(12):345–352, 2017. ISSN 2040-6207. doi: 10.1177/2040620717737701.
- [9] Pierre F. Neuenschwander, Martine M. Fiore, and James H. Morrissey. Factor VII autoactivation proceeds via interaction of distinct protease-cofactor and zymogen-cofactor complexes: Implications of a two-dimensional enzyme kinetic mechanism. *Journal of Biological Chemistry*, 268(29):21489–21492, 1993. ISSN 00219258.

- [10] Sanjeev Palta, Richa Saroa, and Anshu Palta. Overview of the coagulation system. *Indian Journal of Anaesthesia*, 58(5):515–523, 2014. ISSN 00195049. doi: 10.4103/0019-5049.144643.
- [11] J. W M Heemskerk, Edouard M. Bevers, and Theo Lindhout. Platelet activation and blood coagulation. *Thrombosis and Haemostasis*, 88(2):186–193, 2002. ISSN 03406245.
- [12] Brit B. Sørensen, Per Ola Freskgård, Lars Søgaard Nielsen, L. Vijaya Mohan Rao, Mirella Ezban, and Lars C. Petersen. Factor VIIa-induced p44/42 mitogen-activated protein kinase activation requires the proteolytic activity of factor VIIa and is independent of the tissue factor cytoplasmic domain. *Journal of Biological Chemistry*, 274(30):21349–21354, 1999. ISSN 00219258. doi: 10.1074/jbc.274.30.21349.
- [13] Mehmet Sen, Mark Herzik, John W Craft Jr., Andrea L Creath, Sameer Agrawal, Wolfram Ruf, and Glen B Legge. Spectroscopic Characterization of Successive Phosphorylation of the Tissue Factor Cytoplasmic Region. *The open spectroscopy journal*, 3(713):58–64, 2009. ISSN 1874-3838. doi: 10.2174/1874383800903010058.
- [14] J P Burnier, M Borowski, B C Furie, and B Furie. Gamma-carboxyglutamic acid. *Molecular and cellular biochemistry*, 39:191–207, sep 1981. ISSN 0300-8177 (Print). doi: 10.1007/bf00232574.
- [15] James H. Morrissey. Coagulation Factor VIIa. *Handbook of Proteolytic Enzymes*, 277(8):2905–2908, 2013. doi: 10.1016/B978-0-12-382219-2.00641-4.
- [16] Lizbeth Hedstrom, Tiao Yin Lin, and Walter Fast. Hydrophobic interactions control zymogen activation in the trypsin family of serine proteases. *Biochemistry*, 35(14):4515–4523, 1996. ISSN 00062960. doi: 10.1021/bi951928k.
- [17] Joachim Walter, Wolfgang Steigemann, Tej Pal Singh, Hans Bartunik, Wolfram Bode, and Robert Huber. Trypsinogen: Chemical Labelling and Low- Temperature Crystallography. *Acta Cryst*, 38(1982):1462–1472, 1982.
- [18] Prosenjit Sen, Pierre F. Neuenschwander, Usha R. Pendurthi, and L. Vijaya Mohan Rao. Analysis of factor VIIa binding to relipidated tissue factor by surface plasmon resonance. *Blood Coagulation and Fibrinolysis*, 21(4):376–379, 2010. ISSN 09575235. doi: 10.1097/MBC.0b013e328333b084.
- [19] Andrew W. Shaw, Vincent S. Pureza, Stephen G. Sligar, and James H. Morrissey. The local phospholipid environment modulates the activation of blood clotting. *Journal of Biological Chemistry*, 282(9):6556–6563, 2007. ISSN 00219258. doi: 10.1074/jbc.M607973200.
- [20] David W Banner, Allan D’Arcy, Christiana Chene, Fritz K Winkler, Arabinda Guha, William H Konigsberg, Yale Nemerson, and Daniel Kirchhofer. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor, 1996. ISSN 0028-0836.
- [21] Shouichi Higashi, Naomi Matsumoto, and Sadaaki Iwanaga. Molecular mechanism of tissue factor-mediated acceleration of factor VIIa activity. *Journal of Biological Chemistry*, 271(43):26569–26574, 1996. ISSN 00219258. doi: 10.1074/jbc.271.43.26569.

- [22] Kasper D. Rand, Thomas J D Jørgensen, Ole H. Olsen, Egon Persson, Ole N. Jensen, Henning R. Stennicke, and Mette D. Andersen. Allosteric activation of coagulation factor VIIa visualized by hydrogen exchange. *Journal of Biological Chemistry*, 281(32):23018–23024, 2006. ISSN 00219258. doi: 10.1074/jbc.M602968200.
- [23] Egon Persson, Helle Bak, Anette Østergaard, and Ole H. Olsen. Augmented intrinsic activity of Factor VIIa by replacement of residues 305, 314, 337 and 374: Evidence of two unique mutational mechanisms of activity enhancement. *Biochemical Journal*, 379(2):497–503, 2004. ISSN 02646021. doi: 10.1042/BJ20031596.
- [24] E. Persson, L. S. Nielsen, and O. H. Olsen. Substitution of aspartic acid for methionine-306 in factor VIIa abolishes the allosteric linkage between the active site and the binding interface with tissue factor. *Biochemistry*, 40(11):3251–3256, 2001. ISSN 00062960. doi: 10.1021/bi001612z.
- [25] D W Banner, A D’Arcy, C ChŠne, F Vilbois, W H Konigsberg, A Guha, Y Nemerson, and D Kirchhofer. The Crystal Structure of the Complex of Human Factor VIIa With Human Soluble Tissue Factor. *Thrombosis and haemostasis*, 73:41–46, 1995.
- [26] Narjes Tavoosi, Stephanie A. Smith, Rebecca L. Davis-Harrison, and James H. Morrissey. Factor VII and protein C are phosphatidic acid-binding proteins. *Biochemistry*, 52(33):5545–5552, 2013. ISSN 00062960. doi: 10.1021/bi4006368.
- [27] Pierre F. Neuenschwander and James H. Morrissey. Roles of the membrane-interactive regions of factor VIIa and tissue factor: The factor VIIa Gla domain is dispensable for binding to tissue factor but important for activation of factor X. *Journal of Biological Chemistry*, 269(11):8007–8013, 1994. ISSN 00219258.
- [28] Stephen B. Harvey, Matthew D. Stone, Michael B. Martinez, and Gary L. Nelsestuen. Mutagenesis of the  $\gamma$ -carboxyglutamic acid domain of human factor VII to generate maximum enhancement of the membrane contact site. *Journal of Biological Chemistry*, 278(10):8363–8369, 2003. ISSN 00219258. doi: 10.1074/jbc.M211629200.
- [29] Mark A Brown, Leisa M Stenberg, and Johan Stenflo. Coagulation Factor Xa. *Handbook of Proteolytic Enzymes*, 9(1995):2908–2915, 2013. doi: <http://dx.doi.org/10.1016/B978-0-12-382219-2.00642-6>.
- [30] Brian V. Norledge, Ramona J. Petrovan, Wolfram Ruf, and Arthur J. Olson. The Tissue Factor/Factor VIIa/Factor Xa Complex: A Model Built by Docking and Site-Directed Mutagenesis. *Proteins: Structure, Function and Genetics*, 53(3):640–648, 2003. ISSN 08873585. doi: 10.1002/prot.10445.
- [31] Randal J. Kaufman. Post-translational modifications required for coagulation factor secretion and function. *Thrombosis and Haemostasis*, 79(6):1068–1079, 1998. ISSN 03406245.
- [32] Karin Carlsson, Per Ola Freskgård, Egon Persson, Uno Carlsson, and Magdalena Svensson. Probing the interface between factor Xa and tissue factor in the quaternary complex tissue factor-factor VIIa-factor Xa-tissue factor pathway inhibitor. *European Journal of*

- Biochemistry*, 270(12):2576–2582, 2003. ISSN 00142956. doi: 10.1046/j.1432-1033.2003.03625.x.
- [33] L. R. Paborsky and R. J. Harris. Post-translational modifications of recombinant human tissue factor. *Thrombosis Research*, 60(5):367–376, 1990. ISSN 00493848. doi: 10.1016/0049-3848(90)90219-3.
- [34] F. A. Pitlick. Concanavalin A inhibits tissue factor coagulant activity. *The Journal of clinical investigation*, 55(1):175–179, 1975. ISSN 00219738. doi: 10.1172/JCI107908.
- [35] H. Kothari, L. V.M. Rao, and U. R. Pendurthi. Glycosylation of tissue factor is not essential for its transport or functions. *Thrombosis and haemostasis*, 9(8):1511–1520, 2011. doi: 10.1038/mp.2011.182.doi.
- [36] M J Stone, W Ruf, D J Miles, T S Edgington, and P E Wright. Recombinant soluble human tissue factor secreted by *Saccharomyces cerevisiae* and refolded from *Escherichia coli* inclusion bodies: glycosylation of mutants, activity and physical characterization. *The Biochemical journal*, 310 ( Pt 2:605–14, 1995. ISSN 0264-6021.
- [37] Jolanta Krudysz-Amblo, Mark E. Jennings, Kenneth G. Mann, and Saulius Butenas. Carbohydrates and activity of natural and recombinant tissue factor. *Journal of Biological Chemistry*, 285(5):3371–3382, 2010. ISSN 00219258. doi: 10.1074/jbc.M109.055178.
- [38] L. V.M. Rao and U. R. Pendurthi. Regulation of tissue factor coagulant activity on cell surfaces. *Journal of Thrombosis and Haemostasis*, 10(11):2242–2253, 2012. ISSN 15387933. doi: 10.1111/jth.12003.
- [39] K. Hansson and J. Stenflo. Post-translational modifications in proteins involved in blood coagulation. *Journal of Thrombosis and Haemostasis*, 3(12):2633–2648, 2005. ISSN 1538-7933. doi: 10.1111/j.1538-7836.2005.01478.x.
- [40] François Fenaille, Catherine Groseil, Christine Ramon, Sandrine Riandé, Laurent Siret, Sami Chtourou, and Nicolas Bihoreau. Mass spectrometric characterization of N- and O-glycans of plasma-derived coagulation factor VII. *Glycoconjugate Journal*, 25(9):827–842, 2008. ISSN 02820080. doi: 10.1007/s10719-008-9143-7.
- [41] Masaki Iino, Donald C. Foster, and Walter Kisiel. Functional consequences of mutations in Ser-52 and Ser-60 in human blood coagulation factor VII. *Archives of Biochemistry and Biophysics*, 352(2):182–192, 1998. ISSN 00039861. doi: 10.1006/abbi.1998.0595.
- [42] Keisuke Inoue and Takashi Morita. Identification of O-linked oligosaccharide chains in the activation peptides of blood coagulation factor X. *European Journal of Biochemistry*, 218(1):153–163, 1993.
- [43] Likui Yang, Chandrashekhara Manithody, and Alireza R. Rezaie. Functional Role of O-linked and N-linked Glycosylation Sites Present on the Activation Peptide of Factor X. *Thrombosis and Haemostasis*, 7(10):1696–1702, 2009. doi: 10.1038/mp.2011.182.doi.

- [44] J. R. Maynard, C. A. Heckman, F. A. Pitlick, and Y. Nemerson. Association of tissue factor activity with the surface of cultured cells. *Journal of Clinical Investigation*, 55(4): 814–824, 1975. ISSN 00219738. doi: 10.1172/JCI107992.
- [45] H. Kothari, U. R. Pendurthi, and L. V.M. Rao. Analysis of tissue factor expression in various cell model systems: cryptic vs. active. *Thrombosis and Haemostasi*, 11(7):1353–1363, 2013. doi: 10.1038/mp.2011.182.doi.
- [46] P. L. A. Giesen, U. Rauch, B. Bohrmann, D. Kling, M. Roque, J. T. Fallon, J. J. Badimon, J. Himber, M. A. Riederer, and Y. Nemerson. Blood-borne tissue factor: Another view of thrombosis. *Proceedings of the National Academy of Sciences*, 96(5):2311–2315, 2002. ISSN 0027-8424. doi: 10.1073/pnas.96.5.2311.
- [47] Shabbir Ansari, Usha Pendurthi, and L. Rao. Role of Cell Surface Lipids and Thiol-Disulphide Exchange Pathways in Regulating the Encryption and Decryption of Tissue Factor. *Thrombosis and Haemostasis*, pages 860–870, 2019. ISSN 0340-6245. doi: 10.1055/s-0039-1681102.
- [48] J. Fernando Bazan. Structural design and molecular evolution of a cytokine receptor superfamily (hematopoietic system/interferon/tissue factor/fibronectin/immunoglobulin). *Proc. Natl. Acad. Sci. USA*, 87(September):6934–6938, 1990. ISSN 0027-8424. doi: 10.1073/pnas.87.18.6934.
- [49] A M de Vos, M Ultsch, and A A Kossiakoff. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science*, 255(5042):306 LP – 312, jan 1992. doi: 10.1126/science.1549776.
- [50] Oded Livnah, Enrico A Stura, Dana L Johnson, Steven A Middleton, Linda S Mulcahy, Nicholas C Wrighton, William J Dower, Linda K Jolliffe, and Ian A Wilson. Functional Mimicry of a Protein Hormone by a Peptide Agonist: The EPO Receptor Complex at 2.8 Å. *Science*, 273(5274):464 LP – 471, jul 1996. doi: 10.1126/science.273.5274.464.
- [51] S. Roy, L. R. Paborsky, and G. A. Vehar. Self-association of tissue factor as revealed by chemical cross-linking. *Journal of Biological Chemistry*, 266(8):4665–4668, 1991. ISSN 00219258.
- [52] Ronald R. Bach and Charles F. Moldow. Mechanism of Tissue Factor Activation on HL-60 Cells. *Blood*, 9 Suppl 1:S27–35, 1998.
- [53] F. Donate, C. R. Kelly, W. Ruf, and T. S. Edgington. Dimerization of tissue factor supports solution-phase autoactivation of factor VII without influencing proteolytic activation of factor X. *Biochemistry*, 39(37):11467–11476, 2000. ISSN 00062960. doi: 10.1021/bi000986p.
- [54] S D Carson and M E Bromberg. Tissue factor encryption/de-encryption is not altered in the absence of the cytoplasmic domain. *Thrombosis and haemostasis*, 84(4):657–663, oct 2000. ISSN 0340-6245 (Print).

- [55] A S Wolberg, R H Kon, D M Monroe, M Ezban, H R Roberts, and M Hoffman. De-encryption of cellular tissue factor is independent of its cytoplasmic domain. *Biochemical and Biophysical Research Communications*, 272(2):332–336, 2000. ISSN 0006291X. doi: 10.1006/bbrc.2000.2783.
- [56] Alnawaz Rehemtulla, Wolfram Ruf, and Thomas S. Edgington. The integrity of the cysteine 186-cysteine 209 bond of the second disulfide loop of tissue factor is required for binding of factor VII. *Journal of Biological Chemistry*, 266(16):10294–10299, 1991. ISSN 00219258.
- [57] Ramesh Prasad, Suparna Banerjee, and Prosenjit Sen. Contribution of allosteric disulfide in the structural regulation of membrane-bound tissue factor–factor VIIa binary complex. *Journal of Biomolecular Structure and Dynamics*, 0(0):1–14, 2018. ISSN 15380254. doi: 10.1080/07391102.2018.1526118.
- [58] Gerrit Van Meer, Dennis R. Voelker, and Gerald W Feigenson. Membrane lipids: where they are. *Nature Reviews Molecular Cell Biology*, 10(1):1–4, 2009. ISSN 1362-4962. doi: 10.1038/nrm2330.Membrane.
- [59] Akiko Y Amaji Asegawa and Masafumi T Sujimoto. Lipid Dynamics and Pathobiology in Membrane Lipid Rafts Asymmetric Distribution of Phospholipids in Biomembranes. *Biological & Pharmaceutical Bulletin*, 29(August):1547–1553, 2006.
- [60] Gerrit Van Meer. Dynamic Transbilayer Lipid Asymmetry. *Cold Spring Harbor perspectives in biology*, 3:1–11, 2011.
- [61] A. J. Verkleij, R. F A Zwaal, B. Roelofsen, P. Comfurius, D. Kastelijn, and L. L M van Deenen. The asymmetric distribution of phospholipids in the human red cell membrane. A combined study using phospholipases and freeze-etch electron microscopy. *BBA - Biomembranes*, 323(2):178–193, 1973. ISSN 00052736. doi: 10.1016/0005-2736(73)90143-0.
- [62] Drew Marquardt, Barbara Geier, and Georg Pabst. Asymmetric lipid membranes: Towards more realistic model systems. *Membranes*, 5(2):180–196, 2015. ISSN 20770375. doi: 10.3390/membranes5020180.
- [63] Aaron J Marcus. Lipids in Blood Clotting. *JAMA*, 194(1):100–101, 1965. ISSN 0098-7484. doi: 10.1001/jama.1965.03090140108048.
- [64] Yale Nemerson. The phospholipid requirement of tissue factor in blood coagulation. *Journal of Clinical Investigation*, 47(1):72–80, 1968. ISSN 0021-9738. doi: 10.1172/jci105716.
- [65] James H. Morrissey, Rebecca L. Davis-Harrison, Narjes Tavoosi, Ke Ke, Vincent Pureza, John M. Boettcher, Mary C. Clay, Chad M. Rienstra, Y. Zenmei Ohkubo, Taras V. Pogorelov, and Emad Tajkhorshid. Protein-Phospholipid Interactions in Blood Clotting James. *Thrombosis Research*, 2010. ISSN 19326203. doi: 10.1371/journal.pone.0178059.
- [66] L. Vijaya Mohan Rao, H. Kothari, and Usha R. Pendurthi. Tissue Factor: Mechanisms of Decryption. *Frontiers in Bioscience*, 2014. doi: 10.1038/mp.2011.182.doi.

- [67] C D McCallum, R C Hapak, P F Neuenschwander, J H Morrissey, and A E Johnson. The location of the active site of blood coagulation factor VIIa above the membrane surface and its reorientation upon association with tissue factor - A fluorescence energy transfer study. *Journal of Biological Chemistry*, 271(45):28168–28175, 1996. ISSN 0021-9258. doi: 10.1074/jbc.271.45.28168.
- [68] Christine D. McCallum, Bixia Su, Pierre F. Neuenschwander, James H. Morrissey, and Arthur E. Johnson. Tissue factor positions and maintains the factor VIIa active site far above the membrane surface even in the absence of the factor VIIa Gla domain. A fluorescence resonance energy transfer study. *Journal of Biological Chemistry*, 272(48):30160–30166, 1997. ISSN 00219258. doi: 10.1074/jbc.272.48.30160.
- [69] Y. Zenmei Ohkubo and Emad Tajkhorshid. Distinct Structural and Adhesive Roles of Ca<sup>2+</sup> in Membrane Binding of Blood Coagulation Factors. *Structure*, 16(1):72–81, 2008. ISSN 09692126. doi: 10.1016/j.str.2007.10.021.
- [70] Emily K Waters and James H Morrissey. Restoring full biological activity to the isolated ectodomain of an integral membrane protein. *Biochemistry*, 45(11):3769–3774, mar 2006. ISSN 0006-2960 (Print). doi: 10.1021/bi052600m.
- [71] Ke Ke, Jian Yuan, and James H. Morrissey. Tissue factor residues that putatively interact with membrane phospholipids. *PLoS ONE*, 9(2), 2014. ISSN 19326203. doi: 10.1371/journal.pone.0088675.
- [72] Y. Zenmei Ohkubo, James H. Morrissey, and Emad Tajkhorshid. Dynamical view of membrane binding and complex formation of human factor VIIa and tissue factor. *Journal of Thrombosis and Haemostasis*, 8(5):1044–1053, 2010. ISSN 15378276. doi: 10.1158/0008-5472.CAN-10-4002.BONE.
- [73] Daniel Kirchhofer, Michael T. Lipari, Paul Moran, Charles Eigenbrot, and Robert F. Kelley. The tissue factor region that interacts with substrates factor IX and factor X. *Biochemistry*, 39(25):7380–7387, 2000. ISSN 00062960. doi: 10.1021/bi000182+.
- [74] Qiuling Huang, Pierre F. Neuenschwander, Alireza R. Rezaie, and James H. Morrissey. Substrate Recognition by Tissue Factor-Factor VIIa. *Journal of Biological Chemistry*, 271(36):21752–21757, 2002. ISSN 0021-9258. doi: 10.1074/jbc.271.36.21752.
- [75] Pierre F. Neuenschwander, Emma Bianco-Fisher, Alireza R. Rezaie, and James H. Morrissey. Phosphatidylethanolamine Augments Factor VIIa-Tissue Factor Activity: Enhancement of Sensitivity to Phosphatidylserine. *Biochemistry*, 34(43):13988–13993, 1995. ISSN 15204995. doi: 10.1021/bi00043a004.
- [76] Narjes Tavoosi, Rebecca L. Davis-Harrison, Taras V. Pogorelov, Y. Zenmei Ohkubo, Mark J. Arcario, Mary C. Clay, Chad M. Rienstra, Emad Tajkhorshid, and James H. Morrissey. Molecular determinants of phospholipid synergy in blood clotting. *Journal of Biological Chemistry*, 286(26):23247–23253, 2011. ISSN 00219258. doi: 10.1074/jbc.M111.251769.

- [77] James H. Morrissey, Emad Tajkhorshid, Stephen G. Sligar, and Chad M. Rienstra. Tissue factor/factor VIIa complex: Role of the membrane surface. *Thrombosis Research*, 129 (SUPPL. 2):S8–S10, 2012. ISSN 00493848. doi: 10.1016/j.thromres.2012.02.019.
- [78] Jue Wang, Usha R. Pendurthi, and L. Vijaya Mohan Rao. Sphingomyelin encrypts tissue factor: ATP-induced activation of A-SMase leads to tissue factor decryption and microvesicle shedding. *Blood Advances*, 1(13):849–862, 2017. ISSN 2473-9529. doi: 10.1182/bloodadvances.2016003947.
- [79] Dennis J. Dietzen, Keith L. Page, and Tina A. Tetzloff. Lipid rafts are necessary for tonic inhibition of cellular tissue factor procoagulant activity. *Blood*, 103(8):3038–3044, 2004. ISSN 00064971. doi: 10.1182/blood-2003-07-2399.
- [80] Vineet Awasthi, Samir K. Mandal, Veena Papanna, L. Vijaya Mohan Rao, and Usha R. Pendurthi. Modulation of tissue factor-factor VIIa signaling by lipid rafts and caveolae. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 27(6):1447–1455, 2007. ISSN 10795642. doi: 10.1161/ATVBAHA.107.143438.
- [81] Susan A. Maroney, Paul E. Ellery, Jeremy P. Wood, Josephine P. Ferrel, Catherine E. Bonesho, and Alan E. Mast. Caveolae Optimize Tissue Factor-Factor VIIa Inhibitory Activity of Cell Surface Associated Tissue Factor Pathway Inhibitor. *Biochemical Journal*, 443(1):259–266, 2012. ISSN 19326203.
- [82] Samir K. Mandal, Alexei Iakhiaev, Usha R. Pendurthi, and L. Vijaya Mohan Rao. Acute cholesterol depletion impairs functional expression of tissue factor in fibroblasts: modulation of tissue factor activity by membrane cholesterol. *Blood*, 105(1):153–160, 2005. doi: 10.1182/blood-2004-03-0990.Acute.
- [83] Corinne Kunzelmann-Marche, Jean-Marie Freyssinet, and M. Carmen Martinez. Loss of Plasma Membrane Phospholipid Asymmetry Requires Raft Integrity. *Journal of Biological Chemistry*, 277(22):19876–19881, 2002. ISSN 0021-9258. doi: 10.1074/jbc.m200324200.
- [84] J H Morrissey, B G Macik, P F Neuenschwander, and P C Comp. Quantitation of activated factor VII levels in plasma using a tissue factor mutant selectively deficient in promoting factor VII activation. *Blood*, 81(3):734–44, 1993. ISSN 0006-4971.
- [85] Peter Wildgoose and Walter Kisiel. Activation of Human Factor VII by Factors IXa and Xa on Human Bladder Carcinoma Cells. *Blood*, 7:1888–1895, 1989.
- [86] Robert Radcliffe, Andranik Bagdasarian, Robert Colman, and Yale Nemerson. Activation of Bovine Factor VII by Hageman Factor Fragments. *Blood*, 50(4), 1977.
- [87] Robert Radcliffe and Yale Nemerson. Activation Activated Factor Control X and of Factor Thrombin VII by of. *The Journal of biological chemistry*, 250(2):388–395, 1975.
- [88] S. P. Bajaj, S. I. Rapaport, and S. F. Brown. Isolation and characterization of human Factor VII. Activation of Factor VII by Factor X(a). *Journal of Biological Chemistry*, 256 (1):253–259, 1981. ISSN 00219258.

- [89] G. J. Broze and P. W. Majerus. Purification and properties of human coagulation factor VII. *Journal of Biological Chemistry*, 255(4):1242–1247, 1980. ISSN 00219258.
- [90] Emily K. Waters, Subramanian Yegneswaran, and James H. Morrissey. Raising the active site of factor VIIa above the membrane surface reduces its procoagulant activity but not factor VII autoactivation. *Journal of Biological Chemistry*, 281(36):26062–26068, 2006. ISSN 00219258. doi: 10.1074/jbc.M604915200.
- [91] Tomohiro Nakagaki, Donald C Foster, Kathleen L Berkner, and Walter Kisiel. Initiation of the extrinsic pathway of blood coagulation: evidence for the tissue factor dependent autoactivation of human coagulation factor VII. *Biochemistry*, 30(45):10819–10824, nov 1991. ISSN 0006-2960. doi: 10.1021/bi00109a001.
- [92] Qiuling Huang, Pierre F Neuenschwander, Alireza R Rezaie, and James H Morrissey. Substrate Recognition by Tissue Factor-Factor VIIa. *J Biol Chem.*, 271(36):21752–21757, 1996.
- [93] Sabine Dittmar, Wolfram Ruf, and Thomas S. Edgington. Influence of mutations in tissue factor on the fine specificity of macromolecular substrate activation. *Biochemical Journal*, 321(3):787–794, 1997. ISSN 0264-6021. doi: 10.1042/bj3210787.
- [94] B N Violand, M R Schlittler, K L Duffin, and C E Smith. Determination of the disulfide bond pairings in human tissue factor pathway inhibitor purified from Escherichia coli. *Journal of protein chemistry*, 14(5):341–347, 1995.
- [95] Robert J. Baugh, George J. Broze, and Sriram Krishnaswamy. Regulation of extrinsic pathway factor Xa formation by tissue factor pathway inhibitor. *Journal of Biological Chemistry*, 273(8):4378–4386, 1998. ISSN 00219258. doi: 10.1074/jbc.273.8.4378.
- [96] Jeremy P. Wood, Paul E R Ellery, Susan A. Maroney, and Alan E. Mast. Biology of tissue factor pathway inhibitor. *Blood*, 123(19):2934–2943, 2014. ISSN 15280020. doi: 10.1182/blood-2013-11-512764.
- [97] Fillmore Freeman and Christos N Angeletakis. *Il 0 9.*, 73(2):3403–3408, 1982.
- [98] Roger J. S. Preston, Jamie M. O’Sullivan, and James S. O’Donnell. Advances in understanding the molecular mechanisms of venous thrombosis. *British Journal of Haematology*, pages 1–11, 2019. doi: 10.1111/bjh.15869.
- [99] P Stassens, P W Bergum, Y Gansemans, L Jespers, Y Laroche, S Huang, S Maki, J Messens, M Lauwereys, M Cappello, P J Hotez, I Lasters, and G P Vlasuk. Anticoagulant repertoire of the hookworm *Ancylostoma caninum*. *Proceedings of the National Academy of Sciences of the United States of America*, 93(5):2149–2154, 1996. ISSN 00278424. doi: 10.1073/pnas.93.5.2149.
- [100] Sai K. Buddai, Larisa Touloukhonova, Peter W. Bergum, George P. Vlasuk, and Sriram Krishnaswamy. Nematode anticoagulant protein c2 reveals a site on factor Xa that is important for macromolecular substrate binding to human prothrombinase. *Journal of Biological Chemistry*, 277(29):26689–26698, 2002. ISSN 00219258. doi: 10.1074/jbc.M202507200.

- [101] M. T. Murakami, J. Rios-Steiner, S. E. Weaver, A. Tulinsky, J. H. Geiger, and R. K. Arni. Intermolecular Interactions and Characterization of the Novel Factor Xa Exosite Involved in Macromolecular Recognition and Inhibition: Crystal Structure of Human Gla-domainless Factor Xa Complexed with the Anticoagulant Protein NAPc2 from the Hematophagou. *Journal of Molecular Biology*, 366(2):602–610, 2007. ISSN 00222836. doi: 10.1016/j.jmb.2006.11.040.
- [102] Peter W. Bergum, Amy Cruikshank, Steven L. Maki, Curtis R. Kelly, Wolfram Ruf, and George P. Vlasuk. Role of Zymogen and Activated Factor X as Scaffolds for the Inhibition of the Blood Coagulation Factor VIIa-Tissue Factor Complex by Recombinant Nematode Anticoagulant Protein c2. *Journal of Biological Chemistry*, 276(13):10063–10071, 2001. ISSN 00219258. doi: 10.1074/jbc.M009116200.
- [103] Brendan M. Duggan, H. Jane Dyson, and Peter E. Wright. Inherent flexibility in a potent inhibitor of blood coagulation, recombinant nematode anticoagulant protein c2, 1999. ISSN 00142956.
- [104] Timothy H. Bayburt, Yelena V. Grinkova, and Stephen G. Sligar. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Letters*, 2(8):853–856, 2002. ISSN 15306984. doi: 10.1021/nl025623k.
- [105] I. G. Denisov, Y. V. Grinkova, A. A. Lazarides, and S. G. Sligar. Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *Journal of the American Chemical Society*, 126(11):3477–3487, 2004. ISSN 00027863. doi: 10.1021/ja0393574.
- [106] Franz Hagn, Manuel Etzkorn, Thomas Raschle, and Gerhard Wagner. Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins Franz. *Journal of the American Chemical Society*, 135(5):1919–1925, 2014. doi: 10.1021/ja310901f.Optimized.
- [107] Soung Hun Roh, Nicholas J. Stam, Corey F. Hryc, Sergio Couoh-Cardel, Grigore Pintilie, Wah Chiu, and Stephan Wilkens. The 3.5-Å CryoEM Structure of Nanodisc-Reconstituted Yeast Vacuolar ATPase VoProton Channel. *Molecular Cell*, 69(6):993–1004.e3, 2018. ISSN 10974164. doi: 10.1016/j.molcel.2018.02.006.
- [108] Nopnithi Thonghin, Vasileios Kargas, Jack Clews, and Robert C. Ford. Cryo-electron microscopy of membrane proteins. *Methods*, 147(April):176–186, 2018. ISSN 10959130. doi: 10.1016/j.ymeth.2018.04.018.
- [109] Mahmoud L. Nasr, Diego Baptista, Mike Strauss, Zhen Yu J. Sun, Simina Grigoriu, Sonja Huser, Andreas Plückthun, Franz Hagn, Thomas Walz, James M. Hogle, and Gerhard Wagner. Covalently circularized nanodiscs for studying membrane proteins and viral entry. *Nature Methods*, 14(1):49–52, 2016. ISSN 15487105. doi: 10.1038/nmeth.4079.
- [110] Nicolai Tidemand Johansen, Martin Cramer Pedersen, Frederik Grønbaek Tidemand, Kasper Dyrberg Rand, Tam T. T. N. Nguyen, and Lise Arleth. Circularized and solubility-enhanced MSPs facilitate simple and high yield production of stable nanodiscs for stud-

- ies of membrane proteins in solution. *The FEBS Journal*, 2019. ISSN 17424658. doi: 10.1111/febs.14766.
- [111] Vivien Yeh, Tsung Yen Lee, Chung Wen Chen, Pai Chia Kuo, Jessie Shiue, Li Kang Chu, and Tsyrr Yan Yu. Highly Efficient Transfer of 7<sup>TM</sup> Membrane Protein from Native Membrane to Covalently Circularized Nanodisc. *Scientific Reports*, 8(1):1–11, 2018. ISSN 20452322. doi: 10.1038/s41598-018-31925-1.
- [112] Mahmoud L. Nasr and Gerhard Wagner. Covalently circularized nanodiscs; challenges and applications. *Current Opinion in Structural Biology*, 51:129–134, 2018. ISSN 1879033X. doi: 10.1016/j.sbi.2018.03.014.
- [113] Søren A.R. Kynde, Nicholas Skar-Gislinge, Martin Cramer Pedersen, Søren Roi Midtgaard, Jens Baek Simonsen, Ralf Schweins, Kell Mortensen, and Lise Arleth. Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment. *Acta Crystallographica Section D: Biological Crystallography*, 70(2):371–383, 2014. ISSN 09074449. doi: 10.1107/S1399004713028344.
- [114] Nicholas Skar-Gislinge, Søren A.R. Kynde, Ilia G. Denisov, Xin Ye, Ivan Lenov, Stephen G. Sligar, and Lise Arleth. Small-angle scattering determination of the shape and localization of human cytochrome P450 embedded in a phospholipid nanodisc environment. *Acta Crystallographica Section D: Biological Crystallography*, 71(1):2412–2421, 2015. ISSN 13990047. doi: 10.1107/S1399004715018702.
- [115] K Harlos, D M Martin, D P O’Brien, E Y Jones, D I Stuart, I Polikarpov, A Miller, E G Tuddenham, and C W Boys. Crystal structure of the extracellular region of human tissue factor. *Nature*, 371(6499):720–720, 1994. ISSN 0028-0836. doi: 10.1038/371720a0.
- [116] Kaillathe Padmanabhan, K. P. Padmanabhan, A. Tulinsky, Chang H. Park, W. Bode, R. Huber, D. T. Blankenship, A. D. Cardin, and W. Kisiel. Structure of human des(1-45) factor Xa at 2.2 Å resolution, 1993. ISSN 00222836.
- [117] Egon Persson and Ole Hvilsted Olsen. Allosteric activation of coagulation factor VIIa Egon. *Frontiers in Bioscience*, pages 3156–3163, 2011.
- [118] Harvey J. Weiss, William J. Vivic, Bruce A. Lages, and John Rogers. Isolated deficiency of platelet procoagulant activity. *The American Journal of Medicine*, 67(2):206–213, 1979. ISSN 00029343. doi: 10.1016/0002-9343(79)90392-9.
- [119] Elisabetta Castoldi, Peter W. Collins, Patrick L. Williamson, and Edouard M. Bevers. Compound heterozygosity for 2 novel TMEM16F mutations in a patient with Scott syndrome To the editor : Chimerism levels after stem cell transplantation are primarily determined by the ratio of donor. *Blood*, 117(16):4399–4401, 2011.
- [120] K. Khair, M. G. Mazzucconi, R. Parra, E. Santagostino, D. A. Tsakiris, C. Hermans, J. Oldenburg, G. Spotts, K. Steinitz-Trost, and A. Gringeri. Pattern of bleeding in a large prospective cohort of haemophilia A patients: A three-year follow-up of the AHEAD (Advate in HaEmophilia A outcome Database) study. *Haemophilia*, 24(1):85–96, 2018. ISSN 13652516. doi: 10.1111/hae.13361.

- [121] Gerry Dolan, Gary Benson, Anne Duffy, Cedric Hermans, Victor Jiménez-Yuste, Thierry Lambert, Rolf Ljung, Massimo Morfini, and Silva Zupančić Šalek. Haemophilia B: Where are we now and what does the future hold? *Blood Reviews*, 32(1):52–60, 2018. ISSN 15321681. doi: 10.1016/j.blre.2017.08.007.
- [122] Allison P. Wheeler and David Gailani. Why Factor XI Deficiency is a Clinical Concern Allison. *Expert Rev Hemmatol*, 118, 2016. doi: 10.1002/cncr.27633.Percutaneous.
- [123] Huanghe Yang, Andrew Kim, Tovo David, Daniel Palmer, Taihao Jin, Jason Tien, Fen Huang, Tong Cheng, Shaun R. Coughlin, Yuh Nung Jan, and Lily Yeh Jan. TMEM16F Forms a Ca<sup>2+</sup>-Activated Cation Channel Required for Lipid Scrambling in Platelets during Blood Coagulation. *Cell*, 151(1):111–122, 2012. ISSN 19326203. doi: 10.1371/journal.pone.0178059.
- [124] Maria E. Falzone, Mattia Malvezzi, Byoung-Cheol Lee, and Alessio Accardi. Known structures and unknown mechanisms of TMEM16 scramblases and channels. *The Journal of General Physiology*, 150(7):jgp.201711957, 2018. ISSN 0022-1295. doi: 10.1085/jgp.201711957.
- [125] Dung M. Nguyen, Louisa S. Chen, Wei-Ping Yu, and Tsung-Yu Chen. Comparison of ion transport determinants between a TMEM16 chloride channel and phospholipid scramblase. *The Journal of General Physiology*, 151(4):jgp.201812270, 2019. ISSN 0022-1295. doi: 10.1085/jgp.201812270.
- [126] Robert Kliegman and Waldo Nelson. Nelson textbook of pediatrics. In *Nelson textbook of pediatrics*. Elsevier, 2011.
- [127] Vijaya S. Pilli. Understanding the Clotting Cascade, Regulators, and Clinical Modulators of Coagulation, 2018.
- [128] S. He, M. Blombäck, G. Jacobsson Ekman, and U. Hedner. The role of recombinant factor VIIa (FVIIa) in fibrin structure in the absence of FVIII/FIX. *Journal of Thrombosis and Haemostasis*, 1(6):1215–1219, 2003. ISSN 15387933. doi: 10.1046/j.1538-7836.2003.00242.x.
- [129] Marc D. Silverstein, John A. Heit, David N. Mohr, Tanya M. Petterson, W. Michael O’Fallon, and L. Joseph Melton. Trends in the incidence of deep vein thrombosis and pulmonary embolism: A 25-year population-based study. *Archives of Internal Medicine*, 158(6):585–593, 1998. ISSN 00039926. doi: 10.1001/archinte.158.6.585.
- [130] Douglas Kopcke, Ondina Harryman, Emyr W. Benbow, Charles Hay, and Nicholas Chalmers. Mortality from pulmonary embolism is decreasing in hospital patients. *Journal of the Royal Society of Medicine*, 104(8):327–331, 2011. ISSN 01410768. doi: 10.1258/jrsm.2011.100395.
- [131] Kellie R. Machlus, Maria M. Aleman, and Alisa S. Wolberg. Update on Venous Thromboembolism. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 31(3):476–478, 2011. ISSN 1079-5642. doi: 10.1161/atvbaha.111.223008.

- [132] Akihiro Onishi, Kalib St Ange, Jonathan S. Dordick, and Robert J. Lindhardt. Heparin and anticoagulation. *Frontiers in Bioscience*, 21(7):1372–1392, 2016. ISSN 10939946. doi: 10.2741/4462.
- [133] Anne Cornélie J M de Pont, A. H M Moons, E. de Jonge, J. C M Meijers, G. P. Vlasuk, W. E. Rote, H. R. Büller, T. van der Poll, and M. Levi. Recombinant nematode anticoagulant protein c2, an inhibitor of tissue factor/factor VIIa, attenuates coagulation and the interleukin-10 response in human endotoxemia. *Journal of Thrombosis and Haemostasis*, 2(1):65–70, 2004. ISSN 15387933. doi: 10.1111/j.1538-7836.2004.00526.x.
- [134] Xiaoying Chen, Jennica Zaro, and Wei-Chiang Shen. Fusion Protein Linkers: Property, Design and Functionality. *Advanced Drug Delivery Reviews*, 65(10):1357–1369, 2013. doi: 10.1002/9783527678679.dg06857.
- [135] Saul R. Trevino, J. Martin Scholtz, and C. Nick Pace. Amino Acid Contribution to Protein Solubility: Asp, Glu, and Ser Contribute more Favorably than the other Hydrophilic Amino Acids in RNase Sa. *Journal of Molecular Biology*, 366(2):449–460, 2007. ISSN 00222836. doi: 10.1016/j.jmb.2006.10.026.
- [136] Ryan M. Kramer, Varad R. Shende, Nicole Motl, C. Nick Pace, and J. Martin Scholtz. Toward a molecular understanding of protein solubility: Increased negative surface charge correlates with increased solubility. *Biophysical Journal*, 102(8):1907–1915, 2012. ISSN 00063495. doi: 10.1016/j.bpj.2012.01.060.
- [137] Tomohiro Imura, Yohei Tsukui, Kenichi Sakai, Hideki Sakai, Toshiaki Taira, and Dai Kitamoto. Minimum amino acid residues of an  $\alpha$ -helical peptide leading to lipid nanodisc formation. *Journal of Oleo Science*, 63(11):1203–1208, 2014. ISSN 13473352. doi: 10.5650/jos.ess14172.
- [138] Joseph E Tropea, Scott Cherry, and David S Waugh. Expression and Purification of Soluble His6-Tagged TEV Protease BT - High Throughput Protein Expression and Purification: Methods and Protocols. pages 297–307. Humana Press, Totowa, NJ, 2009. ISBN 978-1-59745-196-3. doi: 10.1007/978-1-59745-196-3\_19.
- [139] Yu Tong, Jun Yue, Meng Mao, Qingqing Liu, Jing Zhou, and Jiyun Yang. Recombinant nematode anticoagulant protein c2 inhibits cell invasion by decreasing uPA expression in NSCLC cells. *Oncology Reports*, 33(4):1815–1822, 2015. ISSN 17912431. doi: 10.3892/or.2015.3795.
- [140] Lisa R. Paborsky, Keri M. Tate, Reed J. Harris, Daniel G. Yansura, Louise Band, Glynis McCray, Cornelia M. Gorman, Donogh P. O’Brien, Judy Y. Chang, James R. Swartz, Victor P. Fung, James N. Thomas, and Gordon A. Vehar. Purification of recombinant human tissue factor. *Biochemistry*, 28(20):8072–8077, 1989. ISSN 15204995. doi: 10.1021/bi00446a016.
- [141] Jean Luc Popot. Folding membrane proteins in vitro: A table and some comments. *Archives of Biochemistry and Biophysics*, 564:314–326, 2014. ISSN 10960384. doi: 10.1016/j.abb.2014.06.029.

- [142] Søren Roi Midtgaard, Martin Cramer Pedersen, Jacob Judas Kain Kirkensgaard, Kasper Kildegaard Sørensen, Kell Mortensen, Knud J. Jensen, and Lise Arleth. Self-assembling peptides form nanodiscs that stabilize membrane proteins. *Soft Matter*, 10(5): 738–752, 2014. ISSN 1744683X. doi: 10.1039/c3sm51727f.
- [143] D G Dervichian, G Fournet, and A Guinier. Mise en évidence d’une structure submicroscopique dans les globules rouges par la diffusion des rayons X aux petits angles. *Comptes rendus hebdomadaires des seances de l’Academie des sciences*, 224(26):1848–1850, jun 1947. ISSN 0001-4036.
- [144] Des McMorrow and Jens Als Nielsen. *Elements of Modern X-ray Physics*. John Wiley And Sons Ltd, 2011.
- [145] Feiginm L. A. and Dmitri I. Svergun. *Structure analysis by small-angle X-ray and neutron scattering*. Plenum Press, 1987. ISBN 0306426293 9780306426292.
- [146] O. Glatter. A new method for the evaluation of small-angle scattering data. *Journal of Applied Crystallography*, 10(5):415–421, 1977. ISSN 00218898. doi: 10.1107/S0021889877013879.
- [147] DI Svergun. Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophysical journal*, 76(June):2879–2886, 1999.
- [148] Varley F. Sears. Neutron Scattering Lengths and Cross Sections. *Neutron News*, 3(2), 1992. ISSN 10794042. doi: 10.1016/B978-0-12-398374-9.09989-7.
- [149] Frank Gabel. *Small-angle neutron scattering for structural biology of protein-RNA complexes*, volume 558. Elsevier Inc., 1 edition, 2015. doi: 10.1016/bs.mie.2015.02.003.
- [150] Marta Comellas-Aragonès, Friso D. Sikkema, Guillaume Delaittre, Ann E. Terry, Stephen M. King, Dirk Visser, Richard K. Heenan, Roeland J.M. Nolte, Jeroen J.L.M. Cornelissen, and Martin C. Feiters. Solution scattering studies on a virus capsid protein as a building block for nanoscale assemblies. *Soft Matter*, 7(24):11380–11391, 2011. ISSN 1744683X. doi: 10.1039/c1sm06123b.
- [151] L. He, A. Piper, F. Meilleur, R. Hernandez, W. T. Heller, and D. T. Brown. Conformational Changes in Sindbis Virus Induced by Decreased pH Are Revealed by Small-Angle Neutron Scattering. *Journal of Virology*, 86(4):1982–1987, 2012. ISSN 0022-538X. doi: 10.1128/jvi.06569-11.
- [152] Nicholas Skar-Gislinge and Lise Arleth. Small-angle scattering from phospholipid nanodiscs: Derivation and refinement of a molecular constrained analytical model form factor. *Physical Chemistry Chemical Physics*, 13(8):3161–3170, 2011. ISSN 14639076. doi: 10.1039/c0cp01074j.
- [153] Søren Roi Midtgaard, Tamim A. Darwish, Martin Cramer Pedersen, Pie Huda, Andreas Haahr Larsen, Grethe Vestergaard Jensen, Søren Andreas Røssell Kynde, Nicholas Skar-Gislinge, Agnieszka Janina Zygadlo Nielsen, Claus Olesen, Mickael Blaise, Jerzy Józef

- Dorosz, Thor Seneca Thorsen, Raminta Venskutonytė, Christian Krintel, Jesper V. Møller, Henrich Frielinghaus, Elliot Paul Gilbert, Anne Martel, Jette Sandholm Kastrop, Poul Erik Jensen, Poul Nissen, and Lise Arleth. Invisible detergents for structure determination of membrane proteins by small-angle neutron scattering. *FEBS Journal*, 285(2): 357–371, 2018. ISSN 17424658. doi: 10.1111/febs.14345.
- [154] Selma Maric, Nicholas Skar-Gislinge, Søren Midtgaard, Mikkel B. Thygesen, Jürgen Schiller, Henrich Frielinghaus, Martine Moulin, Michael Haertlein, V. Trevor Forsyth, Thomas Günther Pomorski, and Lise Arleth. Stealth carriers for low-resolution structure determination of membrane proteins in solution. *Acta Crystallographica Section D: Biological Crystallography*, 70(2):317–328, 2014. ISSN 09074449. doi: 10.1107/S1399004713027466.
- [155] Inokentij Josts, Julius Nitsche, Selma Maric, Haydyn D. Mertens, Martine Moulin, Michael Haertlein, Sylvain Prevost, Dmitri I. Svergun, Sebastian Busch, V. Trevor Forsyth, and Henning Tidow. Conformational States of ABC Transporter MsbA in a Lipid Environment Investigated by Small-Angle Scattering Using Stealth Carrier Nanodiscs. *Structure*, 26(8):1–8, 2018. ISSN 18784186. doi: 10.1016/j.str.2018.05.007.
- [156] Jesse B. Hopkins and Robert E. Thorne. Quantifying radiation damage in biomolecular small-angle X-ray scattering. *Journal of Applied Crystallography*, 49(3):880–890, 2016. doi: 10.1107/s1600576716005136.
- [157] Petra Pernot, Adam Round, Ray Barrett, Alejandro De Maria Antolinos, Alexandre Gobbo, Elspeth Gordon, Julien Huet, Jérôme Kieffer, Mario Lentini, Muriel Mattenet, Christian Morawe, Christoph Mueller-Dieckmann, Staffan Ohlsson, Werner Schmid, John Surr, Pascal Theveneau, Louiza Zerrad, and Sean McSweeney. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *Journal of Synchrotron Radiation*, 20(4):660–664, 2013. ISSN 0909-0495. doi: 10.1107/s0909049513010431.
- [158] Ashley Jordan, Mark Jacques, Catherine Merrick, Juliette Devos, V. Trevor Forsyth, Lionel Porcar, and Anne Martel. SEC-SANS: Size exclusion chromatography combined in situ with small-angle neutron scattering. *Journal of Applied Crystallography*, 49(6):2015–2020, 2016. ISSN 16005767. doi: 10.1107/S1600576716016514.
- [159] Clement E. Blanchet, Alessandro Spilotros, Frank Schwemmer, Melissa A. Graewert, Alexey Kikhney, Cy M. Jeffries, Daniel Franke, Daniel Mark, Roland Zengerle, Florent Cipriani, Stefan Fiedler, Manfred Roessle, and Dmitri I. Svergun. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *Journal of Applied Crystallography*, 48(2):431–443, 2015. doi: 10.1107/s160057671500254x.
- [160] Doris Orthaber, Alexander Bergmann, and Otto Glatter. SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *Journal of Applied Crystallography*, 33(2):218–225, 2000. ISSN 0021-8898. doi: 10.1107/s0021889899015216.
- [161] Efstratios Mylonas and Dmitri I Svergun. Accuracy of molecular mass determination of

- proteins in solution by small-angle X-ray scattering. *Journal of Applied Crystallography*, 40(s1):s245—s249, apr 2007. doi: 10.1107/S002188980700252X.
- [162] Dmitri I. Svergun, S Richard, M H J Koch, Z Sayers, S Kuprin, and G Zaccai. Protein hydration in solution : Experimental observation by x-ray. *Proc. Natl. Acad. Sci. USA*, 95(March):2267–2272, 1998.
- [163] D Svergun, C Barberato, and M H J Koch. CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography*, 28(6):768–773, dec 1995. doi: 10.1107/S0021889895007047.
- [164] Maxim V. Petoukhov and Dmitri I. Svergun. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophysical Journal*, 89(2):1237–1250, 2005. ISSN 00063495. doi: 10.1529/biophysj.105.064154.
- [165] Tone Bengtsen, Viktor L Holm, Søren R Midtgaard, Nicolai Tidemand, Sandro Bottaro, Lise Arleth, and Kresten Lindorff-larsen. Structure and dynamics of a lipid nanodisc by integrating NMR , SAXS and SANS experiments with molecular dynamics simulations. *bioRxiv*, pages 1–33, 2019.
- [166] Kyle T Powers, Melissa S Gildenberg, and M Todd Washington. Modeling Conformationally Flexible Proteins With X-ray Scattering and Molecular Simulations. *Computational and Structural Biotechnology Journal*, 17:570–578, 2019. ISSN 2001-0370. doi: 10.1016/j.csbj.2019.04.011.
- [167] Daniel Franke and Dmitri I. Svergun. DAMMIF , a program for rapid ab-initio shape determination in small-angle scattering . *Journal of Applied Crystallography*, 42(2):342–346, 2009. ISSN 0021-8898. doi: 10.1107/s0021889809000338.
- [168] Nicholas Skar-Gislunge, Jens Bæk Simonsen, Kell Mortensen, Stephen G Sligar, Birger Lindberg Møller, and Thomas Bjørnholm. Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins : Casting the Roles of the Lipids and the Protein Preparation of Nanodiscs :. *Journal of Applied Crystallography*, (17):1–4, 2010.
- [169] Martin Cramer Pedersen, Lise Arleth, and Kell Mortensen. WillItFit: A framework for fitting of constrained models to small-angle scattering data. *Journal of Applied Crystallography*, 46(6):1894–1898, 2013. ISSN 00218898. doi: 10.1107/S0021889813026022.
- [170] Ralf P. Richter, Rémi Bérat, and Alain R. Brisson. Formation of solid-supported lipid bilayers: An integrated view. *Langmuir*, 22(8):3497–3505, 2006. ISSN 07437463. doi: 10.1021/la052687c.
- [171] Jakob Andersson, Ingo Köper, and Wolfgang Knoll. Tethered Membrane Architectures—Design and Applications. *Frontiers in Materials*, 5(September):1–11, 2018. doi: 10.3389/fmats.2018.00055.
- [172] James Kurniawan, Joao Francisco Ventrici De Souza, Amanda T. Dang, Gang Yu Liu, and Tonya L. Kuhl. Preparation and Characterization of Solid-Supported Lipid Bilayers

- Formed by Langmuir-Blodgett Deposition: A Tutorial. *Langmuir*, 34(51):15622–15639, 2018. ISSN 15205827. doi: 10.1021/acs.langmuir.8b03504.
- [173] C. A. Keller, K. Glasmästar, V. P. Zhdanov, and B. Kasemo. Formation of Supported Membranes from Vesicles. *Physical Review Letters*, 84(23):5443–5446, 2000. ISSN 00319007. doi: 10.1103/PhysRevLett.84.5443.
- [174] Tania Kjellerup Lind, Marité Cárdenas, and Hanna Pauliina Wacklin. Formation of supported lipid bilayers by vesicle fusion: Effect of deposition temperature. *Langmuir*, 30(25):7259–7263, 2014. ISSN 15205827. doi: 10.1021/la500897x.
- [175] Hudson P. Pace, Jonas K. Hannestad, Antonious Armonious, Marco Adamo, Bjorn Agnarsson, Anders Gunnarsson, Samantha Micciulla, Peter Sjövall, Yuri Gerelli, and Fredrik Höök. Structure and Composition of Native Membrane Derived Polymer-Supported Lipid Bilayers. *Analytical Chemistry*, 90(21):13065–13072, 2018. ISSN 15206882. doi: 10.1021/acs.analchem.8b04110.
- [176] Ralf P. Richter and Alain R. Brisson. Following the formation of supported lipid bilayers on Mica: A study combining AFM, QCM-D, and ellipsometry. *Biophysical Journal*, 88(5):3422–3433, 2005. ISSN 00063495. doi: 10.1529/biophysj.104.053728.
- [177] Volker Kiessling, Sung Tae Yang, and Lukas K. Tamm. *Supported Lipid Bilayers as Models for Studying Membrane Domains*, volume 75. Elsevier Ltd, 2015. doi: 10.1016/bs.ctm.2015.03.001.
- [178] Alessandra Luchini, Achebe N.O. Nzulumike, Tania K. Lind, Tommy Nylander, Robert Barker, Lise Arleth, Kell Mortensen, and Marité Cárdenas. Towards biomimics of cell membranes: Structural effect of phosphatidylinositol triphosphate (PIP<sub>3</sub>) on a lipid bilayer. *Colloids and Surfaces B: Biointerfaces*, 173(March 2018):202–209, 2019. ISSN 18734367. doi: 10.1016/j.colsurfb.2018.09.031.
- [179] Valeria Rondelli, Paola Brocca, Simona Motta, Massimo Messa, Laura Colombo, Mario Salmona, Giovanna Fragneto, Laura Cantu, and Elena Del Favero. Amyloid $\beta$  Peptides in interaction with raft-mimic model membranes: A neutron reflectivity insight. *Scientific Reports*, 6:1–11, 2016. ISSN 20452322. doi: 10.1038/srep20997.
- [180] Giuseppe Vitiello, Annarita Falanga, Ariel Alcides Petruk, Antonello Merlino, Giovanna Fragneto, Luigi Paduano, Stefania Galdiero, and Gerardino D’Errico. Fusion of raft-like lipid bilayers operated by a membranotropic domain of the HSV-type i glycoprotein gH occurs through a cholesterol-dependent mechanism. *Soft Matter*, 11(15):3003–3016, 2015. ISSN 17446848. doi: 10.1039/c4sm02769h.
- [181] Annette Granéli, Jan Rydström, Bengt Kasemo, and Fredrik Höök. Formation of supported lipid bilayer membranes on SiO<sub>2</sub> from proteoliposomes containing transmembrane proteins. *Langmuir*, 19(3):842–850, 2003. ISSN 07437463. doi: 10.1021/la026231w.
- [182] Frank Giess, Marcel G. Friedrich, Joachim Heberle, Renate L. Naumann, and Wolfgang Knoll. The protein-tethered lipid bilayer: A novel mimic of the biological membrane.

- Biophysical Journal*, 87(5):3213–3220, 2004. ISSN 00063495. doi: 10.1529/biophysj.104.046169.
- [183] Robert Cubitt and Giovanna Fragneto. Chapter 2.8.3 – Neutron Reflection:: Principles and Examples of Applications. *Scattering*, pages 1198–1208, 2002. doi: 10.1016/B978-012613760-6/50065-6.
- [184] G. Fragneto-Cusani. Neutron reflectivity at the solid/liquid interface: Examples of applications in biophysics. *Journal of Physics Condensed Matter*, 13(21):4973–4989, 2001. ISSN 09538984. doi: 10.1088/0953-8984/13/21/322.
- [185] Jean Daillant and Alain Gibaud. X-ray and Neutron Reflectivity: Principles and Applications (Lecture Notes in Physics). In *Springer*, page 348. Springer, 2008. ISBN 3540885870.
- [186] Xiao Lin Zhou and Sow Hsin Chen. Theoretical foundation of X-ray and neutron reflectometry. *Physics Reports*, 257(4-5):223–348, 1995. ISSN 03701573. doi: 10.1016/0370-1573(94)00110-O.
- [187] Max Born and Emil Wolf. *Principles of Optics*. Cambridge University Press, 1989. doi: 10.1080/713826373.
- [188] DG Bucknall, J Penfold, JRP Webster, A Zarbakhsh, RM Richardson, A Rennie, JS Higgins, RAL Jones, P Fletcher, RK Thomas, S Roser, and E Dickinson. SURF - A Second Generation Neutron Reflectometer. Technical report, 1995.
- [189] T. Charitat, E. Bellet-Amalric, G. Fragneto, and F. Graner. Adsorbed and free lipid bilayers at the solid-liquid interface. *European Physical Journal B*, 8(4):583–593, 1999. ISSN 14346028. doi: 10.1007/s100510050725.
- [190] David P Hoogerheide, Sergei Yu Noskov, Adam J Kuszak, Susan K Buchanan, Tatiana K Rostovtseva, and Hirsh Nanda. Structure of voltage-dependent anion channel-tethered bilayer lipid membranes determined using neutron reflectivity. *Acta crystallographica. Section D, Structural biology*, 74(Pt 12):1219–1232, dec 2018. ISSN 2059-7983 (Electronic). doi: 10.1107/S2059798318011749.
- [191] Thomas Soranzo, Donald K. Martin, Jean Luc Lenormand, and Erik B. Watkins. Coupling neutron reflectivity with cell-free protein synthesis to probe membrane protein structure in supported bilayers. *Scientific Reports*, 7(1):1–10, 2017. ISSN 20452322. doi: 10.1038/s41598-017-03472-8.
- [192] Hsin-hui Shen, Denisse L Leyton, Takuya Shiota, Matthew J Belousoff, Jingxiong Lu, Stephen A Holt, Khershing Tan, Joel Selkrig, Chaille T Webb, Susan K Buchanan, Lisan-dra L Martin, and Trevor Lithgow. Reconstitution of a nanomachine driving the assembly of proteins into bacterial outer membranes. *Nature Communications*, 5, 2015. doi: 10.1038/ncomms6078.Reconstitution.
- [193] Reg Rehal, Piers R.J. Gaffney, Alasdair T.M. Hubbard, Robert D. Barker, and Richard D. Harvey. The pH-dependence of lipid-mediated antimicrobial peptide resistance in a model

- staphylococcal plasma membrane: A two-for-one mechanism of epithelial defence circumvention. *European Journal of Pharmaceutical Sciences*, 128(November 2018):43–53, 2019. ISSN 18790720. doi: 10.1016/j.ejps.2018.11.017.
- [194] Meng Hsuan Hsieh, Po Tsang Huang, Horng Huei Liou, Po Huang Liang, Pei Ming Chen, Stephen A. Holt, Isaac Furay Yu, Michael James, Yu Shuan Shiau, Ming Tao Lee, Tsang Lang Lin, and Kuo Long Lou. The Penetration Depth for Hanatoxin Partitioning into the Membrane Hydrocarbon Core Measured with Neutron Reflectivity. *Langmuir*, 34(30):9036–9046, 2018. ISSN 15205827. doi: 10.1021/acs.langmuir.8b01076.
- [195] Arwel V. Hughes, Dhilon S. Patel, Göran Widmalm, Jeffery B. Klauda, Luke A. Clifton, and Wonpil Im. Physical Properties of Bacterial Outer Membrane Models: Neutron Reflectometry & Molecular Simulation. *Biophysical Journal*, 116(6):1095–1104, 2019. ISSN 15420086. doi: 10.1016/j.bpj.2019.02.001.
- [196] Frank Heinrich and Mathias Lösche. Zooming in on disordered systems: Neutron reflection studies of proteins associated with fluid membranes. *Biochimica et biophysica acta*, 1838(9):2341–2349, 2014. doi: 10.1016/j.bbamem.2014.03.007.Zooming.
- [197] Candace M. Pfefferkorn, Frank Heinrich, Alexander J. Sodt, Alexander S. Maltsev, Richard W. Pastor, and Jennifer C. Lee. Depth of  $\alpha$ -synuclein in a bilayer determined by fluorescence, neutron reflectometry, and computation. *Biophysical Journal*, 102(3):613–621, 2012. ISSN 00063495. doi: 10.1016/j.bpj.2011.12.051.
- [198] Brilliant Adhi Prabowo, Agnes Purwidyantri, and Kou Chen Liu. Surface Plasmon Resonance Optical Sensor: A Review on Light Source Technology. *Biosensors*, 8(3), 2018. ISSN 20796374. doi: 10.3390/bios8030080.
- [199] Chanjuan Liu, Feichi Hu, Wei Yang, Jiyong Xu, and Yi Chen. A critical review of advances in surface plasmon resonance imaging sensitivity. *TrAC - Trends in Analytical Chemistry*, 97:354–362, 2017. ISSN 18793142. doi: 10.1016/j.trac.2017.10.001.
- [200] Hoang Hiep Nguyen, Jeho Park, Sebyung Kang, and Moonil Kim. Surface plasmon resonance: A versatile technique for biosensor applications. *Sensors (Switzerland)*, 15(5):10481–10510, 2015. ISSN 14248220. doi: 10.3390/s150510481.
- [201] Yustina Yusuf, Julien Massiot, Yu Ting Chang, Pei Hao Wu, Vivien Yeh, Pai Chia Kuo, Jessie Shiue, and Tsyr Yan Yu. Optimization of the Production of Covalently Circularized Nanodiscs and Their Characterization in Physiological Conditions. *Langmuir*, 34(11):3525–3532, 2018. ISSN 15205827. doi: 10.1021/acs.langmuir.8b00025.
- [202] Jonas Miehling, David Goricanec, and Franz Hagn. A Split-Intein-Based Method for the Efficient Production of Circularized Nanodiscs for Structural Studies of Membrane Proteins. *ChemBioChem*, 19(18):1927–1933, 2018. ISSN 14397633. doi: 10.1002/cbic.201800345.
- [203] Xinying Jia, Theo Crawford, Alan H. Zhang, and Mehdi Mobli. A new vector coupling ligation-independent cloning with sortase a fusion for efficient cloning and one-step purifi-

- cation of tag-free recombinant proteins. *Protein Expression and Purification*, 161(April): 1–7, 2019. ISSN 10465928. doi: 10.1016/j.pep.2019.04.004.
- [204] Hongyuan Mao. A self-cleavable sortase fusion for one-step purification of free recombinant proteins. *Protein Expression and Purification*, 37(1):253–263, 2004. ISSN 10465928. doi: 10.1016/j.pep.2004.06.013.
- [205] Robert Warden-Rothman, Ilaria Caturegli, Vladimir Popik, and Andrew Tsourkas. Sortase-tag expressed protein ligation: Combining protein purification and site-specific bioconjugation into a single step. *Analytical Chemistry*, 85(22):11090–11097, 2013. ISSN 00032700. doi: 10.1021/ac402871k.
- [206] Udayar Ilangovan, Hung Ton-that, Junji Iwahara, Olaf Schneewind, and Robert T Clubb. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*.pdf. *Proceedings of the National Academy of Sciences*, 2001.
- [207] Andreas N. Larsen, Kasper K. Sørensen, Nicolai T. Johansen, Anne Martel, Jacob J.K. Kirkensgaard, Knud J. Jensen, Lise Arleth, and Søren Roi Midtgaard. Dimeric peptides with three different linkers self-assemble with phospholipids to form peptide nanodiscs that stabilize membrane proteins. *Soft Matter*, 12(27):5937–5949, 2016. ISSN 17446848. doi: 10.1039/c6sm00495d.
- [208] Y. Sun, X. Du, W. Zhang, L. Sun, and R. Li. Seed germination and physiological characteristics of *amaranthus mangostanus* L. under drought stress. *Advanced Materials Research*, 183-185:1–7, 2011. ISSN 0002-7863. doi: 10.1021/ja310901f.Optimized.
- [209] Melanie P. Muller, Yan Wang, James H. Morrissey, and Emad Tajkhorshid. Lipid Specificity of the Membrane Binding Domain of Coagulation Factor X. *Journal of Thrombosis and Haemostasis*, 15(10):2005–2016, 2017. ISSN 1527-5418. doi: 110.1016/j.bbi.2017.04.008.
- [210] Chang Jun Lee, Vasu Chandrasekaran, Sangwook Wu, Robert E. Duke, and Lee G. Pedersen. Recent Estimates of the Structure of the Factor VIIa (FVIIa)/Tissue Factor (TF) and Factor Xa (FXa) Ternary Complex. *Thrombosis Research*, 125(SUPPL. 1):S7–S10, 2010. ISSN 00493848. doi: 10.1016/j.thromres.2010.01.022.
- [211] Sandro Bottaro, Tone Bengtsen, and Kresten Lindorff-Larsen. Integrating Molecular Simulation and Experimental Data: A Bayesian/Maximum Entropy Reweighting Approach. *bioRxiv*, page 457952, 2018. doi: 10.1101/457952.

## Part III

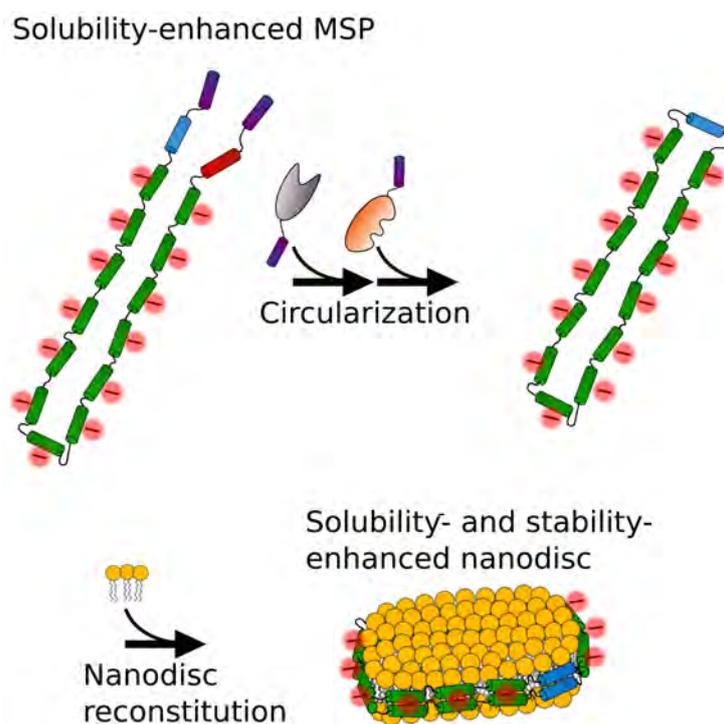
# Manuscripts



## 5 | Paper I

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

*Nicolai T. Johansen, Frederik G. Tidemand, Tam T. T. N. Nguyen, Kasper D. Rand, Martin C. Pedersen, and Lise Arleth*



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

Nicolai Tidemand Johansen<sup>1</sup>, Frederik Grønbæk Tidemand<sup>1</sup>, Tam T. T. N. Nguyen<sup>2</sup>, Kasper Dyrberg Rand<sup>2</sup>, Martin Cramer Pedersen<sup>1</sup> and Lise Arleth<sup>1</sup>

<sup>1</sup> Structural Biophysics, Niels Bohr Institute, University of Copenhagen, Denmark

<sup>2</sup> Protein Analysis Group, Department of Pharmacy, University of Copenhagen, Denmark

## Keywords

membrane protein; nanodisc; protein circularization; small-angle X-ray scattering; thermal and temporal stability

## Correspondence

L. Arleth, Structural Biophysics, Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark.  
E-mail: arleth@nbi.ku.dk

Nicolai Tidemand Johansen and Frederik Grønbæk Tidemand are equally contributing authors

(Received 21 September 2018, revised 10 December 2018, accepted 22 January 2019)

doi:10.1111/febs.14766

Recently, an enzymatic reaction was utilized to covalently link the N and C termini of membrane scaffold proteins to produce circularized nanodiscs that were 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, we designed an optimized membrane scaffold protein (HisMSP1E3D1) with a sortase recognition motif and high abundance of solubility-enhancing negative charges. With these modifications, we show that high protein expression is maintained and that the circularization reaction is efficient, such that we obtain a high yield of circularized membrane scaffold protein (csMSP1E3D1) and downstream circularized nanodiscs. We characterize the circularized protein and corresponding nanodiscs biophysically by small-angle X-ray scattering, size-exclusion chromatography, circular dichroism spectroscopy, and light scattering and compare to noncircularized samples. First, we show that circularized and noncircularized (lsMSP1E3D1) nanodiscs are structurally similar and have the expected nanodisc structure. Second, we show that lsMSP1E3D1 nanodiscs are more stable compared to the template MSP1E3D1 nanodiscs as an effect of the extra negative charges and that csMSP1E3D1 nanodiscs have further improved stability as an effect of circularization. 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 better access to the nanodisc technology and enable applications at physiologically relevant temperatures.

## Introduction

Nanodiscs are discoidal particles composed of a lipid bilayer surrounded by two amphipathic and helical membrane scaffold proteins (MSP) [1–4]. Membrane proteins can be incorporated in nanodiscs, allowing for structural and functional studies [3] and their size

may be modulated [5] to accommodate membrane proteins of different sizes. Compared to other popular membrane protein carrier systems, such as detergent micelles [6], vesicles [7], or bicelles [8], nanodiscs provide a lipid bilayer environment in a relatively small

## Abbreviations

DMPC, 1,2-dimyristoyl-sn-glycero-3-phosphocholine; POPC, 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine.

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, for example, nuclear magnetic resonance (NMR) [2,9], small-angle X-ray and neutron scattering (SAXS and SANS) [10,11], or electron microscopy (EM) techniques including cryo-EM [12,13]. 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 noncircularized 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) [14].

The original MSP1D1 and MSP1D1E3 nanodiscs [1,5] 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 (cNDs) [15], where an evolved variant of the enzyme sortase A (eSrt) from *S. aureus* [16] was utilized to covalently link the N and C termini of the MSP. The main improvements reported on cNDs were a narrower size distribution, that is, reduced polydispersity, increased stability, and better control of the number of membrane proteins incorporated per nanodisc compared to noncircularized nanodiscs [15]. 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 [17]. However, the final yield was still significantly lower than for noncircularized MSPs with an estimated 12.5 mg of cMSP1E3D1 per L culture [17]. 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 *Escherichia coli* host [18]. 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 [15] are all based on the original MSPs [5] 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 being highly expressing homogeneous membrane protein carriers. This leaves room for sequence optimization toward 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 [19,20]. The obvious substitutions are Asn to Asp and Gln to Glu, where the lengths 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 [5]. Furthermore, elevated temperatures facilitate the formation of other types of lipid-MSP particles than the nanodisc [21], for instance by fusion of nanodiscs. These observations 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 [4].

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 toward this goal as it improved the nanodisc stability [15,17,18]. In the present work, we provide a protocol for a large-scale preparation (75 mg per L of bacterial culture) of a novel MSP construct based on MSP1E3D1 [5], which is sequence optimized 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 noncircularized nanodiscs (lsNDs), which are again more stable than standard MSP1E3D1 nanodiscs, 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.

## Results and Discussion

### 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 [5,22]. This template was added a sortase recognition motif, LPETG, in the C terminus to make it compatible with circularization along with an additional His<sub>6</sub>-tag for purification following the work of Nasr and colleagues [15]. Furthermore, solubility-enhancing mutations were made at all possible positions, that is, Asn to Asp, Gln to Glu, and Thr to Ser, respectively [19,20]. We name this construct His-tagged-linear-soluble-MSP1E3D1, in short His-lsMSP1E3D1. A sequence alignment of MSP1E3D1 and His-lsMSP1E3D1 is shown in Fig. 1A with a graphical representation of important sequence elements described in the legend of Fig. 1B. Figure 1B shows the different steps of the circularization of His-lsMSP1E3D1, that is, digestion by tobacco etch virus (TEV) protease to remove the N-terminal His<sub>6</sub>-tag and yield a free N-terminal Gly followed by circularization by eSrt and reverse IMAC purification.

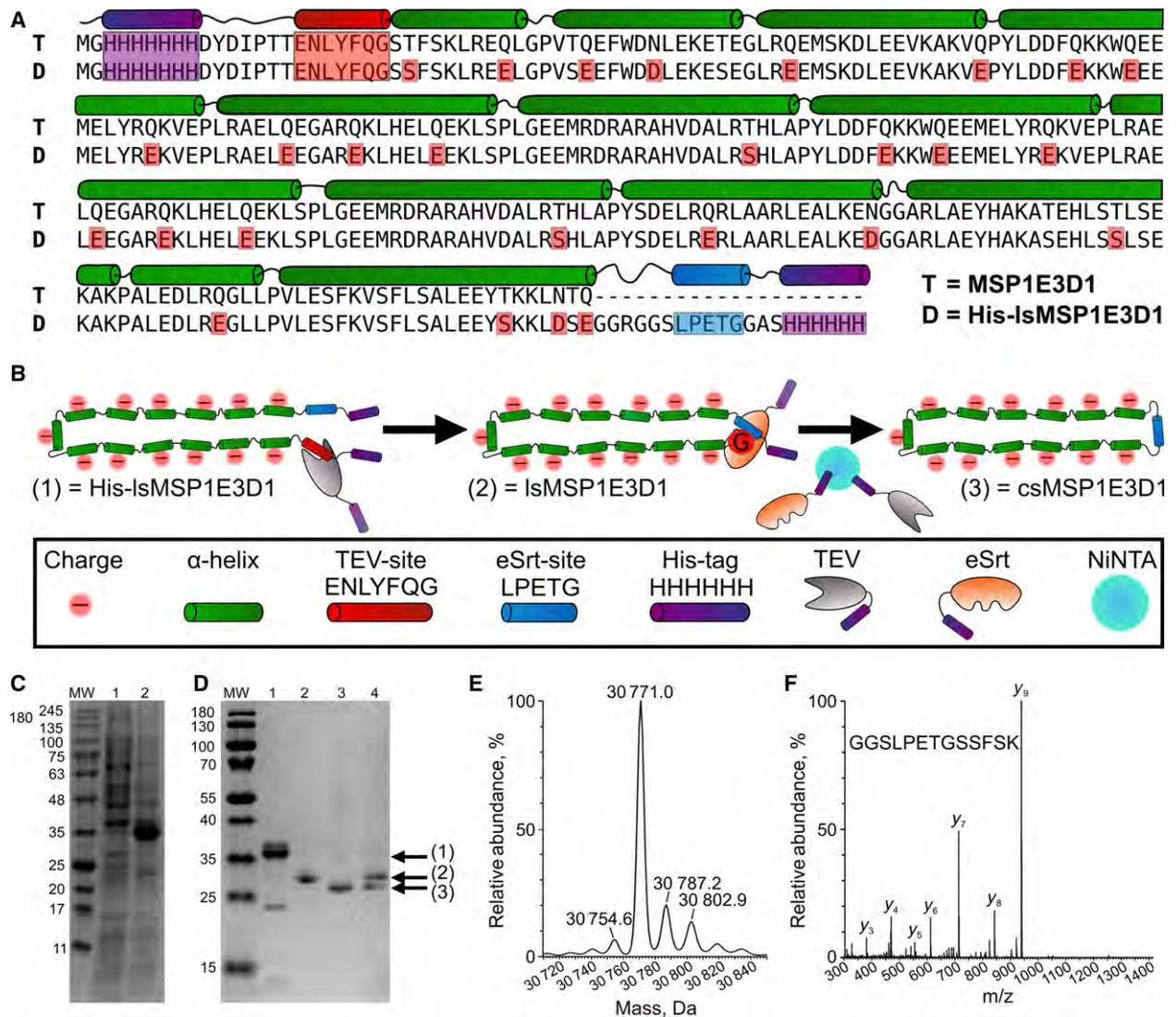
Membrane scaffold proteins have proven to be highly overexpressed in *E. coli*, with MSP1D1 as an example expressing at 250 mg protein per L culture in a fermenter [1], and in our experience around 150 mg per L culture in shaker flasks, which usually yields

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  $\beta$ -D-1-thiogalactopyranoside (IPTG) (Fig. 1C). IMAC purification under denaturing conditions was highly effective, leaving only few impurities and yielding an estimated 100 mg His-lsMSP1E3D1 (Fig. 1B, '(1)' and Fig. 1D, lane 1) per L culture. Subsequent proteolytical digestion of the N terminus with TEV protease yielded lsMSP1E3D1 (Fig. 1B, '(2)' and Fig. 1D lane 2), the substrate for eSrt in the circularization reaction, with undetectable sample loss.

### Circularization efficiency

After IMAC purification, the concentration of lsMSP1E3D1 was approximately 166  $\mu$ M, a factor of 16.6 higher than the concentration used by others for the circularization reaction in order to avoid oligomers by intermolecular circularization [15]. Circularized MSP can be visualized by SDS/PAGE, where it migrates slightly longer than the noncircularized MSP [15]. A series of small reactions with varying concentrations of lsMSP1E3D1 confirmed that a concentration of 10  $\mu$ M was also optimal for our construct with a yield of 95% monomeric circular-soluble-MSP1E3D1 (csMSP1E3D1) (Fig. 2A,C). Although 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 (Fig. 2B,D). 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 (Fig. 1D), eSrt was added to a ratio of 20 : 1 lsMSP1E3D1:eSrt at a lsMSP1E3D1 concentration of 10  $\mu$ M. The formation of csMSP1E3D1 (Fig. 1B, '(3)' and Fig. 1D, lane 3) was verified by SDS/PAGE, observing a small but significant gel shift compared to lsMSP1E3D1 (Fig. 1B, '(2)' and Fig. 1D, lane 2). The significance of the shift was confirmed by loading both lsMSP1E3D1 and csMSP1E3D1 in the same well, which yielded two clearly distinguishable bands (Fig. 1D, lane 4). We note that the samples in Fig. 1D lanes 2–4 were loaded in small amounts to better



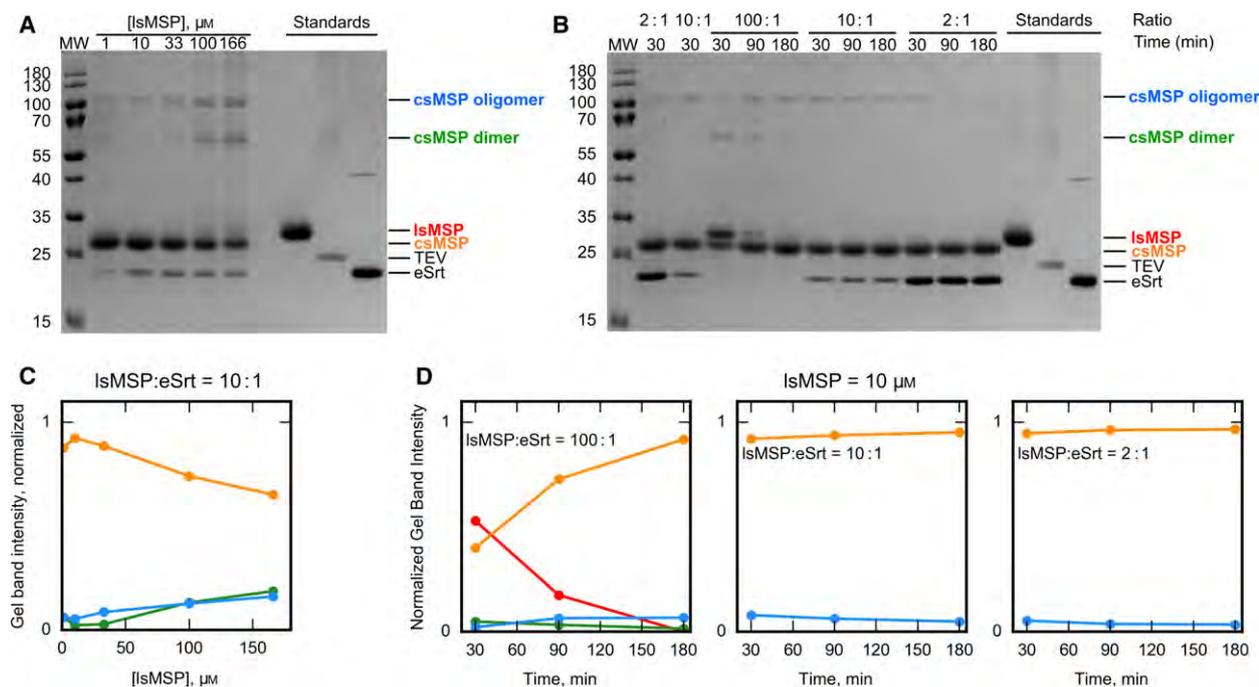
**Fig. 1.** Overview of sequence optimized MSP and its purification. (A) Primary sequence alignment of the template MSP1E3D1 (**T**) and the designed His-IsMSP1E3D1 (**D**), including graphical and highlighted sequence elements. (B) Schematic illustrating the sequence elements of His-IsMSP1E3D1 and the circularization process. (C) SDS/PAGE of the expression of His-IsMSP1E3D1. 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-IsMSP1E3D1, 2 = IsMSP1E3D1, 3 = csMSP1E3D1, 4 = IsMSP1E3D1 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 Figs S2 and S3). (F) CID MS/MS fragmentation spectrum of a tryptic peptide corresponding to the peptide segment linking the N and C terminus of csMSP1E3D1. The precursor ion was  $[M + 2H]^+$  ( $m/z$  627.32).

visualize the mobility shift of IsMSP1E3D1 compared to csMSP1E3D1, which made TEV protease, eSrt, oligomeric csMSP1E3D1, and eventual impurities undetectable.

#### Purity and MS analysis

After circularization and reverse IMAC purification, csMSP1E3D1 was diluted 2.5-fold in Milli-Q water to

a conductivity of approximately  $8 \text{ mS}\cdot\text{cm}^{-1}$  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 (Fig. S1). However, it was also evident that monomeric csMSP1E3D1 was present in both populations, and thus, the best yield and highest concentration was achieved in a step gradient eluting both populations in



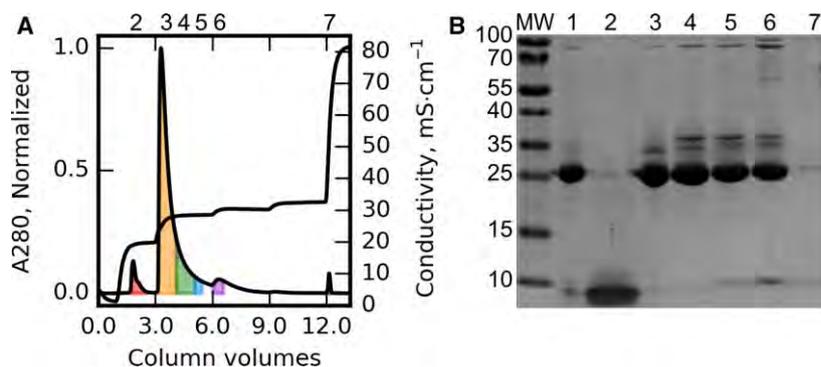
**Fig. 2.** Circularization conditions. (A) SDS/PAGE analysis of a 3-h long circularization reaction with varying concentration of IsMSP1E3D1 but with a fixed IsMSP1E3D1:eSrt ratio of 10 : 1. (B) SDS/PAGE analysis of circularization reactions with three fixed IsMSP1E3D1:eSrt ratios of 100 : 1, 10 : 1, and 2 : 1 at a fixed IsMSP1E3D1 concentration of 10 μM. (C) Quantification of gel band intensities in panel A. (D) Quantification of gel band intensities in panel B.

a single fraction (Fig. 3A). The two fractions of highest purity and concentration (Fig. 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 assess 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 (Fig. 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, that is, noncircularized 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 (Fig. S2). Furthermore, a tryptic peptide corresponding to the peptide segment linking the N and C terminus of csMSP1E3D1 (GGSLPETGSSFSK) was also identified (Fig. 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 (Fig. 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 straightforward manner. All intermediates of the constructs in the different purification steps are stable and highly



**Fig. 3.** Purification and concentration on IEC. (A) Chromatogram of the final purification on IEC. A step gradient was applied from 0 to 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.

soluble, making them compatible with enzymatic reactions at or above room temperature without the need for the addition of detergent or other stabilizing agents. Furthermore, the solubility-enhancing mutations, that is, extra negative charges, of the construct aid in the IEC step, where dilution with Milli-Q water to a relatively high conductivity of  $8 \text{ mS}\cdot\text{cm}^{-1}$  is enough to make the protein bind the column, thus excluding the need for tedious dialysis of very large volumes.

### Structural verification of nanodiscs by SAXS

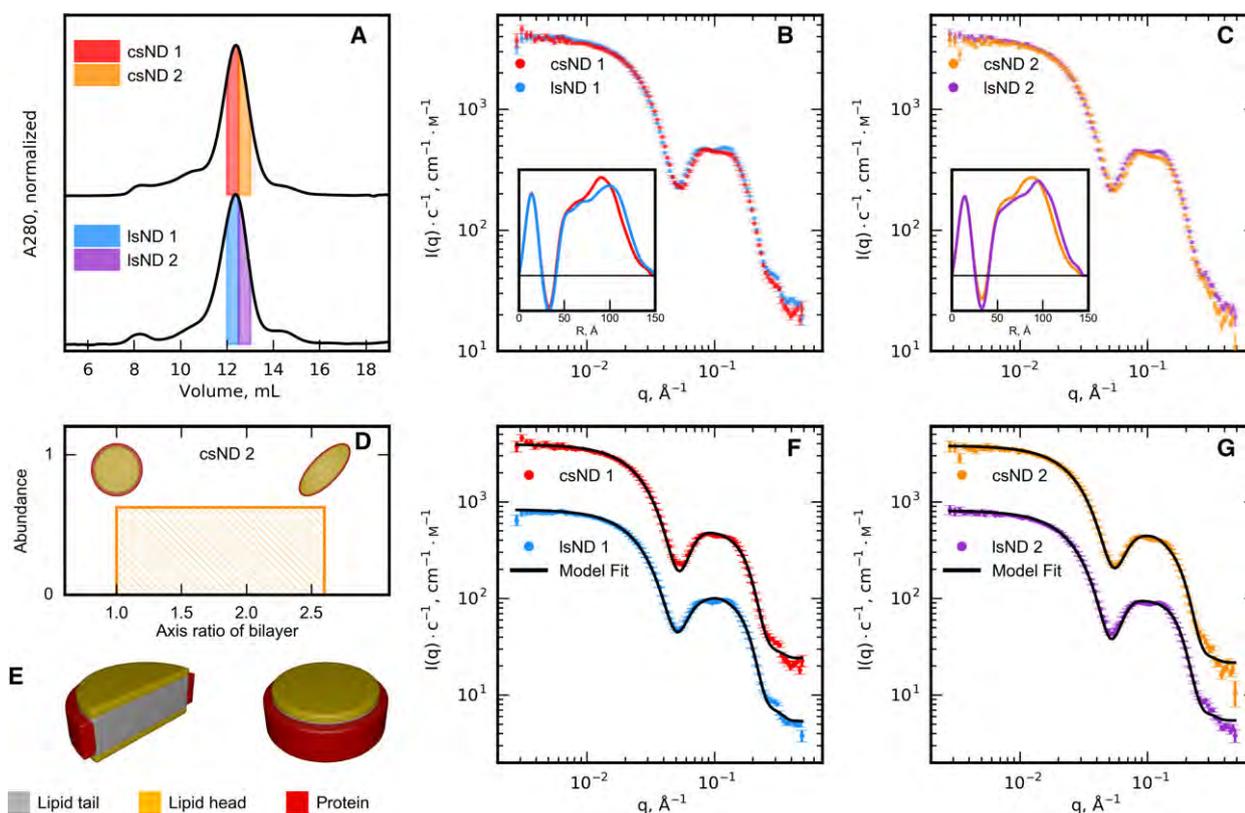
#### Model-free analysis

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  $\theta$  is the scattering angle, and  $\lambda$  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  $10 \text{ \AA}$  to  $1000 \text{ \AA}$ . 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 [10,22,24,25] 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 (Fig. 4A), giving a first indication that nanodisc reconstitution with these constructs was readily done by the usual protocol for other MSPs [14]. 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 (Fig. 4A).

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 (Fig. 4B,C). At first glance, the scattering data from the presented samples of csND and lsND differ only slightly. The same general trends and features are visible in both data sets and for both fractions indicating very similar overall structure of the samples. The  $p(r)$ -distributions refined from the data also show profiles with only small differences (Fig. 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 find that the radius of gyration ( $R_g$ ) and the largest distance in the sample ( $D_{max}$ ) 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



**Fig. 4.** Structural verification and modeling by SAXS. (A) SEC profiles of nanodiscs made from csMSP1E3D1 (csND) or IsMSP1E3D1 (IsND), 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 IsND data are divided by a factor 5 for clarity.

column, 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 IsND (Fig. 4B,C), we note that, despite the great similarity, there are for both fractions slightly more pair distances at higher  $R$ -values for the IsNDs. Furthermore,  $R_g$  and  $D_{max}$  values of the IsNDs (see Table 1) are systematically slightly larger than those of csNDs. This shows that the small differences between the circularized and non-circularized nanodiscs go in the direction of slightly larger noncircularized nanodiscs.

In a recent study, cNDs were reported to be 1–2 nm wider in diameter than corresponding noncircularized nanodiscs as evaluated by SEC and EM [18]. 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

**Table 1.** Parameters from IFT analysis.

Parameter <sup>a</sup>	csND 1	csND 2	IsND 1	IsND 2	CorA csND
Largest distance in sample, $D_{max}/\text{\AA}$	144	140	148	143	150
Radius of gyration, $R_g/\text{\AA}$	58.2	56.2	60.5	58.9	55.6

<sup>a</sup> Parameters determined from the IFT-analysis and corresponding to the  $p(r)$ -distributions shown in Fig. 4B insert, Fig. 4C insert, and Fig. 6B insert.

smaller than for IsNDs (Table 1). In another EM study, circularization did not appear to influence the mean size of the formed nanodiscs [15], and the discrepancy between all of these data could be due to differences between the specific MSP constructs.

## 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 (Fig. 4D) and hence a distribution of different lipid: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 (Fig. 4E). The model fits to the csND and lsND data are plotted on top of the respective data in Fig. 4F,G. 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  $q = 0.08 \text{ \AA}^{-1}$  and  $q = 0.15 \text{ \AA}^{-1}$ . 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  $65.2 \text{ \AA}^2$  and  $60.9 \text{ \AA}^2$ , which is in good accordance with the  $62.7 \text{ \AA}^2$  reported for POPC in vesicles at  $20 \text{ }^\circ\text{C}$  [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 [4] and our previous results [10]. 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.* [17], 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 noncircularized MSPs. Finally, the specific densities of the MSPs were refined to  $1.46 \text{ g}\cdot\text{cm}^{-3}$  for both samples. This value is higher than the  $1.35 \text{ g}\cdot\text{cm}^{-3}$  reported for globular proteins [28] and

implies that the MSPs are slightly more dense than average globular proteins.

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 [15]. 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 Fig. 4B,C.

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 cryocooling, 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.

## Stability enhancement by negative charges and circularization

### 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 [1,15,29]. 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 (Fig. 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, that is, 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

**Table 2.** Model fit parameters.

Refined parameters <sup>a</sup>	csND 1	csND 2	lsND 1	lsND 2
Area per PC headgroup/Å <sup>3</sup>	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 <sup>3</sup>	1270 ± 5.1	1280 ± 6.1	1280 ± 4.5	1270 ± 4.4
Volume of belt protein/Å <sup>3</sup>	34 200 ± 490	35 100 ± 500	33 900 ± 500	36 200 ± 470
Derived parameters <sup>b</sup>				
$\chi^2$	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 <sup>-3</sup>	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

<sup>a</sup> Structural parameters refined directly from the SAXS data in Fig. 4F,G.

<sup>b</sup> Parameters derived from the refined parameters.

magnitude of the CD signals at 222 nm for the nanodiscs (Fig. 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 222 nm/208 nm ratio can indicate whether  $\alpha$ -helices exist as free (222 nm/208 nm  $\approx$  0.9) or in a coiled-coil conformation (222 nm/208 nm  $\approx$  1.1) [30]. Thus, the observed ratio close to 0.9 for csMSP1E3D1 and lsMSP1E3D1 indicates that these proteins contain mainly free  $\alpha$ -helices, which are 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 (Fig. 5A). The negative charges on the solubility-enhanced MSPs are likely repelling each other, which is supported by SEC data that show 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. However, circularization does not perturb secondary structure.

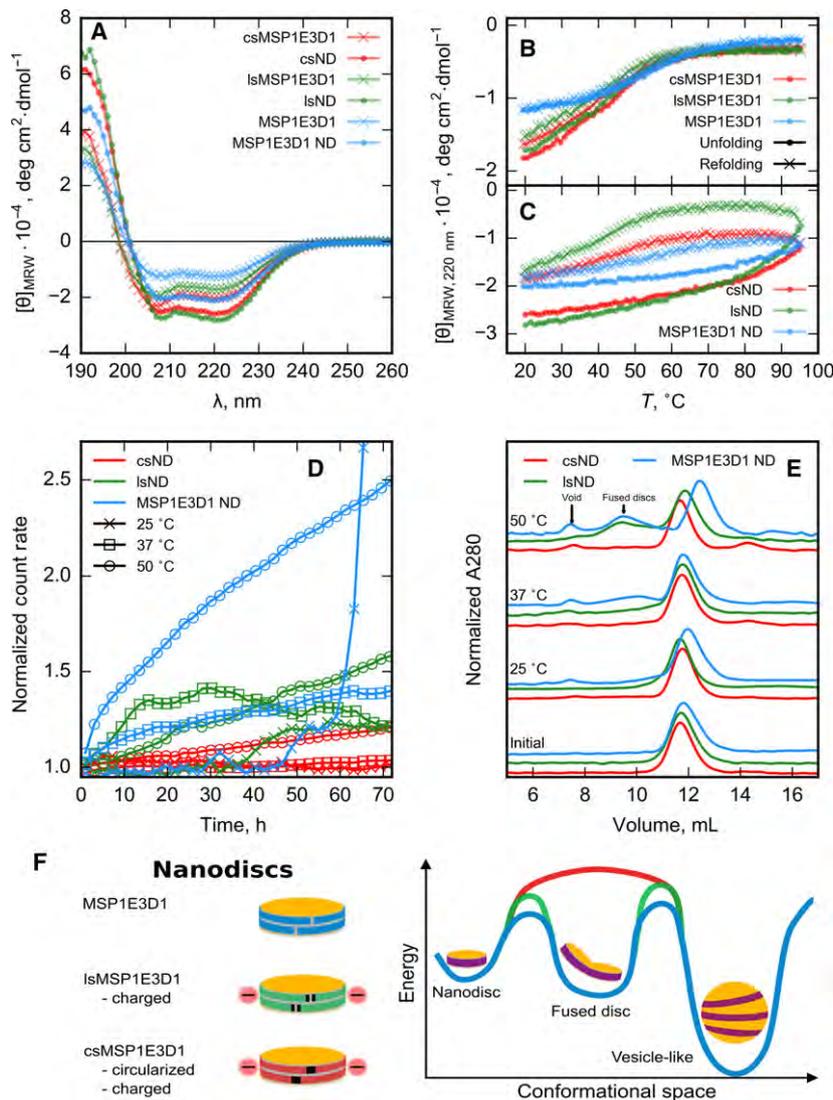
### Thermal denaturation

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 (Fig. 5B), whereas the nanodisc

samples unfold only partially toward the maximum temperature of the experiment (Fig. 5C). In the refolding of the lipid-free MSPs, the curves almost overlay those of the unfolding (Fig. 5B), whereas for the nanodisc samples, the refolding curves are very different from the unfolding curves (Fig. 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 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 calmer slope, for csND than for lsND (Fig. 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 (Fig. 5C). Given 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.



**Fig. 5.** Membrane scaffold proteins 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 nanodisc 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.

Furthermore, the circularization of lsMSP1E3D1 to obtain csMSP1E3D1 appears 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) [15], 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.

### Temporal stability and aggregation propensity

Although not observed from the CD measurements on csMSP1E3D1 (Fig. 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 was also evaluated as a function of time at different temperatures by static light scattering (SLS) and compared to MSP1E3D1 nanodiscs (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 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.*[17] for nanodiscs made from noncircularized 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 similar 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 (Fig. 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 (Fig. 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 (Fig. 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.

### 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 (Fig. 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, such that higher stability is obtained for the formed structures that resemble a vesicle-like structure, where the lipid:MSP ratio is presumably higher than for nanodiscs.

The proposed energy landscape (Fig. 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 cannot 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.

### 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 other membrane proteins [15].

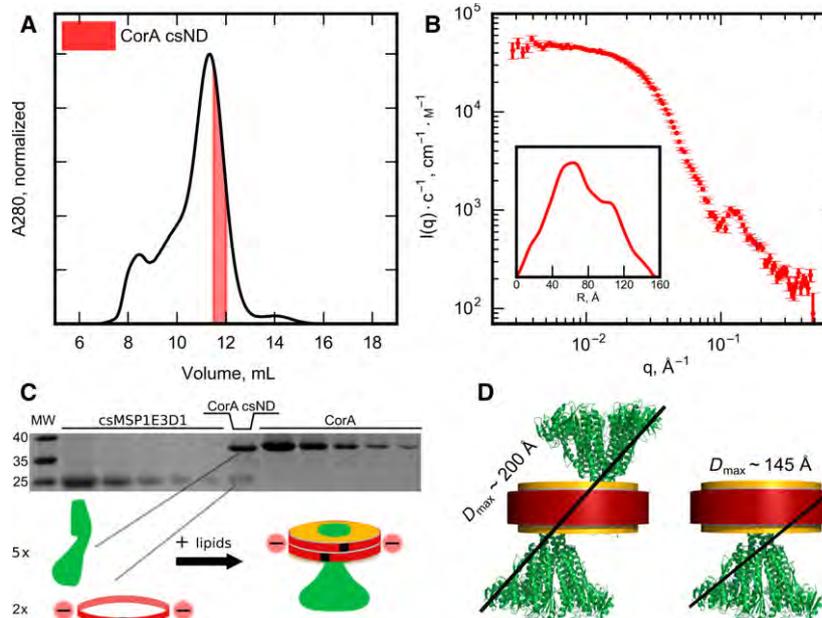
Here, the 200 kDa pentameric membrane protein CorA from *T. maritima* was used as model for testing membrane protein incorporation in csND. The SEC purification of CorA-loaded csND reveals a sample containing several subpopulations, with a minor peak in the void volume trailing into the major peak at approximately 11 mL (Fig. 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 (Fig. 6B). As described for the unloaded nanodiscs, the  $p(r)$ -distribution was determined from the data (Fig. 6B insert), yielding the parameters listed in Table 1. The scattering data show no signs of aggregation or the presence of

contaminating large particles, as seen from the flat Guinier region in the SAXS data and the smooth descent to in the  $p(r)$ -distribution.

The presence of CorA and csMSP1E3D1 in the sample was verified by SDS/PAGE (Fig. 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 agrees 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, for example, a dimeric CorA pentamer would have a  $D_{max}$  of approximately 200 Å independent of the orientation of the dimer (Fig. 6D). These results indicate 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 (Fig. 6B insert) has smaller oscillations than for the empty nanodiscs (Fig. 4B insert and Fig. 4C insert). Furthermore, the SAXS data from the CorA sample (Fig. 6B) have inferior counting statistics compared to the csND sample (Fig. 4B,C). In a model refinement situation, one would likely have to supplement the presented SAXS data with complementary SANS, ideally contrast variation SANS utilizing, for example,



**Fig. 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 Fig. 4E.

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.

## 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 but 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 noncircularized nanodiscs.

Other studies based on the original MSP1E3D1 template report on a gain of structural homogeneity for circularized nanodiscs [15]. 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 noncircularized 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.

## Methods

### Materials

The plasmid harboring the gene for MSP1E3D1 was kindly provided by 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	DDFEKKWEE	MELYREKVEP
LRAELEEGAR	EKLHELEEKL	SPLGEEMRDR
ARAHVDALRS	HLAPYLDDFE	KKWEEEMELY
REKVEPLRAE	LEEGAREKLH	ELEEKLSPLG
EEMRDRARAH	VDALRSHLAP	YSEDLRERLA
ARLEALKEDG	GARLAEYHAK	ASEHLSLSE
KAKPALEDLR	EGLLPVLESF	KVSFLSALEE
YSKCLDSEGG	RGGSLPETGG	ASHHHHHH

#### *eSrt*

MASGGSGGSG	GSGGSGGSQA	KPQIPKDKSK
VAGYIEIPDA	DIKEPVYPGP	ATREQLNRGV
SFAEENESLD	DQNISIAGHT	FIDRPNYQFT
NLKAACKGSM	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, that is, yeast extract and peptone, were from VWR (Radnor, PA, USA) and glycerol from Sigma-Aldrich.

### 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 (OD<sub>600</sub>) 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 chloride. Insoluble material was pelleted by centrifugation at 30 000 *g* 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 cutoff 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  $-80^{\circ}\text{C}$  for making noncircularized reference samples. For the circularization reaction, the remaining sample was transferred to an Erlenmeyer flask and mixed with buffer, salts, and Milli-Q water to a final concentration of 10  $\mu\text{M}$  lsMSP1E3D1 in circularization buffer (30 mM TrisHCl, pH 7.5, 150 mM NaCl, 10 mM CaCl<sub>2</sub>, 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  $37^{\circ}\text{C}$ . 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 min. 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 Milli-Q water to a conductivity of  $8\text{ mS}\cdot\text{cm}^{-1}$  (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). 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 dialyzed against gel filtration buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl) over night (ON) at  $4^{\circ}\text{C}$ , then flash frozen and stored at  $-80^{\circ}\text{C}$ .

### Expression and purification of eSrt

A codon-optimized version of eSrt based on the sequence presented in the work by Chen *et al.* [16] was synthesized and cloned into the pET28a(+) vector by Genscript. The plasmid was transformed into *E. coli* BL21 Star cells and grown at  $37^{\circ}\text{C}$  in TB media until expression was induced with 1 mM IPTG at OD<sub>600</sub> of 0.6 after which cells were grown at  $20^{\circ}\text{C}$  ON. All subsequent steps were performed at  $4^{\circ}\text{C}$ . 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 20 000 *g* 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.

### Expression and purification of CorA

The plasmid-encoding *T. maritima* CorA was kindly provided by Mikaela Rapp (University of Stockholm, Sweden). CorA was expressed and purified as described elsewhere [32,34]. The N-terminal His<sub>6</sub>-tag was left intact to allow IMAC purification of CorA-loaded nanodiscs. The purified protein was flash frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until use.

### Nanodisc preparation

Nanodiscs were prepared as previously described [14]. 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, that is, 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), at 15% w/v for empty nanodiscs and 30% w/v for CorA-loaded nanodiscs. The samples were incubated at  $5^{\circ}\text{C}$  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) 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 twofold in gel filtration buffer, concentrated to less than 1 mL, and purified by SEC on a Superdex 200 10/300 GL (GE Healthcare) 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.

### Mass spectrometry

A ProteinWorks™ eXpress Digest Kit (Waters, Miliford, MI, 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 three-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  $\mu\text{L}$  followed by addition of 25  $\mu\text{L}$  trypsin solution. The samples were subsequently incubated for 2 h at 45 °C. To quench trypsin activity, 5  $\mu\text{L}$  of digestion inactivation reagent was added to the samples and incubated for 15 min at 45 °C. The sample supernatants were collected after centrifugation at 1800  $g$  for 15 min at 10 °C.

The 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  $\mu\text{m}$  VanGuard column, Waters) 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  $\mu\text{L}\cdot\text{min}^{-1}$ . 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  $\mu\text{m}$  VanGuard column, Waters) for 3 min at 150  $\mu\text{L}\cdot\text{min}^{-1}$  of mobile phase A (0.23% FA). The tryptic peptides were eluted across a C18 analytical column (ACQUITY UPLC BEH C18 1.7  $\mu\text{m}$  1.0  $\times$  100 mm column, Waters) 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  $\mu\text{L}\cdot\text{min}^{-1}$ . 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, USA). Criteria for peptide identification were a mass tolerance for precursor 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.

## Gel quantification of MSP circularization

For the quantification of the influence of substrate concentration, lsMSP1E3D1 was mixed with eSrt in 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% trichloroacetic acid to a final concentration of 12%. Precipitated protein was pelleted by centrifugation at 14 500  $g$  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 (50 mM TrisHCl pH 6.8, 10% glycerol, 2% w/v SDS, 0.05% w/v bromophenol blue) to a final concentration of 8  $\mu\text{M}$  MSP. For the quantification of the circularization reaction with varying lsMSP1E3D1:eSrt ratios, reactions were setup with different amounts of eSrt in samples containing 10  $\mu\text{M}$  lsMSP1E3D1. At indicated time points, 20  $\mu\text{L}$  samples were taken out and quenched by the addition of 5  $\mu\text{L}$  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). The images were quantified in ImageLab (Bio-Rad) using standard settings. Relative gel band intensities were calculated to report the relative quantities of species.

## Gel quantification of CorA-loaded nanodiscs

Standards with a known absorption at 280 nm were prepared of CorA and csMSP1E3D1, respectively, and loaded in different amounts on the same gel as the CorA-loaded nanodisc sample. Gels were stained with Coomassie brilliant blue G-250 (Bio-Rad) and subsequently destained in 10% acetic acid, 20% ethanol. Gel images were obtained and quantified as above. The intensities of the standards were fitted with linear regression in order to quantify the amount of CorA and csMSP1E3D1 in the nanodisc sample.

## Static light scattering

Samples of 5  $\mu\text{M}$  nanodiscs were measured on a BI-200SM light scattering instrument (Brookhaven Instruments, Holtsville, NY, USA) equipped with a 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  $\mu\text{m}$  filter (Merck/Millipore, Darmstadt, Germany) to eliminate dust particles. Hereafter, samples were degassed for 15 min 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].

### Circular dichroism spectroscopy

Samples were exchanged into 20 mM sodium phosphate pH 7.0 on a NAP5 column (GE Healthcare) and subsequently diluted to 1.5  $\mu\text{M}$  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 $\cdot$ min $^{-1}$ , 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 ( $[\theta]_{\text{MRW}}$ ) with the relation

$$[\theta]_{\text{MRW}} = \frac{m\theta}{10 \cdot c \cdot N_{pb} \cdot l}$$

where  $c$  is the protein concentration in molar,  $N_{pb}$  is the number of peptide bonds, that is, 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 $\cdot$ min $^{-1}$ . The melting temperatures were estimated by the inflection points of the melting curves. These were computed by the maxima of the first derivatives, which were calculated using the Savitsky–Golay algorithm [36].

### Small angle X-ray scattering

Samples were prepared in the laboratory 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  $\mu\text{M}$ , csND 2 = 10.3  $\mu\text{M}$ , lsND 1 = 12.3  $\mu\text{M}$ , lsND 2 = 11.1  $\mu\text{M}$ , and CorA csND = 0.9  $\mu\text{M}$ . 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  $\theta$  being the scattering angle and  $\lambda$  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 were brought to absolute scale in units of  $\text{cm}^{-1}$  using water as a calibration standard [38]. 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 further 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 [10,41]. The refinement was again done in WillItFit [40], with the source code of the model available in the associated repository. Aside 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 [4].

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.* [10,41]. 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 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,  $I(q)$ :

$$I(q) = \frac{n}{\kappa_{\text{max}} - 1} \int_1^{\kappa_{\text{max}}} d\kappa P(q, \vec{p})$$

where  $\vec{p}$  is the list of parameters in our model for a nanodisc with a particular shape,  $\kappa$  is the aforementioned axis ratio of the bilayer, and  $n$  is the number density of nanodiscs estimated from the absorption at 280 nm. In particular, note that  $\kappa \in \vec{p}$ , whereas  $\kappa_{\text{max}}$ , 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,  $\vec{p}$  [10,41].

### 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 and Jens Berndtsson

(University of Stockholm, Sweden) for the plasmid harboring the gene for CorA and for providing guidance on protein purification, 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 and the University of Copenhagen, bioSYnergy; Novo Nordisk Foundation, Synergy; Lundbeck Foundation, Brainstruc for funding.

## Conflict of interest

The authors declare no conflict of interest.

## 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.

## References

- 1 Bayburt TH, Grinkova YV & Sligar SG (2002) Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Lett* **2**, 853–856.
- 2 Hagn F, Etzkorn M, Raschle T & Wagner G (2013) Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *J Am Chem Soc* **135**, 1919–1925.
- 3 Denisov IG & Sligar SG (2017) Nanodiscs in membrane biochemistry and biophysics. *Chem Rev* **117**, 4669–4713.
- 4 Bibow S, Polyhach Y, Eichmann C, Chi CN, Kowal J, Albiez S, McLeod RA, Stahlberg H, Jeschke G, Güntert P *et al.* (2017) Solution structure of discoidal high-density lipoprotein particles with a shortened apolipoprotein A-I. *Nat Struct Mol Biol* **24**, 187–193.
- 5 Denisov IG, Grinkova YV, Lazarides AA & Sligar SG (2004) Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *J Am Chem Soc* **126**, 3477–3487.
- 6 Seddon AM, Curnow P & Booth PJ (2004) Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta – Biomembr* **1666**, 105–117.
- 7 Jørgensen IL, Kemmer GC & Pomorski TG (2017) Membrane protein reconstitution into giant unilamellar vesicles: a review on current techniques. *Eur Biophys J* **46**, 103–119.
- 8 Dürr UHN, Gildenberg M & Ramamoorthy A (2012) The magic of bicelles lights up membrane protein structure. *Chem Rev* **112**, 6054–6074.
- 9 Hagn F, Nasr ML & Wagner G (2018) Assembly of phospholipid nanodiscs of controlled size for structural studies of membrane proteins by NMR. *Nat Protoc* **13**, 79–98.
- 10 Skar-Gislinge N, Simonsen JB, Mortensen K, Feidenhans'l R, Sligar SG, Lindberg Møller B, Bjørnholm T & Arleth L (2010) 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.
- 11 Kynde SAR, Skar-Gislinge N, Pedersen MC, Midtgaard SR, Simonsen JB, Schweins R, Mortensen K & Arleth L (2014) Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment. *Acta Crystallogr Sect D Biol Crystallogr* **70**, 371–383.
- 12 Shen PS, Yang X, DeCaen PG, Liu X, Bulkley D, Clapham DE & Cao E (2016) The structure of the polycystic kidney disease channel PKD2 in lipid nanodiscs. *Cell* **167**, 763–773.
- 13 Gao Y, Cao E, Julius D & Cheng Y (2016) TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* **534**, 347.
- 14 Ritchie TK, Grinkova Y V, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM & Sligar SG (2009) Chapter eleven – reconstitution of membrane proteins in phospholipid bilayer nanodiscs. In *Liposomes, Part F* (Düzgünes N, ed.), pp. 211–231. Academic Press, Cambridge, MA.
- 15 Nasr ML, Baptista D, Strauss M, Sun Z-YJ, Grigoriu S, Huser S, Plückthun A, Hagn F, Walz T, Hogle JM *et al.* (2017) Covalently circularized nanodiscs for studying membrane proteins and viral entry. *Nat Methods* **14**, 49–52.
- 16 Chen I, Dorr BM & Liu DR (2011) A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc Natl Acad Sci USA* **108**, 11399–11404.
- 17 Yusuf Y, Massiot J, Chang YT, Wu PH, Yeh V, Kuo PC, Shue J & Yu TY (2018) Optimization of the production of covalently circularized nanodiscs and their characterization in physiological conditions. *Langmuir* **34**, 3525–3532.
- 18 Miehl J, Goricanec D & Hagn F (2018) A split-intein-based method for the efficient production of circularized nanodiscs for structural studies of membrane proteins. *ChemBioChem* **19**, 1927–1933.
- 19 Trevino SR, Scholtz JM & Pace CN (2007) 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.
- 20 Kramer RM, Shende VR, Motl N, Pace CN & Scholtz JM (2012) Toward a molecular understanding of protein solubility: increased negative surface charge

- correlates with increased solubility. *Biophys J* **102**, 1907–1915.
- 21 Denisov IG, McLean MA, Shaw AW, Grinkova YV & Sligar SG (2005) Thermotropic phase transition in soluble nanoscale lipid bilayers. *J Phys Chem B* **109**, 15580–15588.
- 22 Huda P, Binderup T, Pedersen MC, Midtgaard SR, Elema DR, Kjær A, Jensen M & Arleth L (2015) PET/CT based in vivo evaluation of <sup>64</sup>Cu labelled nanodiscs in tumor bearing mice. *PLoS ONE* **10**, e0129310.
- 23 Sigalov AB & Stern LJ (2001) Oxidation of methionine residues affects the structure and stability of apolipoprotein A-I in reconstituted high density lipoprotein particles. *Chem Phys Lipids* **113**, 133–146.
- 24 Midtgaard SR, Pedersen MC & Arleth L (2015) Small-angle X-ray scattering of the cholesterol incorporation into human ApoA1-POPC discoidal particles. *Biophys J* **109**, 308–318.
- 25 Graziano V, Miller L & Yang L (2018) Interpretation of solution scattering data from lipid nanodiscs. *J Appl Crystallogr* **51**, 157–166.
- 26 Skar-Gislinge N, Johansen NT, Høiberg-Nielsen R & Arleth L (2018) Comprehensive study of the self-assembly of phospholipid nanodiscs: What determines their shape and stoichiometry? *Langmuir* **34**, 12569–12582.
- 27 Kučerka N, Nieh M-P & Katsaras J (2011) Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim Biophys Acta – Biomembr* **1808**, 2761–2771.
- 28 Mylonas E & Svergun DI (2007) Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J Appl Crystallogr* **40**, s245–s249.
- 29 Larsen AN, Sørensen KK, Johansen NT, Martel A, Kirkensgaard JJK, Jensen KJ, Arleth L & Midtgaard SR (2016) Dimeric peptides with three different linkers self-assemble with phospholipids to form peptide nanodiscs that stabilize membrane proteins. *Soft Matter* **12**, 5937–5949.
- 30 Greenfield NJ & Hitchcock-Degregori SE (1993) Conformational intermediates in the folding of a coiled-coil model peptide of the N-terminus of tropomyosin and  $\alpha$ -tropomyosin. *Protein Sci* **2**, 1263–1273.
- 31 Antos JM, Popp MW-L, Ernst R, Chew G-L, Spooner E & Ploegh HL (2009) A straight path to circular proteins. *J Biol Chem* **284**, 16028–16036.
- 32 Johansen NT, Pedersen MC, Porcar L, Martel A & Arleth L (2018) Introducing SEC–SANS for studies of complex self-organized biological systems. *Acta Crystallogr Sect D Struct Biol* **74**, 1178–1191.
- 33 Maric S, Skar-Gislinge N, Midtgaard S, Thygesen MB, Schiller J, Frielinghaus H, Moulin M, Haertlein M, Forsyth VT, Pomorski TG *et al.* (2014) Stealth carriers for low-resolution structure determination of membrane proteins in solution. *Acta Crystallogr Sect D Biol Crystallogr* **70**, 317–328.
- 34 Palombo I, Daley DO & Rapp M (2012) The periplasmic loop provides stability to the open state of the CorA magnesium channel. *J Biol Chem* **287**, 27547–27555.
- 35 Høiberg-Nielsen R, Fuglsang CC, Arleth L & Westh P (2006) Interrelationships of glycosylation and aggregation kinetics for *Peniophora lycii* Phytase. *Biochemistry* **45**, 5057–5066.
- 36 Savitzky A & Golay MJE (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* **36**, 1627–1639.
- 37 Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, Franke D, Mark D, Zengerle R, Cipriani F *et al.* (2015) Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J Appl Crystallogr* **48**, 431–443.
- 38 Orthaber D, Bergmann A & Glatter O (2000) SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J Appl Crystallogr* **33**, 218–225.
- 39 Hansen S (2014) Update for *BayesApp*: a web site for analysis of small-angle scattering data. *J Appl Crystallogr* **47**, 1469–1471.
- 40 Pedersen MC, Arleth L & Mortensen K (2013) WillItFit : a framework for fitting of constrained models to small-angle scattering data. *J Appl Crystallogr* **46**, 1894–1898.
- 41 Skar-Gislinge N & Arleth L (2011) Small-angle scattering from phospholipid nanodiscs: derivation and refinement of a molecular constrained analytical model form factor. *Phys Chem Chem Phys* **13**, 3161–3170.
- 42 Als-Nielsen J & McMorrow D (2011) *Elements of Modern X-ray Physics*, 2nd edn. Wiley, West Sussex.
- 43 Pedersen JS (1997) Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting. *Adv Colloid Interface Sci* **70**, 171–210.
- 44 Pawitan Y (2001) *In All Likelihood*. Clarendon Press, Cork, Ireland.
- 45 Pedersen MC, Hansen SL, Markussen B, Arleth L & Mortensen K (2014) Quantification of the information in small-angle scattering data. *J Appl Crystallogr* **47**, 2000–2010.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** IEC purification tests.

**Fig. S2.** MS sequence coverage.

**Fig. S3.** Identification of oxidized methionines.

# **Circularized and solubility-enhanced MSPs facilitate simple and high-yield production of stable nanodiscs for studies of membrane proteins in solution**

Nicolai Tidemand Johansen, Frederik Grønbæk Tidemand, Tam T. T. N. Nguyen, Kasper Dyrberg Rand, Martin Cramer Pedersen and Lise Arleth

DOI: 10.1111/febs.14766

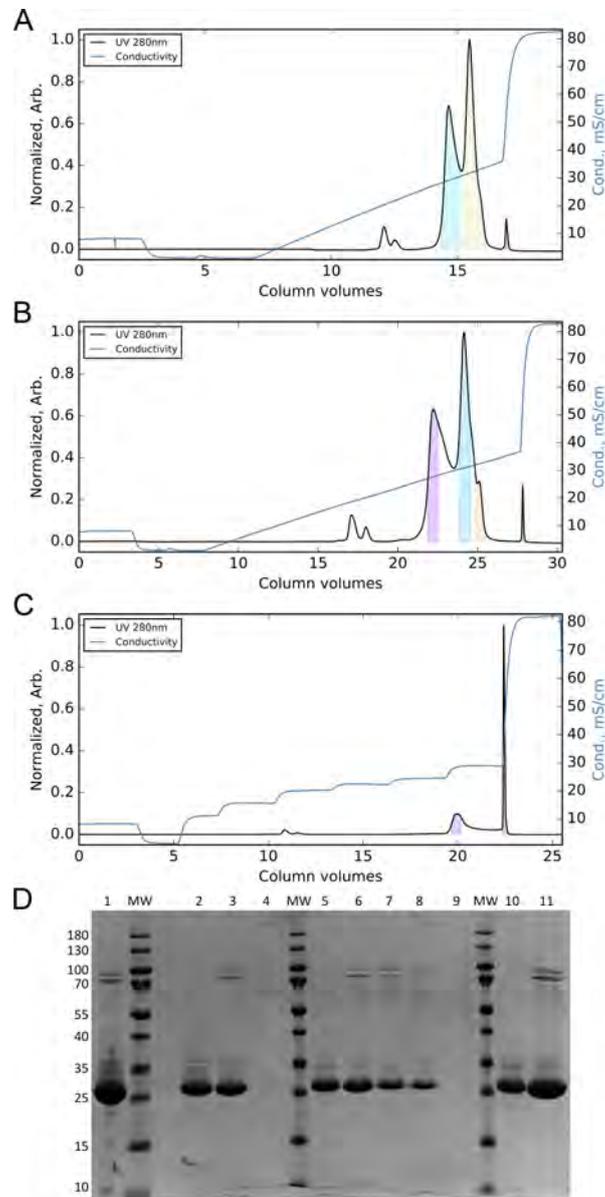
---

**SUPPORTING INFORMATION**

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

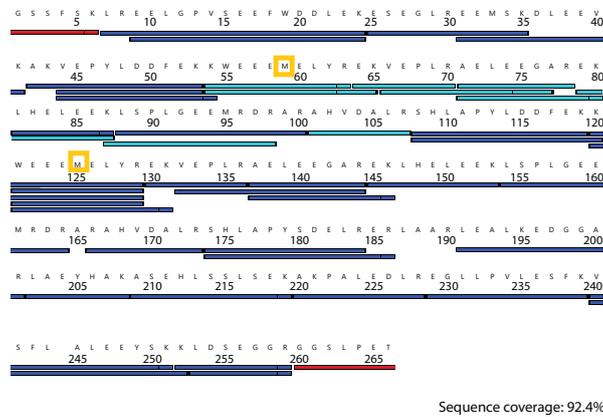
Nicolai Tidemand Johansen<sup>1</sup> | Frederik Grønbæk  
Tidemand<sup>1</sup> | Tam T. T. N. Nguyen<sup>2</sup> | Kasper Dyrberg  
Rand<sup>2</sup> | Martin Cramer Pedersen<sup>1</sup> | Lise Arleth<sup>1</sup>

## Circularization: IEC



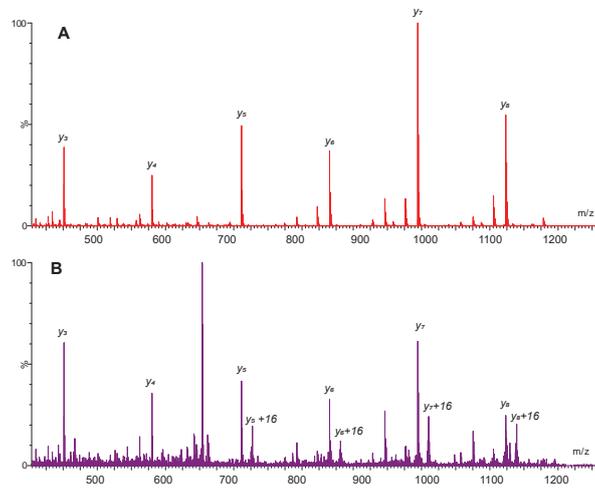
**FIGURE S1** IEC purification tests. A: csMSP1E3D1 elution with 10 CV gradient from 0 to 0.5 M NaCl. B: 20 CV gradient from 0 to 0.5 M NaCl. C: Multi-step gradient from 0 to 0.32 M NaCl. D: SDS-PAGE analysis of the samples collected in A, B and C (colored areas). Lane 1 is the sample before the IEC 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 weight 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

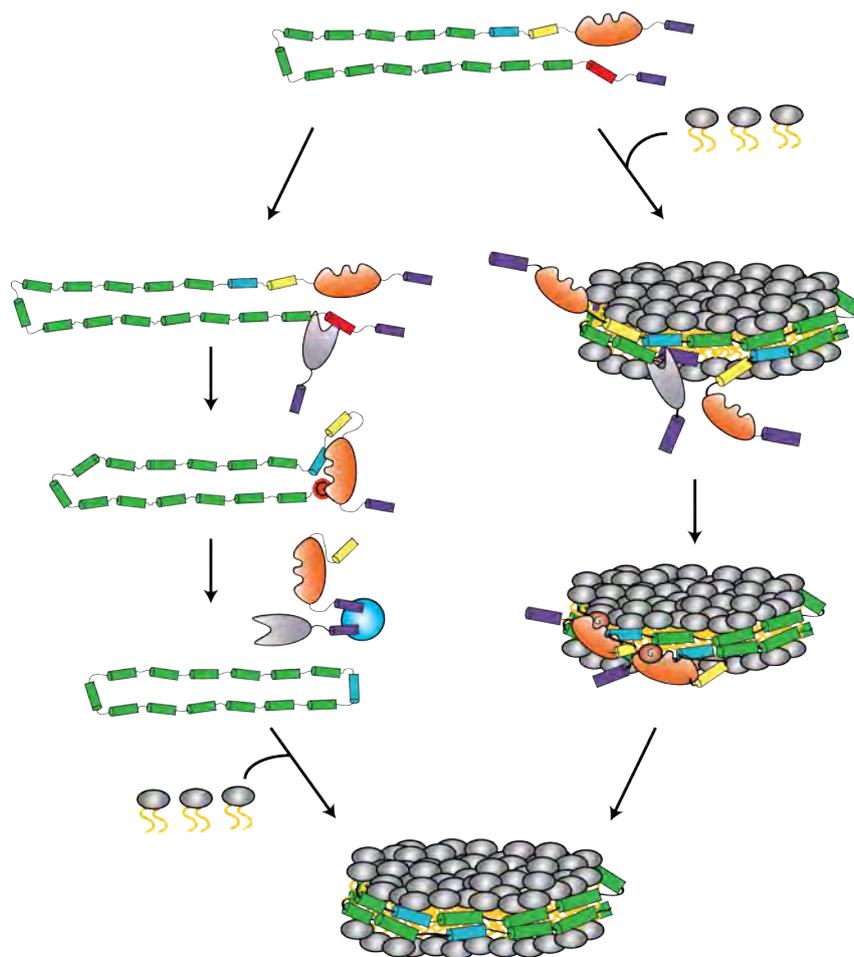


**FIGURE S3 Identification of oxidized methionines.** CID MS/MS fragmentation spectra of (A) non-oxidized and (B) oxidized form of a tryptic peptide 121-129 containing methionine (WEEEMEL<sup>R</sup>YR). The precursor ions were  $[M+2H]^{2+}$  (m/z 642.79) and  $[M+2H+16]^{2+}$  (m/z 650.79), respectively .

## 6 | Paper II

### Circularizing reconstituted nanodiscs using a fusion construct

*Frederik Grønæk Tidemand, Xiaoxuan Chen, Nicolai Tidemand Johansen, Tam T. T. N. Nguyen, Kasper Dyrberg Rand, Birthe B. Kragelund, Lise Arleth*



# Circularizing reconstituted nanodiscs using a fusion construct

Frederik Grønbaek Tidemand<sup>1</sup>, Xiaoxuan Chen<sup>1</sup>, Nicolai Tidemand Johansen<sup>1</sup>, Tam T. T. N. Nguyen<sup>2</sup>, Kasper Dyrberg Randz, Birthe B. Kragelund<sup>3</sup>, Lise Arleth<sup>1</sup>

1: Structural Biophysics, X-ray and Neutron Science, The Niels Bohr Institute, University of Copenhagen, Denmark.

2: Protein Analysis Group, Department of Pharmacy, University of Copenhagen, Denmark

3: Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Denmark

## Abstract

Gaining structural and functional information about membrane proteins is vital for our understanding of many biological processes. To obtain this information, carrier systems optimized for study of membrane proteins (MP) are needed, as they are essential for maintaining a stable MP. The nanodisc (ND) is one of these carrier systems, and using the Sortase enzyme, circularized NDs (cND) were recently developed. Although being more time-consuming to produce, the cNDs have gained ground, as they were both more homogeneous and more stable at physiological temperatures, which are important features for both functional and structural studies. Two of the major bottlenecks for the ND technology, are the production of the membrane scaffolding proteins and the reconstitution of an MP into the ND. In the present study, we developed a fusion construct that is used for an optimized straight forward production of circularized NDs, as it skips the need for producing Sortase in parallel. We also implemented a protocol using circularization after reconstitution of the ND, which opens up the ND technology for membrane proteins that do not incorporate well into pre-circularized MSPs, produced using current protocols. Easier production of MSPs, and a broader range of reconstitution opportunities, will facilitate easier access to the ND technology, and will hopefully result in more successful studies of membrane proteins.

## Introduction

During the last decade, techniques for studying the structure of membrane proteins (MP) have continuously developed and improved. Similarly, the sample environments have improved, with the development of new carrier systems, used to keep the MPs in solution. Detergents are usually employed during purification of the MPs, but for both functional and structural studies, other carrier systems are usually preferred, as in some cases detergents compromise the activity and stability of the MP (1,2). The nanodisc (ND) is a widely used carrier system, because of the relatively low heterogeneity, adjustable size, and the native-like lipid environment that better maintains the native activity of some MPs (1,3). The ND is composed of a lipid bilayer held together by two copies of an amphipathic protein called a membrane scaffolding protein (MSP). The MSP is central in controlling the size, stability, and monodispersity of the final ND, and therefore there have been many studies trying to develop improved and versatile MSPs (2,4,5). A recent major improvement of the ND was the invention of circularized MSPs, which was achieved by making use of the enzyme Sortase that forms a covalent bond between a specific five-residue C-terminal motif (LPXTG) and a free N-terminal Gly (2). An evolved version of the Sortase (eSrt) which has a 140-fold higher activity compared to Sortase (6), was utilized to link the two termini of the MSP, to make a circularized variant, which proved to form more monodisperse and stable NDs (2,6). The enzymatic reaction was performed after the purification of the MSP, and did not seem to affect the following reconstitution of the ND and MP incorporation.

In our group, we have further developed these circularized NDs (cNDs) by mutating the MSP sequence, introducing a substantial number of negatively charged residues to enhance the solubility of the formed ND (5). These NDs were named, circularized and solubility enhanced MSP1E3D1 (csE3), and proved to be possible to produce in high amounts. In addition, they were found to have a much higher temporal stability at physiologically relevant temperatures compared to the original MSP1E3D1 and also the non-circularized, linear counterpart (lsE3) (5). The cNDs have until now only been applied for a few MPs, that all contain only a transmembrane domain or a transmembrane domain and a single extramembrane domain on one side of the membrane, and until now no

problems with incorporation of MPs in the cNDs have been reported (2,5,7,8). However, for more complicated MPs with hydrophilic domains on both sides of the membrane, this may not be as straight forward, as one of the hydrophilic domains would have to go through the hydrophobic interior of the MSP upon insertion. E.g. most G protein-coupled receptors contain both a large folded extracellular domain and an intercellular unfolded domain, responsible for recruiting the G-protein (9). The same overall topology is present for the cytokine receptors and other MP families (10,11), and thus we foresee problems with the incorporation into circularized MSPs for these classes of MPs.

A drawback of the ND technology, and especially the circularized variants is the need for producing the circularized MSP, which is a time-consuming process, consisting of several purification steps. To approach this, an in cell split-intein based production was developed to circularize the MSP *in vivo*, but this method had the drawback that subsequent purification had to be carried out by other means than the very convenient affinity chromatography (8). In addition, the obtained yields ranged from 8 to 25 mg per liter culture depending on the MSP, which is similar to yields obtained by Sortase circularization (7). Therefore, the main approach is still to use the eSrt enzyme. Using this strategy, two enzymes (TEV and eSrt), and the circularized MSP, have to be produced recombinantly in *E. Coli*, and the different enzymatic reactions have to be carried out after the purification to produce the MSPs. Furthermore, the yields of most of the circularized variants have shown to be considerably lower than the original MSPs, in many cases around a factor of 10 lower (8,12,13). This hampers the general utility for this technique, as high amounts of MSPs are usually required, especially for structural studies.

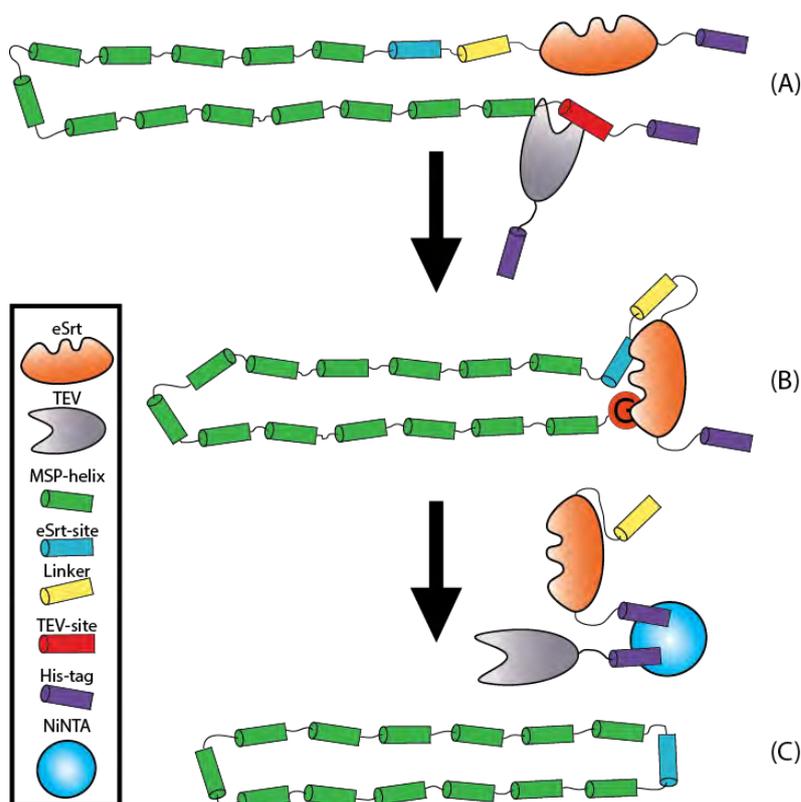


Figure 1: Schematic overview of the mechanism of fE3. A: eSrt and lse3 joint in one construct: fE3. B: After TEV cleavage a free N-terminal Glycine is available for eSrt reaction. C: After circularization both TEV and eSRT

In the present study, we designed a fusion protein (called fE3) which joins the eSrt and the lse3 in one, minimizing the number of proteins to express and purify (see figure 1A for an overview of the construct). Using fE3, the circularized cse3 can in principle be obtained by simply adding TEV, leaving a free Glycine, which is recognized by the eSrt, which will then perform the circularization process (figure 1). The idea of having both the eSrt and the eSrt-site in the same protein has been used in previous studies as a purification strategy for other proteins, and as a molecular tagging strategy, which both proved very efficient (14–16). We show that fE3 can be used for both pre and post circularization of NDs. In addition, fE3 is compatible with MP incorporation, and a combination of functional studies and SAXS measurement confirmed that both the function and structure of the MP were maintained post-circularization.

## Methods

### Expression and purification

All chemicals and lipids are from Sigma Aldrich. lsE3, csE3 and tissue factor (TF) were expressed and purified as described previously (5)(Paper III). Four fusion constructs containing the sequence from lsE3 (containing a TEV site and the eSrt motif), a linker (of variable length: 3, 9 or 18 residues) and a variant of sortase (either WT Sortase or eSrt (6)) were synthesized and cloned into the pET28a vector by Twist Bioscience (see an overview of the constructs in figure S1A+B). The plasmids were transformed into BL21 Star cells, and grown in TB media. Cells were grown to an optical density at 600 nm (OD<sub>600</sub>) at 0.6-0.8 before the expression was induced with 1 mM IPTG either at 20 °C for 20 hours or 37 °C for three hours. Cells were harvested by centrifugation, and the pellet was dissolved in 3.5 ml per grams of cells in lysis buffer (6 M guanidine hydrochloride (GuHCl), 30 mM Tris-HCl, pH 8, 150 mM NaCl, 20 mM imidazole). Cellular debris was pelleted by centrifugation at 20000g for 20 min before adding 3 ml buffer equilibrated NiNTA resin per liter of initial culture, to the supernatant. After incubation for 45 min, the preparation was loaded onto an empty gravity flow column, and the resin was washed with 5 column volumes (CV) of 30 mM Tris-HCl, pH 8, 150 mM NaCl, 40 mM imidazole and eluted with 5 CV of 30 mM Tris-HCl, pH 8, 150 mM NaCl, 250 mM imidazole while fractions were collected. The fractions containing the protein were pooled and dialyzed against 20 mM Tris-HCl, pH 8, 80 mM NaCl, 1 mM EDTA before loading onto a 5 ml HiTrap Q column (GE Healthcare) equilibrated in 20 mM Tris-HCl, pH 8, 1 mM EDTA. The column was washed with 5 CV of 20 mM Tris-HCl, pH 8, 200 mM NaCl, 1 mM EDTA before a step gradient to 250 mM NaCl and 300 mM NaCl was applied while fractions were collected. After visualization by SDS-PAGE, fractions containing pure fusion protein were pooled, concentrated in a spinfilter (10 kDa cutoff), and kept at 4 °C until use. For purification of csE3 from the fusion protein, the elution fractions from the NiNTA purification were diluted to a concentration of 10 μM in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl and CaCl<sub>2</sub> was added to a final concentration of 10 mM. TEV was added to a final concentration of 0.1 μM, and the preparation was incubated at 37 °C for three hours before equilibrated NiNTA resin was added. After another hour of incubation, the preparation was added to an empty gravity flow column, and the flowthrough was collected and loaded onto a 5 ml HiTrap Q column equilibrated in 20 mM Tris-HCl, pH 8. The column was washed with 5 CV of 20 mM Tris-HCl, pH 8, 200 mM NaCl, 1 mM EDTA before a step gradient going to 250 mM NaCl was applied while fractions were collected.

### Refolding tests

To optimize refolding from inclusion bodies (IB) of the fusion construct, four different strategies were applied, using either urea or GuHCl as denaturant and either on column or dialysis as the refolding method. Pellets from expression at 37°C were dissolved in either 8 M urea, 30 mM Tris-HCl, pH 8, 150 mM NaCl or 6 M GuHCl, 30 mM Tris-HCl, pH 8, 150 mM NaCl and centrifuged to remove cellular debris. Half of the supernatant was put in a dialysis bag (10 kDa cutoff) and submerged in 2 liters of 30 mM Tris-HCl, pH 8, 150 mM NaCl placed in a coldroom over night (ON) stirring. The other half was incubated with NiNTA resin for 45 min before loading it onto an empty gravity flow column. The resin was washed with 5 CV of 30 mM Tris-HCl, pH 8, 150 mM NaCl, 40 mM Imidazole and eluted with 5 CV of 30 mM Tris-HCl, pH 8, 150 mM NaCl, 250 mM Imidazole while fractions were collected. Volume normalized samples of the different steps in the refolding were loaded onto a 15 % SDS-PAGE gel.

### Circularization tests

Circularization tests were performed for both the fusion construct and lsE3 with a concentration of 10 μM. The latter was done at a lsE3:eSrt ratio of 1:1 to mimic the ratio of the fusion protein. The reactions were conducted in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl and either 10 mM CaCl<sub>2</sub> or 10 mM EDTA at 37°C with shaking. Samples were extracted after one minute and after one, three, and 24 hours. Samples were loaded onto 15 % SDS-PAGE gel with relevant standards. For the circularization of an already reconstituted ND, the same conditions were applied, except that the MSP concentration was 76 μM as this is the concentration used for the standard ND reconstitution. The NDs were prepared using a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC):MSP ratio of 130:1. POPC was dissolved in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 100 mM cholate to a concentration of 50 mM before mixing with the MSP (either lsE3 or the fusion protein). After incubation for 15 min at 10 °C, detergent absorbing

beads (Amberlite XAD-2) were added at 15% w/v ratio. After incubation ON, the beads were removed and the circularization tests were conducted as described above.

### **Mass spectrometry**

To confirm the protein sequence and the presence of the unique peptide formed by the enzymatic reaction, peptide mapping was performed by tryptic digestion of csE3 (from fE3) and LC-MS/MS analysis on a NanoAQUITY UPLC system with a C18 reversed-phase column (ACQUITY UPLC BEH C18 1.7  $\mu\text{m}$  1.0 x 100 mm column, Waters, Miliford, USA) coupled to a ESI hybrid Q-TOF mass spectrometer (Synapt G2-SI, Waters, Wilmslow, UK) as described in the previous study (5).

### **Reconstitution of nanodiscs and incorporation of a membrane protein**

POPC and 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS) were dissolved in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 100 mM sodium cholate to a final concentration of 50 mM. NDs without TF were assembled with a lipid:MSP ratio of 130:1 and all TF containing NDs were assembled with a lipid:MSP ratio of 120:1 and an MSP:TF ratio of 8. All preparations had a final lipid concentration of 10 mM. Samples were mixed and incubated for 15 min at room temperature (RT) before detergent absorbing beads (Amberlite XAD-2, Sigma) were added to a ratio of 15% w/v. Samples were in a shaking thermomixer at 10°C for three hours before the relevant enzyme was added (TEV for fE3 and eSrt for lsE3). After incubation ON, the samples were separated from the detergent absorbing beads. Samples containing TF were applied to a 1 ml HisTrap column (GE Healthcare) equilibrated in 20 mM Tris-HCl, pH 8, 100 mM NaCl, and 20 mM imidazole and eluted with 20 mM Tris-HCl pH 8, 100 mM NaCl, 300 mM imidazole while fractions of 1 ml were collected. Fractions containing the TF loaded discs were pooled, concentrated to 0.5 ml and applied to a Superdex 200 10/300 GL column (GE Healthcare) equilibrated in gel filtration buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM  $\text{CaCl}_2$ ). Empty NDs were directly applied to the same gel filtration column in the same buffer. Samples were stored at 4°C until they were used for either functional studies or SAXS measurements.

### **Small angle X-ray scattering**

SAXS measurements were performed at the P12 beamline at PETRA III in Hamburg, Germany (17). The samples were measured at 10°C and the intensity was measured as a function of the magnitude of the scattering vector,  $q = 4 \cdot \sin(\theta)/\lambda$ , where  $\theta$  is the scattering angle and  $\lambda$  is the wavelength, for this experiment 1.24 Å. The data was automatically reduced and background subtracted using the software available at the at the beamline (17), yielding the scattering intensities  $I(q)$  in units of  $\text{cm}^{-1}$  for the different samples. Pair distance distribution functions,  $p(r)$ , were obtained using the BayesApp-server (18).

### **Activity assay**

The affinity between TF in NDs and Factor VIIa (provided by Novo Nordisk A/S) (FVIIa) were probed using the increase in FVIIa activity upon TF binding as done in previous studies (19,20). Factor X (FX) (650 nM) activation was allowed for 20 min at room temperature by 10 pM TF in ND and a varying amount of FVIIa. The reaction was stopped by adding EDTA and the generated amount of activated FX (FXa) was quantified using the chromogenic substrate S2765 (from Chromogenix). The slope of the absorbance at 405 nm was evaluated for 20 min and a quadratic binding function was fitted to the slope as a function of FVIIa concentration to obtain the  $K_d$ .

## Results and discussion

### Expression and refolding of fE3

eSrt expression is normally performed at 20°C and MSP at 37°C to maximize yields (5,6), and thus, both conditions were tested for fE3 (figure 2A). The band around 55 kDa (green asterisk) demonstrated that fE3 was expressed at 37°C and present in the crude cellular lysate, while only a minimal amount was present when expressed at 20°C (green asterisk). After sonication, the samples were centrifuged, and the fE3 band appeared in the pellet (figure 2A, teal asterisk), suggesting that it partitioned into IBs. This explained the little amount of protein obtained from NiNTA purification, which was performed on the supernatant (yellow asterisk). However, two other major bands were seen in the analyses of the NiNTA elution at both temperatures (marked with red asterisks). The 25 kDa band corresponds to eSrt in size and the 35 kDa corresponds to the MSP part of the fE3. The development of two bands instead of one has been encountered in previous studies with similar construct designs (14–16).

There are three possible explanations for this problem: 1) the flexible linker between MSP and eSrt is prone to degradation by proteases present in the bacterial cells, and the bands are

degradation products, 2) that eSrt recognizes the eSrt site and a free glycine in the cell and attaches the MSP to this cellular compound or 3) that eSrt forms a bond with the eSrt-motif, and this bond gets hydrolyzed, which in practice means the two will be separated. No solution nor explanation to this problem has been presented in previous studies (14,15). However, in a study by *Warden-Rothman et al.* a construct containing the same linker was used, but the WT sortase was utilized and did not cause any such problems, which might suggest that it was the high activity of the eSrt that was the problem (16). In the present study, the problem seemed to be partly circumvented, when the expression was performed at 37 °C, as a major band was seen in the pellet after sonication which suggested that fE3 partitioned into IBs, and was therefore protected from both degradation and eSrt reaction.

To tests these hypotheses, we designed three new constructs with shorter linkers and with either WT sortase or eSrt (see figure S1A+B). We did similar expression tests of all fusion constructs, and found that expression patterns were very similar, with the same three major bands, and the yield of fE3 expressed at 37 °C was still the highest (figure S1C). Thus, it remains elusive what causes the *in vivo* cleavage of the fusion constructs.

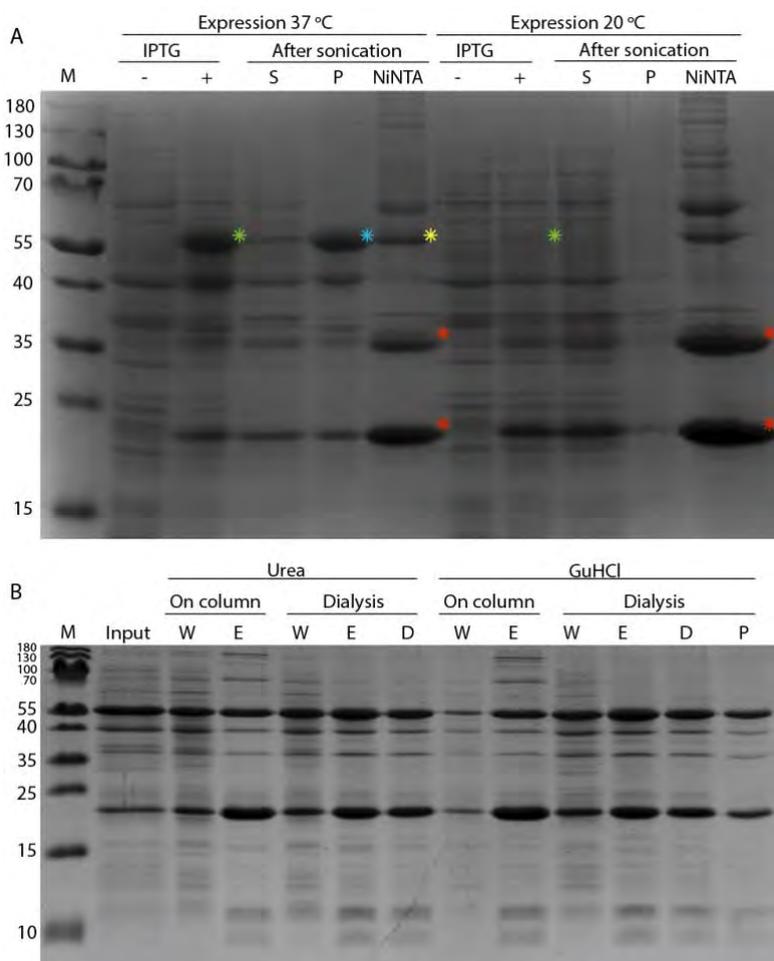


Figure 2: A: Expression of fE3 at two different temperatures and after cell lysis. S: Supernatant. P: Pellet. B: Refolding of the fE3 expressed at 37°C. W: Wash. E: Elution. D: Dialysis supernatant. P: Dialysis pellet.

Consequently, we aimed at optimizing refolding of fE3 from IBs coming from cells expressed at 37 °C. We tested a fast (on column) and a slow (dialysis) refolding protocol, starting from both urea and GuHCl solubilized IBs, respectively (figure 2B). We observed visible aggregates when performing the refolding using GuHCl in dialysis, but in general, all four methods provided soluble fE3, albeit at varying efficiencies (figure 2B, lanes E).

We proceeded with the on column refolding of fE3 from GuHCl solubilized IBs, as it is fast and similar to how MSPs are commonly purified (5,21). The refolded fE3 was subsequently applied to an ion exchange column, to remove free eSrt and impurities (see figure S2A+B). By choosing only the peak spanning fraction 3 and 4, a final yield of the fE3 was estimated to 3.9 mg per liter of culture.

### Obtaining csE3 from fE3

As eSrt activity is reported to be highly dependent on the presence of calcium (15), we investigated the efficiency of both lsE3 (with a 1:1 ratio of eSrt) (figure 3A) and fE3 (figure 3B) to form csE3 with and without calcium present over a timespan of 24 hours. The conversion from the lsE3 band, to the slightly faster migrating csE3, was dependent on calcium, as only a minor fraction was converted to csE3 after three hours without calcium (figure 3A, lane 6). After 24 hours, the majority of the lsE3 was converted to csE3 (figure 3A, lane 7), but at a much slower rate than for the reaction with calcium which was already complete after 1 hour (figure 3A, lane 9). The formation of dimers and oligomers was also evident; a side effect also observed previously using lsE3 (5).

fE3 (figure 3B) was expected to circularize slower than lsE3, as the TEV reaction must occur prior to circularization. Still, in the presence of calcium, the reaction was complete after one hour. More interestingly, the reaction without calcium was also almost complete after 3 hours, which was not the case for lsE3, even though the MSP:eSrt ratio was identical. This higher efficiency could rely on a proximity effect, as the local concentration of the substrate for the eSrt is higher when the MSP and the eSrt are fused together, than when the two have to encounter each other in solution.

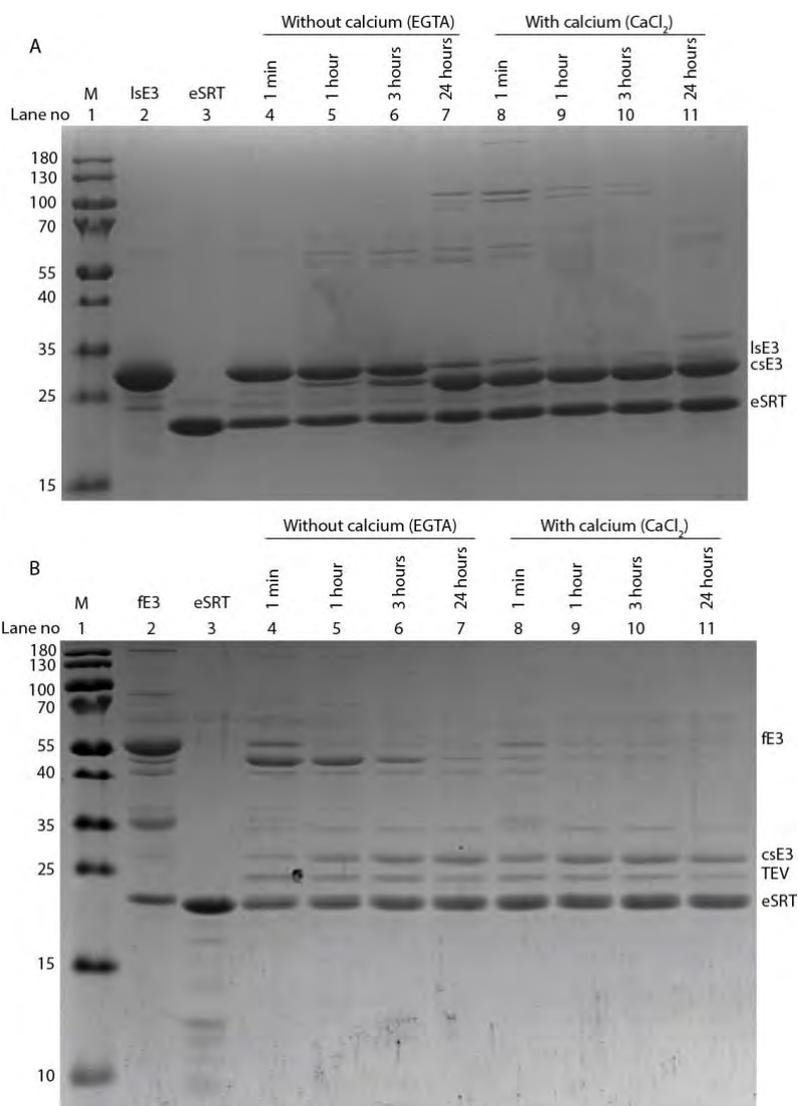


Figure 3: Circularization reaction performed either without calcium (EGTA) or with calcium (CaCl<sub>2</sub>) during a 24 hour time course. A: lsE3. B: fE3.

To estimate the efficiency of using fE3 to produce csE3, large-scale purification of fE3 was performed. TEV and CaCl<sub>2</sub> were added, and after incubation for 3 hours, a reverse NiNTA purification was performed, from which the flowthrough was dialyzed to adjust the salt content before loading it onto a HiTrap Q column (see figure S2C+D). As a fE3 band was still present in the elution (fraction 5), the reaction was incomplete, possibly due to the larger volume, but this was not further investigated. Furthermore, fE3 was not completely removed by the reverse NiNTA, suggesting more resin has to be used in this step. The peak spanning fraction 3 and 4 was the purest, and was selected for further use (figure S2C). This resulted in a final yield of 4 mg csE3 per liter of bacterial culture, which is considerably lower than that obtained by the original circularization protocol of csE3, but not much lower than obtained from circularization of other constructs (13,22). This purification can likely be optimized by longer incubation time, and by adding an extra step in the elution gradient in the IEC purification, as some of the csE3 is lost in the last fraction containing many impurities. The pooled fractions were analyzed by mass spectrometry to confirm circularization, and as seen in figure S3, the unique peptide formed during circularization was indeed detected, confirming the formation of csE3.

### Circularization of an already reconstituted nanodisc

To test whether it is possible to circularize the MSP after ND reconstitution, an ND was formed with fE3, followed by TEV cleavage and reverse NiNTA purification (figure 4A).

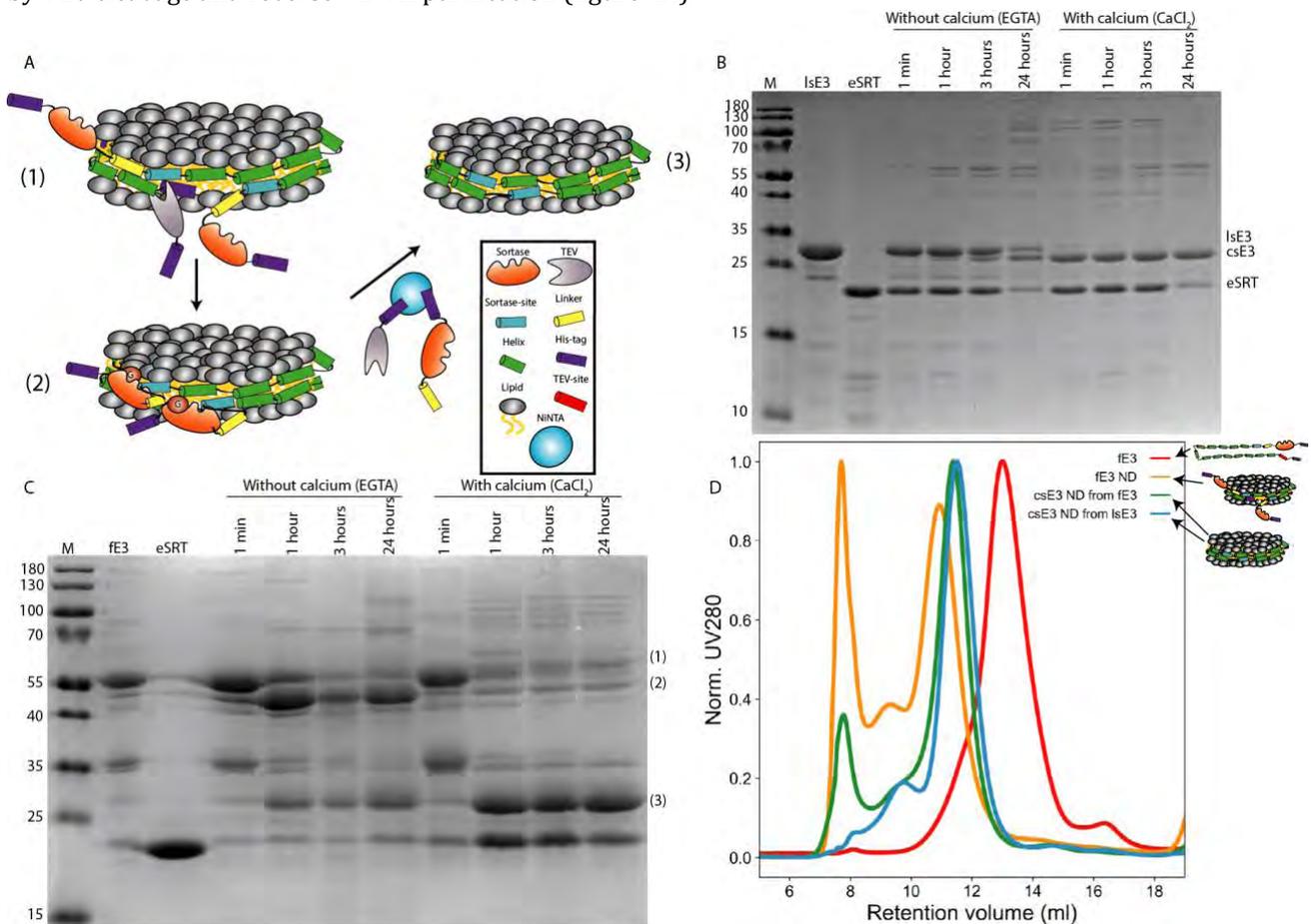


Figure 4: Circularization of a reconstituted ND. A: Schematic overview of circularization of a reconstituted ND formed with the fE3 and subsequent circularization. B: Circularization of a lsE3 ND with and without calcium present. C: Circularization of a fE3 ND with and without calcium present. D: SEC chromatogram of free fE3, a fE3 ND, and two csE3 NDs both circularized after reconstitution either from fE3 or lsE3.

As a reference, an ND was formed with lse3, and a 1:1 ratio of eSrt was added. The reactions of both samples were tested in the absence and presence of calcium. lse3 remained linear in the absence of calcium, while the process was almost instantaneous in the presence of calcium (figure 4B). The same pattern was seen for the ND formed with fe3 (figure 4C). In both cases, the formation of dimers still appeared to be a problem, which was expected, as the two MSPs in each ND are in close proximity. However, the formation of higher oligomers, albeit at a lower population, may suggest that two NDs are being linked together. SEC analysis (figure 4D) demonstrated that circularization after ND reconstitution resulted in particles with the expected size, and that the fe3 ND with the eSrt still attached formed a slightly larger particle. The csE3 NDs formed either by lse3 or fe3, eluted at the same volume, and thus appeared as though they were of similar size. The experiments in figure 4 were all conducted at 37 °C, but for this method to be relevant for more delicate MPs, lower temperatures were investigated as well. As evaluated by SDS-PAGE, it appeared possible to circularize an ND formed by either lse3 or fe3 at both 20 °C and at 10 °C within a timescale of three hours (figures S4A and B). Furthermore, we tested the importance of the eSrt:lse3 ratio, and found that even at a ratio of 1:100 (at 10 °C), the reaction was complete within three hours (figure S4C).

### Structural characterization

To investigate the structural integrity of the samples which were circularized post reconstitution, we prepared samples for small angle x-ray scattering (SAXS). SAXS is a technique providing structural information in the medium to low-resolution regime on particles in solution, and has been used to evaluate the structure of NDs in previous studies (23). The five samples presented in the top of figure 5, were all prepared and purified using SEC before X-ray irradiation. For each of the five different samples, a fraction was selected from the middle of the main SEC peak (figure 5A), and forwarded for SAXS measurement. After irradiation with X-rays and buffer subtraction, the data all contained the double bump in the  $0.1 \text{ \AA}^{-1}$  regime, characteristic for the bilayer patch in the NDs (figure 5) (23,24). This was in particular interesting for the fe3 ND (red, figure 5) as this demonstrated the presence of eSrt on the MSPs did not hinder the formation of an ND. Inspecting the  $p(r)$ -functions (figure 5C), it was evident that the five samples were quite similar, but that the fe3 ND exhibited an extra bump compared to the other four samples. This enlarged the  $D_{max}$  from  $\sim 130 \text{ \AA}$  to  $\sim 180 \text{ \AA}$ , which presumably arose from the two eSrts still attached at the periphery. According to the NMR structure of Sortase (25) the  $D_{max}$  should be in the  $40 \text{ \AA}$  range, but as eSrt is expected to bind to the eSrt-motif, it is likely more compact. This larger size of the ND also manifests in the slightly earlier elution volume, in the SEC chromatogram (figure 5A).

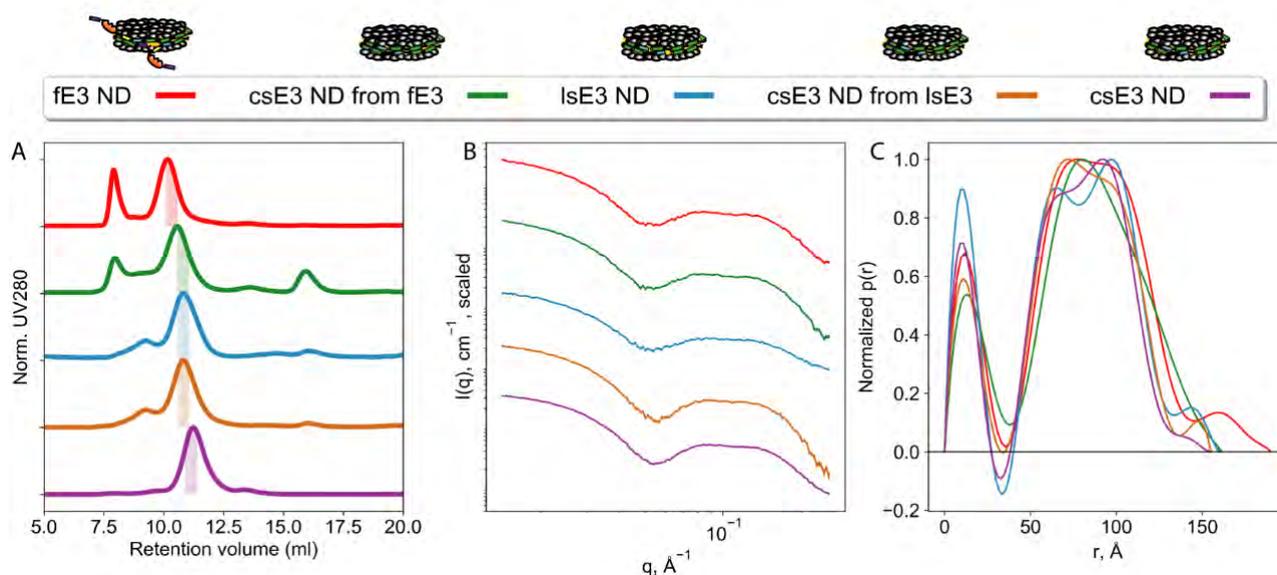


Figure 5: SAXS measurements of nanodiscs. A: SEC chromatograms of the five different samples. Marked areas are fractions selected for SAXS measurements. B: Buffer subtracted SAXS data. Scaled for clarity by a factor of  $10^n$ . C: Normalized pair distribution functions calculated using bayesapp.org.

The four remaining samples all contained MSPs (csE3 or lsE3), which have been shown to produce NDs with similar sizes, and indeed the  $p(r)$ -functions suggested a similar  $D_{max}$ -value around the expected  $\sim 130$  Å (5). It should be noted that the  $p(r)$ -functions did not descend towards 130 Å exactly, but all had a small tail, indicative of slight aggregation in the samples. The reason for the slight shift in elution volume of the csE3 ND (purple, figure 5A) was likely caused by purification on a different Äkta system, and the SAXS data (and other SEC data) also showed that it had a similar size to the three other NDs.

The structural investigation of the different ND samples, demonstrated that circularization after reconstitution indeed provided samples with similar structural characteristics, as NDs made with MSPs that have been circularized before reconstitution. In addition, the ND formed with fE3 also showed characteristics of an intact ND. However, it should be noted that the SEC runs on samples produced using the fE3, revealed considerable amounts of material in the void volume, which probably arose from the higher amount of impurities in the starting material. This can potentially result in a less optimal reconstitution, as the concentration of accessible MSP is challenging to estimate, leading to a suboptimal MSP:lipid ratio, and thus non-optimal reconstitution conditions. Additionally, the presence of impurities may also interfere with the reconstitution process.

### Incorporation of a membrane protein

To evaluate the potential of this new protocol to make circularized NDs with active MP incorporated, we used TF to investigate to which extent the structure and function of the ND was maintained. TF NDs were reconstituted with both fE3 and lsE3 using standard procedures, and after detergent removal, the samples were exposed to the same enzymatic treatment as the empty NDs. A main peak in the SEC (figure 6A) was present at the expected retention volume in all four samples (sample overview, figure 6, top), from which a fraction was selected for SAXS measurements.

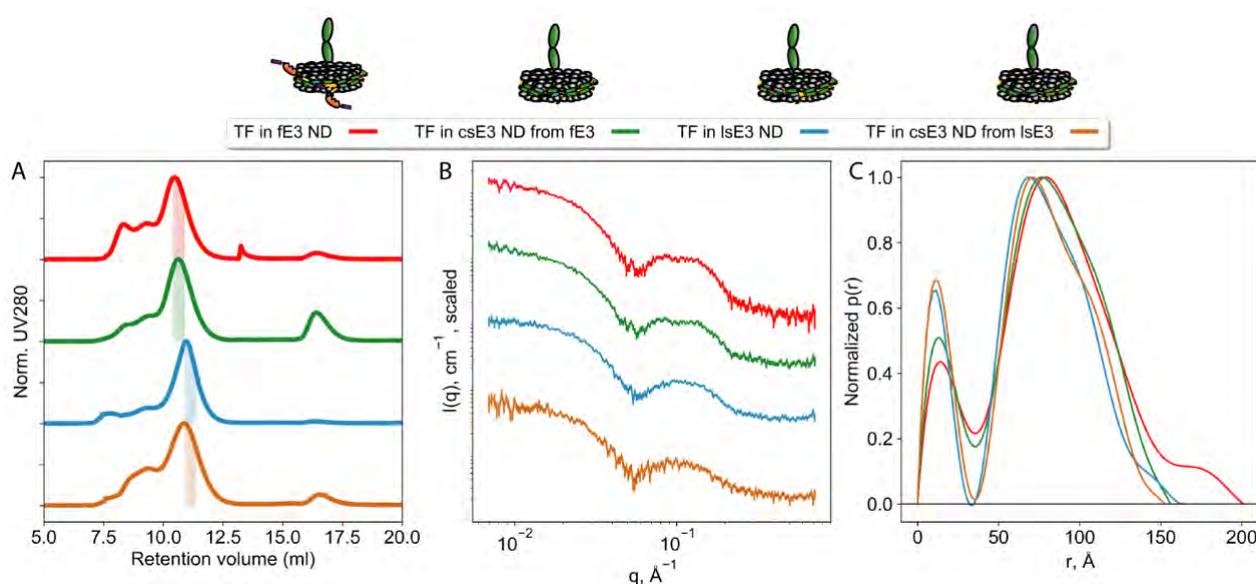


Figure 6: SAXS measurements of different nanodiscs with TF incorporated. A: SEC chromatograms of the five different samples. Marked areas are fractions selected for SAXS measurements. B: Buffer subtracted SAXS data. Scaled for clarity. C: *Normalized p(r)*-distribution functions calculated using bayesapp.org.

Empty and TF-loaded-NDs are impossible to separate during a gel filtration run, and fE3 and lsE3 samples were expected to contain empty NDs (as they still contained the His-tag), which would contribute to the SAXS signal. After buffer subtraction, the data was as presented in figure 6B. Due to a low sample concentration, the SAXS data was noisy, but still exhibited the expected features with the bump in the  $0.1$  Å<sup>-1</sup> region, in line with previous studies on TF in NDs (Paper III). Moreover, the  $D_{max}$ -values from the  $p(r)$  functions (figure 6C) were in accordance with

the expected  $\sim 150$  Å, that TF-containing-NDs have been shown to have (Paper III). Again, the fE3 ND had an increased  $D_{max}$  due to the two eSrts still attached to the MSPs.

To confirm the activity of TF in an ND circularized after reconstitution, an activity assay using the interaction with FVIIa and the complex' ability to activate FX to FXa was used. Samples were made using the same conditions as for the SAXS experiments, but only the two samples which were circularized, and could therefore be separated from empty discs, were measured. As seen in figure 7, the activities of the two samples were almost identical with  $K_a$  values fitted to be  $56 \pm 24$  pM (csE3 from lsE3) and  $94 \pm 26$  pM (csE3 from fE3) which is in agreement with other studies using the same assay (19,20), (Paper III).

## Conclusions

Obtaining high yields of circularized MSPs and incorporation of more challenging MPs are two important factors that have to be improved, in order to make the ND technology more attractive for studies on a broader range of MPs. In the present study, we present a novel method to obtain circularized MSPs and also a novel protocol for reconstituting an MP in an ND, which is subsequently circularized. A fusion construct containing both the Sortase and the MSP was developed, and we showed that it was able to produce the circularized MSP. Furthermore, we demonstrated that the fusion construct could be used to reconstitute an ND, and perform the circularization reaction afterward. Combining SAXS measurements and a functional study, demonstrated that both the structure and function of an incorporated MP is conserved using this new protocol. This protocol opens the ND technology up for MPs that struggle with incorporation into cNDs using currently available protocols. Furthermore, the fusion idea can be optimized, and potentially provide a new and more efficient way of producing circularized MSPs, if the expression and circularization processes can be controlled to a higher degree than shown in this study.

## Supporting Information

S1: Overview of fusion constructs and corresponding expression tests. S2: IEC purification of fE3 and csE3. S3: Mass spectrometry results. S4: Effects of temperature and eSrt ratio on circularization efficiency.

## Acknowledgment

The authors thank EMBL, Germany, Hamburg for the beamtime allocated at the BioSAXS beamline P12 and Hayden Mertens and Al Kikhney for assisting during the experiment. We thank DanScatt for the financial support to go to the beamtime. This work was supported by the Novo Nordisk foundation, Synergy program (LA, BBK).

## References

1. Arana MR, Fiori MC, Altenberg GA. Functional and structural comparison of the ABC exporter MsbA studied in detergent and reconstituted in nanodiscs. *Biochem Biophys Res Commun* [Internet]. 2019;512:7–11. Available from: <https://doi.org/10.1016/j.bbrc.2019.03.069>
2. Nasr ML, Baptista D, Strauss M, Sun ZYJ, Grigoriu S, Huser S, et al. Covalently circularized nanodiscs for studying membrane proteins and viral entry SI. *Nat Methods*. 2016;14(1):49–52.
3. Denisov IG, Sligar SG. Nanodiscs in Membrane Biochemistry and Biophysics. *Arter Thromb Vasc Biol*. 2017;117(6):4669–713.
4. Hagn F, Etzkorn M, Raschle T, Wagner G. Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins *Franz. J Am Chem Soc*. 2014;135(5):1919–25.

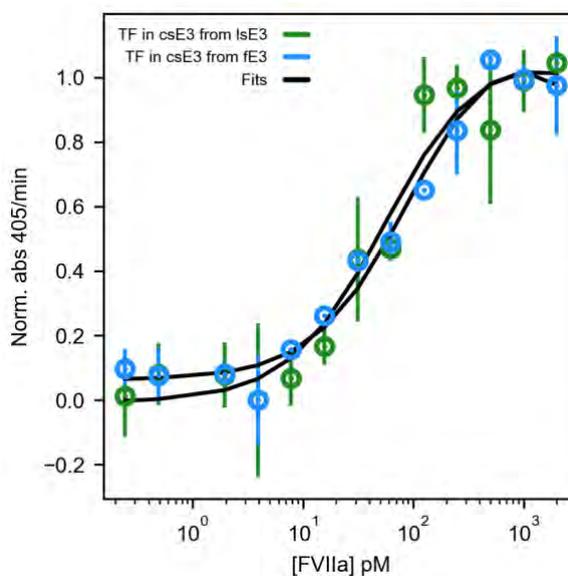


Figure 7: Activity of TF in csE3 from either lsE3 or fE3.

5. Johansen NT, Pedersen MC, Tidemand FG, Rand KD, Nguyen TTTN, Arleth L. Circularized and solubility-enhanced MSPs facilitate simple and high yield production of stable nanodiscs for studies of membrane proteins in solution. *FEBS J.* 2019;
6. Chen I, Dorr BM, Liu DR. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc Natl Acad Sci.* 2011;108(28):11399–404.
7. Yeh V, Lee TY, Chen CW, Kuo PC, Shiue J, Chu LK, et al. Highly Efficient Transfer of 7TM Membrane Protein from Native Membrane to Covalently Circularized Nanodisc. *Sci Rep.* 2018;8(1):1–11.
8. Miehl 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;19(18):1927–33.
9. Langenhan T, Aust G, Hamann J. Sticky signaling - Adhesion class g protein-coupled receptors take the stage. *Sci Signal.* 2013;6(276):1–22.
10. FOXWELL BMJ, BARRETT K, FELDMANN M. Cytokine receptors: structure and signal transduction. *Clin Exp Immunol.* 2008;90(2):161–9.
11. Vinothkumar KR, Henderson R. Structures of membrane proteins. Vol. 43, Quarterly Reviews of Biophysics. 2010. 65–158 p.
12. Denisov IG, Grinkova Y V., Lazarides AA, Sligar SG. Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *J Am Chem Soc.* 2004;126(11):3477–87.
13. 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;34(11):3525–32.
14. Jia X, Crawford T, Zhang AH, Mobli M. A new vector coupling ligation-independent cloning with sortase a fusion for efficient cloning and one-step purification of tag-free recombinant proteins. *Protein Expr Purif.* 2019;161(April):1–7.
15. Mao H. A self-cleavable sortase fusion for one-step purification of free recombinant proteins. *Protein Expr Purif.* 2004;37(1):253–63.
16. Warden-Rothman R, Caturegli I, Popik V, Tsourkas A. Sortase-tag expressed protein ligation: Combining protein purification and site-specific bioconjugation into a single step. *Anal Chem.* 2013;85(22):11090–7.
17. 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;48(2):431–43.
18. Hansen S. Update for BayesApp: a web site for analysis of small-angle scattering data. *J Appl Crystallogr.* 2014 Aug;47(4):1469–71.
19. Shaw AW, Pureza VS, Sligar SG, Morrissey JH. The local phospholipid environment modulates the activation of blood clotting. *J Biol Chem.* 2007;282(9):6556–63.
20. Waters EK, Morrissey JH. Restoring full biological activity to the isolated ectodomain of an integral membrane protein. *Biochemistry.* 2006 Mar;45(11):3769–74.
21. Bayburt TH, Grinkova Y V., Sligar SG. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* 2002;2(8):853–6.
22. Skar-Gislinge N, Johansen NT, Høiberg-Nielsen R, Arleth L. A comprehensive study of the self-assembly of phospholipid nanodiscs: What determines their shape and stoichiometry? *Langmuir.* 2018;34:acs.langmuir.8b01503.
23. 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):3161–70.
24. 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;132(39):13713–22.
25. Ilangovan U, Ton-that H, Iwahara J, Schneewind O, Clubb RT. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*.pdf. *Proc Natl Acad Sci.* 2001;

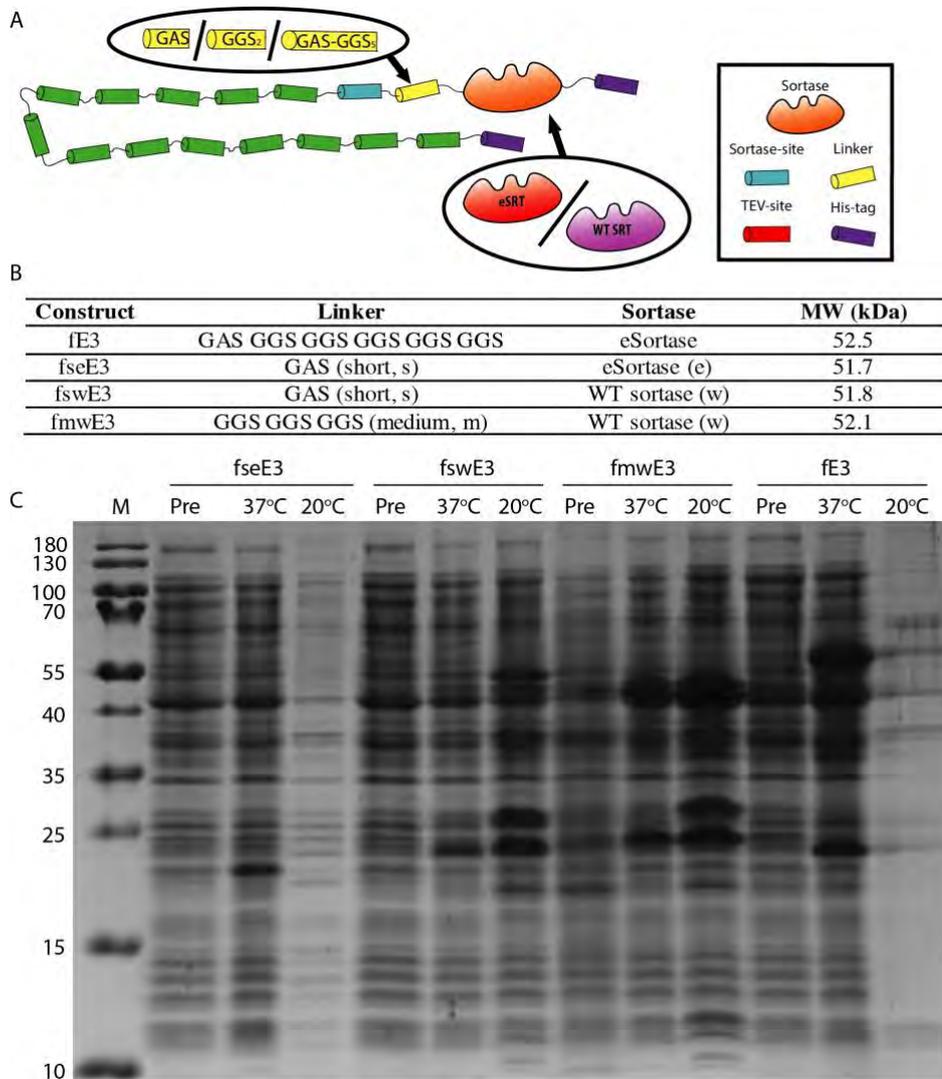
## Supplementary information

### Figure S1: Overview of fusion constructs

A: Illustration of the different fusion constructs.

B: Table of the specific fusion constructs and their molecular weights.

C: Expression gel of the fusion constructs used in this study. Expression was conducted both at 37°C and 20°C. Total cellular lysates were loaded on the gels from the time of harvest or just prior to induction with IPTG.



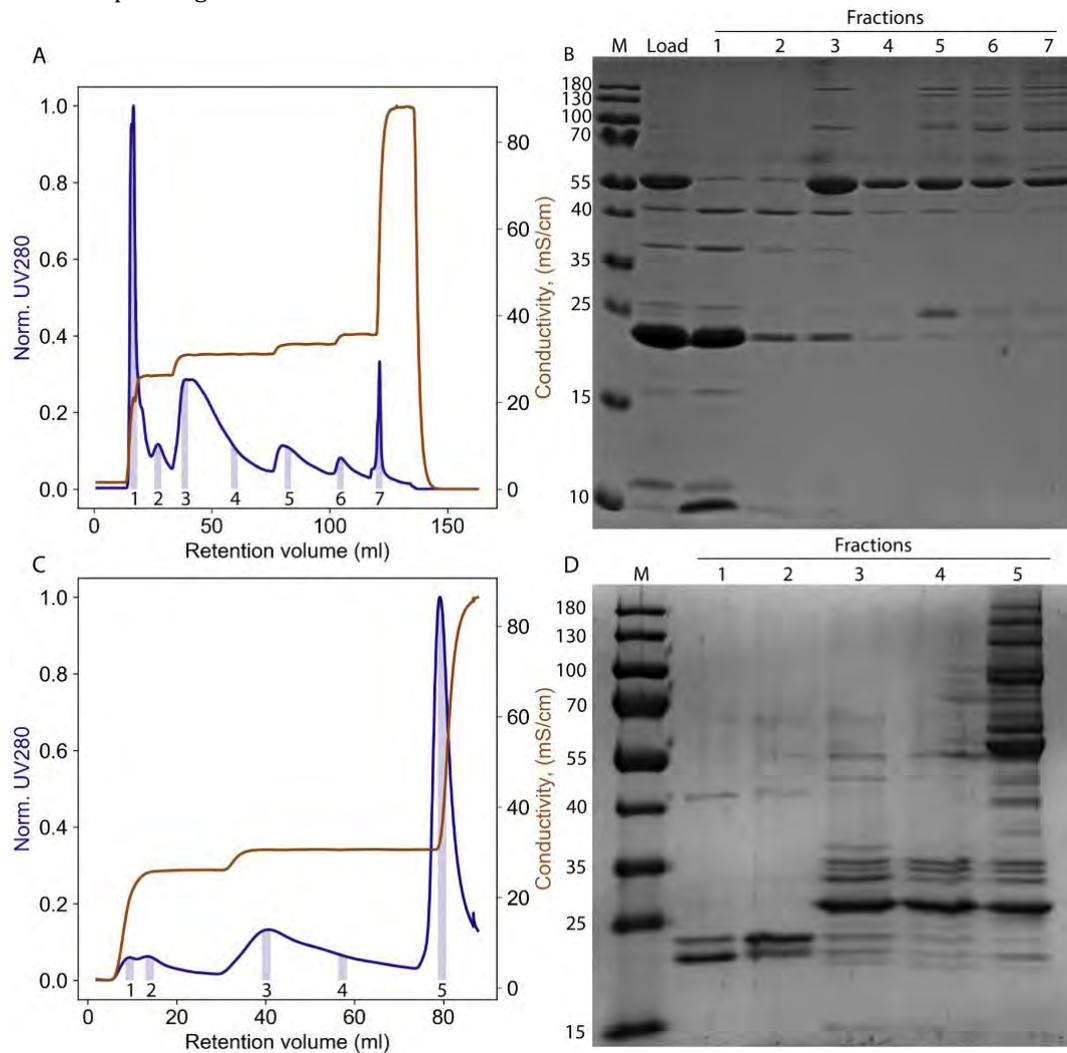
**Figure S2: IEC purification of fE3 and cSE3.**

A: IEC chromatogram of the fE3. Blue areas are fractions loaded on gel in B.

B: Gel of the corresponding fractions from A.

C: IEC chromatogram of the cSE3 from fE3. Blue areas are fractions loaded on gel in D.

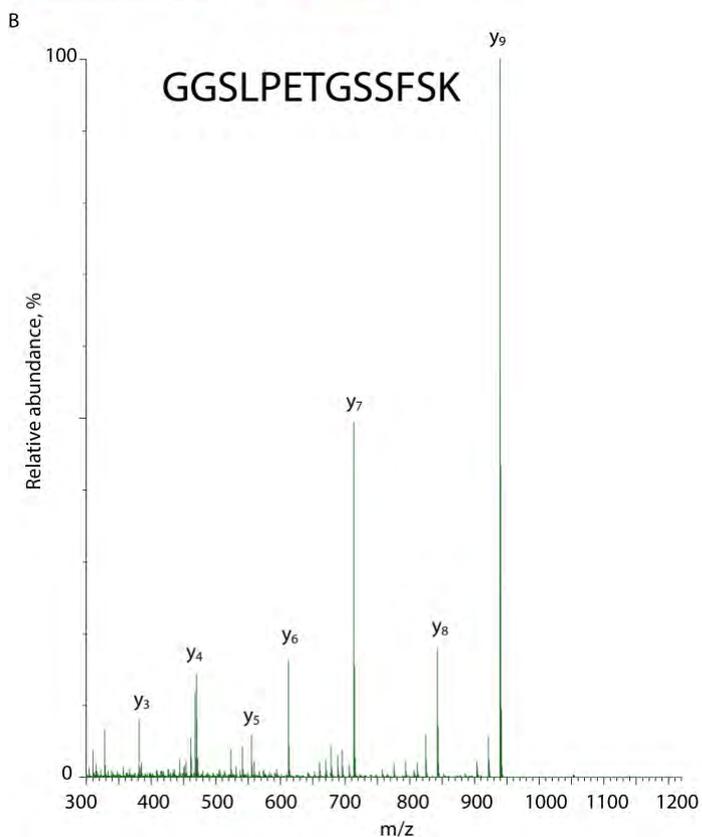
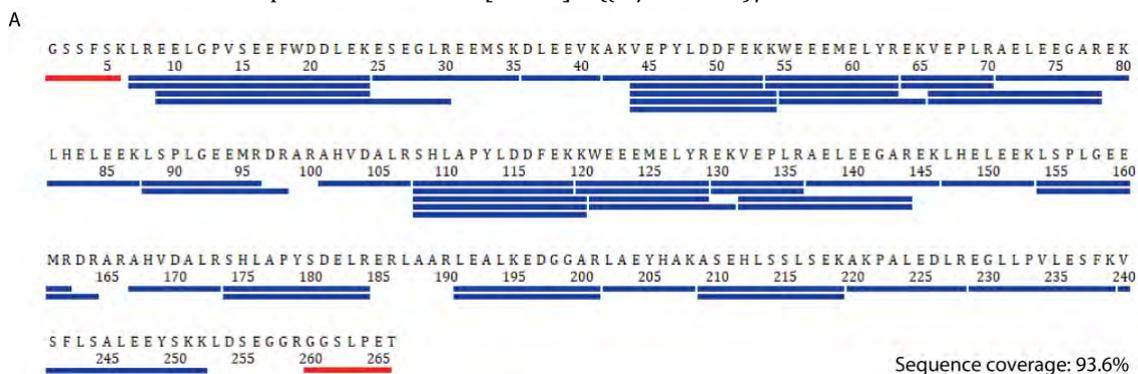
D: Gel of the corresponding fractions from C.



**Figure S3: Mass spectrometry**

A: MS coverage map: Peptide map from LC-MS/MS analysis of csE3 (from fE3) after trypsin digestion. Blue bars indicate peptides identified by MS/MS (CID). Red bar indicate the unique peptide formed by the circularization reaction.

B: MS/MS fragmentation spectrum of the tryptic peptide corresponding to the peptide originating from the circularization reaction. The precursor ion was  $[M+2H]^{2+}$  ( $m/z$  627.32).

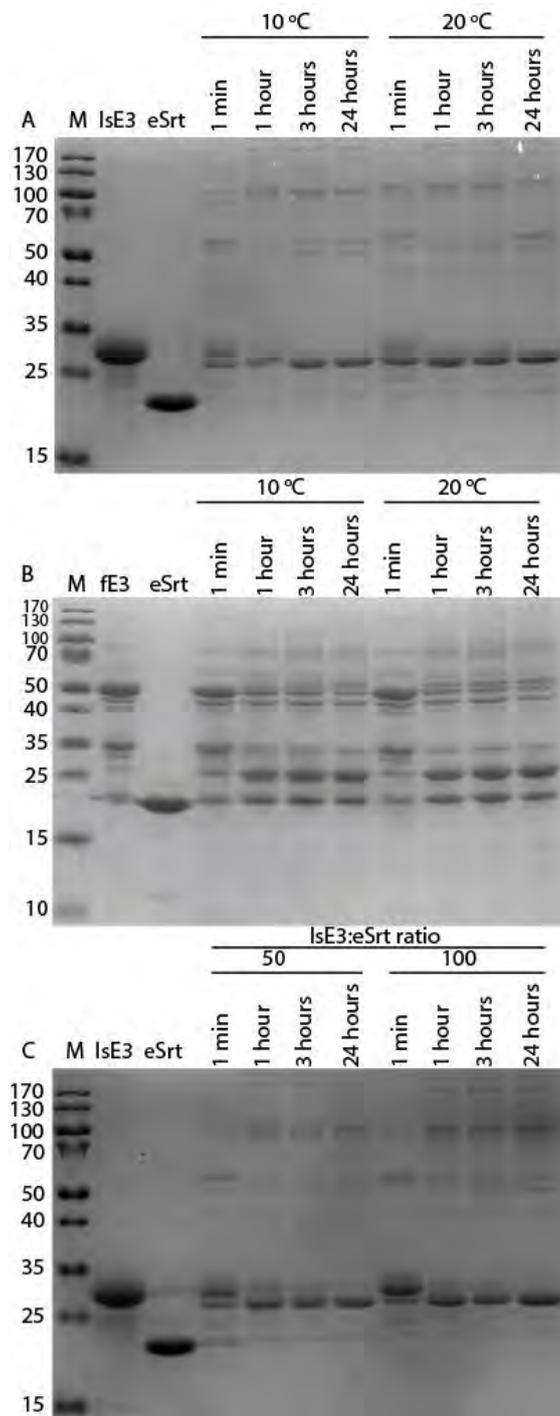


**Figure S4: Effects of temperature and eSrt ratio on circularization.**

A: Circularization of lsE3 with a 1:1 ratio of eSrt at 10°C or 20°C.

B: Circularization of fE3 at 10°C or 20°C.

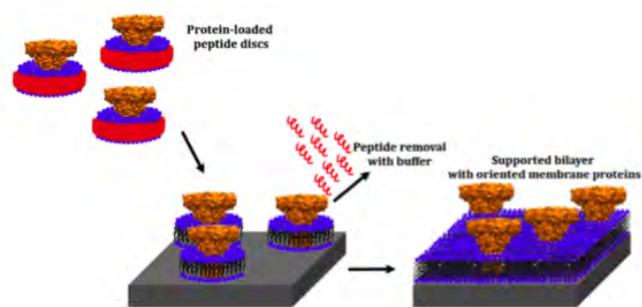
C: Circularization of lsE3 at 10°C with lsE3:eSrt ratio of either 50 or 100.



## 7 | Paper III

### Peptide disc mediated control of membrane protein orientation in supported bilayers for surface-sensitive investigations.

*Alessandra Luchini, Frederik Grønæk Tidemand, Nicolai Tidemand Johansen, Mario Campana, Javier Sotres, Michael Ploug, Marité Cárdenas, Lise Arleth*



## Peptide disc mediated control of membrane protein orientation in supported bilayers for surface-sensitive investigations.

Alessandra Luchini <sup>\*†</sup>, Frederik Grønbæk Tidemand <sup>†</sup>, Nicolai Tidemand Johansen <sup>†</sup>, Mario Campana <sup>‡</sup>, Javier Sotres <sup>⊥</sup>, Michael Ploug <sup>#,||</sup>, Marité Cárdenas <sup>⊥</sup> and Lise Arleth <sup>\*†</sup>

<sup>†</sup> Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

<sup>‡</sup> ISIS-STFC, Rutherford Appleton Laboratory, Chilton, Oxon OX11 0QX, United Kingdom

<sup>⊥</sup> Biofilms Research Center for Biointerfaces and Department of Biomedical Science, Faculty of Health and Society, Malmö University, Per Albin Hanssons Väg 35, 214 32 Malmö, Sweden

<sup>#</sup> Biotech Research & Innovation Center, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark

<sup>||</sup> Finsen Laboratory, Rigshospitalet, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark

**ABSTRACT:** *In vitro* characterization of membrane proteins requires experimental approaches providing mimics of the microenvironment that proteins encounter in native membranes. In this context, supported lipid bilayers provide a suitable platform to investigate membrane proteins by a broad range of surface-sensitive techniques such as Neutron Reflectometry (NR), Quartz Crystal Microbalance with Dissipation monitoring (QCM-D), Surface Plasmon Resonance (SPR), Atomic Force Microscopy (AFM) and fluorescence microscopy. Nevertheless, the successful incorporation of membrane proteins in lipid bilayers with sufficiently high concentration and controlled orientation relative to the bilayer remains challenging. We propose the unconventional use of peptide discs made by phospholipids and amphipathic 18A peptides to mediate the formation of supported phospholipid bilayers with two different types of membrane proteins, CorA and Tissue Factor (TF). The membrane proteins are reconstituted in peptide discs, deposited on a solid surface and the peptide molecules are then removed with extensive buffer washes. This leaves a lipid bilayer with a relatively high density of membrane proteins on the support surface. As a very important feature, the strategy allows membrane proteins with one large extramembrane domain to be oriented in the bilayer, thus mimicking the *in vivo* situation. The method is highly versatile and we show its general applicability by characterizing with the above mentioned surface sensitive techniques two different membrane proteins, which were efficiently loaded in the supported bilayers with ~0.6% mol/mol (protein/lipid) concentration corresponding to 35% v/v for CorA and 8% v/v for TF. Altogether, the peptide disc mediated formation of supported lipid bilayers with membrane proteins represents an attractive strategy for producing samples for structural and functional investigations of membrane proteins, and for preparation of suitable platforms for drug testing or biosensor development.

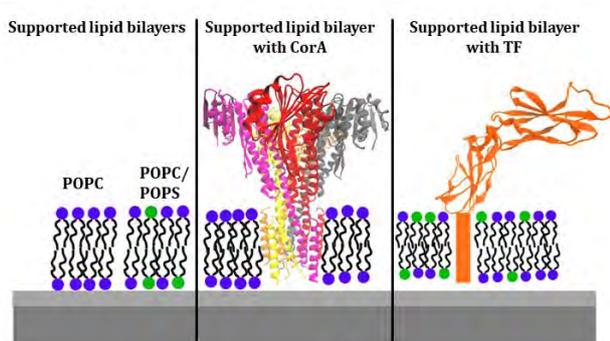
Membrane proteins are fundamental cell membrane components with vital biological functions. For this reason, membrane proteins are relevant therapeutic targets in various diseases and about 60% of the marketed drugs target membrane proteins<sup>1</sup>. In addition, technological applications of membrane proteins, such as receptors or ion channels, are highly attractive for the development bio-inspired functional materials with diagnostic applications<sup>2-3</sup>. In such materials the protein biological functions are exploited for the interaction/capture of a target analyte. Despite their biological relevance and technological importance, relatively little structural and functional information are available on membrane proteins compared to soluble proteins. Membrane proteins are generally difficult to study with standard biophysical methods; they have a poor tendency to crystallize, which complicates X-ray crystallography structural determination<sup>4</sup>. On top of this, it is well-known that the native structure of a membrane protein is strongly dependent on its

surrounding lipid environment, which introduces further complexity in sample preparation and limits the choice of the experimental methods for probing the protein structure<sup>5-6</sup>. Besides the general requirement for a lipid matrix, there are increasing experimental evidences suggesting that specific lipid compositions also drive the functional profile of membrane proteins<sup>7-8</sup>. Membrane protein orientation and the distinction between intra- and extracellular is another important feature of cell membranes. In the cell, all copies of a given membrane protein have the same orientation of their extramembrane domains with respect to the biomembrane<sup>9</sup>. Hence controlling membrane protein orientation is particularly relevant to simulate the native protein arrangement of the cell membrane<sup>10</sup>.

Supported membranes composed of a lipid bilayer in the proximity of a solid support represent an attractive strategy for reproducing the native lipid environment of membrane proteins<sup>11-12</sup>. Supported lipid bilayers are conveniently

formed in the proximity of a given solid support<sup>13-18</sup>, where they enable the assembly of a plethora of different lipid compositions entailing detailed structural and functional characterizations by a broad range of surface-sensitive techniques, e.g. Neutron Reflectometry (NR), Quartz Crystal Microbalance with Dissipation monitoring (QCM-D), Surface Plasmon Resonance (SPR), Atomic Force Microscopy (AFM) and fluorescence microscopy<sup>14, 19-22</sup>.

While conventional applications of supported lipid bilayers are often limited to the characterization of the interaction between lipid membranes and soluble proteins or peptides<sup>23-28</sup>, a central challenge in this scientific field is to produce more complex cell membrane mimics where different lipid species and membrane proteins can be readily included and correctly oriented<sup>29</sup>. Besides providing a suitable membrane model for membrane protein characterization, supported lipid bilayers with membrane proteins are also relevant systems for biosensor design<sup>3</sup>.



**Figure 1:** Schematic representation of the characterized samples. 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) molecules are represented in blue while 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS) molecules are represented in green. The unknown part of the TF structure is schematized as a rectangle inside the lipid bilayer. [Protein Data Bank ID: 4eeb, 1dan].

A few examples are available where membrane proteins were successfully incorporated in lipid bilayers in the proximity of a solid surface<sup>30-33</sup>. The most common strategies are based either 1) on the deposition of proteoliposomes<sup>32, 34</sup>, i.e. protein-loaded liposomes, on solid supports or 2) on the chemical anchoring of detergent-reconstituted proteins on the support surface<sup>35-39</sup>. The first approach is not suitable for proteins exhibiting large extramembrane domains and does not allow the protein orientation in the membrane to be controlled. In the second approach, the two membrane components, i.e. the lipids and the protein, are added in two separate steps, which requires the use of detergents to stabilize the proteins before their reconstitution into the lipid bilayer<sup>33, 35, 40</sup>. This approach requires a proper chemical modification of the solid support surface, which often involves gold coating, and a surface functionalization with specific molecular groups for the protein anchoring<sup>41</sup>. A popular approach is to insert a histidine tag (His-tag) in the primary sequence of the membrane protein such that it can be anchored to the support surface with a defined orientation. One of the main advantages of the method is that a linker can be designed to produce sufficiently large space between the membrane

and the support surface to fit large protein extramembrane domains. Although anchoring of detergent-reconstituted membrane proteins is a successful strategy for controlling their orientation in the lipid bilayer<sup>37-38</sup>, the method involves chemical modifications of both the support surface and the protein sequence, which limits its application to specific surfaces, e.g. using gold-coated supports, and hence specific surface-sensitive techniques and to membrane proteins that can be tagged e.g. through genetic engineering. Other methods for anchoring proteins on solid surfaces involve the use of antibody functionalized surfaces<sup>42</sup>. Besides requiring a specific antibody to be designed, produced and suitably anchored on the support surface, this is also not a general method easily adaptable for all surface-sensitive techniques.

Here, we propose a versatile and efficient method for the preparation of supported lipid bilayers with oriented membrane proteins. The strategy allows membrane proteins with one large extramembrane domain to be oriented in the supported bilayer such that all the reconstituted membrane protein molecules have the large extramembrane domain pointing towards the bulk solvent. It also avoids any sophisticated modification of the support surface or the introduction of a His-tag in the protein sequence.

Supported lipid bilayers with incorporated and oriented membrane proteins are self-assembled through the unconventional use of a specific kind of nanodiscs, named peptide discs<sup>43</sup>, that we use as carriers of both the lipids and the membrane proteins to the support surface. Nanodiscs are discoidal lipid bilayers, typically with a diameter around 8-14 nm, that are surrounded and stabilized by an amphipatic protein belt, the so-called membrane scaffold protein, MSP<sup>44-46</sup>. MSP based nanodiscs are extensively used to stabilize membrane in structural studies<sup>47-50</sup>. In this context, nanodiscs have also been self-assembled to form a film on a solid support surface for the structural investigation of membrane proteins by means of surface-sensitive techniques<sup>51-52</sup>. Although membrane proteins could be successfully oriented in such films, the low surface coverage and the presence of the belt protein within the film were the major limitation of the system.

In contrast to these studies, the present work builds on the application of peptide discs where the nanodisc belt is self-assembled by amphipatic 18A peptide molecules while the interior of the disc still contains a phospholipid bilayer as in the original MSP-based nanodiscs<sup>43</sup>. Such peptide discs have previously been proposed for the reconstitution of membrane proteins<sup>43, 53-54</sup>, but they proved less stable over time than classical MSP-based nanodiscs<sup>43, 53</sup>. While this instability is clearly a limitation in many applications of the nanodiscs, is the key property that we exploit in the present work for the assembly of supported lipid membranes. The formation of the lipid and membrane protein containing peptide disc is a reversible process which allows the lipids and the membrane proteins to be released on the support surface to form a supported lipid bilayer with incorporated membrane proteins. We have used the peptide disc deposition on solid supports to produce and characterize supported lipid bilayers with different lipid compositions as well as the supported lipid bilayers with two structurally and functionally different membrane proteins, i.e. CorA and Tissue Factor (TF) (Figure 1).

CorA is a homopentameric magnesium transporter of known structure<sup>55-57</sup> which is naturally found in bacterial membranes<sup>58</sup>. CorA was used to initially test the effective formation of the supported bilayer with oriented membrane proteins. The other test system, Tissue Factor (TF), is a monomeric membrane protein responsible for the initiation of the blood coagulation process<sup>59</sup>. Structural information is so far available only for the TF extracellular domain<sup>60-61</sup>. Besides a few FRET studies<sup>62-63</sup>, to the best of our knowledge there are no experimental data on the structure of the full-length protein in lipid membranes.

We tested the method for three different types of solid supports; silicon oxide, mica and gold, and show that for these surfaces the method works independently from the specific solid support composition and only requires its surface to be hydrophilic. The method also avoids any sophisticated modification of the support surface or the introduction of a His-tag in the protein sequence which minimizes the time and number of steps required for sample preparation. Overall, we found that the nature and quality of the supported bilayers with CorA or TF comply well with the work-flow and sample requirements of NR, QCM-D, AFM and SPR as examples of different surface sensitive techniques. Hence, the presented method provides a highly flexible experimental strategy, which can easily be accommodated in the biophysical characterization of different membrane proteins with complementary surface-based techniques.

## EXPERIMENTAL SECTION

### Chemicals

POPC ( $\geq 99\%$  purity), 1-palmitoyl-d31-2-oleoyl-*sn*-glycero-3-phosphocholine (dPOPC,  $\geq 99\%$  purity) and POPS ( $\geq 99\%$  purity) were purchased from Avanti Polar Lipids, Inc. (Alabaster, AL) and used without further purification. 18A peptide (DWLKAFYDKVAEKLKEAF acetylated and amidated at the N-term and C-term respectively,  $>95\%$  purity) was purchased from GenScript and used without further purification. Heavy water ( $D_2O$  99.9% purity), chloroform ( $\geq 99.5\%$  purity), ethanol (98% purity), methanol (99.8% purity), were purchased from Sigma-Aldrich. Ultrapure Milli-Q (MQ) water with resistivity of 18.2  $M\Omega \cdot cm$  at 25 °C was used for all cleaning procedures, and preparation of all hydrogenated samples and buffers. The activated factor VII (FVIIa) used in this study was provided by NovoNordisk A/S and used without further purification.

### Protein expression and purification

Professor Mikaela Rapp (University of Stockholm, Sweden) kindly provided the CorA encoding plasmid. Expression and purification was performed essentially as described elsewhere<sup>64</sup>. CorA was concentrated to a concentration sufficient for the peptide disc reconstitution and flash frozen in liquid nitrogen and stored at -80 °C until use.

Recombinant TF (rTF), containing human TF (1-244) with a linker (G<sub>4</sub>S) and a His<sub>6</sub>-tag C-terminally was synthesized and cloned into the pET 28a vector by Genscript. A more in-detail protocol is included in a manuscript in preparation, but in brief the expression and purification were carried out as follows: Expression was carried out in BL21 Star cells, which were grown at 37°C in Terrific Broth medium.

Expression was induced with IPTG to a final concentration of 1 mM at an optical density at 600 nm of 0.8 and the cells were harvested after 3 hours. The pellet was resuspended in 50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 1 mM PMSF and passed through a cell disrupter (E1061, Constant Systems LTD) at 30kPsi. The sample was centrifuged at 25000 *g* for 20 min and the pellet was resuspended in 20 mM Tris, pH 8, 150 mM NaCl, 0.5 M Urea and 0.25 % (w/v) Tween20. After centrifugation at 25000 *g* for 20 min the pellet was resuspended in 6 M GuHCl, 10 mM Tris-HCl, pH 8, 50 mM NaCl, 2 mM EDTA and 0.1% Triton X-100 (TX). After centrifugation at 25000 *g* for another 20 min, the supernatant was dripped into a reservoir of 50mM Tris-HCl, pH 8.5, 250mM NaCl, 0.1% TX containing 60 times the volume of the supernatant under fast magnetic stirring. The solution was left in a cold room for 24 hours under magnetic stirring. Imidazole solution (pH=8) was added to a final concentration of 20 mM and the preparation was centrifuged at 17000 *g* for 20 min. The supernatant was loaded onto a 5 ml HisTrap column (GE Healthcare), washed with 25 ml of 20 mM Tris-HCl, pH 8, 100 mM NaCl, 0.1% TX and 40 mM imidazole and eluted with 25 ml 20 mM Tris-HCl, pH 8, 100 mM NaCl, 0.1% TX and 250 mM Imidazole. Protein containing fractions were pooled and dialyzed against 20 mM Tris, pH 8, 100 mM NaCl, 0.05 % TX.

### Peptide disc preparation and supported bilayer formation

According to the protocol described elsewhere<sup>43</sup>, the peptide discs were prepared by mixing 18A and the phospholipids (POPC or POPC/POPS) in methanol with 1:1 w/w ratio. The methanol solution was dried in a glass vial and in the case of the peptide discs without membrane protein, the film was subsequently re-dissolved with buffer (20mM Tris-HCl pH=7.5, 100mM NaCl, 20mM CaCl<sub>2</sub>). Membrane protein loaded peptide discs were formed by dissolving the 18A-lipid film with a membrane protein solution as described previously<sup>43</sup>. In order to produce the CorA loaded peptide discs a solution of 12 $\mu$ M CorA (20mM Tris-HCl pH=7.5, 150mM NaCl, 0.5mM EDTA, 0.03% DDM) was added to the POPC-18A film. As shown in Figure S11, the peptide discs and the CorA-loaded peptide discs have sufficiently different sizes to be readily separated on a Superdex 200 10/300 increase column. The collected fractions for CorA-loaded peptide disc suspension (see Figure S1) were used for QCM-D and NR experiments. On the other hand, a dilution 1:2 v/v of the sample was applied for the AFM image collection. This allowed for a better visualization in the image of the protein molecules compared to the surrounding lipid membrane.

The above described preparation for the CorA-loaded peptide discs was used as well for the preparation of rTF-loaded peptide discs. The produced samples were successfully used for the collection of the QCM-D and SPR data, which shows a high sensitivity to rTF as the binding constant with FVIIa has very low  $K_D$ .

The peptide discs and rTF-loaded peptide discs eluted together from the Superdex 200 10/300 column in a single broad peak (Figure S1). rTF (~28kDa) is considerably smaller than CorA (~43kDa for the monomer, ~215kDa for the CorA pentamer). Hence, rTF-loaded peptide discs are expected to have a size that is not very different from the empty peptide discs. Injection on solid support of the rTF-

loaded peptide disc solution purified by SEC did not produce a sufficiently high protein loading of the membrane in order to allow rTF structure to be characterized by NR. In order to increase the final TF:lipid ratio, the protocol for the rTF-loaded peptide disc preparation was modified as follows. The TF peptide disc preparation was loaded onto a 1 ml HisTrap column (GE Healthcare) equilibrated in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 20 mM Imidazole and eluted with 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 250 mM Imidazole while collecting fractions of 0.5 ml. Fractions containing loaded peptide discs were pooled and frozen until use at the experiments.

To confirm the presence of CorA and TF in the peptide discs, the samples were run on a 15% acryl amide gel. 40  $\mu$ l of each peptide disc sample were mixed with 10  $\mu$ l of 250 mM Tris-HCl pH 6.8, 50% glycerol, 10% w/v SDS, 0.25% w/v bromophenol blue and boiled for 5 min at 95°C before loading on the gel. Gels were stained with Coomassie brilliant blue G250 (Bio-Rad, Hercules, CA, USA) in 20% ethanol, 10% acetic acid and images of the gels were obtained on a Gel Doc (Bio-Rad). The same protocol for the formation of the supported lipid bilayers was adopted for all of the performed experiments (NR, QCM-D, AFM, SPR). Briefly, the peptide disc solution was injected in the cell containing the solid support. The injected volume varied according to the size of the specific instrument cell, i.e. 2 ml for NR, 0.25 ml for QCM-D, 0.5ml for AFM and 0.150ml for SPR. After 10-15 min incubation, the buffer was flushed through the cell (10-20min) to remove the excess of peptide discs from the solution as well as to remove the peptide molecules from the deposited lipid bilayer. The same protocol was used as well in the case of protein-loaded peptide discs.

### Neutron Reflectometry (NR)

NR experiments were performed on the SURF reflectometer at ISIS neutron source, Chilton, (UK). SURF is a horizontal time of flight reflectometer. Three incoming angles ( $\theta$ ) typically of 0.35°, 0.65° and 1.5° were used to cover the  $q$ -range  $0.01 \text{ \AA}^{-1} < q < 0.2 \text{ \AA}^{-1}$ , where  $q$  is defined as follows:

$$q = \frac{4\pi}{\lambda} \sin(\theta) \quad (2)$$

The measured reflected intensity ( $I(q)$ ) was converted in an absolute reflectivity scale ( $R(q)$ ) by normalization to the direct beam ( $I_0$ ) measured at the same slit settings. Slits were chosen to vary with the incident angle in such a way as to provide a constant illumination of the sample (30  $\times$  60 mm).

$$R(q) = \frac{I(q)}{I_0} \quad (3)$$

The scattering length density profile ( $\rho(z)$ ) can be calculated from the experimentally determined reflectivity profiles (equation 4). This gives information on the composition of the sample along the surface normal ( $z$ )<sup>65-66</sup>.

$$\rho = \sum_i \frac{n_i b_i}{V_m} \quad (4)$$

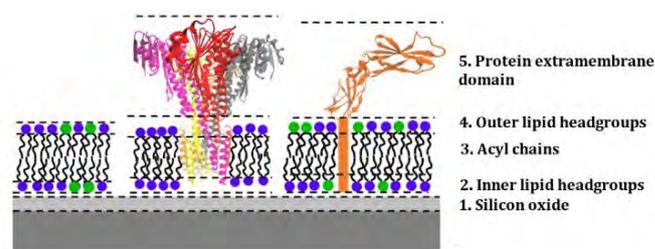
As reported in equation 4,  $\rho$  depends on the chemical and isotopic composition of the sample as the neutrons are sensitive to the nuclei composing the atoms in the molecules, where  $n_i$  is the number of atoms  $i$ ,  $b_i$  is the coherent scattering length, and  $V_m$  is the partial molecular volume.

NR was performed using solid/liquid flow cells with polished silicon crystals (111) with a surface area of 5  $\times$  8 cm. Neutron cell flow modules and O-rings were cleaned by bath sonication in 2%(v/v) Hellmanex and rinsed with MQ. The crystals were cleaned with dilute piranha solution; 1:4:5 H<sub>2</sub>O<sub>2</sub>:H<sub>2</sub>SO<sub>4</sub>:H<sub>2</sub>O at 80 °C for 15 min.

The temperature was maintained at 25 °C by circulating water from a thermostated water bath. Variation of the aqueous solvent contrast was achieved by exchanging the bulk solvent using an HPLC pump set to a flow rate of 2mL/min. Support surfaces were characterized in H<sub>2</sub>O and D<sub>2</sub>O. The samples were characterized in at least 3 isotopic solvent contrasts, i.e. buffer with different ratio of D<sub>2</sub>O to H<sub>2</sub>O. Specifically, the buffers used during the experiments were prepared with pure D<sub>2</sub>O (d-buffer,  $\rho = 6.35 \cdot 10^{-6} \text{ \AA}^{-2}$ ), D<sub>2</sub>O:H<sub>2</sub>O 38:62 w/w (Silicon Matched Water, SMW-buffer  $\rho = 2.07 \cdot 10^{-6} \text{ \AA}^{-2}$ ) and pure H<sub>2</sub>O buffer (h-buffer,  $\rho = -0.56 \cdot 10^{-6} \text{ \AA}^{-2}$ ).

### NR data analysis

NR data were analyzed with RasCAL<sup>67</sup>. In this software environment, the interface between the support and the bulk is considered as a stratified medium composed by different slabs. The reflectivity originating from such an interface can be described according to the optical matrix method<sup>68</sup>, where each of the slabs included in the model is characterized by four output parameters: the layer thickness ( $t$ ), the scattering length density ( $\rho$ ), the volume fraction of solvent ( $\phi_s$ ) and the surface roughness ( $r$ ). Figure 7 shows a schematic representation of the models used for the characterized samples. The experimental data were analyzed by creating a custom model within RasCAL which allows for defining custom fitting parameters that can be directly optimized to the experimental data.



**Figure 7**-Schematic representation of the sample structures as composed by a stack of different layers. POPC molecules are represented with blue colored headgroups, while POPS molecules are represented with green colored headgroups.

In the case of the supported bilayers with rTF and CorA, the  $\rho$  for the acyl chain region was re-parametrized to directly account for the protein TMD volume fraction,  $\phi_{TMD}$  (equation 5).

$$\rho_{Acyl\ chains} = \phi_{TMD} \rho_{TMD} + (1 - \phi_{TMD}) \rho_{lipid\ acyl\ chains} \quad (5)$$

In equation 5,  $\rho_{TMD}$  was calculated according to the protein structure and atomic composition ( $\rho_{TMD} = 1.58 \cdot 10^{-6} \text{ \AA}^{-2}$  for CorA-TMD and  $\rho_{TMD} = 1.15 \cdot 10^{-6} \text{ \AA}^{-2}$  for rTF-TMD), while  $\rho_{lipid\ acyl\ chains}$  was estimated from the data collected on the pure lipid bilayers ( $\rho_{lipid\ acyl\ chains} = -0.29 \cdot 10^{-6} \text{ \AA}^{-2}$ )<sup>69</sup>. In order to reduce the number of free parameters in the model, the

protein volume fraction in layer 3 (acyl chain region) was calculated as reported in equation 6, where  $t_5$  and  $t_3$  are the thicknesses of layer 3 and 5, while  $\phi_{extramembrane}$  and  $V_{extramembrane}$  are respectively the volume fraction and the molecular volume of the protein extramembrane domains.

$$\phi_{TMD} = \phi_{extramembrane} \frac{t_5 V_{TMD}}{t_3 V_{extramembrane}} \quad (6)$$

Equation 6 restricts the protein volume fraction in layer 3 and 5 of Figure 7 to reflect the same number of protein molecules in the two layers.

#### Quartz Crystal Microbalance with Dissipation monitoring (QCM-D)

QCM-D was performed with a Q-Sense E4 instrument (Q-Sense, Biolin Scientific AB, Sweden), using SiO<sub>2</sub>-coated 5 MHz quartz sensors. Crystals and O-rings were placed in Hellmanex 2% for 10 min, extensively flushed with ethanol and MQ, and then dried under a nitrogen flow. Immediately before use, the crystals were treated with a UV ozone cleaner (BioForce Nanosciences, Inc., Ames, IA) for 20 min. Before acquisition, the fundamental frequency and six overtones (3<sup>rd</sup>, 5<sup>th</sup>, 7<sup>th</sup>, 9<sup>th</sup>, 11<sup>th</sup> and 13<sup>th</sup>) were recorded and the system was equilibrated in MQ at 25 °C, until stable baselines were obtained. After further equilibration in buffer, the samples were introduced in the flow cell at 0.2 mL/min.

For homogeneous thin and rigid films fully coupled to the sensor surface, the recorded frequency shifts, normalized to the overtone number, can be simply related to the absorbed mass ( $\Delta m$ ) through the Sauerbrey equation (equation 1), where  $C$  is the mass sensitivity constant corresponding to 17.7 ng·cm<sup>2</sup>·Hz<sup>-1</sup> for the sensors used in this experiment.<sup>70</sup>

$$\frac{\Delta F_n}{n} = -\frac{1}{C} \Delta m \quad (1)$$

During the experiments, real-time shifts in the resonance frequency ( $\Delta F_n$ ) with respect to the calibration value were measured for different overtones indicated as  $F_n$ , with  $n$  representing the overtone number. Simultaneously, also the energy dissipation factor ( $D$ ) was evaluated for all the overtones<sup>71</sup>.

#### Atomic Force Microscopy

Atomic force microscopy measurements were carried out using a commercial setup equipped with a liquid cell (MultiMode 8 SPM with a NanoScope V control unit, Bruker AXS). Samples were visualized in 20mM Tris, 150mM NaCl buffer pH=7 at room temperature by operating the AFM in the PeakForce Tapping mode.

AFM images were acquired at room temperature by using triangular silicon nitride cantilevers with nominal spring constant of 0.7 N m<sup>-1</sup> and tip radius of 20 nm (ScanAsyst-Fluid, Bruker AXS). Analysis and processing of AFM images was performed with WSxM software<sup>72</sup>. Image processing consisted of plane subtraction/flattening and equalization. Freshly cleaved mica supports were imaged in ultrapure water prior to sample injection to ensure a clean and smooth surface. Then, sample injection and subsequent rinsing with buffer (20 mM Tris, 150 mM NaCl buffer pH=7.5) was performed in situ as described previously<sup>73</sup> using a slow gravity-fed flow of approximately 50  $\mu$ L/min.

#### Surface Plasmon Resonance (SPR)

Plain unmodified gold sensors (Sensor Chip Au from GE Healthcare, cat BR100542) and a Biacore3000™ system (GE Healthcare) were used. During the experiments 150  $\mu$ l of rTF-loaded peptide disc solution in 20 mM Tris, 100 mM NaCl, 20 mM EDTA pH 7.2 were injected at a flow rate of 2  $\mu$ l/min at 20°C followed by a stabilization period of 40 min with buffer 20 mM Tris-HCl pH 7.2, 100 mM NaCl, 10 mM CaCl<sub>2</sub>. The control surface was prepared by injection of the corresponding empty peptide discs (without TF). rTF- FVIIa binding was probed injecting 20 nM FVIIa solution at 30  $\mu$ l/min at 20°C in 20 mM Tris, 100 mM NaCl, 10 mM CaCl<sub>2</sub>, pH 7.2 for 250 sec followed by 3,000 sec dissociation. The sensor chip was regenerated by two consecutive 3 min injections of 20 mM EDTA in running buffer (without Ca<sup>2+</sup>). This procedure was repeated at different time intervals to probe the stability of the TF-embedded lipid surface. The evaluation software supplied with the instrument for global fitting (Biaevaluation™ 4.1) was used to calculate the association rate constant ( $k_{on}$ ) and dissociation rate constant ( $k_{off}$ ), the K<sub>D</sub> ( $k_{off}/k_{on}$ ), as well as the binding capacity ( $R_{max}$ ).

## RESULTS AND DISCUSSION

#### Peptide disc mediated formation of supported lipid bilayers

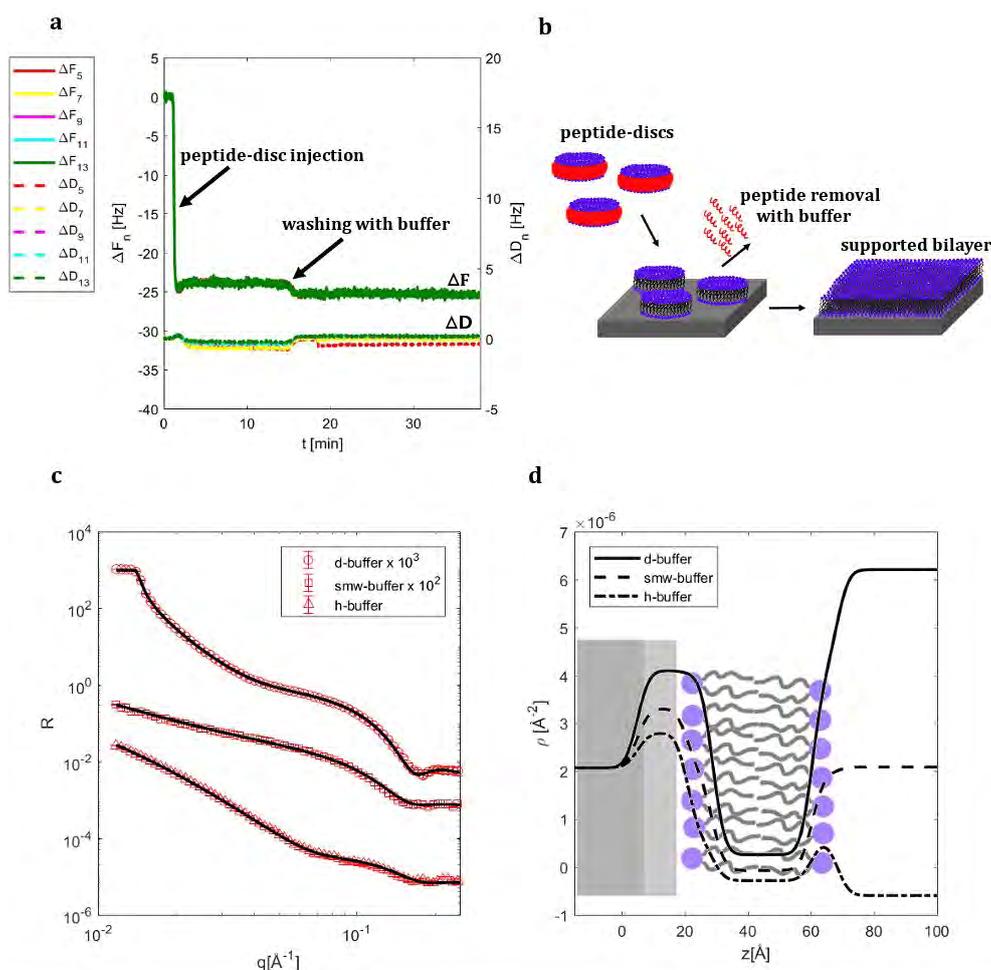
The capacity of the peptide discs to deliver phospholipids on a solid support surface and form a supported lipid bilayer was initially tested by Quartz Crystal Microbalance with Dissipation monitoring (QCM-D). Figure 2a shows the QCM-D data collected during the deposition of 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) loaded peptide discs on a SiO<sub>2</sub> coated quartz sensor and subsequent removal of the peptide. As soon as the peptide discs were injected, a rapid decrease in the frequency shift ( $\Delta F$ ) down to  $\sim -23$ Hz was observed suggesting a rapid adsorption of the peptide discs on the support surface. The dissipation ( $\Delta D$ )  $\sim 0$  also suggests the formation of a rigid layer on the support. About 15 min after the peptide disc injection, the buffer solution was flushed through the QCM-D cell in order to remove the peptide disc excess as well as to facilitate the removal of the peptide molecules from the deposited lipids (See illustration in Figure 2b). Interestingly the  $\Delta F$  for all of the monitored frequencies reached the stable value of  $\sim -25$  Hz, which is the typical frequency shift value observed for the formation of a POPC bilayer on the sensor surface<sup>74</sup> (Figure S2). The variation of the  $\Delta F$  from  $\sim -23$ Hz to  $\sim -25$ Hz together with the variation of  $\Delta D$  from  $\sim -0.6$  to  $\sim -0.1$  suggests small rearrangements in the deposited film towards the formation of a more compact structure. Neutron Reflectometry (NR) was applied to obtain more detailed information about the sample structure and to confirm the effective formation of a lipid bilayer. Figures 2c and 2d shows, respectively, the collected NR data together with the corresponding fitting curves and scattering length density profiles,  $\rho(z)$ , extracted from the fits. The model used for analysing the NR data was composed by 4 layers. One layer described the native silicon oxide layer present on the silicon support surface, two layers described the outer and inner lipid headgroups and one layer

described the hydrophobic region of the bilayer composed by the lipid acyl chains sandwiched between the two headgroup layers. In the initial analysis the two headgroup layers were considered independent. But as the bilayer was confirmed to be symmetric, these layers were constrained to have the same parameter values in the final analysis. Four parameters, thickness ( $t$ ), scattering length density ( $\rho$ ), solvent volume fraction ( $\phi_s$ ) and roughness ( $r$ ), are assigned to each layer. Further details are reported in the Experimental Section.

According to the optimized fitting parameters (Table S1) a lipid bilayer was efficiently formed on the support surface (surface coverage > 80% as calculated from the solvent volume fraction) with the expected structural parameters for a POPC membrane<sup>69</sup>. No indications of remaining peptide molecules were detected among the lipid acyl chains. Similar results were obtained for the supported lipid bilayer prepared using a mixture of POPC and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine (POPS) (70/30 mol/mol) in the peptide discs (Figure S3). To modify the contrast between the two lipid species, a supported lipid bilayer was also produced by using peptide discs prepared with single chain deuterated POPC (dPOPC) and POPS

(70/30 mol/mol). The contrast between the partially deuterated acyl chains of dPOPC and the hydrogenous chains of POPS allows the POPS concentration in the bilayer to be determined from the  $\rho$  value of the bilayer acyl chain region. We found that the POPS content in the supported lipid bilayer produced by the deposition of 18A dPOPC/POPS peptide discs was in perfect agreement with the nominal content of the sample composition, i.e.  $(28 \pm 2)$  % mol/mol (Figure S4 and Table 1).

Altogether the collected data confirms that the peptide discs provide an efficient delivery system to form supported lipid bilayers with relatively high surface coverage and the expected phospholipid composition. The overall performance of the peptide disc as lipid carriers for the formation of supported lipid bilayers is comparable to that of vesicles. However, the deposition of the peptide discs can be conducted in 10-20 s as judged from the QCM-D and is therefore much faster than the adsorption and subsequent fusion of the vesicles which typically takes a few minutes (Figure S1)<sup>20, 74</sup>. Although only data on POPC and POPC/POPS lipid bilayers are presented here (Figure 2, S3 and S4) the method does not contain any intrinsic limitation to the choice of lipid composition for the lipid bilayer preparation.



**Figure 2:** Peptide disc mediated formation of the supported POPC bilayer on SiO<sub>2</sub> surfaces. a) QCM-D data monitoring the formation of the supported POPC bilayer *via* peptide discs deposition (see also supplementary information). b) schematic

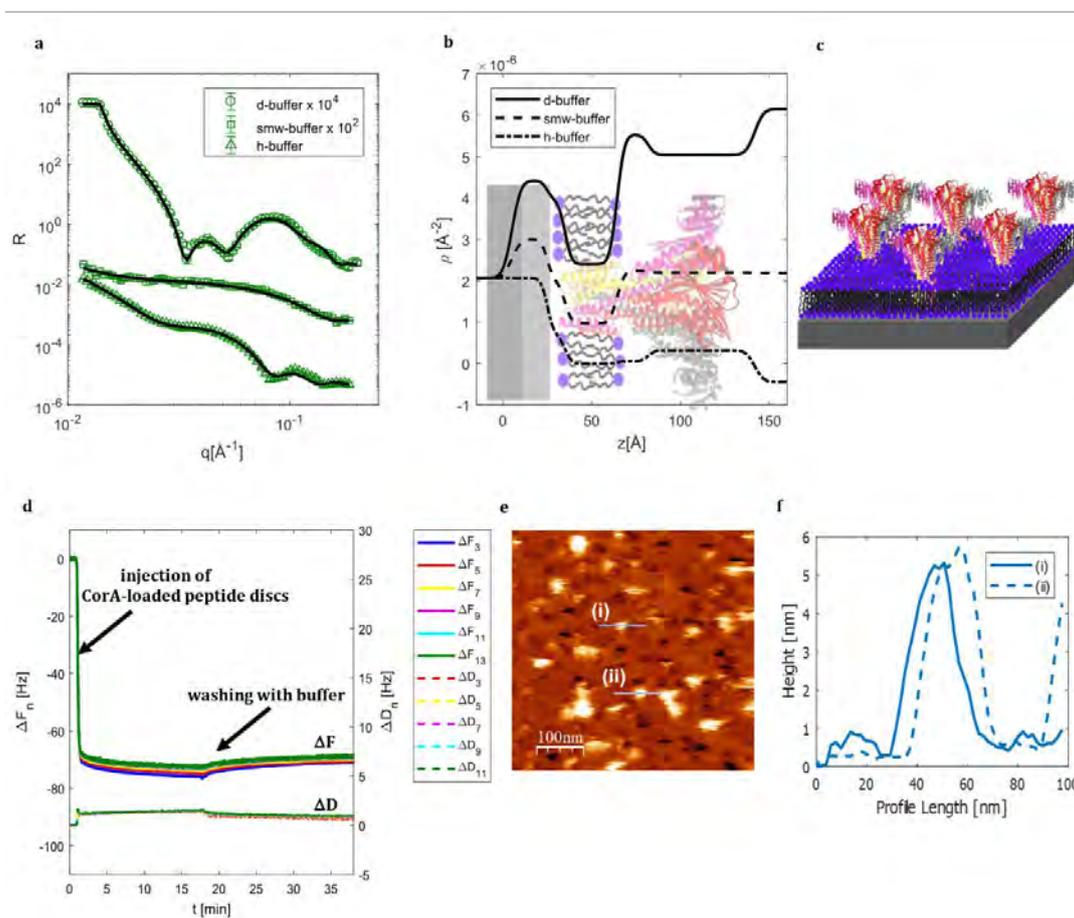
representation of the supported lipid bilayer formation through peptide discs deposition. c) NR experimental data together with the corresponding fitting curves. Data corresponding to the different buffer compositions are multiplied by a scale factor as reported in the legend. d) Scattering length density profile calculated from experimental data analysis. During both the QCM-D and the NR experiment, the peptide discs were prepared with 1.3mM POPC and 0.43mM 18A in 20mM Tris-HCl, 100mM NaCl, 10mM CaCl<sub>2</sub> (pH=7.5) buffer. All measurements were carried out at 25°C

#### Formation of supported bilayers with oriented CorA

The formation of supported lipid bilayers with membrane proteins was first evaluated in the case of the bacterial pentameric magnesium transporter CorA. Detailed structural information is available for CorA, which contains a large extramembrane domain protruding from the intracellular side of the membrane (Figure 1). Furthermore, previous work from our group demonstrated the successful reconstitution of CorA in nanodiscs<sup>75</sup>. In summary, CorA with its large size and good sample control represents an ideal model system to verify that the present method indeed selects a given protein orientation in the deposited lipid bilayer.

Figures 3a and b show the collected NR experimental data with the relative fitting curves and the calculated scattering length density profiles, respectively, for CorA in a supported lipid bilayer. The model used to analyse the NR data considered the CorA molecules as composed by three main regions: the transmembrane domain (TMD), a large

extramembrane domain corresponding to the intracellular domain (ICD) and a short extracellular domain (ECD). More specifically, the model was composed by 5 layers describing 1) silicon oxide layer, 2) inner POPC headgroup layer, which also included CorA ECD, 3) POPC acyl chain layer and CorA TMD, 4) outer POPC headgroup layer, which partially included the CorA ICD and 5) the part of the CorA ICD that protrudes out of the membrane surface. In the model the number of protein molecules contained in layer 3 and 5 was constrained to be the same so that to each TMD corresponds one ECD (see further details in Experimental Section). Because of the very low contrast between the protein and the lipid headgroups (Table S1 and 1), the presence of CorA was not explicitly taken into account in layers 2 and 4. Optimization of the model parameters resulted in good agreement between the collected experimental data and the theoretical curves (Figure 3a).



**Figure 3:** Supported POPC bilayer with CorA. a) NR experimental data together with the relative fitting curves for the membrane deposited on SiO<sub>2</sub> surface. Data corresponding to the different buffer compositions are multiplied by a scale factor as reported in the legend. b) Scattering length density profile calculated from experimental data analysis. c) schematic representation of the supported bilayer with the CorA molecules all sharing the same orientation. d) QCM-D data monitoring

the formation of the supported bilayer with CorA on a SiO<sub>2</sub> surface. e) AFM image collected for the POPC supported bilayer with CorA on mica. f) The height profiles corresponding to the blue lines in the AFM image. For the QCM-D and the NR experiments, the CorA-loaded peptide discs were prepared with 1.3mM POPC, 0.43mM 18A and 12μM CorA in 20mM Tris-HCl, 100mM NaCl, 10mM CaCl<sub>2</sub> (pH=7.5) buffer (further details on the sample preparation are reported in Materials and Methods). For the AFM measurements, the CorA-loaded peptide disc suspension was diluted 1:2 before use. All measurements were carried out at 25°C.

**Table 1:** Structural parameters obtained from NR data analysis for the POPC supported bilayer with CorA. The sample is schematically represented as a stack of layers and four parameters are assigned to each layer: thickness (*t*), scattering length density ( $\rho$ ), solvent volume fraction ( $\phi_s$ ) and surface roughness (*r*).  $\phi_s$  provides information on the defects present in the sample, i.e. area of the support surface not covered by lipids or protein molecules, and it is used to calculate the overall surface coverage. The protein scattering length density was calculated according to the protein molecular structure<sup>55</sup>.

POPC supported bilayer with CorA	
Inner headgroup layer (including CorA ECD):	
<i>t</i> [Å]	15 ± 1
$\rho$ [Å <sup>-2</sup> ]	(2.4 ± 0.1) · 10 <sup>-6</sup>
$\phi_s$	47 ± 3
<i>r</i> [Å]	3 ± 1
Acyl chain layer (including CorA TMD):	
<i>t</i> [Å]	28 ± 1
$\rho$ [Å <sup>-2</sup> ]	(0.07 ± 0.01) · 10 <sup>-6</sup>
$\phi_s$	36 ± 1
<i>r</i> [Å]	3 ± 1
Outer headgroup layer (including CorA ICD):	
<i>t</i> [Å]	18 ± 1
$\rho$ [Å <sup>-2</sup> ]	(3.0 ± 0.3) · 10 <sup>-6</sup>
$\phi_s$	83 ± 2
<i>r</i> [Å]	3 ± 1
CorA ICD	
<i>t</i> [Å]	60 ± 1
$\rho$ [Å <sup>-2</sup> ]	2.91 · 10 <sup>-6</sup> (d-buffer) 2.22 · 10 <sup>-6</sup> (smw-buffer) 1.79 · 10 <sup>-6</sup> (h-buffer)
$\phi_s$	65 ± 1
<i>r</i> [Å]	4 ± 1

As a first important result, NR data analysis confirmed the obtainment of a supported bilayer with all the CorA molecules oriented with the large ICD pointing towards the solvent as reported in Figure 3c (simulation of different orientations of CorA compared to the collected data are reported in Figures S5 and S6). The favorable interactions between the lipid headgroups and the hydrophilic support surface<sup>76</sup> are hypothesized to be the driving force that controls the membrane protein orientation. Only if the large ICD is oriented towards the bulk solvent the lipid headgroups become sufficiently close to support surface and can stably interact with it. Hence, the injection of CorA-loaded peptide disc led to the spontaneous formation of a supported lipid bilayer with oriented membrane proteins. Furthermore, a large volume fraction of the CorA ICD (35% v/v) and TMD (20% v/v) was found in the prepared supported bilayers. Although the overall membrane surface

coverage was 64% v/v, according to the 36% v/v of solvent which is present in the layer containing the lipid acyl chains and the CorA-TMD, the collected data still represents one of the very a few examples of high protein loading in supported lipid bilayers<sup>38</sup>.

Peptide disc formed supported bilayers with CorA were also investigated by means of complementary surface-sensitive techniques such as QCM-D and Atomic Force Microscopy (AFM). As for the peptide discs prepared with lipids only, the deposition of CorA-loaded peptide discs on the quartz sensor is a fast process that, in a few minutes, leads to a stable  $\Delta F$  signal (Figure 3d). The  $\Delta F$  decrease is clearly more pronounced than the one corresponding to the pure lipid bilayer (Figure 2a), as expected from the high protein loading suggested by the NR data. Buffer flushing through the cell produced a small increase of  $\Delta F$ , as observed for the pure POPC supported bilayer, which reached a stable value close to -69 Hz. Altogether, the recorded  $\Delta F$  values are compatible with the formation of the supported bilayer with a high degree of protein loading.

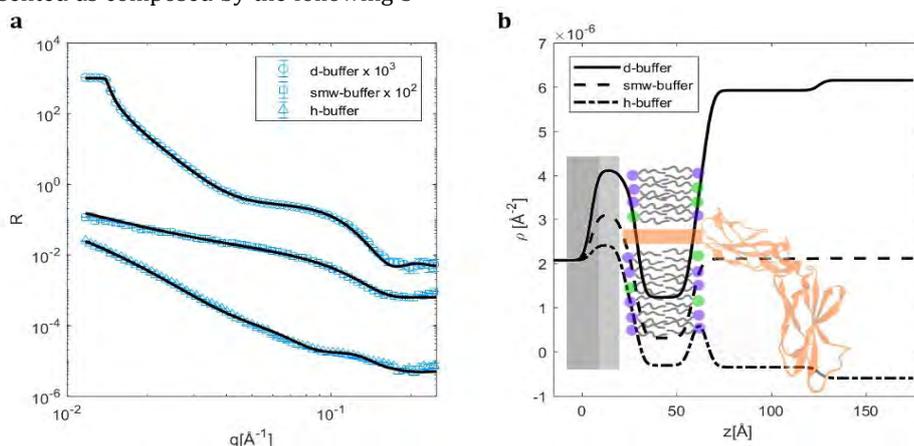
Figure 3e shows an AFM image collected for the POPC supported bilayer with CorA on a mica substrate. The CorA-loaded peptide disc solution was diluted 2-fold compared to the sample used for NR data collection (see more details in Experimental Section) to enable a better visualization of the protein molecules in the supported bilayer. The AFM image clearly shows the presence of features protruding out of the membrane plane. Most of these features exhibited a width at half height of approximately 20 nm and heights in the 4–7 nm range (representative profiles shown in Figure 3f). Height values are similar to those determined from the NR data analysis for the CorA ICD thickness (Table 1). The higher width values can be attributed to tip dilation effects<sup>77</sup> i.e., imaging of features smaller than the AFM tip results in width values similar to those of the tip (approx. 20 nm). Thus, it is reasonable to attribute the smaller visualized features to individual CorA proteins. Larger features were also present in the sample. Higher resolution AFM images (Figure S7) revealed that these larger features consisted on individual CorA, indicating a tendency for self-association, which might occur through lateral diffusion of the CorA molecules in the supported bilayer.

#### Formation of supported bilayers with Tissue Factor (TF)

A modified version of Tissue Factor, named recombinant Tissue Factor rTF, was used in this study (see Experimental section). While the rTF ECD and TMD are the same as in the native TF, the composition of the ICD is different in the two proteins. This modification of the protein sequence is required for an efficient protein purification. Although rTF ICD (~ 12 residues) and TF ICD (~ 19 residues) have different amino acid composition, they still exhibit an overall similar size, which makes rTF a suitable candidate

to mimic TF. Figure 4 shows the NR data collected for the POPC/POPS (70/30 mol/mol) supported bilayer with rTF. Data modelling was based on the same schematic representation of the membrane protein used in the case of CorA. However, in rTF, the ICD is a very short domain while the ECD is a larger domain (~ 219 residues). The peptide disc deposition is hypothesized to favor the rTF orientation with the ICD close to the support surface and rTF ECD pointing towards the bulk solvent. The rTF TMD is expected to be located among the lipid acyl chains. Hence, in analogy to the CorA case, the supported bilayer with rTF was schematically represented as composed by the following 5

layers: 1) SiO<sub>2</sub> layer on the support surface; 2) inner headgroup layers including the rTF ICD, 3) lipid acyl chain including rTF TMD, 4) outer headgroup layer including the loop connecting the rTF-TMD to the rTF ECD, 5) the rTF ECD that is protruding out from the membrane surface. Initial optimization of the model highlighted that the structural parameters for layer 2 and 4 converged to very similar values. Hence, in the final model optimization these two layers were constrained to have the same parameters values, which considerably reduced the number of parameters in the model.



**Figure 4:** a) NR experimental data together with the relative fitting curves for the supported bilayer with rTF and lipid composition POPC/POPS (70/30 mol/mol). Data corresponding to the different buffer composition are reported with different symbols and multiplied by a scale factor as reported in the legend. b) Scattering length density profile calculated from experimental data analysis. Peptide discs were prepared with 1.3 Mm POPC/POPS, 3.3 mM 18A and 12 μM rTF as detailed in the Materials and Method section. During the measurements, a 20 mM Tris, 100 mM NaCl, 20mM CaCl<sub>2</sub> buffer (pH=7) was used. All the experiments were carried out at 25 °C.

**Table 2**-Structural parameters obtained from NR data analysis on POPC/POPS supported bilayer with rTF. The sample is schematically represented as a stack of layers and four parameters are assigned to each layer: thickness ( $t$ ), scattering length density ( $\rho$ ), solvent volume fraction ( $\phi_s$ ) and roughness ( $r$ ).  $\phi_s$  provides information on the defects present in the sample, i.e. area of the support surface not covered by lipids or protein molecules, and it is used to calculate the overall surface coverage. The protein scattering length density was calculated according to the protein molecular structure<sup>61</sup>.

POPC/POPS supported bilayer with rTF	
Inner and outer headgroup layer (including rTF ICD and rTF loop respectively)	
$t$ [Å]	$7 \pm 1$
$\rho$ [Å <sup>-2</sup> ]	$(1.8 \pm 0.2) \cdot 10^{-6}$
$\phi_s$	$38 \pm 4$
$r$ [Å]	$3 \pm 1$
Acyl chain layer (including rTF TMD)	
$t$ [Å]	$29 \pm 1$
$\rho$ [Å <sup>-2</sup> ]	$(-0.22 \pm 0.02) \cdot 10^{-6}$
$\phi_s$	$23 \pm 2$
$r$ [Å]	$3 \pm 1$
rTF ECD	

$t$ [Å]	$59 \pm 5$
$\rho$ [Å <sup>-2</sup> ]	$3.35 \cdot 10^{-6}$ (d-buffer) $2.52 \cdot 10^{-6}$ (smw-buffer) $2.01 \cdot 10^{-6}$ (h-buffer)
$\phi_s$	$92 \pm 1$
$r$ [Å]	$4 \pm 1$

The consistency between the experimental data and fitting curves confirmed the orientation of rTF in the lipid membrane with the ECD exposed to the solvent. This further validated the preparation method as a suitable strategy for orienting structurally and functionally diverse membrane proteins. Indeed, assuming that the rTF is randomly oriented produces theoretical curves that are considerably different as compared to the collected experimental data (see Figures S8 and S9). The optimized structural parameters obtained by fitting the model to the data are reported in Table 2. Inspection of the table shows that the presence of rTF in the lipid bilayer has little impact on the main membrane structural parameters, e.g. the acyl chain thickness, compared to the corresponding pure lipid membrane (Table S1). Concerning rTF, the ECD volume fraction corresponds to 8% v/v, while the rTF TMD corresponds to 2% v/v. rTF is considerably smaller than CorA and is expected to have a smaller volume fraction in the prepared lipid bilayer. Conversion to mol fraction shows that the protein content in the supported bilayers with CorA and rTF are actually quite similar ~0.6% mol/mol (protein-

1 TMD/lipids) for rTF and  $\sim 0.4\%$  mol/mol (protein-  
2 TMD/lipids) for CorA, respectively. In this calculation CorA  
3 is assumed to be folded as a pentamer in the lipid bilayer.

4 Besides validating the peptide disc mediated lipid  
5 deposition with integral membrane proteins, the NR data  
6 collected for the bilayer with rTF also provided structural  
7 information on the orientation of rTF ECD with respect to  
8 the bilayer surface. Indeed, the thickness associated with  
9 the rTF ECD protruding out of the bilayer surface was  
10 refined to  $59 \pm 5 \text{ \AA}$ , which notably is considerably smaller  
11 than the maximum length of the protein domain (i.e.  $75 \text{ \AA}^{61}$ ).  
12 This suggests that the rTF ECD is on average tilted towards  
13 the lipid headgroups as sketched in Figure 4.

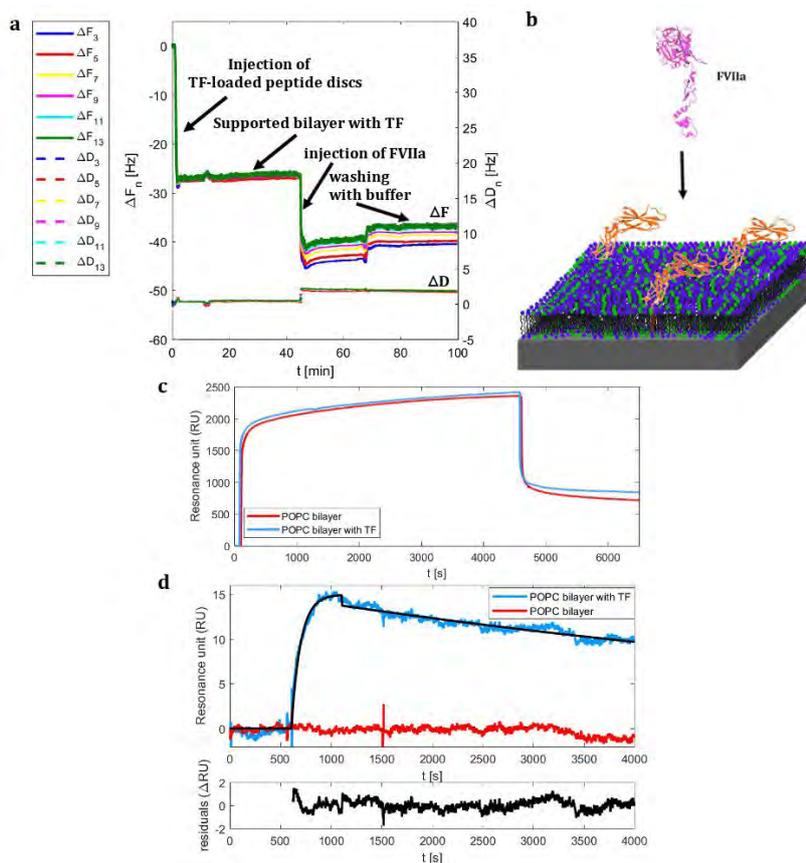
14 The initial step in the extrinsic blood coagulation pathway  
15 involves the formation of the protein complex between the  
16 TF ECD and the activated coagulation factor FVII (FVIIa)<sup>78-  
17 79</sup>. Hence, demonstrating a robust binding of FVIIa to rTF  
18 ECD in the produced supported bilayers with the expected  
19 kinetic rate constants is a fundamental acid test of the  
20 mimicry that is achieved by this surrogate lipid bilayer.

21 QCM-D and Surface Plasmon Resonance (SPR) were used to  
22 monitor the binding of FVIIa to the supported bilayer with  
23 rTF. The signal measured by both QCM-D and SPR is related  
24 to the mass of the adsorbed molecules. In the case of the  
25 supported bilayer with rTF, an increase of the adsorbed  
26 mass is expected compared to the pure lipid bilayer due to  
27 the presence of the membrane protein in the bilayer. A  
28 further mass increase is expected upon FVIIa binding to rTF.  
29 As FVIIa is reported to interact with PS lipids independent  
30 of TF<sup>60, 80</sup> these experiments were performed using a POPC  
31 supported bilayer, which indeed exhibited very little  
32 affinity for FVIIa in the absence of rTF (Figure S9 and S10).  
33 To detect the selective binding of FVIIa to rTF, a supported  
34 bilayer composed by POPC and rTF was used during the  
35 QCM-D and SPR measurements. Figure 5a shows the QCM-

D data for the formation of the supported POPC bilayer with  
rTF and subsequent binding of FVIIa. Addition of FVIIa  
produced a large  $\Delta F$  decrease, which signifies FVIIa binding  
to rTF. Since FVIIa binds to the TF-ECD<sup>78</sup>, this result  
qualitatively confirms that the rTF-ECD was in a binding  
competent stage.

To obtain more quantitative data on the binding kinetics of  
the FVIIa-rTF interaction, we also embarked on real-time  
binding studies with SPR measurements on a gold sensor  
chip. Figure 5c shows the difference in SPR signals during  
the deposition of a POPC supported bilayer in the control  
surface (red curve) and the supported bilayer with rTF in  
the active flow cell (blue curve). The difference in the  
immobilized mass between the reference (supported POPC  
bilayer) and the active surface (supported POPC bilayer  
with TF) was estimated to be 100–200 RU suggesting that  
 $3.5\text{--}7\text{-fmols of TF/mm}^2$  was deposited in the active flow cell.  
This corresponds to a surface coverage of around 2-4% of  
the TF ECD.

Injection of FVIIa over the gold supported lipid bilayer,  
revealed a specific binding of FVIIa to the active surface  
with embedded rTF yielding an equilibrium dissociation  
constant  $K_D$  of  $0.2 \cdot 10^{-9} \text{ M}$  (Figure 5c) in agreement with  
previously reported results<sup>81</sup>. The association rate constant  
( $k_{on}$ ), dissociation rate constant ( $k_{off}$ ), the  $K_D$  ( $k_{off}/k_{on}$ ), and  
binding capacity ( $R_{max}$ ) were derived by fitting of the double  
blank referenced data by non-linear regression to a  
bimolecular interaction model, assuming pseudo-first order  
reaction conditions (see Table S2). The tight binding of  
FVIIa to lipid embedded TF is governed by fast on-rates ( $k_{on}$   
 $6.3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ ) and relatively slow off-rates ( $k_{off}$   $1.4 \times$   
 $10^{-5}\text{s}^{-1}$ ). Repeating that analysis on the same supported  
lipid bilayer revealed that the surface is relatively stable and  
only loses approximately 0.5 % of its binding capacity per  
hour (Table S2).



**Figure 5:** a) QCM-D data collected for the supported POPC bilayer with rTF on a SiO<sub>2</sub> surface; the frequency shift,  $\Delta F$ , (left) and the dissipation,  $\Delta D$ , (right) are monitored over time. b) schematic representation of the interaction between FVIIa and the supported bilayer with rTF. c) and d) SPR data collected for the supported POPC bilayer and supported POPC bilayer with rTF (c); experimental data with the relative fitting curve for the 20 nM FVIIa binding on TF embedded in the supported POPC bilayer (blue curve) and buffer injection (red curve) with double-reference subtraction (d). For both the QCM-D and SPR experiment peptide discs were prepared with 1.3 mM POPC, 0.43 mM 18A and 12  $\mu$ M rTF as detailed in the Materials and Method section. During the measurements a 20 mM Tris-HCl, 100 mM NaCl, 10mM CaCl<sub>2</sub> buffer (pH=7.5) was used. All the experiments were carried out at 25 °C.

## CONCLUSION

The use of phospholipid and membrane protein loaded peptide discs represents a versatile approach for the production of supported lipid bilayers with membrane proteins, which offers several unique advantages compared to other state-of-the-art methods. In particular, no chemical modification of the support is required. Solid supports with different chemical composition can be successfully used, as long as the surface is hydrophilic, and no tagging of the membrane protein is needed to control the protein orientation. Since the membrane proteins are not chemically anchored to the support, the membrane proteins are, in principle, free to diffuse laterally within the bilayer, as they are native cell membranes.

Two different kinds of membrane proteins, CorA and TF, were efficiently loaded in the peptide discs and used to validate the proposed method for the formation of supported lipid bilayers with membrane proteins. CorA was used as a model system to evaluate whether the deposition of CorA-loaded peptide discs could lead to a supported lipid

bilayer with oriented CorA molecules. Indeed, the performed structural characterization provided such information and showed that CorA was efficiently loaded in the supported bilayer and that all the protein molecules shared the same orientation with the ICD pointing towards the bulk solvent (Figure 3).

In order to validate the protocol to other membrane proteins, Tissue Factor was investigated. The proposed method based on the deposition of rTF-loaded peptide disc produced a relevant model system of TF in a native-like lipid environment. The data presented here showed that supported lipid bilayers with rTF were successfully prepared and that they exhibited the expected structure with the rTF extracellular domain (rTF ECD) pointing towards the bulk solvent. To further validate this conclusion, the interaction between rTF and its soluble binding partner FVIIa was also characterized. Both QCM-D and SPR showed FVIIa binding to the supported bilayers consolidating the native-like properties of the embedded rTF.

1 Altogether, the peptide disc mediated formation of  
2 supported lipid bilayers with membrane proteins is a  
3 promising method for the structural and functional  
4 investigation of membrane proteins. The method can be  
5 used for the investigation of diverse membrane proteins as  
6 well as protein-protein complexes, and in principle allows  
7 the composition of the membrane lipids to be varied.  
8 Supported bilayers with oriented membrane proteins can  
9 be prepared on different types of solid supports and thus  
10 the method proposed here allows for the powerful  
11 combination of several surface-sensitive techniques (e.g.  
12 NR, AFM, SPR, QCM-D) for membrane protein  
13 characterization.

## 14 ASSOCIATED CONTENT

### 15 Supporting Information

## 16 AUTHOR INFORMATION

### 17 Corresponding Authors

18 Dr. Alessandra Luchini, Niels Bohr Institute, University of  
19 Copenhagen, Universitetsparken 5, 2100 Copenhagen,  
20 Denmark. a.luchini@nbi.ku.dk

21 Prof. Lise Arleth, Niels Bohr Institute, University of  
22 Copenhagen, Universitetsparken 5, 2100 Copenhagen,  
23 Denmark. arleth@nbi.ku.dk

## 24 ACKNOWLEDGMENT

25 The work was supported by grants from Novo Nordisk  
26 foundation Interdisciplinary Synergy program, the  
27 Lundbeck foundation "BRAINSTRUC" project and Danscatt  
28 for travel support. The authors also thank the ISIS neutron  
29 source for the allocation of beamtime.

## 30 REFERENCES

1. Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L., How many drug targets are there? *Nature Reviews Drug Discovery* **2006**, *5* (12), 993-996.
2. Mukherjee, K.; Bhattacharyya, S.; Peralta-Yahya, P., GPCR-Based Chemical Biosensors for Medium-Chain Fatty Acids. *ACS Synthetic Biology* **2015**, *4* (12), 1261-1269.
3. Mazur, F.; Bally, M.; Städler, B.; Chandrawati, R., Liposomes and lipid bilayers in biosensors. *Advances in Colloid and Interface Science* **2017**, *249*, 88-99.
4. Seddon, A. M.; Curnow, P.; Booth, P. J., Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2004**, *1666* (1), 105-117.
5. Rawlings, A. E., Membrane proteins: always an insoluble problem? *Biochem Soc Trans* **2016**, *44* (3), 790-5.
6. Kang, C.; Li, Q., Solution NMR study of integral membrane proteins. *Curr Opin Chem Biol* **2011**, *15* (4), 560-9.
7. Falkenburger, B. H.; Jensen, J. B.; Dickson, E. J.; Suh, B. C.; Hille, B., Phosphoinositides: lipid regulators of membrane proteins. *J Physiol* **2010**, *588* (Pt 17), 3179-85.

8. McIntosh, T. J.; Simon, S. A., ROLES OF BILAYER MATERIAL PROPERTIES IN FUNCTION AND DISTRIBUTION OF MEMBRANE PROTEINS. *Annual Review of Biophysics and Biomolecular Structure* **2006**, *35* (1), 177-198.
9. Alberts B, J. A., Lewis J, et al., *Molecular Biology of the Cell. 4th edition*. New York: Garland Science: 2002.
10. Lyu, S. W.; Wang, J. F.; Chao, L., Constructing Supported Cell Membranes with Controllable Orientation. *Sci Rep* **2019**, *9* (1), 2747.
11. van Weerd, J.; Karperien, M.; Jonkheijm, P., Supported Lipid Bilayers for the Generation of Dynamic Cell-Material Interfaces. *Advanced Healthcare Materials* **2015**, *4* (18), 2743-2779.
12. Shen, H. H.; Lithgow, T.; Martin, L. L., Reconstitution of Membrane Proteins into Model Membranes: Seeking Better Ways to Retain Protein Activities. *Int J Mol Sci* **2013**, *14* (1), 1589-1607.
13. Kurniawan, J.; Ventrici de Souza, J. F.; Dang, A. T.; Liu, G. Y.; Kuhl, T. L., Preparation and Characterization of Solid-Supported Lipid Bilayers Formed by Langmuir-Blodgett Deposition: A Tutorial. *Langmuir* **2018**, *34* (51), 15622-15639.
14. Richter, R. P.; Berat, R.; Brisson, A. R., Formation of solid-supported lipid bilayers: an integrated view. *Langmuir* **2006**, *22* (8), 3497-505.
15. Andersson, J.; Koper, I.; Knoll, W., Tethered Membrane Architectures-Design and Applications. *Front Mater* **2018**, *5*.
16. Andersson, J.; Köper, I., Tethered and Polymer Supported Bilayer Lipid Membranes: Structure and Function. *Membranes* **2016**, *6* (2), 30.
17. Isaksson, S.; Watkins, E. B.; Browning, K. L.; Kjellerup Lind, T.; Cárdenas, M.; Hedfalk, K.; Höök, F.; Andersson, M., Protein-Containing Lipid Bilayers Intercalated with Size-Matched Mesoporous Silica Thin Films. *Nano Letters* **2017**, *17* (1), 476-485.
18. Jagalski, V.; Barker, R. D.; Thygesen, M. B.; Gotfryd, K.; Krüger, M. B.; Shi, L.; Maric, S.; Bovet, N.; Moulin, M.; Haertlein, M.; Günther Pomorski, T.; Loland, C. J.; Cárdenas, M., Grafted biomembranes containing membrane proteins – the case of the leucine transporter. *Soft Matter* **2015**, *11* (39), 7707-7711.
19. Pace, H. P.; Hannestad, J. K.; Armonious, A.; Adamo, M.; Agnarsson, B.; Gunnarsson, A.; Micciulla, S.; Sjövall, P.; Gerelli, Y.; Höök, F., Structure and Composition of Native Membrane Derived Polymer-Supported Lipid Bilayers. *Anal Chem* **2018**, *90* (21), 13065-13072.
20. Richter, R. P.; Brisson, A. R., Following the formation of supported lipid bilayers on mica: A study combining AFM, QCM-D, and ellipsometry. *Biophysical Journal* **2005**, *88* (5), 3422-3433.
21. Kiessling, V.; Yang, S.-T.; Tamm, L. K., Chapter One - Supported Lipid Bilayers as Models for Studying Membrane Domains. In *Current Topics in Membranes*, Kenworthy, A. K., Ed. Academic Press: 2015; Vol. 75, pp 1-23.
22. Keller, C. A.; Glasmästar, K.; Zhdanov, V. P.; Kasemo, B., Formation of Supported Membranes from Vesicles. *Physical Review Letters* **2000**, *84* (23), 5443-5446.
23. Hollmann, A.; Martinez, M.; Maturana, P.; Semorile, L. C.; Maffia, P. C., Antimicrobial Peptides: Interaction With Model and Biological Membranes and Synergism With Chemical Antibiotics. *Front Chem* **2018**, *6*, 204.
24. Rondelli, V.; Brocca, P.; Motta, S.; Messa, M.; Colombo, L.; Salmona, M.; Fragneto, G.; Cantu, L.; Del Favero, E., Amyloid-beta Peptides in interaction with raft-mime model membranes: a neutron reflectivity insight. *Sci Rep-Uk* **2016**, *6*.
25. Eells, R.; Barros, M.; Scott, K. M.; Karageorgos, I.; Heinrich, F.; Losche, M., Structural characterization of membrane-bound human immunodeficiency virus-1 Gag matrix with neutron reflectometry. *Biointerphases* **2017**, *12* (2), 02D408.
26. Lad, M. D.; Birembaut, F.; Clifton, L. A.; Frazier, R. A.; Webster, J. R.; Green, R. J., Antimicrobial peptide-lipid binding

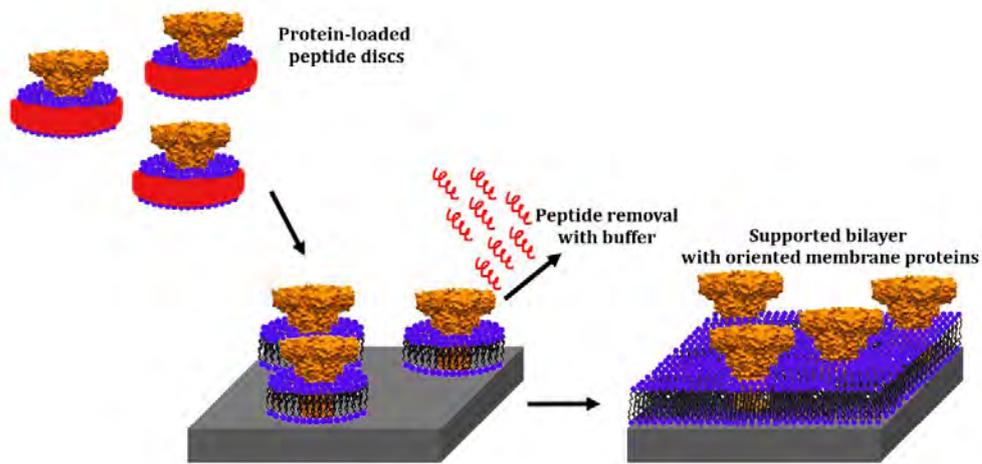
- interactions and binding selectivity. *Biophys J* **2007**, *92* (10), 3575-86.
27. Vitiello, G.; Falanga, A.; Petruk, A. A.; Merlino, A.; Fragneto, G.; Paduano, L.; Galdiero, S.; D'Errico, G., Fusion of raft-like lipid bilayers operated by a membranotropic domain of the HSV-type I glycoprotein gH occurs through a cholesterol-dependent mechanism. *Soft Matter* **2015**, *11* (15), 3003-3016.
28. Merlino, A.; Vitiello, G.; Grimaldi, M.; Sica, F.; Busi, E.; Basosi, R.; D'Ursi, A. M.; Fragneto, G.; Paduano, L.; D'Errico, G., Destabilization of Lipid Membranes by a Peptide Derived from Glycoprotein gp36 of Feline Immunodeficiency Virus: A Combined Molecular Dynamics/Experimental Study. *The Journal of Physical Chemistry B* **2012**, *116* (1), 401-412.
29. Fragneto, G.; Delhom, R.; Joly, L.; Scoppola, E., Neutrons and model membranes: Moving towards complexity. *Current Opinion in Colloid & Interface Science* **2018**, *38*, 108-121.
30. Pace, H.; Nystrom, L. S.; Gunnarsson, A.; Eck, E.; Monson, C.; Geschwindner, S.; Snijder, A.; Hook, F., Preserved Transmembrane Protein Mobility in Polymer-Supported Lipid Bilayers Derived from Cell Membranes. *Anal Chem* **2015**, *87* (18), 9194-9203.
31. Poudel, K. R.; Jones, J. P.; Brozik, J. A., A Guide to Tracking Single Transmembrane Proteins in Supported Lipid Bilayers. In *Lipid-Protein Interactions: Methods and Protocols*, Kleinschmidt, J. H., Ed. Humana Press: Totowa, NJ, 2013; pp 233-252.
32. Graneli, A.; Rydstrom, J.; Kasemo, B.; Hook, F., Formation of supported lipid bilayer membranes on SiO<sub>2</sub> from proteoliposomes containing transmembrane proteins. *Langmuir* **2003**, *19* (3), 842-850.
33. Heinrich, F.; Losche, M., Zooming in on disordered systems: neutron reflection studies of proteins associated with fluid membranes. *Biochim Biophys Acta* **2014**, *1838* (9), 2341-9.
34. Goers, R.; Thoma, J.; Ritzmann, N.; Di Silvestro, A.; Alter, C.; Gunkel-Grabole, G.; Fotiadis, D.; Müller, D. J.; Meier, W., Optimized reconstitution of membrane proteins into synthetic membranes. *Communications Chemistry* **2018**, *1* (1), 35.
35. Giess, F.; Friedrich, M. G.; Heberle, J.; Naumann, R. L.; Knoll, W., The protein-tethered lipid bilayer: a novel mimic of the biological membrane. *Biophys J* **2004**, *87* (5), 3213-20.
36. Theillet, F. X.; Binolfi, A.; Frembgen-Kesner, T.; Hingorani, K.; Sarkar, M.; Kyne, C.; Li, C. G.; Crowley, P. B.; Gierasch, L.; Pielak, G. J.; Elcock, A. H.; Gershenson, A.; Selenko, P., Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem Rev* **2014**, *114* (13), 6661-6714.
37. Le Brun, Anton P.; Haigh, Cathryn L.; Drew, Simon C.; James, M.; Boland, Martin P.; Collins, Steven J., Neutron Reflectometry Studies Define Prion Protein N-terminal Peptide Membrane Binding. *Biophysical Journal* **2014**, *107* (10), 2313-2324.
38. McGillivray, D. J.; Valincius, G.; Heinrich, F.; Robertson, J. W. F.; Vanderah, D. J.; Febo-Ayala, W.; Ignatjev, I.; Lösche, M.; Kasianowicz, J. J., Structure of Functional *Staphylococcus aureus*  $\alpha$ -Hemolysin Channels in Tethered Bilayer Lipid Membranes. *Biophysical Journal* **2009**, *96* (4), 1547-1553.
39. Holt, S. A.; Le Brun, A. P.; Majkrzak, C. F.; McGillivray, D. J.; Heinrich, F.; Losche, M.; Lakey, J. H., An ion-channel-containing model membrane: structural determination by magnetic contrast neutron reflectometry. *Soft Matter* **2009**, *5* (13), 2576-2586.
40. Shen, H. H.; Leyton, D. L.; Shiota, T.; Belousoff, M. J.; Noinaj, N.; Lu, J.; Holt, S. A.; Tan, K.; Selkrig, J.; Webb, C. T.; Buchanan, S. K.; Martin, L. L.; Lithgow, T., Reconstitution of a nanomachine driving the assembly of proteins into bacterial outer membranes. *Nat Commun* **2014**, *5*, 5078.
41. Giess, F.; Friedrich, M. G.; Heberle, J.; Naumann, R. L.; Knoll, W., The protein-tethered lipid bilayer: A novel mimic of the biological membrane. *Biophysical Journal* **2004**, *87* (5), 3213-3220.
42. Liu, Y.; Yu, J., Oriented immobilization of proteins on solid supports for use in biosensors and biochips: a review. *Microchimica Acta* **2016**, *183* (1), 1-19.
43. Midtgaard, S. R.; Pedersen, M. C.; Kirkensgaard, J. J. K.; Sørensen, K. K.; Mortensen, K.; Jensen, K. J.; Arleth, L., Self-assembling peptides form nanodiscs that stabilize membrane proteins. *Soft Matter* **2014**, *10* (5), 738-752.
44. Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G., Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Letters* **2002**, *2* (8), 853-856.
45. 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. *Journal of the American Chemical Society* **2010**, *132* (39), 13713-13722.
46. Hagn, F.; Eitzkorn, M.; Raschle, T.; Wagner, G., Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. *Journal of the American Chemical Society* **2013**, *135* (5), 1919-1925.
47. Denisov, I. G.; Sligar, S. G., Nanodiscs in Membrane Biochemistry and Biophysics. *Chemical Reviews* **2017**, *117* (6), 4669-4713.
48. Yu, T.-Y.; Raschle, T.; Hiller, S.; Wagner, G., Solution NMR spectroscopic characterization of human VDAC-2 in detergent micelles and lipid bilayer nanodiscs. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2012**, *1818* (6), 1562-1569.
49. Kynde, S. A. R.; 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 Crystallographica Section D* **2014**, *70* (2), 371-383.
50. 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 Crystallographica Section D* **2015**, *71* (12), 2412-2421.
51. Bertram, N.; Laursen, T.; Barker, R.; Bavishi, K.; Moller, B. L.; Cardenas, M., Nanodisc Films for Membrane Protein Studies by Neutron Reflection: Effect of the Protein Scaffold Choice. *Langmuir* **2015**, *31* (30), 8386-91.
52. Wadsater, M.; Laursen, T.; Singha, A.; Hatzakis, N. S.; Stamou, D.; Barker, R.; Mortensen, K.; Feidenhans'l, R.; Moller, B. L.; Cardenas, M., Monitoring Shifts in the Conformation Equilibrium of the Membrane Protein Cytochrome P450 Reductase (POR) in Nanodiscs. *J Biol Chem* **2012**, *287* (41), 34596-34603.
53. Larsen, A. N.; Sorensen, K. K.; Johansen, N. T.; Martel, A.; Kirkensgaard, J. J.; Jensen, K. J.; Arleth, L.; Midtgaard, S. R., Dimeric peptides with three different linkers self-assemble with phospholipids to form peptide nanodiscs that stabilize membrane proteins. *Soft Matter* **2016**, *12* (27), 5937-49.
54. Carlson, M. L.; Young, J. W.; Zhao, Z.; Fabre, L.; Jun, D.; Li, J.; Li, J.; Dhupar, H. S.; Wason, I.; Mills, A. T.; Beatty, J. T.; Klassen, J. S.; Rouiller, I.; Duong, F., The Peptidisc, a simple method for stabilizing membrane proteins in detergent-free solution. *Elife* **2018**, *7*.
55. 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 Mg<sup>2+</sup> transporter. *Nature* **2006**, *440* (7085), 833-837.

56. Matthies, D.; Dalmás, O.; Borgnia, Mario J.; Dominik, Pawel K.; Merk, A.; Rao, P.; Reddy, Bharat 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* (4), 747-756.
57. Neale, C.; Chakrabarti, N.; Pomorski, P.; Pai, E. F.; Pomès, R., Hydrophobic Gating of Ion Permeation in Magnesium Channel CorA. *PLOS Computational Biology* **2015**, *11* (7), e1004303.
58. Niegowski, D.; Eshaghi, S., The CorA family: Structure and function revisited. *Cellular and Molecular Life Sciences* **2007**, *64* (19), 2564.
59. Mackman, N., The many faces of tissue factor. *J Thromb Haemost* **2009**, *7 Suppl 1*, 136-9.
60. Ansari, S. A.; Pendurthi, U. R.; Sen, P.; Rao, L. V. M., The Role of Putative Phosphatidylserine-Interactive Residues of Tissue Factor on Its Coagulant Activity at the Cell Surface. *PLOS ONE* **2016**, *11* (6), e0158377.
61. Banner, D. W.; D'Arcy, A.; Chene, C.; Winkler, F. K.; Guha, A.; Konigsberg, W. H.; Nemerson, Y.; Kirchhofer, D., The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature* **1996**, *380* (6569), 41-46.
62. McCallum, C. D.; Hapak, R. C.; Neuenschwander, P. F.; Morrissey, J. H.; Johnson, A. E., The location of the active site of blood coagulation factor VIIa above the membrane surface and its reorientation upon association with tissue factor. A fluorescence energy transfer study. *J Biol Chem* **1996**, *271* (45), 28168-75.
63. McCallum, C. D.; Su, B.; Neuenschwander, P. F.; Morrissey, J. H.; Johnson, A. E., Tissue factor positions and maintains the factor VIIa active site far above the membrane surface even in the absence of the factor VIIa Gla domain. A fluorescence resonance energy transfer study. *J Biol Chem* **1997**, *272* (48), 30160-6.
64. Palombo, I.; Daley, D. O.; Rapp, M., The periplasmic loop provides stability to the open state of the CorA magnesium channel. *The Journal of biological chemistry* **2012**, *287* (33), 27547-27555.
65. Braun, L.; Uhlig, M.; von Klitzing, R.; Campbell, R. A., Polymers and surfactants at fluid interfaces studied with specular neutron reflectometry. *Adv Colloid Interfac* **2017**, *247*, 130-148.
66. Fragneto-Cusani, G., Neutron reflectivity at the solid/liquid interface: examples of applications in biophysics. *J Phys-Condens Mat* **2001**, *13* (21), 4973-4989.
67. <http://sourceforge.net/projects/rscl/>.
68. Emil Wolf, M. B., *Principles of Optics*. Pergamon Press: 1970.
69. Luchini, A.; Nzulumike, A. N. O.; Lind, T. K.; Nylander, T.; Barker, R.; Arleth, L.; Mortensen, K.; Cárdenas, M., Towards biomimics of cell membranes: Structural effect of phosphatidylinositol triphosphate (PIP3) on a lipid bilayer. *Colloids and Surfaces B: Biointerfaces* **2019**, *173*, 202-209.
70. Cans, A. S.; Hook, F.; Shupliakov, O.; Ewing, A. G.; Eriksson, P. S.; Brodin, L.; Orwar, O., Measurement of the dynamics of exocytosis and vesicle retrieval at cell populations using a quartz crystal microbalance. *Anal Chem* **2001**, *73* (24), 5805-5811.
71. Höök, F., *Development of a Novel QCM Technique for Protein Adsorption Studies*. Department of Applied Physics, Chalmers University of Technology: Göteborg, 1997.
72. Horcas, I.; Fernandez, R.; Gomez-Rodriguez, J. M.; Colchero, J.; Gomez-Herrero, J.; Baro, A. M., WSXM: a software for scanning probe microscopy and a tool for nanotechnology. *Rev Sci Instrum* **2007**, *78* (1), 013705.
73. Lind, T. K.; Cardenas, M., Understanding the formation of supported lipid bilayers via vesicle fusion-A case that exemplifies the need for the complementary method approach (Review). *Biointerphases* **2016**, *11* (2), 020801.
74. Lind, T. K.; Cardenas, M.; Wacklin, H. P., Formation of Supported Lipid Bilayers by Vesicle Fusion: Effect of Deposition Temperature. *Langmuir* **2014**, *30* (25), 7259-7263.
75. Johansen, N. T.; Pedersen, M. C.; Porcar, L.; Martel, A.; Arleth, L., Introducing SEC-SANS for studies of complex self-organized biological systems. *Acta Crystallographica Section D* **2018**, *74* (12), 1178-1191.
76. Anderson, T. H.; Min, Y.; Weirich, K. L.; Zeng, H.; Fyngson, D.; Israelachvili, J. N., Formation of supported bilayers on silica substrates. *Langmuir* **2009**, *25* (12), 6997-7005.
77. Margeat, E.; Le Grimmeléc, C.; Royer, C. A., Visualization of trp repressor and its complexes with DNA by atomic force microscopy. *Biophys J* **1998**, *75* (6), 2712-20.
78. Rao, L. V.; Pendurthi, U. R., Regulation of tissue factor coagulant activity on cell surfaces. *J Thromb Haemost* **2012**, *10* (11), 2242-53.
79. Shaw, A. W.; Pureza, V. S.; Sligar, S. G.; Morrissey, J. H., The local phospholipid environment modulates the activation of blood clotting. *J Biol Chem* **2007**, *282* (9), 6556-6563.
80. Tavooosi, N.; Davis-Harrison, R. L.; Pogorelov, T. V.; Ohkubo, Y. Z.; Arcario, M. J.; Clay, M. C.; Rienstra, C. M.; Tajkhorshid, E.; Morrissey, J. H., Molecular Determinants of Phospholipid Synergy in Blood Clotting. *J Biol Chem* **2011**, *286* (26), 23247-23253.
81. Sen, P.; Neuenschwander, P. F.; Pendurthi, U. R.; Rao, L. V., Analysis of factor VIIa binding to relipidated tissue factor by surface plasmon resonance. *Blood Coagul Fibrinolysis* **2010**, *21* (4), 376-9.

---

Insert Table of Contents artwork here

---

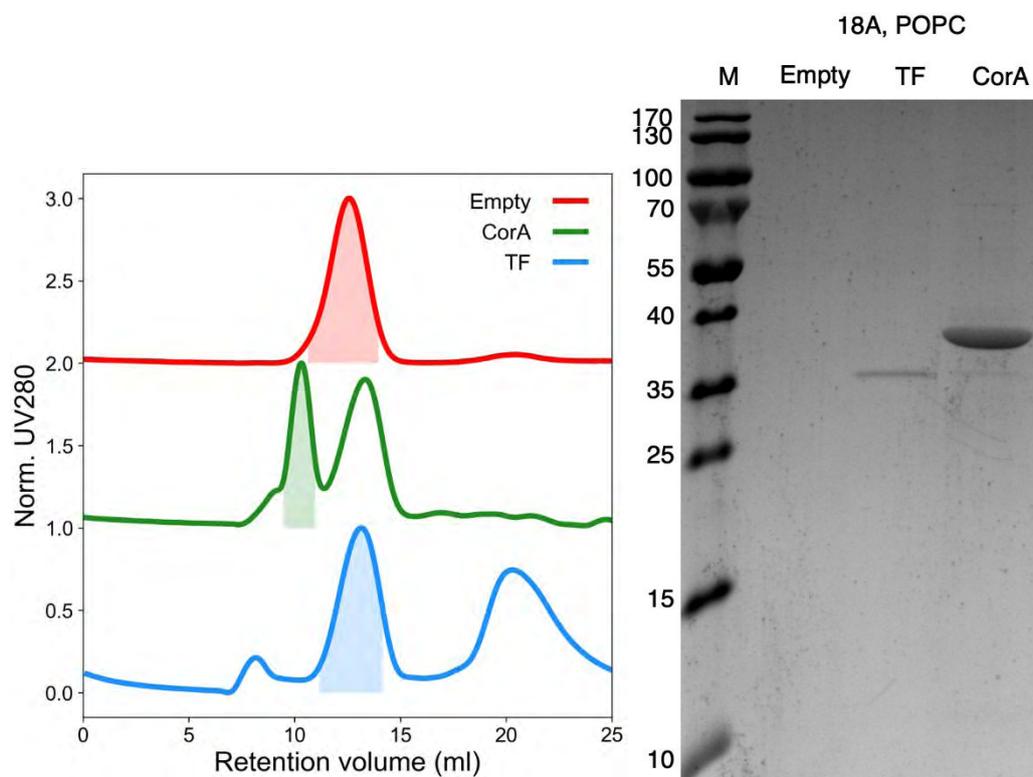


Formation of a supported lipid bilayer with oriented membrane proteins by deposition on a solid support of protein-loaded peptide discs

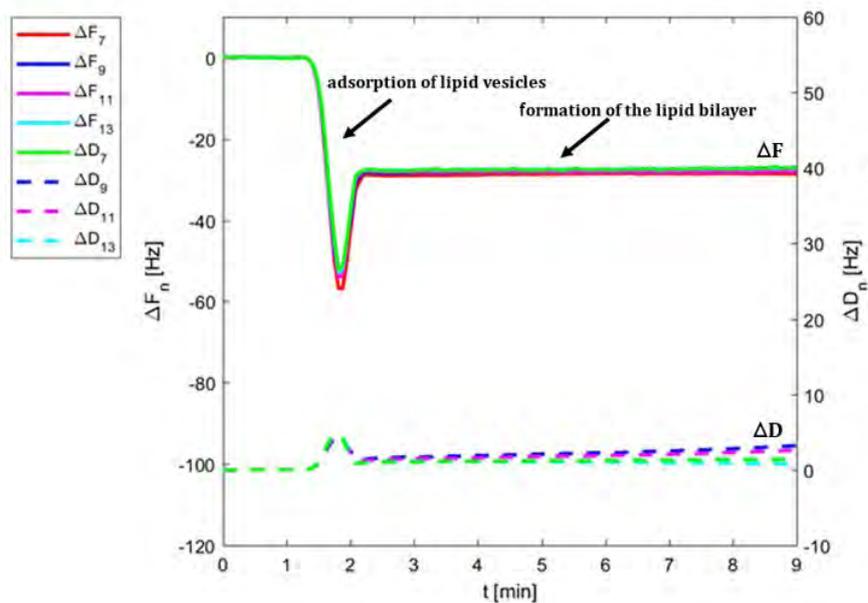
---

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Supporting Information



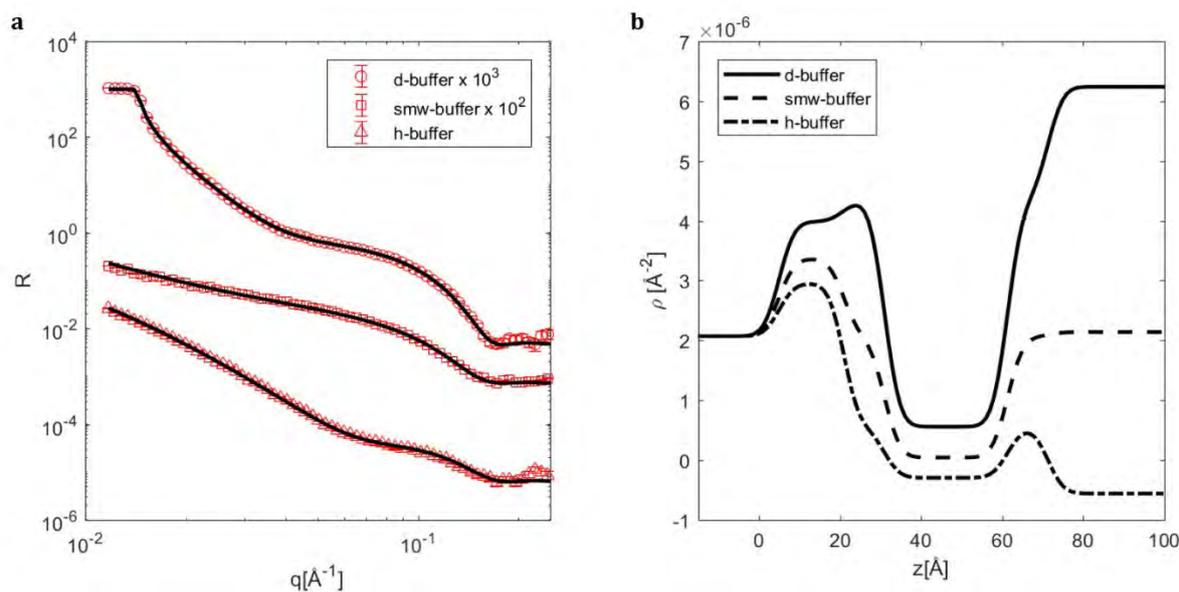
**Figure S1-** Size Exclusion Chromatogram (left) showing the characteristic peaks for the elution of the empty peptide disc, the TF-loaded peptide discs and the CorA-peptide discs. Picture of the ... gel confirming the presence of TF and CorA in the selected fraction from the SEC purification.



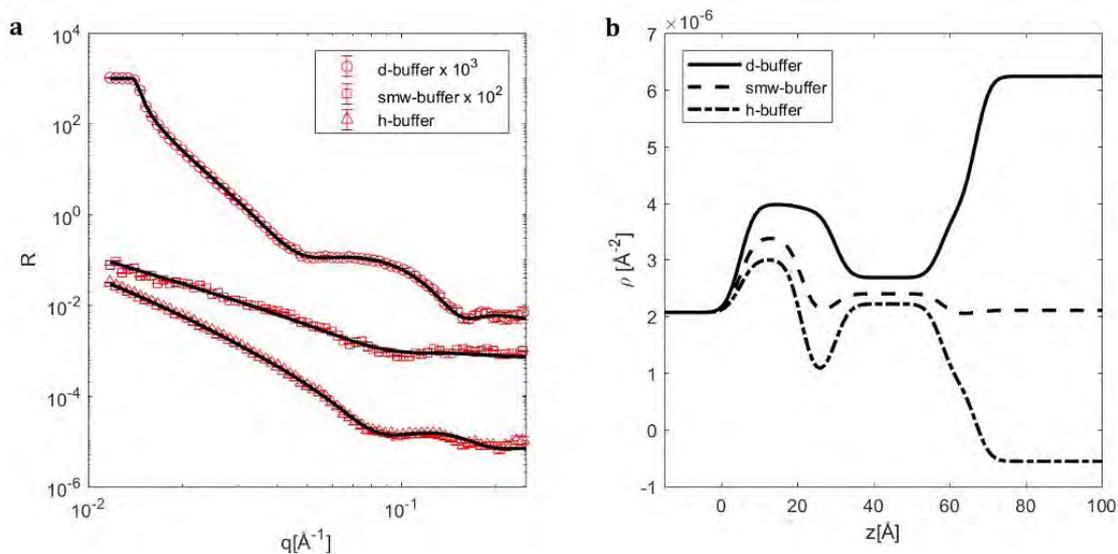
**Figure S2-** QCM-D data monitoring the formation of a POPC lipid bilayer by vesicle fusion. Intact vesicles initially adsorb on the substrate and subsequently fuse to form a supported lipid bilayer ( $\Delta F \approx -25$  Hz). Measurements were performed at 25°C and by using 1.3 mM POPC solution and 20 mM Tris-HCl, 100 mM NaCl, 10 mM CaCl<sub>2</sub> (pH=7.5) buffer.

**Table S1-** Structural parameters obtained from the analysis of the NR data collected for the POPC, POPC/POPS and dPOPC/POPS supported lipid bilayer formed *via* peptide-disc injection.

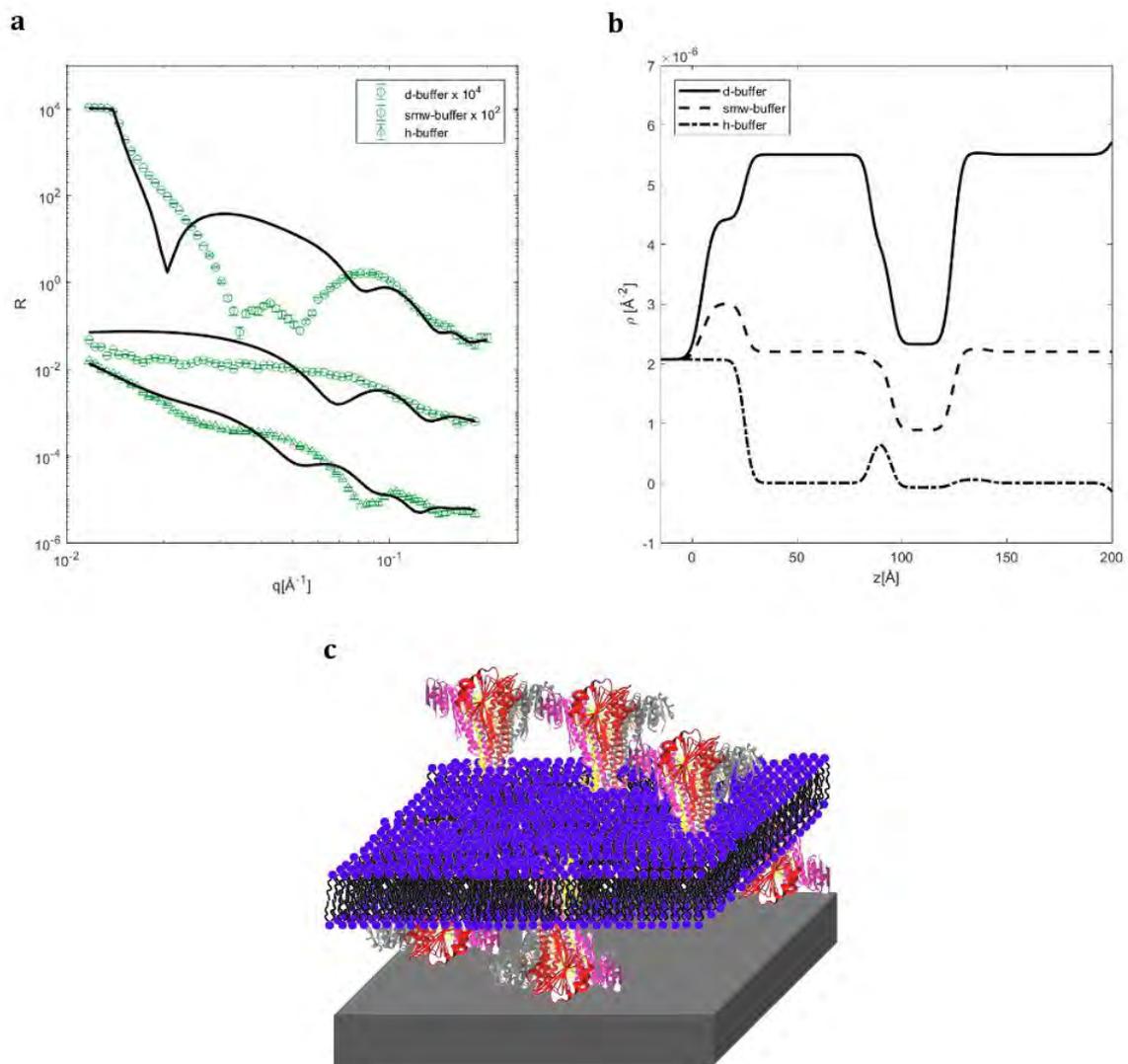
	POPC	POPC/POPS 70/30 mol/mol	dPOPC/POPS 70/30 mol/mol
<b>Lipid headgroup</b>			
Thickness (Å)	8 ± 1	9 ± 1	9 ± 1
Scattering length density (Å <sup>-2</sup> )	1.8 ± 0.1	1.9 ± 0.1	1.9 ± 0.1
Solvent volume fraction (%)	50 ± 3	56 ± 3	45 ± 5
Roughness (Å)	3 ± 1	3 ± 1	3 ± 1
<b>Lipid acyl chain</b>			
Thickness (Å)	31 ± 1	30 ± 1	28 ± 2
Scattering length density (Å <sup>-2</sup> )	-0.29 ± 0.04	-0.27 ± 0.08	2.46 ± 0.07
Solvent volume fraction (%)	8 ± 1	12 ± 1	7 ± 2
Roughness (Å)	3 ± 1	3 ± 1	3 ± 1



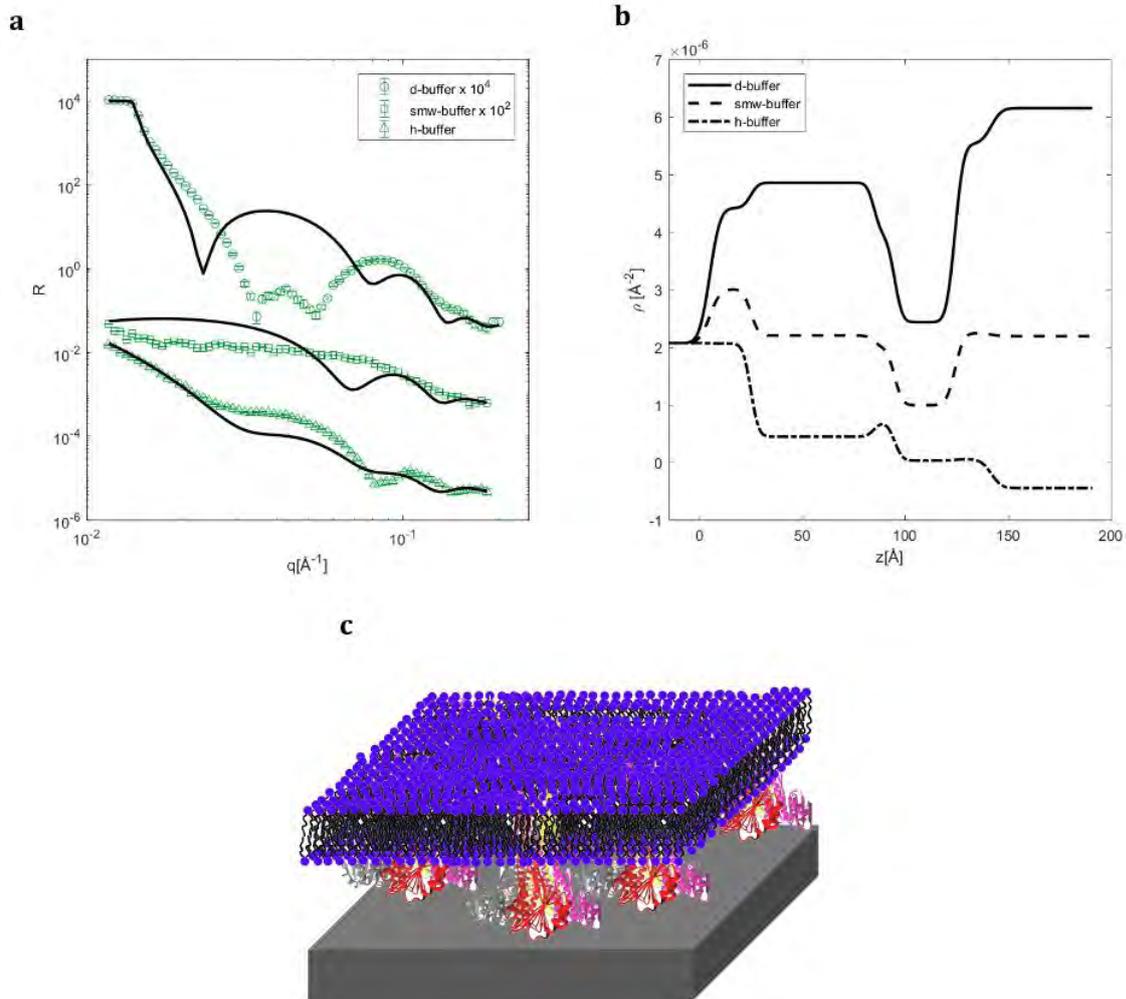
**Figure S3-** a) NR experimental data together with the relative fitting curves for the POPC/POPS bilayer. Data corresponding to the different buffer compositions are reported with different symbols and multiplied by a scale factor as reported in the legend. b) Scattering length density profile calculated from experimental data analysis. Measurements were performed at 25°C with a peptide disc solution prepared with 1.3 mM POPC/POPS (70/30 mol/mol) and 0.43 mM 18A solution and 20mM Tris-HCl, 100mM NaCl, 10mM CaCl<sub>2</sub> (pH=7.5) buffer.



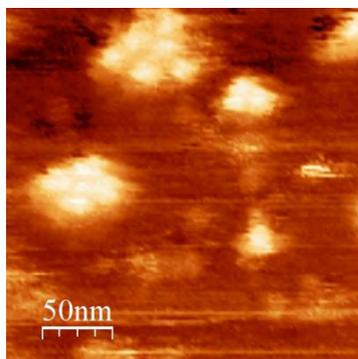
**Figure S4-** a) NR experimental data together with the relative fitting curves for the dPOPC/POPS bilayer. Data corresponding to the different buffer compositions are reported with different symbols and multiplied by a scale factor as reported in the legend. b) Scattering length density profile calculated from experimental data analysis. Measurements were performed at 25°C with a peptide disc solution prepared with 1.3 mM dPOPC/POPS (70/30 mol/mol) and 0.43 mM 18A solution and 20mM Tris-HCl, 100mM NaCl, 10mM CaCl<sub>2</sub> (pH=7.5) buffer.



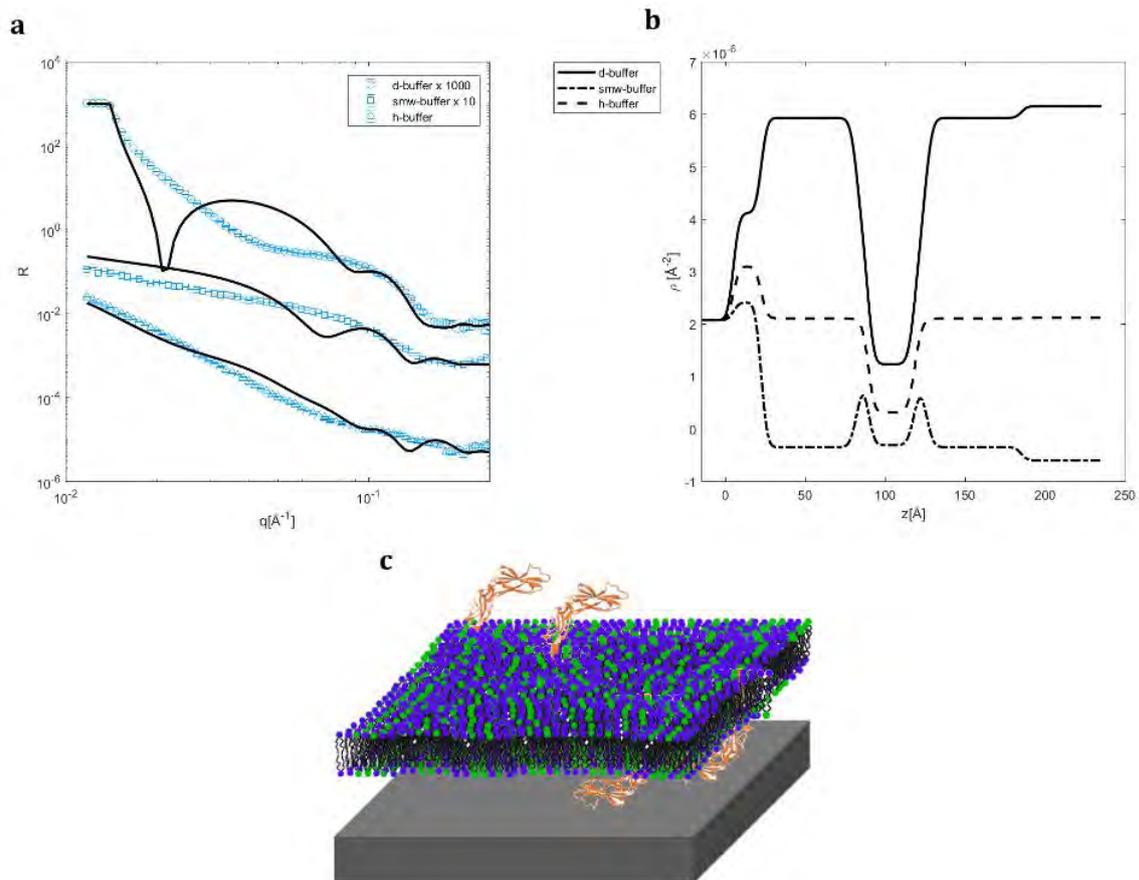
**Figure S5-** Comparison between the NR experimental data for the POPC supported membrane with CorA and the model assuming the protein molecules to be randomly oriented with the ECD both in contact with the substrate and pointing towards the solvent.



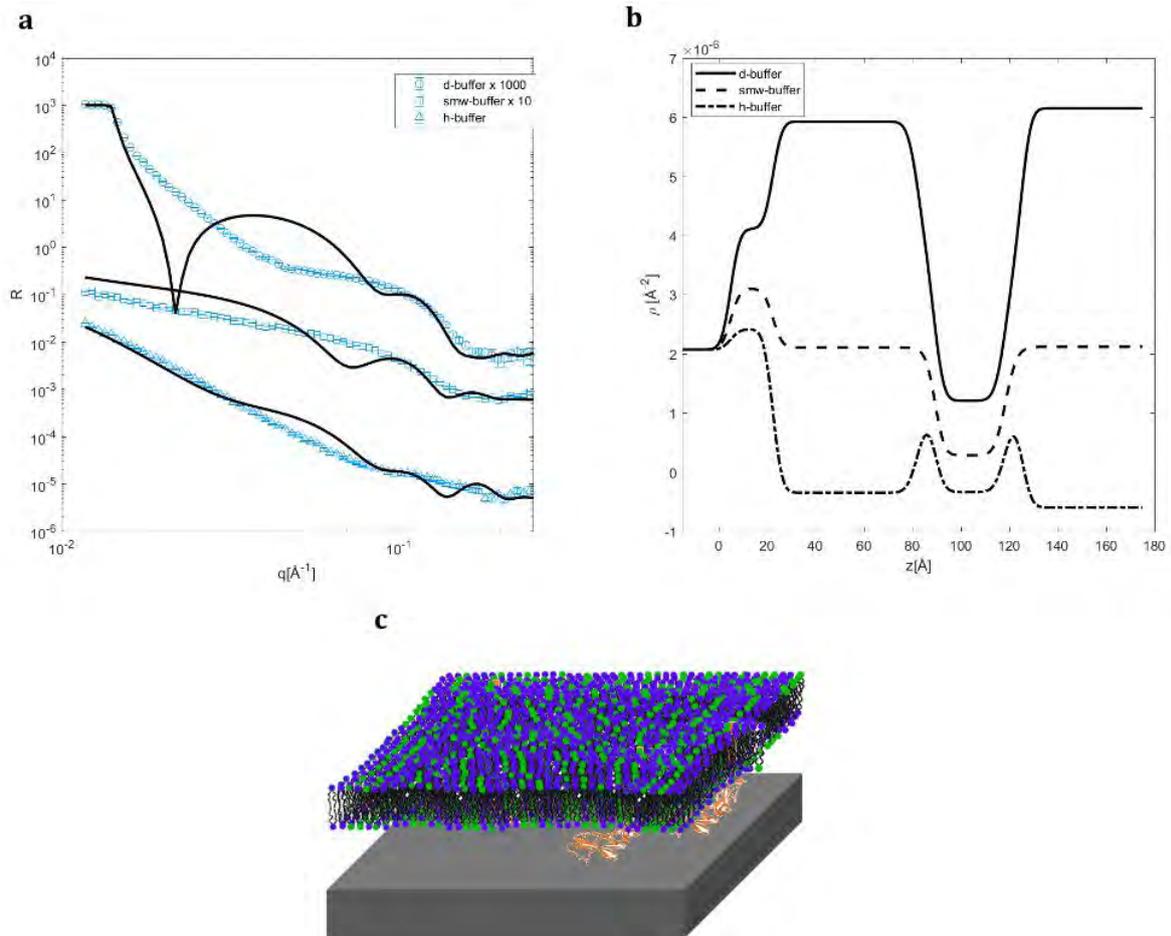
**Figure S6-** Comparison between the NR experimental data for the POPC supported membrane with CorA and the model assuming the protein molecules to be oriented with the ECD in contact with the substrate and pointing towards the solvent.



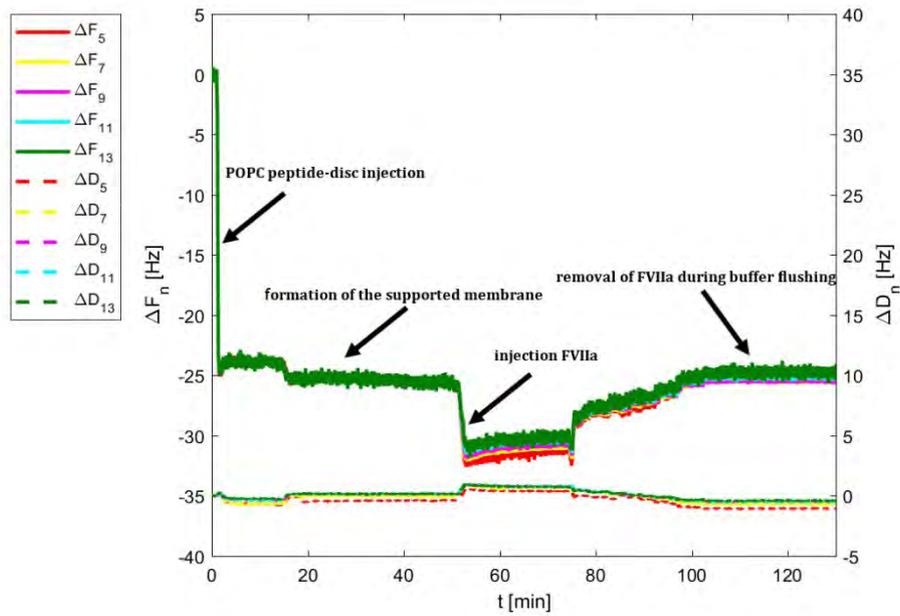
**Figure S7-** AFM image collected for the POPC supported bilayer with CorA in 20mM Tris, 150mM NaCl buffer pH=7 (color height scale 12 nm). It can be seen how larger features consist of self-associated CorA individual protein molecules.



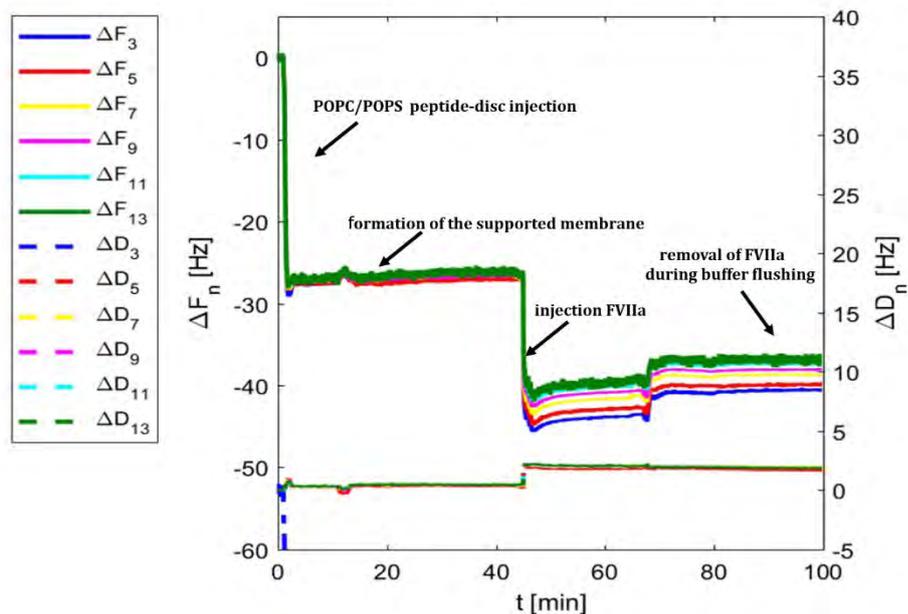
**Figure S8-** Comparison between the NR experimental data for the POPC/POPS supported membrane with TF and the model assuming the protein molecules to be randomly oriented with the ECD both in contact with the substrate and pointing towards the solvent.



**Figure S9-** Comparison between the NR experimental data for the POPC/POPS supported membrane with TF and the model assuming the protein molecules to be oriented with the ECD both in contact with the substrate.



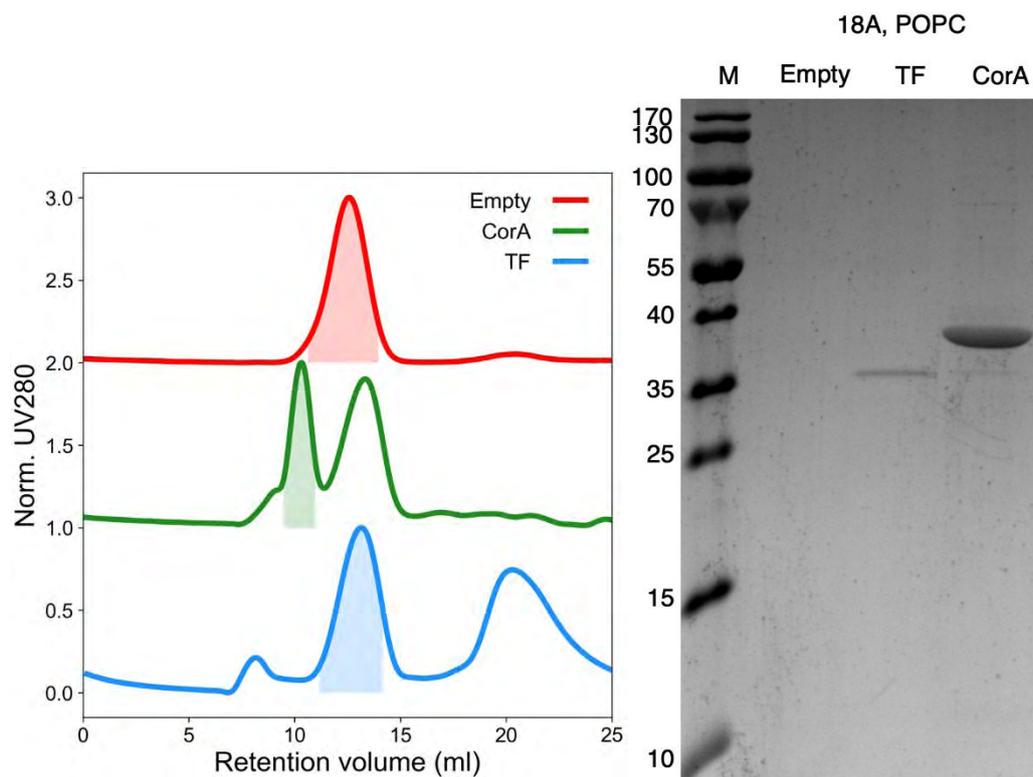
**Figure S9:** QCM-D data monitoring the interaction between FVIIa and the POPC supported membrane which was formed by POPC peptide-disc injection. FVIIa initially interacts with the membrane, but the interaction is very weak as all the adsorbed molecules are washed away during buffer flushing. The measurement was performed at 25°C and the peptide discs were prepared with 1.3 mM POPC, 0.43mM 18A and 20mM Tris-HCl, 100mM NaCl, 10mM CaCl<sub>2</sub> (pH=7.5) buffer. After the formation of the supported membrane a 10 $\mu$ M solution of FVIIa was injected.



**Figure S10:** QCM-D data monitoring the interaction between FVIIa and the POPC/POPS supported membrane which was formed by POPC/POPS peptide-disc injection. FVIIa initially interacts with the membrane, and the adsorbed protein molecules are only partially removed during buffer flushing. The measurement was performed at 25°C and the peptide discs were prepared with 1.3 mM POPC/POPS (70/30 mol/mol), 0.43mM 18A and 20mM Tris-HCl, 100mM NaCl, 10mM CaCl<sub>2</sub> (pH=7.5) buffer. After the formation of the supported membrane a 10 $\mu$ M solution of FVIIa was injected.

**TABLE S2-** FVIIa binding properties of TF-embedded lipid surface

<i>“age” of surface</i>	$k_{on}$ ( $10^5 \text{ M}^{-1}\text{s}^{-1}$ )	$k_{off}$ ( $10^{-5} \text{ s}^{-1}$ )	$K_D$ (nM)	Capacity (RU)
3 hrs	$6.31 \pm 0.08$	$13.6 \pm 0.04$	0.215	$43.6 \pm 0.29$
24 hrs	$6.38 \pm 0.07$	$15.8 \pm 0.08$	0.248	$31.0 \pm 0.19$
48 hrs	$6.94 \pm 0.06$	$8.96 \pm 0.04$	0.129	$19.6 \pm 0.02$
144 hrs	$4.06 \pm 0.06$	$13.1 \pm 0.08$	0.322	$14.9 \pm 0.02$

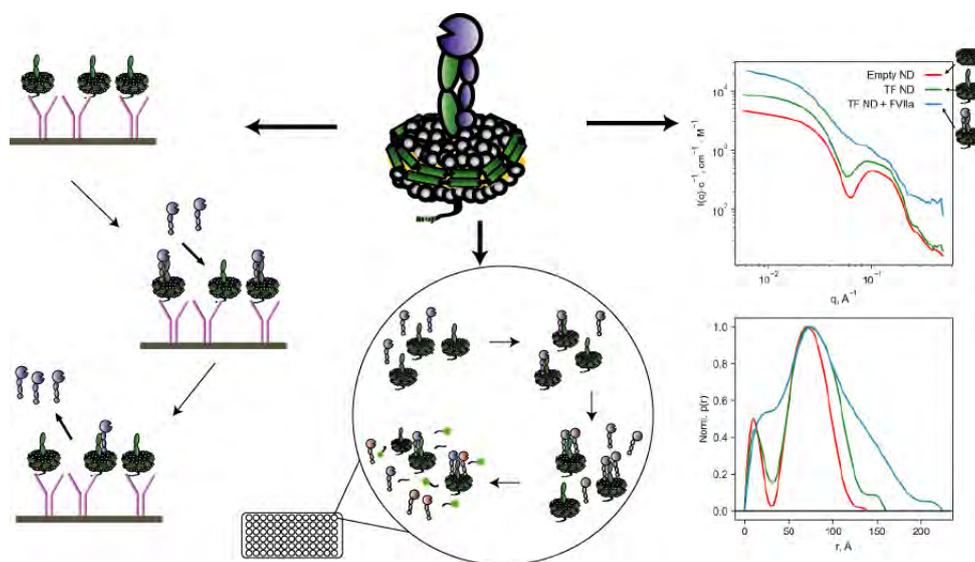


**Figure S11-** Size Exclusion Chromatogram (left) showing the characteristic peaks for the elution of the empty peptide disc, the TF-loaded peptide discs and the CorA-peptide discs. Picture of the ... gel confirming the presence of TF and CorA in the selected fraction from the SEC purification.

## 8 | Paper IV

### Efficient refolding and reconstitution of Tissue Factor into nanodiscs facilitates structural investigation of a multicomponent system on a lipid bilayer

*Frederik Grøn­bæk Tidemand, Henrik Østergaard, Michael Ploug, Birthe B. Kragelund and Lise Arleth*



# Efficient refolding and reconstitution of Tissue Factor into nanodiscs facilitates structural investigation of a multicomponent system on a lipid bilayer

Frederik Grøn­bæk Tidemand<sup>1</sup>, Henrik Østergaard<sup>2</sup>, Michael Ploug<sup>3,4</sup>, Birthe B. Kragelund<sup>5</sup>, Lise Arleth<sup>1</sup>

1: Structural Biophysics, X-ray and Neutron Science, The Niels Bohr Institute, University of Copenhagen, Denmark. 2: Novo Nordisk A/S, Global Research Technologies, DK-2760, Maaloev, Denmark. 3: Finsen Laboratory, Rigshospitalet, DK-2200 Copenhagen N, Denmark. 4: Biotech Research and Innovation Centre (BRIC), University of Copenhagen, DK-2200 Copenhagen N, Denmark. 5: Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Denmark

**KEYWORDS.** *Tissue Factor. Nanodisc. Factor VIIa. SAXS. Refolding. Membrane protein.*

---

**ABSTRACT:** Structural data on membrane proteins in a lipid membrane environment is challenging to obtain but needed to provide information on the often essential protein-lipid interplay. A common experimental bottleneck in obtaining such data is providing samples in sufficient amounts and quality required for structural studies. We developed a new expression and purification protocol for the single-pass membrane protein Tissue Factor (TF), exploiting the high expression level in *E. coli* inclusion bodies and subsequent refolding. This provided more than 5 mg of functional TF per liter bacterial culture. This is substantially more than what was obtained by the classical approaches for expressing TF in the membrane-anchored configuration. We optimized reconstitution into circularized nanodiscs enabling the formation of stable, TF loaded nanodiscs with different lipid compositions and with a limited material waste. The blood coagulation cascade is initiated by the complex formation between TF and Factor VIIa (FVIIa), and we probed this interaction by a functional activity assay and SPR measurements, which revealed similar activity and binding kinetics as TF produced by other protocols, demonstrating that high-yield production does not compromise TF function. Furthermore, the amounts of sample produced permitted initial small angle X-ray scattering studies providing the first structural information about TF and its binding to FVIIa in a lipid environment. This strategy possibly allows for probing the multicomponent complex TF:FVIIa together with its macromolecular substrate Factor X on a lipid bilayer, but may also be relevant as a production strategy for other single pass-membrane proteins for which structural information, in general, is limited.

---

Membrane proteins (MPs) maintain numerous functions in the human body, but despite their biological importance, structures of MPs remain in general sparse. MPs are notoriously difficult to produce and the main roadblock to such progress is the production and sample preparation of the mg-quantities of sample that is needed for structural studies. Even though up to one-third of the human genome encodes MPs, less than 3% of the entries in the PDB database are MPs, highlighting the gap in knowledge on structures of this particular class of proteins (1,2). In particular, the structures of the single-pass transmembrane proteins (SPTMP) remain enigmatic with no high-resolution structure of full-length proteins being available despite constituting 6% of the human genome (1). Many proteins in this class consist of a folded extracellular domain (ECD), an  $\alpha$ -helical transmembrane domain (TMD) and a disordered intracellular domain (ICD) (3).

Experiments on the full-length SPTMPs are generally hampered by the very different nature of the three domains. Therefore, a divide-and-conquer strategy has previously been applied, integrating structural data on the isolated domains using simulations to yield a final model of the entire protein (3).

Tissue factor (TF), which is the focus of the present study, is an example of an SPTMP. TF comprises a folded ECD, termed soluble TF (sTF) when produced separately, a helical TMD, and a short, likely disordered ICD. The crystal structure of the sTF was published in 1996 (4) and showed two fibronectin domains each containing one disulfide bridge. To the best of the authors' knowledge, no structural information exist on the TMD of TF

and little is known about the ICD (5). TF is a central player in the blood coagulation cascade and is located on the surface of cells adjacent to blood vessels. Upon vascular injury, TF is exposed to blood and initiates coagulation by engaging in a high-affinity complex with coagulation factor VIIa (FVIIa) (6). This complex serves to activate the macromolecular substrate Factor X (FX) through a transitory quaternary FVIIa:TF:FX:phospholipid membrane complex (7,8). The lipid environment plays a vital role in this complex, and especially the negatively charged phosphatidylserine (PS) lipids are essential for the activation of FX by the TF:FVIIa complex. PS lipids are usually distributed in the inner leaflet of the membrane but is upon injury flipped to the outer leaflet (9,10). Formation of the TF:FVIIa complex increases the catalytic activity of FVIIa by a factor  $10^5$  compared to free FVIIa, and increasing the PS content also increases the FVIIa activity by more than a factor of 10 (11). The affinity between recombinantly produced ECD-TMD of TF and FVIIa measured by surface plasmon resonance (SPR) showed that the  $K_d$  between TF in a membrane and FVIIa was in the pM range, but was not altered by changing the lipid composition (11,12). On the other hand, the affinity between the sTF and FVIIa was in the nM range emphasizing the importance of lipids for the interaction.

Even though a high-resolution structure of sTF was published more than 20 years ago, there is to date no structural information on TF in a lipid microenvironment even though the affinity and activity of the complex are profoundly affected by

lipids (13). The only structural experiments performed on TF in a bilayer are FRET studies investigating the distance between the membrane and the active site of FVIIa. This distance was measured to 82Å for membrane-bound FVIIa and changed upon TF binding to 74Å, suggesting that TF promotes a slightly tilted orientation of FVIIa (14,15). These results might be important for elucidating the mechanism behind the increase in activity upon TF binding, but more structural data is needed to do so.

Upon activation of FX by the TF:FVIIa complex, it is released from the complex and can activate prothrombin to thrombin, which effectively results in fibrin and clot formation (16). FVIIa and FXa have the same overall structure including a C-terminal Gla-domain and an N-terminal catalytic domain containing the active site (see figure 1 for an illustration of the ternary complex on a lipid bilayer) (4,17). The Gla-domain has been shown to have specific lipid binding motifs, and both FVIIa and FXa have affinities for a membrane containing PS in the  $\mu\text{M}$  range (11,18).

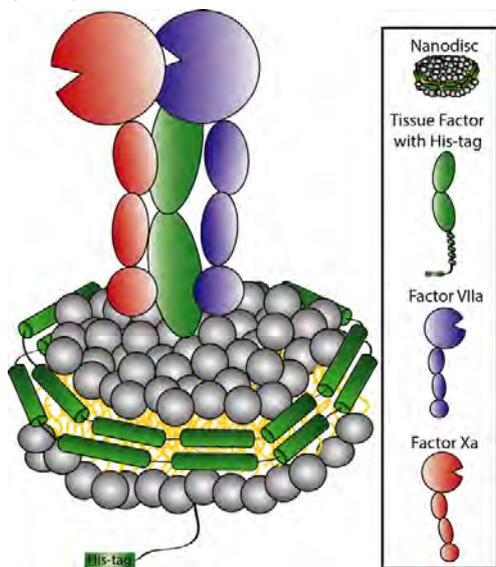


Figure 1: Multicomponent complexes in membranes. Model of a nanodisc with TF incorporated and FVIIa and FXa bound.

The structure of the complex sTF:FVIIa was published in 1996 (4) but to date there exists no published high-resolution structural information of the ternary TF:FVIIa:FXa complex. This complex can only be formed on a lipid bilayer, and still only transiently, which poses enormous experimental challenges. Furthermore, producing TF, in sufficient amounts for structural studies represents a severe roadblock for dual reasons. *i)* Expression levels in a membrane anchored conformation are typically low, as the plasma membrane already before expression is very crowded and consists of about 50 % MPs and *ii)* the purification has to be performed using a detergent to keep TF in solution (19). The identification of a non-denaturing, but still solubilizing detergent, depends on the individual MP and can be tricky. In addition, the polydisperse nature of detergent micelles often impose limitations on structural experiments, and therefore, besides producing TF in sufficient amounts, an optimal carrier system for maintaining the TF in a structurally and functionally relevant form must be found for such studies. For TF a bilayer is needed for the assembly of the ternary complex with FVIIa and FXa, requiring

a lipid containing carrier system. The nanodisc (ND), which consists of a lipid bilayer surrounded by two amphipathic proteins called membrane-scaffolding proteins (MSPs) provides exactly this bilayer. Besides that NDs generally has better performance than detergents with respect to maintaining the activity of an MP(20), it is also very well suited for many types of structural studies due to a monodisperse nature and the possibility to control the size and lipid contents (21,22). Recently, a circularized MSP was developed using the Sortase enzyme, which covalently links the two termini of the MSP enabling even less heterogeneity (23). These covalently linked circularized discs were further improved with respect to solubility and expression providing yields in the range of 75 mg MSP per liter culture, ready for assembly of more stable, homogeneous and soluble nanodiscs (24). Thus, this new generation of NDs provides an optimal platform for investigating lipid-dependent multicomponent systems. The system is easy to tune regarding lipid composition and the size of the bilayer, which allows binding of other proteins, and confers the high stability and homogeneity that is required for structural studies.

In this study, we present an effective, high-yield recombinant expression and subsequent *in vitro* refolding of the ECD and TMD of human TF. The obtained TF was subsequently successfully reconstituted into circularized nanodiscs, allowing us to probe the FVIIa affinity through activity assays and surface plasmon resonance (SPR). Finally, we show that the amounts and quality of the samples were sufficient for structural studies using size exclusion chromatography coupled to small angle X-ray scattering (SEC-SAXS). Our results show how the SPTMP TF can be produced with optimized yield, reconstituted in nanodiscs and structurally studied both on its own and in complex with FVIIa. This is a central prerequisite for performing structural studies of the multicomponent TF-FVIIa and TF-FVIIa-FXa systems. The successful preparation of these difficult accessible samples will enable closing the gap in our knowledge of the structural biology of the blood clotting cascade in particular and provide a model system and generalizable experience for the handling of membrane bound multicomponent systems more generally.

## Materials and methods

Unless otherwise stated all chemicals and lipids are from Sigma Aldrich.

### Expression and purification of Tissue factor

The gene encoding the extracellular and transmembrane domain of human TF (Uniprot ID: P13726) (1–244), a C-terminal linker (G<sub>4</sub>S), and a His<sub>6</sub>-tag was synthesized and cloned into the pET 28a vector by Genscript. The plasmid harboring the TF gene was transformed into BL21 Star cells, which were grown at 37°C in Terrific Broth medium. After an overnight inoculation, the cells were transferred to 2 L of preheated TB media and protein expression was induced with IPTG to a final concentration of 1 mM at an optical density at 600 nm of 0.8. The cells were harvested by centrifugation after 3 hours. The pellet was resuspended in 3.5 ml lysis buffer (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 1 mM PMSF) per gram of cells and passed through a cell disrupter (E1061, Constant Systems LTD) at 30kPsi. Inclusion bodies (IB) were harvested by centrifugation at 25000 g for 20 min at 4°C, and resuspended in the same volume as before using a homogenization glass piston in 20 mM Tris, pH 8, 150 mM NaCl, 0.5 M Urea and

0.25 % (w/v) Tween20. The IBs were harvested by centrifugation (25,000 g for 20 min) and subsequently dissolved in the same volume as before, using the same homogenization apparatus in 6 M GuHCl, 10 mM Tris-HCl, pH 8, 50 mM NaCl, 2 mM EDTA and 0.1% Triton X-100 (TX). Cellular debris was removed by centrifugation (25000 g for 20 min), and the supernatant was dripped with a flow of 1 ml/min into a reservoir containing cold (4°C) refolding buffer (50mM Tris-HCl, pH 8.5, 250mM NaCl, 0.1% TX) in 60 times the volume of the supernatant, under fast magnetic stirring. The solution was left in a cold room for 24 hours under magnetic stirring. Imidazole was added to a final concentration of 20 mM, and aggregated protein was removed by centrifugation at 17,000 g for 20 min. The supernatant was loaded onto a pre-equilibrated 5 ml HisTrap (GE Healthcare) (20 mM Tris-HCl, pH 8, 100 mM NaCl, 0.1% TX and 20 mM Imidazole), washed with 5 column volumes (CV) of washing buffer (20 mM Tris-HCl, pH 8, 100 mM NaCl, 0.1% TX and 40 mM Imidazole) and eluted with 5 CV elution buffer (20 mM Tris-HCl, pH 8, 100 mM NaCl, 0.1% TX and 250 mM Imidazole) while fractions of 1 ml were collected. Fractions were evaluated by SDS-PAGE, and those containing the most concentrated TF were pooled and dialyzed against 100 times the volume of 20 mM Tris-HCl, pH 8, 10 mM NaCl, 0.05% TX. As TX absorbs light at 280 nm, TF concentrations were determined either by gel quantification using sTF standards (provided by Novo Nordisk A/S) or by BCA kit (Sigma Aldrich). The solution containing TF was aliquoted into 0.5 mL fractions and flash frozen in liquid nitrogen and stored at -80°C until use.

#### Estimation of refolding yields

Each small-scale refolding experiment was carried out using 1 ml of the GuHCl preparation containing TF and diluting this in 5 to 59 ml of 4°C refolding buffer (50mM Tris-HCl, pH 8.5, 250mM NaCl, 0.1% TX) under magnetic stirring in a cold-room. Samples of relevant volume (between 6 and 60 ml) were extracted at timepoints placed between 2 min and 24 hours, centrifuged at 12,000 g for 10 min and 200 µl of equilibrated NiNTA resin (20 mM Tris-HCl, pH 8, 100 mM NaCl, 0.1% TX and 20 mM Imidazole) was added to the supernatant. The sample was incubated for 45 min at room temperature (RT) and loaded onto an empty gravity flow column and washed with 2 ml of washing buffer (20 mM Tris-HCl, pH 8, 100 mM NaCl, 0.1% TX and 40 mM Imidazole) and eluted with 1 ml elution buffer (20 mM Tris-HCl, pH 8, 100 mM NaCl, 0.1% TX and 250 mM Imidazole). 40 µl of each elution was mixed with 10µl SDS loading buffer (250 mM Tris-HCl pH 6.8, 50% glycerol, 10% w/v SDS, 0.25% w/v bromophenol blue) and boiled at 95°C for 5 min before centrifugation at 13,000g for 5 min. Standards of sTF (kindly provided by the company Novo Nordisk A/S) were mixed in the same ratio and treated similarly before all samples were run on 15% 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). The images were quantified in the ImageLab software (Bio-Rad) using default settings. Relative gel band intensities were used to calculate refolding yields using sTF samples for normalization. Western blots (using primary TF antibody, TF8, subclone 5G9, provided by Novo Nordisk A/S) were used to estimate the ratio between TF in solution and TF in the pellet, using the same imaging software. All NiNTA elutions were further analyzed for homogeneity and folding using a Superdex200 10/300 GL equilibrated in 20 mM Tris,

pH 8, 100 mM NaCl, 0.1% TX. The ratio of soluble protein and monodisperse protein was estimated as the area under the main peak divided by the sum of the area under the main peak and the area under the void peak.

#### Nanodisc preparation

1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS) were dissolved in cholate buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 100 mM sodium cholate) to a final lipid concentration of 50 mM. The MSP used in this study is the circularized and solubility enhanced MSP1E3D1 (csE3) which was expressed and purified as previously described (24). All MP-free nanodiscs were assembled with a lipid:MSP ratio of 130:1 and all TF containing nanodiscs were assembled with a lipid:MSP ratio of 120:1 and a MSP:TF ratio ranging from 20 to 4. All preparations had a final lipid concentration of 10 mM. Samples were mixed and incubated for 15 min at room temperature (RT) before detergent absorbing beads (Amberlite XAD-2, Sigma) were added at a ratio of 15% w/v. Samples were incubated at 10°C while shaking at 800 rpm overnight to remove detergent. The samples were separated from the detergent absorbing beads by making a small hole in the bottom and top of the Eppendorf tube, putting the tube in the top of a 15 ml falcon tube (without lid) and centrifugate at 300 g for 3 min. The samples with TF added were applied to a 1 ml HisTrap column (GE Healthcare) equilibrated in 20 mM Tris-HCl pH 8, 100 mM NaCl, and 20 mM imidazole to bind TF loaded nanodiscs, which subsequently were eluted with 300 mM imidazole while fractions of 1 ml were collected. Fractions containing the TF were pooled, diluted 2-fold in MiliQ water, concentrated to 0.5 ml and applied to a Superdex 200 10/300 GL column (GE Healthcare) equilibrated in gel filtration buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM CaCl<sub>2</sub>). MP-free nanodiscs were directly applied to the Superdex 200 column after the removal of detergent absorbing beads. Samples were stored at 4°C until they were used for functional studies. Samples used for structural studies were prepared in our home lab in Copenhagen and transported to the SAXS beamline on ice and they were then gel filtrated directly into the X-ray beam as described later.

#### Activity assay

The binding of FVIIa (Uniprot ID: P08709) to TF in NDs was measured using the increase in FVIIa activity upon TF binding as described by others (11,25). FX (Uniprot ID: P00742) (650 nM) activation was allowed for 20 min by 10 pM TF in ND and a varying amount of FVIIa (provided by Novo Nordisk A/S). The reaction was stopped by adding EDTA and the amount of FXa generated was quantified using the chromogenic substrate S2765 (from Chromogenix), as the slope of the absorbance at 405 nm was evaluated for 20 min. A quadratic binding function was fitted to the slope as a function of FVIIa concentration to obtain the K<sub>d</sub>.

#### Surface plasmon resonance

The binding kinetics were determined for FVIIa binding to TF in NDs with different lipid compositions using surface plasmon resonance (SPR) measurements on a Biacore T200® system (GE Healthcare). To capture TF-containing nanodiscs we exploited that TF has a C-terminal His-tag. An anti-His antibody (GE Healthcare Life Sciences, cat BR100530) was immobilized on a CM5 sensor chip using N-hydroxysuccinimide and N-ethyl-N-(3-(diethylamino)propyl)-

carbodiimide chemistry. Excess reactive sites were inactivated by injecting 1 M ethanolamine. 100 nM TF-containing nanodiscs in 20 mM Tris, 100 mM NaCl, 20 mM EDTA pH 7.2 were injected, which resulted in a stable capture level of 100–200 RU. The reference flow cell was mock coupled and contained neither immobilized anti-His nor captured nanodiscs. Kinetic rate constants for the various TF:FVIIa interactions were determined with single-cycle and multi-cycle protocols by which serial 2-fold dilutions of the interaction partner (FVIIa) were injected at 50  $\mu$ l/min in 10 mM HEPES, 150 mM NaCl, 10 mM CaCl<sub>2</sub> and 0.05% (v/v) surfactant P-20 at pH 7.4 at 20°C. In the single-cycle protocol, 5 different FVIIa concentrations were injected for 90 s without intervening regeneration and followed by a longer dissociation phase of 2000 s. Two consecutive injections of EDTA buffer (20 mM EDTA in 10 mM HEPES, pH 7.4, 150 mM NaCl, and 0.05% (v/v) surfactant P-20) at the end of each single cycle regenerated the chip. For complete regeneration of the chip, including the TF capture level, two injections of 10 mM glycine-HCl, pH 1.5 for 60 s was used. For multi-cycle protocols, 2-fold dilutions of FVIIa were injected for 300 s followed by a dissociation phase of 5,000 s. Bound FVIIa on TF was removed at the end of each cycle by two consecutive injections of EDTA buffer as above. The association rate constant ( $k_{on}$ ) and dissociation rate constant ( $k_{off}$ ), the  $K_d$  ( $k_{off}/k_{on}$ ), as well as the binding capacity ( $R_{max}$ ) were derived by the fitting of the double blank referenced data by non-linear regression to a bimolecular interaction model, assuming pseudo-first order reaction conditions. We used the evaluation software supplied with the instrument for global fitting (BiacoreT200 Evaluation® 3.0).

#### **Small angle X-ray scattering**

Nanodiscs with and without TF were assembled as described and concentrated to a volume of 0.5 ml. The SAXS measurements were performed in Grenoble at ESRF at BM29 by using the SEC-SAXS setup (26). A Superdex 200 10/300 Increase column equilibrated in 20 mM Tris, pH 7.5, 100 mM NaCl, 10 mM CaCl<sub>2</sub>, 1 mM DTT, was attached to the setup, and the sample injected with a flowrate of 0.7 ml/min. Throughout the column volume, the intensity of the scattered X-rays was measured as a function of the magnitude of the scattering vector,  $q = 4\pi \sin(\theta) / \lambda$ .  $\theta$  is the scattering angle and  $\lambda$  the wavelength. The data were reduced using the software available at the beamline, and the average intensity of a given frame was plotted as a function of frame number(26). Relevant frames were selected from the peak and averaged and subtracted a buffer from the region before the peak. The data was converted to absolute scale using a water measurement to finally obtain the scattering intensity,  $I(q)$ , of the sample. The BayesApp-server was used to perform the indirect Fourier transformation of the  $I(q)$  to provide the pair distance distribution functions,  $p(r)$ .

## Results and discussion

### Expression and purification

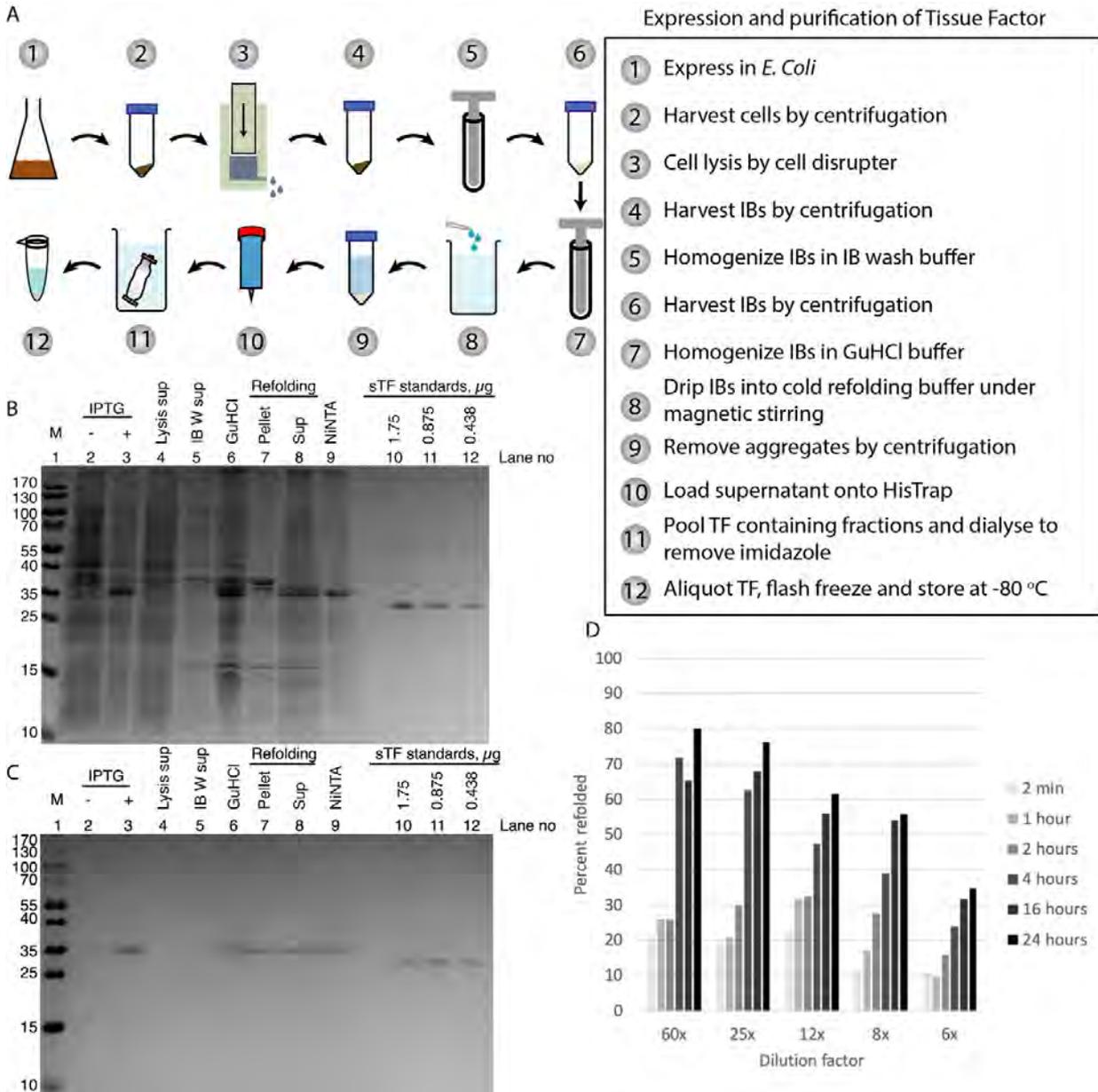


Figure 2: A: Schematic overview of the expression and purification of Tissue factor. B: SDS-PAGE gel, Coomassie stained. C: SDS-PAGE gel, Western blot with primary Tissue factor antibody. D: Refolding yields as a function of dilution factor and time. See discussion in the main text.

To produce full-length TF, we pursued a strategy involving a recombinant expression of TF ECD-TMD in IBs with subsequent refolding (see overview in figure 2A). As seen by coomassie staining (figure 2B) and Western blot performed using a primary TF antibody (figure 2C), the expression of TF was visible in the cellular lysate (figure 2B and C, lane 3) at the time of harvest. Furthermore, the background expression was very low compared to that of TF, which is essential for the subsequent purification steps. After cells lysis, minimal amounts of TF were present in the supernatant (fig. 2B and C, lane 4), indicating that almost the entire expression goes into inclusion bodies. During the IB washing step, a fraction of *E.*

*coli* proteins was present in the supernatant (figure 2B, lane 5), and almost no TF was lost as only a slight TF band is seen in the WB (figure 2C, lane 5). Almost no pellet was present after this step. The supernatant in GuHCl showed a substantial fraction of TF but also the presence of several host proteins (visualized in lane 6 on both gels, after ethanol precipitation). The next step was to refold the obtained protein. Refolding by dilution into a reservoir of the buffer without GuHCl is the most commonly used refolding technique (14), and was also applied here. Since refolding of a protein *in vitro* is usually described as a process on the millisecond to the second scale (27), the protein solution was dripped into refolding buffer under fast

stirring. This ensured rapid dilution with a low local concentration of TF to avoid dimer formation, as the formation of two disulfide-bridges needs to occur as well as folding of secondary structure elements. We investigated the effect of the refolding time and the dilution of the TF preparation. A distinct time dependence of refolding was observed (figure 2D) suggesting a much slower equilibrium than anticipated from the typically reported timescales of folding. As an example, the yields for the samples taken immediately after initiating the refolding process ranged from 10 to 20%, while leaving the refolding process for 24 hours yielded up to 80%. Furthermore, changing the dilution factor from 6x to 60x (1M to 0.1M GuHCl) increased the yield from 35% to 80%, which could both be an effect of lower GuHCl concentration and lower protein concentration. To ensure that the protein detected on the gels was in a soluble and folded state, we performed size exclusion chromatography (SEC) on each sample. The fraction of soluble TF was between 0.93 and 0.99 for all samples, indicating that the samples were monomeric, had a similar size and that the samples were relatively pure (see a representative example in figure S1D).

After establishing this protocol, we conducted a large-scale purification where we expressed and purified IBs from two liters of bacterial culture. We performed a 60x dilution for 24 hours and found that only a small amount of TF was present in the pellet, likely as a result of aggregation (figure 2A and B,

lane 7), while the majority of TF was in the supernatant (figure 2A and B, lane 8). After NiNTA purification (figure 2A and B, lane 9) and dialysis, the most concentrated fractions were selected, yielding 5 mg/liter of culture, which is much higher than the 0.4 mg/liter of culture reported in the previous purification protocol (28). This optimized yield provides ample TF both for reconstitution optimization and for structural studies.

#### **Nanodiscs assembly**

The reconstitution of TF in ND requires several steps and several different components. Figure 3A highlights the processes, where detergent absorbing beads are used to pull out the detergents from the mixed micelles containing lipids and cholate and the MP solubilized by Triton X-100, resulting in the formation of NDs with and without TF. Several parameters can be varied when assembling NDs, including the MSP:MP ratio, which is usually kept at 20:1 to decrease the probability that two MPs become incorporated in the same ND (22,29). However, using a high ratio such as this, results in a waste of lipids and MSP, since only a small fraction of the used material is carried through to the next step. In the present study, we used a circularized and solubility enhanced version of MSP1E3D1 (csE3) previously developed by our group (24). It yields a relatively large nanodisc with a diameter of approximately 13 nm, which is suitable in order to have an excess bilayer facilitating the later binding of FVIIa.

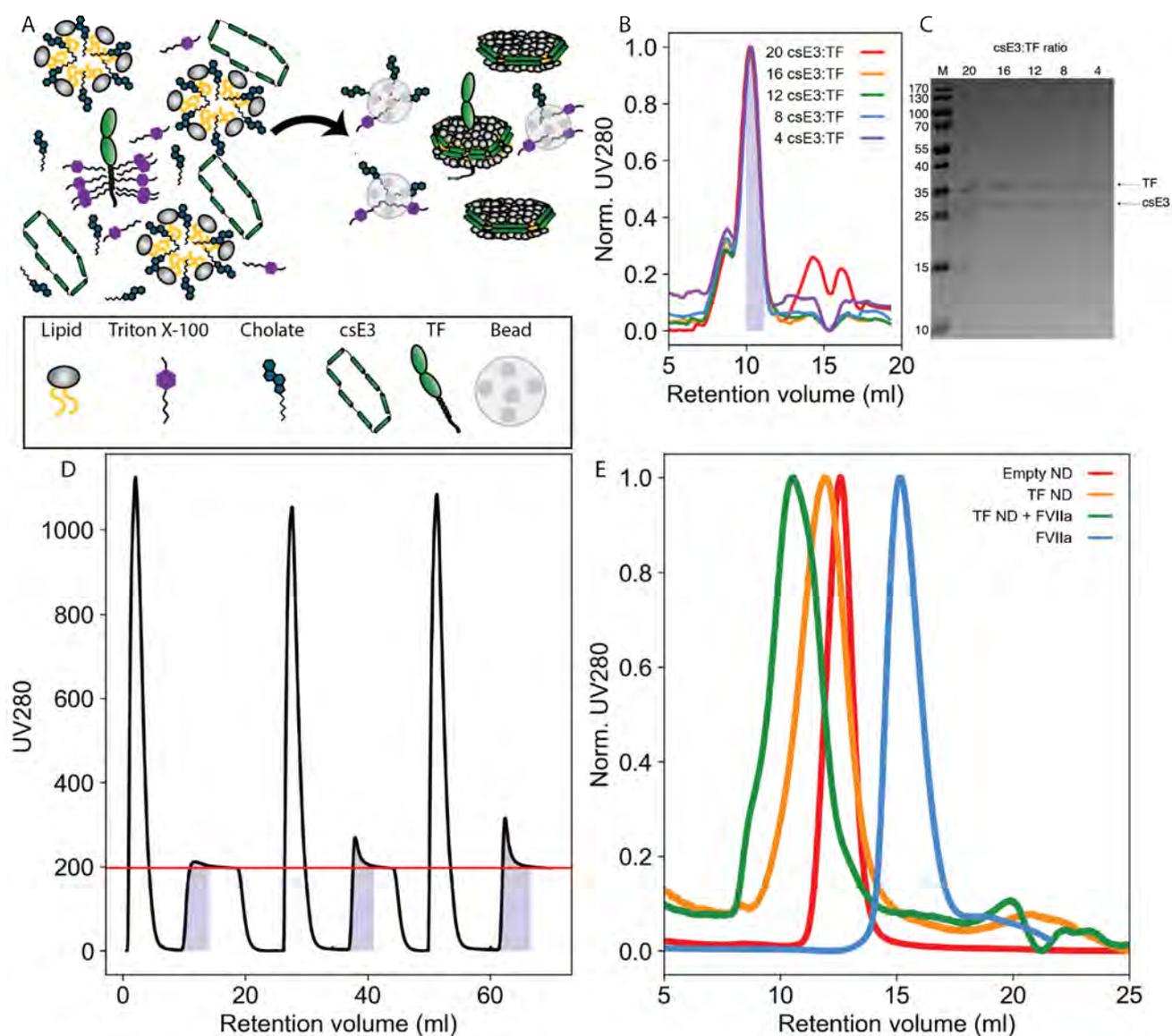


Figure 3: A: Schematics of the reconstitution process of a TF embedded nanodisc. B: SEC chromatogram of TF in NDs reconstituted with different csE3:TF ratios. Blue area indicate fraction selected for further work. C: Gel of the peak fraction from B. D: HisTrap chromatogram of nanodisc preparation with TF and increasing POPS content (0%, 15% and 30%). The blue area indicates the fractions selected for further work. The red line indicates the UV280 absorption from the imidazole buffer. The three large peaks corresponds to empty NDs. E: SEC chromatograms: Red: Empty nanodisc, orange: nanodisc with TF incorporated, blue: FVIIa, and green: nanodisc with TF incorporated and a ratio of 1:1 FVIIa added.

We tested different ratios of csE3:TF, and all tested ratios provided a single main peak at precisely the same elution volume in the SEC (figure 3B). The fractions spanning the main peaks (marked by blue) were loaded onto a gel to assess the csE3:TF ratio. By gel quantification, these were found to be virtually identical (figure 3C) and correspond to one TF per disc. However, it should be noted that the reconstitution ratio of 4, for unknown reasons, consistently yielded slightly less final ND than the other ratios, and therefore a ratio of 8 was identified as the most optimal, as the yield was high and excess lipids and MSP low.

As it is well established that the activity of the TF:FVIIa complex increases with the POPS contents, nanodiscs with increasing amounts of POPS were assembled. Using HisTrap purification of assembled NDs with TF and an increasing amount of POPS, we observed a change in the quantity of reconstituted protein. In the chromatograms (figure 3D) the three largest peaks were empty NDs not captured by the HisTrap column and the blue areas correspond to the samples selected for further work. The red line indicates the UV280 background of the imidazole buffer, and hence the UV signal in the blue areas above the red line roughly corresponds to NDs with TF incorporated.

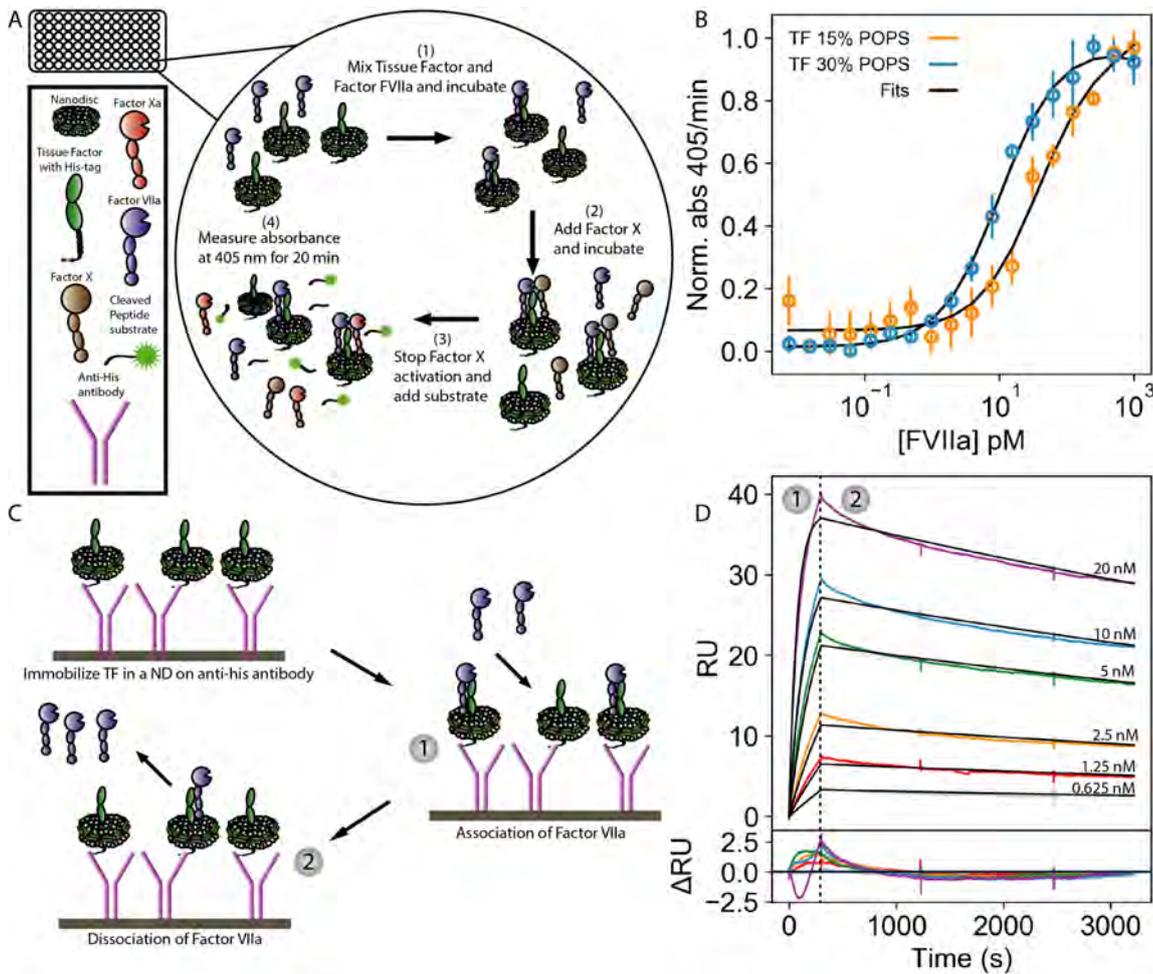


Figure 4: A: Schematic overview of the experimental setup of the activity assay. B: Activity assay: The slope of the absorbance at 405 nm is plotted as a function of the FVIIa concentration added. The data is fitted by a quadratic binding curve to obtain the  $K_d$ . C: Experimental setup of the SPR experiments. D: SPR data obtained on FVIIa binding to TF in NDs with 30 % POPS. The applied FVIIa concentrations indicated in the plot.

Table 1	Sample	$K_d$ (pM)	$k_{on}$ ( $10^5 M^{-1} s^{-1}$ )	$k_{off}$ ( $10^{-5} s^{-1}$ )
Activity assay	0 % POPS	ND	ND	ND
	15 % POPS	48±8	ND	ND
	30 % POPS	10±1	ND	ND
SPR, multi cycle	0 % POPS	126	5.60 ± 0.01	7.08 ± 0.01
	15 % POPS	127	5.58 ± 0.01	7.08 ± 0.01
	30 % POPS	125	6.77 ± 0.01	8.47 ± 0.01
SPR, single cycle	0 % POPS	238	5.70 ± 0.01	13.5 ± 0.01
	15 % POPS	118	7.34 ± 0.01	8.66 ± 0.02
	30 % POPS	168	8.31 ± 0.01	13.9 ± 0.01

Table 1: Overview of data from the activity assay and from the SPR experiments. ND = Not determined.

As seen in figure 4D, the incorporation was much more efficient when POPS was present and, as evaluated by area under the curve (light blue area lying above the red line indicating the background), the preparation with 0% POPS only yielded 17% of the UV signal obtained in the sample with 30% POPS, while the sample with 15% POPS yielded 69%. Thus, the presence of POPS clearly enhanced TF reconstitution into NDs. The lower

yield of the 0% POPS sample naturally resulted in lower yield after the SEC purification, but besides the weaker UV signal, the SEC chromatograms were virtually identical (see figure S2). To the best of the authors' knowledge, this apparent POPS preferential behavior in the reconstitution has not been described elsewhere. It has however been suggested that TF has specific residues that interact with POPS in a bilayer, and the

formation of these interactions could potentially stabilize the system and aid the reconstitution (30).

### Interaction with FVIIa

To validate that TF is indeed present in the ND and in an active state, a SEC of an empty ND (figure 3E, red), a ND with TF (30% POPS) incorporated (figure 3E, orange) alone and with FVIIa added in a 1:1 ratio (figure 3E, green) was analyzed on a Superdex 200 column. As evident from the elution profiles, the particle was larger when TF was incorporated and even further so when FVIIa was bound to TF in the ND. This indirectly confirms the presence and capability of recombinantly produced and refolded TF to bind FVIIa.

To further evaluate the activity of TF produced using this protocol, we performed an activity assay with TF in ND (see schematics in figure 4A). The assay relies on the fact that FX activation is a factor of  $10^5$  higher by the TF:FVIIa complex than by free FVIIa (31). The activity of TF in 30% POPS was slightly higher than in 15% POPS (figure 4B), which coincides with previous studies using similar experiments (11). The  $K_d$  of FVIIa binding to TF in ND was determined to  $48 \pm 8$  pM (15% POPS) and  $10 \pm 1$  pM (30% POPS) (see table 1), which is within range of previously reported results (11). It is not feasible to measure the affinity of FVIIa towards TF in an ND containing 0% POPS using this assay as the activity is orders of magnitude lower, and the background signal from activation by free FVIIa contributes significantly to the overall signal (31).

To probe the affinity between FVIIa and the TF ND with 0% POPS and to gain information about how the POPS contents affected the binding kinetics of the interactions, we performed surface plasmon resonance (SPR) experiments on the same three TF samples (see the experimental setup in figure 4C). The kinetic rate constants were largely identical in both experimental setups (single and multi cycle) and did not change with the lipid compositions of the ND. The recorded data and results were in accordance with previous studies (7). It should nonetheless be emphasized that the fitting of the data is not perfect with some curvatures in the residuals (see fig 4D), however, this is not anticipated to affect the overall conclusions. The reason for the higher affinity determined by the functional assay is not known, but could potentially arise from the assisted complex assembly by the FX present in the assay

### Samples suitable for structural studies

The monodisperse nature of nanodiscs and in particular the circularized version makes this an ideal system for small X-ray angle scattering (SAXS) experiments. SAXS provides structural information down to  $10 \text{ \AA}$  resolution on particles in solution and is therefore ideal to structurally investigate the lipid-protein interplay in a membrane-like context.

We performed SEC-SAXS of TF-loaded ND and after background subtraction, calibration to absolute scale and concentration normalization, we obtained the data plotted in figure 5A.

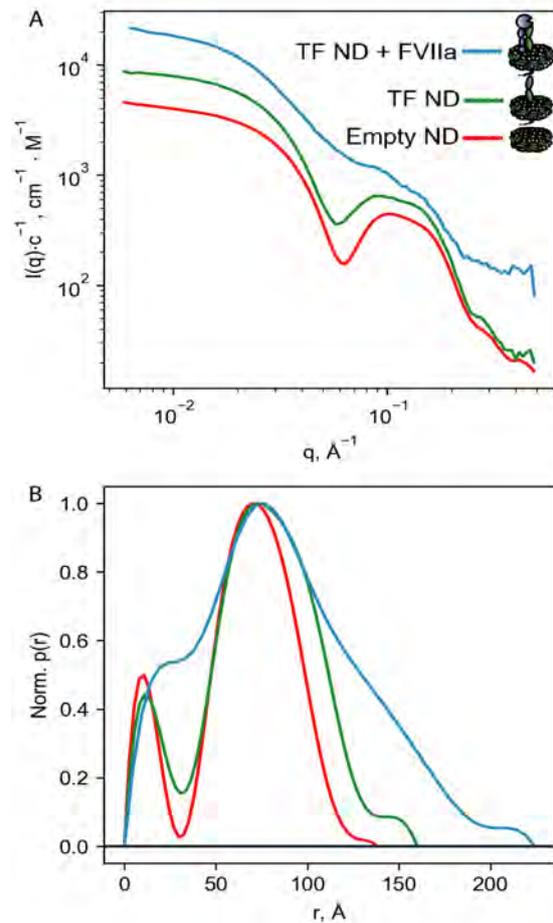


Figure 5: A: SAXS data of the empty ND (red), TF in ND (green) and TF in ND and FVIIa (blue). All data on absolute scale and normalized by the concentration in molar. B: Normalized pair distribution function of the same samples.

The oscillation in the  $0.1 \text{ \AA}^{-1}$  regime is a characteristic feature of SAXS data from an ND (21,24,32), and both the empty ND and the ND containing TF has this feature. The feature arises from the oscillating sign of the contrasts (excess scattering length density) from the lipid bilayer within the ND. Due to the small size of TF compared to the ND, the oscillation is also seen in this sample. However, when FVIIa is added, the curve changed considerably, due to the large size of FVIIa now located on the 'top' of the ND. As visualized by the pair distance distribution function in figure 5B, the largest distance ( $D_{\max}$ ) in the measured particle matched well with the  $\sim 13 \text{ nm}$  for the empty ND and then increased as first TF is embedded in the ND and again when FVIIa is bound to TF. To further investigate the nature of the complex on a lipid bilayer, more advanced modeling is needed, which was not within the scope of the present study. However, high-quality samples made possible by the optimized protocol for production of full-length TF lay the foundation for access to structure information of this multicomponent complex. The protocols developed in this study can produce samples in the amounts and quality needed for both functional and structural studies and may be readily adaptable for the production of other single - or multiple pass transmembrane proteins. This is a prerequisite for future experiments and gives unique structural data necessary for modeling of the TF complex. Combined this will provide a

deeper understanding of the structural effects of lipid-protein interaction in the blood coagulation.

### Conclusion

We have shown that a final yield of more than 5 mg of pure and concentrated TF per liter of culture can be obtained from a high-level expression in IBs and a subsequent efficient refolding *in vitro*. A significantly prolonged equilibration time turned out to be crucial to optimize the yield. After reconstitution into circularized nanodiscs with different lipid contents, we probed the activity and the binding kinetics to FVIIa and found them to be similar to TF expressed in a membrane-anchored state (11,12) hence documenting that the higher yielding protocol provides samples on the functionally relevant form. Lastly, we investigated the samples by small angle X-ray scattering and show that both the quality and amount of the samples are sufficient to obtain structural information about the interplay between TF and the surrounding lipids as well as with binding partners in multicomponent complexes. The approach may be readily adaptable to the production of other single-pass membrane proteins fueling an increase in the number of membrane protein structures.

### ASSOCIATED CONTENT

#### Supporting Information

S1: A-C:SDS-PAGE gels used for estimating refolding yield. D: Representative SEC chromatogram of refolded TF. S2: SEC chromatograms of TF containing NDs with varying amounts of POPS. S3: SPR data of the single cycle experiments of FVIIa binding to TF NDs.

#### Protein accession codes

Tissue factor: P13726. Factor VII: P08709 Factor X: P00742.

### AUTHOR INFORMATION

#### Corresponding Author

Lise Arleth, Arleth@nbi.ku.dk.

#### Conflict of interest

The authors declare no conflict of interest.

#### Author Contributions

All authors contributed to the experiment design, FGT conducted all experiments and data treatment with input from HØ, BBK and LA, except SPR measurements which were conducted by MP. FGT wrote the manuscript with contributions and input from all coauthors.

### ACKNOWLEDGMENT

The authors thank ESRF, Grenoble, France for the beamtime allocated at the BM29 BioSAXS beamline and Mark Tully for assisting during the experiment. We thank DanScatt for the financial support to go to the beamtime in Grenoble. This work was supported by the Novo Nordisk foundation, Synergy program (LA, BBK).

### REFERENCES

1. Bugge K, Lindorff-Larsen K, Kragelund BB. Understanding single-pass transmembrane receptor signaling from a structural viewpoint—what are we missing? *FEBS J.* 2016;283(24):4424–51.
2. Fagerberg L, Jonasson K, Von Heijne G, Uhlén M, Berglund L. Prediction of the human membrane proteome. *Proteomics.* 2010;10(6):1141–9.
3. Bugge K, Papaleo E, Haxholm GW, Hopper JTS, Robinson C V., Olsen JG, et al. A combined computational and structural model of the full-length human prolactin receptor. *Nat Commun.* 2016;7(May):11578.
4. Banner DW, D'Arcy A, Chene C, Winkler FK, Guha A, Konigsberg WH, et al. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature.* 1996. p. 41–6.
5. Sen M, Herzik M, Craft Jr. JW, Creath AL, Agrawal S, Ruf W, et al. Spectroscopic Characterization of Successive Phosphorylation of the Tissue Factor Cytoplasmic Region. *Open Spectrosc J.* 2009;3(713):58–64.
6. Morrissey JH, Macik BG, Neuenschwander PF, Comp PC. Quantitation of activated factor VII levels in plasma using a tissue factor mutant selectively deficient in promoting factor VII activation. *Blood.* 1993;81(3):734–44.
7. MacKman N. The role of tissue factor and factor VIIa in hemostasis. *Anesth Analg.* 2009;108(5):1447–52.
8. Davie E, Fujikawa K, Kisiel W. The Coagulation Cascade: Initiation, Maintenance, and Regulation. *Perspectives Biochem.* 1991;30(43).
9. Van Meer G, Voelker DR, Feigenson GW. Membrane lipids: where they are. *Nat Rev Mol Cell Biol.* 2009;10(1):1–4.
10. Kodigepalli KM, Bowers K, Sharp A, Nanjundan M. Roles and regulation of phospholipid scramblases. *FEBS Lett.* 2015;589(1):3–14.
11. Shaw AW, Pureza VS, Sligar SG, Morrissey JH. The local phospholipid environment modulates the activation of blood clotting. *J Biol Chem.* 2007;282(9):6556–63.
12. Sen P, Neuenschwander PF, Pendurthi UR, Rao LVM. Analysis of factor VIIa binding to relipidated tissue factor by surface plasmon resonance. *Blood Coagul Fibrinolysis.* 2010;21(4):376–9.
13. Muller Y a., Ultsch MH, de Vos, Abraham M. The crystal structure of the extracellular domain of human tissue factor refined to 1.7 Å resolution. *J Mol Biol.* 1996;256(1):144–59.
14. McCallum CD, Hapak RC, Neuenschwander PF, Morrissey JH, Johnson AE. The location of the active site of blood coagulation factor VIIa above the membrane surface and its reorientation upon association with tissue factor: A fluorescence energy transfer study. *J Biol Chem.* 1996;271(45):28168–75.
15. McCallum CD, Su B, Neuenschwander PF, Morrissey JH, Johnson AE. Tissue factor positions and maintains the factor VIIa active site far above

- the membrane surface even in the absence of the factor VIIa Gla domain. A fluorescence resonance energy transfer study. *J Biol Chem.* 1997;272(48):30160–6.
16. Zur M, Nemerson Y. The Esterase Activity of Coagulation Factor VII. *J Biol Chem.* 1977;253(7):2203–9.
  17. Maignan S, Guilloteau JP, Pouzieux S, Choi-Sledeski YM, Becker MR, Klein SI, et al. Crystal structures of human factor Xa complexed with potent inhibitors. *J Med Chem.* 2000;43(17):3226–32.
  18. Tavoosi N, Smith SA, Davis-Harrison RL, Morrissey JH. Factor VII and protein C are phosphatidic acid-binding proteins. *Biochemistry.* 2013;52(33):5545–52.
  19. Popot JL. Folding membrane proteins in vitro: A table and some comments. *Arch Biochem Biophys.* 2014;564:314–26.
  20. Arana MR, Fiori MC, Altenberg GA. Functional and structural comparison of the ABC exporter MsbA studied in detergent and reconstituted in nanodiscs. *Biochem Biophys Res Commun [Internet].* 2019;512:7–11. Available from: <https://doi.org/10.1016/j.bbrc.2019.03.069>
  21. Denisov IG, Grinkova Y V., Lazarides AA, Sligar SG. Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *J Am Chem Soc.* 2004;126(11):3477–87.
  22. Bayburt TH, Grinkova Y V., Sligar SG. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* 2002;2(8):853–6.
  23. Nasr ML, Baptista D, Strauss M, Sun ZYJ, Grigoriu S, Huser S, et al. Covalently circularized nanodiscs for studying membrane proteins and viral entry SI. *Nat Methods.* 2016;14(1):49–52.
  24. Johansen NT, Pedersen MC, Tidemand FG, Rand KD, Nguyen TTTN, Arleth L. Circularized and solubility-enhanced MSPs facilitate simple and high yield production of stable nanodiscs for studies of membrane proteins in solution. *FEBS J.* 2019;
  25. Waters EK, Morrissey JH. Restoring full biological activity to the isolated ectodomain of an integral membrane protein. *Biochemistry.* 2006 Mar;45(11):3769–74.
  26. Pernot P, Round A, Barrett R, De Maria Antolinos A, Gobbo A, Gordon E, et al. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J Synchrotron Radiat.* 2013;20(4):660–4.
  27. Naganathan AN, Muñoz V. Scaling of folding times with protein size. *J Am Chem Soc.* 2005;127(2):480–1.
  28. Paborsky LR, Tate KM, Harris RJ, Yansura DG, Band L, McCray G, et al. Purification of recombinant human tissue factor. *Biochemistry.* 1989;28(20):8072–7.
  29. Skar-Gislinge N, Johansen NT, Høiberg-Nielsen R, Arleth L. A comprehensive study of the self-assembly of phospholipid nanodiscs: What determines their shape and stoichiometry? *Langmuir.* 2018;34:acs.langmuir.8b01503.
  30. Ke K, Yuan J, Morrissey JH. Tissue factor residues that putatively interact with membrane phospholipids. *PLoS One.* 2014;9(2).
  31. Neuenschwander PF, Morrissey JH. Roles of the membrane-interactive regions of factor VIIa and tissue factor: The factor VIIa Gla domain is dispensable for binding to tissue factor but important for activation of factor X. *J Biol Chem.* 1994;269(11):8007–13.
  32. 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;132(39):13713–22.

## Supplementary information

### Figure S1

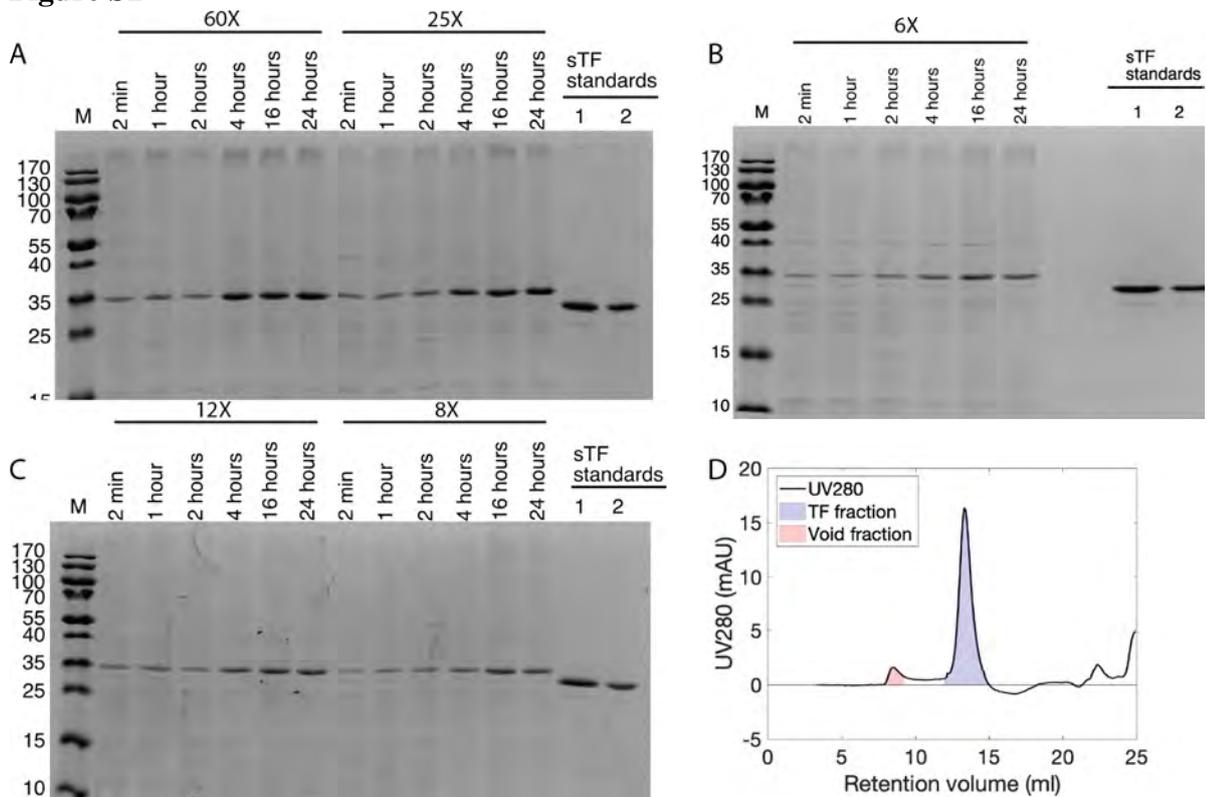


Figure S1: Gels used for quantification of refolding yields. On all gels two standards of sTF were loaded for normalization. Band intensities were quantified using the Image Lab software (BioRad) using default settings. Samples were taken out after 6 different time points (2 min, 1, 2, 4, 16 and 24 hours) A: Samples diluted in 60x buffer and 25x buffer B: Samples diluted in 6x buffer. C: Samples diluted in 12x buffer and 8x buffer. D: Gel filtration chromatogram of refolded and NiNTA purified TF on a Superdex200 equilibrated in 20 mM Tris, pH 8, 100 mM NaCl, 0.1% Triton X-100. The areas under the two peaks (red and blue areas) were calculated and the fraction correctly folded is calculated to be 93%.

**Figure S2**

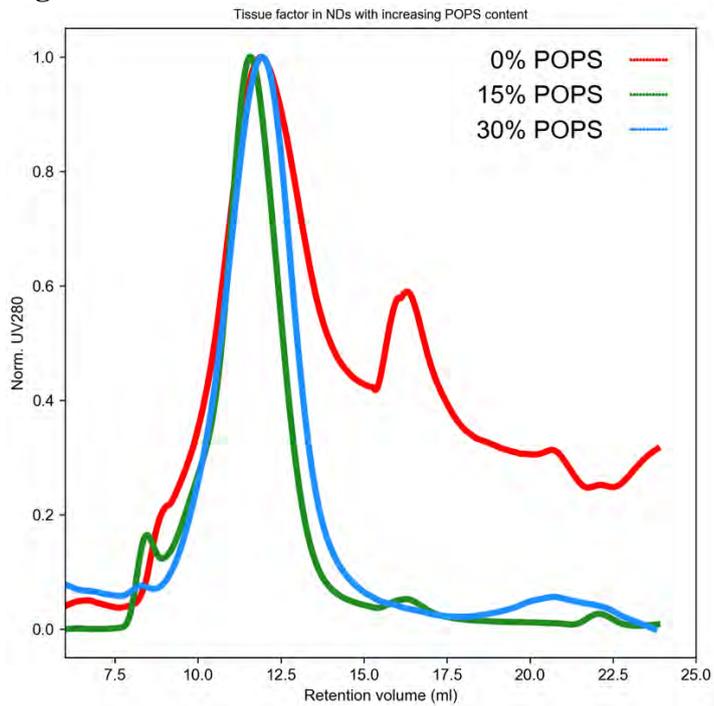


Figure S2: Gel filtration chromatogram (Superdex 200) of TF NDs assembled with different amount of POPS. The low UV signal for the 0% POPS sample provided an unstable baseline after the peak.

**Figure S3**

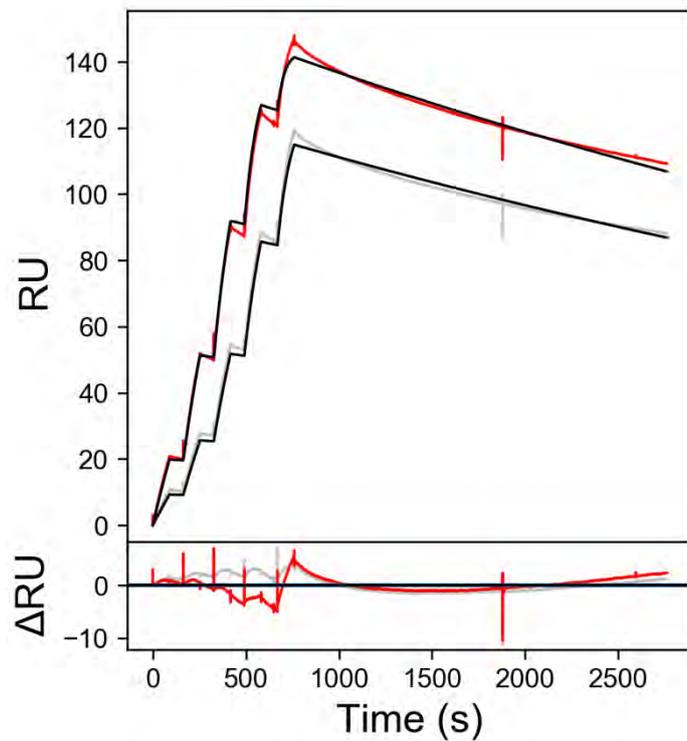
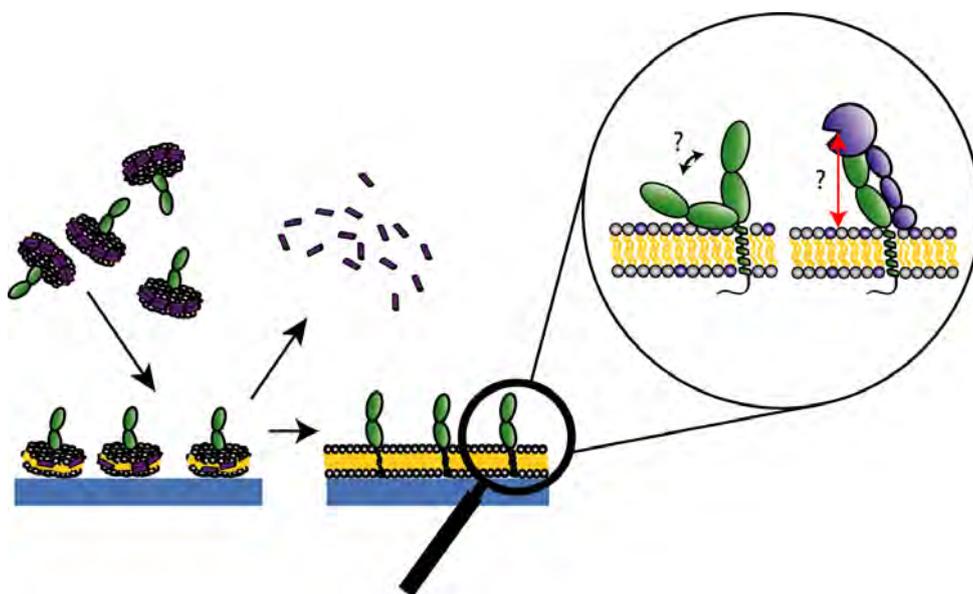


Figure S3: SPR data: Single cycle analysis of the interaction between TF in a nanodisc comprising 30% POPS lipids and FVIIa. FVIIa concentrations: Grey: 1 nM, 2 nM, 4 nM, 8 nM, 16 nM. Red: 2 nM, 4 nM, 8 nM, 8 nM, 16 nM, 32 nM.

## 9 | Paper V

### Probing the orientation of the tissue factor:Factor VIIa complex in a lipid bilayer

*Frederik Grønæk Tidemand, Alessandra Luchini, Raul Araya-Secchi, Katrine Bugge, Michael Ploug, Birthe B. Kragelund, Lise Arleth1*



# Probing the orientation of the tissue factor:Factor VIIa complex in a lipid bilayer

Frederik Grøn­bæk Tidemand<sup>1</sup>, Alessandra Luchini<sup>1</sup>, Raul Araya-Secchi<sup>1</sup>, Katrine Bugge<sup>2</sup>, Michael Ploug<sup>3,4</sup>, Birthe B. Kragelund<sup>2</sup>, Lise Arleth<sup>1</sup>

1: Structural Biophysics, X-ray and Neutron Science, The Niels Bohr Institute, University of Copenhagen, Denmark.

2: Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Denmark

3: Finsen Laboratory, Rigshospitalet, DK-2200 Copenhagen N, Denmark.

4: Biotech Research and Innovation Centre (BRIC), University of Copenhagen, DK-2200 Copenhagen N, Denmark.

## Abstract

The blood coagulation system has been investigated for decades, and high resolution structures of many of the implicated proteins are available. However, structural information on lipid-protein interplay is scarce, and this is especially true for the tissue factor (TF):Factor VIIa (FVIIa):Factor Xa (FXa) complex which initiates blood coagulation. This is probably due to the lack of experimental possibilities in making samples containing both lipids and protein. In the present study, we used a recently developed methodology to deposit a lipid bilayer containing TF on a surface. We used this bilayer to obtain the heights of TF, TF:FVIIa and FXa on a membrane, by neutron reflectometry, and based on these measurements we speculated on the orientations of the proteins on the membrane. Using this information, the height of the FVIIa active site and the FX scissile bond, with respect to the membrane, were calculated to be 77 Å and 78 Å, respectively. These equivalent heights suggest that one function of the membrane is to ensure the correct orientation of the two proteins, which could be responsible for the increased activity of TF:FVIIa upon the presence of a lipid bilayer.

## Introduction

The extrinsic part of the human blood coagulation cascade is initiated by the rupture of a blood vessel and subsequent formation of the high affinity complex between the membrane protein: tissue factor (TF), and the soluble protein: Factor VII (FVII). TF consists of a folded extracellular domain (ECD), followed by a single pass transmembrane domain (TMD) and a short, likely unfolded, intracellular domain (ICD) (1), while FVII consists of an N-terminal Gla-domain, followed by two EGF-like domains and a peptidase domain (see illustration of both in figure 1). After formation of the TF:FVII complex, FVII is activated to FVIIa, and the TF:FVIIa complex can then activate Factor X (FX) to FXa. Once activated, FXa participates in the Prothrombinase complex, which activates Prothrombin, ultimately resulting in blood clotting. An essential co-factor in all of the above mentioned processes, is the lipid bilayer, on which most of the reactions in the blood coagulation pathway occur (2). High resolution structures of most soluble parts of the TF:FVIIa:FXa complex was solved decades ago, but structural information about the tertiary complex and the interplay with the lipids is scarce (3–5). It is well established that the presence of the membrane increases the activity of the TF:FVIIa complex by more than a 1000 fold (6) and the presence of the negatively charged Phosphatidylserine (PS) lipids even further increases the activity by a 100 fold (7). However, the structural feature responsible for the increased activity is still not completely clear. The Gla-domain of FVIIa is known to interact with the membrane with a  $K_D$  in the low  $\mu\text{M}$  range (8), and it has been hypothesized that this interaction could induce a change in FVIIa that increases the catalytic activity. However, later studies have shown that the activity of FVIIa and a mutant of FVIIa lacking the Gla-domain (des-gla, DG-FVIIa) had the same activity towards a peptide substrate, indicating that the catalytic domain of FVIIa is not affected by the lipid interaction (6). On the other hand, the activity towards FX was a 100 times lower for DG-FVIIa, demonstrating that the Gla-domain is critical for the activity (6). To probe this effect, the height of the FVIIa active site situated above the membrane, was measured in a FRET study, where it was found that the height decreased from  $83 \pm 3.3$  Å for the free FVIIa (figure 1A) to  $75 \pm 1.8$  Å for the TF bound FVIIa (figure 1B),

indicating that the role of TF and the membrane is to locate FVIIa in the correct height for catalytic efficiency (9,10). In addition, it was found that the height of TF bound DG-FVIIa (figure 1C) was  $73 \pm 1.8 \text{ \AA}$  and thus not significantly different from the active site of TF bound FVIIa, suggesting that it might just be TF introducing the change in orientation and not the Gla-domain of FVIIa. It should be mentioned that all these experiments were conducted using a PS containing bilayer, and thus the effect of PS presence, was not addressed by these studies.

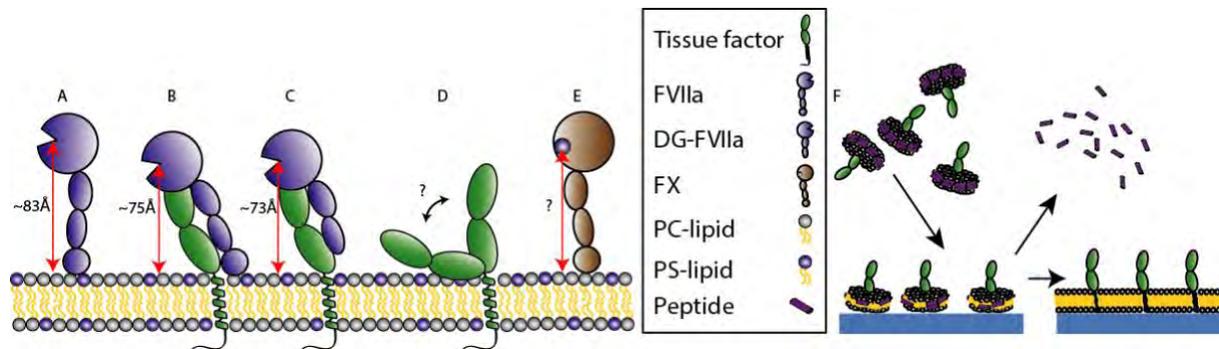


Figure 1: A-D: Schematic overview of TF and FVIIa on a membrane. All numbers presented are found in (9,10) and represent the distance between the active site of FVIIa and the membrane as measured by FRET. A: Free FVIIa binding to the membrane. B: The FVIIa:TF complex on a membrane. C: The DG-FVIIa:TF complex on a membrane. D: Orientation of TF in a membrane. E: FX on a membrane. Indication of the FX scissile bond (blue circle). F: Schematic overview of the deposition of TF in a peptidedisc and the subsequent flushing with buffer to remove peptides, leaving a membrane with TF incorporated in a unidirectional way.

To investigate the effects of the TF-TMD and the linker region between the TMD and the ECD of TF, the ECD was attached to a modified lipid head group. This setup was shown to maintain the native activity towards FX, demonstrating that the activity is not dependent on the specific sequence of the linker nor the TMD, but solely the ECD of TF being close to the membrane surface (11). This has led to the idea that the orientation of the ECD of TF is potentially guided by interactions with the lipid membrane, and especially interactions between the ECD and PS lipids have been investigated, as these could potentially also explain the increase in activity. But as it is for many systems involving both lipids and proteins, the lack of an experimental setup to investigate the structures is a bottleneck. Therefore, mutational studies have been used to identify a number of residues located in the C-terminal of the TF ECD, proximal to the membrane, which was found to be necessary for the activation of FX (12). However, as this activity is dependent on TF substrate recognition (interaction with FX) and potentially also TF interacting with lipids, it was not possible to distinguish these two effects. Subsequently, a molecular dynamics (MD) simulation with the crystal structure of the TF ECD (called soluble TF, sTF), placed in the proximity of a membrane containing PS-lipids was used to investigate this further. The simulation showed an interaction between the lipids and sTF, and some residues in sTF were in contact with PS more than 70% of the time, indicating a relatively strong interaction. A fraction of these residues were the same residues proved important for FX activation in the mutational study, again suggesting an essential role of the membrane and PS lipids specifically (13,14). However, these findings are still to be confirmed by experimental data, and in addition, a simulation of TF situated in the membrane by the TMD would provide a more native-like setup for the simulations.

A piece of additional information lacking to describe the mechanism behind the function of the membrane is the position of the scissile bond on FX when bound to the membrane (figure 1E), as this should be within reach of the FVIIa active site. FX is known to have an affinity towards the membrane in the nM-range (7), but the orientation of FX when bound to the membrane is not yet determined.

The lack of experimental data in a membrane environment of this system, points towards a more general problem of obtaining structural information about proteins in a lipid context. Neutron reflectometry (NR) has emerged as a technique able to probe the thickness of layers on a surface, and has recently also been shown to provide valuable information about heights of membrane proteins in bilayers. Additionally,

information on the interaction between soluble proteins and a membrane with regards to penetration depth and protein height over the surface can be obtained (15–18). The main problem for this technique, has been sample preparation, as depositing a bilayer with membrane proteins all oriented in the same way is difficult. However, a recently developed method utilizing the so-called peptidedisc as carrier for a patch of membrane containing the membrane protein of interest, presents a solution to this problem (Paper III). After injecting the peptidediscs into the sample cell, the protruding ECDs of TF will all be oriented away from the surface, and a subsequent flushing with buffer removes the amphipathic peptides, leaving a membrane with TF incorporated in a unidirectional manner (figure 1F).

In the present study, we show that a combination of NR data on TF in lipid membranes, and MD simulations on the same system, can provide information about the orientation of TF in a membrane. In addition, we show how the binding of FVIIa affects this system, investigate the orientation of FXa on a lipid bilayer and speculate about the possible correlations between the heights of these.

## Materials and methods

### Sample preparation for NR

All chemicals are from Sigma Aldrich. The 18A peptide was synthesized by Genscript. Tissue factor (1-224) was expressed and purified as described previously (Paper IV). Samples were assembled with 12  $\mu\text{M}$  TF, 1 mg/ml 18A and 1 mg/ml lipid (30 % 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 70% 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS)). After incubation for 1 hour at 10 °C the preparation was loaded onto a 1 ml HisTrap column (GE Healthcare) equilibrated in loading buffer (20 mM Tris-HCl, pH 8, 100 mM NaCl, 20 mM imidazole). After loading, the column was washed with 3 column volumes (CV) of loading buffer and eluted with 20 mM Tris-HCl, pH 8, 100 mM NaCl, 250 mM imidazole while fractions were collected. Fractions spanning the main peak were collected (typically 2 ml) and frozen and kept at -20 °C until use. Factor VIIa and FXa were provided by Novo Nordisk A/s and were diluted to 10  $\mu\text{M}$  in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM  $\text{CaCl}_2$  before experiments.

### Neutron reflectometry

The NR experiments were performed at ISIS, Chilton, UK, at the SURF beamline. SURF is a time of flight reflectometer and three angles were used to cover the desired  $q$ -range: 0.35°, 0.65° and 1.5°. All samples were measured under three different buffer conditions (100%  $\text{D}_2\text{O}$ , 0%  $\text{D}_2\text{O}$  and in silicon matched water (SMW) (38%  $\text{D}_2\text{O}$ ). The samples for the experiments and the data reduction was performed as described previously (Paper III). After measuring the TF containing bilayer, 10  $\mu\text{M}$  FVIIa was injected and the cell was washed with buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM  $\text{CaCl}_2$ ) before measuring the TF:FVIIa complex. The data analysis for the TF sample was conducted as previously described (Paper III). For the TF:FVIIa analysis, the same model was used, except that FVIIa was now included in the TF ECD layer.

For the FXa sample, a pure POPC/POPS (30% POPS) bilayer was prepared using the methodology described previously (Paper III). 10  $\mu\text{M}$  FXa was injected and after a buffer wash (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM  $\text{CaCl}_2$ ) the measurements were conducted.

### Small angle X-ray scattering

FVIIa (from Novo Nordisk A/S) was measured using the SEC-SAXS setup at the BM29 beamline, Grenoble, France (19). The buffer was 20 mM Tris-HCl, pH 7.5, 100 mM NaCl and either 10 mM  $\text{CaCl}_2$  (with calcium) or 1 mM EDTA (without calcium). A Superdex 200 10/300 column equilibrated in relevant buffer was attached to the setup, and the sample was injected with a flowrate of 0.5 ml/min. The intensity of the scattered X-rays was measured for 48 min. The scattering vector,  $q = 4\pi \sin(\theta) / \lambda$  where  $\theta$  is the scattering angle and  $\lambda$  the wavelength. The data were reduced using the software available at the beamline, and the average intensity of a given frame was plotted as a function of frame number (19). Relevant frames were selected from the peak and averaged and subtracted a buffer from the region before the peak. The data was converted to absolute scale using a water measurement to finally obtain the scattering intensity,  $I(q)$ , of

the sample. The BayesApp-server was used to perform the indirect Fourier transformation of the  $I(q)$  to provide the pair distance distribution ( $p(r)$ ) functions.

### Nuclear magnetic resonance

$2D,_{15}N$ -labeled sTF (provided by Novo Nordisk A/S) was used to collect  $^1H,_{15}N$  -spectra. The protein concentration was 50  $\mu$ M, buffer conditions: 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM CaCl<sub>2</sub> and the experiments were performed at 25 °C on a Varian INOVA 800-MHz ( $^1H$ ) spectrometer equipped with a room temperature probe. The obtained spectra could be assigned by assignments provided by Novo Nordisk A/S.

To assess the lipid binding of sTF, nanodiscs (NDs) made using the novel MSP called csdH5 was utilized. csdH5 is designed in the same way as csE3 (described in detail in (20)) but with MSP1 $\Delta$ H5D1 as template instead of MSP1E3D1. The csdH5 sequence is:

```
MGHHHHHHHD YDIPTTENLY FQGSSSSFSK LREELGPVSE EFWDDLEKES EGLREEMSKD
LEEVKAKVEP YLDDFEKKWE EEMELYREKV EPLGEEMRDR ARAHV DALRS H LAPYSDEL R
ERLAARLEAL KEDGGARLAE YHAKASEHLS SLSEKAKPAL EDLREGLLPV LESFKVSFLS
ALEEYSKKLD SEGGRRGSLP ETGGASHHHH HH
```

The expression was conducted at 20 °C over night, but otherwise was the protocol identical to the one for csE3 which is described in detail in (20). NDs were assembled with a lipid:MSP ratio of 45 with either 100% POPC or 80% POPC and 20 % POPS. The reconstitution was performed as described in (20) but the NDs were concentrated using a spinfilter (cutoff 30 kDa) after the SEC run to obtain a sufficiently high concentration.

Titration experiments were performed with an ND:sTF ratio of 5 as an initial experiment. Titration points with ratios of 5, 4 and 2.5 was measured for both the pure POPC NDs and the mixed POPC/POPS NDs. Intensity changes (based on the volumes of the peaks) were determined at the different ratios and compared to free sTF intensities.

### Surface plasmon resonance

Binding kinetics of the interactions between TF and FVIIa and DG-FVIIa (both provided by Novo Nordisk A/S) were measured using SPR. TF incorporated in NDs with different lipid compositions were prepared as described previously (*Paper IV*). Measurements were performed using a Biacore T200® system (GE Healthcare) using an anti-His antibody (GE Healthcare Life Sciences, cat BR100530) which was immobilized on a CM5 sensor chip using N-hydroxysuccinimide and N-ethyl-N-(3-(diethylamino)propyl)-carbodiimide chemistry. 100 nM TF NDs were injected in a buffer containing 20 mM Tris-HCl, pH 7.2, 100 mM NaCl, 20 mM EDTA pH 7.2. Similarly 100 nM of empty discs containing a His-tag on the MSP were injected in the reference cell (NDs called lsE3, prepared as described in (20)). The binding kinetics were determined for three different lipid compositions (0%, 15%, and 30% POPS, with POPC as a counterbalance). Single-cycle experiments were performed with 2-fold dilution series of the FVIIa variants in 10 mM HEPES, pH 7.4, 150 mM NaCl, 10 mM CaCl<sub>2</sub> and 0.05% (v/v) surfactant P-20 at 20°C at a flowrate of 50  $\mu$ l/min.  $k_{on}$  and  $k_{off}$  values were fitted by a non-linear regression to a bimolecular interaction model, assuming pseudo-first order reaction conditions. Subsequently, the affinity,  $K_D$ , could be calculated as  $K_D = k_{off}/k_{on}$ .

### Quartz crystal microbalance with dissipation monitoring

The experiments and subsequent data treatment were conducted as described previously (*Paper III*). In brief, a bilayer (30% POPS, 70% POPC) was formed using peptide discs as described (Paper III). 10  $\mu$ M sTF was injected in QCMD buffer: 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM CaCl<sub>2</sub>. For the FVIIa interaction experiment, a TF containing bilayer (30% POPS, 70% POPC) was formed using peptidediscs. After initial

wash with QCMD buffer, 10 $\mu$ M FVIIa (also in QCMD buffer) was injected. Subsequently, the surface was washed with QCMD buffer again.

### Molecular dynamics simulations

The models used for simulations were build based on the crystal structure of sTF (pdb-ID: 1BOY). The TMD was build using the MOLEFACTURE plugin in VMD (21). 5000 structures were created using the 'mp\_domain\_assembly' protocol to join the ECD and TMD of TF by a flexible linker (22). The best 10 models were selected and of these two main conformations were identified, which were used for the simulations: Non-canonical (NC) and Canonical-like (CL) (figure S1). A 30% POPS: 70% POPC bilayer with TF incorporated was build using the CHARMM-GUI suite (23). Simulations were performed for at least 600 ns using an ionic strength corresponding to 150 mM NaCl.

## Results

### Lipid interactions of the extracellular domain of Tissue factor

To investigate the potential lipid interaction between sTF and PS lipids, two different experimental setups were used. First, QCMD was performed in a setup where a membrane, containing PS-lipids, was formed using peptidediscs (figure 1E), after which sTF was injected (figure 2A). The data (figure 2B) showed no sign of adsorption to the lipid surface under these experimental conditions, as no change in  $\Delta F$  or  $\Delta D$  was observed.

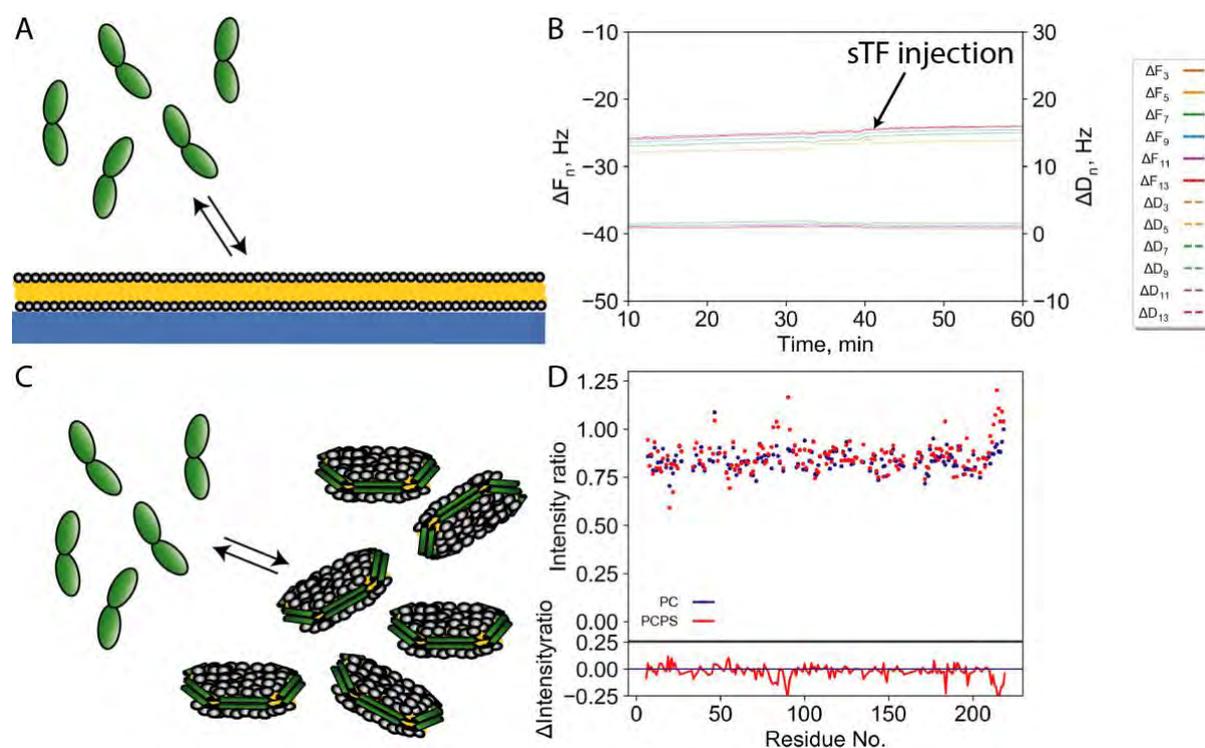


Figure 2: Lipid interactions of sTF. A: Experimental setup for monitoring lipid interactions by QCMD. B: QCMD data on the interaction between sTF and a POPC:POPS bilayer. C: Schematics of NMR experiment. D: Intensity changes of sTF upon addition of NDs (either POPC or POPC:POPS) relative to free sTF. Lower panel: Difference between intensity ratios of PC and PCPS.

To further investigate this, and to potentially map the residues involved in the lipid interaction, nuclear magnetic resonance (NMR) was conducted using a labeled version of sTF and the novel csdH5 NDs, which are relatively small (diameter  $\sim$ 9 nm, schematics in figure 2C). HSQC spectra were recorded, and upon addition of NDs with either pure POPC or a mix of POPC and POPS, lipid interactions were expected to be observed as either chemical shift changes or intensity changes. As assignments of sTF were available, these

could in principle readily be transformed to residues specific information. However, no chemical shift changes were observed upon addition of PS containing NDs (figure S1) and therefore, intensity changes were monitored (figure 2D). A drop was observed in the intensity upon addition of NDs, but the change was very similar for the two different NDs (figure 2D, lower panel) and seemed to affect the intensities globally, suggesting that the intensity change was not related to lipid binding but rather viscosity changes. Thus neither of the experimental setups confirmed the lipid binding suggested by MD simulations (12). However, it is important to underline that this does not necessarily mean that TF in a membrane does not have any lipid interactions; they are just not present under these experimental conditions. If an interaction is present, the affinity must be low (as we did not observe any sign of it) and it is likely that this affinity will be higher due to a proximity effect when TF is anchored to the membrane. This underlines the need for a model of FL TF incorporated in a membrane, to investigate these interactions and the general influence of the bilayer, both *in vitro* and *in silico*.

### Orientation of tissue factor in a bilayer

Using the recently developed protocol (Paper III) to form a lipid bilayer (POPC/POPS) containing TF, NR data was obtained (figure 3A). The calculated scattering length densities (SLDs) curves (figure 3B), and the fitting parameters (figure 3C), showed that the values of the bilayer were consistent with the previously reported numbers (Paper III) and that the height of the TF ECD layer was  $61 \pm 3$  Å. By simply measuring the largest distance in the crystal structure of TF ECD (pdb-ID: 1BOY), it was evident that the average position of TF in a bilayer had an angle of  $41^\circ$ , as illustrated in figure 3B.

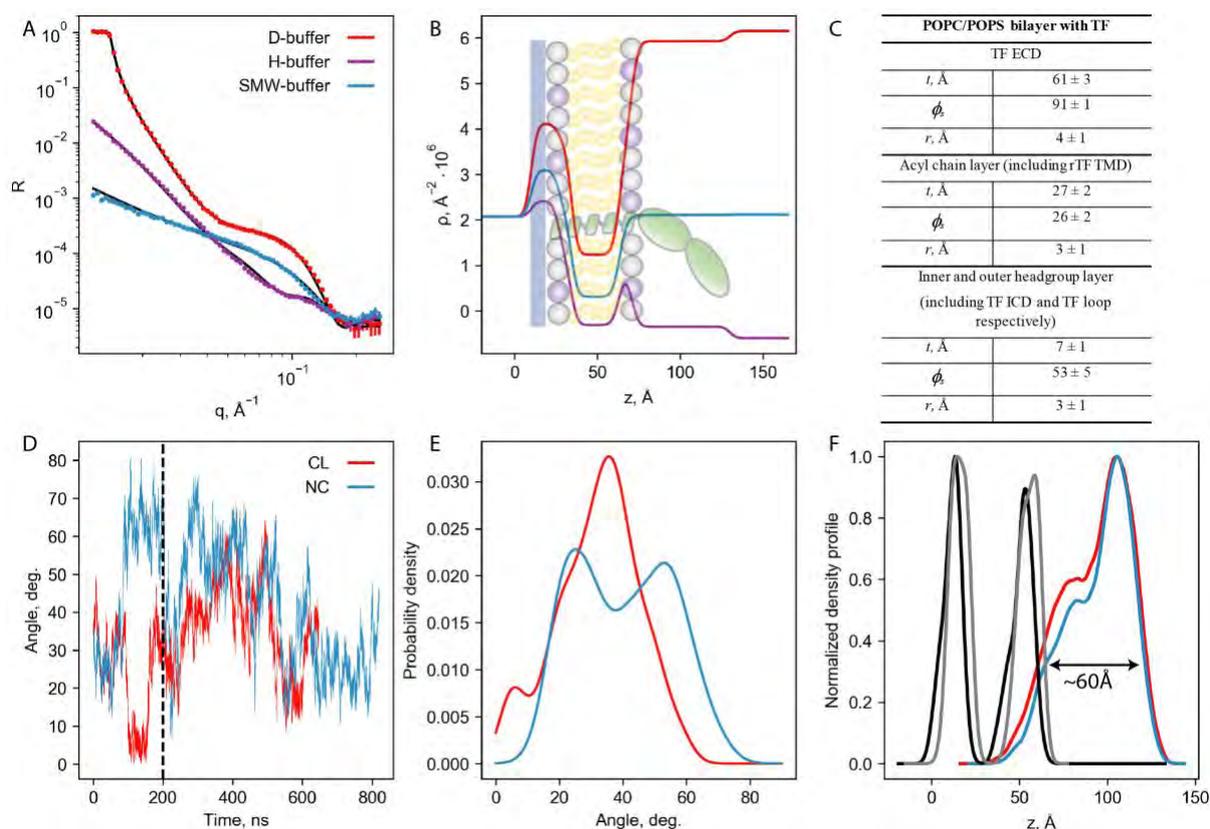


Figure 3: TF in a lipid bilayer. A: NR data with corresponding fits. Data is scaled for visualization. B: Scattering length densities calculated from the fits presented in B. C: Corresponding fitting parameters. D: Temporal evolution of the angles of Non-canonical (NC) and the Canonical-like (CL) starting conformations. Dashed line: equilibration. E: Probability density as a function of angle. The data is from after equilibration time (after vertical line in D). F: Density profile. Black: Phosphates from the CL simulation. Grey: Phosphates from the NC simulation.

To support these findings, we developed a computational model of TF in a bilayer, to investigate the dynamics of the system. The model of TF was constructed based on the crystal structure of sTF (3), and a novel approach to build flexible linkers between ECDs and TMDs (22). 5000 models were generated, and from these, two commonly occurring conformations could be identified: Non-canonical (NC) and Canonical-like (CL) (figure S2). To investigate the flexibility and orientation of the TF, simulations were performed using these two conformations in a POPC:POPS membrane. The orientation of both starting orientations seemed to reach an equilibrium after around 200 ns (dashed line, figure 3D) but generally showed a flexible nature. This was also demonstrated by probability density functions (figure 3E) which showed two relatively broad distributions. The average angles after equilibration were calculated to be  $40.14 \pm 15.29^\circ$  and  $35.96 \pm 11.42^\circ$  for the NC and the CL, respectively. Again, the large standard deviations reflects that TF occupies many different conformations during the simulation. However, both values are relatively close to the  $41^\circ$ , as probed by NR, suggesting a general accordance between experiments and the simulations. Lastly, the density profile was evaluated for the equilibrated part of the simulation (figure 3F). First of all, it seemed that also the density profiles were independent of the starting conformation. A minor fraction of the TF density was present among the headgroups of the outer leaflet, while the majority was around 40 Å above the lipid headgroups. The maximal distance in the density profile between the lipid headgroups and TF was 60-65 Å (depending on starting conformation), which was in agreement with the  $61 \pm 3$  Å found by NR. The fact that a fraction of the density was located among the lipid headgroups, suggested that lipid interactions with the TF-ECD were present. To investigate these interactions on a residue specific level, the lipid contacts of TF during the simulation were evaluated (figure S3). The linker regions was found to be in contact with the lipid headgroups during almost the entire simulation, but interestingly also three other regions in the ECD were found to have contacts in around 50% of the time. These regions (113-122, 154-167 and 174-184) span many of the residues previously found to confer lipids interaction, and intriguingly, many of the residues shown to be important for FX activation (12). Lastly, it should be mentioned that both contacts with POPC and POPS were observed to an equal degree, and thus no apparent PS preferential behavior was observed in this simulation.

### **Binding of FVIIa**

To assess the orientation of the TF:FVIIa complex, we first evaluated the ability of FVIIa to bind to TF on the surface. QCMD data (figure 4A) showed a clear drop in  $\Delta F$  upon injection of FVIIa, indicating complex formation. After washing with buffer, it seemed that a fraction of FVIIa was removed, which was most likely the fraction of the FVIIa bound to the membrane and not TF, however the majority was stable. A sample for NR was prepared in the same manner, and the recorded data (figure 4B) demonstrated different features compared to the TF membrane (figure 3A). The data was fitted using the same layers, and the calculated SLD curves (figure 4C) with their corresponding fitting parameters (figure 4D) showed an increase in the height of the protein layer. The height increased from  $61 \pm 3$  Å to  $99 \pm 3$  Å consistent with the presence of a taller protein (FVIIa).

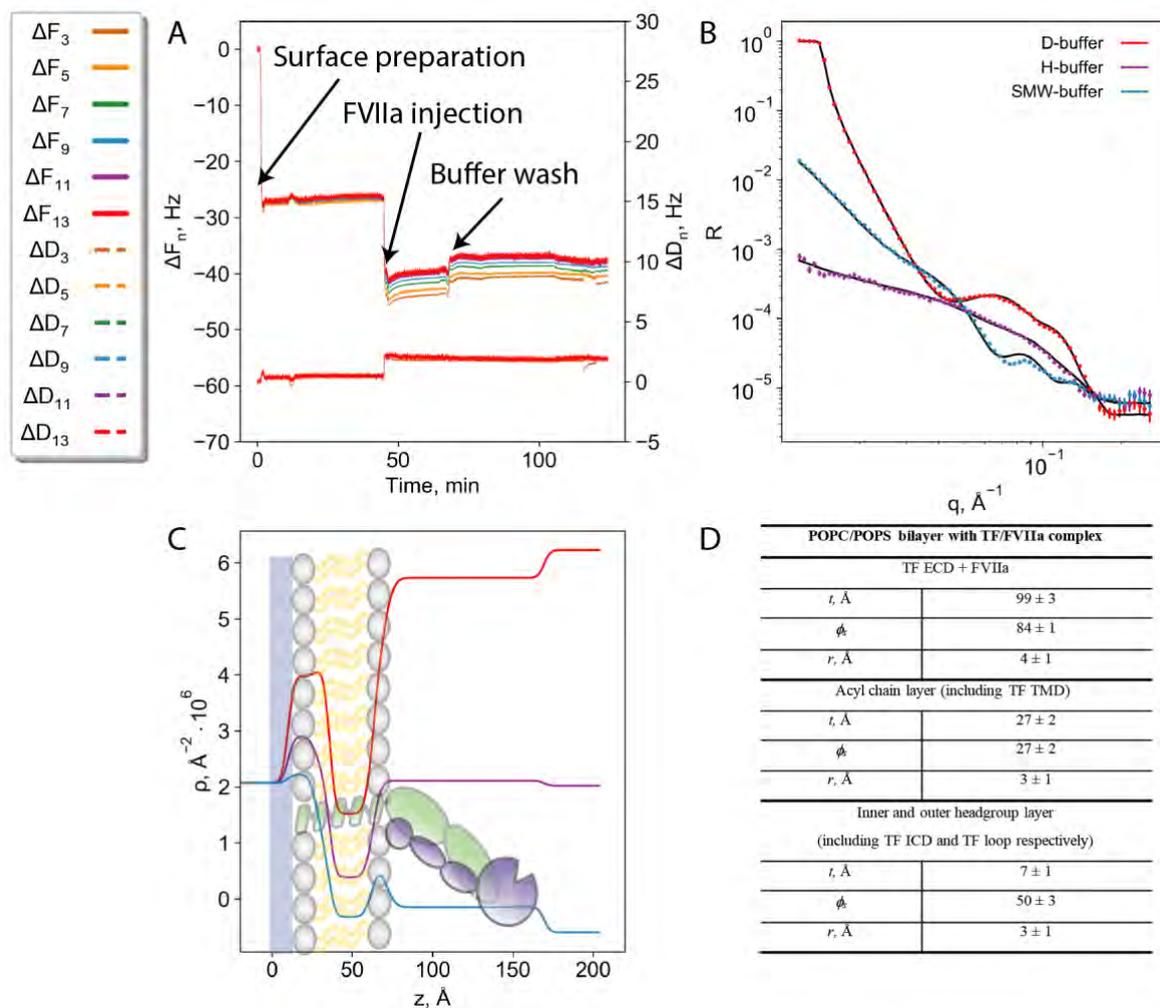


Figure 4: Formation of the TF:FVIIa complex. A: QCMD showing the formation of first the bilayer containing TF and subsequent injection of FVIIa followed by a buffer wash. B: NR data with corresponding fits. Data is scaled for visualization. C: Scattering length densities calculated from the fits presented in B. D: Corresponding fitting parameters.

### Unfolding of the Gla-domain affects complex formation

As mentioned, the Gla-domain of FVIIa is reported to interact with a lipid bilayer with an affinity in the  $\mu\text{M}$ -range (8), and is believed to play an essential role in the activity of the TF:FVIIa complex on a membrane. However, DG-FVIIa is also able to bind TF, albeit with an affinity 1000 times lower than FL FVIIa. This was confirmed by SPR measurements, where the dependency of PS lipids also was probed (figure S4). Neither of the two interactions seemed to be profoundly affected by the PS content in the membrane.

It is well established that calcium is needed for proper folding of the Gla-domain, so we tested how removal of calcium would affect the system, by forming the TF:FVIIa complex and subsequently flushing the surface with an EDTA containing buffer. The QCMD data (figure S5) showed recovery of the same  $\Delta F$  as before FVIIa injection, indicating complete removal of FVIIa. Performing the same experiment using NR (figure 5A), showed that the curves for the TF bilayer, and the same TF bilayer after FVIIa removal, were quite similar. However, some changes were observed in the raw data, but when looking at the calculated SLD profiles (figure 5B) and the corresponding fitting parameters (figure 5C), they were almost identical.

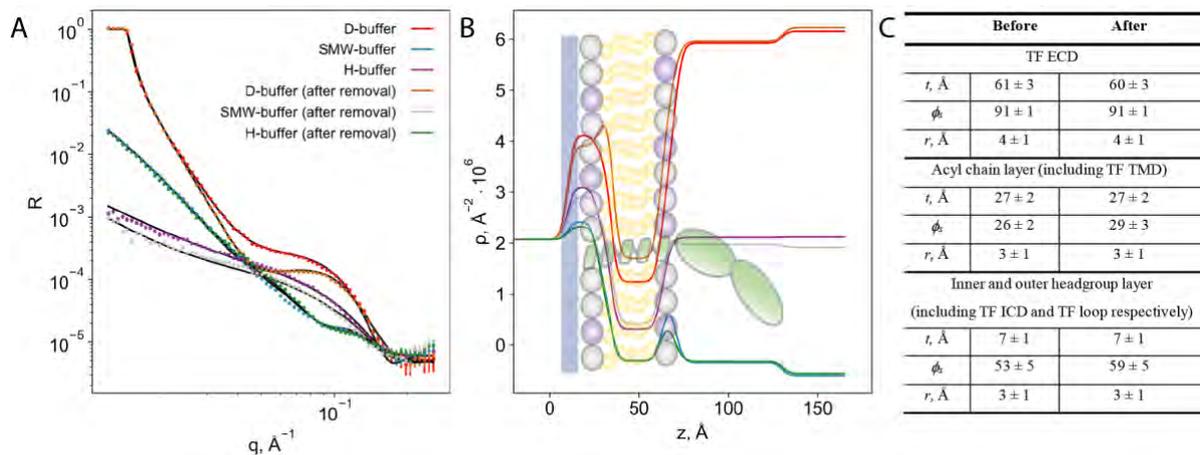


Figure 5: Removal of FVIIa by removing calcium. First measurements were conducted with TF in a membrane. Then FVIIa was added and measured (data from figure 4) and lastly, calcium was removed and the same membrane was measured again (called 'after' in this figure) A: Experimental NR data with fits for each buffer condition. Data is scaled for visualization purposes. B: SLD profiles calculated from the NR fits, with a schematic illustration of TF in a membrane overlaid.

As the DG-FVIIa is able to bind TF in a bilayer, the results above indicated that the unfolded Gla-domain prevented the formation of the complex. We investigated this by performing SAXS measurements on FVIIa in the presence and absence of calcium (figure S6). The data showed that FVIIa was clearly elongated upon calcium removal, as the  $D_{max}$  increases by around 25 Å. As the binding interface between TF and FVIIa is quite rigid, it is likely that the elongated structure of FVIIa without calcium prevents the complex formation. This is probably due to a sterically hindrance posed by the membrane, and the elongated Gla-domain, which distances FVIIa from the TF interface. These findings clearly underline the importance of proper folding and positioning of the Gla-domain, but also possess an important experimental feature in regenerating a 'clean TF membrane' for a new experiment.

### Factor Xa on a lipid membrane

Similarly to FVIIa, FXa is also known to bind the membrane, though with a higher affinity (7). We examined the interaction of FVIIa and FXa with the membrane using NR, to investigate the height and orientation of the proteins. For FVIIa, unfortunately we were not able to obtain useful results, as the buffer change during the experiment washed away the majority of FVIIa, leaving an undetectable amount for NR. The same experiments were conducted using QCMD (data not shown), which confirmed the presence of FVIIa on the membrane, even after buffer washing, but QCMD is much more sensible than NR. On the other hand, reasonable amounts of FXa stayed bound to the membrane, and NR data was collected (figure 6). The data showed a bilayer with the expected characteristics, and with an FXa layer with a height of  $92 \pm 3$  Å, resulting in a slightly tilted orientation, as illustrated in figure 6B. A previous MD study suggested that the Gla-domain of FX penetrates the bilayer, in such a way that the calcium ions were able to interact with the phosphates, and the three hydrophobic residues of FX interacted with the lipid acyl chains (24). However, the current NR data neither confirmed nor invalidated this, as it would require a higher contrast between the lipid headgroup and FXa, to distinguish these.

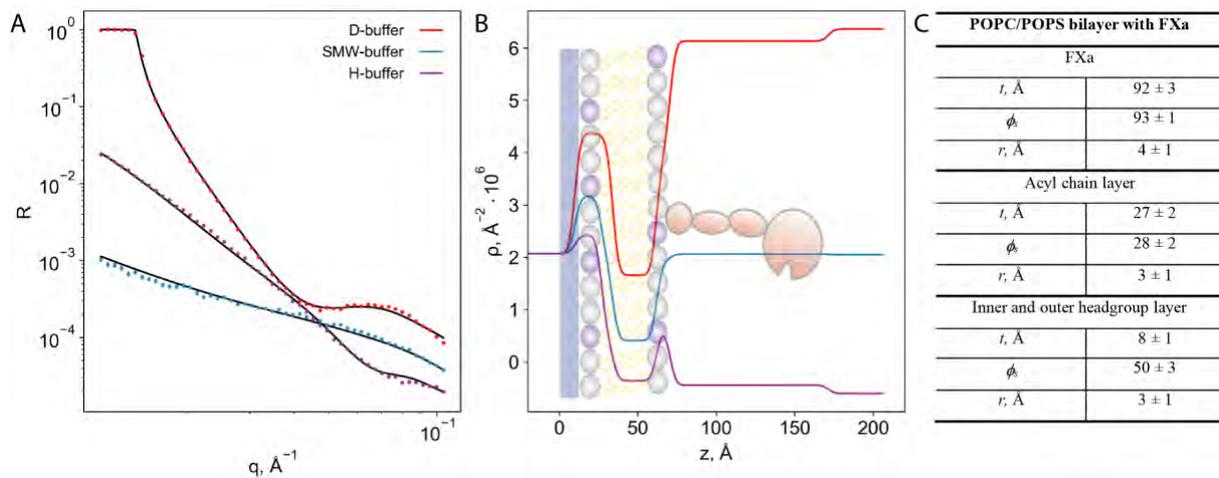


Figure 6: FXa binding to a lipid membrane. A: Experimental NR data with fits for each buffer condition. Data is scaled for visualization. B: SLD profiles calculated from the NR fits, with a schematic illustration of FXa on a membrane overlaid. C: Parameters for the SLD fits.

## Discussion

With all the measured distances at hand, it is possible to speculate about the functional implications. The crystal structure of sTF is available, and the longest distance in the ECD was measured to be 80 Å (figure 7A). The structure of FVIIa is also available, albeit with a few lacking residues, and the structure of full length FXa is not solved. The individual proteins in the computational model presented in (25) of the tertiary sTF:FVIIa:FXa complex, was built based on the available crystal structures and completed by homology modeling. From this model, the maximal distance in FVIIa was measured to be 120 Å (figure 7B) while it was 110 Å for FXa (figure 7C). As the two proteins' Gla-domains are expected to be buried in the bilayer, this height (8 Å) was subtracted from the total height (24,26). Using these distances and the heights measured by NR, simple trigonometry was used to calculate the angles of TF (figure 7D), TF:FVIIa (figure 7E) and FXa (figure 7F) in regards to the membrane.

The active site of FVIIa is located approximately 94 Å from the bottom of the Gla-domain (red residues, figure 7B) which, assuming the angle of 27°, means that it is located approximately 77 Å above the membrane (figure 7E). This is in agreement with the  $75 \pm 3$  Å determined previously by FRET (9), and also close to the  $\sim 79$  Å found by a MD simulation (26). All of these numbers are vitiated with uncertainties: NR only provides the height of the layer, and the exact orientation of the complex is unknown, and thus also the precise location of the active site is unknown. On the other hand, the FRET experiments were conducted using a probe located directly inside the active site of FVIIa, however FRET measurements are highly dependent on the orientation of the two probes, and only small changes in this orientation can affect the measured distances dramatically (27). Lastly, the simulations are highly dependent on the used force field, and are thus usually supported by experimental data. However, as all three values are very similar, these numbers seem trustworthy enough to continue with further analysis.

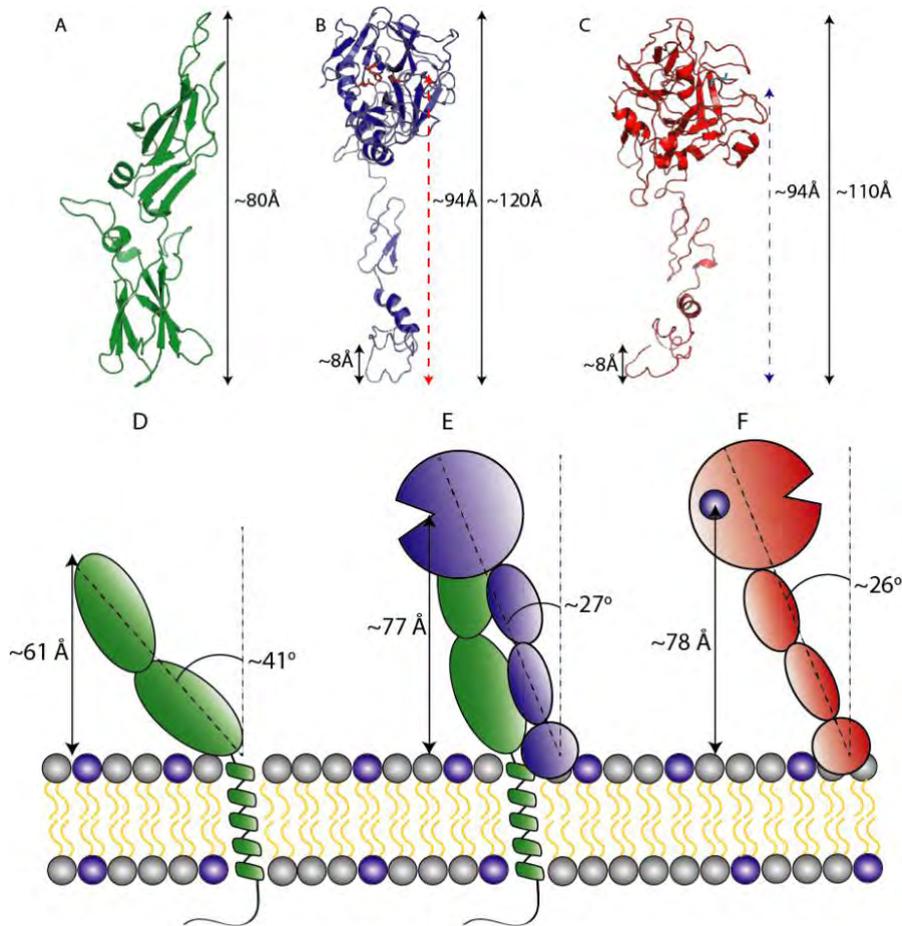


Figure 7: Important distances. A: sTF, pdb-ID: 1BOY. B: FVIIa: Model from (25). Active site residues in red. C: FXa: Model from (25). New terminal residue after cleavage by FVIIa in teal. D: TF in a membrane with related height from NR and calculated angle. E: TF:FVIIa complex in a membrane with calculated angle and height of active site over the membrane. F: FXa on a membrane with calculated angle and height of cleavage site over the membrane.

The model of FXa of course does not contain the scissile bond, and thus the position of the new terminal residue (left after cleavage, teal residue, figure 7E), was used to estimate the position of the scissile bond. Using this approximate height and the calculated angle, the distance of the scissile bond above the membrane was calculated to be 78 Å. Even though this distance is associated with uncertainties, it is quite curious that the FVIIa active site height (in the TF complex) and the scissile bond on FX were estimated to be the same (77 Å and 78 Å, respectively). This positioning of the proteins, is likely one of the functions of the membrane, supporting the increase in activity mediated by the presence of a bilayer. The increase in TF:FVIIa activity upon the presence of PS-lipids, might therefore not only be a question of recruiting FX, but also a question of forming the exact membrane interaction that will facilitate the correct position of the FX scissile bond in relation to the FVIIa active site.

The Gla-domain interaction with a pure PC lipid bilayer is expected to be different (or non-existing), as no serine is available for contact (26), and the headgroup of PC is too bulky to allow penetration into the lipid tails. Therefore, it is possible that the increase in TF:FVIIa activity upon PS exposure is mediated by the Gla-domain dependent localization of the FVIIa active site. Although the FRET study (10) suggested a similar height of the active site of DG-FVIIa, it is still possible that a Gla-domain, unable to interact with a pure PC membrane, will enforce a changed orientation of the TF:FVIIa complex. At least, an unfolded Gla-domain will not support activation, as it hinders TF interaction, which serves as an example of how delicate the position and conformation of the Gla-domain is.

Using the heights from NR measurements, it is important to keep in mind that they represent the average height of the protein layer, and in practice a population of different heights ranging around the average is most likely the case. This was also reflected in the simulations of TF in a membrane, which showed a range of conformations around the measured heights. Investigating the flexibility of the TF:FVIIa on a membrane, using similar simulations, would provide very interesting information about the mechanism behind the activity. With the results currently at hand, we hypothesize that the flexible nature of TF, functions as an anchor for FVIIa, which rigidifies the complex upon binding, and fixes the FVIIa active site in the correct height for catalytic efficiency. In contrast to the experiments conducted with sTF, our simulation suggested that the TF-ECD possesses lipid interactions, but whether these confer a functional mechanism cannot be determined at this point.

## Conclusion

Structural data on the TF:FVIIa complex in a lipid bilayer might be the missing piece needed to explain the PS induced increase in activity of the complex. In the present study, we utilized a novel deposition technique, allowing us to utilize NR to measure the heights of the proteins above the membrane. Using these heights and the available structure, we could calculate the height of the FVIIa active site and the FX scissile bond. These two proved to be very similar, suggesting that the effect of the lipid bilayer (containing PS-lipids) is to localize the two sites in the same height. The role of the lipids and TF binding was also investigated, as we observed no interactions between sTF and lipids. However, incorporating TF in a membrane, allowed us to determine the angle with respect to the membrane, which we found to be 41°. MD simulations supported this tilted angle, but suggests that the angle reflects quite a flexible nature, with a distribution of conformation around this angle. The simulation also pointed towards TF interactions with the membrane in a minor fraction of the time, and these interactions are likely to be mediated by a proximity effect, as TF is anchored to the membrane, contrary to the experiments performed with sTF.

## References

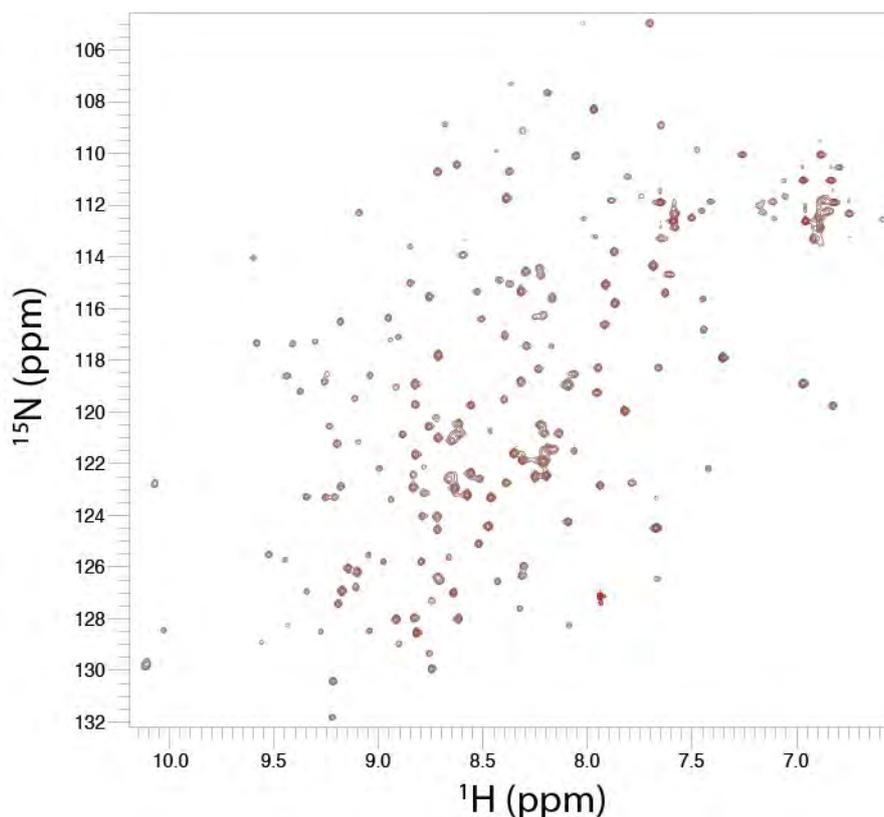
1. Sen M, Herzik M, Craft Jr. JW, Creath AL, Agrawal S, Ruf W, et al. Spectroscopic Characterization of Successive Phosphorylation of the Tissue Factor Cytoplasmic Region. *Open Spectrosc J.* 2009;3(713):58–64.
2. Nemerson Y. The phospholipid requirement of tissue factor in blood coagulation. *J Clin Invest.* 1968;47(1):72–80.
3. Harlos K, Martin DM, O'Brien DP, Jones EY, Stuart DI, Polikarpov I, et al. Crystal structure of the extracellular region of human tissue factor. *Nature.* 1994;371(6499):720–720.
4. Banner DW, D'Arcy A, Chene C, Winkler FK, Guha A, Konigsberg WH, et al. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. Vol. 380, *Nature.* 1996. p. 41–6.
5. Padmanabhan K, Padmanabhan KP, Tulinsky A, Park CH, Bode W, Huber R, et al. Structure of human des(1-45) factor Xa at 2.2 Å resolution. Vol. 232, *Journal of Molecular Biology.* 1993. p. 947–66.
6. Neuenschwander PF, Morrissey JH. Roles of the membrane-interactive regions of factor VIIa and tissue factor: The factor VIIa Gla domain is dispensable for binding to tissue factor but important for activation of factor X. *J Biol Chem.* 1994;269(11):8007–13.
7. Shaw AW, Pureza VS, Sligar SG, Morrissey JH. The local phospholipid environment modulates the activation of blood clotting. *J Biol Chem.* 2007;282(9):6556–63.
8. Tavoosi N, Smith SA, Davis-Harrison RL, Morrissey JH. Factor VII and protein C are phosphatidic acid-binding proteins. *Biochemistry.* 2013;52(33):5545–52.
9. McCallum CD, Hapak RC, Neuenschwander PF, Morrissey JH, Johnson AE. The location of the active site of blood coagulation factor VIIa above the membrane surface and its reorientation upon association with tissue factor - A fluorescence energy transfer study. *J Biol Chem.*

- 1996;271(45):28168–75.
10. McCallum CD, Su B, Neuenschwander PF, Morrissey JH, Johnson AE. Tissue factor positions and maintains the factor VIIa active site far above the membrane surface even in the absence of the factor VIIa Gla domain. A fluorescence resonance energy transfer study. *J Biol Chem*. 1997;272(48):30160–6.
  11. Waters EK, Morrissey JH. Restoring full biological activity to the isolated ectodomain of an integral membrane protein. *Biochemistry*. 2006 Mar;45(11):3769–74.
  12. Ke K, Yuan J, Morrissey JH. Tissue factor residues that putatively interact with membrane phospholipids. *PLoS One*. 2014;9(2).
  13. Ohkubo YZ, Morrissey JH, Tajkhorshid E. Dynamical view of membrane binding and complex formation of human factor VIIa and tissue factor. *J Thromb Haemost*. 2010;8(5):1044–53.
  14. Kirchhofer D, Lipari MT, Moran P, Eigenbrot C, Kelley RF. The tissue factor region that interacts with substrates factor IX and factor X. *Biochemistry*. 2000;39(25):7380–7.
  15. Hsieh MH, Huang PT, Liou HH, Liang PH, Chen PM, Holt SA, et al. The Penetration Depth for Hanatoxin Partitioning into the Membrane Hydrocarbon Core Measured with Neutron Reflectivity. *Langmuir*. 2018;34(30):9036–46.
  16. Rehal R, Gaffney PRJ, Hubbard ATM, Barker RD, Harvey RD. The pH-dependence of lipid-mediated antimicrobial peptide resistance in a model staphylococcal plasma membrane: A two-for-one mechanism of epithelial defence circumvention. *Eur J Pharm Sci*. 2019;128(November 2018):43–53.
  17. Hoogerheide DP, Noskov SY, Kuszak AJ, Buchanan SK, Rostovtseva TK, Nanda H. Structure of voltage-dependent anion channel-tethered bilayer lipid membranes determined using neutron reflectivity. *Acta Crystallogr Sect D, Struct Biol*. 2018 Dec;74(Pt 12):1219–32.
  18. Soranzo T, Martin DK, Lenormand JL, Watkins EB. Coupling neutron reflectivity with cell-free protein synthesis to probe membrane protein structure in supported bilayers. *Sci Rep*. 2017;7(1):1–10.
  19. Pernot P, Round A, Barrett R, De Maria Antolinos A, Gobbo A, Gordon E, et al. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J Synchrotron Radiat*. 2013;20(4):660–4.
  20. Johansen NT, Pedersen MC, Tidemand FG, Rand KD, Nguyen TTTN, Arleth L. Circularized and solubility-enhanced MSPs facilitate simple and high yield production of stable nanodiscs for studies of membrane proteins in solution. *FEBS J*. 2019;
  21. Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. *J Mol Graph*. 1996;14(October 1995):33–8.
  22. Koehler Leman J, Bonneau R. A Novel Domain Assembly Routine for Creating Full-Length Models of Membrane Proteins from Known Domain Structures. *Biochemistry*. 2018;57(13):1939–44.
  23. Wu EL, Cheng X, Jo S, Rui H, Song KC, Dávila-Contreras EM, et al. CHARMM-GUI Membrane Builder Toward Realistic Biological Membrane Simulations. *J Comput Chem*. 2014;35(27):1994–2004.
  24. Muller MP, Wang Y, Morrissey JH, Tajkhorshid E. Lipid Specificity of the Membrane Binding Domain of Coagulation Factor X. *J Thromb Haemost*. 2017;15(10):2005–16.
  25. Norledge B V., Petrovan RJ, Ruf W, Olson AJ. The Tissue Factor/Factor VIIa/Factor Xa Complex: A Model Built by Docking and Site-Directed Mutagenesis. *Proteins Struct Funct Genet*. 2003;53(3):640–8.
  26. Ohkubo YZ, Tajkhorshid E. Distinct Structural and Adhesive Roles of Ca<sup>2+</sup> in Membrane Binding of Blood Coagulation Factors. *Structure*. 2008;16(1):72–81.
  27. Vogel SS, Van der Meer BW, Blank PS. Estimating the distance separating fluorescent protein FRET pairs. *Methods*. 2014;15(66):131–8.

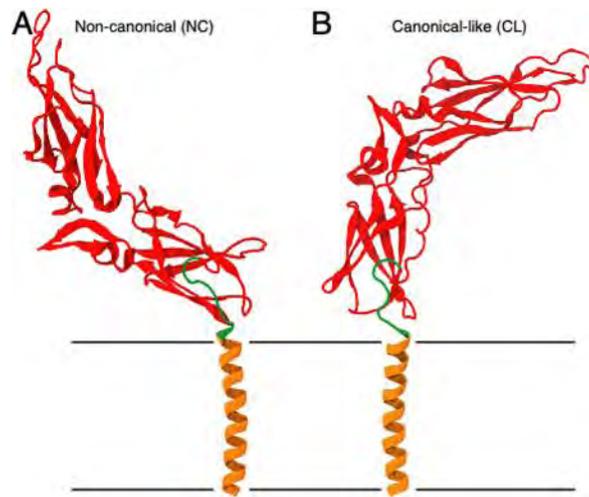
28. Sen P, Neuenschwander PF, Pendurthi UR, Rao LVM. Analysis of factor VIIa binding to relipidated tissue factor by surface plasmon resonance. *Blood Coagul Fibrinolysis*. 2010;21(4):376–9.
29. Huang M, Rigby AC, Morelli X, Grant MA, Huang G, Furiel B, et al. Structural basis of membrane binding by Gla domains of vitamin K-dependent proteins. *Nat Struct Biol*. 2003;10(9):751–6.

### Supplementary information

**S1:** Lipid interactions of sTF as probed by NMR.  $2D,^{15}N$ -labeled sTF used for recording  $^1H,^{15}N$ -HSQC spectra. Blue: sTF (50  $\mu M$ ). Red: sTF (50  $\mu M$ ) + csdH5, POPC,POPS (250  $\mu M$ ). In conclusion, no chemical shift changes were observed. Changes in intensity of the peaks were also monitored as presented in main text.

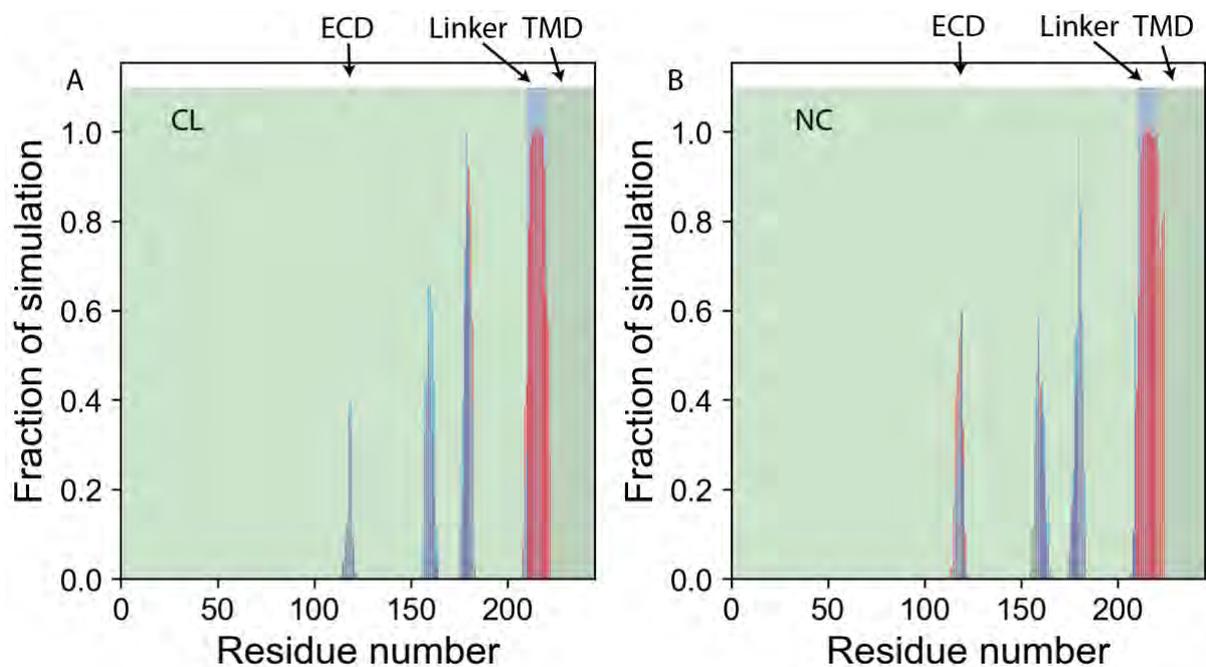


**S2:** Input conformation for simulations of TF in a membrane. Red: sTF from pdb-ID: 1BOY. Orange: Generated helical TMD. Green: Flexible linker generated using the 'mp\_domian\_assembly' protocol implemented in Rosetta\_MP (22). 5000 conformations were generated and two commonly occurring conformations were identified and used for subsequent simulations: A: Non-canonical (NC). B: Canonical-like (CL).



**S3:** TF lipid contacts. Red: POPC. Blue: POPS. CL: Canonical-like. NC: Non-cononical. ECD, linker and TMD as defined in S2.

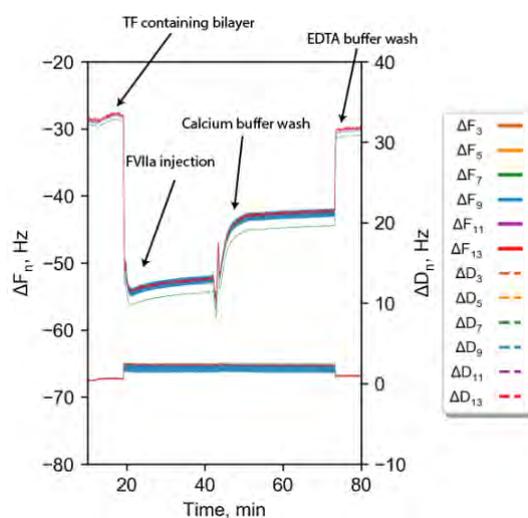
The contacts between are very similar for POPS and POPC and are seen in four different regions. The linker is almost during the entire simulation in contacts with lipids, while three other regions (in the ECD) also have many contacts.



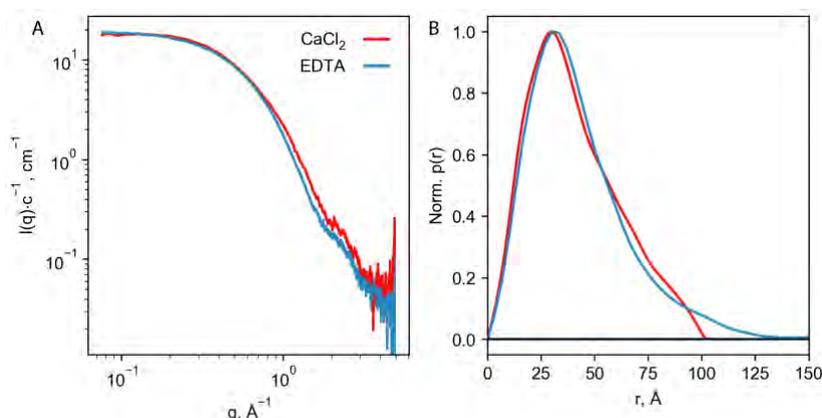
**S4:** Binding kinetics of the interaction between TF and FVIIa or DG-FVIIa. SPR measurements on TF in NDs immobilized on anti-his antibodies. Injection of FVIIa or DG-FVIIa provided the binding kinetics found in the table. The TF NDs had three different lipid compositions with an increasing amount of POPS (and decreasing of POPC). All affinities are within range of previously reported (6,28).

Sample	DG-FVIIa			FVIIa		
	$k_{on}$ (M <sup>-1</sup> s <sup>-1</sup> )	$k_{off}$ (s <sup>-1</sup> )	$K_D$ (nM)	$k_{on}$ (M <sup>-1</sup> s <sup>-1</sup> )	$k_{off}$ (s <sup>-1</sup> )	$K_D$ (pM)
0% POPS	$0.35 \pm 0.006$	$29.7 \pm 0.5$	84.9	$3.28 \pm 0.002$	$0.60 \pm 0.05$	180
15% POPS	$0.54 \pm 0.002$	$22.8 \pm 0.5$	42.2	$4.84 \pm 0.008$	$1.00 \pm 0.02$	210
30% POPS	$0.65 \pm 0.005$	$24.3 \pm 0.1$	37.5	$3.30 \pm 0.003$	$1.02 \pm 0.07$	310

**S5:** QCMD data on FVIIa binding to a membrane. A change in  $\Delta F$  was observed upon FVIIa injection indicating that interaction occurred. After washing with the calcium containing buffer a fraction of the FVIIa left the surface, but the new plateau was clearly different from the initial level. When the sample was flushed with EDTA containing buffer, the initial  $\Delta F$  level was reached again, indicating a complete regeneration of the pure membrane.



**S6:** Calcium dependent folding of the FVIIa's Gla-domain. SAXS measurements were performed to investigate the calcium dependent fold of the Gla-domain of FVIIa. SEC-SAXS runs were performed in the presence and absence (EDTA) of calcium, and the final data was as presented in A. Upon removal of calcium (EDTA), FVIIa clearly elongates as both indicated by the earlier descent in the raw data and the larger  $D_{max}$  from the  $p(r)$ -function in B. This coincides with previously reported results indicating that calcium is necessary to fold the Gla-domain, which otherwise has a more flexible and elongated structure (29).

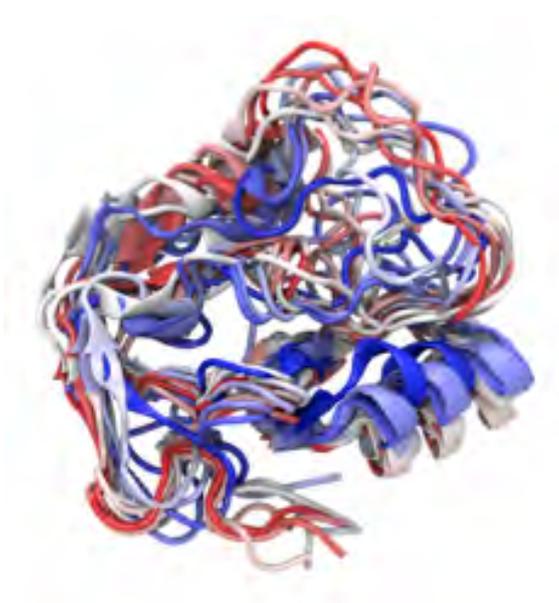


## 10 | Paper VI

### Probing the flexibility of the blood coagulation inhibitor, NAPc2

*Frederik Grønæk Tidemand, Katrine Bugge, Raul Araya-Secchi, Birthe B. Kragelund, Lise Arleth*

**Note:** This manuscript lacks the final NMR experiments before the draft can be finalized. Thus the NMR part is currently incomplete.



# Probing the flexibility of the blood coagulation inhibitor, NAPc2

Frederik Grønbaek Tidemand<sup>1</sup>, Raul R Araya-Secchi<sup>1</sup>, Katrine Bugge<sup>2</sup>, Birthe B. Kragelund<sup>2</sup>, Lise Arleth<sup>1</sup>

1: Structural Biophysics, X-ray and Neutron Science, The Niels Bohr Institute, University of Copenhagen, Denmark.

2: Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Denmark

## Abstract

NAPc2 is an inhibitor of the intrinsic coagulation pathway, in which it inhibits the tissue factor:Factor VIIa (FVIIa) complex in a Factor X dependent mechanism. The structure of NAPc2 is unusual as it is mainly constrained by five disulfide bridges, has no hydrophobic core and only has a few small secondary structure elements. We investigated the solution structure of NAPc2 by a combination of small angle X-ray scattering (SAXS) and molecular dynamics simulations, from which we found NAPc2 to be much less compact than previously reported. The simulations highlighted the most flexible regions of the structure, and in general, the overall size of the ensemble converged towards the size described by the SAXS data. The flexibility of NAPc2 is likely to be important for its promiscuous binding behavior, as it only functions in other organisms than the nematode it originates from. In this context, its interaction with human FVIIa was investigated, using nuclear magnetic resonance (NMR). Initial NMR experiments were conducted, confirming the binding to FVIIa although complete assignments are still lacking.

## Introduction

It is estimated that more than one out of a thousand individuals suffer from thrombosis each year (1), making this the third most common vascular disease, and resulting in a long history of research within this field.

In the search for drugs against thrombosis, the nematode *Ancylostoma caninum* was investigated, as it had previously proven to have strong anticoagulant properties (2). The main contributions to the anticoagulant properties were narrowed down to the three proteins, NAP5, NAP6, and NAPc2, which are all very potent inhibitors of blood clotting. More specifically, they inhibit the two clotting factors, Factor VIIa (FVIIa) and Factor Xa (FXa) both of which play an important role in the intrinsic part of the human blood coagulation pathway. The interaction between the integral membrane protein tissue factor (TF) and FVIIa initiates the intrinsic coagulation pathway, as it increases the catalytic activity of FVIIa by a factor  $10^5$  (3,4). FVIIa cleaves Factor X (FX) at a specific site, resulting in its activated counterpart: FXa. FXa then goes into the next step of the coagulation pathway, ultimately leading to fibrin formation and blood clotting. The tertiary TF:FVIIa:FXa complex is still structurally uncharacterized, even though high-resolution structures of the individual parts have been available for many years (5–7). NAPc2 was found to inhibit the TF:FVIIa complex with a very high affinity, by a mechanism in which FXa is crucial (see illustration of the quaternary complex in figure 1A). While NAPc2 binds the active site of FVIIa, it binds to an exosite on FXa underlining the ‘carrier’ effect of FXa. This results in an FXa that is active on peptide substrates, but inactive on its native substrate, Prothrombin, because it cannot form the Prothrombinase complex with factor V when it is ‘trapped’ by NAPc2 in the quaternary complex (8). As seen in the illustration (figure 1A), both FVIIa and FXa interact with the lipid bilayer, which in general is an essential component in the coagulation pathway (9,10). NAPc2 interacts with FX and FXa with a  $K_d$  in the low pM range, while the interaction with FVIIa has a  $K_d$  in the low  $\mu$ M range (11). In the mechanism of NAPc2 binding, it is generally accepted that NAPc2 and FXa form an initial complex, but whether this complex binds to the membrane, and then translates until encountering a TF:FVIIa complex, or simply binds a TF:FVIIa complex from solution is still unclear (11,12).

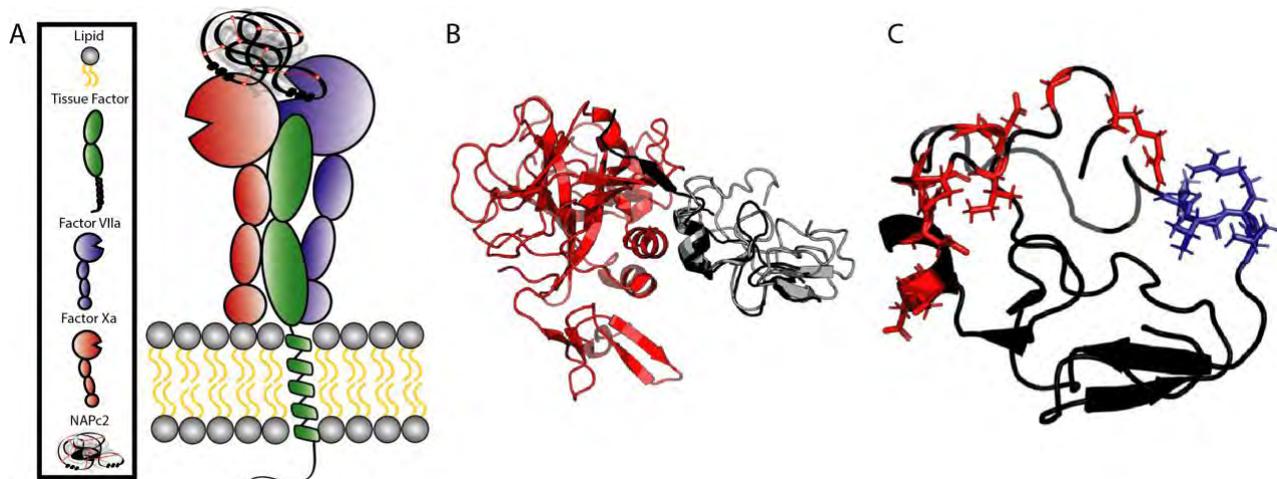


Figure 1: A: Schematic illustration of the quaternary TF:FVIIa:FXa:NAPc2 complex in/on a lipid bilayer. B: Crystal structure of the FXa:NAPc2 complex (2H9E) overlaid with the NMR structure of NAPc2 (1COU). FXa: Red. NAPc2: Black. NMR structure 1COU: Grey. C: NMR structure of NAPc2 with the FXa interface marked with red and the proposed FVIIa binding interface marked with blue.

On a structural level NAPc2, is a curious protein with 84 residues, five disulfide bridges and almost no secondary structure suggesting that the primary constraint of the size is the disulfide bonds. A nuclear magnetic resonance (NMR) ensemble showing a very flexible nature with only a few smaller secondary structure elements was published in 1999 (see figure 1B+C) (13). In addition, this NMR structure demonstrates that the NAPc2 does not contain a hydrophobic core, but is mainly supported by salt bridges, placed in the interior of the protein due to the high amount of charged residues in the sequence (40% of the residues are charged). The flexibility of the protein is probably critical, since the protein functions in organisms other than the nematode itself, making a promiscuous binding behavior essential. In different mammals, blood coagulation factors are similar but not identical, and the flexibility of NAPc2 is likely to embrace these differences. Until now, only the interactions between NAPc2 and human and bovine FXa has been probed, but both of these affinities are very tight (8,14).

The interaction between NAPc2 and FXa is well described, both in terms of affinity (8) and structure (14), but the structural features of the FVIIa interaction are not yet well understood. The crystal structure of NAPc2 and FXa (figure 1B) demonstrates an interface of interaction consisting of a small helix, a short beta strand, and a few disordered contact points. Looking at the overlay of the NAPc2 crystal and NMR structures (figure 1B), two things are revealed: *i)* folding upon binding must happen, since the NAPc2 NMR structure does not entirely overlay the crystal structure, and *ii)* almost no secondary structure elements are left for the interaction with FVIIa, which either means that the interaction purely relies on disordered contact points, or that a folding upon binding event also occurs in this interaction. Folding upon binding can either occur directly in the binding event, or as a result of transient structure elements forming in the apoprotein, and this conformation having a higher affinity for the binding partner (15). The binding interface between NAPc2 and FVIIa is not structurally determined, but biochemical assays suggest that it is the flexible region spanning Leu<sub>42</sub> to Val<sub>45</sub> in NAPc2 that binds in the active site of FVIIa (11) (blue residues, figure 1C).

The samples for most experiments performed on recombinant NAPc2 are produced in *Pichia pastoris*, as the yeast facilitates the formation of disulfide bridges (2). However, a more recent study showed that it is possible to produce NAPc2 in *E.coli*, in an active state using a 20 kDa intein and chitin-binding domain, as an affinity tag for purification.

In the present study, we demonstrate the production of a His-tagged version of NAPc2 in *E. coli* and also set out to investigate the solution structure and flexibility of NAPc2 using small angle X-ray scattering (SAXS) and molecular dynamics (MD). In addition, the interaction between NAPc2 and FVIIa was investigated using NMR.

## Materials and methods

### Expression and purification of NAPc2

All chemicals are from Sigma Aldrich unless otherwise stated.

A condon optimized NAPc2 containing a His<sub>6</sub>-tag and a TEV-site N-terminally was synthesized and cloned into the pET28b+ vector by GenScript. The plasmid was transformed into competent BL21 Origami cells, and expression was performed in TB media. The culture was grown at 37°C until it reached an optical density at 600 nm (OD<sub>600</sub>) of 0.8. The culture was then cooled to 20°C and induced with 1 mM IPTG over night (ON). The cells were harvested, and the pellet was resuspended in 3.5 ml lysis buffer (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 20 mM Imidazole, 1mM PMSF) per gram of cells and passed through a cell disrupter at 30kPsi (E1061, Constant Systems LTD). Cellular debris was pelleted with centrifugation at 20000g for 20 min, and the supernatant was incubated for 1 hour with NiNTA resin (equilibrated in lysis buffer) on a tilting table. The preparation was loaded onto an empty disposable column, and the flowthrough was collected. The resin was washed with 5 column volumes (CV) of washing buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 40 mM Imidazole) and eluted in 3 CV of elution buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 250 mM Imidazole). Fractions were collected and analyzed by SDS-PAGE. Fractions containing NAPc2 were pooled, and TEV was added to a 1:100 ratio. The sample was put in a 3.5MW cutoff dialysis bag, submerged in 2 l of TEV buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM DTT and 0.1 mM EDTA) and placed at 4°C ON under magnetic stirring. The sample was put in a falcon tube, and equilibrated NiNTA (lysis buffer) was added to the sample and incubated for 1 hour on a tilting table. The sample was poured onto a disposable column, and the flowthrough was collected. 1.5 CV of lysis buffer was added, and the flowthrough was again collected. The preparation was concentrated using a spinfilter (cutoff 3 kDa) and loaded onto a Superdex 200 HiLoad (26/600) equilibrated in 20 mM Tris-HCl, pH 7.5 and 100 mM NaCl and the peak fractions were collected.

### Expression and purification of labeled NAPc2

Expression of <sup>15</sup>N labeled NAPc2 was performed similarly to unlabeled NAPc2, except that the media was <sup>15</sup>N-labelled M9-media (3 g l<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub>, 7.5 g l<sup>-1</sup> Na<sub>2</sub>HPO<sub>4</sub> · H<sub>2</sub>O, 5 g l<sup>-1</sup> NaCl, 1mM MgSO<sub>4</sub>, 1 ml M2 trace solution, 4 g l<sup>-1</sup> glucose, 1.5 g l<sup>-1</sup> (<sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, (ISOTEC)). The purification was identical to that of unlabeled-NAPc2, until the TEV-cleavage, which was not performed on labeled NAPc2. Instead, the sample was freeze-dried and dissolved in 0.5 ml, which was loaded onto a Superdex 200 increase (10/300) equilibrated in 20 mM Tris, pH 7.5, 100 mM NaCl. The fractions spanning the main peak were collected and freeze-dried, and resuspended in 300 µl water. Microdialysis was performed against 20 mM Tris, pH 7.5, 100 mM NaCl, 10 mM CaCl<sub>2</sub>.

### Inhibition of FVIIa

The assay was conducted in 50 mM Hepes, pH 7.3, 100 mM NaCl, 10 mM CaCl<sub>2</sub>, 0.1% PEG8000 and 1 mg/ml BSA. 200nM sTF (extracellular domain of TF, called soluble TF, sTF) and 15nm FVIIa (both kindly provided by Novo Nordisk) were mixed with NAPc2 to the desired concentration and incubated for 10 min at RT. S2288 (from Chromogenix) was added to a final concentration of 1 mM, and after thorough mixing the absorbance at 405 nm was read out for 15 min. The slope of the absorbance was determined for the first 10 minutes of the measurement, and the slope as a function of NAPc2 concentration was fitted to a quadratic binding function to obtain the K<sub>i</sub>.

### Inhibition of FVIIa in the presence of FX

10 pM TF incorporated in a nanodiscs with a POPC content of 70% and POPS content of 30% (prepared as described in Paper IV) were mixed with 2.5 nM FVIIa (kindly provided by Novo Nordisk A/S) and a varying NAPc2 concentration and incubated for 10 min on a tilting table at room temperature. 50 nM FX (provided by Novo Nordisk A/S) was added and incubated for 20 min. The reaction was stopped with Quench buffer (50 mM Hepes, pH 7.3, 100 mM NaCl, 80 mM EDTA). 0.5 mM S2765 was added just before placing the plate in the plate reader and

measuring the absorbance at 405 nm for 10 minutes. The slope of the absorbance was plotted as a function of the NAPc2 concentration and fitted to a quadratic binding function to obtain the  $K_i$ .

### Small angle X-ray scattering

The SAXS measurements were performed at the P12 beamline, DESY, Hamburg using the standard solution-SAXS setup (16). The sample was run on SEC (20 mM Tris-HCl, pH 7.5, 150 mM NaCl) just prior to measurements and the peak fraction was selected. The exposure temperature was 10 °C and the intensity of the scattered X-rays was measured as a function of the magnitude of the scattering vector,  $q = 4\pi \sin(\theta) / \lambda$ .  $\theta$  is the scattering angle and  $\lambda$  the wavelength. The data was subsequently reduced using the software available at the instrument, and a background measured on the same buffer as used for the final SEC run was subtracted. The data was converted to absolute scale using a water measurement and normalized by the concentration, to finally obtain the scattering intensity,  $I(q)$ , of the sample. The BayesApp-server (BayesApp.org) was used to perform the indirect Fourier transformation of the  $I(q)$  to provide the pair distribution functions,  $p(r)$ .

### Molecular dynamics and SAXS fitting

All fitting of SAXS data was performed using Pepsi SAXS (17). The MD simulations were performed using the first model in the NMR ensemble (13). All-atom simulations were performed in the presence of 150 mM NaCl for 550 ns.  $R_g$ -values and root mean square fluctuations (RMSF) (which is the deviation between the C $\alpha$  and their average position) were obtained as a function of time and residue number, respectively. After the simulations, the generated ensemble of 3500 frames (taken every 100 ps from the last 350 ns) was post-refined by the use of a Bayesian/Maximum Entropy procedure (18,19). A  $\theta$ -value of 100 was chosen which resulted in the use of 22% of the frames for reweighting of the ensemble.

### NMR

$^1\text{H},^{15}\text{N}$ -HSQC spectra were obtained at 25°C on a Varian INOVA 800-MHz ( $^1\text{H}$ ) spectrometer equipped with a room temperature probe. The  $^{15}\text{N}$ -NAPc2 concentration was 150  $\mu\text{M}$  as measured by UV280 absorbance in triplicate on a Nanodrop 1000 just before experiments. The buffer conditions were: 10% D $_2$ O, 20 mM Tris, pH 7.5, 100 mM NaCl and 1 mM DSS. A FVIIa titration was performed using FVIIa dialyzed against the same buffer as  $^{15}\text{N}$ -NAPc2. A sample with 300  $\mu\text{M}$  FVIIa and 150  $\mu\text{M}$   $^{15}\text{N}$ -NAPc2 was measured initially, and by mixing this sample and the free  $^{15}\text{N}$ -NAPc2 sample the titration points in between were obtained (at FVIIa concentrations of 30, 60, 90, 120 and 150  $\mu\text{M}$ ).

Initial assignments of the  $^1\text{H},^{15}\text{N}$ -HSQC spectra were performed based on the assignments performed by *Duggan et al.* which is deposited in the Biological magnetic resonance bank (BMRB)(13). However, probably due to the very different buffer conditions, the two  $^1\text{H},^{15}\text{N}$ -HSQC spectra were quite different, and thus only 25 of the assignments could reliably be transferred to the newly obtained data. Hence the rest of the data treatment was conducted on these 25 residues. Both chemical shift perturbations and intensity changes (based on the volumes of the peaks) were probed at the different titration points and evaluated with respect to the free  $^{15}\text{N}$ -NAPc2 values.

## Results and discussion

### Production and activity of NAPc2

To be able to perform structural studies on NAPc2, a protocol providing a high yield of pure protein had to be established. Previously NAPc2 has most commonly been expressed in yeast (2,13) but has also been expressed as a fusion protein in *E. coli* (20). To ease the purification, we added a His-tag and a TEV-site to the N-terminal of NAPc2, and performed the expression in *E. Coli* with a subsequent NiNTA purification. To ensure formation of the five disulfide bonds, the expression was carried out in the BL21 Origami cells, which provide an oxidative environment. No NAPc2 was visible in the cellular extract after induction with IPTG (figure 2A, lane 2 and 3). However, the protein clearly became visible after the NiNTA purification, while still containing other impurities (lane 4). The migration of the band changed after TEV cleavage (lane 5), indicating successful removal of the His-tag. The sample was concentrated and applied to a gel filtration column to remove remaining impurities and aggregates. The chromatogram (figure 2B) indicated that the sample contained some aggregates and a dimeric fraction of NAPc2. The fractions marked with blue were visualized in lane 6 (figure 2A), and were used for all subsequent studies. This new protocol yielded 5 mg NAPc2 per liter of bacterial culture and a final purity suitable for structural studies.

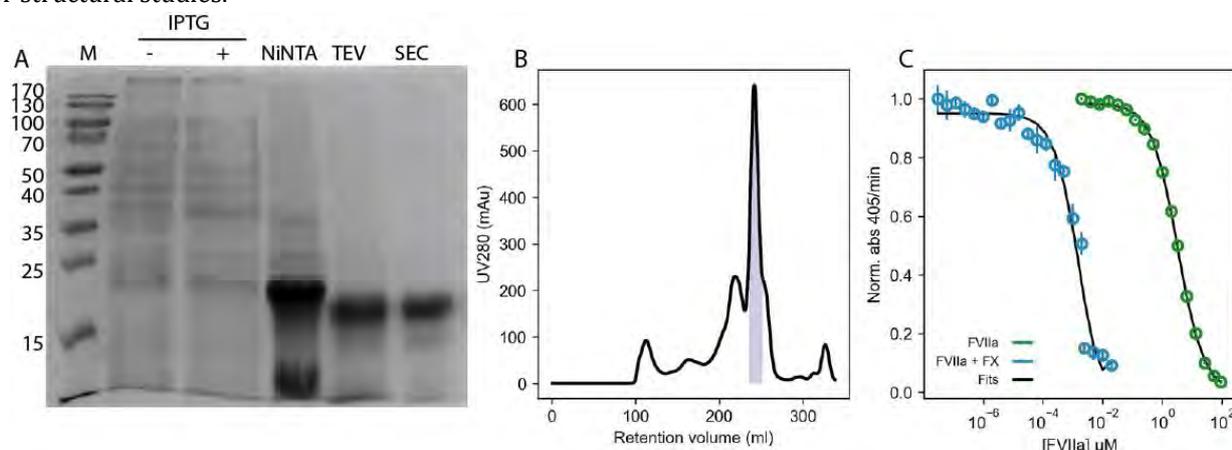


Figure 2: A: SDS-PAGE of the expression and purification of NAPc2. Lane 1: Molecular ladder. Lane 2: Cell lysate before induction. Lane 3: Cell lysate after induction with IPTG. Lane 4: Elution from NiNTA column. Lane 5: After TEV cleavage. Lane 6: After SEC purification. B: Chromatogram from NAPc2 purification on Superdex200 HiLoad. Fractions marked in blue were selected for further work and visualized in lane 6 in A. C: Inhibition of NAPc2 by FVIIa (green) or FVIIa and FX (blue).

To confirm the activity of the NAPc2 produced, using this new protocol, the inhibition of FVIIa in the absence and presence of FX was probed. After formation of the sTF:FVIIa complex, the ability to cleave the chromogenic substrate was measured at different NAPc2 concentrations (green, figure 2C). The  $K_i$  of the interaction was determined to be  $3.2 \pm 0.2 \mu\text{M}$ , which coincides with the one previously determined of  $6.8 \pm 0.7 \mu\text{M}$  (11). The inhibitory efficiency of NAPc2 is markedly increased when FX is present, as FX functions as a carrier of NAPc2 to the TF:FVIIa complex. Therefore, the inhibition of the TF:FVIIa (in a nanodisc (ND)) by NAPc2 in the presence of FX was investigated (figure 2C) and the  $K_i$  was found to be  $2.3 \pm 0.4 \text{ nM}$ . This experiment was conducted using TF incorporated in an ND containing 30 % POPS, which is close to the 25 % used by *P. Bergum* and coworkers, who report a  $K_i$  of  $9.6 \pm 6.7 \text{ pM}$ . The  $K_i$  obtained in this study is closer to the  $K_i$  ( $6.5 \pm 1.4 \text{ nM}$ ) for the sample with 100% PC as reported by *P. Bergum* and colleagues. This difference could have been due to the fact that the previous  $K_i$  was measured using TF in vesicles, while the  $K_i$  reported here was measured in NDs. As the FX:NAPc2 complex is suggested to move around on the lipid surface to dock TF:FVIIa, it could explain the lower  $K_i$ , as the ND surface is minimal compared to a vesicle, and thus the nearby pool of FX:NAPc2 is lower.

### Size and flexibility of NAPc2

To evaluate the overall size and flexibility of NAPc2, SAXS measurements were performed. As an initial step, the 18 models deposited in the NMR ensemble (13) were fitted to the data, but the overall trend was that the models

did not fit the data well with  $\chi^2$ -values ranging from 34 to 99 (figure S1). This was also reflected in the  $R_g$ -values which had an average of  $13.3 \pm 0.5$  Å for the 18 NMR models, while the experimentally determined  $R_g$  was  $15.3 \pm 0.1$  Å, indicating a more extended nature of the SAXS sample. To explore the flexibility of NAPc2 the first model in the NMR ensemble was used for all-atom MD simulations. The simulation was performed for 550 ns, and the superposition of the conformation taken for every 40 ns, demonstrates the flexibility of the NAPc2 (figure 3A). The  $R_g$  was found to increase during the first  $\sim 200$  ns of the simulation, while an equilibrium was reached for the last  $\sim 350$  ns, providing an average  $R_g$  of the last 350 ns of  $14.1 \pm 0.2$  Å (figure S2A). This suggests a less compact structure, compared to the initial NMR ensemble, but still more compact compared to the experimental data. The generated MD ensemble was fitted to the data, and a considerable decrease in  $\chi^2$  was obtained (MD models, purple, figure 3B).

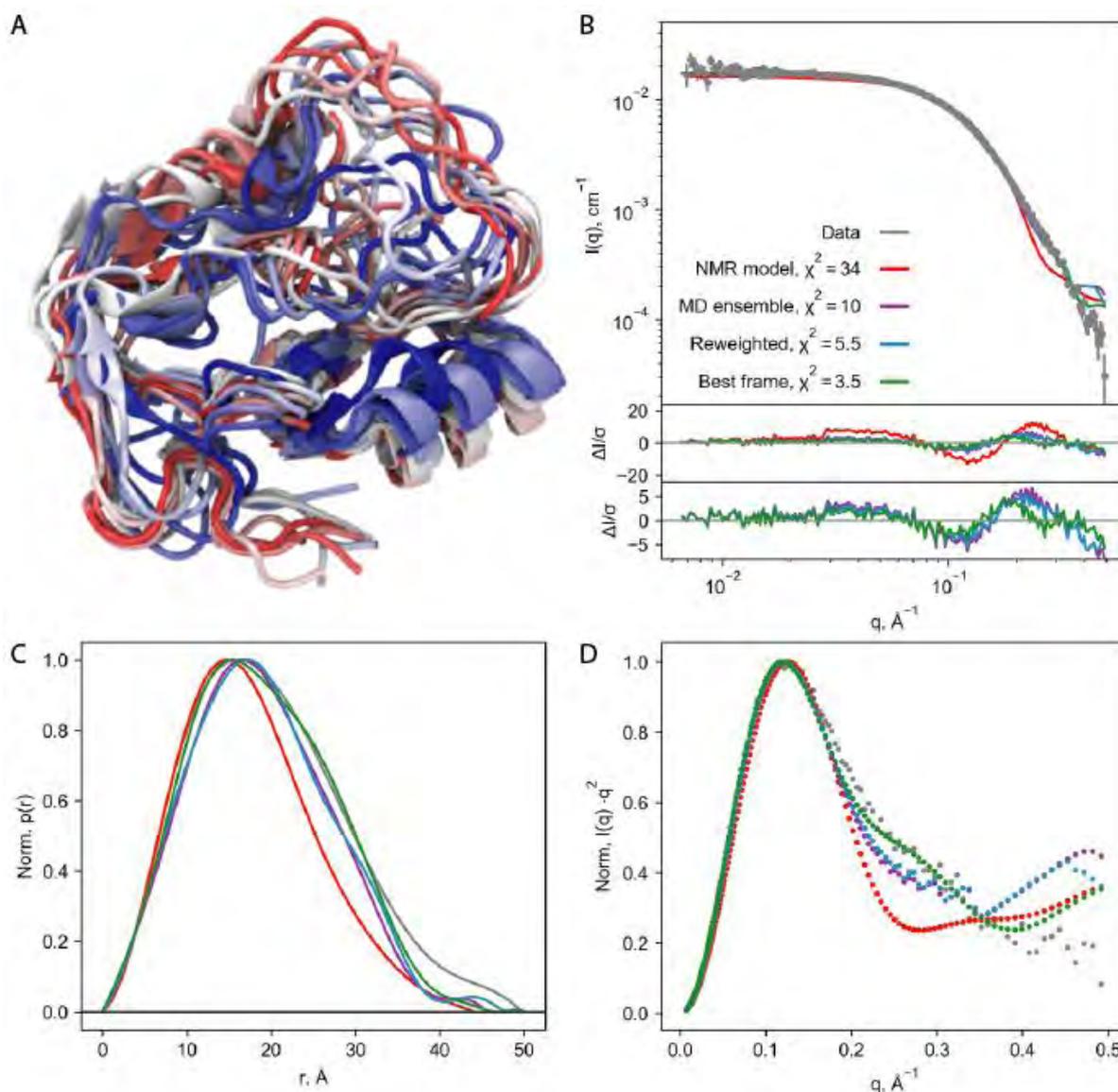


Figure 3: A: Overlaid structures taken every 40 ns of the simulation. B: SAXS data including fitted models. C:  $p(r)$ -functions of the same samples as in A. D: Kratky plot of the same samples as in A.

To further optimize the model, a Bayesian/Maximum Entropy (BME) approach was used to reweigh the ensemble, generating a final ensemble that describes the data (18,19). Selecting 22% of the frames, using this approach, yielded an average  $R_g$  of  $14.3$  Å, which closed the gap to the experimental value even further. The reason for the

continuously smaller  $R_g$ , compared to the one determined from SAXS, could be that NAPc2 is mainly stabilized by salt bridges, which tends to be over-stabilized in MD simulations using current force fields (21).

The different steps were fitted to the experimental data (figure 3B), showing a substantial decrease in the  $\chi_2$ -values from the initial NMR model, to the final best frame of the reweighted models. However, the residuals still revealed that even the best frame failed to reproduce the features at high q-values, which could potentially be explained by the over stabilized salt bridges. The  $p(r)$ -functions (figure 3C) revealed a similar trend, with a large discrepancy between the NMR model and the experimental data, and the best frame approaching the features of the experimental data. The Kratky plot (figure 3D) demonstrated a similar behavior of all five curves, although the NMR structure might look slightly more rigid. Generally, the Kratky plot did not highlight the flexibility of the system, which might be because NAPc2 is restrained globally by the disulfide bridges, and the local flexibility is not visible in the Kratky plot. The root mean square fluctuations (RMSF) on a residue-specific level (figure S2B) revealed that especially the two termini of NAPc2 were very flexible. The FXa binding region is located in the C-terminal, possessing a high degree of flexibility, while the FVIIa binding region is located in a more rigid part of NAPc2.

### Mapping the interaction with FVIIa

In order to examine the interaction between NAPc2 and FVIIa on a residues specific level,  $^{15}\text{N}$ -labeled NAPc2 was produced and a  $^1\text{H},^{15}\text{N}$ -HSQC spectrum was recorded (figure 4A). Many of the peaks were well distributed, but a fraction of the peaks clustered together, suggesting that the protein contains both structured regions and disordered regions. This is in contrast to the  $^1\text{H},^{15}\text{N}$ -HSQC spectra obtained by *Duggan et al.* (13) in which the peaks showed a wider spread, suggesting that their sample preparation introduced a higher degree of rigidity compared to the sample preparation used in the present study.

Due to time limitations, data using  $^{13}\text{C},^{15}\text{N}$ -NAPc2 was not collected, and hence assignments for the protein under these buffer conditions have not yet been performed. Therefore, the assignments from the earlier study (13) were overlaid with the HSQC spectra (figure 4A), and 25 of the assignments could be transferred, while the rest were markedly different and could not be assigned.

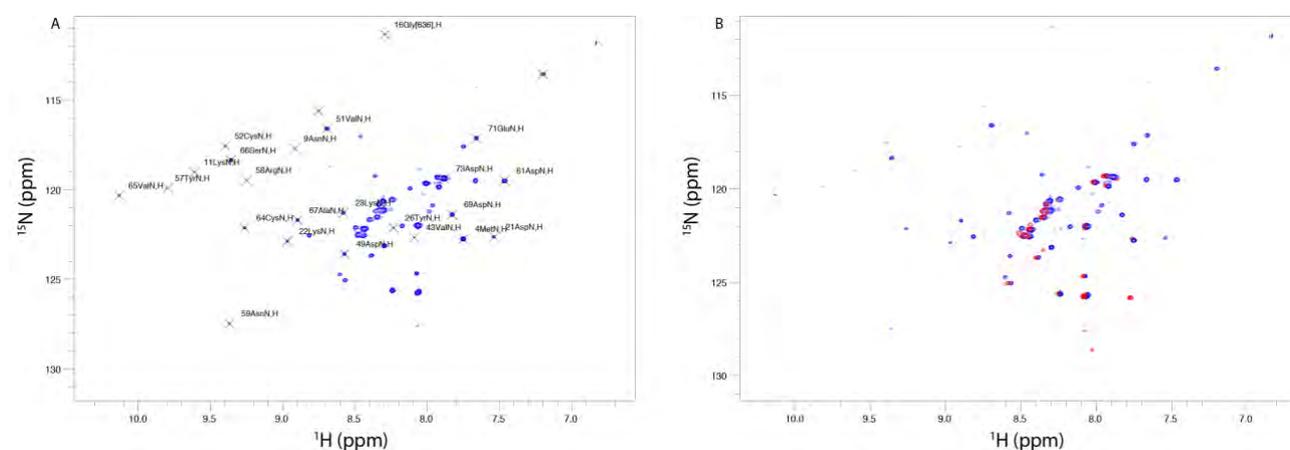


Figure 4: A:  $^{15}\text{N}$ -NAPc2 HSQC with initial assignments obtained from the published 1COU structure. B:  $^{15}\text{N}$ -NAPc2 HSQC (blue) (same as in A). Red: HSQC of  $^{15}\text{N}$ -NAPc2:FVIIa in a ratio of 1:2.

To investigate the FVIIa interaction, a HSQC with a 2:1 ratio of FVIIa:NAPc2 was performed (figure 4B). Evidently, the majority of the peaks now either disappeared or moved, indicating that NAPc2 and FVIIa formed a complex. It is interesting to note that all the most distant peaks disappeared, while the only peaks left were present in what is assumed to be the location of the flexible residues. This indicates that the relatively large size of FVIIa (around 55 kDa) made the tumbling of the FVIIa:NAPc2 complex so slow that the peaks of the rigid parts of NAPc2 disappeared, while the flexible parts on their own, were still present.

To investigate the changes on a residues specific level, the intensity changes for the 25 residues were followed during a titration series, and two representative plots are seen in figure 5. For the 1:1 ratio (figure 5A) a fraction of the peaks was already reduced to 40% of the initial intensity, while another fraction was relatively unaffected and still contained more than 80% of the intensity. For the 2:1 ratio, all residues were reduced to less than half of the initial intensity, and the majority to less than 25%. It seems that while the initial changes were seen locally, the 2:1 ratio seemed to affect the protein globally. However, it should be noted that it was almost only possible to assign the residues that were well spread, which were also the residues taking part in structured elements. This means that these residues were also the residues most likely to be rigid in complex with FVIIa, and hence would tumble too slow to obtain a proper signal.

In conclusion, assignments for the buffer conditions used in the present study must be obtained in order to evaluate the FVIIa interaction by NMR, and will be performed before submission of this manuscript.

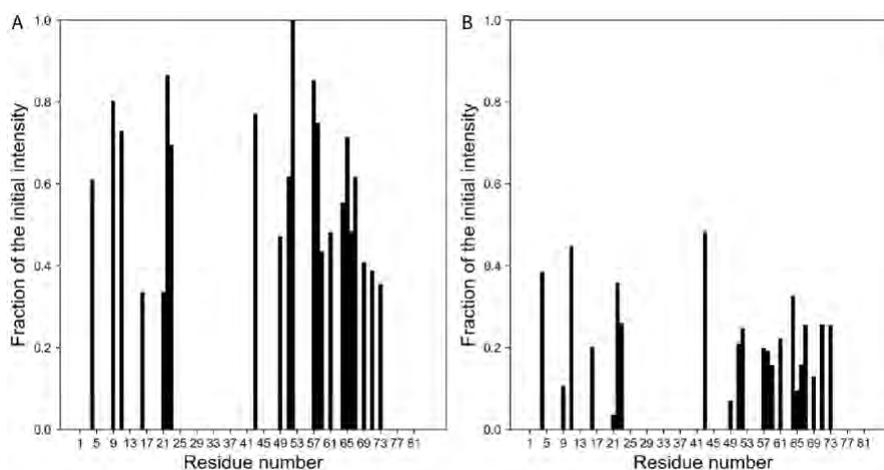


Figure 5: Intensity changes of the assigned residues of  $^{15}\text{N}$ -NAPc2 upon titration with FVIIa compared to free  $^{15}\text{N}$ -NAPc2. A: NAPc2:FVIIa in a ratio of 1:1. B: NAPc2:FVIIa in a ratio of 1:2.

## Conclusions

Flexibility and disorder in proteins are observed at an increasing number of systems, and is being recognized as an important parameter in maintaining cellular functions. As the majority of the protein structures deposited in the protein data bank, are solved using X-ray crystallography, which is unable to describe flexible regions, other methods have to be employed. We have devised a combination of SAXS data and MD simulations to describe the overall size and flexibility of the blood coagulation inhibitor NAPc2. Our SAXS measurements demonstrated a considerably larger size of NAPc2 compared to the NMR structure previously published, which can probably be explained by the very different buffer conditions used in the two experiments. The SAXS measurements were performed at physiological pH, and showed a  $D_{max}$  of about 50 Å compared to the 42 Å for the NMR structure. MD simulations highlighted the most flexible regions in NAPc2, and especially the region conferring the binding interface with FXa showed high flexibility. Lastly, initial NMR experiments were performed in an attempt to map the FVIIa binding interface. However, assignments of the HSQC are lacking and have to be performed, before the data analysis can be finalized.

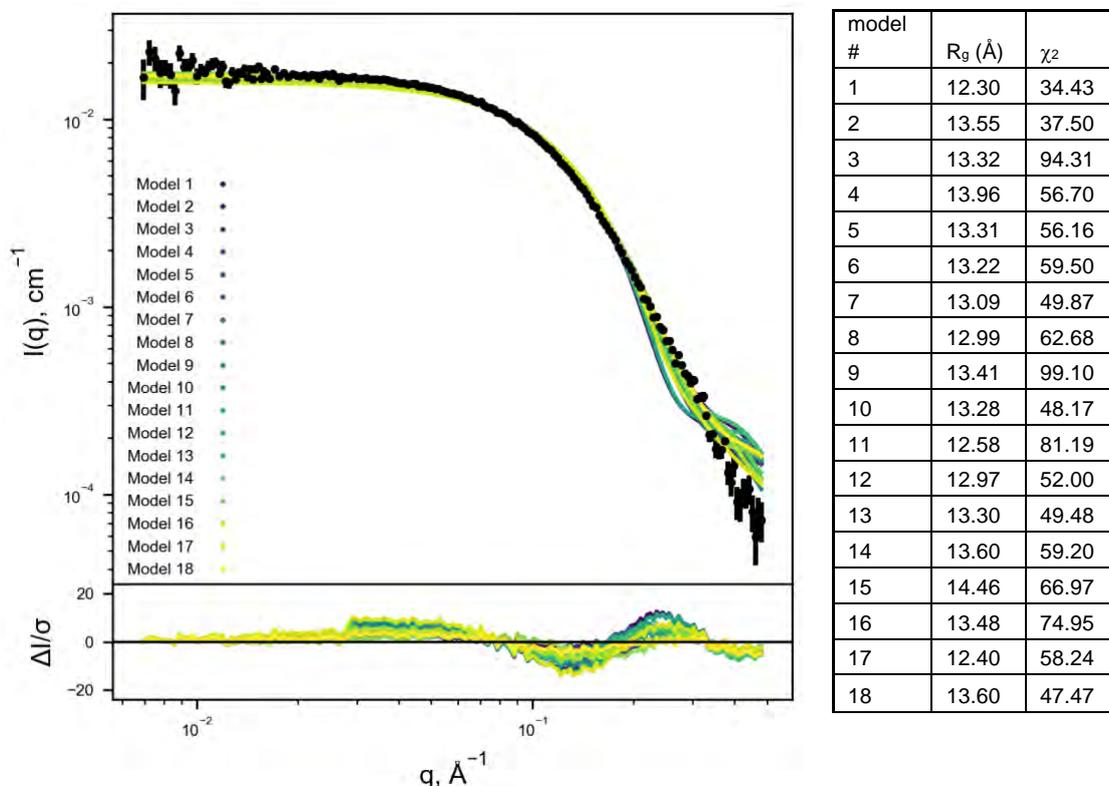
## References

1. Preston RJS, O'Sullivan JM, O'Donnell JS. Advances in understanding the molecular mechanisms of venous thrombosis. *Br J Haematol.* 2019;1–11.
2. Stassens P, Bergum PW, Gansemans Y, Jespers L, Laroche Y, Huang S, et al. Anticoagulant repertoire of the hookworm *Ancylostoma caninum*. *Proc Natl Acad Sci U S A.* 1996;93(5):2149–54.
3. Nemerson Y. The phospholipid requirement of tissue factor in blood coagulation. *J Clin Invest.* 1968;47(1):72–80.
4. Neuenschwander PF, Morrissey JH. Roles of the membrane-interactive regions of factor VIIa and tissue

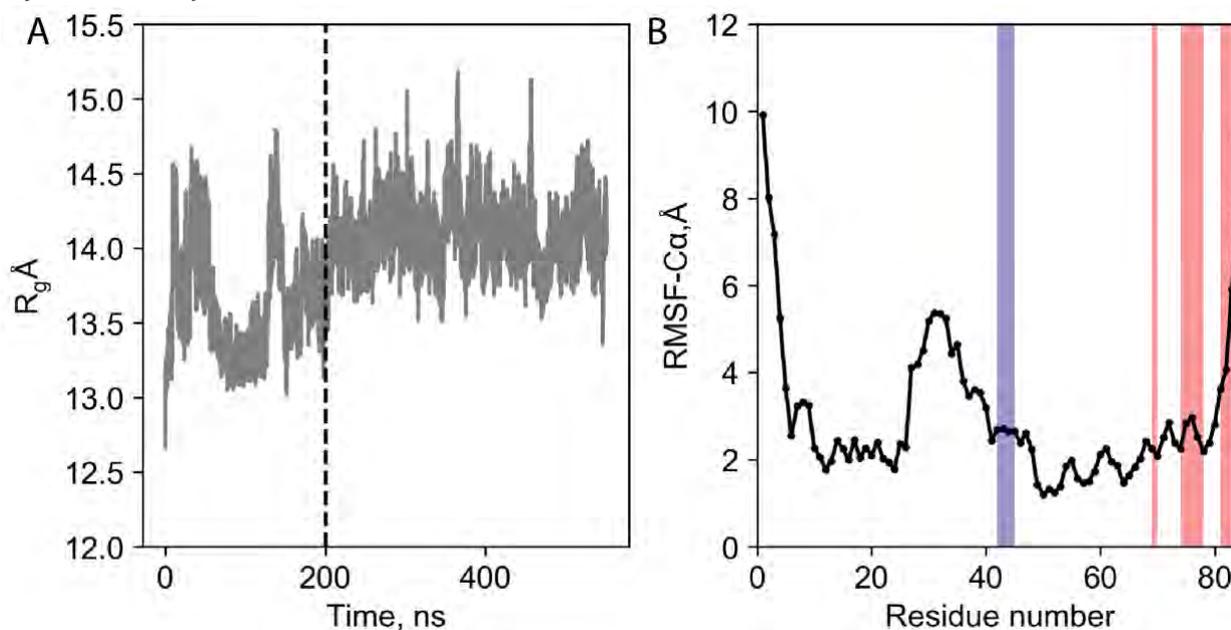
- factor: The factor VIIa Gla domain is dispensable for binding to tissue factor but important for activation of factor X. *J Biol Chem.* 1994;269(11):8007–13.
5. Harlos K, Martin DM, O'Brien DP, Jones EY, Stuart DI, Polikarpov I, et al. Crystal structure of the extracellular region of human tissue factor. *Nature.* 1994;371(6499):720–720.
  6. Banner DW, D'Arcy A, Chene C, Winkler FK, Guha A, Konigsberg WH, et al. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. Vol. 380, *Nature.* 1996. p. 41–6.
  7. Padmanabhan K, Padmanabhan KP, Tulinsky A, Park CH, Bode W, Huber R, et al. Structure of human des(1-45) factor Xa at 2.2 Å resolution. Vol. 232, *Journal of Molecular Biology.* 1993. p. 947–66.
  8. Buddai SK, Touloukhonova L, Bergum PW, Vlasuk GP, Krishnaswamy S. Nematode anticoagulant protein c2 reveals a site on factor Xa that is important for macromolecular substrate binding to human prothrombinase. *J Biol Chem.* 2002;277(29):26689–98.
  9. Shaw AW, Pureza VS, Sligar SG, Morrissey JH. The local phospholipid environment modulates the activation of blood clotting. *J Biol Chem.* 2007;282(9):6556–63.
  10. Tavoosi N, Smith SA, Davis-Harrison RL, Morrissey JH. Factor VII and protein C are phosphatidic acid-binding proteins. *Biochemistry.* 2013;52(33):5545–52.
  11. Bergum PW, Cruikshank A, Maki SL, Kelly CR, Ruf W, Vlasuk GP. Role of Zymogen and Activated Factor X as Scaffolds for the Inhibition of the Blood Coagulation Factor VIIa-Tissue Factor Complex by Recombinant Nematode Anticoagulant Protein c2. *J Biol Chem.* 2001;276(13):10063–71.
  12. Lee AYY, Vlasuk GP. Recombinant nematode anticoagulant protein c2 and other inhibitors targeting blood coagulation factor VIIa/tissue factor. *J Intern Med.* 2003;254(4):313–21.
  13. Duggan BM, Dyson HJ, Wright PE. Inherent flexibility in a potent inhibitor of blood coagulation, recombinant nematode anticoagulant protein c2. Vol. 265, *European Journal of Biochemistry.* 1999. p. 539–48.
  14. Murakami MT, Rios-Steiner J, Weaver SE, Tulinsky A, Geiger JH, Arni RK. Intermolecular Interactions and Characterization of the Novel Factor Xa Exosite Involved in Macromolecular Recognition and Inhibition: Crystal Structure of Human Gla-domainless Factor Xa Complexed with the Anticoagulant Protein NAPc2 from the Hematophagou. *J Mol Biol.* 2007;366(2):602–10.
  15. Fuxreiter M, Simon I, Friedrich P, Tompa P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol.* 2004;338(5):1015–26.
  16. 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;48(2):431–43.
  17. Grudinin S, Garkavenko M, Kazennov A. Pepsi-SAXS: An adaptive method for rapid and accurate computation of small-angle X-ray scattering profiles. *Acta Crystallogr Sect D Struct Biol.* 2017;73(5):449–64.
  18. Ahmed MC, Crehuet R, Lindorff-Larsen K. Analyzing and comparing the radius of gyration and hydrodynamic radius in conformational ensembles of intrinsically disordered proteins. *bioRxiv.* 2019;679373.
  19. Bottaro S, Bengtson T, Lindorff-Larsen K. Integrating Molecular Simulation and Experimental Data: A Bayesian/Maximum Entropy Reweighting Approach. *bioRxiv.* 2018;457952.
  20. Tong Y, Yue J, Mao M, Liu Q, Zhou J, Yang J. Recombinant nematode anticoagulant protein c2 inhibits cell invasion by decreasing uPA expression in NSCLC cells. *Oncol Rep.* 2015;33(4):1815–22.
  21. Ahmed MC, Papaleo E, Lindorff-Larsen K. How well do force fields capture the strength of salt bridges in proteins? *PeerJ.* 2018;6:e4967.

## Supplementary information

**S1:** Fit of all 18 NMR models in the deposited 1COU (13) to SAXS data. Related  $R_g$  and  $\chi^2$  in table. Fitting performed using Pepsi-SAXS (17).



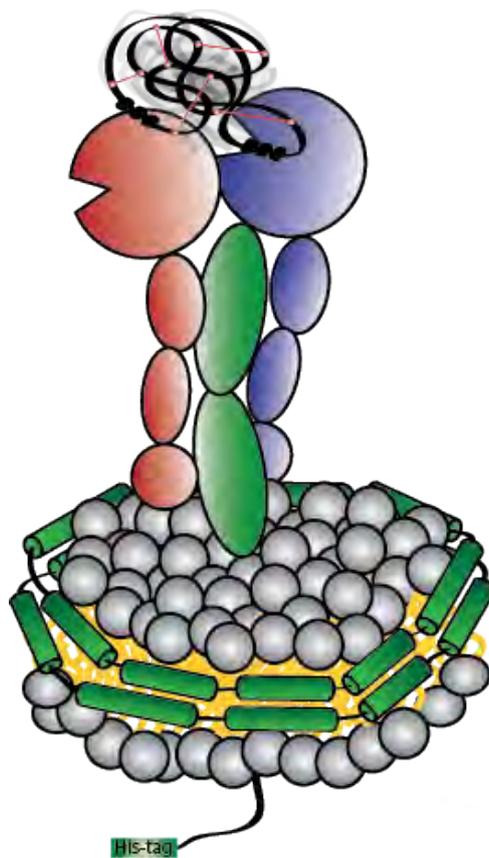
**S2:** Data from MD simulations. A: Temporal evolution of  $R_g$ . Dashed line indicates equilibration time. The frames after equilibration time is used for further analysis. B: Root mean square fluctuations (RMSF) of the  $C\alpha$  as a function of residue number. The higher the RMSF, the higher degree of flexibility. The region proposed to be involved in FVIIa binding is marked with blue, while the region involved in the binding with FXa (from the crystal structure) are marked with red.



# 11 | Report I

Structural investigation of the tissue factor:Factor VIIa:Factor Xa:NAPc2 complex on a lipid bilayer

Frederik Grønæk Tidemand



# Structural investigation of the tissue factor:Factor VIIa:Factor Xa:NAPc2 complex on a lipid bilayer

Frederik Grønbaek Tidemand

Structural Biophysics, X-ray and Neutron Science, The Niels Bohr Institute, University of Copenhagen, Denmark.

## Abstract

High-resolution structures of the individual soluble parts of the tertiary tissue factor (TF):Factor VIIa (FVIIa):Factor Xa (FXa) complex, were solved decades ago, but the structure of the entire complex still remains elusive. TF is an integral membrane protein which, together with the soluble FVIIa, initiates the blood coagulation pathway by forming a complex activating FX to FXa. The structure of the complex between the extracellular domain of TF and FVIIa was solved in the 1990s, but the interaction with FXa is transitory, and highly dependent on the presence of lipids, making the tertiary complex challenging to approach experimentally. In the present study, we employed the inhibitor of the complex, NAPc2, together with a nanodisc (ND) to 'lock' the complex on a lipid bilayer. Using this approach, we could obtain small angle scattering data using both X-rays (SAXS) and neutrons, to confirm the overall sizes of the complex, which increased upon addition of binding partners, as expected. In addition, the individual binding partners were measured using SAXS, and compared to the published structures all cases revealed deviations from the SAXS data. Lastly, a matched-out-deuterated ND (dND) was developed, which was found to be practically invisible to neutrons in a D<sub>2</sub>O based buffer. TF was incorporated in the dND, and the quaternary complex was assembled, enabling SANS data on the complex only, without contributions from the ND. Collectively this data mediates the ongoing modeling of the tertiary complex on a lipid bilayer.

## Introduction

Tissue factor (TF) and the interaction with Factor VII (FVII) initiates the extrinsic blood coagulation cascade by forming a high-affinity complex which after activation of FVII to FVIIa, can activate Factor X (FX) to FXa (1) (figure 1A). This tertiary complex is of paramount importance to the human blood coagulation, and deficiencies in any of the interaction partners lead to hemophilia. FVII is present in the bloodstream while TF is located on the surface of cells adjacent to the vessels. Upon vascular rupture, FVII is able to bind to TF, and subsequently, FX (also present in the blood) will bind the complex (1).

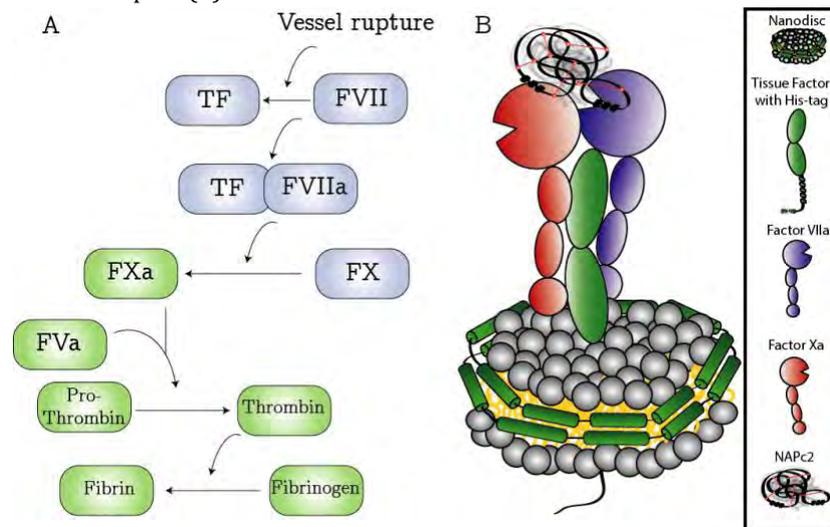


Figure 1: A: Overview of the extrinsic (blue) and common (green) blood coagulation pathway. B: Illustration of the quaternary TF:FVIIa:FXa:NAPc2 complex on a nanodisc.

After activation by the TF:FVIIa complex, FXa participates in the common coagulation pathway by forming the Prothrombinase complex with Factor Va (FVa), activating Prothrombin to Thrombin. Lastly, Thrombin cleaves Fibrinogen to Fibrin which participates in the actual clotting as structural elements (figure 1A) (1). The nature of the TF:FVIIa:FXa complex is multifaceted, as it involves both soluble (FVIIa and FXa), membrane-anchored (TF) proteins, stable (FVIIa + TF) and transient (TF:FVIIa + FXa) interactions and finally, the dependency on the presence of a lipid bilayer. In summary, this makes the tertiary complex challenging to study, especially in terms of structure. However, the investigation of the structures of the human blood clotting pathway has been going on for several decades, and many high-resolution structures of the individual proteins have been published. The structure of the extracellular domain of TF (called sTF), was published in 1994 (2) and two years later, the structure of the complex between sTF and FVIIa was published (3). In addition, an incomplete crystal structure of FXa was published in 2000 (4), but to date, there are still no published experimental data on the tertiary structure, although many studies combining mutagenesis and activity measurements have tried to map the binding interface between FXa and TF:FVIIa (5,6). Both FVIIa and FXa contain a Gla-domain, which is a common feature of the soluble blood clotting proteins (7,8). This domain chelates calcium ions, to obtain a folded state, and mediates the interaction with phospholipids (9). The interplay between the proteins and the lipids is not fully understood, but it is known that the negatively charged phospholipids, Phosphatidylserine (PS), are essential for obtaining a high activity. Also increasing the PS content in the membrane from 10% to 40% was shown to increase the catalytic activity of the TF:FVIIa complex 50 fold (10). In addition, only the interaction with TF, increases the activity of FVIIa by a factor 10s, highlighting the many factors important in this multicomponent system (10).

The main bottleneck in the structural investigation is the need for the presence of lipids, which hampers many of the available structural techniques. This is reflected by the fact that the only published structural data in a lipid environment are two FRET studies measuring the height of the FVIIa active site in different conditions (11,12). Most of the published studies are conducted using TF in vesicles, which is readily obtained using existing protocols (13,14), but for structural studies, other carrier systems are usually preferred. One of these is the nanodisc (ND) which is a self-assembled molecule consisting of a lipid bilayer surrounded by two amphipathic membrane scaffolding proteins (MSPs). The ND is particularly relevant for structural studies because of its relatively small size (compared to a vesicle) and its monodisperse nature. Especially the development of circularized NDs (cNDs) (15) and the further addition of negatively charged residues in the sequence provides an ND (circularized and solubility enhanced, csND) with relatively narrow size distribution and high temporal stability (Paper I). The csNDs provide a foundation for a structural investigation using techniques such as Cryo-electron microscopy and small angle scattering (SAS), as both of these techniques are capable of probing the relatively large size of the system, and also handle the lipid content. Specifically for SAS measurements, which are highly impeded by even a minor fraction of aggregates or other larger particles, the high stability and monodispersity of the csNDs are essential features. As both FVIIa and FXa binds to the lipid membrane through their Gla-domains (9,16), an ND with a reasonably sized bilayer has to be used to assemble the complex. As mentioned, the tertiary complex is only formed transiently with the quick release of FXa, as the biological function is to participate in the Prothrombinase complex after activation. However, NAPc2 has been reported to inhibit FVIIa in an FX dependent mechanism, which effectively locks the quaternary complex (17) (figure 1B). The NAPc2 binding site on FXa is located on an FXa exosite (18), while the binding interface between FVIIa and NAPc2 is still not determined (17). On its own NAPc2 is a 10 kDa protein with a high degree of flexibility, five disulfide bridges restraining the structure, and only a few secondary structure elements (19)(Paper VI).

In the present study, we utilized csNDs and the inhibitory function of NAPc2 to be able to probe the structure of the quaternary TF:FVIIa:FXa:NAPc2 complex on a bilayer. This was done using neutrons and X-rays for SAS measurements, which provided complementary information, which will aid future modeling of the complex.

## Methods

All chemicals and lipids are from Sigma Aldrich unless otherwise stated.

### Expression and purification of proteins

The expression and purification of TF, csE3, and NAPc2 were performed exactly as described elsewhere (Paper IV, (20), Paper V, respectively). The deuterated csE3 (dcsE3) was expressed by the D-lab at ILL in Grenoble (21). The expression and purification were identical to csE3. FVIIa and FXa were kindly provided by Novo Nordisk A/S and no further purification was performed before use.

### Assembly of nanodiscs

The TF containing nanodiscs used in this study were produced exactly as described previously (Paper IV). In brief 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS) were dissolved in cholate buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 100 mM sodium cholate) to a final lipid concentration of 50 mM. All TF containing nanodiscs were assembled with a lipid:csE3 ratio of 120:1, a csE3:TF ratio of 8 and a lipid ratio (POPC:POPS) of 4:1. All preparations had a final lipid concentration of 10 mM. Samples were mixed and incubated for 15 min at room temperature (RT) before detergent absorbing beads (Amberlite XAD-2, Sigma) were added to a ratio of 15% w/v. Samples were incubated at 10°C while shaking at 800 rpm overnight. The samples were applied to a 1 ml HisTrap column (GE Healthcare) equilibrated in 20 mM Tris-HCl pH 8, 100 mM NaCl, and 20 mM and subsequently eluted with 250 mM imidazole. Fractions containing the TF loaded NDs were pooled and diluted 2-fold in MiliQ water and concentrated to 0.5 ml. For SAXS and SANS measurements, the samples were stored on ice until use.

The dND was assembled using the same protocol as described above, but with the use of a novel version of POPC (dPOPC) which has a 94% deuteration of the tail-group and a 71% deuteration in the head-group (22). This lipid was synthesized by the National deuteration facility (NDF) at ANSTO, Australia. As no deuterated POPS was available, the dNDs were prepared with 100% dPOPC.

The empty NDs were all reconstituted with a csE3:lipid ratio of 130 and were applied to the gel filtration column directly after removal of the detergent absorbing beads.

### Small angle scattering

For all SAS data the scattering intensity,  $I(q)$ , was recorded, where  $q = \frac{4\pi \cdot \sin(2\theta)}{\lambda}$ , where  $\lambda$  is the wavelength of the incoming neutrons or X-rays and  $\theta$  is the scattering angle.

All SAXS measurements were performed in Grenoble, France, at the BM29 at ESRF. All experiments were performed such that the sample was passed through a gel filtration column just before it was irradiated with either X-rays or neutrons, as described previously (23,24). For the TF in ND sample with binding partners, the respective binding partners were added to the sample in 3x the concentration of the TF ND together with CaCl<sub>2</sub> to a final concentration of 10 mM before injection onto the column. All experiments were performed using a Superose 6 increase 10/300 GL column (GE Healthcare) equilibrated in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM CaCl<sub>2</sub>. For the SEC-SAXS experiments the flow rate of the column was 0.7 ml/min throughout the experiment.

The SANS measurements on the TF in ND and the TF in ND with binding partners were performed at D22, ILL, Grenoble. The measurements were performed in the SEC-SANS setup using a Superose 6 increase 10/300 GL column (GE Healthcare) equilibrated in a 100% D<sub>2</sub>O buffer 20 mM Tris-DCl, pH 7.5, 150 mM NaCl, 10 mM CaCl<sub>2</sub>. The initial flow was 0.4 ml/min and when the peak of interest eluted, the flow was decreased to 0.1 ml/min and then increased to 0.4 ml/min after the peak had eluted. For both the SEC-SAXS and SEC-SANS experiments, a background was chosen from the regions before the peaks, and frames from the peak were averaged and subtracted this background. The concentration of the selected peaks was estimated using the UV cell attached to the instruments, and the data were normalized by this concentration. In the SANS experiment the sample was run twice in order to have two different sample-to-detector distances (2 m and 11.2 m) to cover a larger q-range.

The SANS measurement on the empty ND was performed at SANS 2, FRMII, Munich. The sample was purified using SEC (on a Superdex 200 increase 10/300 GL, equilibrated in 100% D2O buffer 20 mM Tris-DCl, pH 7.5, 150 mM NaCl, 10 mM CaCl<sub>2</sub>) before measurement, and a peak fraction was put in a cuvette for the measurement.

The dND and the corresponding hydrogenated ND (hND) samples were measured at D22, ILL, Grenoble, and were gel filtrated on a Superose 6 increase 10/300 offline in a 100% D2O buffer (20 mM Tris-DCl, pH 7.5, 150 mM NaCl, 10 mM CaCl<sub>2</sub>). A fraction from the main peak was selected and measured in a normal cuvette SANS setup at sample-to-detector distance of 2 m and 6 m. The concentrations were 2.3  $\mu$ M and 2.1  $\mu$ M for the hND and the dND samples, respectively. The buffer used for the gelfiltration was measured in a separate cuvette in the same settings to obtain a background measurement.

## Results and discussion

### Assembly of the complex on a ND

In a previous study (Paper IV) we showed that we were able to produce TF incorporated circularized nanodiscs in the amounts needed for structural studies. To further explore the utility of this, we tried to assemble the tertiary TF:FVIIa:FXa complex on an ND. In order to obtain an SAS dataset only describing the complex of interest, the experiments need to be carried out as an SEC-SAS experiment to get rid of excess binding partners. An SEC-SAS experiment sets a certain requirement for the binding kinetics within the complex, as excess binding partners are continuously removed when the sample passes through the column. We tested the binding of different binding partners on a TF loaded ND (figure 2), and it was evident that when adding FVIIa and FXa (blue) to the TF ND, the size of the particle in the main peak did not increase compared to only adding FVIIa (green). Consequently, we employed an inhibitor of the complex, NAPc2, which has been shown to bind both FVIIa and FXa and in this study it was used to stabilize the quaternary complex. As shown by the purple curve, the main peak eluted slightly earlier indicating that the quaternary complex was stable on an ND during an SEC run, which was required in order to obtain high quality SAS data.

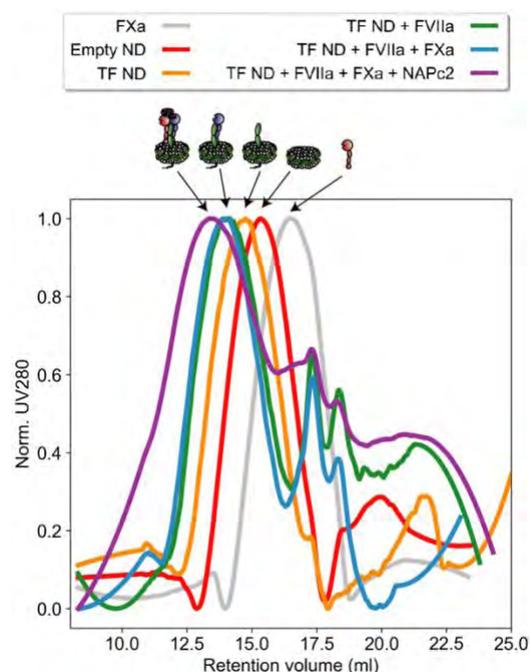


Figure 2: SEC chromatograms of the assembly of the quaternary complex TF:FVIIa:FXa:NAPc2 on a nanodisc.

### Probing the structure of the individual interaction partners

Before initializing the investigation of the entire complex, the solution structures of the individual parts of the complex were probed using SAXS. The data on sTF (figure 3A) appeared to agree with the overall features of the crystal structure but showed deviations especially at higher  $q$ -values. This was expected, as the crystal structure is lacking a few residues, furthermore it is possible that the solution structure is different from the crystal structure e.g. due to crystal packing effects. Adding the missing residues and performing a short simulation to relax the structure, will most likely fix this issue. Looking at the FVIIa data and fit (figure 3B), the model presented by *Norledge et al.* (25) seemed to capture the overall features, but substantial deviations were also present. First of all, the model used for fitting was based on a crystal structure lacking several residues, which have been constructed by homology modeling. However, the large deviations seen around  $0.1 \text{ \AA}^{-1}$  was likely in part due to the lack of glycosylations of the FVIIa model. FVIIa contains four glycosylations (26) which are not present in any published model, but they will indeed contribute to the scattering patterns. The flexible nature of these glycosylations makes them challenging to study using X-ray crystallography and electron microscopy, but a combination of MD simulations and SAXS data have proved valuable in describing the structural features of glycosylations, and such a strategy should be pursued for the present analysis (27,28).

The sTF:FVIIa complex was also investigated (figure 3C), and the crystal structure (pdb-ID: 1DAN) was fitted to the data. As expected, the fit was yet again hampered by incomplete structures and the lack of glycosylations, which was particularly visible at high  $q$ -values, but the overall features of the complex were well represented. Lastly, SAXS data on NAPc2 had previously been shown to deviate significantly from the published NMR structure (Paper VI) (teal, figure 4D). However, in a previous study (Paper VI) using MD simulations, we were able to refine a model that better described the flexibility and overall size observed by the SAXS data. This study resulted in the final model presented in figure 4D, with the corresponding fit (black, figure 4D) describing the SAXS data considerably better than the NMR structure ( $\chi^2=3.5$  compared to  $\chi^2=35$  for the NMR model). There are slight deviations present in the fit at high  $q$ -values, which are likely due to overstabilization by the simulation of the many salt bridges in NAPc2.

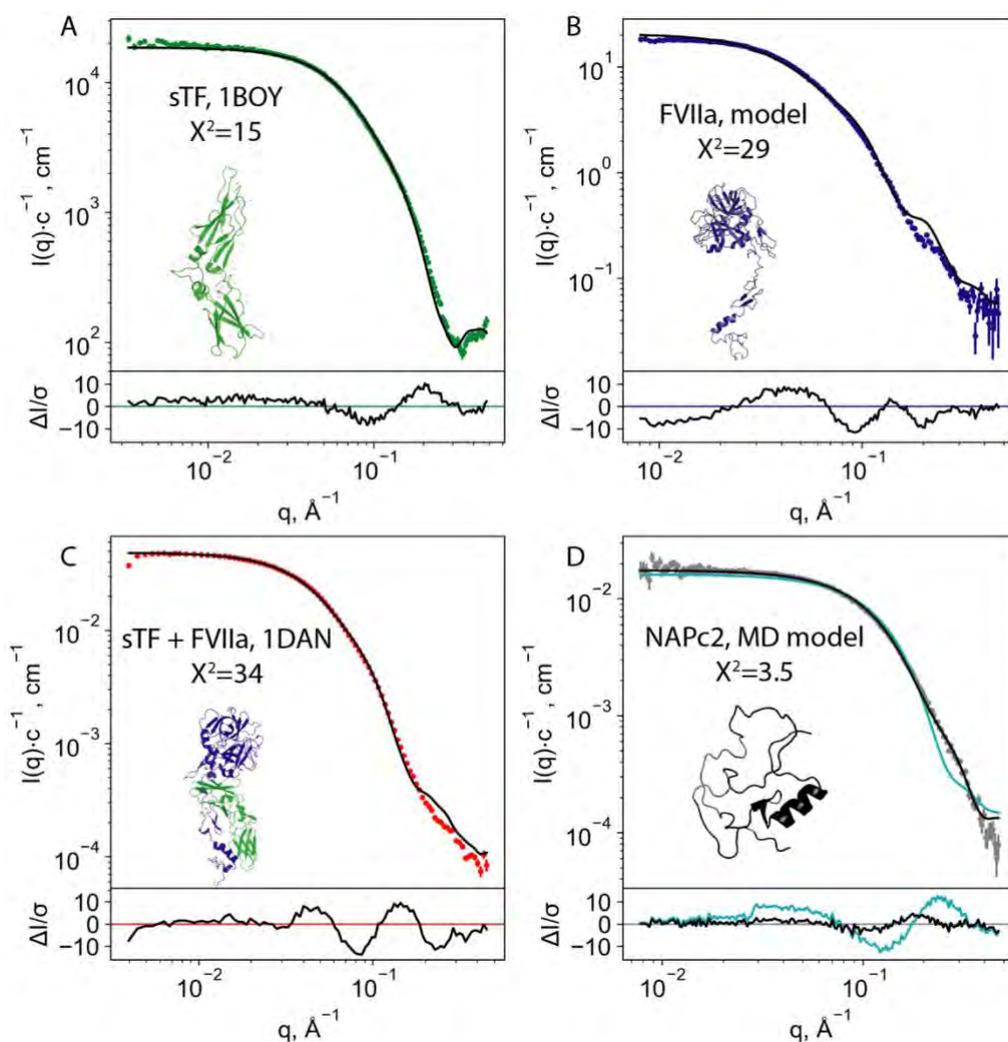


Figure 3: SAXS measurements of individual proteins from the complex. Chi2-values from the respective fits and fitted model as inserts. A: sTF. Crystal structure (pdb-ID: 1BOY) fitted. B: FVIIa. FVIIa from the model presented in (25) fitted. C: sTF:FVIIa complex. Crystal structure (pdb-ID: 1DAN) fitted. D: NAPc2. NMR structure (pdb-ID: 1COU) (teal) and MD model (Paper VI)(black) fitted (data and model from Paper VI).

### Structural investigation of the complex formation on a nanodisc

The structure of the TF complex in a nanodisc was investigated using both SAXS and SANS. For both techniques, four samples were measured (top panel, figure 4): Empty ND, TF incorporated ND, TF incorporated ND + FVIIa, and TF incorporated ND + FVIIa + FXa + NAPc2. SEC-SAXS measurements were conducted on all four samples, providing information about the  $R_g$ -values over the peak of interest (figure S1). As evident from all four figures,

the  $R_g$ -values decreased substantially over the peak, and a difference of more than 20 Å was observed for some of the samples. This is commonly seen for NDs measured using SEC-SAXS (24), and reflects the dynamic nature of the discs. A previous study, showed a change in both the number of lipids and the axis ratio of the NDs over the peak (29). For the empty ND, the TF ND and the TF ND + FVIIa samples, a slight plateau was observed, and frames spanning this plateau were selected for further analysis. For the ND with the entire complex assembled, the main peak seemed to have a small ‘shoulder’ on the right-hand side, which could represent a minor fraction of NDs with only TF and FVIIa bound, and thus frames in the left-hand side of the peak were selected.

The resulting SAXS data (figure 4A) showed an increase in  $I(0)$  upon addition of binding partners, which was expected, as the molecular weight of the complex increased. Both the empty ND and the TF-containing-ND had an oscillation around 0.1 Å<sup>-1</sup>, which is line with previously reported observations for both the ND and this system ((30,31) and Paper IV). When FVIIa was also bound, this feature disappeared, while when the entire complex was assembled, a new feature emerged in the same q-region. However, the increase in  $I(0)$  was larger than expected upon addition of FXa and NAPc2, but as the concentration estimates from the SEC-SAXS measurements are related with quite substantial errors, this might be the reason for this increase.

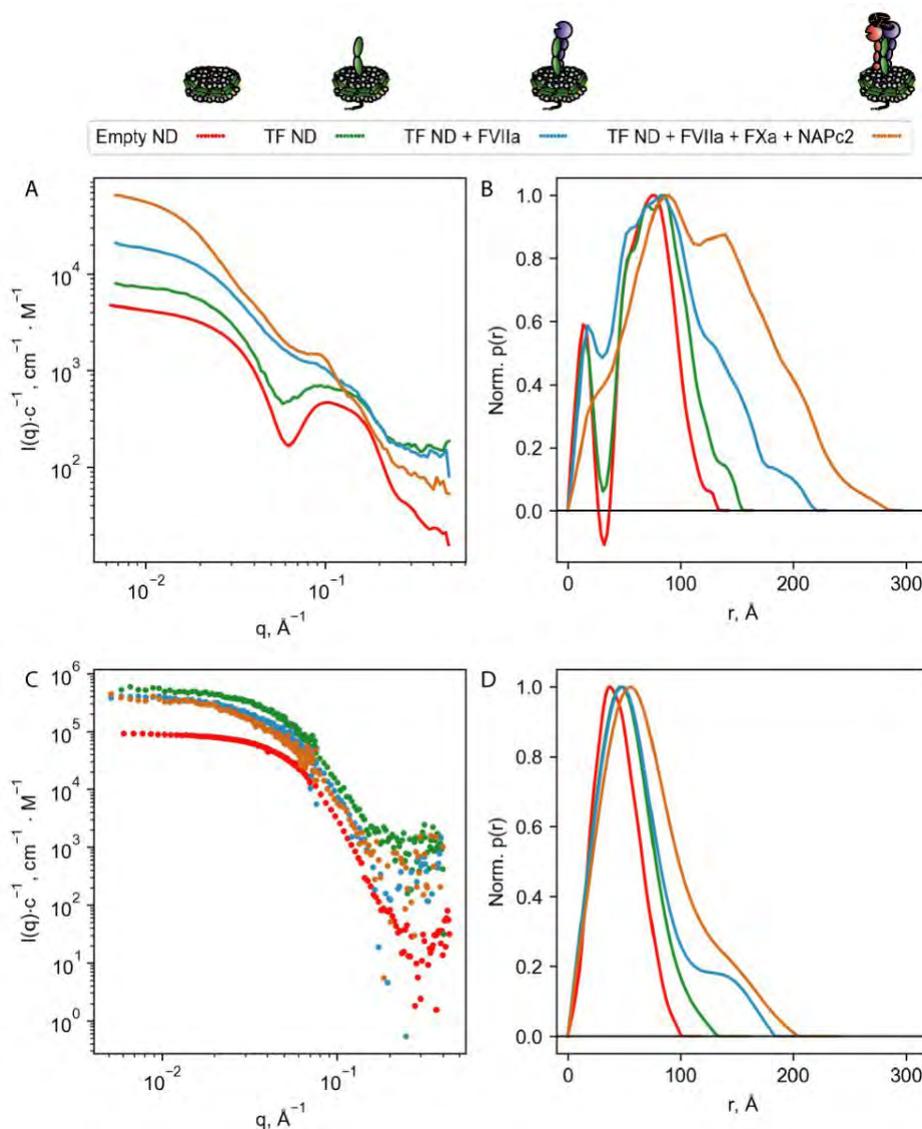


Figure 4: SAS data on four different samples: Empty ND (red), TF containing ND (green), TF containing ND + FVIIa (blue) and TF containing ND + FVIIa + FXa + NAPc2. A: SAXS data on absolute scale and normalized by concentration. B:  $p(r)$ -functions of the data presented in A. C: SANS data on absolute scale and concentration normalized. D:  $p(r)$ -functions of the data presented in C.

The  $p(r)$ -functions (figure 4B) showed the same overall trends, as the  $D_{max}$  increased for every binding partner added to the complex, and the downward trend around 30 Å almost disappeared when FVIIa was bound. SANS measurements were also performed to support the SAXS data. The empty ND was purified using SEC (figure S2A), and static SANS was measured on the selected fraction (figure 4C). The other three samples were measured using SEC-SANS, in which the flowrate had to be lowered during the run to obtain useful statistics on the final scattergrams. Unfortunately, this resulted in chromatograms with broad peaks (as they were only available as a function of time, figure S2B-D), which were difficult to assess. However, as the SAXS and SANS samples were from the same batch, we tried to select frames from the same place in the peaks, as we did for the SAXS measurements. This resulted in SANS data (figure 4C) which looked quite similar, and especially the data with TF in NDs also showed roughly the same  $I(0)$ . This is probably also due to difficulties in estimation of the concentrations, as an increase in  $I(0)$  was expected. When investigating the  $p(r)$ -functions (figure 4D), it was once again evident that the  $D_{max}$  values increased, as more binding partners were added to the complex. The  $D_{max}$  values obtained from SANS measurements were considerably smaller than those obtained by SAXS, but this is a commonly seen feature (31). The difference relies on the different contrast situations, as the lipid tails have the largest contrast in SANS, while the MSPs have the largest contrast in SAXS. This means that the interior of the ND scatters the most in SANS, while it is the periphery in SAXS. When calculating  $p(r)$ -functions, the curve is 'pushed' towards 0 by the algorithm. This means that if the signal at large distances is low, it will be forced to zero and hence a low  $D_{max}$  will be obtained, as in this case for SANS. As the large distances in SAXS scatter more, the curve will not be forced to zero as fast, and hence a larger  $D_{max}$  is obtained.

### Invisible nanodiscs to investigate the complex

To be able to investigate the structure of the complex, without contribution from the ND, a so-called 'matched-out' deuterated ND (dND) was developed. The dND was composed of the same constituents as the NDs used for the other experiments, except that both lipids and MSPs were partially deuterated. The partial deuteration of the lipids was obtained by synthesis, while the MSPs were expressed in *E. Coli* grown in a D<sub>2</sub>O containing media. To investigate the dND in a SANS context, a sample was made in buffers with different D<sub>2</sub>O contents, and the background subtracted data (figure 5A) showed a decreasing tendency of the  $I(0)$ . To evaluate the exact match point of the dND, the square root of  $i(0) \cdot c$  as a function of D<sub>2</sub>O content was plotted (figure 5B). The data showed a linear relationship, and the intersect with the x-axis provided the match-point, 101% D<sub>2</sub>O, which means that the ND was practically invisible in 100% D<sub>2</sub>O. To confirm this, the raw data from the ND at 100 % D<sub>2</sub>O and the corresponding background were plotted (figure 5C) demonstrating that the ND data was almost within the noise of the background.

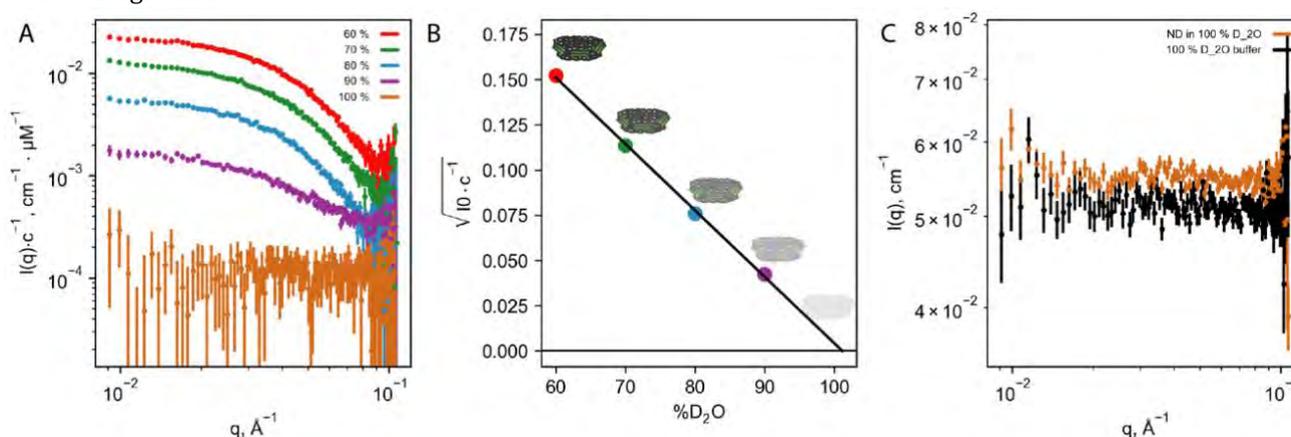


Figure 5: Match-out deuterated nanodisc. A: SANS measurements of the nanodisc in buffers with varying D<sub>2</sub>O content. B: The square root of the  $I(0)$ -values obtained from A times the concentration as a function of the D<sub>2</sub>O content shows a linear relationship. C: Raw data of the nanodisc in 100% D<sub>2</sub>O and the corresponding buffer.

TF was incorporated, and the quaternary complex was assembled on the novel dND, and a SEC chromatogram (figure 6A, red) showed the same elution volume as the corresponding normal ND (hydrogenated, hND) also with the complex assembled (figure 6A, blue). The resulting SANS measurements (figure 6B) showed a flat Guinier region for the hND, while the dND seemed to ascent even at low  $q$ -values. This was also reflected in the  $p(r)$ -distributions (figure 6C), as the hND had approximately the same  $D_{max}$  as seen for the SEC-SANS data (figure 4D). On the other hand, the dND had an even larger  $D_{max}$ , which was actually in the range of the  $D_{max}$  obtained from the SAXS measurements. This was unexpected, as the dND sample, in theory, should have a smaller or maximally equally large  $D_{max}$  as the hND sample. However, it should be mentioned that the concentration of the dND sample was not optimal (identical to the hND, but with much less scattering due to the contrast), and preferably, this sample should be remeasured at a higher concentration. The fact that no scattering was obtained from the dND was also seen in the markedly lower  $I(0)$  of the dND + complex sample, compared to the hND + complex sample.

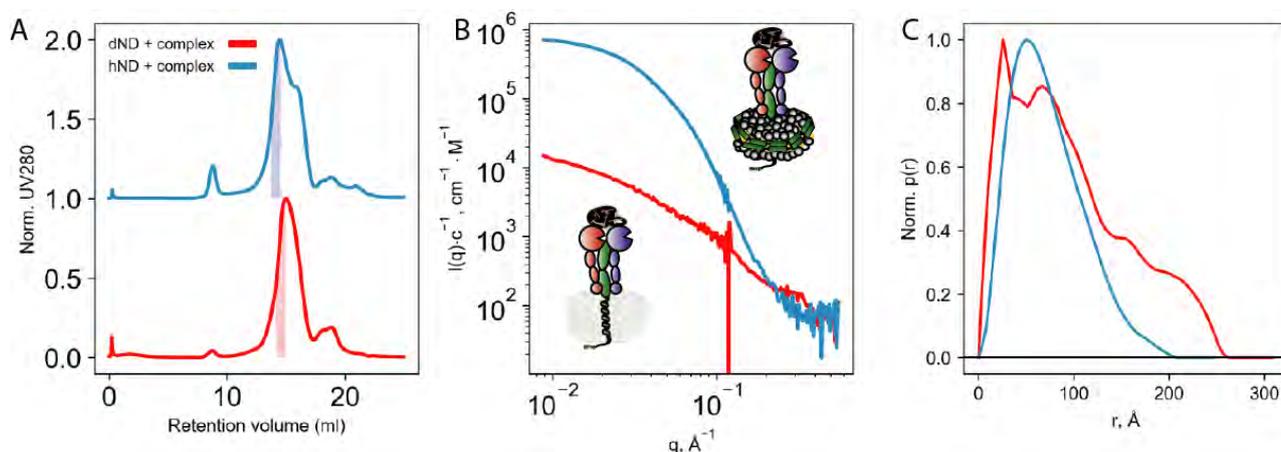


Figure 6: Comparison of the quaternary complex TF:FVIIa:FXa:NAPc2 assembled on a dND and a regular ND. A: SEC-chromatograms of the two different samples. Marked areas are the fractions selected for SANS measurements. B: SANS data on absolute scale and concentration normalized. C:  $p(r)$ -functions obtained using the Bayesapp server.

## Conclusions and outlook

In the present study, we assembled the quaternary TF:FVIIa:FXa:NAPc2 complex on an ND. It was evident from analysis by SEC that NAPc2 was necessary to form a stable complex between TF, FVIIa, and FXa. After SAXS measurements of the soluble separate parts of the complex, the individual steps in the assembly of the complex were measured using both SAXS and SANS. Both datasets showed the expected increase in  $D_{max}$ , suggesting enlargement of the particle for each addition of new binding partners. Lastly, the matched out deuterated NDs were utilized to obtain SANS data on the complex alone. Unfortunately, these data were hampered by a low concentration, but the method proved to be applicable for this system.

The next step in this study is modeling of the individual proteins to fit the SAXS data, before assembling these parts to fit the complex on a bilayer. The previous work on NAPc2 clearly showed how MD simulations could improve a model to fit the SAXS data, and similar simulations are to be performed on both TF, FVIIa, and FXa. In addition, glycosylations need to be attached to FVIIa to represent the solution structure. On the data side, SAXS measurements of FXa needs to be performed to be able to refine the model of FXa before assembling the complex. Furthermore, the SANS measurements on the dND data are to be remeasured, perhaps using a fraction from the middle of the peak, as the concentration needs to be much higher, because there is no scattering contribution from the lipids in the sample. In summary, the present study lays the foundation for the first experimentally supported model of the tertiary TF:FVIIa:FXa model in a lipid context, which will hopefully provide a deeper understanding of the initial steps of the blood coagulation pathway in humans.

## Acknowledgments

Tamim Darwish and Rao Yepuri (both from NDF, ANSTO, Sydney) are acknowledged for their synthesis of the deuterated POPC. The D-lab in Grenoble is acknowledged for the growth of deuterated MSP. The SANS experiments at D22, ILL were performed under great guidance from Anne Martel and Lionel Porcar, while the SAXS measurements were performed with help from Mark Tulley (BM29, ESRF). A big thank you to all three. The Novo Nordisk Foundation is acknowledged for the funding of the project, and DanScatt is acknowledged for the financial support to go to the many beamtimes.

## References

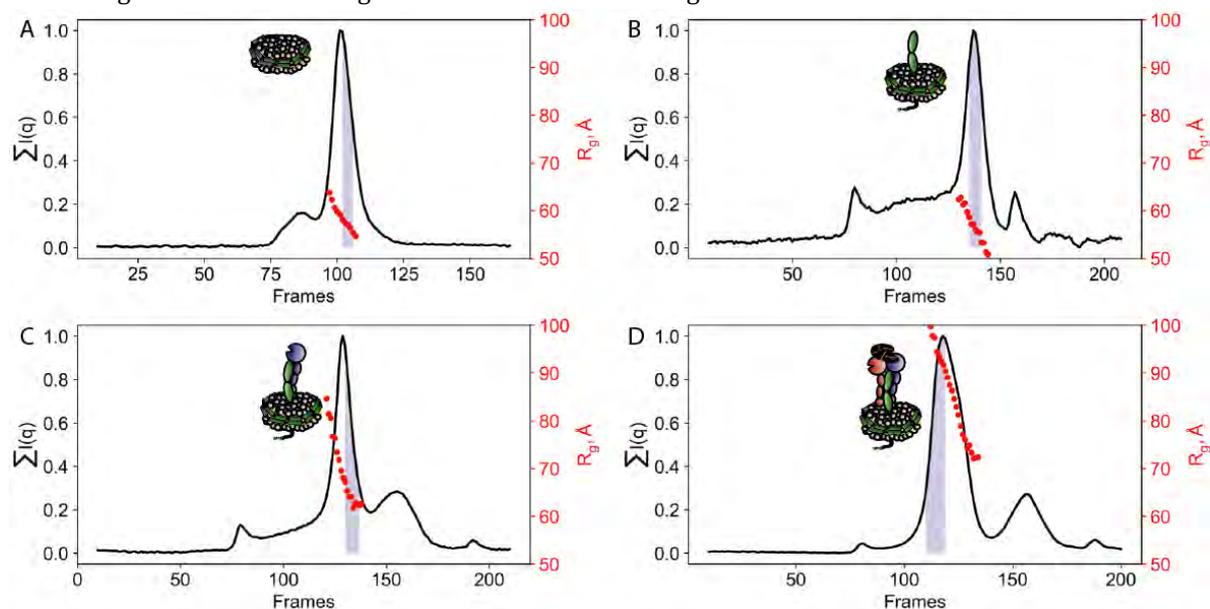
1. Furiel B, Furie BC. The Molecular Basis of Blood Diseases. *Cell*. 1988;53:505–18.
2. Harlos K, Martin DM, O'Brien DP, Jones EY, Stuart DI, Polikarpov I, et al. Crystal structure of the extracellular region of human tissue factor. *Nature*. 1994;371(6499):720–720.
3. Banner DW, D'Arcy A, Chene C, Winkler FK, Guha A, Konigsberg WH, et al. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. Vol. 380, *Nature*. 1996. p. 41–6.
4. Maignan S, Guilloteau JP, Pouzieux S, Choi-Sledeski YM, Becker MR, Klein SI, et al. Crystal structures of human factor Xa complexed with potent inhibitors. *J Med Chem*. 2000;43(17):3226–32.
5. Carlsson K, Freskgård PO, Persson E, Carlsson U, Svensson M. Probing the interface between factor Xa and tissue factor in the quaternary complex tissue factor-factor VIIa-factor Xa-tissue factor pathway inhibitor. *Eur J Biochem*. 2003;270(12):2576–82.
6. Ke K, Yuan J, Morrissey JH. Tissue factor residues that putatively interact with membrane phospholipids. *PLoS One*. 2014;9(2).
7. Huang M, Rigby AC, Morelli X, Grant MA, Huang G, Furiel B, et al. Structural basis of membrane binding by Gla domains of vitamin K-dependent proteins. *Nat Struct Biol*. 2003;10(9):751–6.
8. Ohkubo YZ, Tajkhorshid E. Distinct Structural and Adhesive Roles of Ca<sup>2+</sup> in Membrane Binding of Blood Coagulation Factors. *Structure*. 2008;16(1):72–81.
9. Tavoosi N, Smith SA, Davis-Harrison RL, Morrissey JH. Factor VII and protein C are phosphatidic acid-binding proteins. *Biochemistry*. 2013;52(33):5545–52.
10. Shaw AW, Pureza VS, Sligar SG, Morrissey JH. The local phospholipid environment modulates the activation of blood clotting. *J Biol Chem*. 2007;282(9):6556–63.
11. McCallum CD, Hapak RC, Neuenschwander PF, Morrissey JH, Johnson AE. The location of the active site of blood coagulation factor VIIa above the membrane surface and its reorientation upon association with tissue factor - A fluorescence energy transfer study. *J Biol Chem*. 1996;271(45):28168–75.
12. McCallum CD, Su B, Neuenschwander PF, Morrissey JH, Johnson AE. Tissue factor positions and maintains the factor VIIa active site far above the membrane surface even in the absence of the factor VIIa Gla domain. A fluorescence resonance energy transfer study. *J Biol Chem*. 1997;272(48):30160–6.
13. Neuenschwander PF, Fiore MM, Morrissey JH. Factor VII autoactivation proceeds via interaction of distinct protease-cofactor and zymogen-cofactor complexes: Implications of a two-dimensional enzyme kinetic mechanism. *J Biol Chem*. 1993;268(29):21489–92.
14. Mimms LT, Zampighi G, Nozaki Y, Tanford C, Reynolds JA. Phospholipid Vesicle Formation and Transmembrane Protein Incorporation Using Octyl Glucoside. *Biochemistry*. 1981;20(4):833–40.
15. 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*. 2016;14(1):49–52.
16. Muller MP, Wang Y, Morrissey JH, Tajkhorshid E. Lipid Specificity of the Membrane Binding Domain of Coagulation Factor X. *J Thromb Haemost*. 2017;15(10):2005–16.
17. Bergum PW, Cruikshank A, Maki SL, Kelly CR, Ruf W, Vlasuk GP. Role of Zymogen and Activated Factor X as Scaffolds for the Inhibition of the Blood Coagulation Factor VIIa-Tissue Factor Complex by Recombinant Nematode Anticoagulant Protein c2. *J Biol Chem*. 2001;276(13):10063–71.
18. Murakami MT, Rios-Steiner J, Weaver SE, Tulinsky A, Geiger JH, Arni RK. Intermolecular Interactions and Characterization of the Novel Factor Xa Exosite Involved in Macromolecular Recognition and Inhibition: Crystal Structure of Human Gla-domainless Factor Xa Complexed with the Anticoagulant Protein NAPc2 from the Hematophagou. *J Mol Biol*. 2007;366(2):602–10.
19. Duggan BM, Dyson HJ, Wright PE. Inherent flexibility in a potent inhibitor of blood coagulation, recombinant nematode anticoagulant protein c2. Vol. 265, *European Journal of Biochemistry*. 1999. p. 539–48.
20. Johansen NT, Pedersen MC, Tidemand FG, Rand KD, Nguyen TTTN, Arleth L. Circularized and solubility-

enhanced MSPs facilitate simple and high yield production of stable nanodiscs for studies of membrane proteins in solution. *FEBS J.* 2019;

21. Dunne O, Weidenhaupt M, Callow P, Martel A, Moulin M, Perkins SJ, 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.* 2017;46(5):425–32.
22. Yepuri NR, Darwish TA, Krause-Heuer AM, Leung AE, Delhom R, Wacklin HP, et al. 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(3):315–21.
23. Pernot P, Round A, Barrett R, De Maria Antolinos A, Gobbo A, Gordon E, et al. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J Synchrotron Radiat.* 2013;20(4):660–4.
24. 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 [Internet].* 2018 Dec;74(12):1178–91. Available from: <https://doi.org/10.1107/S2059798318007180>
25. Norledge B V., Petrovan RJ, Ruf W, Olson AJ. The Tissue Factor/Factor VIIa/Factor Xa Complex: A Model Built by Docking and Site-Directed Mutagenesis. *Proteins Struct Funct Genet.* 2003;53(3):640–8.
26. Fenaille F, Groseil C, Ramon C, Riandé S, Siret L, Chtourou S, et al. Mass spectrometric characterization of N- and O-glycans of plasma-derived coagulation factor VII. *Glycoconj J.* 2008;25(9):827–42.
27. Guttman M, Weinkam P, Sali A, Lee KK. All-atom ensemble modeling to analyze small angle X-ray scattering of glycosylated proteins. *Structure.* 2013;21(3).
28. Awad W, Adamczyk B, Örnros J, Karlsson NG, Mani K, Logan DT. Structural aspects of N-glycosylations and the C-terminal region in human glypican-1. *J Biol Chem.* 2015;290(38):22991–3008.
29. Skar-Gislinge N, Johansen NT, Høiberg-Nielsen R, Arleth L. A comprehensive study of the self-assembly of phospholipid nanodiscs: What determines their shape and stoichiometry? *Langmuir.* 2018;34:acs.langmuir.8b01503.
30. 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;132(39):13713–22.
31. 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):3161–70.

## Supplementary information

**S1:** SEC-SAXS measurements. Normalized total intensity as a function of frame number.  $R_g$ -values over the peak on the second y-axis. Blue areas are the frames selected for analysis as presented in main text. A: Empty ND. B: TF containing ND. C: TF containing ND + FVIIa. D: TF containing ND + FVIIa + FXa + NAPc2.



**S2:** SANS measurements. A: Empty ND was measured at FRM II, Munich, in normal cuvette SANS. The sample was purified using SEC before measurement. UV280 chromatogram in A, and the blue are spanning the fractions selected for the sample. B-D: SEC-SANS measurements performed at D22, ILL, Grenoble. Normalized UV280 form as a function of time. The reason for the weird-looking chromatograms is that the flow rate was decreased when reaching the peak and speeded up again after the peak. The blue areas are the frames selected for analysis as presented in main text. B: TF containing ND. C: TF containing ND + FVIIa. D: TF containing ND + FVIIa + FXa + NAPc2.

