



**HAL**  
open science

**Direct structural/functional characterization of  
therapeutically relevant membrane protein complexes  
and soluble oligomers by NALIM (Native Liquid  
MALDI) - TOF MS, an original mass spectrometry  
approach**

Edison Zhamungui Sanchez

► **To cite this version:**

Edison Zhamungui Sanchez. Direct structural/functional characterization of therapeutically relevant membrane protein complexes and soluble oligomers by NALIM (Native Liquid MALDI) - TOF MS, an original mass spectrometry approach. Biochemistry [q-bio.BM]. Université d'Orléans, 2023. English. NNT : 2023ORLE1020 . tel-04326450

**HAL Id: tel-04326450**

**<https://theses.hal.science/tel-04326450v1>**

Submitted on 6 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# UNIVERSITÉ D'ORLÉANS

*ÉCOLE DOCTORALE Santé, Sciences Biologiques et Chimie du Vivant - SSBCV*

CENTRE DE BIOPHYSIQUE MOLECULAIRE CBM, UPR 4301

**THÈSE** présentée par :

**Edison ZHAMUNGUI SANCHEZ**

soutenue le : 05 juillet 2023

pour obtenir le grade de : **Docteur de l'Université d'Orléans**

Discipline/ Spécialité : Biologie - Spectrométrie de Masse

**Direct structural/functional characterization of  
therapeutically relevant membrane protein complexes  
and soluble oligomers by NALIM (Native Liquid MALDI) –  
TOF MS, an original Mass Spectrometry approach**

**THÈSE dirigée par :**

**Martine CADÈNE**

Directeur de Recherche, CBM UPR4301- CNRS

**RAPPORTEURS :**

**Mme. BOERI-ERBA Elisabetta**

Chargée de recherche, IBS - Grenoble

**M. MARCOUX Julien**

Chargé de recherche, IPBS - Toulouse

**JURY :**

**M. JAULT Jean-Michel**

Directeur de recherche, IBCP - CNRS/Université de Lyon  
Président du jury

**Mme. BOERI-ERBA Elisabetta**

Chargée de recherche, IBS - Grenoble

**M. MARCOUX Julien**

Chargé de recherche, IPBS - Toulouse

**Mme. NEHME Reine**

Professeure des universités, Université d'Orléans/ICOA

**Mme. BEAUFOUR Martine**

Maître de Conférences, Université d'Orléans/CBM - CNRS

**Mme. CADENE Martine**

Directeur de Recherche, Université d'Orléans/CBM - CNRS

**Invité : Mme. MORISSET-LOPEZ Severine** Chargée de recherche, CBM UPR4301- CNRS

**Invité : Mme. TOGBE Dieudonnée** Professeure des universités, Université d'Orléans/INEM

# ACKNOWLEDGMENTS

This work was done at “Centre de Biophysique Moléculaire” (CBM) UPR4301-CNRS in Orleans.

First and foremost, I would like to express my heartfelt gratitude to my thesis director, Martine Cadène, for giving me the opportunity to work in her team as well as for her expert guidance and support that were essential in ensuring the success of my research project. I am also particularly indebted to my co-advisor, Martine Beaufour, for her guidance, insightful feedback, and constant encouragement. Martine(s), your invaluable support, and mentorship have been instrumental in shaping my research and shaping me as a researcher.

I would like to extend my sincere gratitude to Guillaume Gabant, for his exceptional technical support, time, and good humor during these three years. Vraiment, merci Guillaume.

I would like to thank the expert reviewers who have generously agreed to review and provide feedback on my doctoral thesis, Dr. Julien Marcoux and Dr. Elisabetta Boeri Erba.

Thanks, Elisabetta because you were the first person who opened the doors in your lab to the fascinating field of mass spectrometry to me, I am deeply grateful for all that I learned from you.

I would also like to thank the other members of my jury, Dr. Severine Morisset-Lopez, Pr. Reine Nehmé, Pr. Dieudonnée Togbe, and Dr. Jean-Michel Jault for agreeing to judge this work.

I would like to express my deep gratitude to CBM directors during these three years, Dr. Matthieu Refregiers and Dr. Eva Jakab Toth, as well as my appreciation to the CBM lab staff, technicians, and administrative personnel for their help and cooperation.

I would like to thank all my friends who shared part of your life with me during this small journey in Orleans.

Finally, but not least, to my family who has been the driving force behind my success. All my love and gratitude to them, and to whom I dedicate this thesis.

# ABBREVIATIONS

3-AQ	3-AminoQuinoline
AA	Ammonium Acetate
ABC	ATP-Binding Cassette
ATP	Adenosine TriPhosphate
BmrA	Bacillus multidrug resistance ATP
BTB	Broad-complex, Tramtrack and Bric-à-brac
CD	Circular Dichroism
CFT	Cell-Free Translation
ChTx	CharybdoToxin
CMC	Critical Micelle Concentration
CNS	Central Nervous System
Cryo-EM	Cryo-Electron Microscopy
CTD	C-Terminal Domain
Da	Daltons
EPR	Electron Paramagnetic Resonance
ER	Endoplasmic Reticulum
ESI	ElectroSpray Ionization
FRET	Fluorescence Resonance Energy Transfer
FWI	Formic acid/Water/Isopropanol (3/1/2)
GPCR	G Protein-Coupled Receptors
GPD	Gas Phase Dissociation
HAG	HCCA-3Aq-Glycerol matrix solution
HCCA	$\alpha$ -cyano-4-hydroxycinnamic acid
HDX	Hydrogen-Deuterium Exchange
HEPES	2-(4-(2-HydroxyEthyl)-1-Piperazinyl)Ethane Sulfonic acid
IF	Inward-Facing
Ig(s)	Immunoglobulin(s)

IMS	Ion Mobility Spectrometry
$K_D$	equilibrium dissociation constant
LC	Liquid Chromatography
LP	Laser Power
$m/z$	mass/charge ratio
mAbs	monoclonal Antibodies
MALDI	Matrix Assisted Laser Desorption Ionization
MDR	MultiDrug Resistance
$M_{obs}$	observed mass
MP	Membrane Protein
MP(s)	Membrane Protein(s)
MS	Mass Spectrometry
MSP	Membrane Scaffold Protein
$M_{theo}$	theoretical mass
MV	Membrane Vesicles
NALIM	NActive LIquid MALDI
NBD	Nucleotide-Binding Domains
NTD	N-Terminal Domain
NMR	Nuclear Magnetic Resonance
nMS	native Mass Spectrometry
OF	Outward-Facing
OmpA	Outer membrane protein A
$P_i$	inorganic phosphate
P-L	Protein-Ligand complex
PLi	Protein-Ligand interactions
PNGase F	Peptide-N-Glycosidase F
P-P	Protein-Protein complex
PPi	Protein-Protein interactions

PTM	Post-Translational Modification
QC	Quality Control
RTK	Receptor Tyrosine Kinase
SAXS	Small-Angle X-ray Scattering
SDR	Structurally Diverse Region
Sh-KcsA	chimeric Shaker-KcsA
SMALPs	Styrene Maleic Acid Lipid Particles
SOLVE	SONicated Lipid VESicles
SRP	Single-Recognition Particle
TMD	TransMembrane Domains
$t_{de}$	evaporation time of the remaining water in the matrix/sample solution droplet
TEA	TetraEthylAmmonium
TF(s)	Transcription Factor(s)
$t_i$	incubation time
$t_{ivr}$	ion vacuo residence time
$T_m$	melting temperature
TOF	Time Of Flight
Tx7335	Toxin 7335
UTL	UltraThin Layer
$V_i$	inorganic vanadate
w/w	weight/weight
XL	Chemical Crosslinking
$\alpha 1AT$	alpha-1 AntiTrypsin

# TABLE OF CONTENTS

<b>ACKNOWLEDGMENTS .....</b>	<b>1</b>
<b>ABBREVIATIONS .....</b>	<b>2</b>
<b>INTRODUCTION.....</b>	<b>9</b>
<b>A. MOLECULAR ASSEMBLIES IN BIOLOGY .....</b>	<b>10</b>
<i>MEMBRANE PROTEINS.....</i>	<i>10</i>
1.General introduction.....	10
2.Membrane proteins functions .....	12
2.1. Signaling receptors .....	12
2.2. Transporters .....	13
2.2.1. Ion channels .....	13
2.2.2. Molecule transporters.....	14
2.3. Enzymes.....	15
3.MP complexes characterized by NALIM in the present work .....	15
3.1. Potassium ion (K <sup>+</sup> ) channels and toxin ligands.....	15
3.1.1. KcsA potassium ion channel.....	16
3.1.2. Chimeric Shaker-KcsA (Sh-KcsA) potassium ion channel .....	17
3.1.3. Charybdotoxin ligand.....	17
3.1.4. Tx7335 ligand .....	17
3.2. ATP-binding cassette (ABC) transporters .....	17
3.2.1. Bacillus subtilis multidrug resistance (BmrA) .....	19
3.3. LINGO-1.....	20
4.Membrane proteins solubilization methods.....	21
4.1. Solubilization in detergent micelles .....	21
4.2. Proteoliposomes.....	25
4.3. Nanodiscs.....	26
4.4. Styrene maleic acid lipid particles (SMALPs) .....	27
4.5. Sonicated lipid vesicles (SOLVE) .....	27

<i>LARGE SOLUBLE COMPLEXES</i> .....	29
1.General introduction.....	29
2.Soluble protein complexes characterized in NALIM in the present work.....	30
2.1. Immunoglobulins (Igs).....	30
2.2. Alpha1-antitrypsin (aka alpha1-protease inhibitor) .....	33
2.3. Transcription factors .....	34
2.3.1. Rho helicase and NusG transcription factors .....	34
2.3.2. ZBTB8A transcription factor .....	35
<b>B. CHARACTERIZATION OF PROTEINS IN NATIVE STATE</b> .....	<b>36</b>
<i>BIOPHYSICS TOOLS FOR THE STUDY OF PROTEINS IN THEIR NATIVE STATE</i> .....	36
1.Non-MS tools .....	37
1.1. X-ray crystallography .....	37
1.2. Small-angle X-ray scattering (SAXS).....	38
1.3. Nuclear magnetic resonance (NMR) spectroscopy .....	39
1.4. Electron paramagnetic resonance (EPR).....	41
1.5. Cryo-electron microscopy (cryo-EM).....	42
1.6. Fluorescence resonance energy transfer (FRET) .....	43
1.7. Circular dichroism (CD) spectroscopy.....	43
1.8. Mass photometry.....	44
2.Mass Spectrometry (MS) tools for the study of proteins in their native state .....	44
2.1. Hydrogen-deuterium exchange (HDX) MS .....	45
2.2. Chemical crosslinking (XL) MS .....	46
2.3. Intact protein crosslinking MALDI (XL-MALDI) MS.....	47
2.4. Ion mobility spectrometry .....	48
2.5. Fast photochemical oxidation of proteins (FPOP) .....	49
2.6. Charge Detection Mass Spectrometry (CDMS).....	50
2.7. Nano-Electromechanical Resonator-Based Mass Spectrometry (NEMS-MS) .....	50
2.8. Native MS (nMS).....	51
2.8.1. Noncovalent interactions .....	51
2.8.2. Ionization methods used in nMS.....	54



3. Native MS (nMS) advantages for structural characterization of MPs and large molecular complexes that motivated this work .....	61
---	----

## **OBJECTIVES ..... 63**

1. To obtain structural/functional information about membrane proteins and their complexes in native state such as stoichiometry, stability, and ligand binding .....	64
2. To characterize large soluble molecular complexes in order to determine their degree of oligomerization and the effect of ligand binding .....	64

## **RESULTS AND DISCUSSION ..... 66**

### **CHAPTER 1 ..... 70**

<i>CONTEXT</i> .....	70
----------------------	----

#### *NALIM ANALYSIS OF LARGE SOLUBLE PROTEIN OLIGOMERS*..... 71

INTRODUCTION .....	72
--------------------	----

MATERIALS AND METHODS .....	76
-----------------------------	----

Materials .....	76
-----------------	----

Sample preparation .....	77
--------------------------	----

NALIM MS .....	77
----------------	----

RESULTS.....	78
--------------	----

Searching for a high m/z calibrant for large complexes in NALIM .....	78
---	----

Delay optimization .....	78
--------------------------	----

IgA as a potential calibrant.....	78
-----------------------------------	----

Alpha-1-antitrypsin ( $\alpha$ 1AT) as an alternative calibrant for high masses .....	84
---	----

Alpha-1-antitrypsin ( $\alpha$ 1AT) as calibrant for a high m/z range .....	88
---	----

Structural insight into soluble protein complexes by NALIM .....	89
--	----

NusG-Rho complex in NALIM.....	89
--------------------------------	----

Oligomerization of ZBTB8A transcription factor .....	93
--	----

DISCUSSION .....	95
------------------	----

Overcoming limitations to study large molecular complexes NALIM .....	95
---	----

Exploring high molecular complexes through NALIM .....	97
--	----

CONCLUSION .....	99
------------------	----

<b>CHAPTER 2 .....</b>	<b>101</b>
<i>CONTEXT</i> .....	101
<i>NALIM ANALYSIS OF MEMBRANE PROTEINS (MPs) COMPLEXES</i> .....	103
INTRODUCTION .....	104
MATERIAL AND METHODS .....	106
Materials .....	106
Sample preparation .....	107
NALIM MS .....	107
Calculation of oligomer percentages .....	108
RESULTS .....	109
Evidence of BmrA dimer and stabilization of the dimer in NALIM .....	109
Instrumental optimizations .....	109
Matrix/sample ratio .....	111
Sample composition .....	113
BmrA substrate binding for further dimer stabilization through ligand binding .....	117
Localizing the Tx7335 binding site on KcsA .....	123
Tetramer stabilization .....	123
Competition binding experiment to determine the Tx7335 toxin binding site on KcsA .....	125
Membrane proteins from membrane preparations or liposomal membrane mimics .....	134
LINGO1 in liposomes .....	134
SpABC in cell membranes vesicles .....	137
Cell membrane vesicles (Spiking experiment) .....	140
DISCUSSION .....	146
Observation of specific membrane protein oligomers in NALIM .....	146
Stability of MP complexes in NALIM .....	146
Protein-protein, and protein-ligand interaction .....	148
Direct analysis from membrane vesicles .....	150
CONCLUSION .....	153
<b>GENERAL CONCLUSION AND PERSPECTIVES .....</b>	<b>154</b>
<b>REFERENCES .....</b>	<b>158</b>

# **INTRODUCTION**

## **A. MOLECULAR ASSEMBLIES IN BIOLOGY**

To perform a function, proteins form complexes with other biomolecules, including nucleic acids, carbohydrates, lipids, and/or other proteins. For instance, DNA-protein interactions, such as in the chromatin, plays a key role in gene expression. Carbohydrate-protein interactions are important for cell-cell communication and adhesion. Lipid-protein interactions, as found in the plasma membrane, are essential for cell integrity and communication. Furthermore, since most cellular processes rely on protein-protein interactions (PPI), molecular assemblies are of central importance in biology. Improving our knowledge about them can give us a detailed understanding of molecular functions and dynamics in the cell, either in the normal state or in pathologies, which could open the way to drug discovery and new therapies. Thus, tools that provide information about their structure, conformation, and interaction, while preserving their integrity in native-like conditions, need to be developed. Next to existing structural biology methods such as crystallography, NMR and cryo-EM, native mass spectrometry (nMS) is coming of age as a set of methods to assist in this sometimes-difficult task. In the present thesis, we developed a technique called NALIM (Native Liquid MALDI) to help characterize complexes formed around membrane proteins as well as soluble oligomers.

In the next section, we will give background on the protein complexes we focused on, based on their relevance as therapeutic targets and/or therapeutic proteins.

## **MEMBRANE PROTEINS**

### **1. General introduction**

The cell and/or its organelles are delimited by a lipid bilayer, also known as the membrane. This membrane protects and selectively controls the exchange of gases (e.g., CO<sub>2</sub>, O<sub>2</sub>) or small uncharged polar molecules (e.g., H<sub>2</sub>O, C<sub>2</sub>H<sub>5</sub>OH), between the delimited structure with its environment. However, to survive, the cell needs better and more sophisticated ways to communicate and exchange a wide variety of ions and molecules through the membrane. Thus, the specialized machinery responsible for these key roles is a family of proteins called membrane proteins (MPs). The relevance of MPs in disease and therapeutics is made clear by the fact that two-thirds of all potential drug targets are MPs, even though they represent only 30% of the proteome (Gabant & Cadene, 2008; Marty, 2020). Yet over 85% of these potential targets are still unexploited. This is due in part to the difficulties associated with their handling.

MPs can be classified into two categories; integral MPs, which are embedded in the membrane, and peripheral or membrane-associated MPs, which interact externally with the membrane or an integral MP. Since most MPs are integral and this work is mainly focused on this category, “MP(s)” will be used herein to refer exclusively to integral MP(s).

Given the amphiphilic/amphipathic nature of membrane lipids, the bilayer resembles a sandwich of 3 nm across with an outward hydrophilic part and an inward hydrophobic part. MPs share this amphipathic nature, possessing mainly hydrophobic domains and some hydrophilic domains. Thus, unlike soluble proteins, MPs present hydrophobic domains outside their core.

Structurally, alpha helices are the most common secondary structure of hydrophobic transmembrane domains. Around 20 consecutive hydrophobic residues can constitute an alpha helix that is the right length to orthogonally span the width of the membrane. This fingerprint is easily recognized in a protein sequence and has been used as a diagnostic of MPs by sequence analysis. However, since the advent of MP structural biology, the idea that alpha helices are just the right length and cross the membrane at a 90° angle has been found to have many exceptions, as the orientation, length, and position of alpha-helices in MPs relative to the membrane can differ from this simple rule (Dutzler et al., 2002; Y. Jiang et al., 2002). This has in turn shown the limitations of sequence-based algorithms and hydropathy plots to identify membrane-spanning MP domains.

Beta sheets are also found in MPs. This secondary structure has been mainly found in MPs with a barrel structure, which are placed in the outer membrane of organelles and gram-negative bacteria. Compared to alpha helices, identification of MPs based on sequence recognition of beta-sheets is even more difficult, because these beta-sheet sequences share similar characteristics with soluble proteins. The more MP structures are determined experimentally however, the more precisely AI-based algorithms will be able to assist in this task.

Until now, no MP with both alpha helix and beta-sheet has been described (Petsko & Ringe, 2004; Rodwell et al., 2018).

In MPs, hydrophobic domains directly interact with membrane lipids' nonpolar groups. Pathways leading to the insertion of hydrophobic domains inside the membrane are initiated in the endoplasmic reticulum (ER).

As with other proteins, the ribosome is in charge of synthesizing the new MP. In the nascent-polypeptide chain, a specific amino acid sequence (signal sequence) is recognized by a ribonucleoprotein, known as the single-recognition particle (SRP). SRP first binds the ribosome-nascent chain complex and subsequently binds the SRP receptor in the ER's membrane. Soon after, the ribosome interacts with a protein-conducting channel in the ER's membrane (Rapoport, 2007).

This conducting channel is created by a heterotrimeric MP complex, known as the Sec61 complex in eukaryotes and the SecY complex in bacteria and Archaea (for a review of these proteins, see Rapoport et al., 1996) At this point, the ribosome continues synthesizing the new MP while inserting the nascent chain into the protein-conducting channel. Hydrophilic domains cross entirely the channel (i.e., the membrane), while hydrophobic domains exit the channel by its lateral walls, entering thus into the lipid bilayer. The orientation of transmembrane segments in a multi-spanning MP is frequently determined by the initial transmembrane segment; although there are exceptions (Rapoport et al., 2004; Wessels & Spiess, 1988).

Once the synthesis finishes, all the machinery is dissociated, and the MP remains embedded in the membrane. This MP will be transported by an ER's membrane vesicle until its ultimate membrane, where it will carry out its function (Rapoport, 2007; Rodwell et al., 2018).

Membrane proteins assemble as complexes to carry out different functions in the cell, as detailed below.

## **2. Membrane proteins functions**

The classification of MPs according to their function can be redundant and not universally accepted, mainly because a MP can have multiple functions. For example, an ion channel can be ligand gated, which makes it a signaling receptor, while also being obviously a transporter. Bearing in mind the multifunctional nature of MPs, in this section, we will focus on the most common functions for MPs, whether in eukaryotic or prokaryotic cells, which fall within three main categories as: signaling receptors, transporters, and enzymes.

### **2.1. Signaling receptors**

In a process known as transduction, membrane protein acting as receptors receive a chemical, physical, or electrical signal, and transfer it through the membrane to generate a specific response. In these receptors, some agonist drugs can bind them and activate the physiological response, while antagonist drugs can bind them without any activation, thereby blocking or inhibiting the effects of certain molecules.

Members of the family of the G protein-coupled receptors (GPCR) represent a major group of MPs with signaling activity. GPCRs are a group of membrane receptors also known as seven-(pass)-transmembrane domain receptors. As this alternate name implies, their main structural feature consists of seven alpha helices that span the membrane and are linked by three loops on the extracellular side and three loops on the intracellular side. The extracellular loops interact with the

ligand, while the three intracellular loops interact with a protein that acts as a signal transmitter (i.e., the multi-subunit G protein).

These receptors are found only in eukaryotes, and their ligands include light-sensitive compounds, odors, pheromones, hormones, and neurotransmitters. Of the *circa* 800 GPCRs that have been discovered until now, more than 300 of them are non-olfactory receptors, which are the main target of most drugs in clinical use, although only a small fraction of them is being exploited now (Alexander et al., 2011).

Another well-known example of MP receptors is the receptor tyrosine kinase (RTK) family, which includes receptors for growth factors, hormones, and cytokines. Thus, the RTK family is important for cell growth, differentiation, survival, and dysregulation of these membrane receptors can lead to cancer and other diseases (Alexander et al., 2011).

## **2.2. Transporters**

Some MPs are specialized in moving ions or molecules through the membrane by either active transport or facilitated diffusion. The main difference between these mechanisms is that active transport requires energy, - generally obtained from ATP hydrolysis -, to transport molecules against the concentration gradient, whereas facilitated diffusion is a type of passive transport which does not directly require energy and instead uses an electrochemical gradient to move molecules from a higher to a lower concentration area.

According to the type of transport, transporters can be broadly sub-categorized as ion channels and molecule transporters.

### **2.2.1. Ion channels**

Ion channels control the passage of ions (e.g.,  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$ , and  $\text{Cl}^-$ ) through the membrane and there are several subfamilies of these channels, each one with a specific structure, and mechanism of activation. Ion channels can be (i) simple ion channels, (ii) ligand-gated ion channels (iii) voltage-gated ion channels, and (iv) mechanically triggered ion channels (Alexander et al., 2011).

Archaeobacteria can express simple ion channels, such as KcsA, which do not possess a regulatory domain. These can be regulated through pore-blocking or other mechanisms affecting the pore (Jiang et al., 2002).

Ligand-gated ion channels can respond to the change in concentration of an intracellular ion such as  $\text{Ca}^{2+}$  (Lemoine et al., 2012). Voltage-gated ion channels allow the selective transport of ions across

the cell membrane in response to changes in the transmembrane difference of potential. They are involved in electric signal propagation, pH regulation, and have an important role in excitable cells such as neurons. Finally, mechano-sensitive ion channels or stretch-gated ion channels respond to changes in local pressure and are involved in sensory perception and balance, for example. Their role in biology has been explored more recently than other ion channels (Jiang et al., 2002; Alexander et al., 2011).

### 2.2.2. Molecule transporters

In this category, we find uniporters, symporters, and antiporters. These types of transporters differ by the mechanism by which they move different molecules either in favor or against the concentration gradient.

Uniporters move a single type of molecule across the membrane. They use either active transport or facilitated diffusion.

As active transporters, we can mention the ABC transporters (explained in detail in section 3.2), which are ubiquitous MPs mainly involved in drug and xenobiotic disposal, or in the transport of endogenous solutes such as lipids, peptides, and steroids. Thanks to their role in removing external substances, particularly chemotherapeutic agents, and their involvement in antibiotic resistance, these transporters have generated significant biological interest (Alexander et al., 2011).

As facilitated diffusion transporters, we can find carriers and channels. The main difference between carriers and channels is their functional mechanism. A carrier coordinates the opening on one side with the closing on the other side. In contrast, a channel remains open at both ends while allowing specific molecules to flow, for instance porins. Porins form large channels pathways called pores and are mainly engaged in the transport of hydrophilic molecules across the membrane. Structurally, their transmembrane domains can be made exclusively with alpha helixes (e.g., aquaporins) or beta sheets (e.g., beta barrels).

Aquaporins transport water molecules across the membranes along an osmotic gradient. Until now, 13 isoforms have been identified in humans and rodents. Some of them can be considered exclusively as water channels, while others additionally let glycerol pass and are known as aquaglyceroporins (King et al., 2004).

One example of a beta-barrel porin is the bacterial protein called “outer membrane protein A” (OmpA). OmpA is found in gram-negative bacteria such as *Escherichia coli*, where it is one of the four most abundant outer membrane proteins, counting *circa* 100 000 copies per cell (Wang, 2002). Its main function is to contribute to the outer membrane’s structural integrity and the transport of molecules. However, it has been described as a bacteriocin and phages receptor (Foulds & Barrett,



1973; Morona et al., 1984), as well as a mediator in biofilm formation and in the interaction of pathogenic bacteria with eukaryotic cells (Gaddy et al., 2009; Nie et al., 2020). All this makes OmpA an interesting target for the development of vaccines and antibiotics.

Symporters and antiporters are also known as cotransporters because they couple the movement of one molecule against its concentration gradient to the movement of another one in favor of its concentration gradient. While symporters move the two molecules in the same direction, antiporters move the two molecules in opposite directions.

One example of a symporter is the sodium-glucose linked transporter or SGLT. This transporter moves in sodium into the cytoplasmic space using a favorable concentration gradient while also moving in glucose against its concentration gradient.

One example of antiporter is the sodium-calcium exchanger. This exchanger moves out calcium against its concentration gradient thanks to the movement of sodium using its favorable concentration gradient in an opposite direction than calcium.

### **2.3. Enzymes**

Some MPs act as catalysts in a variety of chemical reactions important for many cellular processes.

Many MPs acting as enzymes require cofactors such as metal ions or small molecules (e.g.,  $Mg^{2+}$ , ATP, NAD, riboflavin). These cofactors can assist in the chemical reaction, for instance accepting or donating functional groups, or they can promote specific conformational changes.

Additionally, most drugs that target enzymes are inhibitors. Based on kinetic assays, these inhibitors can be classified as competitive (i.e., targeting the same binding site as the endogenous ligand), non-competitive (i.e., targeting a different binding site than the endogenous ligand), and uncompetitive (i.e., unable to bind the free enzyme but binding to the enzyme-substrate complex).

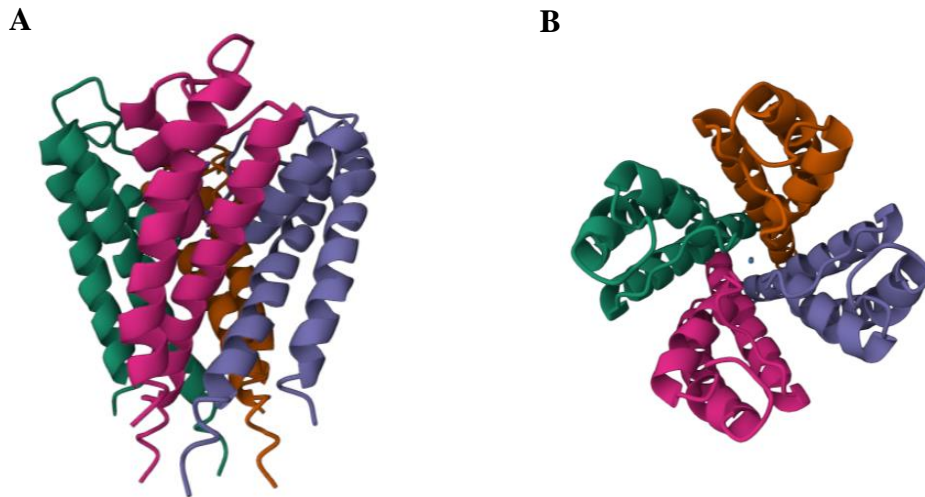
## **3. MP complexes characterized by NALIM in the present work**

### **3.1. Potassium ion ( $K^+$ ) channels and toxin ligands**

Charges carried out by inorganic ions, such as  $Na^+$ ,  $K^+$ ,  $Ca^{2+}$ , and  $Cl^-$ , across the cell membranes create bioelectric impulses involved with electrical signaling in neuronal, renal, and cardiac functions. Potassium ion ( $K^+$ ) channels are extremely selective transporters of one of these biologically abundant ions (Doyle et al., 1998; Yu et al., 2005). Generally, their malfunction can be related to diseases such as arrhythmia, epilepsy, diabetes, and cancer, which makes  $K^+$  channels important therapeutic targets.

### 3.1.1. KcsA potassium ion channel

An archetype of  $K^+$  channels is KcsA from the Gram+ bacterium *Streptomyces lividans*. The first crystal structure of this MP was obtained by X-ray crystallography (Doyle et al., 1998), where four identical monomers positioned as an inverted cone were identified, the broader end facing the extracellular side (Figure 1).



**Figure 1.** KcsA 3D structure, each monomer is represented with a different color. **A.** Lateral view. **B.** Top view of the KcsA-channel entrance with  $K^+$  ion represented as a blue dot (Figure adapted from Doyle et al., 1998 [PDB code 1BL8]).

The molecular basis for the selection and permeation of  $K^+$  was provided thanks to the revealed structure. According to Doyle *et al.* (1998), the selectivity filter has a length of *circa* 120 nm. KcsA is a small, tightly folded protein typical of Archaea proteins, and although it does not possess a regulatory domain, it is a voltage-dependent channel. After a voltage stimulus, rearrangement among amino acids in the selectivity filter opens the entrance to a size perfectly fitting for  $K^+$ . These amino acids keep the filter opened at a constant size, which prevents a reorganization of the selectivity filter and its interaction with other ions (e.g.,  $Na^+$ ). When a hydrated  $K^+$  ion from the medium takes contact with the selectivity filter, the ion automatically dehydrates by the aromatic residues present in the filter enter. To compensate for this dehydration, a sequential row of oxygen replaces the water and coordinates  $K^+$  entry into the filter. The channel inactivation is independent of a stimulus and occurs after 1 or 3 seconds of the channel opening (Rivera-Torres et al., 2016). Molecules known to block the pore include tetraethylammonium (TEA) ions, which bind on the external side. KcsA is a model  $K^+$  channel for conduction, however it does not bind other pore blockers such as venom toxins, so that it cannot be used directly to investigate the effect of these other pore blockers on activity.

### 3.1.2. Chimeric Shaker-KcsA (Sh-KcsA) potassium ion channel

Unlike KcsA, the Shaker K<sup>+</sup> channel is a potassium channel found in eukaryotes. This voltage-gated channel is encoded by the “*Shaker*” gene, and it was first identified in fruit flies, where mutations caused the flies to shake uncontrollably (Jan et al., 1977; Tempel et al., 1987).

The ability to clone and mutate this gene opened the way to study eukaryotic K<sup>+</sup> channels. One of the approaches to studying the channel function has been its interaction with venom toxins, small proteins which usually inhibit its normal function. Although all K<sup>+</sup> channels share a critical amino acid sequence (signature sequence) in their K<sup>+</sup> selectivity filter, residues lining the channel entrance are quite variable. Thus, site-directed mutagenesis experiments allowed the determination of the key amino acid residues directly related to the interaction between the Shaker K<sup>+</sup> channel and toxins such as agitoxin and charybdotoxin.

In order to mimic the Shaker K<sup>+</sup> channel pore entrance in bacteria channel (i.e., KcsA), key punctual mutations (i.e., Q58A, T61S, and R64D) in KcsA have been performed (MacKinnon et al., 1998), thus expressing a chimeric Sh-KcsA K<sup>+</sup> channel sensitive to these toxins.

### 3.1.3. Charybdotoxin ligand

Charybdotoxin (ChTx) is a small disulfide-rich protein of 37 residues which was isolated for the first time by Miller and collaborators (1985) from the *Leiurus quinquestriatus* scorpion's venom. ChTx acts as an antagonist toxin of the Shaker K<sup>+</sup> channel by binding and blocking the extracellular side of the channel pore (channel vestibule). It has been shown that one ChTx molecule binds one K<sup>+</sup> channel. Furthermore, its high affinity with the chimeric Sh-KcsA (K<sub>d</sub> = 900 nM) has been leveraged to perform investigations about the channel's structure and function (Goldstein, 1994; Yu et al., 2005).

### 3.1.4. Tx7335 ligand

Tx7335 is a small disulfide-rich protein of 63 residues isolated for the first time from the green mamba, *Dendroaspis angusticeps* in the early 2000s and characterized by Rivera-Torres and collaborators (2016). It is a three-finger toxin capable of activating pH-gated K<sup>+</sup> channels in a dose-dependent manner. Until it was discovered, no other activating toxin, a.k.a. actitoxin, had been described, as all other known toxins were inhibitors or blockers. The exact mechanism of action and details of the toxin-channel interaction, in particular the toxin binding localization, remain unknown.

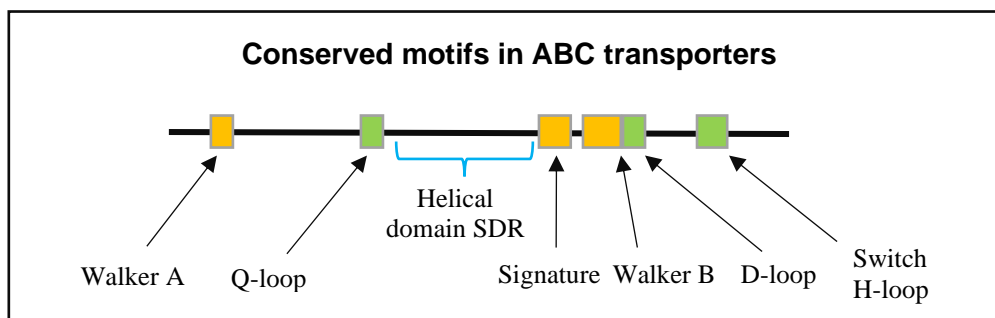
## 3.2. ATP-binding cassette (ABC) transporters

ATP-Binding Cassette (ABC) -containing systems are a superfamily of proteins found in all living organisms. The main characteristic of these proteins is their ATPase activity coupled with the

performance of important physiological functions (Davidson et al., 2008). For instance, ABC systems are known as key transmembrane transporters of a wide variety of molecules. Furthermore, their contribution to other roles such as DNA repair and translation elongation has also been described (Chakraborty, 2001; Goosen & Moolenaar, 2001). However, only ABC systems acting as transporters have transmembrane domains (i.e., MPs) (Davidson et al., 2008).

All ABC transporters (i.e., importers or exporters) share a similar topology characterized by two hydrophobic transmembrane domains (TMDs), which form the substrate transmembrane pathway, and two hydrophilic nucleotide-binding domains (NBDs) (Javed et al., 2022; Orelle et al., 2019).

The NBD can be found in every ABC protein (i.e., either soluble or MP). This NBD has a highly conserved ATP-hydrolyzing domain (Figure 2), which can be identified by three short sequence motifs such as Walker A, Walker B, and the “signature motif”. The signature motif is unique for ABC systems and it’s responsible for the protein-ATP interaction. Hence, the NBD is responsible for the ATP hydrolysis (Davidson et al., 2008; Orelle et al., 2008).



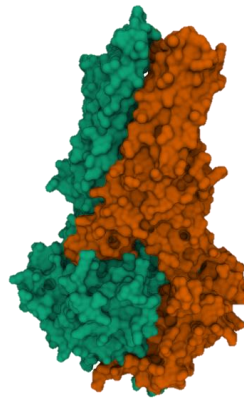
**Figure 2.** Scheme of the conserved motifs in ABC transporters, which is a highly conserved ATP-hydrolyzing domain of the NBD. The region between the Q-loop and the Signature motif forms a helical domain, known as the Structurally Diverse Region (SDR), which contains critical residues that play a key role in their interaction with the membrane (Figure adapted from Davidson et al., 2008).

ABC transporters are involved in important physiological functions such as maintaining the osmotic balance of cells, transporting essential nutrients, and in drug absorption and distribution inside the cell. Their malfunction is linked to diseases such as cystic fibrosis (due to mutation in the CFTR chloride/bicarbonate exporter) or Dubin-Johnson syndrome. Thus, they represent an important drug target family. Moreover, there is a greater interest in studying them since they can act as drug-expelling pumps, as is the case for the mammalian P-glycoprotein or the bacterial transporter BmrA, which can prevent any therapy by conferring multidrug resistance (MDR) in the context of cancer or bacterial infection (Javed et al., 2022).

### 3.2.1. *Bacillus subtilis* multidrug resistance (BmrA)

BmrA is an ABC transporter found in *Bacillus subtilis* and works as an efflux pump of cell-damaging molecules. This MP makes bacteria resistant to a broad range of drugs such as cervimycin C (Krügel et al., 2010). It is a model system for MDR transporters involved in pathogenic bacteria's resistance to antibiotics. BmrA is also relevant for its high homology with the P-glycoprotein found in mammals, which is a MP related to drug resistance in cancer therapies.

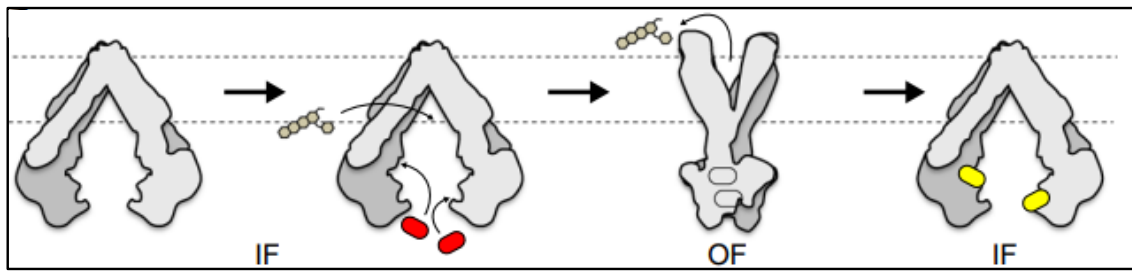
Structurally, BmrA is a homodimer (Figure 3), where each monomer has one TMD and one NBD fused. TMDs are responsible for recognizing and transporting substrates across the membrane, while the NBDs hydrolyze ATP and provide the energy for transport (Chaptal et al., 2022; Javed et al., 2022; Lacabanne et al., 2019).



**Figure 3.** BmrA 3D structure, each monomer is represented with a different color (Figure adapted from Chaptal et al., 2022 [PDB code 7OW8]).

BmrA allows the translocation of different molecules outward across the cell membrane thanks to a conformational change between an inward-facing (IF) and an outward-facing (OF) conformation (Figure 4).

The apo-state or inward-facing (IF) conformation is characterized by the two NBDs being separated from each other. In the IF conformation, the translocation pathway is accessible for the binding of the molecule to be expelled (e.g., drug, shown as hexagons in grey in Figure 4). Then, the binding of one ATP molecule on each NBD triggers a reorganization of the MP which brings the two NBDs closer in an outward-facing (OF) conformation. The OF conformation opens the pathway for the extracellular side. At that point, the NBD hydrolyzes ATP, which provides the energy required for the release of the molecule and subsequently the transmembrane domain returns to its initial conformation. How exactly the MP gets again the IF conformation after ATP hydrolysis and molecule release is not well established (Chaptal et al., 2022; Javed et al., 2022; Lacabanne et al., 2019).

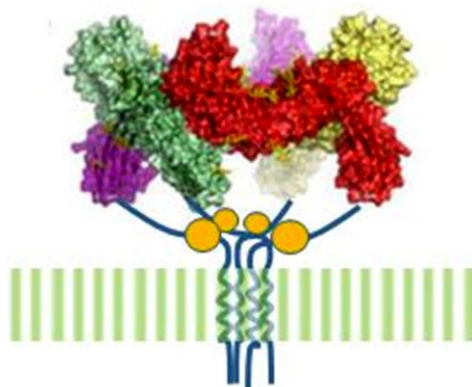


**Figure 4.** Scheme of BmrA (in grey) expelling a drug (chain of small hexagons in grey) through conformational changes between the inward-facing (IF) and the outward-facing (OF). In red ATP. In yellow ADP (Figure obtained from Lacabanne et al., 2019).

The structure of BmrA has been extensively studied using X-ray crystallography and cryo-electron microscopy. Different conformations can be induced by different mechanisms. For instance, BmrA WT can be trapped in the OF conformation induced by ATP and magnesium ions ( $Mg^{2+}$ ) binding using vanadate ( $V_i$ ). In this case,  $Mg^{2+}$  acts as a cofactor for ATP binding, while vanadate acts as a trapping agent for the ADP-BmrA in the OF conformation after ATP hydrolysis.  $V_i$  is an analog of the inorganic phosphate  $P_i$ , a coproduct of ATP hydrolysis, with a lower  $k_{off}$ . Moreover, specific mutations such as the non-hydrolytic mutant BmrA E504A, can be trapped in a pre-hydrolytic OF conformation after ATP: $Mg^{2+}$  binding (Chaptal et al., 2022; Javed et al., 2022; Lacabanne et al., 2019).

### 3.3. LINGO-1

LINGO-1 is a MP selectively expressed in neurons and oligodendrocytes in the central nervous system (CNS). LINGO-1 consists of a single transmembrane domain and a large extracellular part that contains leucine-rich repeat (LRR) motifs and an immunoglobulin-like (Ig) domain (Figure 5). The LRR motifs are involved in protein-protein interactions, while the Ig domain could play a role in signal transduction (Agúndez et al., 2015; Guillemain et al., 2020; Pepinsky et al., 2014).

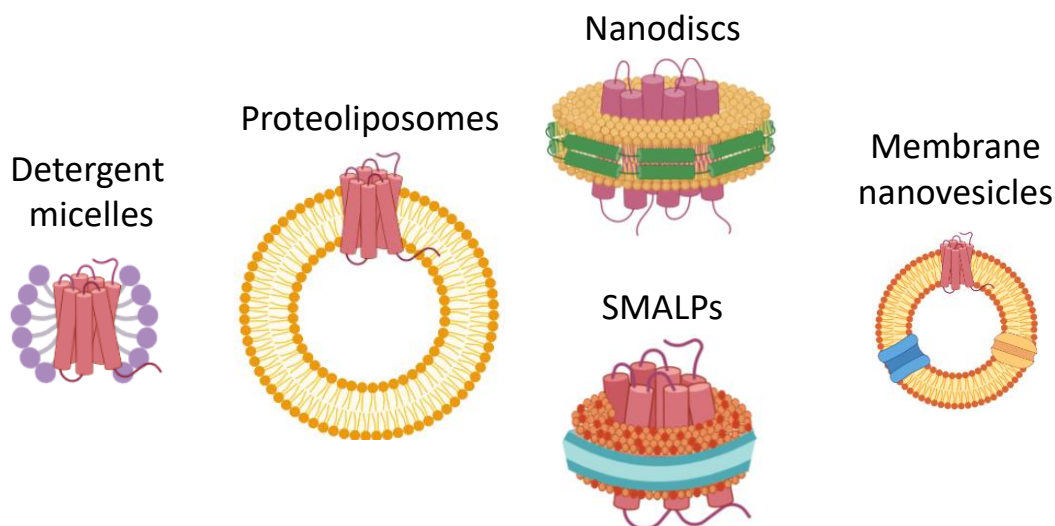


**Figure 5.** Scheme of LINGO-1 MP.

LINGO-1 is described as a negative regulator because it blocks or inhibits the action of growth factor receptors. This negative regulation affects the ability of neurons to survive and regenerate after an injury. Moreover, it interferes with the development of oligodendrocytes and their production of myelin, which is a substance that insulates axons and facilitates electrical impulses conduction in the central nervous system (CNS). Thus, LINGO-1 is linked to various neurological disorders, such as multiple sclerosis and Parkinson' disease. At the moment, several approaches are being explored to develop anti-LINGO-1 therapies to treat these disorders, based on promising effects *in vitro* and in animal models (Agúndez et al., 2015; Guillemain et al., 2020).

#### 4. Membrane proteins solubilization methods

MPs as potential targets must be characterized to effectively develop drugs and new therapies. To this end, the MPs of interest must be solubilized before analysis by biochemical and/or biophysical methods. However, due to their hydrophobic nature, MPs present a challenge for solubilization, and their study *in vitro* has always been a difficult task. For instance, only 3% of all high-resolution structures correspond to MPs (Marty 2020). Thus, different methods have been conceived to solubilize MPs while keeping them functional and intact (Figure 6). Here, we will talk about solubilization methods mainly used in structural biology approaches, including mass spectrometry.



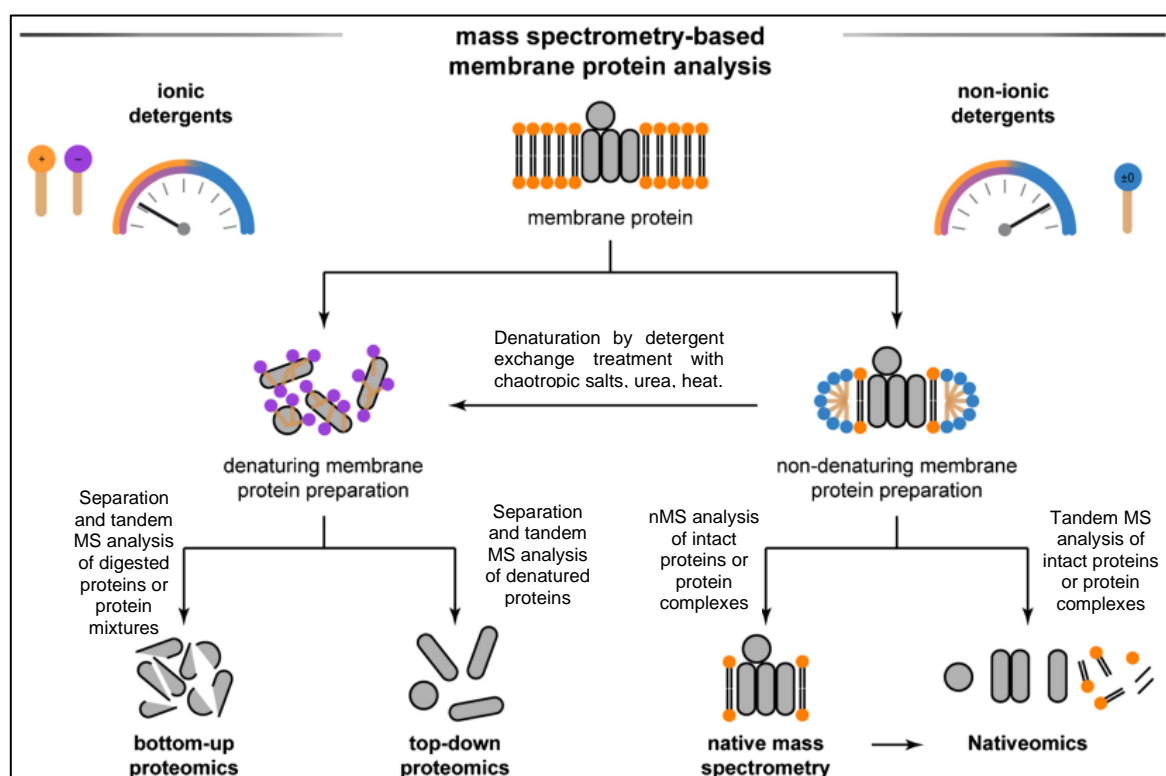
**Figure 6.** Solubilization methods for membrane proteins and their complexes.

##### 4.1. Solubilization in detergent micelles

The formation of detergent micelles is the most common method used for solubilizing MPs. As detergent molecules are amphipathic, they insert their hydrophobic tail in the lipid bilayer while keeping their polar head on the surface. Unlike lipids found in the membrane, detergent molecules have a conical geometry given by a single hydrophobic tail. Thus, increasing detergent concentration

until its critical micelle concentration (CMC) is reached will disrupt the membrane, which will then form pseudo-spherical assemblies known as micelles, in which the MP is solubilized. The MP can then be isolated using classical biochemical separation steps, such as tag-based affinity purification and various other chromatographic methods. Once a MP is solubilized, detergent concentration must be kept above the CMC at all times to keep micelles stable. At detergent concentrations below the CMC, micelles start breaking up into free detergent molecules, which leads to protein precipitation (Stetsenko & Guskov, 2017).

According to the chemical properties of the detergent, it can be classified as non-ionic, ionic, and zwitterionic. The choice of detergent is critical and will depend on each MP' characteristics as well as the method employed. While ionic detergents can be employed in conjunction with denaturing methods to completely extract a membrane protein from the membrane (Figure 7), they are incompatible with mass spectrometry (MS) due to ion suppression. Thus, non-ionic and zwitterionic detergents are preferred for native MS methods.



**Figure 7.** Use of detergents in membrane proteins investigation according to the mass spectrometry approach employed (Figure obtained from Behnke & Uerner, 2023).

Non-ionic detergents (e.g., Triton X-100, n-dodecyl- $\beta$ -D-maltoside) have no charged groups in their structure. They can disrupt protein-lipid interactions while preserving protein-protein interactions. This makes non-ionic detergents a good choice for solubilizing MPs without altering their quaternary



structure, which is crucial for native MS (Figure 7). Indeed, using n-dodecyl- $\beta$ -D-maltoside in MS, it was shown for the first time that a MP complex (i.e., transmembrane and cytoplasmic subunits) could be kept intact from the solution to the gas phase (Barrera et al., 2008).

In contrast, ionic detergents (e.g., sodium dodecyl sulfate) have either positive or negative charged groups, which makes them more likely to disrupt the quaternary structure of the MP. Zwitterionic detergents (e.g., lauryl dimethylamine-N-oxide) have both positive and negative charged groups, making them electrically neutral overall. Thus, zwitterionic detergents fall in between ionic and non-ionic detergents in terms of their mildness (Stetsenko & Guskov, 2017).

Here we summarize information about some of the most common detergents used in the investigation of membrane proteins (Table 1).

**Table 1.** Summarized information about some detergents used in membrane proteins' investigation. The aggregation number is the number of detergent molecules in a micelle at the CMC (Table adapted from Stetsenko and Guskov, 2017).

Detergent name	Abreviation	MW (Da)	CMC (%/mM)	Aggregation number	Micelle MW (kDa)
Polyoxyethylene 8 dodecyl Ether	C <sub>12</sub> E <sub>8</sub>	538.7	0.005/0.09	90–120	66
Polyoxyethylene 9 dodecyl Ether	C <sub>12</sub> E <sub>9</sub>	582.8	0.003/0.05	90	83
3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate	CHAPS	614.9	0.5/8–10	10	6
3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate	CHAPSO	630.9	0.5/8–10	11	7
Cymal-5		494.6	0.12/2.5	47	23
Cymal-6		508.6	0.028/0.56	91	32
Digitonin		1229.3	0.002/0.5	60	70
n-Dodecyl-β-D-Maltopyranoside	DDM	510.6	0.0087/0.17	80–150	65–70
n-Decyl-β-D-Maltopyranoside	DM	482.6	0.087/1.8	69	40
n-Undecyl-β-D-maltopyranoside	UDM	496.6	0.029/0.59	71	50
Lauryldimethylamine-N-Oxide	LDAO	229.4	0.023/1-2	76	21.5
Lauryl Maltose Neopentyl Glycol	LMNG	1069.2	0.001/0.01	~400	91–393
n-Nonyl β-D-Glucopyranoside	NG	306.4	0.20/6.5	133	85
n-Octyl-β-D-Glucopyranoside	OG	292.4	0.53/20	30–100	25
Octyl phenol ethoxylate	Triton X-100	624.8 (av.)	0.01/0.2	75–165	60–90

## 4.2. Proteoliposomes

Liposomes are artificial spherical vesicles made of lipids and with an aqueous interior. They have been used in the last years as nanocarriers in mRNA and DNA vaccines, or for a variety of biomolecules such as nutrients or drugs (Sejwal et al., 2017; Torchilin, 2006). When MPs are inserted in liposomes, they are called proteoliposomes.

To reconstitute proteoliposomes, different approaches can be used. In the first approach, MPs solubilized in detergent are mixed with lipids solubilized also in detergent. Then, by dialysis, gel filtration, or gradual dilution, the detergent concentration is reduced, which leads to the formation of proteoliposomes with MPs in them. A variant of this method uses unilamellar liposomes previously made. Then, detergent is added which leads to a liposome reorganization. At this point, MPs solubilized in detergent are added to the solution. Finally, detergent is removed by any of the methods mentioned above, which leads to the formation of proteoliposomes (Rigaud & Lévy, 2003; Sejwal et al., 2017).

In a second approach, liposomes previously made are mixed with MPs solubilized in detergent. This resulting solution is frozen with liquid N<sub>2</sub>, which causes small cracks in liposomes. Later, at zero degrees, the solution is slowly thawing which lets MPs insert into the liposomes' cracks. Finally, by a mild sonication, the liposomes with MPs already inserted are sealed, which gives as a result proteoliposomes (Scalise et al., 2013).

Since proteoliposomes are membrane-mimetic more stable than detergent micelles, they preserve in a better way the MP's structure and functions. However, they have some disadvantages. For instance, in a solution, there are proteoliposomes with different sizes, morphology, and number of MPs inserted. It is not possible to control the orientation of the MPs in the liposome. Moreover, the preparation of proteoliposomes still requires a pre-solubilization of MPs in detergent, which remove native lipids from the cell membrane (Rigaud & Lévy, 2003; Sejwal et al., 2017).

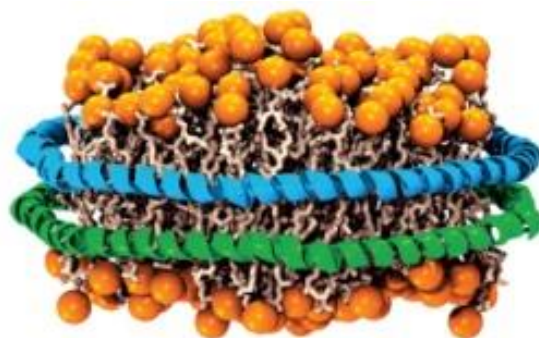
In the cell-free translation (CFT) of MPs approach, proteoliposomes are produced by addition of liposomes to the ribosome/tRNA/mRNA solution. The protein is translated and directly inserted into liposomes, as they would in native Golgi membranes, to produce proteoliposomes, which can be later purified by ultracentrifugation or ultrafiltration (Goren et al., 2009; Goren & Fox, 2008). Thus, issues linked with the use of detergents are circumvented by using proteoliposomes.

In recent months, prokaryotic and eukaryotic MPs, spanning a wide range of oligomeric states (from monomers to pentamers) and masses (from 12 kDa to 226 kDa), have been successfully ejected

preserving their native state and characterized using ESI-based nMS by Panda and coworkers (2023). These MPs were incorporated into liposomes of diverse lipid compositions. While some liposomes were designed to replicate the lipid composition of eukaryotic organelles and bacterial inner membranes, others were assembled using lipids directly extracted from *E. coli*. A critical step of the process involved modifying the surface charge of the proteoliposomes using supercharging agents, which facilitated the ejection of these complexes (Panda et al., 2023). This confirms that it is possible to study MPs and lipids directly from bilayers with specific lipid compositions using nMS.

### 4.3. Nanodiscs

Nanodiscs are self-assembled lipids, which are stabilized by two circular amphipathic  $\alpha$ -helical proteins, also known as membrane scaffold protein (MSP) (Figure 8). They were developed in 2002 by Sligar's group (Denisov et al., 2004). Since then, they have been employed to solubilize and characterize many MPs, for instance, the ABC transporters ABCB1/P-glycoprotein, and a GPCR, the human apelin receptor (A. Alam et al., 2019; Y. Ma et al., 2020).



**Figure 8.** Schematic representation of a nanodisc. In green and blue the two amphipathic  $\alpha$ -helical proteins (MSP). In orange, lipid polar heads. In grey, lipid hydrophobic groups. (Figure obtained from Shih et al., 2007)

Nanodiscs have some advantages over other solubilization methods. Thanks to their MSP, nanodiscs are quite stable, which guarantees MPs solubilization and integrity all along their analysis. Furthermore, since nanodisc are made of a lipid bilayer, they mimic in a better manner the membrane conditions, which is of great biological and physiological importance. In contrast, in detergent, MPs can be embedded in annular structures rather than actual micelles (Figure 7). For instance, Banerjee *et al.*, (2008), show better preservation of Rhodopsin in nanodisc than detergent micelles, while Ritchie *et al.*, (2009), show better ATPase activity of P-glycoprotein in nanodiscs than in liposomes. However, they still present some disadvantages such as: (i) the necessary use of detergents to pre-solubilize the MP, (ii) the impossibility of performing transport assays since both sides of the nanodisc are easily accessible, and (iii) the fairly high cost of implementation. Moreover, although it

contains a lipid bilayer, this synthetic membrane does not have the exact lipid composition and features of a natural membrane (Lee et al., 2016; Marty, 2020).

#### 4.4. Styrene maleic acid lipid particles (SMALPs)

There is a diversity of lipids that are heterogeneously distributed in the membrane. MPs are adapted to this complex environment and have established fundamental relationships with their close surrounding environment. The use of detergent micelles or nanodiscs with a synthetic lipid bilayer cannot accurately mimic the conditions that the membrane provides to the MPs for optimum folding and function. Thus, approaches that can provide more native-like conditions have been developed.

One of these approaches is the styrene maleic acid lipid particles (SMALPs), which is a type of native nanodisc. Unlike nanodisc solubilization, the MP and its surrounding lipid bilayer will be directly extracted and solubilized by an amphiphilic styrene maleic acid (SMA) copolymer, avoiding thus the use of detergents. Styrene forms a copolymer with the maleic acid, in which styrene groups can interact with lipid hydrophobic groups, while maleic acid groups face the solvent. This copolymer combines with membrane lipids to form styrene maleic acid lipid particles or SMALPs (Hellwig et al., 2018; Lee et al., 2016).

In spite of the notorious advantages of being more stable and providing more native conditions, SMALPs have some constraints. For instance, the maximal size of a MP that can be analyzed is given by the maximal SMA diameter (*circa* 15 nm), which corresponds to a MP with a maximum mass of *circa* 400 kDa (Jamshad et al., 2015; Lee et al., 2016). Another limitation is that SMA can only form SMALPs above a pH of 6.5, being insoluble at lower pH. Moreover, SMA is a good chelator of divalent cations such as magnesium ( $Mg^{2+}$ ) and calcium ( $Ca^{2+}$ ), the chelate form being insoluble (Lee et al., 2016).

Despite these limitations, since their apparition in 2009, SMALPs have been used to investigate different types of MPs, for instance, the KcsA  $K^+$  channel, the *E. coli* multidrug transporter AcrB, and the prototypical seven-transmembrane receptor protein bacteriorhodopsin (Dörr et al., 2014; Orwick-Rydmark et al., 2012; Postis et al., 2015). However, this solubilization method did not gain wide acceptance and literature on SMALPs peaked around 2020, maybe due to their heterogeneous nature (Kamilar et al., 2023).

#### 4.5. Sonicated lipid vesicles (SOLVE)

In 2018, C.V Robinson and colleagues took the next step and demonstrated the feasibility of directly ejecting intact MPs and their complexes from native membrane vesicles. These nano-sized vesicles

are made by preparation of natural membranes from whole cells followed by controlled sonication of large vesicles/cell ghosts using a method called “sonicated lipid vesicles” (SOLVE) (Chorev et al., 2018). Hence, SOLVE rules out the use of any chemical (i.e., detergent) or external agents (i.e., MSP, SMA) to isolate and solubilize the MP. Furthermore, SOLVE circumvents the limitations of the low levels of MPs expression. For instance, SOLVE coupled to native MS analysis on an orbitrap instrument was used to characterize a number of *E. coli*'s MP complexes. Some of the complexes identified were: (i) a beta-barrel assembly machinery with evidence of a chaperone-porin association and lipid interactions, (ii) efflux pumps connecting inner and outer membranes, and (iii) a pentameric pore of TonB in the inner membrane (Chorev et al., 2018).

Despite the advantages of SOLVE, this method has some limitations. For instance, since the regulation of lipids distribution between the two membrane layers performed by the lipid flippases does not occur during membrane vesicle preparation, sonication could cause that lipids become disordered, or they translocate between membrane layers. Moreover, the full lateral forces present in a membrane won't be the same in a small vesicle (Chorev et al., 2020).

Overall, SOLVE in combination with other omics such as lipidomic and proteomics can give us a complete view of a MP complex and its interactions in native conditions. This method opens the way for the development of new and better technologies, which will allow the study of MPs directly from a whole cell or a microorganism. To this end, the adaptation of the MS instruments, as well as the development of new techniques and specialized software will be required.

## LARGE SOLUBLE COMPLEXES

In the course of developing a method based on native mass spectrometry for the analysis of membrane protein complexes, the need for high mass proteins or protein complexes that could be used as calibrants appeared. This is due to the fact that at the beginning of this work, calibrants of high molecular weight for MALDI were not commercially available, although a calibrant which can reach a mass range of up to 450 000, was brought to market by CovalX during the course of this work. Calibration kits of high molecular weights of up to 1.2 MDa, such as ThermoFisher's NativeMark™, were available for electrophoresis. However, as they were not conceived for mass spectrometry analysis, they may contain additives that would make them incompatible with ionization in nMS, and proteins may bear micro-heterogeneity, which could lead to excessively broad peaks. Thus, all this sent us on a quest to find an appropriate system, as described in Chapter 1.

Since soluble proteins are generally easier to handle than MPs, we looked for soluble complexes of high mass, which then led us to attempt analysis of larger and larger oligomer distributions. As these oligomers perform essential functions in the cell, they are interesting biological objects in their own right, for which a new tool such as NALIM may prove useful to obtain structural information. Here we will expose their place in biology in general, describe the potential calibrant systems we attempted to develop, and finally give context for the oligomeric systems we studied once we were armed with the right "molecular ladder" tool.

### 1. General introduction

In biology, proteins are known to form complexes with other proteins or other biomolecules, and to make polymeric structures by repetition of a single unit. Oligomers, from the greek "oligo" which means "a few", and "mer" which means "parts", are molecular complexes that result from the association of three or more monomeric subunits. According to Goodsell and Olson (2000), in a cell, oligomers could represent *circa* 35% of the total protein content. From this percentage, almost 79% are made up of identical subunits (i.e., homooligomers), while the remaining 21% are composed of different subunits (i.e., heterooligomers). Interestingly, structural dimerization is very common, as is tetramerization, whereas odd stoichiometries constitute only a small portion of oligomeric structures (Ali & Imperiali, 2005; Goodsell & Olson, 2000). Besides this type of stable complexes, a number of oligomers form distributions with dynamic exchange between oligomeric forms.

Proteins can be inherently present in the cell as oligomers, or their oligomerization can be induced by external agents such as specific ligands (e.g., peptides, nucleotides) or some environmental conditions (e.g., pH, temperature, salt concentration). From an evolutionary point of view, oligomerization can provide some advantages over monomeric proteins, for instance: (i) the placement of enzyme active

sites at the interface between subunits, (ii) the presence of additional allosteric sites, and (iii) protection against denaturation or degradation thanks to a smaller exposed area for each monomer (Ali & Imperiali, 2005; Goodsell & Olson, 2000).

Oligomers are involved in important cellular functions, such as the regulation of gene expression, anabolic and catabolic enzyme activity, signaling transduction, and cell-cell adhesion processes (Hashimoto & Panchenko, 2010).

The disruption or abnormal formation of oligomers can lead to biological dysfunction, and the formation of high-ordered oligomers has been associated with several pathologies, such as Alzheimer's and Parkinson's diseases (Scheibel & Buchner, 2006). Typically, misfolded proteins aggregate to form high-ordered oligomers which accumulate in the cell or in the extracellular matrix. This accumulation interferes with normal cellular processes and eventually contributes to pathology development (Kumari, 2023).

In summary, understanding the role of oligomers in biology and their effect in pathologies is of major interest. However, there are experimental limitations that make their study difficult. For instance, just a low proportion of protein structures in the Protein Data Bank correspond to oligomers. Hence, new methods that preserve these structures and favor their study in native conditions are necessary.

## **2. Soluble protein complexes characterized in NALIM in the present work**

We first embarked on a search for large assemblies to use as calibrant. Their biological properties are described hereafter.

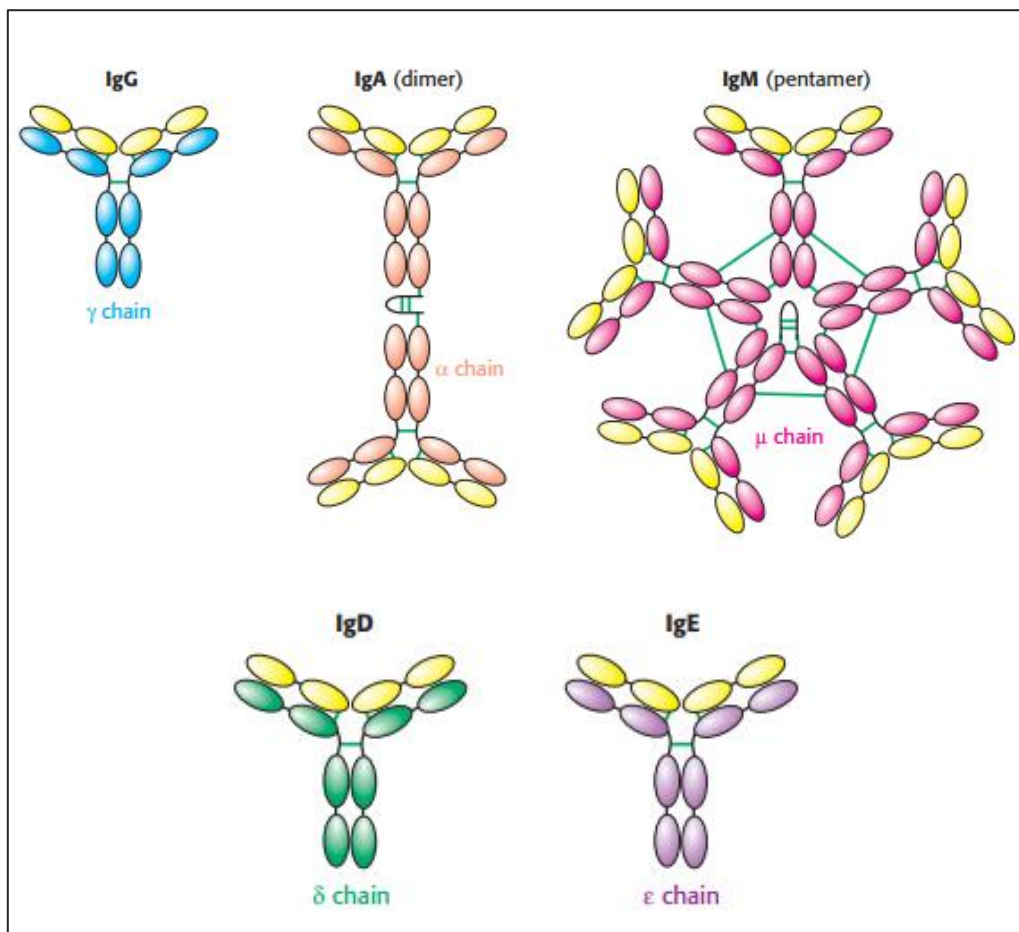
### **2.1. Immunoglobulins (Igs)**

Antibodies, also known as immunoglobulins (Igs), are glycoproteins produced by immune system cells to neutralize any pathogenic agent (e.g., bacteria, virus, toxin, allergen). This neutralization initiates by the recognition of an antigen, which is any molecule able to trigger an immune response, by an antibody.

The formation of Igs in immune system cells and its role in the immune response has been studied for years. However, in the last decades, the production of engineered Igs in the laboratory for medical applications has acquired a major interest. These Igs are known as monoclonal antibodies (mAbs) and are produced from a single type of immune cell (i.e., B cell), which is cloned to produce identical mAb copies. They show extraordinary specificity against a specific molecule or protein of a virus, bacteria, or cell, which brings an enormous range of biological applications. For instance, when



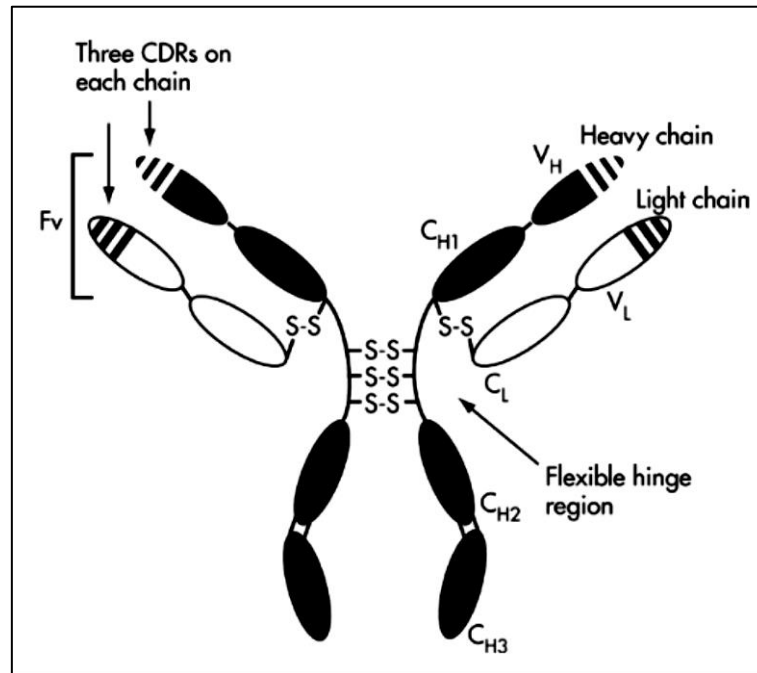
transferred into an organism, they provide a passive immunization, which can be used to treat cancer, infectious diseases, or neurodegenerative and autoimmune disorders (Nevoltris & Chames, 2018). In 2006, the global pharmaceutical market of Ab drugs was \$20.6 billion, in fields such as oncology, infectious diseases, transplantation, and cardiovascular medicine. The pharmaceutical industry is developing Abs as a promising alternative to small molecules thanks to the Ab's high specificity and fewer side effects, with an increasing number of Abs approved as drugs by FDA and EMA (Zhang et al., 2009). Today, the global market for therapeutic antibodies is more than 210 billion (Grand View Research, 2023).



**Figure 9.** Schematic representation of the different immunoglobulin classes. The light chain is represented by yellow spheres, which are combined with different heavy chains (i.e.,  $\gamma$ ,  $\alpha$ ,  $\mu$ ,  $\delta$ , and  $\epsilon$ ), each represented by a different color. Disulfide bonds are represented by green lines. IgA dimer and IgM pentamer have a junction chain (J-chain) represented by a black curve line (Figure obtained from Berg et al., 2012).

In mammals, Igs can be categorized into five main classes, namely IgG, IgA, IgM, IgD, and IgE, which have different functions, characteristics, and subunits numbers (Figure 9). Moreover, each class presents a number of subclasses (Berg et al., 2012). However, all of them share a basic structure similar in shape to a “Y”, which is formed by two identical “heavy” polypeptide chains (H) paired in

each arm of the “Y” with two shorter “light” chains (L) (Figure 10). Each H and L chain has a mass of *circa* 25 000 and 50 000 Da respectively (Amzel & Poljak, 1979; Liddell, 2013).



**Figure 10.** Schematic representation of a basic immunoglobulin IgG. Colored in black is the pair of heavy chains, which presents three constant domains (CH) and one variable domain (VH). Colored in white are the light chains, which consist of one constant domain (CL) and one variable domain (VL). The variable regions of H and L chains have three hypervariable parts (complementary-determining regions, CDRs) (Figure obtained from Liddell, 2013).

Structurally H and L chains consist of two very packed beta sheets, divided into domains which are constants (C) and variables (V). All these chains are held together by a variable number of disulfide bonds, depending on the class and subclass of Ig.

C domains get their name thanks to the amino acid sequence that is constant among all Igs. Here, glycans made of fucose, hexose, hexosamine, and sialic acid building blocks are attached to the Ig structure. This glycosylation is responsible for binding to the complement and cells related to immune response, and its degree differs according to each Ig class and subclass.

In contrast, V domains have a variable amino acid sequence. Furthermore, they present three hypervariable loops known as complementary-determining regions (CDRs), which are in charge to recognize the antigen in the host. In human beings, there can be *circa* 1010 different variations of CDRs (Y. Huang & Orlando, 2017; Liddell, 2013).

Briefly, the IgG is the most abundant Ig in serum and other fluids. It is responsible for providing immunity when facing previous disease-causing agents (adaptive immunity) and can also pass

through the placenta to protect the fetus. IgAs are a specialized type of antibody that is mainly used in the secretory pathway of adaptive immunity. They can be thus considered the first line of defense since it is the major Ig found in mammalian mucosa and secretions such as saliva, tears, and breast milk. The IgA is a dimeric Ig where two monomers are linked by a small polypeptide chain called the "junction chain" (J-chain). IgM is a pentameric Ig and is the first antibody produced after antigen recognition. Hence, it can be used as a marker for recent infection. IgD is found in B cells and could be involved in immune response activation. IgE is involved in the allergic response and parasite infections (Berg et al., 2012; Kumar Bharathkar et al., 2020).

## 2.2. Alpha1-antitrypsin (aka alpha1-protease inhibitor)

Alpha-1 antitrypsin ( $\alpha$ 1AT) is a protein of *circa* 47 kDa encoded by the *SERPINA1* gene. This protein is synthesized in hepatocytes and secreted to the plasma, from where it can diffuse to tissues. Per day, the liver can produce *circa* 34 mg of protein per kilogram of body weight, which results in a plasma concentration between 0.9 and 1.75 mg/mL. Hence  $\alpha$ 1AT is the most abundant serine protease inhibitor in the blood (Bollen et al., 1983; M. Hunt & Tudor, 2012; Strnad et al., 2020).

As its name implies,  $\alpha$ 1AT is an inhibitor of proteases and has as its main target the neutrophil elastase, which is an enzyme that proteolyzes collagen-IV and elastin in the extracellular matrix. Thus,  $\alpha$ 1AT has an important role in protecting tissues from enzymatic or inflammatory damage. Moreover, thanks to a moderate affinity for plasmin and thrombin, it may play a role in the fibrinolysis pathway, possibly helping to maintain a balance between clot formation and dissolution (Bollen et al., 1983; Talens et al., 2013).

More than 100 distinct alleles of the *SERPINA1* gene have been identified. Notably, severe  $\alpha$ 1AT deficiency is a phenotype associated with the Z mutation (E342K). This mutation induces the formation of  $\alpha$ 1AT oligomers, which are later retained in hepatocytes' endoplasmic reticulum. As a result of this oligomerization,  $\alpha$ 1AT levels decrease in the blood and hence, in tissues, which allows tissue damage by neutrophil elastase. Thus, this renders the organism vulnerable to emphysema. Moreover, as  $\alpha$ 1AT oligomers accumulate in liver cells, a high  $\alpha$ 1AT concentration can be toxic and lethal for the cell, which can lead to neonatal hepatitis and adult cirrhosis (Haq et al., 2013; M. Hunt & Tudor, 2012).

*In vitro*, the formation of  $\alpha$ 1AT oligomers can be induced by high temperatures, proteolytic digestion, low concentrations of denaturant, or acidic pH. At the moment, this strategy is being used to study

the role of  $\alpha$ 1AT oligomers in different pathologies so as to develop therapies, as well as to use  $\alpha$ 1AT anti-inflammatory properties against other diseases (Haq et al., 2013; Irving et al., 2011).

Once we identified a calibrant for use over a wide mass range (see Chapter 2), we trialed the characterization of oligomers.

### **2.3. Transcription factors**

Transcription factors (TFs) are proteins that bind specific DNA sequences to regulate the transcription of genes. They are involved in essential physiological roles, for instance, cell development, cell differentiation, and the immune response. Their malfunction triggers regulatory problems that may promote diseases such as cancer or autoimmune disorders. Although their significance, understanding how TFs control gene expression is still challenging, mainly by the formation of dynamic protein complexes and the high complexity of the regulatory network.

#### **2.3.1. Rho helicase and NusG transcription factors**

NusG/Spt5 proteins belong to a family of TFs conserved across all organisms. While NusG (N-Utilization substance G) protein is found in bacteria, Spt5 is its corresponding homolog in Archaea and eukaryotes (Lawson et al., 2018). NusG is an essential TF that can play two opposite roles in transcription. Its N-terminal domain (NTD) can bind the RNA polymerase (RNAP) to enhance the transcription activity (i.e., antitermination), while the C-terminal domain (CTD) can recruit accessory factors to promote the end of the transcription (i.e., termination) (Burmam et al., 2010; Burova et al., 1995).

Rho protein is an ATP-dependent helicase that is involved in the termination of transcription. This protein is an archetype of factor-dependent termination exclusively found in bacteria, and it was first identified by Roberts (1969) in *E. coli*. Rho protein recognizes specific sequence motifs in the RNA transcript called Rho utilization (rut) sites. These sequences are of *circa* 80 nucleotides and are characterized by being rich in guanine and cytosine. When Rho encounters the rut site, it binds to the RNA and moves towards the transcription bubble, where it uses its helicase activity to unwind the RNA-DNA hybrid, eventually releasing the RNA polymerase and terminating transcription (Lawson et al., 2018; Peters et al., 2012).

In the absence of ATP and prior to RUT recognition, Rho is a homohexamer with an open-ring conformation. In this open conformation, Rho binds RUT sites, and it drives the nascent RNA inside the ring. In the presence of ATP, Rho switches to a closed-ring conformation, trapping the RNA

inside the central pore of the ring. Then, Rho moves downstream, using energy from ATP hydrolysis, until it encounters the RNAP, which will stop transcription (Lawson et al., 2018).

Interestingly, Rho can work on RNAs that bear few RUT sites or even no site, and the mechanism for this is not well understood. It has been suggested that NusG plays a key role by allowing the Rho-dependent termination of transcription in a sequence-independent manner. According to Peters *et al.* (2012), in RNAs with few RUT sites, NusG could facilitate Rho-RNA binding by bringing Rho closer to the RNAP exit channel. Thus, there are still many questions to be answered about the mechanisms of NusG modulation of Rho activity.

### **2.3.2. ZBTB8A transcription factor**

The ZBTB8A transcription factor, also known as BOZF1, belongs to the ZBTB family of TFs, which is defined by an N-terminal BTB (Broad-complex, Tramtrack and Bric-à-brac) domain and C-terminal zinc fingers. According to Kim *et al.* (2013), ZBTB8A is a transcriptional repressor that acts on proximal promoters of the cyclin-dependent kinase gene, CDKN1A, inhibiting its transcription. This leads to downstream transcription repression of some regulatory proteins related to apoptosis, cell differentiation, cell cycle progression, and oncogenesis, suggesting a possible link with tumors and metastasis. Indeed, according to Kim *et al.* (2013), ZBTB8A is highly expressed in different cancer tissues. Furthermore, ZBTB proteins are frequently substrates for SUMOylation in which small ubiquitin-like modifiers (SUMO) are covalently linked to the substrate protein. This reversible post-translational modification (PTM) is directly associated with key cellular processes such as DNA replication and reparation, transcriptional regulation, cell cycle progression, signal transduction, and more (Hendriks et al., 2014; Matic et al., 2010).

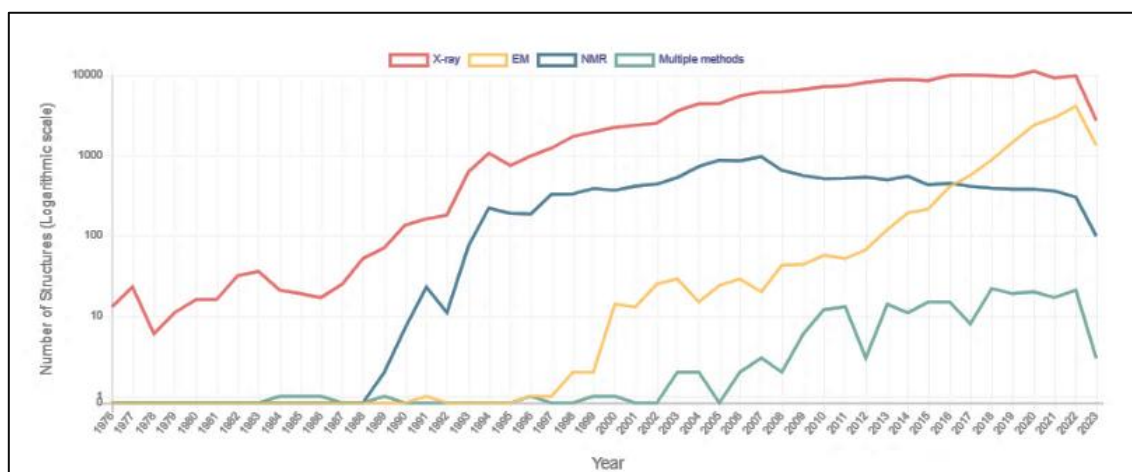
ZBTB proteins can dimerize thanks to the BTB domain, which facilitates the DNA-protein binding mediated by the CTD's zinc fingers (Kim et al., 2013). In some ZBTB proteins, particularly those from insects, interaction between dimers has been suggested. Recent findings by Mance and collaborators (currently in publication) suggest that the BTB domain of ZBTB8A can generate large filaments with unknown stoichiometry and its role in the cell.

## B. CHARACTERIZATION OF PROTEINS IN NATIVE STATE

### BIOPHYSICS TOOLS FOR THE STUDY OF PROTEINS IN THEIR NATIVE STATE.

In order to answer complex biological questions through structural biology, a researcher needs the right tools to obtain information. The choice of these tools will be based on different factors, for instance, the physicochemical characteristics of the sample (e.g., hydrophobicity, size), the type of resolution wanted (e.g., molecular or cellular level), as well as the type of information required (e.g., stoichiometry, binding interactions).

There are several biophysics tools that can be used for the study of proteins in their native state, such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, cryo-electron microscopy (cryo-EM), mass spectrometry (MS), electron paramagnetic resonance (EPR), small-angle X-ray scattering (SAXS), fluorescence resonance energy transfer (FRET), and circular dichroism (CD) spectroscopy.



**Figure 11.** Number of released PDB structures per year according to different structural methods. In red, by X-ray crystallography (X-ray). In blue by nuclear magnetic resonance (NMR). In yellow by electron microscopy (EM). In teal other techniques or multiple techniques (Figure obtained from RCSB.org, 2023).

Nowadays, X-ray crystallography is the most used structural technique for solving protein structure at atomic level. This fact can be confirmed by the number of structures released in the protein data bank (PDB) by this method (Figure 11). However, the combination of different structural techniques provides a comprehensive understanding of the three-dimensional (3D) protein structure, its

complexes and dynamics, as well as a set of tools to explore the protein functions while maintaining a native state (Table 2).

**Table 2.** Characteristics and structural information obtained using the three main non-MS-based structural techniques.

	<b>X-ray crystallography</b>	<b>Nuclear magnetic resonance spectroscopy</b>	<b>Cryo-electron microscopy</b>
<b>Structural Information</b>	Molecular structure	Molecular structure and dynamics	Molecular and supra-molecular structure
<b>Resolution</b>	1 to 3 Å	1 to 3 Å	4 – 5 Å in average
<b>Sample size</b>	Small to large-sized molecules	Small to medium-sized molecules	Molecules > 60 kDa
<b>Sample Preparation</b>	Crystallization	Solution-based techniques	Vitrification

## 1. Non-MS tools

As mentioned above, many structural techniques exist to study proteins in their native state. In this point, we focus on general aspects of non-based-MS structural techniques for native protein characterization.

In paragraph 3, we will present a comprehensive analysis of the benefits of MS, and in particular nMS, compared to the three main conventional structural methods (i.e., X-ray crystallography, NMR spectroscopy, and cryo-EM) in the characterization of MPs and large molecular complexes.

### 1.1. X-ray crystallography

X-ray crystallography is a powerful technique for determining the 3D structure of proteins and other biomolecules at an atomic resolution level (i.e., 1 to 3 Å). The pioneers of X-ray diffraction experiments on crystals were the group of Max von Laue and Braggs father and son, at beginning of 1910 (W. H. Bragg & Bragg, 1913; W. L. Bragg, 1913; Friedrich et al., 1913). Almost 20 years later, this technique began to be applied to the study of proteins, and the first protein structure determined by X-ray crystallography was published 30 years later (Kendrew et al., 1960; Maveyraud & Mourey, 2020).

All crystallographic experiments begin with the production of high-quality protein crystals since atoms arranged in a regular and repeating pattern will diffract X-rays in a specific pattern of spots on a detector. The pattern of spots generated provides information about the electron density map, which will be used to reconstruct the protein structure. However, obtaining good crystals is not always easy, especially with impure samples, and proteins that have low solubility, high heterogeneity, high mass, and highly flexible domains or regions (Dauter & Wlodawer, 2016; Maveyraud & Mourey, 2020; Neutze et al., 2004). Generally, proteins need to be engineered to excise flexible domains or regions or be fused with soluble proteins such as the maltose-binding protein (MBP) to improve protein solubility and stability (Sun et al., 2011). For instance, Kobilka and colleagues fused a GPCR with an easy crystallizable T4 lysozyme, which allowed the GPCR's crystallization (Zou et al., 2012).

Generally, diffraction experiments are conducted using synchrotrons of the third generation as X-ray ion sources. Given that X-ray radiation can cause significant damage to protein crystals, a flow of cold inert gas (e.g., nitrogen) at approximately 100 K is used during data acquisition. However, this requires the preparation of crystals in suitable cryo-solutions to prevent ice formation (Dauter & Wlodawer, 2016).

In the last two decades, the development of a fourth generation of X-ray ion sources, known as XFELs (X-ray Free Electron Lasers), has played a crucial role in atomic-level studies of proteins. These new XFEL sources generate intense pulses of around  $10^{13}$  photons, over ten times the photon yield of third-generation synchrotrons. Moreover, XFEL reduces crystal exposure time to just a few femtoseconds, thus mitigating radiation-induced damage and allowing data acquisition at room temperature. Furthermore, XFEL sources produce brighter X-rays within a smaller area, thereby substantially enhancing resolution (Fuller et al., 2021; N. Huang et al., 2021; Pellegrini, 2020).

In summary, the automation of high-throughput crystallization methods, improved X-ray sources (i.e., brighter synchrotrons), faster detectors, and more powerful software enable the determination of protein structures within a few days. However, the structural characterization of MPs and large biomolecular complexes still presents a big challenge for X-ray crystallography due to the interference of membrane mimetics and the size of the complexes during crystallization (Neutze et al., 2004).

## **1.2. Small-angle X-ray scattering (SAXS)**

Small-angle X-ray scattering (SAXS) is a biophysical technique that uses the same principle of X-ray scattering as X-ray crystallography. However, unlike X-ray crystallography, SAXS analyzes



proteins directly in solution, thus circumventing the need for protein crystallization, a process that can be laborious, time-consuming, and may require protein modifications. Although the scattering pattern produced by SAXS is simpler and has lower resolution compared to X-ray crystallography, it provides valuable insights into the size, shape, and arrangement of biomolecules and their complexes directly in solution (Da Vela & Svergun, 2020; Kikhney & Svergun, 2015).

In SAXS, a collimated monochromatic X-ray beam is directed at a sample solution placed in a quartz capillary. The X-rays are scattered, and the detector records their intensity and scattered angles. The generated 2D scattering pattern is then typically radially averaged, this yields a 1D intensity curve that can be used to extract information about the size, shape, and other information of the sample (Da Vela & Svergun, 2020; Kikhney & Svergun, 2015).

Additionally, SAXS can quantitatively characterize the different coexistent protein conformations in the sample. Furthermore, in time-resolved studies, SAXS provides information about kinetics of protein interactions, and conformational changes (Pollack & Doniach, 2009). This makes SAXS a useful tool to study heterogeneous systems or intrinsically disordered proteins (IDP), which fluctuate between different conformations when they interact with other proteins or biomolecules (Kikhney & Svergun, 2015)

In contrast, SAXS has some disadvantages over other structural techniques to study MPs. For instance, the presence of detergent molecules around the hydrophobic domains can strongly scatter X-rays and contribute significantly to the overall scattering signal. This fact creates challenges for an accurate determination of the protein shape (Molodenskiy et al., 2020). Moreover, detergent-solubilized MPs require a detergent concentration above the CMC in the buffer, which results in a coexistence of free detergent micelles and micelles containing MPs. Since both species have similar size and scattering power, it is difficult to distinguish the contribution of each component in order to isolate just the curve from micelles containing MPs (Berthaud et al., 2012).

### **1.3. Nuclear magnetic resonance (NMR) spectroscopy**

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful biophysical technique applied to determine the structure and interactions of proteins at atomic resolution level - either in solid or solution state-.

NMR takes advantage of the magnetic properties of certain atomic nuclei such as hydrogen-2 ( $^2\text{H}$ ), carbon-13 ( $^{13}\text{C}$ ), nitrogen-15 ( $^{15}\text{N}$ ), and fluorine-19 ( $^{19}\text{F}$ ), which exist in one of the two possible nuclear spin states, commonly referred to as "up" or "down" states. These spin states correspond to different levels of energy. Under a strong magnetic field, spins of atomic nuclei will orient themselves

either in the same direction as the magnetic field or in the opposite direction. Then, when a radiofrequency pulse is applied, spins flip from a lower energy state to a higher energy state (i.e., absorbing electromagnetic radiation). The frequency required to flip the spins is known as the resonant frequency. This resonant frequency depends on the strength of the external magnetic field, the type of atomic nuclei, and the chemical environment surrounding each nucleus. When the radiofrequency pulse ends, spins relax to their original state, emitting electromagnetic radiation at the same resonant frequency as the absorbed radiation. Thus, NMR analysis records these resonant frequencies, which gives access to detailed information about the position and orientation of atomic nuclei in the molecular structure of a protein (Hu et al., 2021; Raja & Barron, 2023).

NMR can be performed in different dimensions depending on the complexity of the sample and the experimental goals.

One-dimensional (1D) NMR shows chemical shifts and intensities of resonances. It is used to identify chemical and functional groups in molecules. Moreover, by analyzing changes in the chemical shifts and intensities under different experimental conditions (e.g., the addition of chemical compounds or biomolecules), 1D NMR can be used to determine protein folding or interactions with other molecules. In drug discovery, it can be used to determine compound hits and characterize their binding affinity (Cala et al., 2014; Gossert & Jahnke, 2016).

Two-dimensional (2D) NMR correlates the chemical shifts and intensities of two different nuclei in the sample. Here we can find different methods such as correlation spectroscopy (COSY), J-spectroscopy, exchange spectroscopy (EXSY), and nuclear Overhauser effect spectroscopy (NOESY). Correlating these methods, information about molecular conformation, connectivity, and dynamics can be obtained. Last but not least, 2D NMR allows for the determination of the 3D structure of a protein (Emwas, 2015).

In general, NMR experiments follow a four-stage protocol, (i) protein production with isotopic labeling, (ii) data collection, (iii) structural calculation and refinement, and (iv) structural quality assessment (Hu et al., 2021)

NMR has some advantages over other structural techniques. For instance, NMR: (i) analyzes biomolecules directly from solution, preserving their native state, (ii) studies protein dynamics in real-time, from picoseconds to seconds (or even days), (iii) does not require crystallization, which circumvents issues related to this step necessary for X-ray crystallography. However, NMR has some limitations. For example, NMR: (i) requires protein isotopic labeling, which can be laborious and expensive, (ii) can be sensitive to sample contaminants, and (iii) 2D NMR experiments take several hours to several weeks, which can make them unsuitable for the analysis of biological samples that

may undergo changes or degradation during the course of the experiment. Moreover, the complexity of (i) protein labeling, (ii) data-acquisition methods, and (iii) data processing increases with the size of the analyzed biomolecule (Emwas, 2015; Hu et al., 2021; Marchand et al., 2017).

#### **1.4. Electron paramagnetic resonance (EPR)**

EPR is a biophysical technique used to study paramagnetic species, which are molecules or atoms that have unpaired electrons, such as free radicals and transition metal ions. The principle of EPR closely resembles that of NMR. However, EPR measures the interaction between an external magnetic field and the magnetic moment of unpaired electron spins, in contrast to NMR which is based on the magnetic moment of isotopic nuclei.

EPR can provide information about the dynamics and orientation of biomolecules in solution or in membranes. For instance, it has been used to obtain information about the topology of the human voltage-gated potassium channel KCNQ1-VSD with respect to the lipid bilayer of the proteoliposome (Sahu et al., 2015). This information is important for understanding how these molecules interact with their environment and how they carry out their biological functions.

Additionally, one technique widely utilized to measure distance in EPR spectroscopy is the Double Electron-Electron Resonance (DEER). DEER uses two microwave pulses separated by a time delay, known as “the evolution time”. When two paramagnetic centers are in a range between 2 and 7 nm, their electron spins experience a dipolar interaction. Since this interaction modulates the echo amplitude, the DEER experiment measures these modulation patterns as a function of the evolution time. Thanks to these patterns, the distance and relative orientation between the two paramagnetic centers can be extracted. Thus, DEER can be used to obtain information about the structure and distances of a wide range of biomolecular complexes including protein-protein or protein-ligand complexes in different physiological conditions such as in solution, in the lipid bilayer, or in membrane mimics (Klare, 2017; Stein et al., 2015).

However, EPR has disadvantages. First, EPR is less sensitive than other biophysical techniques, since the interaction between the unpaired electrons in a paramagnetic species and an external magnetic field is weaker than the interactions that generates signals in other biophysical techniques. Second, since biological systems are devoid of unpaired electrons, EPR requires the introduction of stable radicals which provide unpaired electrons. For instance, the introduction of cysteines by site-directed mutagenesis which are then reacted with a sulfhydryl-specific nitroxide reagent is required to produce stable spin-labeled cysteine side-chains (Sahu et al., 2015).

Additionally, since relaxation times are too short for membrane proteins at room temperature, these EPR experiments typically require extremely low temperatures, which can affect the biological activity of the system being studied (Sahu et al., 2015).

### **1.5. Cryo-electron microscopy (cryo-EM)**

Cryogenic electron microscopy (cryo-EM) is a technique used to determine the structure of proteins and their complexes in the native state with a resolution that tends to near-atomic (i.e., higher than 2.5 Å) (De Oliveira et al., 2021), 5.5 Å being the resolution average of resolved structures in 2022 (Electron Microscopy Data Bank [EMDB], 2023).

In cryo-EM, 3D density maps are a reconstruction based on thousands of “single particle” images (i.e., 2D projections of a macromolecule) of the sample in random orientations. These 2D images are obtained using an electron microscope that projects a beam of high-energy electrons through the sample. These electrons will be scattered by the atoms in the sample, and the thicker the sample region, the more the scattering (i.e., darker zones). Then these images are aligned, classified, and processed in computational software. For proteins, imaging data collection is made from a flash-frozen sample in liquid nitrogen or ethane, which prevents crystal formation and preserves the protein structure and its conformations to such a degree that atoms' positions differing by only a few angstroms can be overlaid (Elmlund & Elmlund, 2015; Sigworth, 2016). Moreover, by rapidly freezing the sample, this technique can be used to get images at different time points of biological processes such as ligand binding or enzymatic catalysis. Thus, cryo-EM can reveal 3D structures, conformations, and interactions of a protein with other (bio)molecules either at a given time point or in a dynamic manner.

Unlike x-ray crystallography, cryo-EM does not require protein crystallization, and compared with NMR, it is more suitable for larger biomolecular complexes. However, this technique has various disadvantages. For instance, protein heterogeneity and/or the presence of contaminants make images complex and difficult to interpret due to the limitations of computational software, so that cryo-EM requires highly pure and homogenous samples (Elmlund & Elmlund, 2015). Moreover, since this technique produces radiation damage, the data acquisition may be made by reducing the amount of electron radiation (i.e., low-dose) in order to preserve samples sensitive to this problem. Nevertheless, this low-dose imaging generates poor signal-to-noise ratio images, which demands advanced computational techniques and a lot of time to extract useful information (Sigworth, 2016). Finally, proteins smaller than 60 kDa do not generate enough contrast to get high-resolution structures, so that their structural characterization is still a challenge for this technique (Fan et al., 2019).

### 1.6. Fluorescence resonance energy transfer (FRET)

Fluorescence resonance energy transfer (FRET) is a technique based on the transfer of energy between two fluorescent molecules, known as donor and acceptor, which allows to follow inter or intramolecular interactions and to obtain spatial measurements on the nanometer scale (Fang et al., 2023).

FRET involves the excitation of the donor fluorophore molecule with a specific wavelength of light. In order to recover its basal energy state, a fluorophore usually emits light at a longer wavelength. In FRET, the donor fluorophore can transmit a quantum of energy (a "virtual photon"), to an acceptor molecule if the acceptor is in close proximity (from 1 to 10 nm) to the donor. This phenomenon is called non-radiative dipole-dipole coupling. The transferred energy puts the acceptor in an excited state, from which in turn it emits light at another wavelength. The efficiency of transfer is a function of the 6<sup>th</sup> power of the distance between the 2 fluorophores. Thus, by monitoring the light emission of the acceptor, using a fluorimeter, one can measure the ratio of the energy transferred from the donor to the acceptor over the total energy absorbed by the donor (i.e., FRET efficiency) and estimate the donor-acceptor distance (Clegg, 1995; Fang et al., 2023).

FRET is commonly used in biological research to study biomolecular interactions either *in vitro* or in living cells, such as the interaction between nucleic acids and proteins (Alam et al., 2023), or protein-protein interactions (Park et al., 2019). Moreover, since FRET can be used to follow interactions along the time, this technique can be used to determine dissociation constants (Jiang et al., 2019), or structural rearrangements such as in the host cells during the fusion and entry of the viruses. This latter is useful to develop strategies to fight against viral infections and develop vaccines (M. Lu, 2021). FRET has also made it possible to study membrane proteins, such as their interaction with lipids, their insertion and folding in the membrane, as well as their interaction with other proteins (for a review, see Krainer et al., 2019). However, due to the diffraction of light, FRET has a limited spatial resolution in nearby molecules. Moreover, FRET generally requires protein modification to add a fluorescent label, which can affect the structure and function of the molecules of interest.

### 1.7. Circular dichroism (CD) spectroscopy

Circular dichroism (CD) spectroscopy measures the difference in the absorption of left- and right-circularly polarized light by a chiral molecule, such as amino acids in proteins, which interact differently with the two types of polarized light due to their asymmetric geometry (Büyükköroğlu et al., 2018). Thus, CD provides information about the protein secondary structure (e.g., the content of alpha-helices, beta-sheets, and random coils). CD provides a rapid analysis of the effect of different

buffer components, environmental conditions, and ligands on protein stability and conformational changes involving secondary structure (Greenfield, 2006; Miles & Wallace, 2016; Siligardi et al., 2014). However, one of the main disadvantages of CD is its limited utility for MPs since membrane mimetics create spectral artifacts that interfere with CD analysis. For instance, the absorption spectrum of the empty membrane mimetics (i.e., liposome, nanodisc, detergent micelle) can overlap with those containing the MP. Thus, it is difficult to accurately quantify the amount of light absorbed by the sample, which interfere with the data interpretation (Miles & Wallace, 2016).

### **1.8. Mass photometry**

Mass photometry was developed by Prof. Kukura's team at Oxford University (Cole et al., 2017). This method involves the precise illumination of the sample solution using a laser light. When the light interacts with molecules present in the solution, it becomes scattered in various directions. This scattered light is then re-directed towards a detector. The amount of light scattered by a particle increases proportionally with both the particle's size and its refractive index. Since the optical properties and density of biomolecules are nearly similar, the scattering signal can be correlated with the mass of the molecules (Refeyn Ltd., 2022). Thanks to mass photometry, Prof. Kukura and colleagues could quantify protein–protein interactions and the heterogeneity of macromolecular machines, directly from a solution (Soltermann et al., 2020; Sonn-Segev et al., 2020). Over the course of this work, mass photometry was applied to membrane protein complexes (Olerinyova et al., 2021). However, in terms of resolution, mass photometry is closer to MALS than to nMS methods.

## **2. Mass Spectrometry (MS) tools for the study of proteins in their native state**

Mass Spectrometry (MS) is a set of analytical techniques that measure the mass-over-charge ( $m/z$ ) ratio of ionized molecules. Ions are separated based on this  $m/z$  ratio using an electric or magnetic field. From  $m/z$ , knowing the charge of species, molecular masses can be calculated in order to identify and characterize different molecules. Its high sensitivity and low sample-analysis time consumption have made MS an essential technique in structural biology. For instance, MS can easily verify the protein's primary sequence so as to detect the presence of mutations, post-translational modifications (PTMs), chemical modifications, and/or proteolytic cleavages, which have been a fundamental basis for many findings and research (Bolbach, 2005; De Hoffmann & Stroobant, 2011). However, to better comprehend protein mechanisms and functions in biological systems, one of the main MS objectives is to gain reliable information about their tertiary and quaternary structure, as well as their interactions with other biomolecules. Thus, MS scientists have developed approaches for gaining structural insights into proteins and their interactions (Table 3).

**Table 3.** Information obtained by MS-based tools for structural studies:

	<b>Hydrogen-deuterium exchange</b>	<b>Chemical crosslinking</b>	<b>Ion mobility spectrometry</b>	<b>Fast photochemical oxidation of proteins</b>
<b>Acronym</b>	HDX-MS	XL-MS	IMS-MS	FPOP-MS
<b>Structural Information</b>	Protein folding, interactions, and dynamics	Intra or inter protein interactions, and distance constraints	Protein shape and size	Protein folding, interactions, and dynamics

### 2.1. Hydrogen-deuterium exchange (HDX) MS

HDX-MS is a footprinting technique that requires protein labeling with deuterium or through exposure to a deuterated environment. As a result, there will be an exchange between hydrogen and deuterium (H-D) atoms from easily exchangeable sites such as the amid groups on the protein the protein backbone. To “quench” (i.e., stop) the reaction and preserve the information until the MS analysis, H-D exchange can be controlled by modifying environmental factors such as the temperature or the pH (Reading, 2018). Prior to MS analysis, the protein is proteolyzed. MS analysis of cleavage peptides is then used to determine the degree and rate of atoms exchange, which depends on the accessibility of the solvent to certain areas of the protein. Thus, HDX provides structural insights into protein folding, conformation, and interactions, (De Vera et al., 2017; Hamdi et al., 2017; Shukla et al., 2014).

HDX can be used through two distinct labeling methods: pulsed labeling and continuous labeling. In pulsed labeling, the sample is subjected to specific conditions. Subsequently, at various time points, a small portion of the sample is exposed to deuterium for a brief time. This enables the monitoring in real-time of the impact of experimental conditions on the sample (Dafun & Marcoux, 2022). For instance, Khanal and coworkers (2012) applied this approach for the first time to an integral membrane protein, bacteriorhodopsin, which comprises seven transmembrane helices along with a covalently bound retinal cofactor. Authors investigated the refolding process of an SDS-denatured bacteriorhodopsin MP within bicelles over a 10 milliseconds to 24 hours timeframe. This method yielded insights into the time required for the protein to regain its secondary and tertiary structures. In contrast, the continuous labeling technique involves prolonged exposure of the sample to deuterium and calculates the average (H-D) exchange reaction taking place across all molecules in a specific condition, such as the presence of a ligand or variations in pH. Subsequently, comparing the

labeling of the tested condition with the labeling of the negative control provides valuable insights into protein dynamics (Dafun & Marcoux, 2022).

For instance, Busenlehner and coworkers (2004) applied the continuous labeling method to investigate the structural and dynamic behaviors of microsomal glutathione (GSH) transferase 1 (MGST1), a functional integral membrane protein. MGST1 is a trimeric protein containing a cytosolic GSH binding domain, and a chemical and oxidative stress-sensing cysteine residue. By analyzing the kinetics of hydrogen-deuterium exchange, authors observed that binding of a single GSH molecule to one trimer unit triggered conformational changes in specific spatial regions. This cooperative conformational transition involved coordinated movements of the transmembrane helices as well as rearrangements within the cytosolic domain. Moreover, the modification (alkylation) of the stress-sensing cysteine residue prearranges the helical structures, enhancing the cooperative conformational transition. Thus, this conformational transition leads to the protein's catalytic activation, which is crucial for its functional response to chemical and oxidative stress conditions.

Hence, HDX method can be applied to study MPs, providing invaluable insights into protein folding, conformation, and dynamics.

HDX methods require a high level of skill and resources. For example, sample preparation can be complex and time-consuming. Moreover, deuterated samples are susceptible to back-exchange to hydrogen, so that the sample pH and in-source atmosphere have to be very well controlled. Alternate methods based on reverse exchange exist but further complicated preparation. For all these reasons, HDX is not widely used, and the technique is mostly practiced by select labs that have invested resources in mastering it.

## **2.2. Chemical crosslinking (XL) MS**

XL-MS is a technique to study either intra or inter-protein interactions. To this end, XL-MS requires a crosslinking agent that covalently links binding moieties through reaction of each of its ends to an amino acid. For instance, the Bis(sulfo-succinimidyl)glutarate or BS2G is an homobifunctional cross-linking agent that can react with nucleophilic groups such as primary amines in lysine residues. The cross-linking reaction must be performed in mild experimental conditions to preserve the protein structure, although protein with few solvent-accessible lysines may lose solubility upon reaction. Then, the protein is proteolytically digested prior to MS analysis, using MALDI or ESI methods to identify stretches of protein sequence around the crosslinked amino acids (Reading, 2018; Sinz et al., 2015).



XL-MS can be employed for the characterization of MPs. For instance, Rehkamp and colleagues (2021) combined XL-MS with computational modeling to reveal the first 3D structure of a specific MP called rod-outer-segment guanylyl cyclase 1 (ROS-GC1). ROS-GC1 plays a key role in converting light into electrical signals in the retina, and mutations in this protein have been associated with various retinal diseases that frequently precede blindness.

Additionally, XL-MS has also been used to gain insights into the physical interactions between proteins within living organisms. For instance, Zheng and colleagues (2011) created the largest collection of *in vivo* crosslinked peptide sequences, providing *in vivo* topological data on many protein interactions. The acquired data exhibit strong concordance with known protein and complex structures determined by X-ray crystallography. What is particularly interesting about this study is that it provides unbiased data, as it was performed without prior information regarding specific interactions. Instead, a broad range of protein interactions was analyzed systematically. This approach can provide new insights into proteins where there is no high-resolution structural information yet, and/or insights that has not been previously considered.

Hence, crosslinked MS provides information about topology, interacting regions, and distance constraints that are used to re-construct 3D structures and reveal protein functions (Dafun & Marcoux, 2022; Reading, 2018; Sinz et al., 2015).

Unfortunately, XL-MS has some disadvantages such as non-productive reaction of the reagent (when only one end reacts), nonspecific binding, or disruption of the protein's native structure (Reading, 2018; Sinz et al., 2015).

### **2.3. Intact protein crosslinking MALDI (XL-MALDI) MS**

As mentioned above, in crosslinking experiments, samples are enzymatically digested and then analyzed by MALDI and/or ESI to identify interacting interfaces in proteins in a complex. Only a few publications have shown the use of XL-MALDI MS, i.e., chemical crosslinking (XL), without enzymatic digestion, combined with MALDI MS to identify protein partners in molecular complexes, as shown by Evans and Rohrmann (1997) or Borchers and Tomer (1999). In these studies, the main constraints were instrument limitations to detect ions in the high-mass range.

In order to counteract these limitations in characterizing large molecular complexes, a novel approach presented by Zenobi's group uses a commercial kit of crosslinking agents (i.e., CovalX™) along with a high-mass MALDI detection system. This high-mass MALDI detection system employs a cryodetector (Yanes et al., 2006), which is a type of detector that operates at very low temperatures, typically below 2 K. This cooling detector significantly enhances the detection of larger and heavier ions (Frank et al., 1996; Twerenbold, 1996).

Thus, the XL-MALDI approach provides information about protein-protein or protein-ligand interactions as they were in their native state prior to analysis. However, in the denaturing conditions used for MALDI-MS, the crosslinked species can survive, while non-crosslinked complexes are dissociated.

For instance, Zenobi and colleagues (2018) provided insights into the specificity, stoichiometry, and binding sites of five different nanobodies as they interacted with the PglK membrane protein, an ABC transporter involved in the N-glycosylation pathway of *Campylobacter jejuni*. The utilization of nanobodies as stabilizing agents for membrane protein crystallization has gained relevance in recent years. However, certain nanobodies exhibit low selectivity and affinity, resulting in the formation of heterogeneous complexes in solution, which can impact the crystallization process. Therefore, the use of cross-linking agents and high-mass MALDI-MS as a rapid screening tool has the potential to enhance the utility of nanobodies as stabilizing agents.

Although this new method offers numerous advantages for analyzing large molecular complexes, there are still some limitations inherent to crosslinking, such as the possibility of nonspecific reactions, the loss of solubility of complexes upon crosslinking, and changes to the structure and function of proteins compared to their native, non-crosslinked state.

#### **2.4. Ion mobility spectrometry**

Ion mobility spectrometry (IMS) is a technique to separate gas-phase ions based on their size, shape, and charge into an inert-gas-filled chamber, where a high-voltage electric field is applied. This fact causes ions colliding with inert-gas molecules to vary their velocity according to the parameters before mentioned, which produces separation of ions of the same mass that vary in their shape. IMS coupled with MS can be used to improve the spectral resolution and S/N ratio to analyze complex sample mixtures (Reading, 2018).

IM-MS can be used to run collision-induced unfolding (CIU) experiments. In CIU experiments, the energy applied to gas-phase protein ions is gradually increased, leading these proteins to unfold. This technique can identify the effect of protein interactions (e.g., ligand binding such as lipid binding in the case of MPs) or subtle changes in protein structure caused by post-translational modifications such as disulfide bonds on protein stability (Fantin et al., 2021). For instance, Laganowsky and coworkers (2014) investigated how the folding and structure of MPs are influenced by lipid environments. They conducted experiments on the aquaporin Z (AqpZ) and the ammonia channel (AmtB) from *E coli*, and the mechanosensitive channel of large conductance (MscL) from *M tuberculosis*. In this work, authors observed that MP complexes maintain their folded structures in the gas phase. Moreover, by applying CIU, they screened non-ionic detergents and could rank the

bound lipids based on their capacity to stabilize the membrane protein structure against gas-phase unfolding.

Similarly, Fantin and coworkers (2021) employed IM-MS and CIU to discern significant differences in the stability of various isoforms of the peripheral myelin protein 22 (PMP22). PMP22 is a MP situated in neurons and plays a key role in promoting myelination. Mutations in PMP22 can give rise to diseases such as Charcot-Marie-Tooth and Dejerine-Sottas syndrome. Thanks to IM-MS and CIU, authors observed that mutated proteins showed lower monomer conformational stability compared with the wild-type protein. Moreover, they showed how these PMP22 proteins can interact to form homodimers. Thus, these findings offer valuable insights into the mechanisms underlying PMP22 dysfunction related to diseases.

### **2.5. Fast photochemical oxidation of proteins (FPOP)**

Fast photochemical oxidation of proteins (FPOP) was first developed by Hambly & Gross (2005) as a proteomics technique that can be used to label the surface of a protein exposed to the solvent with hydroxyl radicals ( $\cdot\text{OH}$ ). This labeling can be generated by irradiation at 248 nm of a solution made of protein, hydrogen peroxide, and hydroxyl scavengers, which causes the oxidative modification of amino acid residues. The labeled protein is then proteolyzed and analyzed by MS (Johnson et al., 2019; Reading, 2018).

FPOP can be a good method for mapping MP exposed surfaces in the membrane or in membrane mimetics. For instance, using FPOP, the light-harvesting complex 2 (LH2) from *Rhodobacter sphaeroide* (solubilized in nanodiscs) exhibited significant oxidative modification in its solvent exposed regions. In contrast, domains buried in the lipid bilayer showed no response to this labeling process (Y. Lu et al., 2016). Thus, by measuring the degree of oxidative modification of peptides, FPOP-MS can provide information about the arrangement of the LH2 in the nanodisc.

Similarly, Pan and colleagues (2009) applied FPOP to bacteriorhodopsin (BR) in its natural lipid bilayer. They found that oxidative modification exclusively occurred at the nine methionine (Met) residues of BR. This contrasted with BR solubilized in water through membrane mimetics, where various types of amino acids were labelled by hydroxyl radicals. These results show that FPOP is a useful tool for studying MPs structure and conformation within their native environment. Moreover, this highlights that membrane mimetics can distort MPs natural structure and conformation, or at the minimum, that they do not protect all the protein areas covered by the lipid bilayer (Pan et al., 2009). Additionally, the diffusion of hydrogen peroxide across the membrane allows proteins to be labeled directly from cells. For instance, while highlighting that the majority of cells are viable, Espino and coworkers (2015) reported the oxidative modification of 105 proteins, including MPs, in cells.

However, FPOP requires a costly excimer laser for H<sub>2</sub>O<sub>2</sub> photolysis, making it a confidential technique. Moreover, buffers need to be compatible with hydroxyl radicals, and hydroxyl radicals display varying reactivity towards different amino acids, which increase the difficulty of interpreting the data (Johnson et al., 2019; Reading, 2018).

## **2.6. Charge Detection Mass Spectrometry (CDMS)**

CDMS is a technique that simultaneously detects both the mass-to-charge ratio ( $m/z$ ) and the charge of molecular ions, enabling an accurate ion mass determination in a sample. One main advantage of this technique is that it does not require extensive ion separation, which is particularly useful for complex mixtures and molecular assemblies larger than several hundred kilodaltons. In CDMS, ions are confined within an electrostatic linear ion trap (ELIT), where they move back and forth, generating a signal that is analyzed using fast Fourier transforms. The frequency of their oscillation corresponds to the ions'  $m/z$ , and the amplitude of oscillation relates to their charge. This technique offers insights into the dynamics and behavior of ions within the trap, contributing to a better understanding of ion-ion interactions (Botamanenko & Jarrold, 2019). However, this technique presents some disadvantages such as low resolution, moderate sensitivity, and the need to build customized instruments in the laboratory, as commercial instruments are not yet available.

## **2.7. Nano-Electromechanical Resonator-Based Mass Spectrometry (NEMS-MS)**

NEMS-MS is a cutting-edge analytical technique developed by Christophe Masselon's team, which combines elements of nanotechnology, electromechanics, and mass spectrometry, and can measure the mass of very large objects such as whole viruses. NEMS are tiny devices on the nanometer scale that act as both detectors and mass analyzers simultaneously. They are designed to vibrate at specific frequencies under particular electric fields or mechanical forces. These vibrations are highly sensitive to changes in mass within the device; for instance, when a virus binds to the NEMS surface, the added mass induces a shift in the resonant frequency of the device. This mass-induced resonant frequency shifts enables NEMS to directly determine the mass of these particles. The resolution of the method, however, is still extremely low with values hovering around 1. The resonators need to be custom-made, and can be used for few analyses, making the method's running cost very high. In spite of these limitations, NEMS occupies a unique niche for determining the mass of extremely large molecular complexes, nanoparticles, and even viruses (Clement et al., 2021; EDYP LAB, 2019).

## 2.8. Native MS (nMS)

Native MS (nMS) groups different methods specially developed to ionize proteins while preserving their tertiary and quaternary structure in the gas phase. To this end, noncovalent interactions must be preserved using mild experimental conditions that avoid denaturation by staying as close as possible to physiological conditions. Thanks to nMS, structural insights about proteins and their complexes such as stoichiometry, stability, conformation, and interaction with other molecules can be obtained. This information improves our understanding of protein functions and dynamics in the cell, either in a normal state or in a pathology, which could open the way to new therapies and drugs (Leney & Heck, 2017).

### 2.8.1. Noncovalent interactions

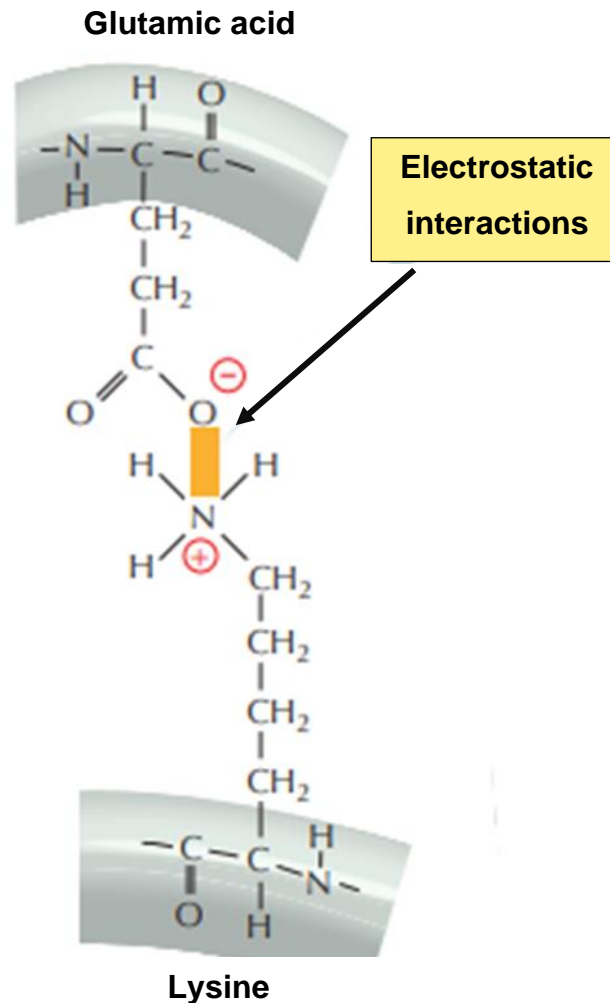
Unlike covalent bonds, non-covalent bonds do not involve the permanent sharing of electrons between the atoms of bound molecules. These interactions play a crucial role in maintaining the folded structure of a protein as well as in mediating dynamic and transient interactions with other biomolecules such as DNA, other proteins, and various types of ligands. There are four noncovalent interactions commonly identified as salt bridges, Van der Waals forces, hydrogen bonds, and hydrophobic interactions. Considered separately, these noncovalent interactions result in relatively low energy levels. However, the abundance, sum and cooperativity of all these interactions raise their global strength to a significant level compared with the energy of covalent bonds.

#### *a) Salt bridges or electrostatic interactions*

Salt bridges involve interactions between oppositely charged amino acids, either inter or intra-protein (Figure 12). This interaction is also known as electrostatic interaction, and its energy is given by Coulomb's law:

$$E = \frac{k \cdot q_1 \cdot q_2}{D \cdot r},$$

where  $k$  is a proportionality constant,  $q_1$  and  $q_2$  are the atoms' charge,  $r$  is the distance between atoms, and  $D$  is the dielectric constant. For instance, the energy of two ions of opposite charges at a 3 Å distance in water, which has a dielectric constant of 80, is *circa* 6 kJ/mol (Berg et al., 2012; Rodwell et al., 2018).

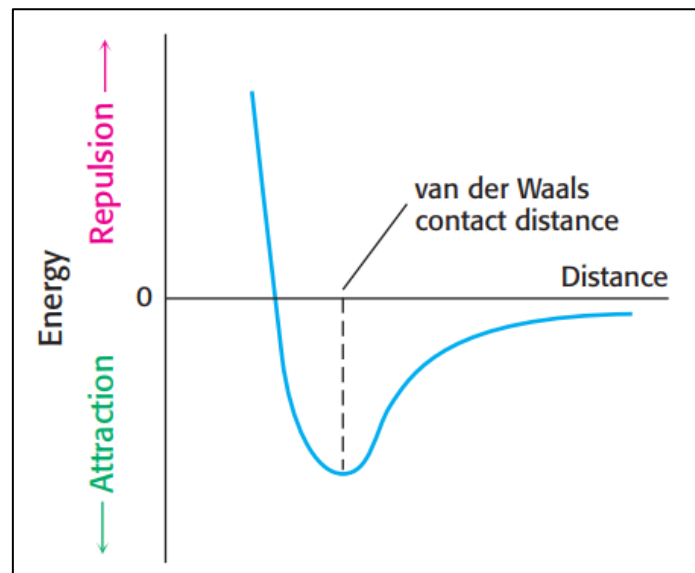


**Figure 12.** Scheme of an electrostatic interaction (represented in yellow) between two amino acids, a lysine and a glutamic acid (Figure obtained from Alberts, 2013).

***b) Van der Waals forces***

Van der Waals forces refer to interactions between neutral molecules through dipole-dipole interactions. A dipole refers to the separation of the electrical charge in a molecule, which means that it has a partial positive charge ( $\delta^+$ ) and a partial negative charge ( $\delta^-$ ) at opposite ends. The dipole can be permanent or induced by a neighboring molecule. In any case, the dipole-dipole interaction happens between the  $\delta^+$  end of one with  $\delta^-$  end of the other.

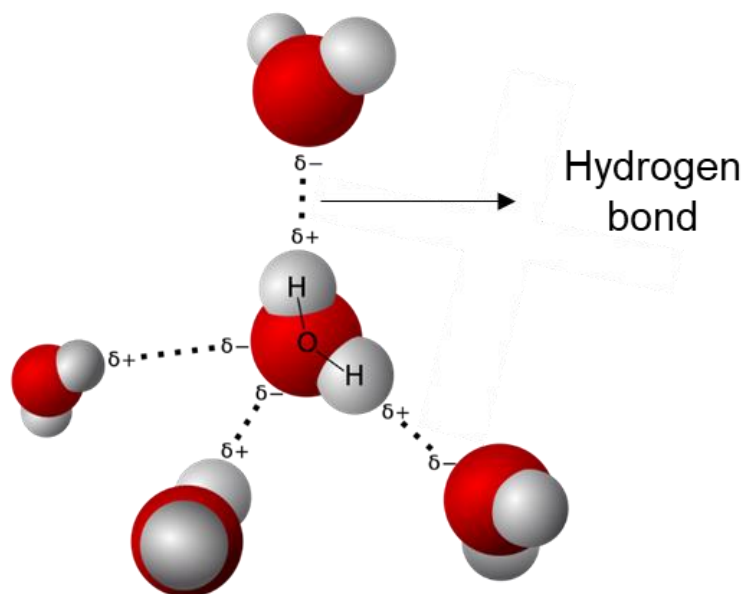
Additionally, Van der Waals forces are distance-dependent and occur in what is known as “Van der Waals contact distance” (Figure 13). Thus, as two atoms come closer to each other, the Van der Waals forces increase in strength. However, if these atoms are too close, strong repulsive forces become dominant which leads to repulsion among them (Berg et al., 2012; Rodwell et al., 2018).



**Figure 13.** Effect of the distance on the energy of the Van der Waals force (Figure obtained from Berg et al, 2009).

### *c) Hydrogen bonds*

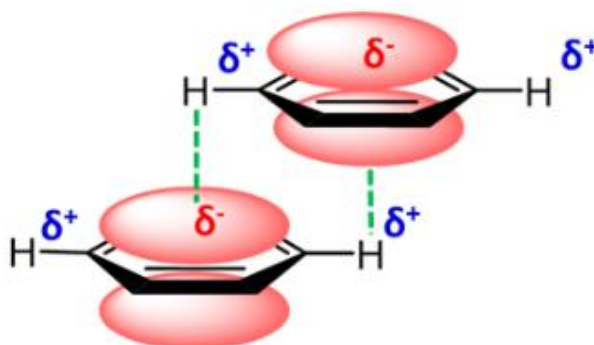
Hydrogen bonds are interactions between a hydrogen atom, which is covalently bonded to an electronegative atom (such as nitrogen or oxygen), with an unshared pair of electrons of a nearby electronegative atom (Figure 14). These interactions are fundamental in biology. For instance, the formation of specific base pairs in the DNA double helix is attributed to hydrogen bonds. On average a hydrogen bond length ranges between 2.4 Å and 3.5 Å with an energy ranging from 4 to 20 kJ/mol. These interactions are relatively strong compared to other noncovalent interactions (Berg et al., 2012; Rodwell et al., 2018).



**Figure 14.** Scheme of hydrogen bonds between water molecules. In red, oxygen atoms. In white, hydrogen atoms (Figure obtained from Wikipedia, 2023).

#### d) *Pi-stacking*

Another form of electrostatic binding is  $\pi$ -stacking, which occurs either in staggered or in perpendicular configuration between aromatic amino acids and/or nucleic acid bases, at a distance that is superior to Van der Waals forces. (Figure 15)



**Figure 15.** Pi-stacking between benzenes molecules. In red, partial negative charges ( $\delta^-$ ) and in blue partial positive charges ( $\delta^+$ ) of benzene atoms (Figure obtained from LibreTexts.org, 2022).

#### e) *Hydrophobic interactions*

Hydrophobic interactions arise because water molecules (i.e., polar) form hydrogen bonds with each other and other polar molecules, while they cannot form hydrogen bonds with nonpolar molecules. Nonpolar substances tend to minimize their contact with water by clustering together and reducing their surface area of interaction. This reduces the overall energy of the system. For instance, membrane proteins devoid of micelles tend to aggregate to avoid contact with water. Moreover, this phenomenon explains the formation of the micelles (Berg et al., 2012; Rodwell et al., 2018).

Aromatic and aliphatic amino acids contribute to the hydrophobic core of proteins and to the formation of hydrophobic patches. However, cooperativity effects with other types of non-covalent interactions means that hydrophobic interactions are not necessarily all lost in the gas phase within the time-frame of a mass spectrometry experiment, which explains how proteins can be kept folded and why native mass spectrometry is actually possible.

### 2.8.2. Ionization methods used in nMS

#### a) *ESI*

Electrospray ionization or ESI is a soft ionization method that was first coupled to mass spectrometry by Dole and coworkers in the late 60s (Dole et al., 1968). It was developed into a full-fledged method for the analysis of biomolecules, in particular proteins, by John Fenn in the late 80s, which revolutionized the field of MS and for which he was awarded the 2002 Nobel Prize in Chemistry (Fenn et al., 1989). Nowadays, ESI is one of the most widely used ionization methods applied in MS



due to its ability to produce ions of large and complex molecules such as proteins, peptides, and nucleic acids directly from the solution.

To generate gas-phase ions, an electrolyte solution passed through a capillary tube (1–10  $\mu\text{L}/\text{min}$ ) under a strong electrical field (in the order of  $10^6$  V/m) at atmospheric pressure. This strong electrical field produces an electrophoretic separation of charges with an accumulation of specific ones at the end of the capillary tube, generating thus a conical shape known as the Taylor Cone. Here, the accumulation of similar charges produces that repulsion forces overcome the surface tension force, which will generate the dispersion of charged droplets in the form of a spray. The spatial dispersion of this spray is restricted by injecting gas coaxially. Moreover, the remaining liquid present in the drop will then be quickly evaporated when this crosses a curtain of a hot inert gas (e.g., nitrogen) or by heating the capillary tube (De Hoffmann & Stroobant, 2011; Wilm, 2011).

Two models try to explain the final production of gas-phase ions. First, the “ion evaporation model” (IEM) proposes that as the solvent evaporates, the droplet reaches a specific size where the electric field strength at the surface is sufficiently large to facilitate ion release. Second, the “charge residual model” (CRM) assumes only one analyte molecule per droplet, which will undergo series of Coulomb fissions until a complete desolvation of the analyte, obtaining thus a charged molecule (Wilm, 2011). Nowadays, it is accepted that small ions (e.g., salts) are produced mainly by the IEM, while larger ions (e.g., proteins) are produced by the CRM.

In addition to ionizing samples directly from a solution, ESI has the following advantages: (i) it can be coupled to liquid chromatography (LC) to separate and analyze complex mixtures of (bio)molecules, (ii) it can be used with different types of analyzers such as quadrupole, time-of-flight (TOF), or ion trap instruments, and (iii) it can be used as a quantitative method since the intensity of the signal is directly related to the quantity of the analyte in the sample. However, ESI is a technique that is very sensitive to adduct formation and signal suppression by sample contaminants such as salts and detergents. Furthermore, the high number of charges can make it difficult for data interpretation and can lead to protein unfolding by Coulombic repulsion. In native ESI, however, the smaller number of charges is an advantage that has been leveraged for fragmentation, as spectra are less complex, which facilitates charge attribution, precursor selection, and fragment interpretation.

The first ESI-based nMS characterization of protein complexes experiments were conducted in the early 1990's (Katta & Chait, 1991). Nowadays, ESI is the main ionization method used for studying biomolecules in native conditions. For instance, based on the charge distribution of a protein, its

conformation in solution can be inferred (Chowdhury et al., 1990). This information can be used to analyze the effect of different conditions on the conformation of a protein or a complex. One of the key applications of ESI in nMS is the study of the interaction of proteins with other biomolecules (e.g., oligosaccharides, peptides, proteins), which are essential for many biological processes and medical applications. For instance, ESI-nMS allows the screening of drug candidates by directly detecting the noncovalent complexes (for a review see Hofstadler & Sannes-Lowery, 2006).

ESI-nMS has been also used to obtain important information about the structure and conformation of MPs, which are not easily accessible to other biophysics methods due to the presence of detergent. In the instrument, MPs can be stripped of micelles or membrane mimetics so that they can be analyzed in isolation while preserving their native state. For instance, Barrera and coworkers (2008) used ESI-nMS to obtain key information about MP complex stoichiometry. They first characterized the intact heteromeric transmembrane complex BtuC2D2, an ABC transporter responsible for transporting vitamin B12 in *E. coli*. Subsequently, the complex was dissociated in subunits by subjecting the MP complex to maximum acceleration voltages, thereby identifying the transmembrane (BtuC) and the cytoplasmic (BtuD) subunits by separating them. Furthermore, they were able to observe changes in the transmembrane subunits and identify the cooperative binding of ATP (Barrera et al., 2008).

Similarly, ESI-nMS has been employed to investigate the role that lipid binding plays in the structure and function of MPs. For instance, Bolla and coworkers (2018) studied the interaction between lipid II, which works as a precursor in the formation of peptidoglycan, and two the lipid II flippases MurJ and FtsW. This research revealed that MurJ exhibits a stronger binding affinity for lipid II in comparison to FtsW. Furthermore, authors identified a specific lipid called cardiolipin (CDL) that, similarly to non-annular lipids, does not form part of the membrane structure. Interestingly, CDL influences the binding of lipid II to MurJ, suggesting a potential role for CDL in modulating the activity of MurJ (Bolla et al., 2018). Hence, bound lipids analysis provides key information about MPs' interaction with the membrane and the implications of these lipids on MPs' structure, stability, and function (Robinson, 2019).

Finally, ESI-nMS can measure binding affinities directly in the instrument, with  $K_D$  values that directly reflect values found by solution-based biophysical methods (Greig et al., 1995; Jaquillard et al., 2012; Sannes-Lowery et al., 2000).

#### ***b) Nanospray-Desorption Electrospray Ionization (nano-DESI):***

DESI is a soft ionization technique developed by Cook's team in 2004 (Takáts et al., 2004), which enables the direct analysis of biomolecules from solid biological samples, such as tissues. In DESI, a

flowing-solvent bridge is made between two capillaries over the sample surface. One capillary supplies the solvent, while the second capillary absorbs this solvent, with the recently dissolved biomolecules, and carries it into the mass spectrometer. A variation of DESI called nano-DESI was developed around 2008, and protein analysis demonstrated by the Laskin group in 2012. The interaction between the solvent and the sample, combined with the use of a solvent containing nondenaturing ammonium acetate, allows molecules desorption while preserving the native state of biological molecules (Hale & Cooper, 2021; Roach et al., 2010).

Applied to membrane proteins, nano-DESI has demonstrated the capability to identify intact MPs such as the porin OmpF in octyl glucoside (OG) micelles, and the dimer of the sugar transporter SemiSWEET from *V. splendidus*. Furthermore, nano-DESI provided insights into the interactions of OmpF with the lipid phosphatidylglycerol (POPG) and the drug kanamycin. Finally, an experiment conducted with a Class A G-protein-coupled receptor (P2Y<sub>1</sub>) allowed the determination of selective binding between this MP and a specific antagonist (MRS2500) in a mixture containing multiple components (Ambrose et al., 2017).

Hence, Nano-DESI emerges as a soft ionization technique with the potential to identify and spatially map membrane protein complexes, on sample surfaces.

### ***c) Liquid Extraction Surface Analysis (LESA) coupled to MS***

LESA is an ambient ionization technique for solid surface analysis which can be coupled to ESI-based nMS and was developed by Van Berkel's group in 2010 (Kertesz & Van Berkel, 2010). LESA works by using a small droplet of liquid solvent to extract molecules from specific sites in biological samples, such as tissue sections, bacterial colonies on agar, or dried blood spots, and re-aspirating the droplet after extraction. Subsequently, the extracted molecules in the droplet are ionized using ESI and then analyzed by MS. Thus, LESA enables the identification and mapping of biomolecules from the sample surface while preserving their native state (Hale & Cooper, 2021). LESA has been already employed for MPs studies. For instance, Cooper's team detected the intact Ammonium transporter (i.e., AmtB), which is a trimeric membrane protein of *circa* 140 kDa, from dried spots from a solution that contained the membrane protein solubilized with octyltetraglycol (C8E4) in ammonium acetate (Mikhailov et al., 2017). Thus, LESA can provide the opportunity to investigate membrane proteins and their interactions with other molecules, such as drugs or lipids from biological samples (Griffiths et al., 2015; Hale et al., 2020). Although the sensitivity of LESA is limited, it provides spatial information not accessible by other methods.

#### ***d) Laser Induced Liquid Bead Ion Desorption (LILBID)***

LILBID (Laser Induced Liquid Bead Ion Desorption) is a soft ionization technique introduced in 2006 by Morgner and colleagues (Morgner et al., 2006). In LILBID, an on-demand droplet generator produces small droplets (approximately 50  $\mu\text{M}$ ) of the sample solution containing the analyte. These droplets are then exposed to IR laser pulses. The laser energy is absorbed by water molecules, leading to an 'explosive expansion' of the droplet, releasing the analyte as ions into the gas phase (Morgner et al., 2006, 2007).

LILBID has shown to be a valuable tool for analyzing MPs in their native conditions. A proof of concept was done by Morgner *et al.* (2007), where intact complexes such as the Complex III (cytochrome bc1 complex) and Complex IV (cytochrome c oxidase) from the respiratory chain of *Paracoccus denitrificans* were ionized by LILBID. In recent years, LILBID, in conjunction with nanodiscs, has enabled to obtain structural details of complexes such as the yeast phospholipid methyltransferase Opi3 and the *Krokinobacter eikastus*' ion pump, rhodopsin-2 (KR2) (Henrich et al., 2017, 2018).

LILBID offers several advantages. For example, it can be coupled with a Time-of-Flight (TOF) analyzer to characterize large molecular assemblies. Additionally, thanks to new on-demand droplet generators, it requires only small sample volumes. Moreover, the IR laser energy can be tuned to provide insights into the intact complex (at low laser power), partially dissociated complexes (at intermediate laser power), or individual components present in the complex (at high laser power) (Hellwig et al., 2022). On the other hand, this technique is not widely applied in laboratories due to the requirement for in-house instruments that are not commercially available (Hellwig et al., 2022).

#### ***e) MALDI***

Matrix-Assisted Laser Desorption Ionization or MALDI is a soft ionization technique developed by Karas and Hillenkamp starting with analysis of alanine and then peptides (Karas et al., 1985; Karas & Hillenkamp, 1988). MALDI is a powerful method to ionize with a high sensitivity several non-volatiles and thermally labile analytes (e.g., proteins and oligonucleotides) of high molecular weight. The first breakthrough report of a ~34 kDa protein analysis came from Tanaka and coworkers (Tanaka et al., 1988) who used cobalt nanoparticles in glycerol as 'matrix', quickly followed by the demonstration by the Karas group of MALDI ionization and analysis of proteins up to ~67 kDa using a small organic molecule as matrix (Karas et al., 1989). In the last 40 years, the MALDI principle has not changed, however, the number of based-MALDI applications for analyzing biomolecules has dramatically increased (Tholey & Heinzle, 2006).

The physicochemical principles that govern this technique have been a challenge to explain. However, there are no doubts about the role that the matrix plays in the ionization process. In conventional solid spot MALDI, the “matrix” is a small organic acid or base that is mixed with the analyte and forms crystals as it dries. The matrix has an aromatic moiety that gives it the ability to absorb UV energy at the wavelength of a laser, usually with a fairly high absorption coefficient. Once spotted onto a sample stage, the matrix-analyte co-crystal is shot with a laser pulse, which lasts for nanoseconds. The absorbed photon energy creates a hot state of matter resulting in sputtering, i.e., a microexplosion that effectively propels the matrix into the gas-phase in a "plume" of excited ions and molecules, a phenomenon called desorption. The analyte is thus sublimated along with the matrix. The matrix lowers the threshold of laser energy that is necessary to ionize the analyte, thus preserving the analyte as an intact molecule through the analysis (Bolbach, 2005; Hillenkamp & Karas, 2007).

There are two competing models to explain how the charge is transferred to the analyte in MALDI. (i) the earliest model proposes that, inside the plume, protonated matrix molecules (primary ions) transfer the charge to neutral analyte molecules, which results in protonated analyte molecules (secondary ion). (ii) the later “lucky survivor or cluster model” proposes that analyte molecules incorporated into the matrix are pre-charged species surrounded by matrix counterions, as found in solution. After the laser pulse, clusters in the plume separate into positively and negatively charged ions. However, most of the analyte molecules become re-neutralized by recombining with photoelectrons or counter ions. Thus, the ions detected are only those that could maintain their charge until their arrival at the detector, i.e., "lucky survivors" (Bolbach, 2005; Hillenkamp & Karas, 2007). In any case, since ions are generated by a laser pulse, this technique is suitable to be coupled with a TOF analyzer. This fact allows us to sort and detect analytes in a wide mass range, from small and simple organic molecules to large biomolecules. For instance, MALDI-TOF instruments complemented with a special high mass detector for (e.g., from CovalX) have allowed for the detection of molecules up to 2 MDa (Shimadzu Corporation, 2022).

Although less widespread in MS instruments and less versatile than ESI sources in terms of hyphenated methods such as LC-MS and CE-MS, MALDI presents some advantages over ESI. First, it has a higher tolerance to salts, detergents, and other buffer contaminants, which makes a technique suitable for MPs (Beaufour et al., 2018; Gabant & Cadene, 2008). Second, it generates fewer charge states, which is an advantage for large proteins as it keeps the signal distributed over a narrower range of peaks, potentially resulting in a higher sensitivity of detection. Lastly, thanks to its ability to provide high spatial resolution images of biological tissues, MALDI is often used for imaging mass spectrometry (IMS), where a spectrum is acquired at each "pixel" of a tissue section.

Classically, both ESI and MALDI are applied in denaturing conditions and use organic solvents to facilitate solubilization and ionization of analytes. In addition, MALDI uses acidic or basic pH matrices which can affect the conformation of a protein. For these reasons, the direct detection of noncovalent interactions is the exception rather than the rule. So far, the analysis of biological macromolecules in the native state has been mainly developed with electrospray ionization (Weidmann et al., 2013). In principle, however, MALDI can be made to be non-denaturing just as is done for ESI for nMS. An original MALDI method based on solid spots was developed in the early 2000s for the detection of native complexes (Bolbach, 2005). However, because the analytes are in solution initially, this requires the transition through first a solid phase, followed by a sublimation into the gas phase. This dual phase transition raises questions about the integrity of the conformation of the protein, as well as of complexes. Possibly for this reason, solid spot-based MALDI methods have not gained traction in the field of nMS.

#### *f) NALIM*

To overcome the limitation of solid-spot MALDI for nMS, a new method called NALIM for NAtive LIquid MALDI was conceived by Beaufour and coworkers in 2018, although the acronym was conceived after this publication. NALIM leverages MALDI's unique properties into a new nMS method based on a liquid spot. Thus, unlike conventional MALDI, NALIM avoids the transition through a solid state.

The success of this method for native conditions relies on an ionic liquid matrix made of HCCA/3-AQ/Glycerol (1:4:6 w/w) with a pH near physiological conditions, which circumvents the acidic pH of traditional matrices. Moreover, glycerol provides stability to the analyte under vacuum conditions (Beaufour et al., 2018). The fact that MALDI is more tolerant to detergents, combined with the ability of NALIM to keep proteins in their native state, opens up new possibilities and provides the chance to characterize MPs and their complexes. In addition, since NALIM is coupled to a TOF analyzer, there is no theoretical limit on the size of the analyte. Thus, NALIM also opens the door to analyzing large complexes.

In summary, NALIM could be an important biophysical tool that, in complement with other structural biology techniques, can help to reveal important details of difficult biological systems.

### **3. Native MS (nMS) advantages for structural characterization of MPs and large molecular complexes that motivated this work.**

As detailed previously, there is a variety of different structural techniques that provide a wide range of information about biomolecules and their complexes. When it comes to membrane proteins, it's worth mentioning that - despite their limitations - NMR, cryo-EM, and, especially, X-ray crystallography, have been the techniques that have enabled significant progress in allowing for a comprehensive understanding of their structure and function. For instance, the structure of the first potassium ion channel was elucidated thanks to X-ray crystallography (Doyle et al., 1998). The first structure of the F<sub>1</sub>-ATPase was revealed by cryo-EM in 1989 by Gogol and colleagues, and in the subsequent years, it was complemented with more molecular details obtained thanks to NMR (Wilkins & Capaldi, 1998) and X-ray crystallography (Abrahams et al., 1994; Bianchet et al., 1991). Similarly, the first structure of bacteriorhodopsin was revealed by cryo-EM in 1990 by Henderson and colleagues, and then complemented with structures at atomic resolution (1.55 Å) level by X-ray crystallography such as the work performed by Luecke and colleagues in 1999. In terms of resolution, cryo-EM has recently been making great strides in catching up with X-ray crystallography.

While these structural biophysics methods have indeed offered invaluable structural insights into large molecular complexes and membrane proteins, the advantages of nMS as a set of straightforward, fast, and sensitive methods make it a precious complementary tool for characterizing these biomolecules.

In the cell, proteins form dynamic and transient biomolecular complexes, which can range from a few Kilodalton to the Megadalton scale. These complexes can present different stoichiometries or conformations. Furthermore, since each of their subunits can have different isoforms (e.g., different PTMs), large molecular complexes can be quite heterogeneous. Unlike other structural techniques, where the size of these large samples can arise issues such as low-resolution structures or preventing the crystallization, nMS can analyze these large complexes in a straightforward manner. For instance, using nMS, heterogeneous protein complexes larger than 10 Megadaltons have been successfully characterized (Shoemaker et al., 2010; Snijder et al., 2013).

nMS has also been an invaluable tool to characterize MPs, either in detergent micelles, liposomes, nanodiscs, or directly from native membrane vesicles (Barrera et al., 2008; Chorev et al., 2020; Ma et al., 2017). This tolerance to detergents and lipids has even been used to show the role of these molecules in the structure and functioning of MPs (Chorev et al., 2018; Gupta et al., 2017). In contrast to nMS, other structural techniques can be sensitive to membrane mimetics, which can interfere with the obtention of data or affect the structure resolution.

Finally, MPs are generally difficult to express and purify. This can be true of large complexes as well. Unlike traditional structural techniques, nMS requires only small amounts of samples, although the required quantities are still larger than MS in denaturing conditions. Thus, nMS is a suitable method to study these biomolecules.

In summary, all biophysical tools should be seen as complementary to each other in structural biology, with nMS being a crucial complementary one. For example, the information obtained through nMS can improve the experimental conditions for X-ray crystallography or NMR and also can help with the refinement of structures. Similarly, nMS can use other techniques to complement its results. For instance, nMS can determine the stoichiometry and composition of the complex but needs other techniques such as SAXS to obtain information about the shape and size that have this complex.

We will see in this work the complexity of studying membrane proteins and large molecular complexes in native conditions, the relevant biological information that can be obtained by nMS, as well as the different problems that we had to face to reach this goal.



# **OBJECTIVES**

***The general aim of this work*** is to get structural information about membrane protein complexes that is not easily, dynamically, or rapidly accessible by other biophysical methods in Structural Biology, and which can help to complete the picture obtained through these other methods. By gaining insight about the structure of these proteins and their complexes, we can better understand the mechanisms of action that underlie their role in key biological processes, such as cellular signaling, nutrient uptake, and drug transport, and help develop drugs to treat various diseases.

For this purpose, specific developments of a method leveraging the advantages of native MS and the MALDI ion source were necessary. The following objectives were defined at the start of the thesis and refined along the way:

**1. To obtain structural/functional information about membrane proteins and their complexes in native state such as stoichiometry, stability, and ligand binding**

To this end, a pre-established MALDI-based nMS method called NALIM needs to be developed specifically for membrane proteins and their interacting partners.

In addition, the following secondary objectives will help us to achieve our main objective.

- To evaluate the potential applicability of this method to different types of membrane protein complexes, such as transporters, ion channels, and receptor proteins.
- To extend the use of NALIM for a membrane protein complexes characterization in an *in vivo* lipid environment, since membrane mimics cannot provide all the characteristics of the native lipid bilayer.
- To optimize the method for resolution and sensitivity in a relatively high mass range.
- To establish a reliable calibrant based on a soluble system that will be easier to handle than a membrane protein and cover an adequate mass range up to 100 000 Da and higher.

*While pursuing these objectives, the search for a high mass calibrant led to the discovery of larger range of applicability than originally anticipated, leading to a new field of application for the NALIM method. Since large soluble oligomers are an essential class of complexes for biology, we decided to develop and take advantage of NALIM to get structural information on these systems. Thus, my second major objective was:*

**2. To characterize large soluble molecular complexes in order to determine their degree of oligomerization and the effect of ligand binding.**

Oligomers are essential structures through which a number of therapeutically relevant proteins perform their functions within the cell. Our objective was to take advantage of the TOF analyzer, which in principle has a limitless mass range, and of NALIM, to provide structure/function

information about these molecular complexes despite their large size, heterogeneity, and low available quantity.

## **RESULTS AND DISCUSSION**

During the development of this thesis, I participated in various conferences where I presented our results either as a poster or an oral communication.

### Posters:

**“Search for a high-mass calibrant for NALIM (Native Liquid MALDI MS): the case of mouse IgA.” Journées Français de Spectrométrie de Masse (JFSM) 2021.**

**Authors :** Zhamungui E<sup>1</sup>, Beaufour M<sup>1</sup>, Cadene M<sup>1</sup>

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR4301, affiliated with Université d'Orléans, Orléans 45071, France

Since we mainly observe low charge states (1<sup>+</sup> and 2<sup>+</sup>) in NALIM, the monocharged state being largely predominant, the application of NALIM to large membrane protein-containing assemblies requires calibrants for high  $m/z$  ratios. Here, we present the suitability and use of an IgA as a potential calibrant. The mass range we needed to cover was up to 320 kDa. Moreover, we presented the optimizations of the biochemical sample preparation, control of the degree of oligomerization, and instrumental setup for NALIM.

**“Large intact biomolecular assembly analysis by NALIM (Native liquid MALDI MS) shows non-covalent Lc binding in mouse IgA.” Biotechnocenter 2021.**

**Authors :** Zhamungui E<sup>1</sup>, Beaufour M<sup>1</sup>, Cadene M<sup>1</sup>

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR4301, affiliated with Université d'Orléans, Orléans 45071, France

The NALIM method was conceived for the study of membrane protein complexes in native state with a mass ranging from 100 to 300 kDa. Since we mainly observe low charge states (1<sup>+</sup> and 2<sup>+</sup>) in NALIM and due to the lack of commercial calibrants for native MALDI, we decided to create our own calibrant. Consequently, we focused on immunoglobulin A (IgA) because of its monomeric and dimeric forms, potentially covering a wide mass range (i.e., up to 320 kDa). Here, we describe how optimizations of the delay and matrix concentration allow for improved observation of dimer species at a high  $m/z$  range.

**“NALIM: Native Liquid MALDI Mass Spectrometry as a new structural method to directly characterize membrane protein complexes.” International Mass Spectrometry Conference (IMSC) 2022.**

**Authors :** Zhamungui E<sup>1</sup>, Orelle C<sup>2</sup>, Jault JM<sup>2</sup>, Arizala JD<sup>3</sup>, Poget SF<sup>3</sup>, Beaufour M<sup>1</sup>, Cadene M<sup>1</sup>.

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR4301, affiliated with Université d'Orléans, Orléans 45071, France.

<sup>2</sup>Molecular Microbiology & Structural Biochemistry (MMSB) UMR 5086, CNRS/University of Lyon, Lyon 69367, France

<sup>3</sup>Department of Chemistry, College of Staten Island, City University of New York, 2800 Victory Boulevard, New York, United States.

This poster talks about the versatile applications of NALIM to various types of membrane proteins. For instance, we show how NALIM provides structural information in a straightforward manner about the complex stoichiometry and ligand binding effect using the *Bacillus* multidrug resistance ABC transporter (i.e., BmrA). We showed preliminary data about a competition experiment to localize the binding site of the actitoxin Tx7335 onto the KcsA potassium channel.

### Oral communications:

#### “Structural biology of membrane proteins and their complexes using Native Liquid MALDI (NALIM) Mass Spectrometry (MS)”. *Analytica2022*

**Authors :** Zhamungui E, Orelle C, Jault JM, Arizala JD, Poget SF, Beaufour M, Cadene M.

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR4301, affiliated with Université d'Orléans, Orléans 45071, France.

<sup>2</sup>Molecular Microbiology & Structural Biochemistry (MMSB) UMR 5086, CNRS/University of Lyon, Lyon 69367, France

<sup>3</sup>Department of Chemistry, College of Staten Island, City University of New York, 2800 Victory Boulevard, New York, United States.

In this conference we presented how NALIM was successfully adapted and applied to address challenges related to the analysis of a multimeric ABC transporter which is an archetype for multi-drug resistance transporters (i.e., BmrA). In addition, we talked about a competition experiment which gives us insight into the localization of the actitoxin Tx7335 onto the KcsA potassium channel. Finally, we presented the preliminary data of the direct ejection and identification of BmrA dimer from nanovesicles, and how this provided insights for future work.

### Future publications:

The chapters of this thesis were written in the form of manuscripts, which will be publications in the future.

#### NALIM ANALYSIS OF LARGE SOLUBLE PROTEIN OLIGOMERS

**Authors :** Zhamungui E<sup>1</sup>, Mance L<sup>1</sup>, Suskiewicz MJ<sup>1</sup>, Boudvillain M<sup>1</sup>, Irving JA<sup>2</sup>, Beaufour M<sup>1</sup>, Cadène M<sup>1</sup>

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR4301, affiliated with Université d'Orléans, Orléans 45071, France.

<sup>2</sup>Cambridge Institute for Medical Research, Department of Medicine, University of Cambridge, Cambridge CB2 0XY, U.K.

The search for a NALIM's high-mass calibrant led us to the discovery of a protein that forms a nice molecular ladder through oligomerization, thus extending the range of analysis of NALIM. Here, we show the establishment of this new calibrant for high  $m/z$  ranges, and the instrumental and sample preparation optimizations that made it possible to get structural information about large soluble molecular complexes, which notably, perform essential functions in the cell, and are worthwhile objects of study in their own right.

## **NALIM (NAtive LIquid MALDI)-TOF MS for the characterization of membrane protein complexes.**

**Authors :** Zhamungui E<sup>1</sup>, Jaubert J<sup>1</sup>, Rostovsky I<sup>2</sup>, Arizala JD<sup>2</sup>, Poget SF<sup>2</sup>, Morisset Lopez S<sup>1</sup>, Orelle C<sup>3</sup>, Jault JM<sup>3</sup>, Beaufour M<sup>1</sup>, Cadene M<sup>1</sup>

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR4301, affiliated with Université d'Orléans, Orléans 45071, France.

<sup>2</sup>Molecular Microbiology & Structural Biochemistry (MMSB) UMR 5086, CNRS/University of Lyon, Lyon 69367, France.

<sup>3</sup>Department of Chemistry, College of Staten Island, City University of New York, 2800 Victory Boulevard, New York, United States.

Membrane proteins have different roles in the cell such as enzymes, transporters, or signaling receptors, which are of vital significance for the cell. As a consequence, membrane proteins represent 2/3rds of potential targets for drugs. Here we show NALIM development for membrane protein complexes. Moreover, we extend NALIM applications to the analysis of membrane protein complexes not only from membrane mimetics, but also directly from membrane preparations, opening the door to analyses in conditions that are ever closer to the live cell.

# CHAPTER 1

## CONTEXT

Proteins establish permanent or dynamic interactions with small organic molecules and larger biomolecules (e.g., nucleic acids, other proteins), either inside or outside the cell in order to perform different functions. Thus, these complexes can reach large sizes, show great heterogeneity, and be present in very low amounts. These features limit their structural characterization by classical biophysical methods such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, or cryo-EM.

Interestingly, the mass range of NALIM coupled with a Time of Flight (TOF) analyzer is in principle unlimited. However, since NALIM generates mainly monocharged ions, we were faced with the lack of a commercial calibrant for the high mass range (i.e., higher than 100 000 Da). Hence, part of our work focused on finding and validating a new calibrant that can be used in this mass range. Moreover, in order to improve the performance of this calibrant, instrumental parameters, and sample preparation conditions need to be optimized to improve the sensitivity and resolution of NALIM in the high mass range. Here we show some of the optimization strategies that we chose to apply.

The search for a high-mass calibrant led us to the discovery of a protein that forms a nice molecular ladder through oligomerization, thus extending the range of analysis. Hence, it appeared we could use NALIM to obtain structural information about large molecular complexes, which notably perform essential functions in the cell, and are worthwhile objects of study in their own right. To test the applicability of NALIM to large soluble complexes, we aimed to get insight into the interaction of Rho, a hexameric bacterial ATP-dependent helicase, with NusG, a transcription factor, as NusG may play a role in modulating the activity of the Rho protein. Modulating this interaction could be useful for drug development. Next, we aimed to analyze the ZBTB8A protein, a transcriptional repressor which acts on proximal promoters of the cyclin-dependent kinase gene, CDKN1A. Some proteins of the ZBTB family could play a role in cancer by affecting DNA repair. It has been suggested that in proteins of this family, particularly those from insects, monomers not only form dimers, but these dimers can interact to form oligomers. Hence, we endeavored to use NALIM to determine the degree of oligomerization that the ZBTB8A protein can achieve and to provide insight into the implications that this oligomerization can have on its functions.



# **NALIM ANALYSIS OF LARGE SOLUBLE PROTEIN OLIGOMERS**

## **Authors**

**Zhamungui E<sup>1</sup>, Mance L<sup>1</sup>, Suskiewicz MJ<sup>1</sup>, Boudvillain M<sup>1</sup>, Irving JA<sup>2</sup>, Beaufour M<sup>1</sup>, Cadène M<sup>1</sup>**

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR4301, affiliated with Université d'Orléans, Orléans 45071, France.

<sup>2</sup> Cambridge Institute for Medical Research, Department of Medicine, University of Cambridge, Cambridge CB2 0XY, U.K.

## INTRODUCTION

In order to fulfill its function(s), a protein needs to interact with small organic molecules and larger biomolecules (e.g., nucleic acids, other proteins), either inside or outside the cell. Protein function is generally carried out through involvement in dynamic noncovalent complexes that range from small to large sizes. Any method that purports to dissect such complexes with the aim of understanding their basic mechanisms must preserve these noncovalent interactions. In other words, such methods must operate under conditions that closely resemble the complexes' natural environment (Leney & Heck, 2017).

In mass spectrometry (MS), different approaches have been developed to directly measure complexes in the instrument using controlled, non-denaturing *in vitro* conditions to better mimic the protein's biological environment. These methods are collectively known as native MS (nMS) and have highly contributed to biological and biophysical information of proteins and their complexes (Barrera et al., 2008; Chorev et al., 2020; Fitzgerald et al., 1996; Hernández & Robinson, 2007).

MS analysis of biomolecules was made possible thanks to the invention of soft ionization methods (i.e., ESI and MALDI). The first ESI-based method for nMS was developed in the early 1990s (Katta & Chait, 1991) and over time this type of method became the cornerstone of nMS by generating ions directly from a non-denaturing solution. In contrast, MALDI methods for nMS took a lot longer to develop. This can be explained by a greater set of hurdles to be overcome in order to make MALDI conditions non-denaturing. Classical MALDI sample preparation requires dissolving a small organic acid (the "matrix") in an organic, denaturing solvent, followed by the co-crystallization of analytes and matrix. Moreover, early developments of MALDI nMS focused on solid deposits, which imply two phase transitions from the sample solution to the gas-phase ions (Leney & Heck, 2017). Thus, Beaufour *et al.* (2018), developed a MALDI nMS method based on an ionic liquid formed of an acidic and a basic matrix to better mimic native conditions and to skip one phase transition.

Even though ESI is the go-to method for nMS, not all samples are easily adaptable to this ionization. In a previous work, this method was extended to the study of membrane proteins (MPs) and renamed as NALIM for Native Liquid MALDI. In the work described here, NALIM is further developed for the characterization of large biomolecular complexes.

Since large biomolecular complexes are frequently heterogeneous, present in low quantities in the cell, and their interaction could be transient, they are not easily characterized by classical biophysical methods such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, or cryo-EM (Ruotolo et al., 2008). Furthermore, since these methods require higher quantities of sample than NALIM, the NALIM sensitive may present advantages to characterize large biomolecular complexes.

Additionally, nMS based on ESI is generally coupled with analyzers with a limited  $m/z$  range by design. For example, quadrupoles scan using sets of RF/DC voltages that keep ions stable in the path of the quadrupole only within set  $m/z$  limits. In contrast, the mass range of NALIM coupled with a Time of Flight (TOF) analyzer is in principle unlimited, so that biomolecular complexes can be analyzed simultaneously over a wide mass range. However, it is known that with MALDI or NALIM coupled with TOF analyzers, instrumental parameters must be adjusted for optimum ionization and detection of high-mass ions. Furthermore, as NALIM generates mainly monocharged ions, this faced us with the lack of commercial calibrants for native MS in MALDI-TOF instruments.

Currently the market offers just one product designed for denaturing conditions MALDI, i.e., CovalX which covers a mass range of up to 450 000. However, this product was not available at the start of this work. Besides this product, a calibration kit made for electrophoresis, i.e., ThermoFisher's NativeMark™, covers masses up to 1.2 MDa. This type of marker was however not designed for mass spectrometry, which means that it may lack the homogeneity needed for calibration, and it may contain additives that are incompatible with ionization in nMS. Thus, we endeavored to search for appropriate soluble protein complexes on which instrumental parameters could be optimized to cover the desired mass range. Two proteins were considered as potential high-mass calibrants in NALIM: a mouse IgA, which in theory could reach a mass range of up to 320 000, and a serine protease inhibitor, alpha-1-antitrypsin ( $\alpha$ 1AT) with a mass of 45 550 Da, which can form oligomers under certain conditions.

In the present work, we have decided to focus on large biomolecular complexes involved in the regulation of transcription, such as transcription factors (TFs) and a DNA helicase. TFs can form dynamic complexes with other biomolecules to control gene expression, which will influence differentiation, cycle, and the fate of the cell (Lambert et al., 2018). One TF can control different genes in different cells in one organism (Gertz et al., 2012), and its malfunctioning or mutation is associated with diseases for the organism (Lambert et al., 2018). However, despite their main importance for the cell and extensive structural and functional knowledge about them, our understanding of fundamental aspects of their assembly into larger structures and the role these structures play is still limited. Thus, ZBTB8A transcription factors were studied by NALIM to get structural insight. The NusG TF is a bacterial protein which may play a role in modulating the hexameric Rho helicase, a process that could be targeted for antibiotic purposes. Thus, we also focused on the Rho/NusG complex.

NusG/Spt5 proteins belong to a family of TFs conserved across all organisms. While NusG is found in bacteria, Spt5 is its corresponding homolog in archaea and eukaryotes (Lawson et al., 2018). NusG

can play two opposite roles in transcription. Its N-terminal domain (NTD) can bind the RNA polymerase (RNAP) in order to enhance the RNAP transcription activity (i.e., antitermination). In contrast, the C-terminal domain (CTD) can recruit accessory factors to promote the end of the transcription (i.e., termination) (Burmam et al., 2010).

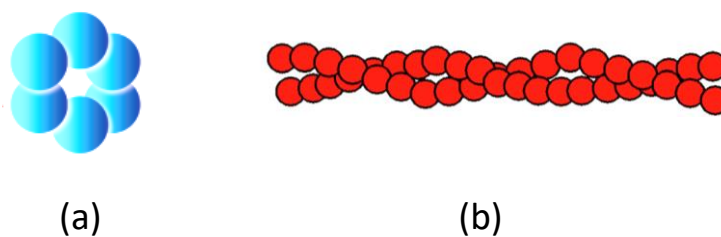
On the other hand, Rho is an ATP-dependent helicase associated with a process for the termination of transcription found exclusively in bacteria. Rho recognizes and binds sequences of *circa* 80 nucleotides with a high pyrimidine content (especially cytosines) in the nascent RNA. These sequences are known as Rho utilization (RUT) sites (Lawson et al., 2018; Peters et al., 2012). In the absence of ATP and prior to RUT recognition, Rho is a homohexamer with an open-ring conformation. In this open conformation, Rho binds RUT sites, and it drives the nascent RNA inside the ring. In the presence of ATP, Rho switches to a closed-ring conformation, trapping the RNA inside the central pore of the ring. Then, Rho moves downstream, using energy from ATP hydrolysis, until it encounters the RNAP, which will stop transcription (Lawson et al., 2018).

Interestingly, Rho can work on RNAs that bear few RUT sites or even no site, and the mechanism for this is not well understood. It has been suggested that NusG plays a key role by allowing the Rho-dependent termination of transcription in a sequence-independent manner. If one considers a stoichiometry of one NusG protein by Rho hexamer, then the mass of this molecular complex in the native state is close to 305 000 Da.

The ZBTB8A transcription factor, also known as BOZF1, belongs to the ZBTB family of TFs, which is defined by an N-terminal BTB (broad-complex, tramtrack and bric-à-brac) domain and C-terminal zinc fingers. According to Kim *et al.* (2013), ZBTB8A is a transcriptional repressor that acts on proximal promoters of the cyclin-dependent kinase gene, *CDKN1A*, inhibiting its transcription. This leads to downstream transcription repression of some regulatory proteins related to apoptosis, cell differentiation, cell cycle progression, and oncogenesis, suggesting a possible link with tumors and metastasis. Indeed, according to Kim *et al.* (2013), ZBTB8A is highly expressed in different cancer tissues. Furthermore, ZBTB proteins are frequently substrates for SUMOylation in which small ubiquitin-like modifiers (SUMO) are covalently linked to the substrate protein. This reversible post-translational modification (PTM) is directly associated with key cellular processes such as DNA replication and reparation, transcriptional regulation, cell cycle progression, signal transduction, and more (Hendriks et al., 2014; Matic et al., 2010).

ZBTB proteins can dimerize thanks to the BTB domain, which facilitates the DNA-protein binding mediated by the CTD's zinc fingers (Kim et al., 2013). In some ZBTB proteins, particularly those from insects, interaction between dimers have been suggested. Recent findings by Mance and collaborators (currently in publication) suggest that the BTB domain of ZBTB8A can generate large

filaments with unknown stoichiometry. Here, these assemblies were analyzed with NALIM (Figure 1).



**Figure 1.** Many biological processes are driven by oligomers resulting from noncovalent association between proteins, which can be divided into 2 classes based on composition: 1. non-covalent homo-oligomerization of identical protein subunits 2. formation of non-covalent protein heterocomplexes built of different proteins. Oligomers can organize into higher-order structures such as rings (a) or filaments (b), which can themselves form super-structures such as fibrils.

In summary, this work shows the instrumental optimization as well as the search for a high mass calibrant suitable for high-mass nMS, and how information about the structural stability of the studied systems was obtained from NALIM analysis of these large molecular complexes.

## MATERIALS AND METHODS

### Materials

Mouse immunoglobulin A (IgA) anti-CD19 was obtained from antibodies-online GmbH (Aachen, Germany). Rho and NusG proteins were provided by Dr. Marc Boudvillain. They were expressed in MG1655 *E. coli* cells and purified as previously described (Nadiras et al., 2018). ZBTB8A WT and mutant S103R proteins were provided by Dr. Marcin Suskiewicz. They were expressed in Rosetta<sup>TM</sup> 2 *E. coli* cells and purified as previously described (Mance et al., 2023 [currently in publication]). Alpha-1-antitrypsin ( $\alpha$ 1AT)  $\alpha$ 1AT by Dr. James Irving. It was expressed in BL21 (DE3) cells and purified as previously described (Irving et al., 2011). The quality control of proteins was performed in denaturing conditions by MALDI-TOF MS using the UltraThin Layer (UTL) method. In brief, we first prepared a 1:2 solution of water and acetonitrile. This solvent solution was then saturated with the HCCA matrix. After homogenizing, this solution was further diluted with isopropyl alcohol at a ratio of 1/4. About 20  $\mu$ L of this ultrathin-layer solution were applied on the gold plate and spread across its entire surface. This layer of diluted matrix was then air-dried until only minimal moisture remained, and the excess matrix removed by wiping the plate with a tissue. As a result, an 'ultrathin layer' of HCCA remains at the surface of the gold plate, which is hardly visible but serves as a seeding layer for deposits.

Upon this ultrathin layer, half a microliter of sample diluted in a saturated matrix solution is deposited. The excess liquid containing contaminants can be aspirated when the so-called matrix-sample 'co-crystal' begins to form on the surface of the plate. The plate is then introduced into the MALDI source for sample analysis. The UTL method generates multicharged species generally centered on the 7+ to 3+ charge state, depending on protein size and ionizability (Cadene & Chait, 2000; Fenyó et al., 2007; Gabant & Cadene, 2008).

Peptide-N-glycosidase F (PNGase F) was from Roche (Basel, Switzerland). Ammonium acetate (AA) was from Merck (Darmstadt, Germany). The 4-hydroxy- $\alpha$ -cinnamic acid (HCCA) matrix was from Bruker Daltonics (Bremen, Germany). The 3-aminoquinoline (3-AQ) matrix, ATP disodium hydrate, urea, glycine, and glycerol were purchased from Sigma-Aldrich (St. Louis, MA, USA). The organic solvents acetonitrile and methanol were purchased from Biosolve (Dieuze, France). All aqueous solutions and buffers were prepared using 18 M $\Omega$  water purified with Milli-Q reagent grade system from Millipore (Bedford, MA, USA), herein referred to as "ultrapure water."

## Sample preparation

The mouse IgA was denatured with 2M urea and 0.2M glycine for 10 minutes at 60 °C. Then, it was incubated with 300 U/mg of PNGase F at 37 °C for 24 hours. Finally, the IgA was concentrated, and the urea was removed by buffer exchanged with 100 mM AA, pH 7.5, using an Amicon device with a cutoff of 100 kDa. The centrifugal concentrator and filter devices (Amicon®) were from Merck (Darmstadt, Germany)

Rho/NusG/ATP complex was prepared as detailed in the results section and then desalted by buffer exchanged with 10 mM ammonium acetate (glycerol/water 50:50, v/v) pH 7.9, using a Zeba™ device of 75 µL with a cutoff of 7 kDa. The Spin Desalting Columns (Zeba™) was from Fisher Scientific (Ottawa, Canada).

Intermediate dilutions of ZBTB8A proteins (i.e., WT and mutant) and  $\alpha$ 1AT were respectively prepared in 50 mM AA (glycerol/water 50:50) and 100 mM AA buffers, both at pH 7.5. Dilutions were made until achieve a concentration four times higher than the final expected to be analyzed.

## NALIM MS

The HCCA/3AQ/Glycerol (1:4:6 w/w) ionic liquid matrix, herein referred to as “HAG”, was prepared according to Beaufour *et al.* (2018). All samples prepared as described above were diluted at a ratio 1/4 in HAG (herein referred to as the “matrix/sample solution”). A 1 µL aliquot of the matrix/sample solution was spotted onto a MTX stainless steel sample stage (Bruker Daltonics, Germany). The drop was left for 15 minutes at room temperature to evaporate the water part of the solution. This time is defined as droplet evaporation time ( $t_{de}$ ). The time that the matrix/sample spot spends in the source vacuum before the acquisition, defined as *in vacuo* residence time ( $t_{ivr}$ ) by Beaufour *et al.* (2018), was 15 minutes. NALIM MS analyses were performed using an UltrafleXtreme MALDI-TOF/TOF mass spectrometer from Bruker equipped with a 2 kHz smartbeam-II laser and PAN™ technology. Spectra were acquired with a 20-kV acceleration voltage in linear positive mode using a single-shot pattern and accumulating of 200 or 400 laser shots. An ion deflection of 2 000 Da below the starting  $m/z$  acquisition value was applied. A delay of 2 700 ns was applied for large soluble complex experiments. The instrument was controlled using Bruker FlexControl software v4. NALIM MS spectra were processed using FlexAnalysis software v4 (Bruker). For proteins quality control, Protein A, apomyoglobin, betalactoglobulin, cytochrome c (all from Sigma-Aldrich), and a deglycosylated IgG of known mass were used for calibration in the 4 000 and 145 000  $m/z$  range. For NALIM experiments external calibration in the 45 000 and 300 000  $m/z$  range was performed using alpha-1-antitrypsin ( $\alpha$ 1AT).

## RESULTS

### Searching for a high $m/z$ calibrant for large complexes in NALIM

In order to develop the NALIM method so that it will help biologists gain a better understanding of large molecular soluble complexes, such as oligomers involved in the regulation of transcription, the first requirement is to have appropriate calibrants in a higher  $m/z$  range than is routinely observed in MALTI-TOF MS. Since molecular ions are mainly monocharged in NALIM, calibrants must have a molecular mass of 100 000 Da or higher so as to be compatible with the analysis of oligomer complexes. For native mass spectrometry, to the best of our knowledge, such calibrants are not commercially available. Thus, extending NALIM applications for a high  $m/z$  range required discovering new high-mass calibrants, and optimizing instrumental parameters for their implementation.

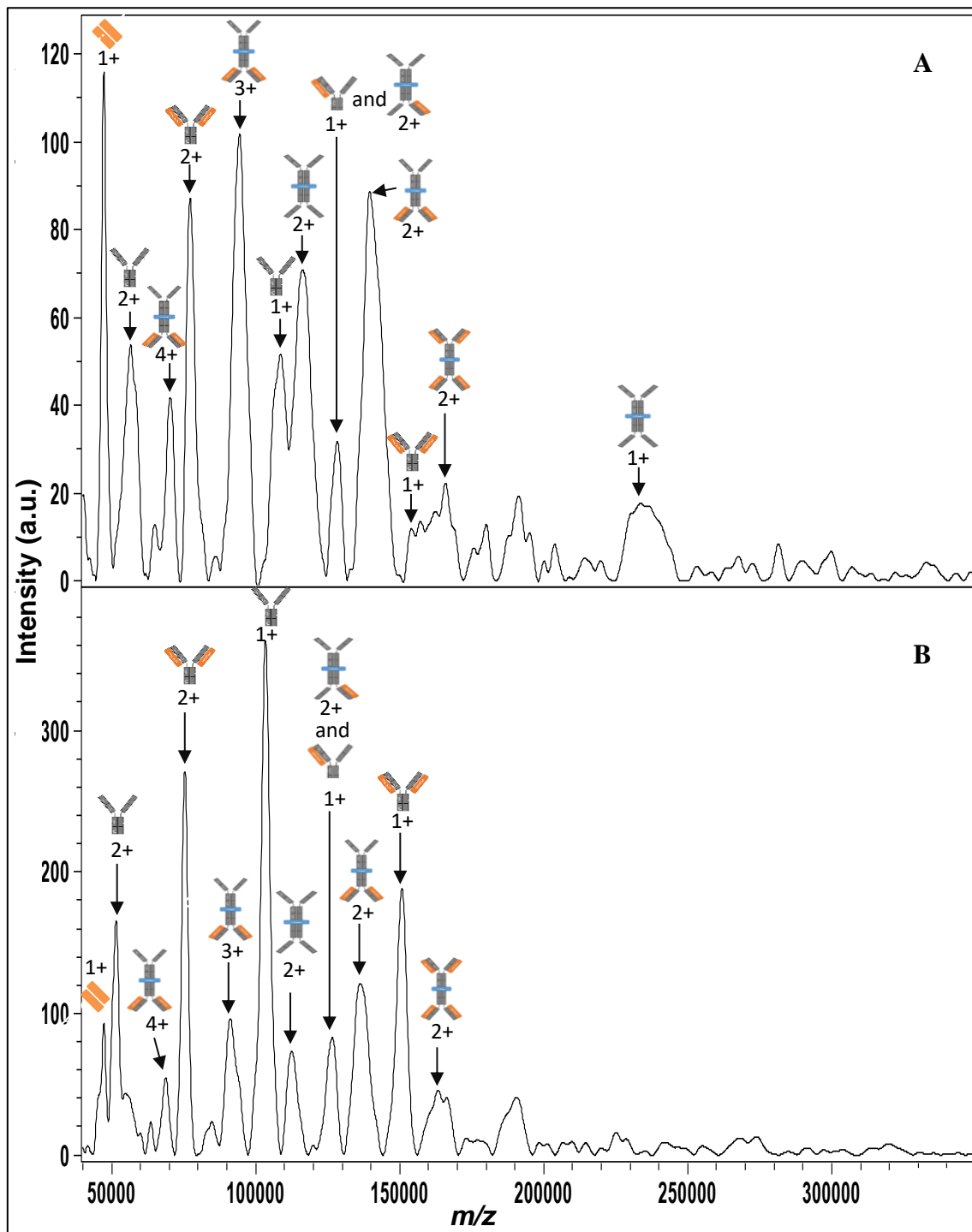
### Delay optimization

In MALDI, intense, resolved, and homogenous calibrant peaks ensure proper calibration. However, instrumental limitations lead to degraded intensity and resolution of peaks in the high  $m/z$  range. Causes for the limitations in this range include constraints of instrument components such as detectors, and/or inadequate tuning of instrumental parameters. Hence, we attempted to improve the signal-to-noise (S/N) ratio and the resolution of high  $m/z$  peaks by focusing on delay optimization.

### IgA as a potential calibrant

Since assemblies of proteins studied in this work were expected to reach a theoretical mass of 300 000 Da, we considered different proteins as potential high-mass calibrants. IgA, as a dimer of an immunoglobulin (Ig), was thought to readily reach this range. However, commercial IgAs available at this time were glycosylated, prompting us to attempt deglycosylation of the protein prior to optimization experiments. Recombinant IgGs are routinely deglycosylated to completion using a 24h incubation with Peptide-N-glycosidase F (PNGase F) (Y. Huang & Orlando, 2017; Qian et al., 2007; Tarentino et al., 1985). Accordingly, we attempted deglycosylation of a mouse IgA (anti-CD19) with PNGase F and assessed the degree of deglycosylation using the UTL method in denaturing conditions (Figure 2) (Cadene & Chait, 2000; Gabant & Cadene, 2008).













**Figure 2.** UTL - MALDI TOF MS analysis in denaturing conditions of the IgA deglycosylation with PNGase F in HCCA/FWI. **A.** Control: IgA before deglycosylation at 0.3  $\mu$ M **B.** IgA at 0.3  $\mu$ M deglycosylated with PNGase F (300 U/mg of protein) for 24h in borate buffered saline at pH 8.3. Spectra were accumulated over 200 shots.







As can be seen in panel A, the mouse IgA is found as monomer or dimer. Furthermore, heavy, and light chains can be bound in a covalent or noncovalent manner. This explains the different IgA species found in the quality control under denaturing conditions (Table 1).

**Table 1.** Schematic representation of all possible IgA species.

<b><i>Schematic representation</i></b>	<b>Description</b>
	IgA monomer
	IgA monomer + 1 light chain
	IgA monomer + 2 light chains
	IgA dimer
	IgA dimer + 1 light chain
	IgA dimer + 2 light chains
	IgA dimer + 3 light chains
	IgA dimer + 4 light chains

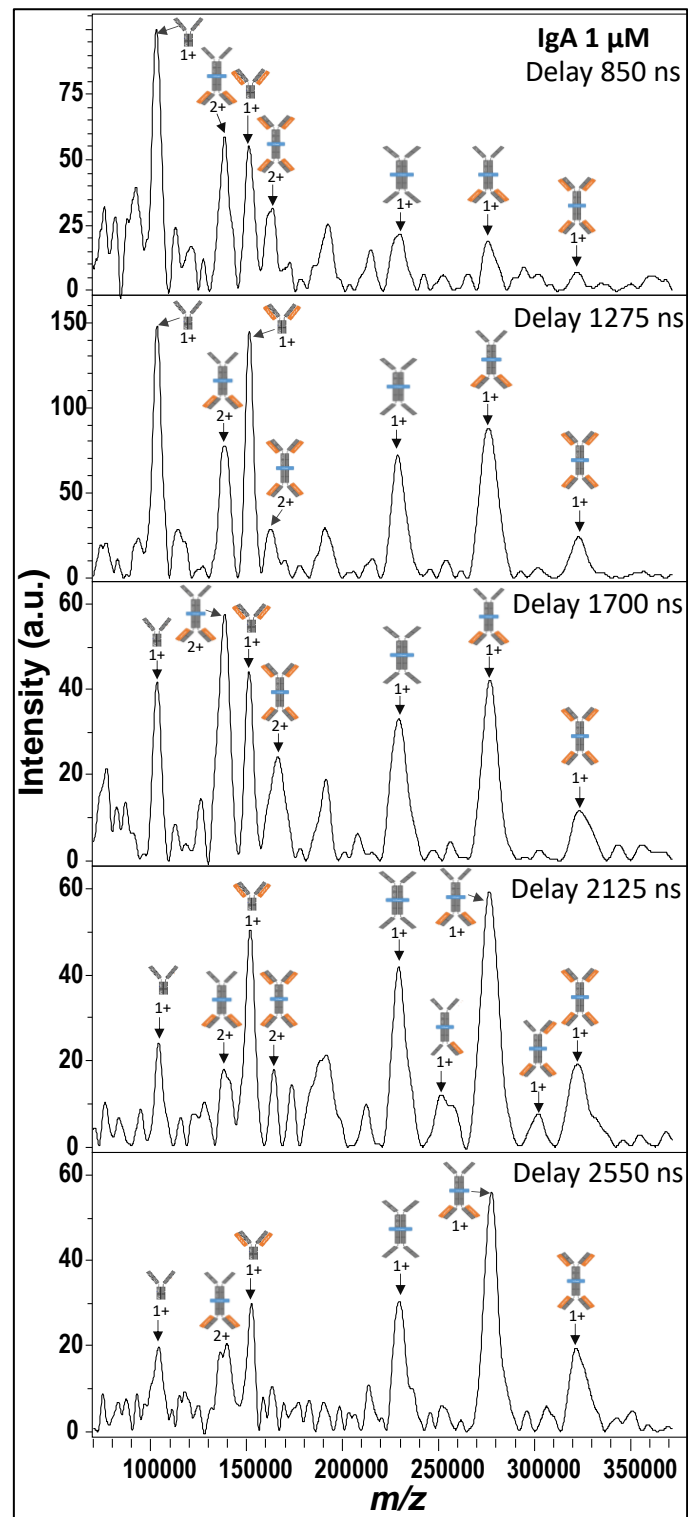
Panel B in Figure 2 shows that IgA monomer peaks become narrower and more intense upon deglycosylation, indicating that PNGase F is active against monomers. However, IgA dimer peaks stay wider than IgA monomer peaks. While the mass of the monomer plus two light chains decreased by 3 458 Da after deglycosylation, the corresponding dimer decreased by just 4 994 Da. Thus, although the dimer has twice as many glycans, the dimer lost only 44% more mass than the monomer (Table 2). Furthermore, the mass decrease after deglycosylation shows that the higher the number of light chains, the lower the number of glycans removed. Consequently, the light and heavy chains stoichiometry is directly related to the degree of IgA deglycosylation, which suggests an incomplete deglycosylation of dimer species. One possible explanation is that the IgA dimer conformation prevented the glycosidase PNGase F from reaching some glycosylation sites due to steric hindrance. Additionally, monomer IgA peaks become the highest intensity peaks in the spectrum. This could just be a reflection of the higher ionization of deglycosylated species, but it could also indicate that deglycosylation affects the stability of dimers in solution, which can be observed as a decrease in dimers' peak intensities.

**Table 2.** IgA mass determination before and after deglycosylation.

<b>Schematic representation</b>	<b>Mass before deglycosylation (Da)</b>	<b>Mass after deglycosylation (Da)</b>	<b>Mass shift (Da)</b>	<b>Estimation* of glycans removed</b>
	47349	47342	7	0.0
	108657	103460	5197	2.2
	154288	150829	3458	1.4
	232836	225177	7660	3.2
	280190	273213	6977	2.9
	331829	326835	4994	2.1

\*The calculation of the number of removed glycans is based on the average mass of glycans commonly present in the mouse IgA (i.e., 2 400 Da).






**Delay optimization using a deglycosylated IgA.** After deglycosylation, we tested the effect on the S/N ratio and the resolution of five delay values ranging from 850 to 2 550 ns in 425 ns increments. Four IgA species were selected for optimization to cover the 100 000 to 320 000  $m/z$  range. As they were consistently observed in all spectra, the effect of the delay could be assessed in parallel with a wide  $m/z$  range (Figure 3).



**Figure 3.** NALIM – MALDI TOF MS analysis of the effect of the delay on the S/N ratio and resolution of high  $m/z$  peaks. A deglycosylated IgA was analyzed at 1  $\mu\text{M}$ . The sample/matrix solution contains HAG 50%, and water 50%. Spectra were accumulated over 200 shots.

As seen in Figure 3, the higher the delay value, the better the sensitivity for large molecular complexes (i.e., IgA dimer species), which is confirmed by an increase in the S/N ratio value (Table 3).

**Table 3.** Signal to noise (S/N) ratio of four IgA species' peaks in NALIM.






<i>m/z</i>	<i>Antibody observed</i>	DELAY (ns)				
		850	1275	1700	2125	2550
105,000		8	10	6	4	5
150,000		5	10	6	8	6
230,000		6	11	19	7	9
275,000		6	18	20	23	21
320,000		4	7	5	9	8

In yellow. The highest S/N ratio value for the most frequently assigned *m/z* peaks.

At 1 275 ns, ions lower than 150 000 *m/z* reach the highest S/N ratio value. Delays higher than 1 275 ns are detrimental to their S/N ratio. This suggests that, in spite of the ramping feature in the UltrafleXtreme® instrument, at a high *m/z* range the optimum delay is more specific to a defined *m/z* range.

For ions higher than 150 000 *m/z*, although the delay improves the S/N ratio, values show no clear general trend. This result may be directly related to the incomplete deglycosylation of dimer IgA species. The heterogeneity of these peaks affects their intensity and resolution (Table 4)

**Table 4.** Resolution of IgA peaks in NALIM.

<i>m/z</i>	<i>Antibody observed</i>	DELAY (ns)				
		850	1275	1700	2125	2550
105 000		19	20	28	24	21
150 000		24	31	33	28	30
230 000		27	30	23	31	29
275 000		36	25	33	29	34
320 000		49	38	27	32	35

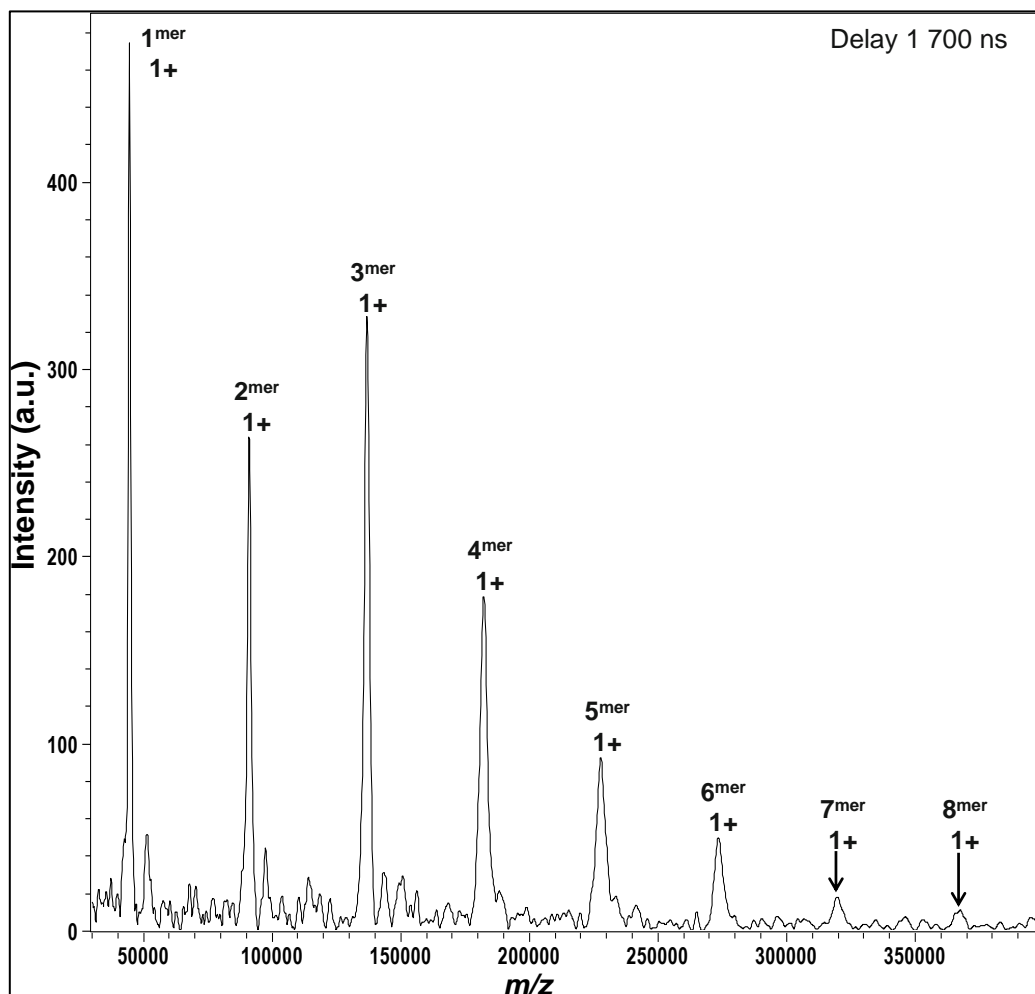
In blue. The highest resolution value for the most frequently assigned *m/z* peaks.

Interestingly and contrary to the S/N ratio, a delay higher than 1 275 ns improves the resolution of peaks lower than 150 000 *m/z*, 1 700 ns being the most suitable. Thus, for a *m/z* range from 100 000 to 150 000, the delay that guarantees the best compromise between intensity and resolution is in a range from 1 275 ns to 1 700 ns. However, for ions over 150 000 *m/z*, a delay higher than 1 700 ns is still needed.

Unfortunately, as above mentioned, the mouse IgA deglycosylation was incomplete. This fact limited its use as a calibrant for a mass range higher than 150 000 Da and prompted us to look for an alternative calibrant.

### Alpha-1-antitrypsin ( $\alpha$ 1AT) as an alternative calibrant for high masses

Based on previous results with IgA, we focused our work on recombinant proteins expressed in bacteria to circumvent glycosylation. Alpha-1-antitrypsin ( $\alpha$ 1AT), an inhibitor of serine proteases expressed in *E. coli*, was evaluated as a calibrant for NALIM. As our goal was to detect ions with  $m/z$  ratios higher than 150 000, the initial delay tested with  $\alpha$ 1AT was 1 700 ns (Figure 4).

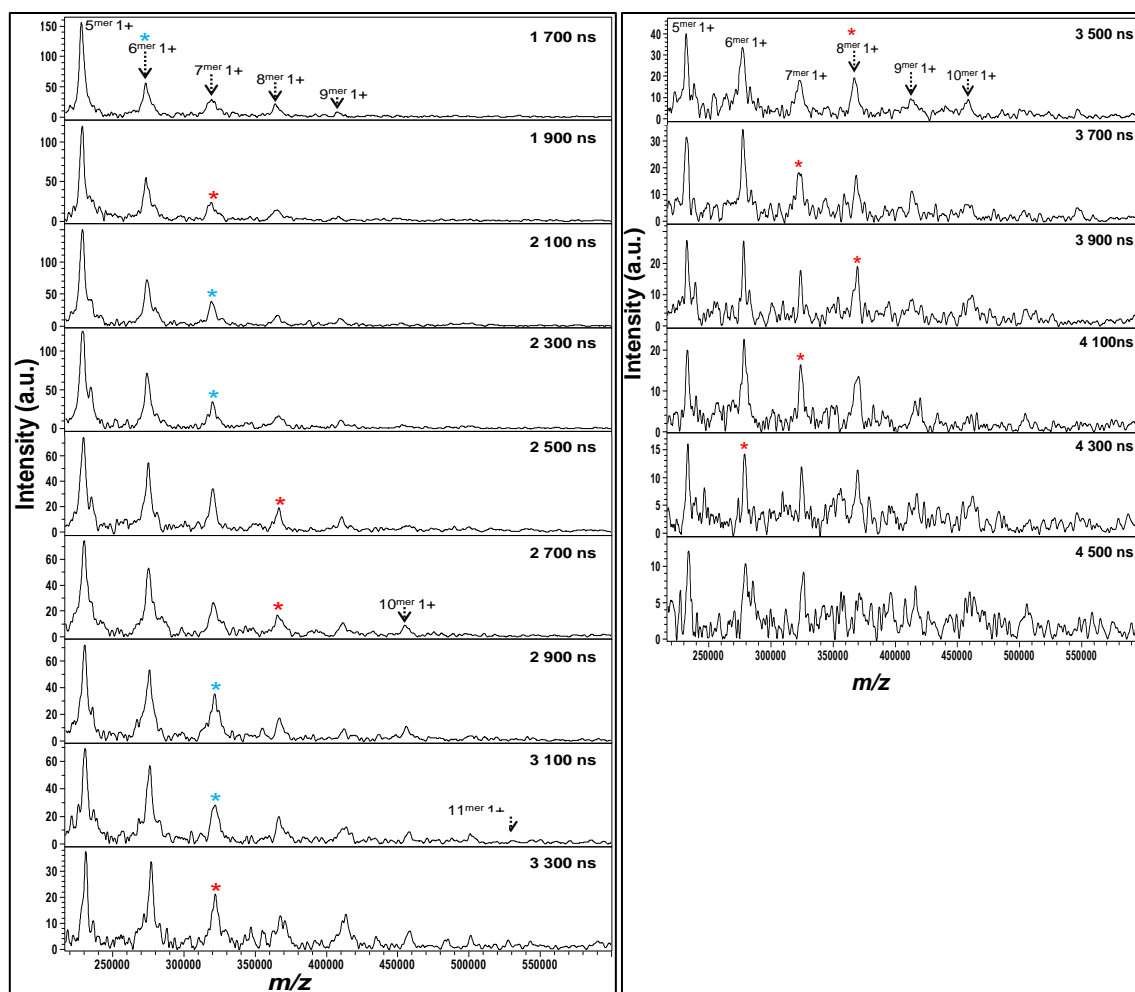


**Figure 4.** NALIM - MALDI TOF MS analysis of  $\alpha$ 1AT at 4  $\mu$ M. The matrix/sample solution contains HAG 75%, water 12.5%, and glycerol 12.5%. Spectra were accumulated over 400 shots.

As seen in Figure 4, this protein forms oligomers in a noncovalent manner, which give homogenous and resolved peaks. Peak intensities do not follow a typical exponential decay. Given that the monomer is the most abundant species in solution, the relative intensities of monomer, dimer and trimer peaks could indicate that monomers and dimers are used as building blocks for higher

oligomeric states. At 1 700 ns, the octamer (*circa* 370 kDa) is the highest oligomer observed in the spectrum. Thus, we proceed to determine the effect of delay values higher than 1 700 ns on the sensitivity and the resolution for ions with  $m/z$  higher than 150 000.

**Delay optimization using  $\alpha$ 1AT.** We assessed the effect on the S/N ratio and the resolution of fourteen delay values ranging from 1 700 to 4 500 ns (maximum achievable value) in 200 ns increments (Figure 5).



**Figure 5.** NALIM - MALDI TOF MS analysis of the effect of the delay on the S/N ratio and resolution of  $m/z$  peaks higher than 150 000, using  $\alpha$ 1AT at 9  $\mu$ M as analyte. Spectra are zoomed in the 212 000 to 600 000  $m/z$  range. The matrix/sample solution contains HAG 75%, water 12.5%, and glycerol 12.5%. Spectra were accumulated over 400 shots. Asterisks show the highest oligomeric state peak for each delay.

As can be seen in Figure 5, increasing the delay improves the sensitivity for ions higher than  $m/z$  150 000. At 3 100 ns, the undecamer becomes the highest observable oligomeric state (Table 5), which is a gain of 100 000 in  $m/z$  range compared to ions observed at 1 700 ns (i.e., nonamer).

**Table 5.** Signal to noise (S/N) ratio of each  $\alpha$ 1AT oligomer peak at different delays.

Oligomeric state	m/z	DELAY (ns)														
		1700	1900	2100	2300	2500	2700	2900	3100	3300	3500	3700	3900	4100	4300	4500
<b>Tetramer</b>	<b>182328</b>	42	48	24	26	25	22	24	21	12	12	11	10	13	8	6
<b>Pentamer</b>	<b>227910</b>	22	21	18	20	17	15	16	13	15	7	8	8	5	4	7
<b>Hexamer</b>	<b>273492</b>	14	13	16	23	13	15	16	12	8	9	9	9	10	4	5
<b>Heptamer</b>	<b>319074</b>	7	8	13	14	14	11	10	12	5	7	4	4	5	3	3
<b>Octamer</b>	<b>364656</b>	8	8	7	7	12	7	7	7	6	8	7	6	4	4	N/D
<b>Nonamer</b>	<b>410238</b>	4	4	5	7	7	7	4	5	6	3	4	N/D	N/D	N/D	N/D
<b>Decamer</b>	<b>455820</b>	N/D	N/D	N/D	N/D	N/D	7	8	4	4	4	N/D	N/D	N/D	N/D	N/D
<b>Undecamer</b>	<b>501402</b>	N/D	N/D	N/D	N/D	N/D	N/D	N/D	5	4	N/D	N/D	N/D	N/D	N/D	N/D

**In yellow.** The three highest S/N ratio values of each oligomeric state. **N/D.** No peak was detected.



Interestingly, as seen in Table 5, the positive effect of the delay on the S/N ratio seems to reach a limit at 3 100 ns. This observation is consistent with the behavior found with IgA. This confirms that the optimization of the delay - which positively impacts the S/N ratio at a high  $m/z$  range - is limited to a defined mass range. Thus, 3 100 ns appears as the maximum usable delay for  $\alpha$ 1AT.

Additionally, peaks at S/N ratio < 10 are heterogeneous, which prevents their use as calibrants. A higher limit of detection may help mitigate peak heterogeneity. Hence, in order to guarantee homogenous peaks and the best S/N ratio, we selected useable peaks above  $m/z$  150 000 based on two criteria: (i) S/N >10, and (ii) 3 100 ns as the maximum delay (Table 6).

**Table 6.** Resolution of each oligomeric state peak at different delays.

Oligomeric state	$m/z$	DELAY (ns)							
		1700	1900	2100	2300	2500	2700	2900	3100
Tetramer	182328	57	55	59	56	68	76	79	68
Pentamer	227910	73	65	73	52	74	67	65	64
Hexamer	273492	78	77	59	63	94	55	94	70
Heptamer	319074			76	80	82	66	88	57

**In Yellow.** The three highest S/N ratio values of each oligomeric state. **In Blue.** The highest resolution of each oligomeric state.

As seen in Table 6, an increase in the delay improves the resolution of high  $m/z$  peaks. Interestingly, the best delay for resolution does not correspond with the best delay for the S/N ratio value. This observation is consistent with the behavior found with IgA. Thus, to guarantee the best resolution and S/N ratio, we propose two delays: one for calibration (Table 7) and another for peak detection in a specific  $m/z$  range (Table 8).

**Table 7.** Delay optimization for calibration.

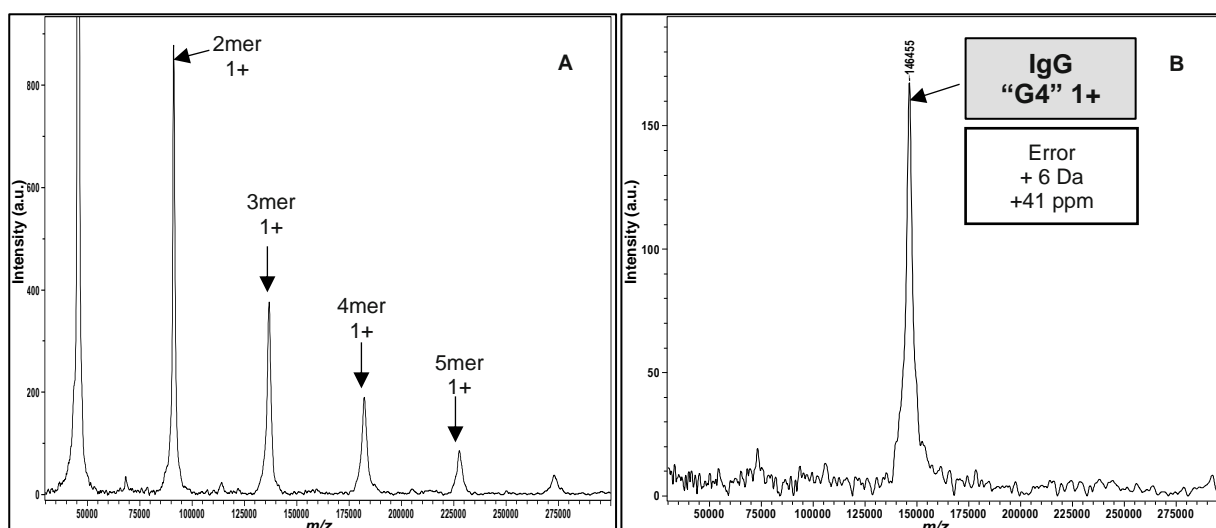
Oligomeric state	$m/z$	Recommended delay (ns)
Tetramer	182328	1 700
Pentamer	227910	2 100
Hexamer	273492	2 500
Heptamer	319074	2 900

**Table 8.** Delay optimization for peak detection in a specific  $m/z$  range.

$m/z$ range	Recommended delay (ns)
150 000 - 210 000	1 900
170 000 - 260 000	2 100
200 000 - 310 000	2 300
250 000 - 360 000	2 500
300 000 - 410 000	2 700
350 000 - 460 000	2 900
400 000 - 510 000	3 100

### Alpha-1-antitrypsin ( $\alpha$ 1AT) as calibrant for a high $m/z$ range

After delay optimizations were performed, we tested  $\alpha$ 1AT as a calibrant for high-mass ions ( $m/z \geq 100\,000$ ). The goal of this experiment was to check the accuracy of the mass determination using  $\alpha$ 1AT as a calibrant in NALIM. To this end, a deglycosylated IgG was used as a test protein. Since the sequence of the IgG was not known, its mass determined by HRMS was used as a proxy for the theoretical mass. The mass determined in NALIM for a deglycosylated IgG was then compared to the mass determined by HRMS (Figure 6).



**Figure 6.** Validation of the calibration for a high  $m/z$  range by NALIM - MALDI TOF MS. **A.** Calibrant:  $\alpha$ 1AT at 16  $\mu$ M. **B.** Sample tested: IgG "G4" at 0.2  $\mu$ M. The matrix/sample solution contains HAG 75%, water 4%, and glycerol 21%. Spectrum was accumulated over 200 shots.

As seen in Figure 6,  $\alpha$ 1AT as a calibrant gives great accuracy in the mass determination for high  $m/z$  peaks. The average mass determined in NALIM is 6 Da or 41 ppm higher than the mass determined

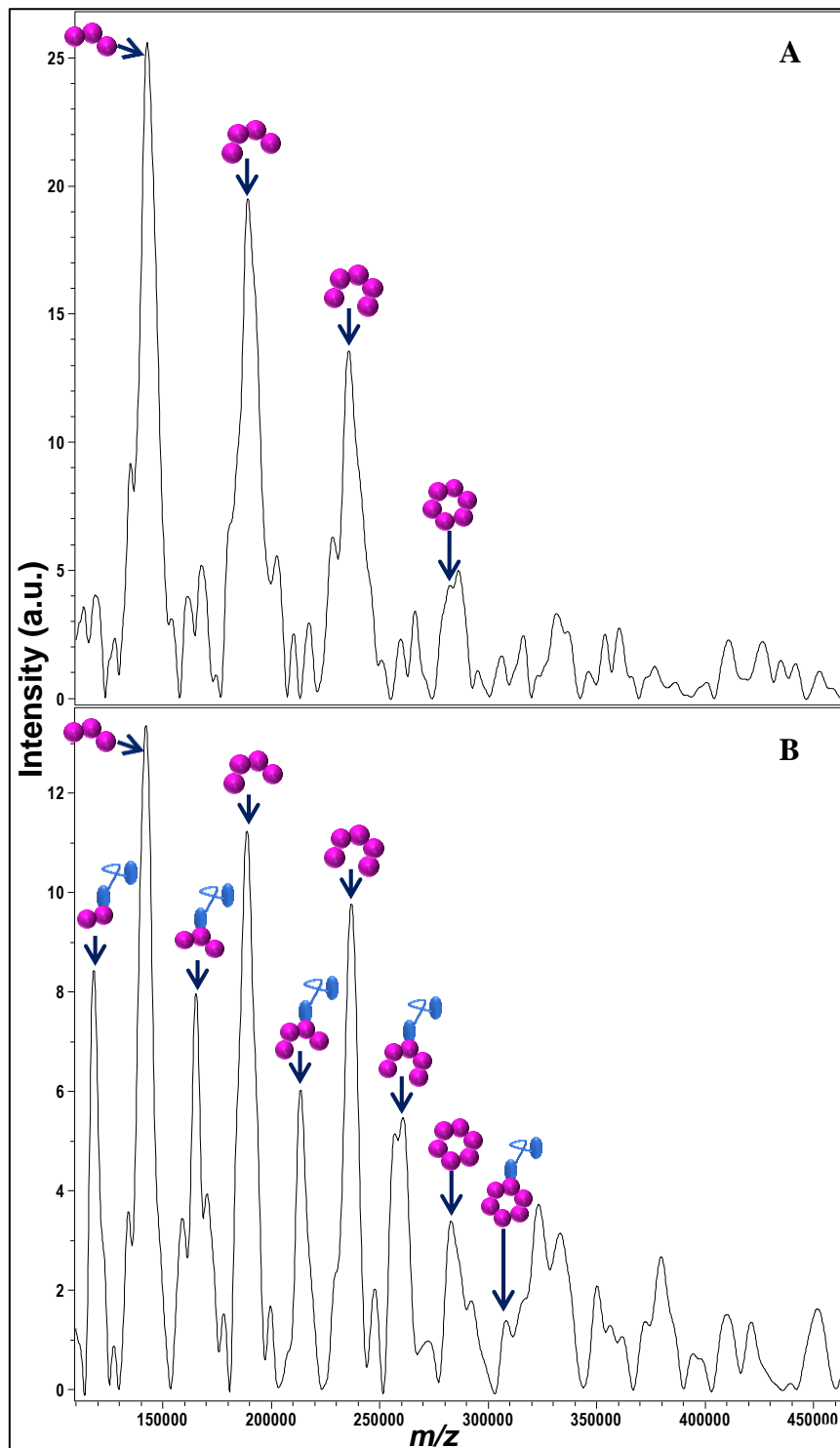
by HRMS (146 448 Da). Considering the limited resolution obtained in native conditions, this relatively low error confirms that  $\alpha$ 1AT can be used as a calibrant for high mass determination.

## **Structural insight into soluble protein complexes by NALIM**

### **NusG-Rho complex in NALIM**

The Rho-NusG complex stoichiometry from *E. coli* has previously been determined. These proteins form a stable complex, where one NusG binds one Rho hexamer, with  $K_d = 15$  nM (Pasman & von Hippel, 2000). To further characterize this complex, we designed an experiment to evaluate whether NusG affects the stability of the Rho hexamer using the NALIM method.

Within the constraints of the initial concentration of proteins, a Rho hexamer concentration of  $3.4 \mu\text{M}$  was set as the maximum attainable prior to  $\frac{1}{4}$  dilution in the HAG matrix. Using this maximum attainable concentration, we chose the NusG concentration so as to reach the saturation point for protein-protein binding. Rho/NusG thus was incubated at a 1:2.5 (n/n) ratio in a glycerol/water solution, which should give 100% saturation. Furthermore, in order to improve the Rho hexamer stability in solution, we added ATP to the incubation as it promotes the closed-ring conformation of Rho. According to Xu *et al.* (2003), the Rho hexamer has six binding sites for ATP, where three sites have a high affinity ( $K_d = 0.33 \mu\text{M}$ ), and three have a low affinity ( $K_d = 9 \mu\text{M}$ ). Thus, a 1000-fold excess of ATP over Rho was added to achieve saturation. As a control, the Rho/ATP complex was incubated in the absence of NusG.



**Figure 7.** NALIM - MALDI TOF MS analysis of the effect of NusG binding on Rho hexamer. **A.** Rho hexamer at 0.85  $\mu\text{M}$ , and ATP at 850  $\mu\text{M}$  **B.** Rho-NusG complex: Rho hexamer at 0.85  $\mu\text{M}$ , NusG at 2.12  $\mu\text{M}$ , and ATP at 850  $\mu\text{M}$ . The matrix/sample solution contains HAG 75%, water 12.5%, and glycerol 12.5%. Spectra were accumulated over 400 shots. Representation of the Rho/NusG complex was adapted from Peters *et al.* (2012).













As seen in Figure 7.A, the Rho hexamer is observed in NALIM. As it is unlikely that stoichiometries lower than the Rho hexamer are present in solution, the observation of peaks corresponding to species ranging from monomers to pentamers suggests that gas phase dissociation (GPD) occurs.

As seen in Figure 7.B, the Rho hexamer/NusG complex is observed in NALIM. The fraction of Rho hexamer bound to NusG, calculated as  $f_{\text{bound}} = \text{Rho-NusG}/(\text{Rho} + \text{Rho-NusG})$ , is 30%. As the experiment was designed to reach 100% saturation, this lower bound fraction suggests that significant GPD of NusG from Rho occurs as well. Furthermore, Rho monomers to pentamers are found associated with one NusG subunit. These complexes are likely the result of GPD of the Rho hexamer/NusG complex, as previously seen with the control.

Interestingly, the resolution measured on Rho/NusG complex peaks is higher than that of Rho peaks (Table 9). The widening of free Rho peaks compared to Rho/NusG complex peaks may indicate that Rho is prone to metastable decay. If this is the case, the comparatively narrow complex peaks would suggest that NusG can prevent metastable decay of Rho, which may be interpreted as the stabilizing effect of NusG on the Rho protein.

Additionally, as seen in table 9 in “Rho/ATP in presence of NusG”, the mass shift ( $M_{\text{obs}} - M_{\text{theo}}$ ) corresponding to peaks of Rho/NusG complex is higher than the mass shift of Rho alone peaks. The conformational changes induced by NusG binding to Rho seems to lead to Rho stabilization and could also increase the affinity of the ATP binding sites.

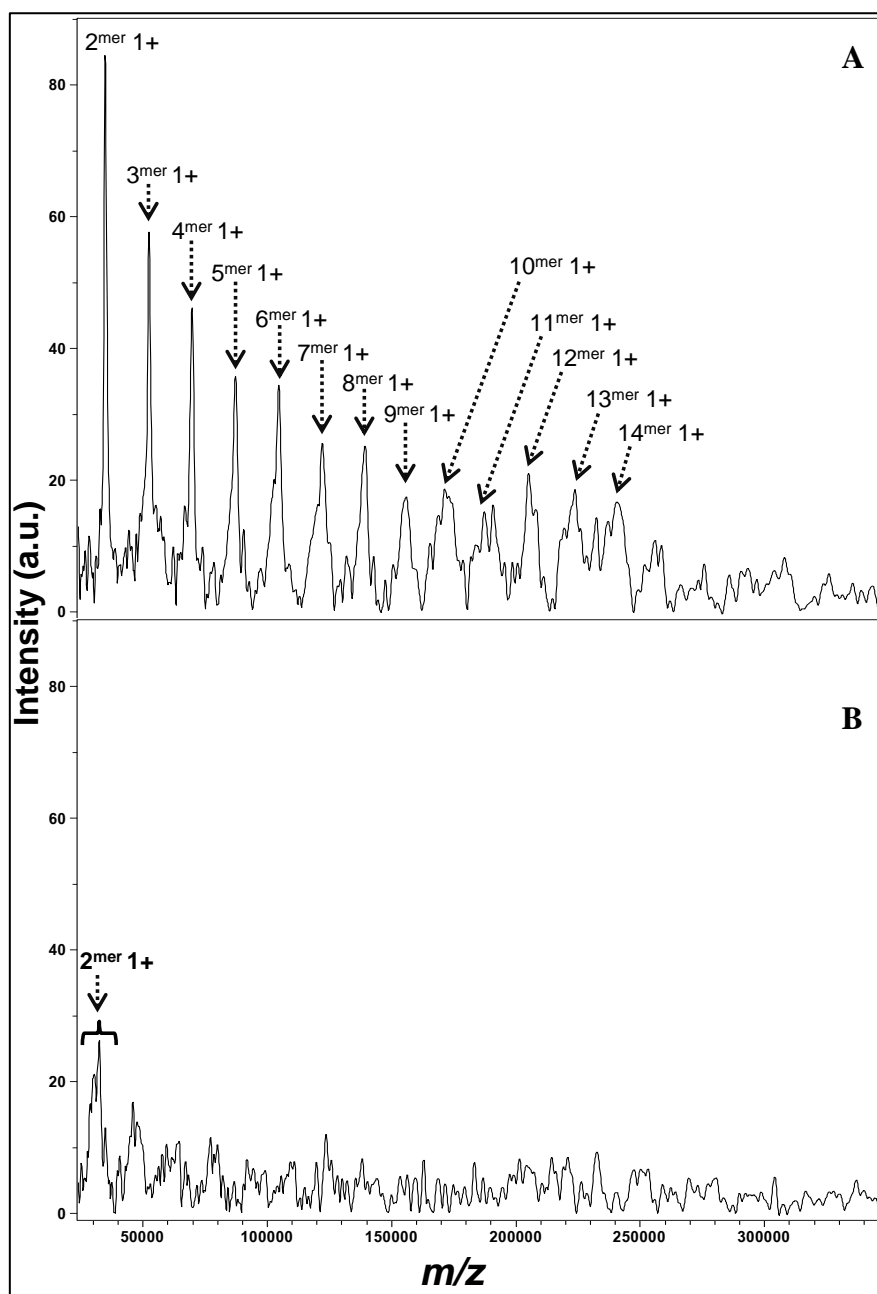
**Table 9.** Mass determination of Rho and Rho-NusG complex.

Rho/ATP in the absence of NusG						Rho/ATP in the presence of NusG					
Stoichiometry	Hypothetical number of ATP	Mass Observed (Da)	$M_{\text{osb}} - M_{\text{theo}}$ (Da)	Experimental error (ppm)	Resolution	Stoichiometry	Hypothetical number of ATP	Mass Observed (Da)	$M_{\text{osb}} - M_{\text{theo}}$ (Da)	Experimental error (ppm)	Resolution
	2	142933	322	2256	19		1	142292	188	1324	21
							2	165470	296	1786	47
	1	189626	324	1709	20		0	188961	166	881	22
							4	213770	383	1790	36
	0	235894	-99	-418	27		2	237073	66	278	33
							4	260739	154	589	63
	0	282648	-543	-1921	51		0	283368	177	623	42
							5	308678	387	1253	74

\* Theoretical mass ( $M_{\text{theo}}$ ) of: Rho monomer = 47 198.4 Da, NusG = 22 563.6 Da, and ATP = 507.18 Da.

### **Oligomerization of ZBTB8A transcription factor**

Despite the importance of ZBTB8A for the cell, structural and functional knowledge about this protein is still limited. Protein oligomerization was first suggested by the identification by X-ray crystallography of a contact interface between dimers through hydrophobic surfaces, and confirmed through electron microscopy (Mance et al., 2023 [currently in publication]). However, the degree of oligomerization could not be precisely determined using these techniques. It is also not clear how oligomerization impacts ZBTB8A protein function. Hence, we endeavored to obtain structural insight about ZBTB8A oligomerization using the NALIM method. For this, a ZBTB8A WT protein (residues 1-146) from *Xenopus laevis* was produced. Based on the above mentioned structure, serine 103 was identified as a key amino acid for the ability to form dimer-dimer interfaces and high-order oligomers. A mutated ZBTB8A protein (S103R), in which serine 103 is replaced with an arginine, was thus also produced as a negative control for the detection of oligomers in NALIM experiments (Mance et al., 2023 [currently in publication]).



**Figure 8.** NALIM - MALDI TOF MS analysis of ZBTB8A protein oligomerization. **A.** ZBTB8A WT at 10  $\mu$ M. **B.** ZBTB8A S103R mutant at 10  $\mu$ M. Spectra are zoomed in a  $m/z$  range from 24 000 to 350 000. The matrix/sample solution contains HAG 75%, water 12.5%, and glycerol 12.5%. Spectra were accumulated over 400 shots.

As seen in Figure 8.A, oligomers of ZBTB8A WT are observed in NALIM. The highest oligomeric state identified is the tetradecamer (14-mer), which has a molecular mass *circa* 246 kDa. As no oligomers higher than the dimer are observed with the ZBTB8A S103 mutant (Figure 8.B), the observed WT oligomers can be attributed to specific association, rather than gas-phase aggregation.

Taken together, these results confirm that ZBTB8A WT forms specific high-order oligomers based on noncovalent interactions of monomers and/or dimers.



## DISCUSSION

### **Overcoming limitations to study large molecular complexes NALIM.**

NALIM generates molecular ions with low charge states (1+ and 2+), the monocharged state being largely predominant. Even though having only monocharged peaks in the high  $m/z$  range could be a potential limitation in terms of resolution and mass accuracy, this fact facilitates data interpretation. Moreover, contrary to ESI where ions are formed with several charges, there is no risk of Coulombic repulsion, so that unwanted effects such as protein unfolding and the loss of noncovalent interactions are avoided (Mehmood et al., 2015; Petroff et al., 2020). Thus, the size of the complexes and the fact that are mainly in a monocharged state featured two main challenges: (i) instrumental parameters inadequately tuned (i.e., delay), (ii) the lack of calibrants for this mass range (i.e., 100 000 Da or higher). As a consequence, we went in search of appropriate protein or protein complexes so as to extend the usability of the NALIM method for the study of large molecular complexes. In turn, this approach prompted us to optimize parameters for the new  $m/z$  range.

In the proteome of most organisms, a limited number of proteins have a molecular mass of 100 000 Da or higher. In this category of large proteins, few are expressed as sufficiently homogeneous species to produce well-resolved peaks, which makes the search for the appropriate calibrant difficult. In native conditions, broad peaks may originate from the interaction or adduction of proteins with components of the sample solution and/or the matrix, as well as micro-heterogeneity resulting from variable post-translational modifications (PTMs).

Our search for calibrant first focused on proteins than can produce peaks over 300 000 Da, so as to cover the mass range of several protein complexes we were considering as good case scenarios for NALIM of oligomers. As analytical techniques are being developed at an accelerated pace for the characterization of diagnostic and therapeutic immunoglobulins, our first idea was to evaluate IgA, as it is an immunoglobulin dimer and fits our mass criterion. The lack of commercial recombinant IgA drove us to attempt analysis of purified mouse IgA from hybridoma cells. Upon quality control, we found the protein to produce more peaks than expected. It turns out that in mouse IgA, contrary to human IgA, the heavy and light chains are bound in a noncovalent manner, consistent with our observations. Moreover, it was found in a both monomer and dimer forms, as was previously described by Corte and Parkhouse (1973). However, we also observed a high degree of heavy chain glycosylation, as was shown previously (Chintalacharuvu & Emancipator, 1997). In spite of our efforts to deglycosylate the dimeric mouse IgA, wide peaks indicating heterogeneity remained. A structure of mouse IgA, which was published over the course of this work, shows that the junction chain (J chain) binds the two monomers in a covalent manner while producing a bend on one face of

the IgA (Kumar Bharathkar et al., 2020). This forms a concave surface that may cause the J-chain to mask glycosylation site(s) in positions Asn437 and/or Asn452 of the heavy chains, thus preventing the PNGase F glycosidase from reaching those glycosylation sites due to steric hindrance. Unfortunately, the steric hindrance is such that full deglycosylation was not expected to be reached through the use of glycosidases with alternative specificity of bond cleavage. In spite of an interesting distribution of light chain dimers, IgA monomer, and IgA dimers with and without light chains (Figures 2 and 3), this intractable glycosylation precludes the use of mouse IgA as a calibrant for masses over 150 000 Da.

The  $\alpha$ 1AT protein can form oligomers in certain conditions. In contrast to our eukaryotic expressed mouse IgA, recombinant  $\alpha$ 1AT expressed in *E. coli* is not glycosylated, which should give rise to more homogeneous peaks than the IgA. Indeed, we readily found that  $\alpha$ 1AT generated quite homogeneous oligomers peaks. This can partly be attributed to the oligomeric nature of  $\alpha$ 1AT, as homo-oligomers form narrower isotopic distribution than hetero-oligomers such as immunoglobulins. Furthermore, oligomers have the intrinsic property to generate molecular ladders made of peaks at constant intervals. This property is useful for calibration over a wide mass range from *circa*  $m/z$  45 500 to 320 000 (Table 6).

Thanks to this new tool, we were able to optimize the delay with the aim of improving peak resolution, although this may happen at the expense of peak S/N ratio (Tables 5 and 6). Improving these two peak properties as a positive effect on mass accuracy, is an important consideration when  $\alpha$ 1AT is used as a calibrant. The  $\alpha$ 1AT oligomer used as calibrant was validated using a deglycosylated IgG and the mass determined in NALIM was compared with the mass determined in HRMS (146 448 Da). The error found was +6 Da or +41 ppm. This error can be considered as low and acceptable for goals such as stoichiometry determination, protein-protein, or protein-ligand interactions, where errors of a few Daltons do not seriously affect the interpretation of the results. To our knowledge, the protein  $\alpha$ 1AT is the first example of a calibrant that can be used with native mass spectrometry.

Additionally, the optimization of the delay showed that the optimum  $m/z$  range for a given delay is limited by the trade-off between resolution and S/N ratio. Similar behavior has been described by Edwards *et al.* (2003) for the high mass of carbonaceous pitches, albeit at a lower mass range. Thus, the delay must be adapted to fit the size of the biological complex to analyze, as well as the experimental goal.

### **Exploring high molecular complexes through NALIM.**

After optimizations were performed, the analysis of soluble proteins provided information related to their stoichiometry and complex stability.

It was possible to determine both the highest degree of oligomerization detectable in NALIM for ZBTB8A WT and  $\alpha$ 1AT proteins, which were respectively the tetradecamer (*circa* 243 kDa) and undecamer (*circa* 500 kDa), and the distribution of oligomers. NALIM thus affords complementary information to structural techniques, such as X-ray crystallography or cryo-EM, which tend to give an averaged picture of oligomers that are present. To fully leverage structural elements brought by all 3 techniques, it would be useful to create models of oligomerization, which is the purpose of subsequent studies.

ZBTB proteins form dimers through the interaction of BTB domains (Bonchuk et al., 2023; Kim et al., 2013; Perez-Torrado et al., 2006). The ability of BTB domains to mediate higher-order oligomers has long been debated. This phenomenon is best characterized for the *Drosophila melanogaster* ZBTB protein GAGA (Stogios et al., 2005), with recent analyses suggesting that an arthropod-specific subclass of BTB domains can hexamerize. For vertebrate ZBTBs, BTB-mediated oligomerization has been suggested for ZBTB6 (also known as ZID) based on the effect of BTB deletion on its behavior, but a conclusive proof has been lacking (Bardwell & Treisman, 1994; Bonchuk et al., 2023). In addition, ZBTB17 (also known as Miz-1 for Myc-interacting zinc finger), has been suggested to form tetramers based on crystal contacts, but these require high concentration in solution (Stead et al., 2007). In preliminary X-ray crystallography analysis of the BTB domain of the ZBTB8A protein from *Xenopus laevis*, dimer-dimer contacts were found, leading to formation of ‘infinite’ filaments across the crystal lattice, which could reflect the propensity to form filaments in solution (Mance et al., 2023 [currently in publication]). To confirm the presence of a ZBTB8A oligomer distribution and start testing the NALIM method's ability to detect larger complexes than previously achieved (Beaufour et al., 2018), we performed NALIM on wild-type ZBTB8A. This allows for the detection of this expected dimer, but also of several oligomer peaks, in keeping with the interface observed in X-ray crystallography. The formation of filaments was then confirmed by electron microscopy. Analysis of the S103R mutant protein shows that residue S103 is critical for the formation of the dimer-dimer interface, as oligomer formation seems greatly diminished in this mutant. This is also consistent with the abolished filament formation found by analytical SEC conducted with this particular mutated protein (Mance et al., 2023 [currently in publication]). Moreover, this result also allows for gas-phase aggregation to be ruled out as a source of oligomers. In principle, NALIM should allow us in future work to assess the dynamics of ZBTB8A oligomers formation, and to explore the

possible role of this oligomerization in interactions with DNA, as oligomerization of the transcription factor would provide the basis for specific interactions with multiple sites in the DNA (Bonchuk et al., 2023).

Additionally, based on peaks corresponding to even and odd ZBTB8A stoichiometries, we can infer differences in affinities of monomer-monomer versus dimer-dimer interfaces, and their effect on both oligomerization and GPD (i.e., complex stability). For instance, in stoichiometries higher than the pentamer, peaks of even stoichiometries have a slightly higher intensity than peaks of odd stoichiometries. This can be indicative of either a higher probability of GPD between dimers than between monomers, or of lower interface affinities in solution. In both cases, an oligomer will more readily dissociate between dimers than monomers, meaning that the interaction established through the BTB domain is stronger than the interaction established by the hydrophobic interface.

NusG is another transcription factor known to interact with the bacterial Rho helicase. NALIM analysis shows that NusG increases the stability of the Rho/ATP complex. Based on the widening of free Rho/ATP complex peaks compared to Rho/ATP/NusG complex peaks, Rho may be prone to metastable decay, while upon NusG binding, metastable decay of Rho is prevented. This may be interpreted as the first clue that NusG has a stabilizing effect on the Rho protein (Figure 7). The smaller the Rho oligomer, the stronger this effect seems to be. It may be that Rho subunits gain a more folded conformation upon cooperative binding, so that dimers and trimers are less stable and benefit more from NusG stabilization. Moreover, within the uncertainties of mass determination associated with this experiment due to the presence of salts, the Rho/ATP complex shows a mass increase when it interacts with NusG, which can be attributed to a greater number of bound ATP molecules on Rho hexamer. Thus, NusG may have a second, indirect stabilization mechanism on Rho hexamer, *via* the stabilization of the Rho protein-ligand interaction.

## CONCLUSION

In the present work, we show for the first time that by leveraging the TOF analyser mass range, the NALIM method can make large noncovalent complexes ranging from 100 000 to 500 000 Da amenable to nMS analysis. Our work helped to show that the ZBTB8A transcription factor forms oligomers in solution, following observations of an infinite polymer pattern from X-ray crystallography. The formation of ZBTB8A filaments was then confirmed by electron microscopy. NALIM show a distribution of oligomers with clearly defined numbers of subunits. Differences in affinities of monomer-monomer versus dimer-dimer interactions may underlie a specific pattern of odd- an even-numbered subunits peaks. In parallel, the NALIM method successfully provided a picture of the interaction between the NusG transcription factor and the Rho helicase hexamer. Here, we found an increase in the stability of both the Rho-hexamers and the Rho-hexamer/ATP complexes upon NusG binding.

This structural insight was made possible by optimizing instrumental parameters and by overcoming the lack of calibrant for the high  $m/z$  range explored for these systems. As shown here, optimizing the delay is essential for studying high-mass complexes, because it improves the sensitivity for high-mass ions (i.e., an increase in the S/N ratio). Finding an adequate calibrant for the high-mass range, was also a straightforward process, as large proteins tend to be heterogeneous due to modifications. Here, we show the example of a mouse IgA dimer which cannot be completely deglycosylated, likely due to steric hindrance.

In contrast to immunoglobulins,  $\alpha$ 1AT generate homogenous, resolved peaks which form a molecular ladder from 150 000 Da to at least 500 000 Da, making it appropriate as a calibrant for high masses. Although NALIM produces less resolved peaks than would MALDI in denaturing conditions, it is suitable for the determination of complex stoichiometries of protein interactions.

In summary, NALIM shows promise as a viable alternative to ESI for the structural investigation of large molecular complexes in their native state, to complement structural biology methods such as X-ray crystallography, NMR and cryo-EM. Here, we lay down the groundwork for future development which could open the door to the analysis of even larger biomolecules, contributing to our understanding of structure-function relationships for molecular machines and assemblies in conditions that are ever closer to their *in vivo* environment.



## CHAPTER 2

### CONTEXT

Membrane proteins have different roles in the cell such as enzymes, transporters, or signaling receptors, which are of vital significance for the cell. As a consequence, membrane proteins represent 2/3rds of potential targets for drugs. However, only 10% of membrane proteins have been targeted for the development of drugs so far. The reason for this is that the characterization of these proteins remains a challenge because they are present in low quantity in the cell membrane and especially because, due to their hydrophobic nature, they require the use of membrane mimetics to solubilize them. Unfortunately, in most structural methods, these membrane mimetics can completely prevent the analysis of membrane proteins or at least impair data acquisition. As a result, only 2% of the high-resolution structures in the PDB correspond to membrane proteins and *circa* 90% of these potential drug targets are still unexploited. Since NALIM is based on MALDI, it is more tolerant to contaminants such as detergents. Thus, NALIM provides the chance to characterize MPs and their complexes in their native state and to extend the understanding of them. For instance, we focused on an ATP-binding cassette containing protein named Bacillus multidrug resistance ATP (BmrA), thanks to a collaboration with Jean-Michel Jault and Cédric Orelle from MMSB in Lyon, who purified ABC transporters for us to study. BmrA is a model for the study of MDR efflux pumps that could be targeted to lift antibiotics resistance and make a number of existing drugs active again towards a number of resistant bacteria. Another interesting system is the protein-protein interaction of Tx7335, novel activator toxin (a.k.a. an actitoxin) with KcsA, an archetypal potassium ion (K<sup>+</sup>) channel discovered in collaboration with Sebastien Poget from CUNY in New York, who kindly sends us purified channels and the TX7335 toxin. Understanding how this toxin exerts this unusual effect (all other toxins being inhibitor), starts with identifying the type of binding site on channel. We estimated that NALIM could be leveraged to answer questions about the binding site and the stoichiometry of the KcsA-Tx7335 complex in the native state.

Part of the challenge was to optimize instrumental parameters and sample preparation to guarantee the integrity of the structure of MPs and their stability throughout their analysis.

Additionally, we aim to explore the possibility of using NALIM in more *in vivo*-like conditions. To this end, the use of detergents needs to be circumvented. Two types of samples were chosen to explore this aspect: proteoliposomes, which mimic the membrane bilayer micelle, with LINGO protein sample provided by Séverine Morisset-Lopez from CBM, and membrane vesicles (MVs), which we made from purified membranes sent by collaborators, and which contain the membrane protein

expressed by the cell in its natural lipid environment. Preliminary results on BmrA-expressing MVs are encouraging and have since been reproduced.



# **NALIM ANALYSIS OF MEMBRANE PROTEINS (MPs) COMPLEXES**

## **Authors**

**Edison Zhamungui<sup>1,2</sup>, Jade Jaubert<sup>1,2</sup>, Irene Rostovsky<sup>3</sup>, Joekeem D. Arizala<sup>3,4</sup>, Sébastien F. Poget<sup>3,4</sup>, Séverine Morisset Lopez<sup>1,2</sup>, Cédric Orelle<sup>5</sup>, Jean-Michel Jault<sup>5</sup>, Martine Beaufour<sup>1,2</sup>,  
Martine Cadene<sup>1,2</sup> \***

1 Centre de Biophysique Moléculaire CBM CNRS UPR4301, Orléans, France

2. Université d'Orléans, Orléans, France

3. Department of Chemistry, CUNY College of Staten Island, New York, USA

4. PhD Program in Biochemistry, The Graduate Center of CUNY, New York, USA

5. Molecular Microbiology and Structural Biochemistry (MMSB), UMR 5086 CNRS/University of Lyon, Lyon, France.

\* Corresponding author

Martine Cadene, PhD, PharmD, CBM CNRS, Rue Charles Sadron, CS 80054, 45071 Orleans cedex 2, France.

[martine.cadene@cnr-orleans.fr](mailto:martine.cadene@cnr-orleans.fr)

## INTRODUCTION

Native mass spectrometry (nMS) methods have now reached the status of powerful tools for the analysis and characterization of biomolecular complexes, providing key information in support of high-resolution methods such as X-ray crystallography, NMR, and cryo-EM. For nMS, samples must be analyzed in carefully controlled non-denaturing conditions to preserve the tertiary and quaternary structure. The direct analysis of intact biomolecular complexes can provide insight into their stoichiometry, identity of binding partners, complex dynamics, and even binding affinities (Boeri Erba & Petosa, 2015; Gan et al., 2017; Jaquillard et al., 2012; Leney & Heck, 2017)

Membrane proteins (MPs) are related to important biological processes inside the cell. They act as signaling receptors, transporter of chemical compounds, and/or enzymes (Depping et al., 2022; Marty, 2020). Indeed, MPs constitute more than 50% of potential drug targets. However, our understanding of the structure and function of MPs is still limited compared to soluble proteins due to the difficulty in characterizing them. For instance, performing biochemical and structural studies require large quantities of pure and stable MPs. Unfortunately, many MPs are present in low quantity in the cell membrane. In some cases, the induced overexpression of a recombinant MP can trigger toxicity problems for the expressing organism or cell (Depping et al., 2022; Di Cesare et al., 2022). Moreover, X-ray crystallography of MPs requires their solubilization using carefully chosen detergents or nanodiscs, with variable rates of success. This explains why only 2 or 3% of all the available high-resolution structures are of MPs, despite the fact that almost 30% of the proteome is composed of MPs (Di Cesare et al., 2022; Marty, 2020). Understandably, nMS characterization of MPs also lags behind that of soluble proteins for similar reasons.

So far, the go-to method for nMS is based on electrospray ionization (ESI). Over the past decades, several breakthrough studies of MP complexes have been published. The first success using native ESI-MS was accomplished in 2008 by conserving the ATP binding cassette transporter BtuC<sub>2</sub>D<sub>2</sub> in the gas-phase (Barrera et al., 2008). Since then, nMS approaches were applied to determine for example the stoichiometry of assemblies containing transmembrane proteins (Barrera et al., 2009), or to probe the interactions of membrane proteins with lipids (for a review, see Bolla et al., 2019). To better mimic cell membranes and/or facilitate structural studies, different alternative solubilization methods such as amphipathic polymers (Leney et al., 2012), nanodisc (Marty et al., 2012), and sonicated membrane vesicles (Chorev et al., 2020) were developed and used in MS. Nevertheless, analysis by native ESI-MS can still be challenging due to the need for additives and salts to stabilize proteins in aqueous buffers. These additives cannot always be removed without impairing complex structure and function, while leaving them in the analysis medium can trigger ion suppression, the

formation of adducts, and the deposition of residue on instruments inner surfaces. This is especially true for nMS analysis of MPs.

By contrast, MALDI is more tolerant to contaminants and consumes lower quantities of samples than ESI. Moreover, MALDI coupled to a TOF instrument gives access a large range of masses. We previously leveraged MALDI's unique properties to create a new native MS method based on a liquid spot (Beaufour et al., 2018), which we are renaming NALIM for Native Liquid MALDI. In comparison with the classical solid-spot MALDI method, the success of NALIM relies on an ionic liquid matrix to transfer biomolecular complexes to the gas phase directly from solution. Proof-of-concept on protein-ligand and protein-protein complexes was published for soluble systems (Beaufour et al., 2018). Compared to ESI-based nMS methods, NALIM has the potential to show better tolerance to additives such as salts and solubilizing agents which are essential for MPs, with minimal sample consumption. As was demonstrated with soluble proteins, critical information such as proof of protein-protein or protein-ligand binding as well as the stoichiometry of complexes can be obtained. MALDI coupled to a TOF analyzer analyzes ions over a wide  $m/z$  range, making it suitable for a diverse gamut of biomolecular assemblies.

The aim of this work was to extend NALIM to the analysis of MPs. To this end, different detergent-solubilized membrane proteins were chosen to cover a wide range of biological functions. This included KcsA, an archetype of potassium ion ( $K^+$ ) channels from *S. lividans*, and BmrA, a xenobiotic transporter of the ABC family from *B. subtilis*. Moreover, in an attempt to bring NALIM to more *in vivo*-like conditions, MPs in liposomes and membrane vesicles (MV) obtained directly from cell membranes were also considered. This included LINGO1, a human leucine-rich co-receptor involved in neuronal regulation, and SpABC, a BceAB-type ABC transporter from in *S. pneumoniae*.

Here, we devised and applied instrumental and sample preparation optimizations to improve the signal and stability of MPs complexes in NALIM. Furthermore, we show that NALIM can be used to monitor protein-protein interactions (PPi) and protein-ligand interactions (PLi), and give structural information such as the complex stoichiometry, the effect of different factors on their stability and interactions with binding partners in the native state.

## MATERIAL AND METHODS

### Materials

The centrifugal concentrator and filter devices (Amicon®) and ammonium acetate were from Merck (Darmstadt, Germany). The 4-hydroxy- $\alpha$ -cinnamic acid (HCCA) matrix was from Bruker Daltonics (Bremen, Germany). The 3-aminoquinoline (3-AQ) matrix, ATP disodium hydrate, magnesium chloride hexahydrate, and glycerol were purchased from Sigma-Aldrich. Sodium orthovanadate (inorganic vanadate, hereafter  $V_i$ ) was from Acros organic (Geel, Belgium). The organic solvents acetonitrile and methanol were purchased from Biosolve (Dieuze, France). All aqueous solutions and buffers were prepared using 18 M $\Omega$  water purified with Milli-Q reagent grade system from Millipore (Bedford, MA, USA), herein referred to as “ultrapure water.”

Wild type (WT) KcsA, and chimeric Shaker KcsA (bearing Q58A, T61S, and R64D mutations) with 6xHis tag were expressed in X11-blue *E. coli* cells, and purified in decylmaltoside (DM) as previously described (MacKinnon et al., 1998). SpABC, WT BmrA and the K380A BmrA mutant were expressed in the *E. coli* BL21 (DE3) strain, and purified with methods previously reported by Steinfels *et al.* (2004), and Di Cesare *et al.* (2022) respectively. Cell membranes from *E. coli* BL21 (DE3) expressing BmrA, and OmpA-deleted SpABC expressing *E. coli* BL21 (DE3), were obtained following the protocol previously described by Di Cesare *et al.* (2022). LINGO1 in liposomes was provided by Dr Severine Lopez-Morrisset. Tx7335 was isolated using an affinity step with WT KcsA as a “bait” and purified by HPLC as described by Rivera-Torres et al. (2016). Lyophilized charybdotoxin (ChTx) was provided by Professor Michel de Waard.

The quality control of MPs and toxins was performed in denaturing conditions by MALDI-TOF MS using the UltraThin Layer (UTL) method. In brief, we first prepared a 1:2 solution of water and acetonitrile. This solvent solution was then saturated with the HCCA matrix. After homogenizing, this solution was further diluted with isopropyl alcohol at a ratio of 1/4. About 20  $\mu$ L of this ultrathin-layer solution were applied on the gold plate and spread across its entire surface. This layer of diluted matrix was then air-dried until only minimal moisture remained, and the excess matrix removed by wiping the plate with a tissue. As a result, an 'ultrathin layer' of HCCA remains at the surface of the gold plate, which is hardly visible but serves as a seeding layer for deposits.

Upon this ultrathin layer, half a microliter of sample diluted in a saturated matrix solution is deposited. The excess liquid containing contaminants can be aspirated when the so-called matrix-sample 'co-crystal' begins to form on the surface of the plate. The plate is then introduced into the MALDI source for sample analysis. The UTL method generates multicharged species generally centered on the 7+

to 3+ charge state, depending on protein size and ionizability (Cadene & Chait, 2000; Fenyo et al., 2007; Gabant & Cadene, 2008).

### Sample preparation

For the analysis of MPs, intermediate dilutions were prepared in glycerol (referred to herein as the “glycerol solution”) to achieve a concentration between 0.5 and 4  $\mu\text{M}$ , depending on the experiment and the MP complex under study.

Membrane vesicles (MVs) were prepared using 10 mg of cell membranes from BmrA expressing *E. coli* BL21 (DE3), or OmpA-deleted SpABC expressing *E. coli* BL21 (DE3) according to the protocol previously described by Chorev *et al.*, (2020). Lyophilized ChTx was reconstituted in ultrapure water then desalted and concentrated by buffer exchanged with 100 mM ammonium acetate (AA), pH 7.5, using an Amicon device with a cutoff of 3 kDa. Lyophilized Tx7335 was reconstituted in ultrapure water.

For competition experiments, ChTx and/or Tx7335 were incubated for 30 min with the Sh-KcsA glycerol solution. For ligand binding experiments, WT BmrA and K380A BmrA in glycerol solution were incubated with ATP/Mg<sup>2+</sup>/V<sub>i</sub> at the ratios and incubation times ( $t_i$ ) detailed in the results section.

### NALIM MS

The HCCA/3AQ/Glycerol (1:4:6 w/w) ionic liquid matrix, herein referred to as “HAG”, was prepared according to Beaufour *et al.*, (2018). The glycerol solution was diluted at a ratio 1/2 or 1/4 in HAG (herein referred to as the “matrix/sample solution”). For MPs, the detergent concentration was adjusted to remain at twice the critical micelle concentration (CMC). A 1  $\mu\text{L}$  aliquot of the matrix/sample solution was spotted onto a MTX stainless steel sample stage (Bruker Daltonics, Germany). The drop was left for 15 minutes at room temperature to evaporate the water part of the solution. This time is defined as droplet evaporation time ( $t_{de}$ ). The time that the matrix/sample spot spends in the source vacuum before the acquisition, defined as in vacuo residence time ( $t_{ivr}$ ) by Beaufour *et al.* (2018), was 15 minutes. NALIM MS analyses were performed using an UltrafleXtreme MALDI-TOF/TOF mass spectrometer from Bruker equipped with a 2 kHz smartbeam-II laser and PAN<sup>TM</sup> technology. Spectra were acquired with a 20-kV acceleration voltage in linear positive mode using a single-shot pattern and accumulating of 200 or 400 laser shots. An ion deflection of 2 000 Da below the starting  $m/z$  acquisition value was applied. The instrument was controlled using Bruker FlexControl software v4. NALIM MS spectra were processed using FlexAnalysis software v4 (Bruker). For toxins quality control, the Peptide Calibration Standard II from Bruker was used for calibration in a mass range between 700 and 3 200 Da. Protein A,

apomyoglobin, betalactoglobulin, and cytochrome c, from Sigma-Aldrich, were used for calibration in the 4 000 and 44 000  $m/z$  range. External calibration up to *circa*  $m/z$  150 000 was performed using a deglycosylated IgG of known mass.

### Calculation of oligomer percentages

KcsA and BmrA can dissociate into lower-number oligomers and monomer. For KcsA, the percentage of the conserved tetramer (%T) was calculated as follows:

$$\%T = \frac{A_T}{\frac{A_M}{4} + \frac{A_D}{2} + \frac{3 \times A_{Tri}}{4} + A_T} \times 100$$

Where,  $A_M$  = monomer peak area,  $A_D$  = dimer peak area,  $A_{Tri}$  = trimer peak area, and  $A_T$  = tetramer peak area.

For BmrA, the percentage of the conserved dimer (%D) was calculated as follows:

$$\%D = \frac{A_D}{\frac{A_M}{2} + A_D} \times 100$$

Where,  $A_M$  = monomer peak area, and  $A_D$  = dimer peak area.

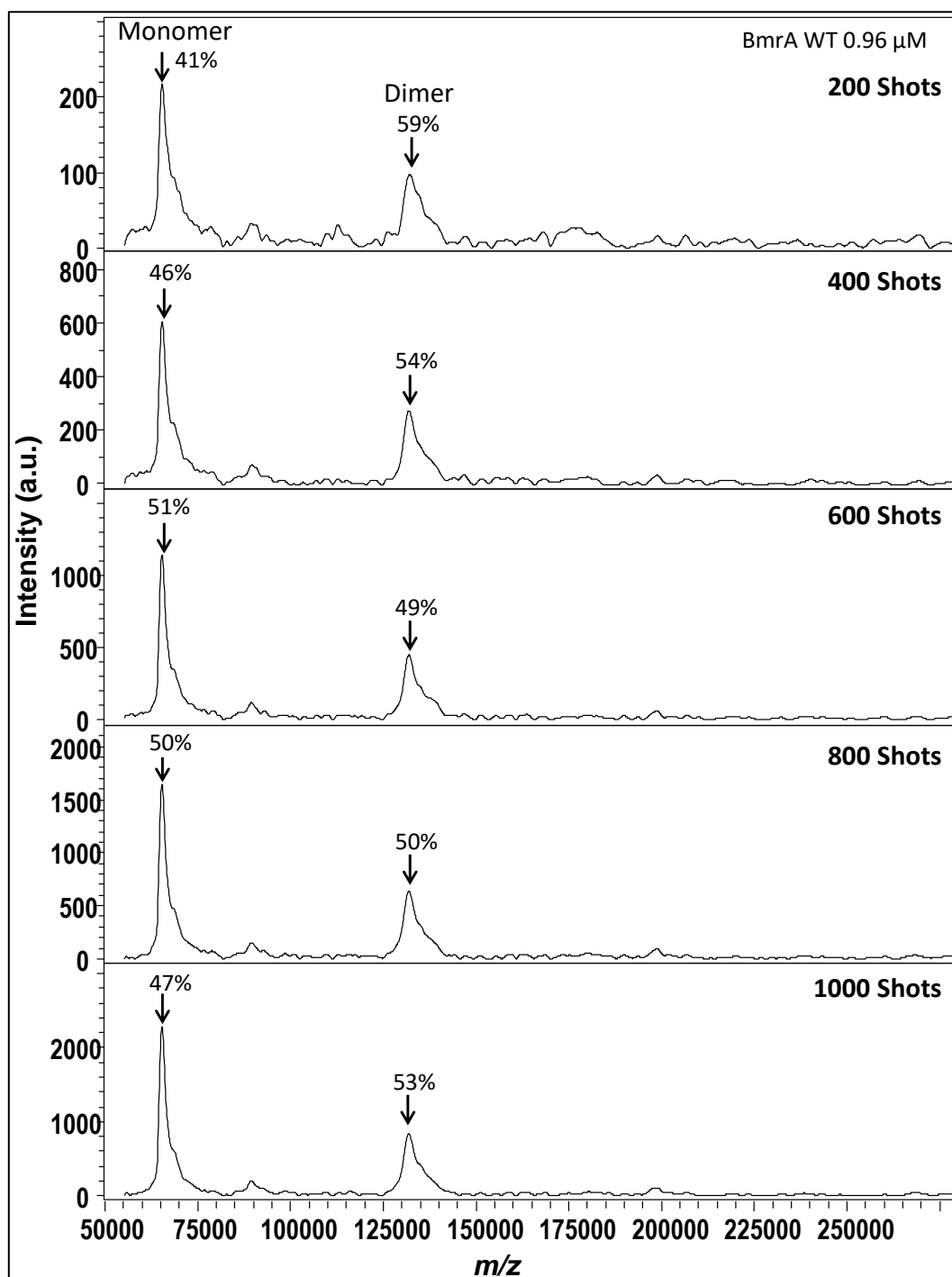
## RESULTS

### **Evidence of BmrA dimer and stabilization of the dimer in NALIM**

NALIM was originally conceived to get relevant structural information about complexes formed by proteins and their partners in native conditions. It was developed on model protein-protein and protein-ligand complexes (Beaufour et al., 2018). While the proof of concept was obtained for soluble protein systems, to make the NALIM methodology relevant to a wider range of applications in biology, we have endeavored to expand it to membrane proteins. The stability of an oligomer, and in particular a dimer such as BmrA, can be affected by instrumental parameters, the matrix/sample ratio, or sample composition.

### **Instrumental optimizations**

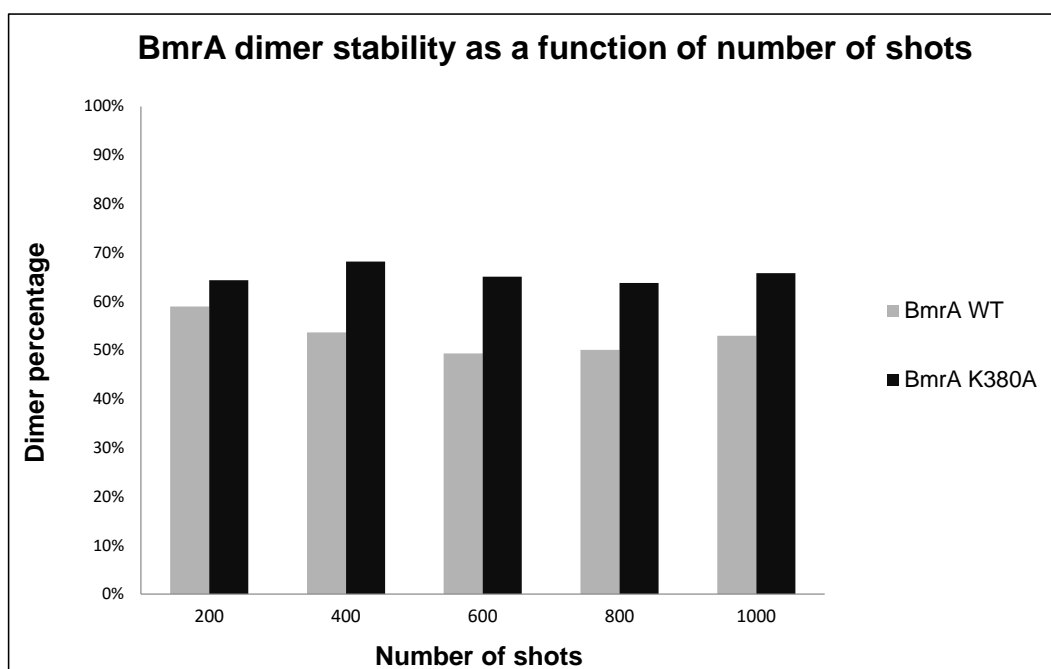
*Number of shots.* As previously reported by Beaufour *et al.* (2018), the spatial/temporal laser shot pattern can influence the stability of a protein complex in NALIM. The effects of laser irradiation acquire even more relevance with membrane proteins because these require a higher laser power to reach the ionization threshold. This prompted us to investigate the impact of the number of shots on the BmrA dimer stability.



**Figure 1.** Effect of the number of shots on the WT dimer stability. The matrix/sample solution was HAG 50%, water 15%, and glycerol 35%.

As seen in Figure 1, the stability of the WT dimer is highest at 200 shots. Between 400 and 1000 shots, the dimer percentage ranges from 49% to 53%, with no clear general trend.



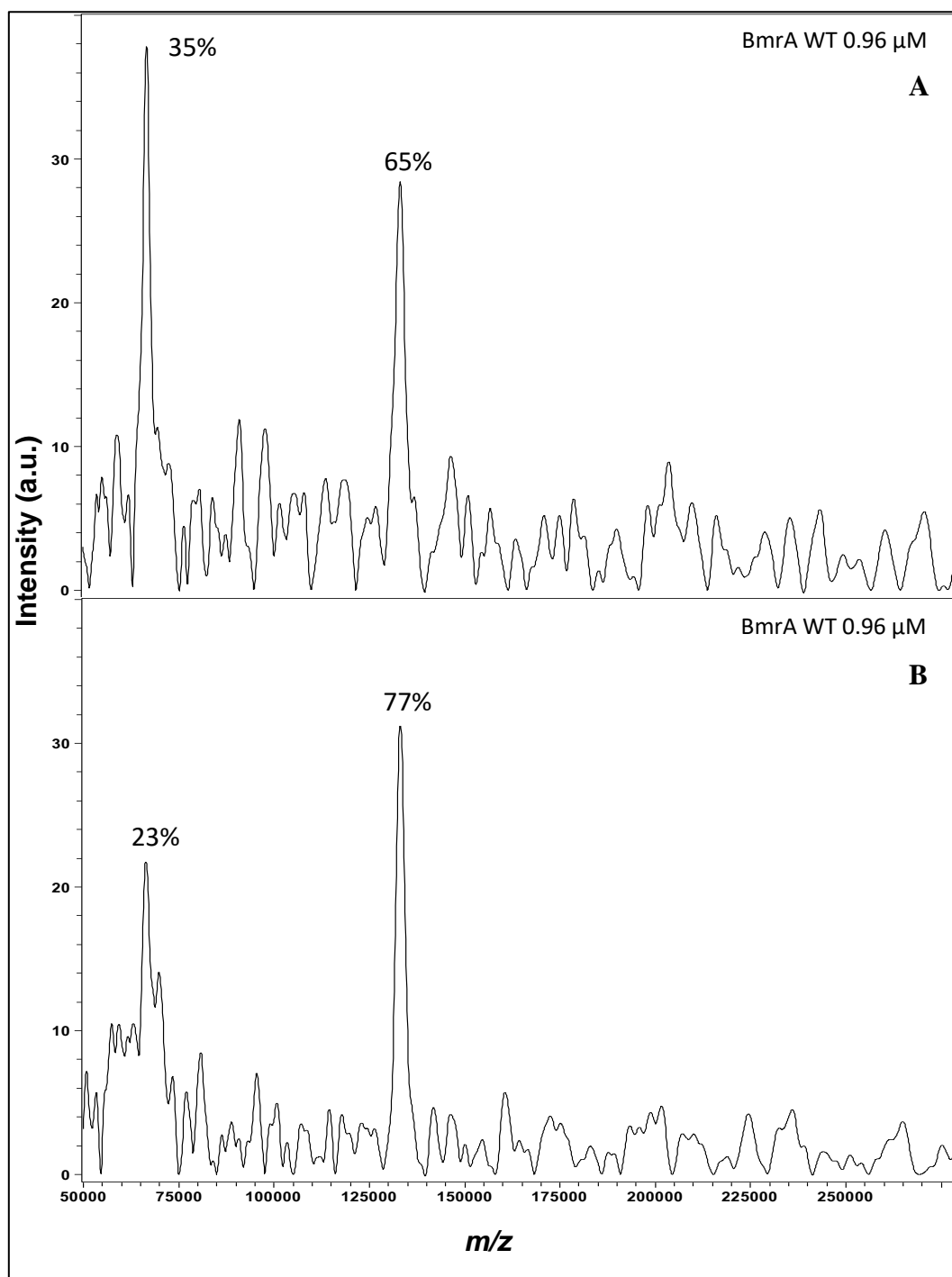


**Figure 2.** Histogram of the dimer stability as a function of the number of shots. **In gray WT. In black K380A.**

For the K380A (Figure 2), the dimer percentage ranges from 64% to 68%, with no clear general trend. Within the error margins of the experiment (protein concentration, measurement error on the percentage...), the mutant seems slightly more stable than the WT.

### Matrix/sample ratio

**Effect of the matrix (HAG) percentage.** It is well known that matrices play a key role in MALDI experiments. For this, different strategies were employed in the search for a liquid matrix solution which allowed to observe noncovalent protein complexes (Beaufour et al., 2018). The resulting liquid matrices showed to have a strong influence on the protein complex stability. Among them, the most suitable to observe noncovalent complexes was an ionic liquid matrix solution of HCCA/3-Aq/Glycerol (1:4:6 w/w). Consequently, there is a limitation in the number of liquid matrices alternatives. Due to this limitation, we focused on the matrix/sample ratio in order to improve the dimer stability.



**Figure 3.** Effect of the matrix/sample ratio on the WT dimer stability. The matrix/sample solution was **A.** HAG 50%, water 7.5%, and glycerol 42.5%. **B.** HAG 75%, water 7.5%, and glycerol 17.5%. Spectra were accumulated over 200 shots.

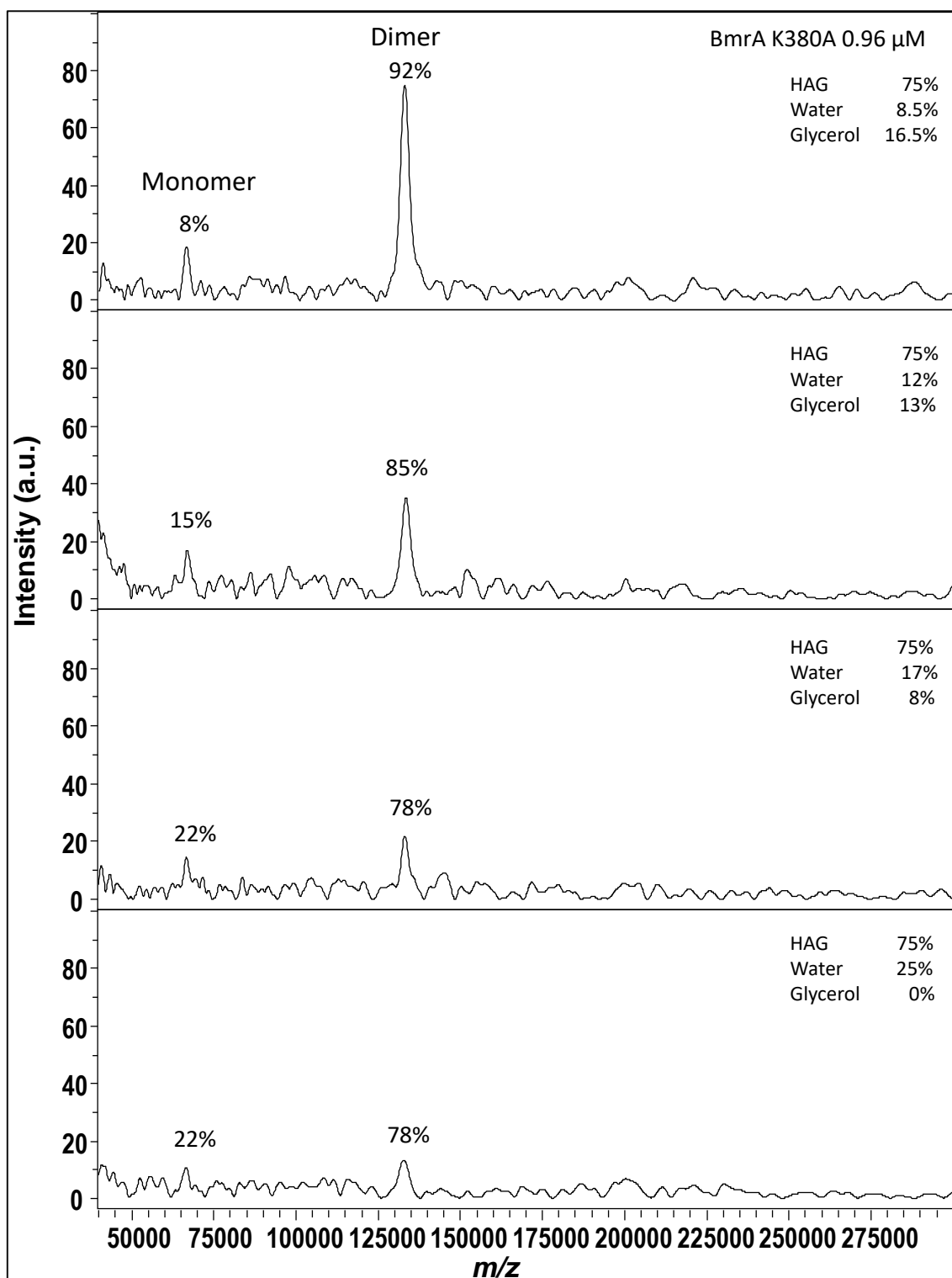
As seen in Figure 3, there is higher dimer preservation at a higher matrix/sample ratio. This suggests that the dimer complex stability is influenced by the matrix/sample ratio.

In addition, at a higher matrix/sample ratio the background noise decreases. This is corroborated by the signal-to-noise (S/N) ratio. For instance, at a higher matrix/sample ratio, this value (S/N = 10) is higher than the value of the lower matrix/sample ratio (S/N = 6).

**Sample composition**

In order to stabilize the dimer, one can attempt to strengthen the interaction at the interface between monomers, which is likely based on hydrophobic interactions according to our current knowledge of the structure of BmrA. To this end, we can modify different parameters which affect hydrophobic interactions, such as the water percentage and the concentration in salt(s).

*Effect of water percentage.* Within the constraints of the matrix-to-sample ratio which were found in the above experiment (Figure 3), the water percentage cannot be more than 25%. However, the water percentage can be decreased by substituting part of the water with glycerol in the final matrix/sample solution.

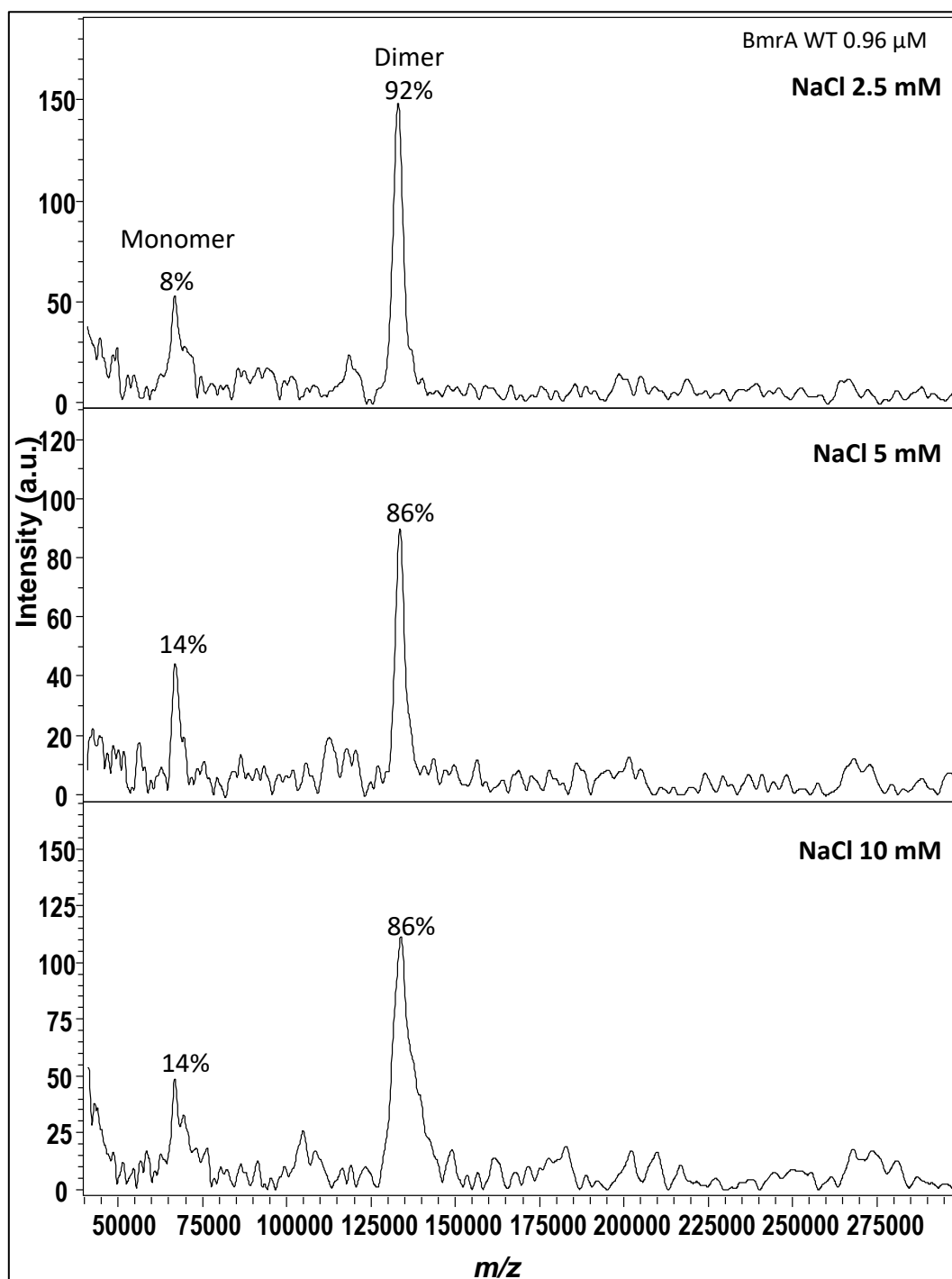


**Figure 4.** Effect of the final water percentage in the matrix/sample solution on the K380A dimer stability. Spectra were accumulated over 200 shots.

Given the protein concentration in this experiment, the water percentage could be set lower than 8.5%. Figure 4 shows that the dimer percentage increases when the water percentage decreases below 12%. This shows that the dimer stability can be improved by glycerol.

Additionally, the water percentage has an effect on the S/N ratio. At the lowest water percentage, S/N ratios of the monomer and dimer are above the limit of the detection ( $S/N > 3$ ). In contrast, the highest water percentage, despite maintaining the same dimer percentage as the second highest water concentration, has an effect in the ionization efficiency, which is reflected in a decrease in the S/N ratio ( $S/N = 3$ ). A possible explanation is that the ionization efficiency decreases because water molecules can compete with the matrix for interaction with the protein, which interferes with the formation of protein ions.

***Effect of salt concentration.*** In biochemistry, the use of nonvolatile salts to mimic the physiological ionic strength brought by salts (150 – 200 mM) is quite extensive. Unfortunately, in mass spectrometry, due to the formation of salt adducts, the presence of nonvolatile salts can degrade the quality of the spectrum. These adducts distribute the signal over many peaks, thus decreasing the detection limit, resolution, and mass accuracy (Susa et al., 2017). We evaluated the tolerance of NALIM to NaCl in the matrix/sample by varying the salt concentration from 2.5 to 10 mM. Salts in a buffer can also have a strong influence on the stability of proteins and their complexes. Hence, this experiment was also designed to evaluate the effect of NaCl on BmrA dimer stability.



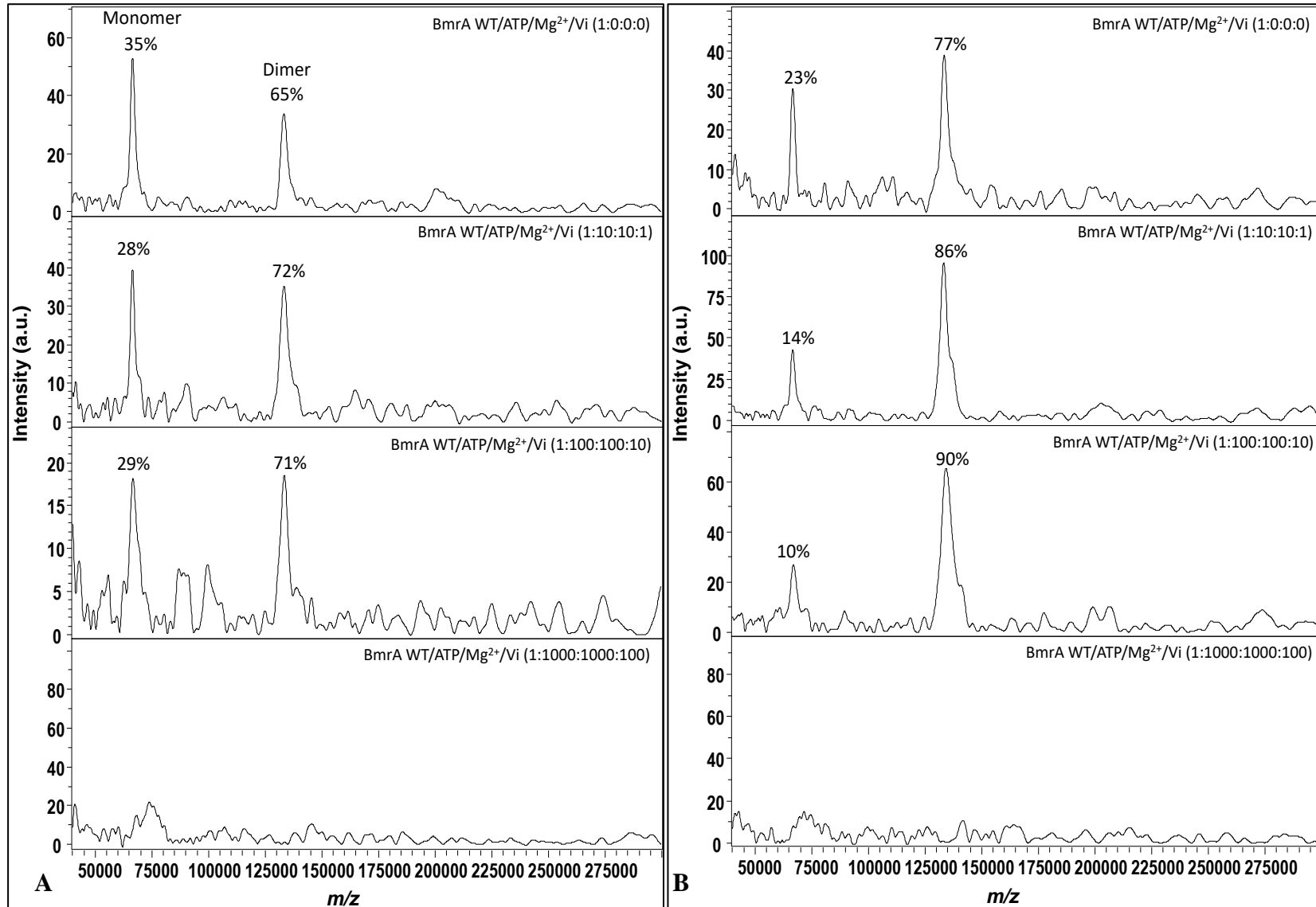
**Figure 5.** Effect of NaCl concentration in the matrix/sample solution on the WT dimer stability. The matrix/sample solution was HAG 75%, water 7.5%, and glycerol 17.5%. Spectra were accumulated over 200 shots.

As seen in Figure 5, the stability of the dimer is highest at 2.5 mM of NaCl. This could be a direct effect of sodium on the protein. In addition, at the highest salt concentration, peaks are broader and the background noise increases, which indicates probable  $\text{Na}^+$  adduct formation. Indeed, the S/N value of the dimer decreases as the NaCl concentration increases, suggesting that the tolerance of the method to NaCl is limited to a final concentration of 5 mM in the matrix/sample solution.

### **BmrA substrate binding for further dimer stabilization through ligand binding**

After optimizations were performed, we designed an experiment to evaluate the effect of ligand binding on the WT dimer stability. To this end, four different Protein-Ligand (P-L) ratios were tested.  $Mg^{2+}$  was added as a cofactor for ATP binding.  $V_i$  acts as a trapping agent for the ADP-BmrA complex after ATP hydrolysis, resulting in a more stable dimer conformation than the apoprotein or the ATP-BmrA complex (Javed et al., 2022).

In NALIM, the protein and ligands are incubated in a glycerol/water solution in order to limit the water percentage in the final matrix/sample solution. In Javed *et al.* (2022), the incubation time ( $t_i$ ) between protein and ligand was set at 15 minutes before the thermal unfolding analysis. The differences in viscosity between the aqueous buffer used in the thermal unfolding analysis and the glycerol/water solution used in our incubation means that diffusion is slower in the presence of glycerol, with a consequent lowering of the kinetic constant of association. To compensate for this effect, we extended the incubation time to 60 minutes prior to dilution with the matrix. As a control, spectra were acquired prior to incubation ( $t_i = 0$ ). Additionally, samples spent 15 minutes air-drying ( $t_{de} = 15$  min) before introduction into the source chamber. BmrA at a concentration of  $0.9 \mu M$  was incubated in the presence or absence of ATP/ $Mg^{2+}$ / $V_i$ . Three different P/L ratios (1:10, 1:100, 1:1000 n/n) were tested, with ATP and  $Mg^{2+}$  in excess by a factor of 10 over  $V_i$ .



**Figure 6.** Effect of ligand (ATP/Mg<sup>2+</sup>/Vi) concentration on the stability of the WT dimer (0.9 μM). **A.** t<sub>i</sub> = 0 **B.** t<sub>i</sub> = 60 minutes. The matrix/sample solution was HAG 75%, water 7%, and glycerol 18%. Spectra were accumulated over 200 shots.



As seen in Figure 6.A, at  $t_i = 0$ , dimer stability increases when the ligand is added at the 1:10:10:1 (BmrA/ATP/Mg<sup>2+</sup>/V<sub>i</sub>) ratio. This shows that, within the dead-time of the experiment ( $t_{de}+t_{ivr}$ ), the ligand associates with the protein in the matrix/sample solution and provides some stabilization of the WT dimer. However, the highest ligand concentration prevents the ionization of the complex due to the presence of sodium salts in the ATP/Mg<sup>2+</sup>/V<sub>i</sub> solution. At  $t_i = 60$  minutes (Figure 6.B), we observe the same trend: the higher the ligand concentration, the higher the dimer stability. Furthermore, the stabilizing effect as measured through dimer percentage increases by factor of 1.17 for the 1:100 (n/n) P/L ratio (Table 1). Thus, the ligand stabilizes the WT dimer.

When we compare Figure 6.A with Figure 6.B, we observe that increasing the time of incubation increases the dimer percentage, even in the absence of ligand. This could be caused by the evaporation of some of the residual water content within the glycerol solution after one hour, although sample volume was not visibly affected at this time-point. Lowering the water content would enhance the stability of the BmrA WT dimer.

**Table 1.** Effect of ligand (ATP/Mg<sup>2+</sup>/V<sub>i</sub>) concentration on the BmrA WT dimer stabilization.

Ligand ratio	t <sub>i</sub> = 0		t <sub>i</sub> = 60 min		Relative % increase over time
	Dimer %	Relative % increase	Dimer %	Relative % increase	
0	65	1.00	77	1.00	1.18
10	72	1.11	86	1.12	1.19
100	71	1.09	90	1.17	1.27
Saturation*	100	1.54	100	1.30	1.00

\* Theoretical estimation of the relative increase at the saturation point to achieve 100% of dimer stabilization.

The mass of the BmrA dimer measured during quality control of the sample is 132 837 Da, which corresponds to the [2-607] sequence (data not shown). The theoretical mass of the WT BmrA dimer without initiator methionine is 132 822 Da. In NALIM, the mass is shifted by 578 Da at  $t_i = 0$ . In the presence of ligand, the mass shifts further up by a mass lower than the mass of the ligand (Table 2). At  $t_i = 60$ , a mass shift of 890 is observed only for the 1:100 ratio. This could reflect the binding of ADP molecules to the dimer. However, due to the presence of Na<sup>+</sup> and Mg<sup>2+</sup> ions and the low resolution of the technique in this high  $m/z$  range, it is not possible to precisely determine the number of ADP molecules and salt ions bound to the WT dimer.

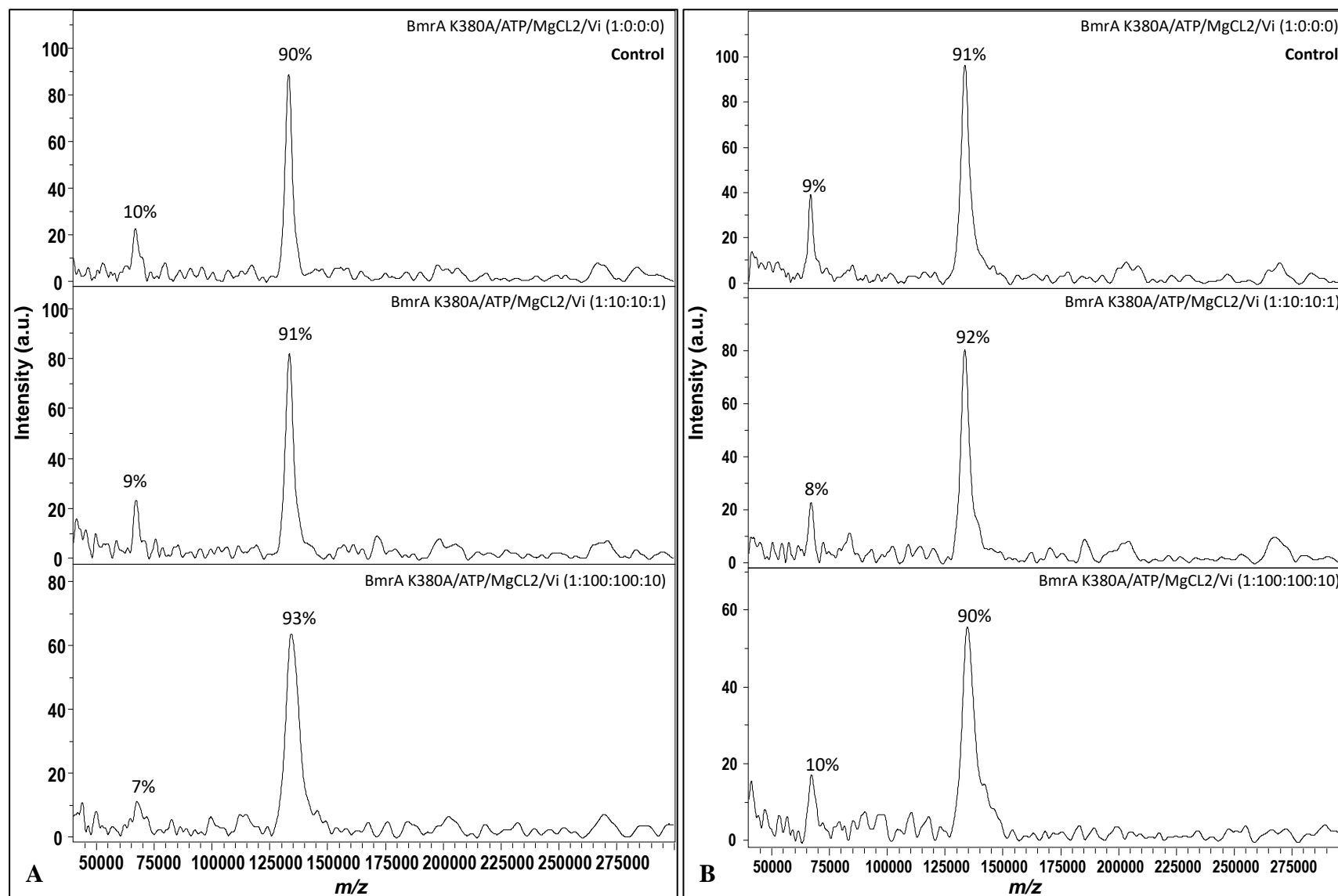
**Table 2.** Mass found for the BmrA WT in the P-L experiment.

Ligand ratio	ti = 0			ti = 60 min		
	Mass Observed (Da)	$M_{\text{obs}} - M_{\text{theo}}^*$	Delta Mass	Mass Observed (Da)	$M_{\text{obs}} - M_{\text{theo}}^*$	Delta Mass
0	66563	152		66684	273	
	133400	578		133646	824	
10	66518	107		66643	232	
	133451	629	52	133436	614	-210
100	66778	367		66991	580	
	133577	755	178	134536	1714	890

\* Theoretical mass ( $M_{\text{theo}}$ ) of BmrA A[2-607]H = 132 822 Da.

In order to confirm that the increase in WT dimer stability is an effect of ligand binding, we needed to control for ATP hydrolysis using a mutant of BmrA that has lost its ATPase activity, so that the ADP-BmrA complex cannot be trapped by  $V_i$  for stabilization.

In the K380A mutant, the replacement of the invariant lysine of the Walker-A motif in the WT protein with an alanine prevents ATP hydrolysis (Javed et al., 2022). Thus, the addition of ligand and trapping agent to the incubation is not expected to improve the K380A dimer stability as it did in the WT.



**Figure 7.** Effect of ligand (ATP/Mg<sup>2+</sup>/Vi) concentration on the stability of the K380A dimer (0.9  $\mu$ M). **A.**  $t_i = 0$  **B.**  $t_i = 60$  minutes. The matrix/sample solution was HAG 75%, water 7%, and glycerol 18%. Spectra were accumulated over 200 shots.

The highest ligand ratio was omitted in this experiment as it prevents ionization. As can be seen in Figure 7, the K380A dimer percentage is roughly constant as a function of ligand concentration at both  $t_i = 0$  min and  $t_i = 60$  min. This confirms that the BmrA WT dimer is stabilized by ADP binding in presence of  $V_i$ , while the K380A dimer cannot benefit from this effect.

Within the error margins of the experiment, we also observe that the stability of the K380A mutant is higher than the WT protein at  $t_i = 0$  and  $[L] = 0$ . Possible interpretations for this result are included in the discussion.

**Table 3.** Mass found for the BmrA K380A in the P-L experiment.

Ligand ratio	$t_i = 0$		$t_i = 60$ min	
	Mass Observed (Da)	Delta Mass	Mass Observed (Da)	Delta Mass
0	66846		66863	
	133275		133724	
10	67118		67096	
	133579	303	133679	-45
100	67723		67289	
	134610	1335	134842	1118

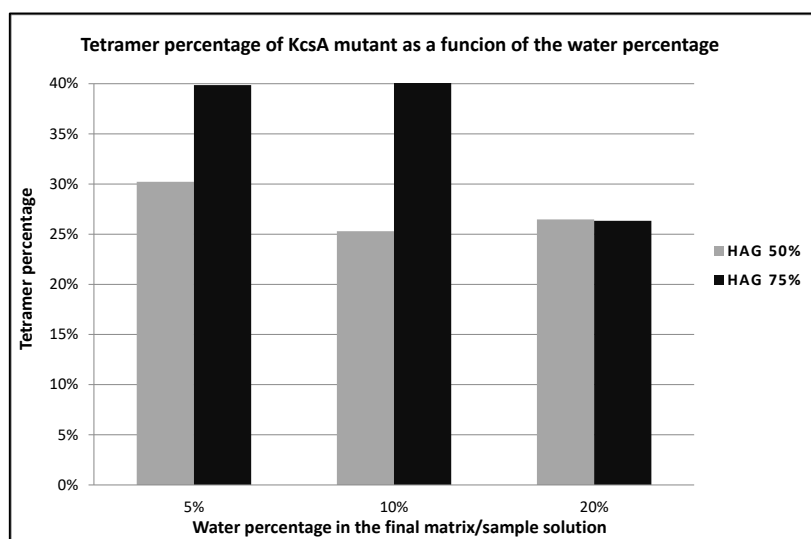
## Localizing the Tx7335 binding site on KcsA

To determine if the Tx7335 toxin binds to the ion pore of the KcsA K<sup>+</sup> channel or at a distant site, we planned a competition experiment with a known pore ligand. To this end, a chimera of KcsA with the pore structure of the Shaker eukaryotic K<sup>+</sup> channel is available to test the ability of a toxin to compete out charybdotoxin (ChTX). In the absence of information about the potential destabilizing effect of the toxin on the channel, trying to achieve the best possible tetramer stability was considered as a prerequisite to this study.

## Tetramer stabilization

### *Matrix/sample solution composition*

The water percentage and the matrix/sample ratio were shown in the above section to have a direct influence on the BmrA dimer stability. Thus, we investigated the combined effect of these parameters on the KcsA tetramer stability. To this end, three water percentage values were tested at two matrix/sample ratios.



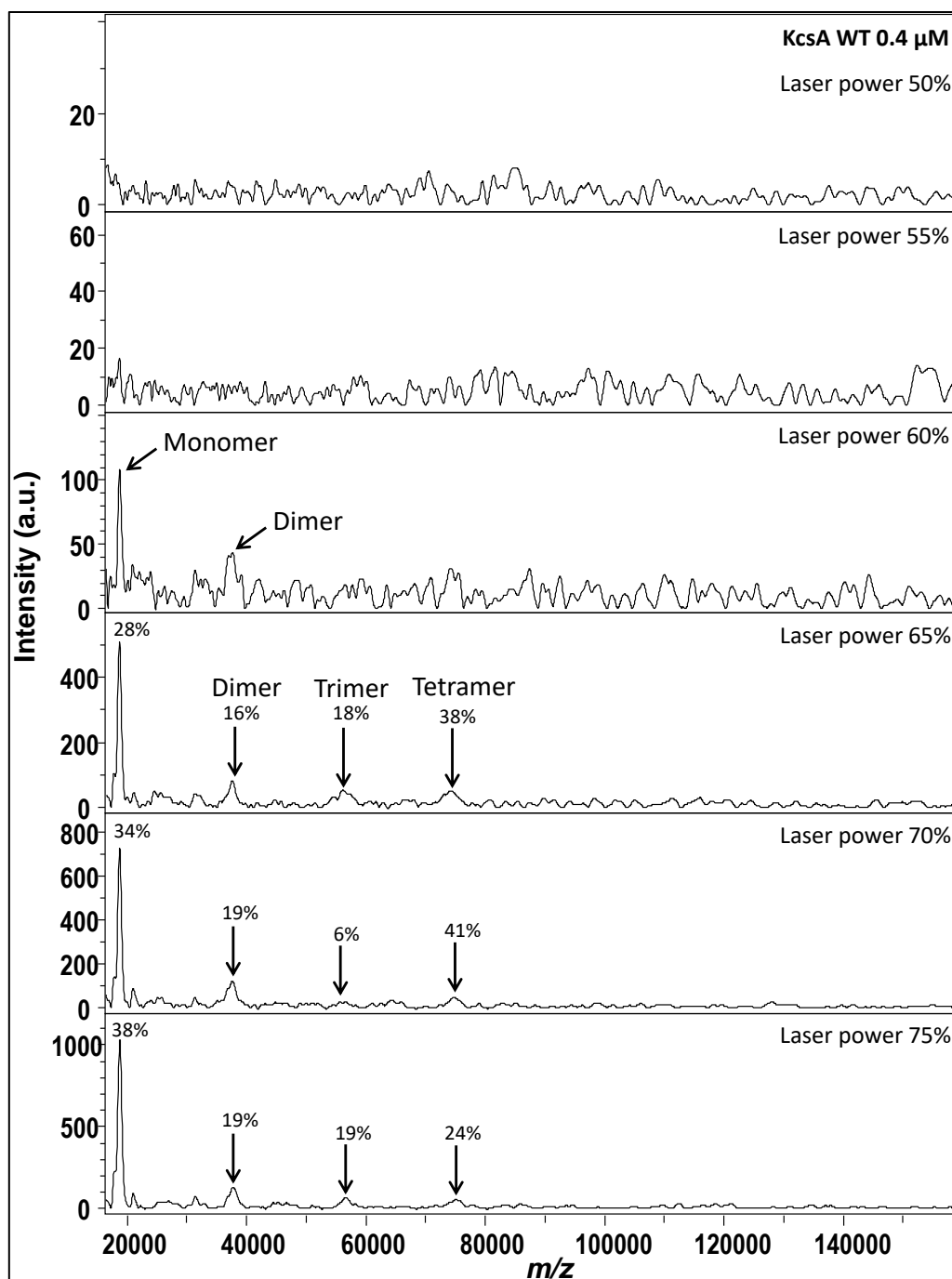
**Figure 8.** Effect of the water percentage on the stability of the mutant tetramer (0.4  $\mu$ M). **Grey.** HAG= 50%. **Black.** HAG= 75%. Spectra were accumulated over 400 shots. Laser power 70%.

Figure 8 shows that the tetramer percentage decreases when the water percentage increases. For both matrix/sample ratios, a high water percentage is detrimental to the tetramer stability.

Additionally, better tetramer stability is achieved with a higher matrix/sample ratio. Up to 10% water, the KcsA tetramer seems to better tolerate the presence of water at a higher matrix/sample ratio, although this effect is lost at 20% water.

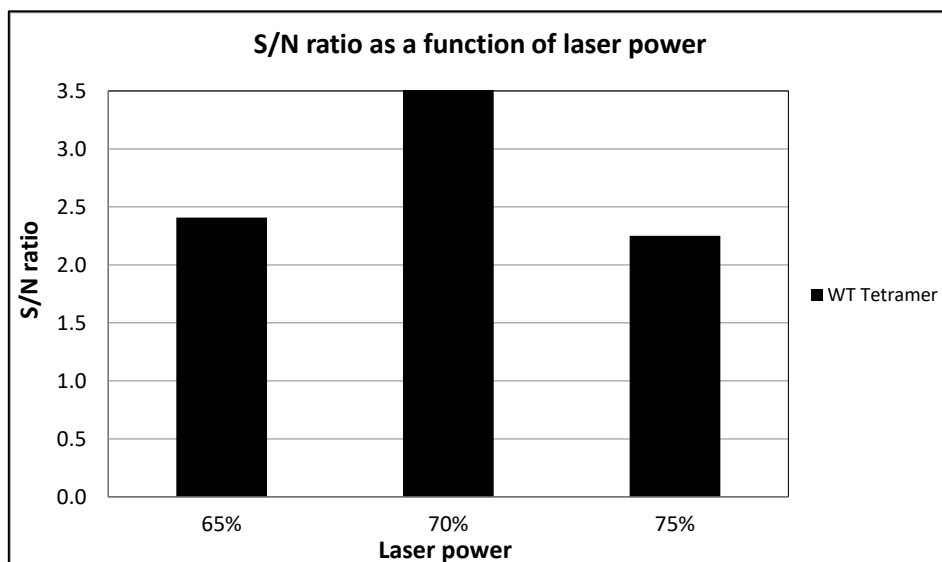
### ***Instrumental optimizations: Laser power.***

When working with KcsA, we observed that the energy necessary to reach the threshold of ionization was slightly different from the energy required for other membrane proteins. It is known that additives present in the sample solution can affect ionization efficiency. Since the stability of a complex depends on the amount of energy that is injected into it, it stands to reason that it may be affected by the laser fluence. We thus investigated the effect of the laser power parameter, a proxy for laser fluence, on tetramer stability.



**Figure 9.** Effect of the laser power on the WT tetramer stability. The matrix/sample solution was HAG 75%, water 6.5%, glycerol 18.5%. Spectra were accumulated over 400 shots.

As seen in Figure 9, the threshold of ionization is reached at 60% laser power for the monomer and dimer, and at 65% for the trimer and tetramer. Moreover, the highest percentage of tetramer is achieved at 70% laser power. Thus, the laser fluence exhibits an influence on the ionization and stability of the tetramer.



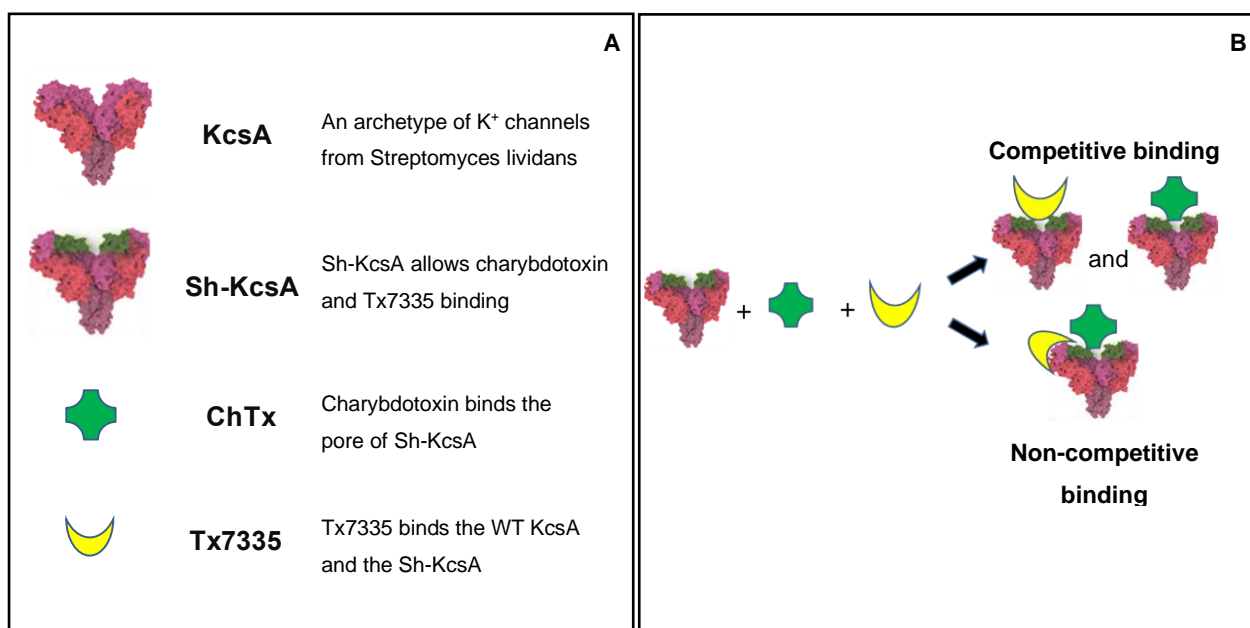
**Figure 10.** S/N ratio of the WT tetramer peak as a function of the laser power.

Figure 10 shows that this effect is clearly measured by looking at the S/N ratio. This confirms that laser power plays a key role in the ionization of the WT tetramer.

### **Competition binding experiment to determine the Tx7335 toxin binding site on KcsA**

The goal of this experiment is to figure out which binding site(s) is (are) targeted by Tx7335 on the K<sup>+</sup> channel. It relies on the knowledge that Tx7335 binds the WT protein, as this toxin was isolated through a pull-down assay using WT KcsA as a “bait” (Rivera-Torres et al., 2016). Charybdotoxin (ChTx) binds Sh-KcsA, a chimeric KcsA protein which mimics the pore structure of the eukaryotic Shaker channel so that it can bind this toxin (MacKinnon et al., 1998). Tx7335 also binds the charybdotoxin-sensitive Sh-KcsA channel (Rivera-Torres et al., 2016).

To figure out if Tx7335 targets the channel pore or a distant site, we designed a competition experiment between the Tx7335 and ChTx toxins for the chimeric protein. If the toxins do not share the same binding site on KcsA, we should be able to see concomitant binding of the toxins on the chimeric protein (Figure 11).



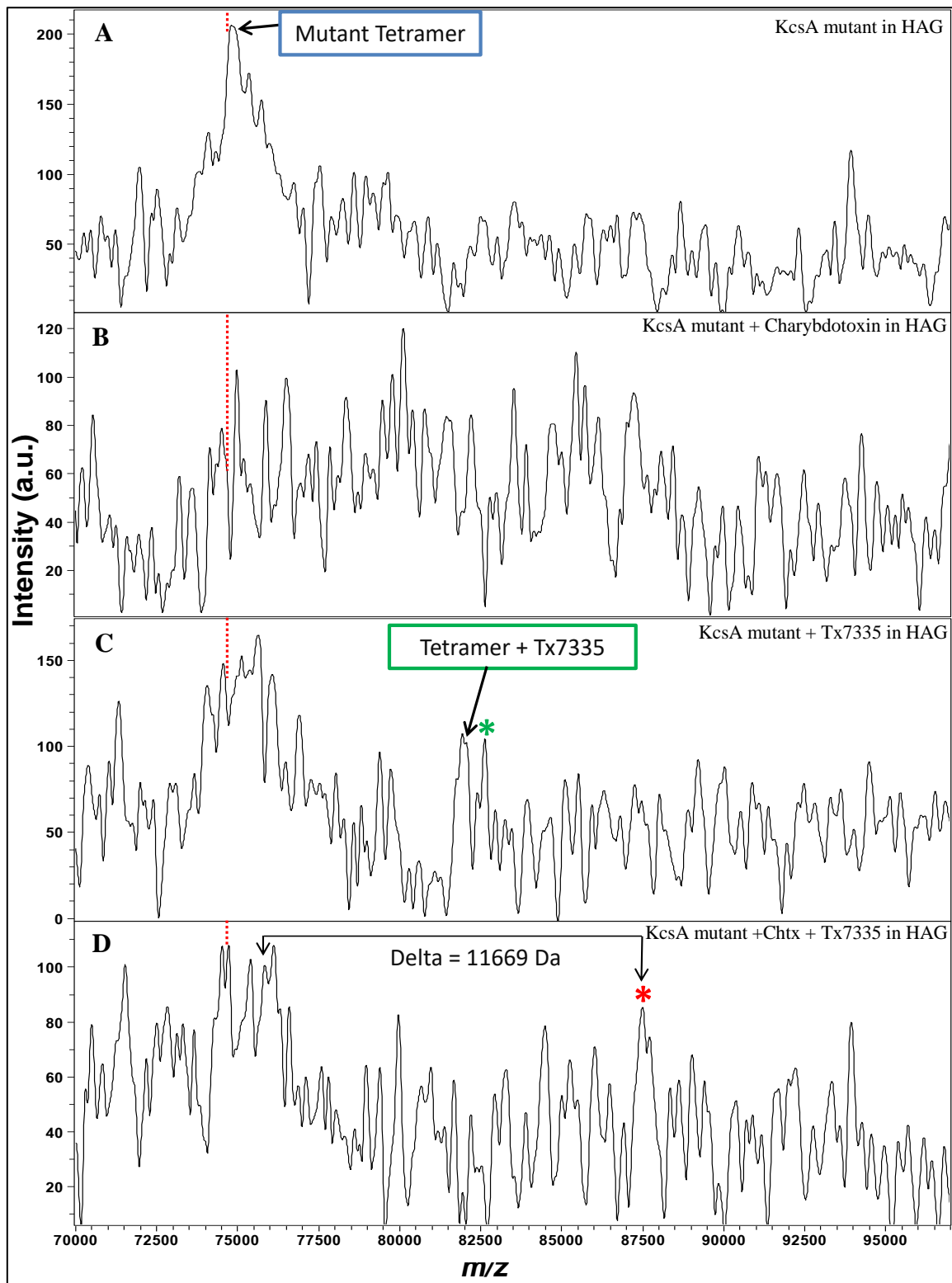
**Figure 11.** Schematic representation of the competition experiment. **A.** Channel variants and toxins used in the experiment. **B.** Competitive *versus* non-competitive binding for both toxins, depending on whether Tx7335 binds to the pore the same way as charybdotoxin, or to a distant site. The actual Tx7335 toxin binding site has not been localized so far.

To facilitate data interpretation, four different conditions can be tested:

1. Chimeric protein without toxins (negative control).
2. Chimeric protein with ChTx (ChTx control).
3. Chimeric protein with Tx7335 (Tx7335 control).
4. Chimeric protein with both toxins (competition experiment).

Conditions 2 and 3 are not absolutely necessary as this information is known. They can be used to assess the effect of the toxin buffers individually.





**Figure 12.** MALDI-TOF MS analysis of the competition experiment using the NALIM method. Spectra are zoomed in a  $m/z$  range from 70 000 to 97 000. **A. Negative control** ShKcsA (0.4  $\mu\text{M}$ ). **B. ChTx control** ShKcsA (0.4  $\mu\text{M}$ ) and ChTx (3.6  $\mu\text{M}$ ). **C. Tx7335 control** ShKcsA (0.4  $\mu\text{M}$ ) and Tx7335 (3.6  $\mu\text{M}$ ). **D. Competition experiment** ShKcsA (0.4  $\mu\text{M}$ ) with ChTx (3.6  $\mu\text{M}$ ) and Tx7335 (3.6  $\mu\text{M}$ ). The matrix/sample solution was HAG 75%, water 8%, and glycerol 17%. Spectra were accumulated over 200 shots. **Red dotted lines** show the position of the theoretical  $m/z$  for ShKcsA tetramer in spectra. See the text for the meaning of asterisks.

In the negative control (Figure 12.A), we observe the ShKcsA tetramer peak, which is heterogeneous. The main peak has a mass shift of *circa* +106 Da with respect to the theoretical mass (74 704 Da). Additionally, we observe two satellite peaks with delta masses of +546 Da and +935 Da with respect to the tetramer peak (Table 4).

**Table 4.** Mass determination, experimental errors and delta masses of peaks observed in the ChTx/Tx7335 toxins competition for ShKcsA binding experiment.

Condition tested	Hypothetical subunits				$M_{\text{obs}} - M_{\text{theo}}^*$ (Da)	Delta mass
	Channel monomers	ChTX	Tx7335	K <sup>+</sup>		
ShKcsA control	4				106	
	4				652	546
	4				1041	935
ShKcsA+ Tx7335	4				126	
	4				283	158
	4				675	549
	4				943	818
	4		1		43	7252
	4		1		590	547
Competition experiment	4				16	
	4				234	218
	4				1131	1115
	4	1	1		1169	11669

\* Theoretical mass ( $M_{\text{theo}}$ ) of: ShKcsA = 74 704 Da, ChTX = 4 319 Da, and Tx7335 = 7 335 Da.

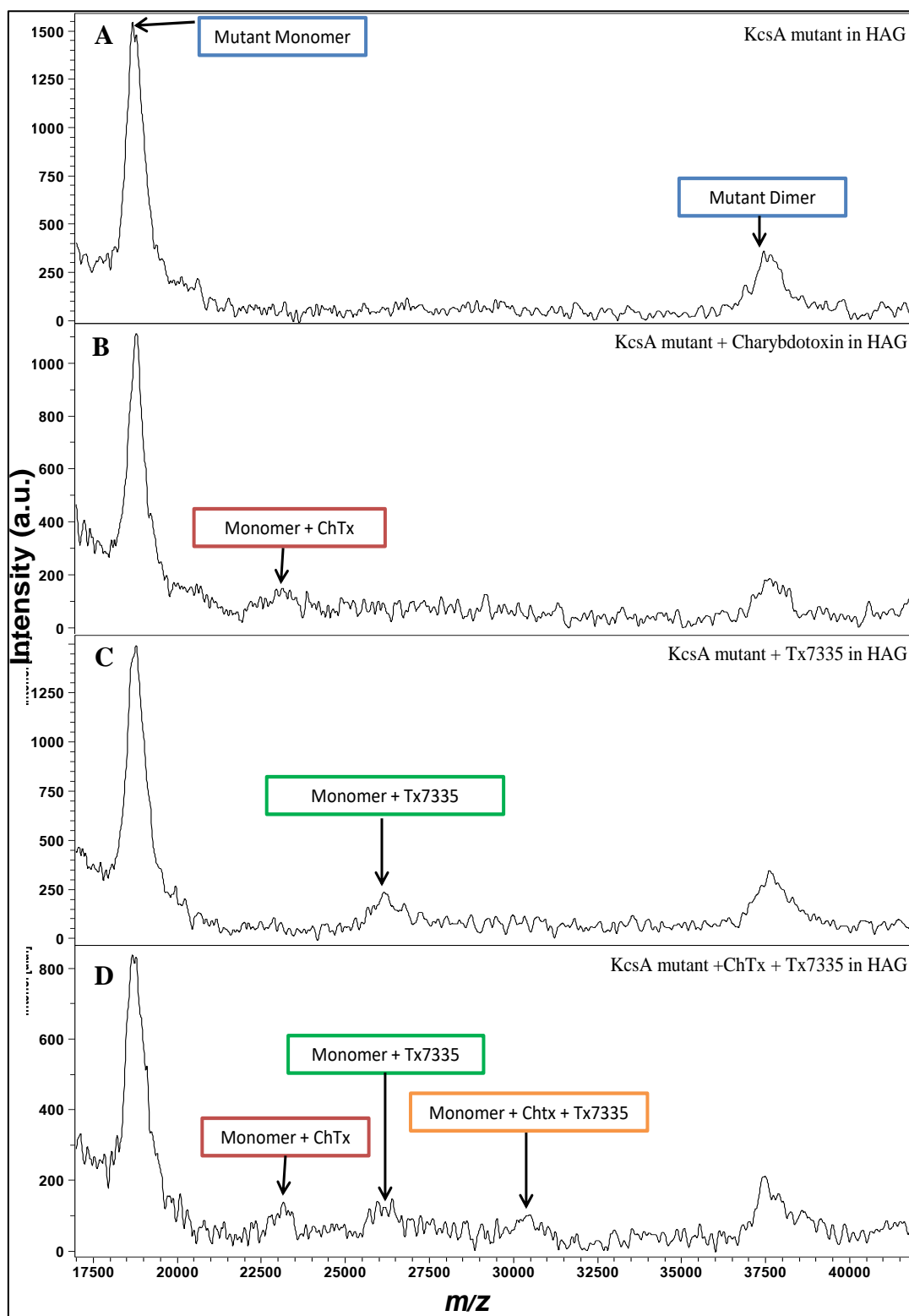
In the ChTx control (Figure 12.B), it was not possible to clearly identify any peak related to the tetramer or the tetramer bound to the toxin in their expected  $m/z$  range. We hypothesize that salts coming from the toxin solution prevent an efficient ionization.

In the Tx7335 control (Figure 12.C), we observe a peak with a mass shift of +43 Da with respect to the theoretical mass of the tetramer bound to the toxin (82 039 Da). Additionally, we find another peak with a delta mass of +547 Da with respect to this peak (labeled with a green asterisk).

In the competition experiment (Figure 12.D), we see a series of peaks potentially corresponding to adducts of ShKcsA (Table 4). We could not clearly identify a peak related to the theoretical mass of the tetramer in complex with a single Tx7335 or ChTx molecule. However, we observe a peak at 87 504  $m/z$  (labelled with a red asterisk), with a delta mas of 11 669 Da with respect to one of the tetramer adduct peaks. This delta mass is +40 Da higher than the sum of the theoretical mass of both

toxins (11 631 Da). This +40 difference could correspond to the mass of a  $K^+$  ion. The satellite tetramer peak from which this 11 669 delta mass was measured, has a mass shift of +1 131 Da.

When focusing on the lower  $m/z$  range, we found trimer, dimer, and monomer forms of ShKcsA, as was observed in KcsA experiments (Figure 9). Since these forms are not present in solution, they can be assumed to be produced by gas-phase dissociation of the tetramer form. These dissociated forms show more evidence of the protein toxin interaction, as shown for the monomer in Figure 13.



**Figure 13.** MALDI-TOF MS analysis of the competition experiment using the NALIM method. Spectra are zoomed in a  $m/z$  range from 17 000 to 42 000. **A. Negative control** ShKcsA (0.4  $\mu\text{M}$ ). **B. ChTx control** ShKcsA (0.4  $\mu\text{M}$ ) and ChTx (3.6  $\mu\text{M}$ ). **C. Tx7335 control** ShKcsA (0.4  $\mu\text{M}$ ) and Tx7335 (3.6  $\mu\text{M}$ ). **D. Competition experiment** ShKcsA (0.4  $\mu\text{M}$ ) with ChTx (3.6  $\mu\text{M}$ ) and Tx7335 (3.6  $\mu\text{M}$ ). The matrix/sample solution was HAG 75%, water 8%, and glycerol 17%. Spectra were accumulated over 200 shots.

In the negative control (Figure 13.A), we observe a peak with a mass shift of +11 Da with respect to the theoretical ShKcsA monomer mass (18 676 Da) (Table 5), which itself has a satellite peak with a delta mass of +101 Da.

In the ChTx control (Figure 13.B), we observe a peak with a mass shift of +120 Da with respect to the ShKcsA monomer peak. A peak with a delta mass of 4 340 Da, which is 44 Da higher than the theoretical mass of ChTx is present.

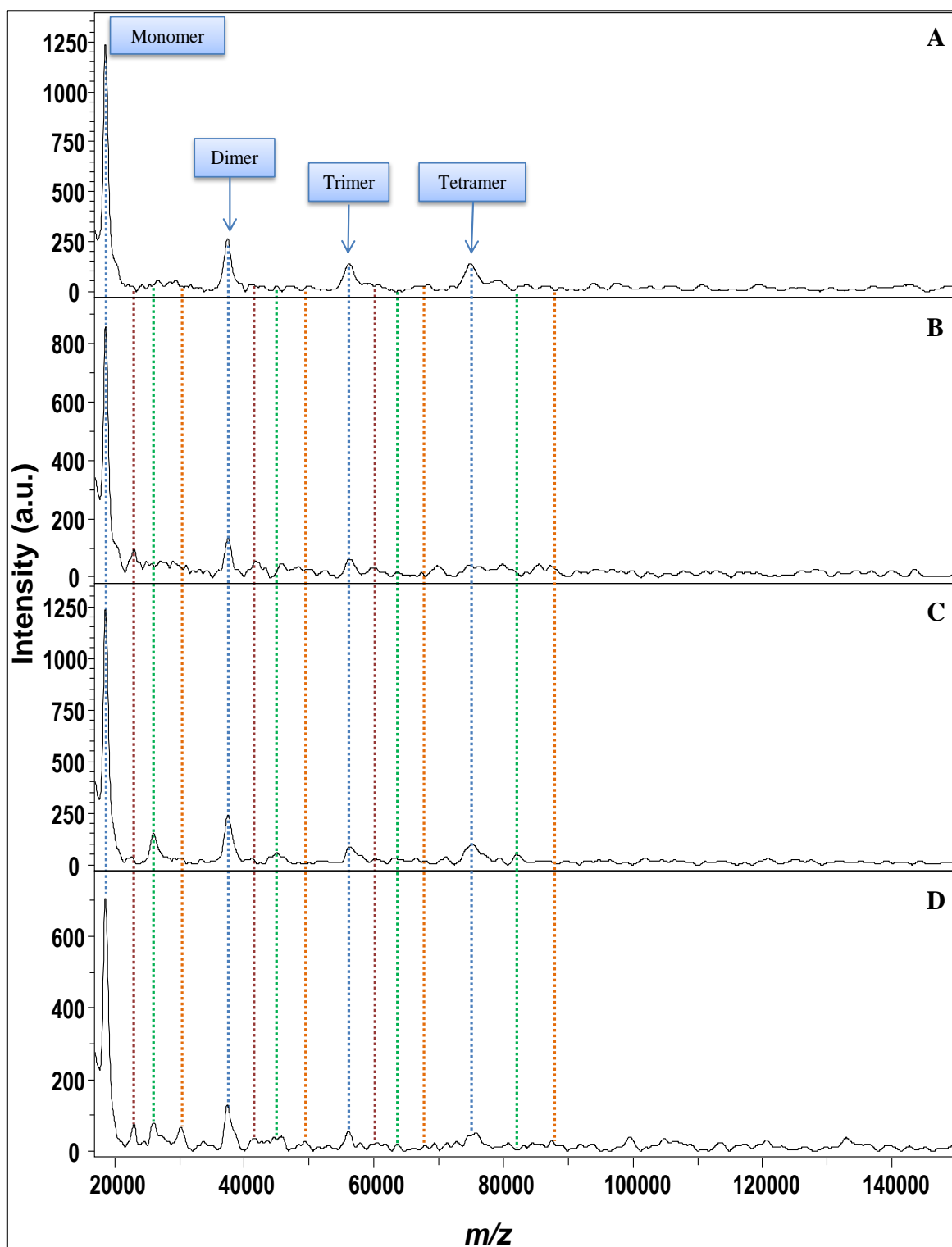
**Table 5.** Mass determination of the mutant monomer and the mutant monomer bound to toxin(s).

Condition tested	Observed $m/z$	Hypothetical subunits				$M_{\text{obs}} - M_{\text{theo}}^*$ (Da)	Delta mass
		Channel monomers	ChTX	Tx7335	$K^+$		
ShKcsA control	18688	1				11	
	18789	1				112	101
ShKcsA + ChTx	18797	1				120	
	23136	1	1			164	4340
ShKcsA + Tx7335	18782	1				105	
	26153	1		1		141	7371
Competition experiment	18670	1				-7	
	18783	1				107	114
	23160	1	1			187	4376
	26157	1		1		145	7373
	30409	1	1	1		101	11625

\* Theoretical mass ( $M_{\text{theo}}$ ) of: (a) KcsA monomer = 18 676 Da, (b) ChTX = 4319 Da, and (c) Tx7335 = 7335 Da.

In the Tx7335 control (Figure 13.C), we observe a peak with a mass shift of +105 Da with respect to the monomer mass. At a delta mass of 7 371 Da is another peak, which is 36 Da higher than the theoretical mass of the toxin.

In the competition experiment (Figure 13.D), we see a mass shift of -7 Da compared to the monomer mass. Next to this peak is a satellite peak with a delta mass of +114 Da. Based on this satellite peak, we further identify peaks that may correspond to the monomer bound to one or both toxins as follows. The peak with a delta mass of 4 376 Da is 80 Da higher than the theoretical mass (4 296 Da) of ChTx. The peak with a delta mass of 7 373 Da is 38 Da higher than the theoretical mass of Tx7335, which could correspond to the monomer associated with Tx7335 and one  $K^+$  adduct. The peak with a delta mass of 11 625 is 6 Da lower than the sum of the theoretical mass of both toxins (11 631 Da). Additionally, we do not identify other peaks that would suggest a different stoichiometry than 1:1:1 (KcsA/ChTx/Tx7335). This peak shows that ShKcsA monomer is bound to both toxins.



**Figure 14.** MALDI-TOF MS analysis of the competition experiment using the NALIM method. Spectra are zoomed in a  $m/z$  range from 15 000 to 150 000. **A. Negative control** KcsA mutant (0.4  $\mu\text{M}$ ). **B. ChTx control** KcsA mutant (0.4  $\mu\text{M}$ ) and ChTx (3.6  $\mu\text{M}$ ). **C. Tx7335 control** KcsA mutant (0.4  $\mu\text{M}$ ) and Tx7335 (3.6  $\mu\text{M}$ ). **D. Competition experiment** KcsA mutant (0.4  $\mu\text{M}$ ) with ChTx (3.6  $\mu\text{M}$ ) and Tx7335 (3.6  $\mu\text{M}$ ). The matrix/sample solution was HAG 75%, water 8%, and glycerol 17%. Spectra were accumulated over 200 shots. **Blue dotted lines** show the position of KcsA. **Red dotted lines** show the position of KcsA bound to ChTx. **Green dotted lines** show the position of KcsA bound to Tx7335. **Orange dotted lines** show the position of KcsA bound to ChTx and Tx7335.

A similar analysis was performed with the dimer and trimer forms of ShKcsA observed in NALIM. Figure 14 shows peaks possibly related to the dimer and trimer bound to the toxins as was hypothesized with the monomer. As it is unlikely that the monomer, dimer, or trimer are present in solution, these peaks could originate from the gas-phase dissociation of the tetramer complex formed with either or both toxins. In addition, the data suggest the presence of  $K^+$  adducts and/or bound ions on the channel with or without bound toxin.

Taken together, this data shows the concomitant binding of both toxins to the chimeric ShKcsA, indicating that Tx7335 and ChTx do not share the same binding site (Figures 12, 13 and 14). The stoichiometry of the ShKcsA-Tx7335 complex seems to be 1:1. This could further be confirmed through a titration experiment.

The identity of the molecule(s) corresponding to the ~546 Da added mass is not known. It should be noted that the 1 131 Da is compatible with two 546 Da "compounds" with an additional  $K^+$  adduct or ion bound to a non-protonated site.

### **Membrane proteins from membrane preparations or liposomal membrane mimics**

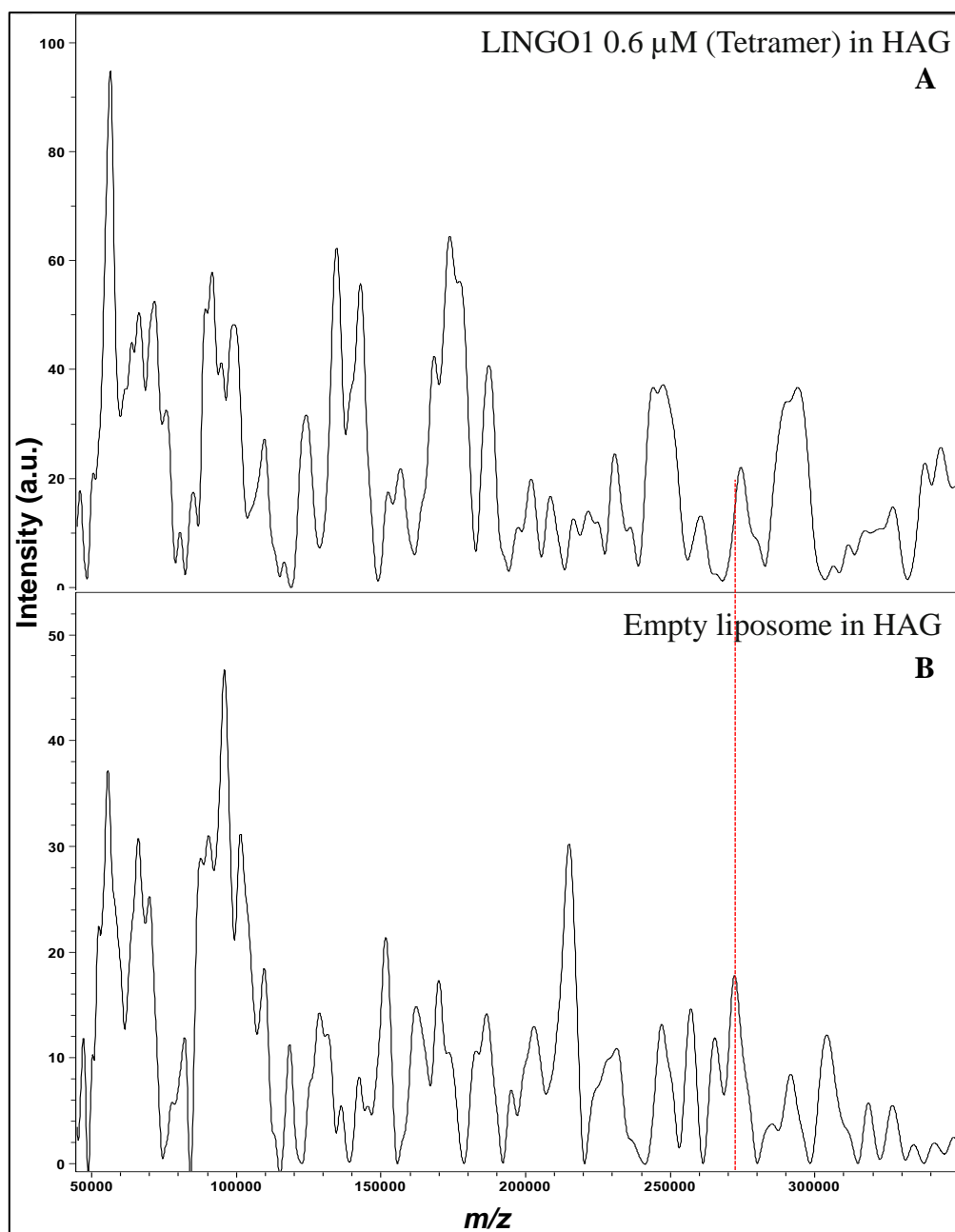
The above results were obtained on detergent-solubilized membrane proteins. To extend the usability of the method for sample preparations that are closer to the *in vivo* context of a membrane protein, we attempted to do NALIM analysis on three types of samples: (i) a Leucine-rich co-receptor called LINGO1 solubilized in liposomes, (ii) an ABC transporter called SpABC in cell membranes vesicles, (iii) BmrA WT in cell membranes vesicles.

#### **LINGO1 in liposomes**

In order to gain information about the stoichiometry of LINGO1 prepared in liposomes, we analyzed the membrane protein in NALIM using optimized conditions previously defined with BmrA. Due to the low initial concentration of the sample (0.9  $\mu\text{M}$  for the tetramer), we performed a buffer exchange/concentration step into AA using ultrafiltration. However, the protein concentration after this step (2.5  $\mu\text{M}$  for the tetramer) was still too low to be compatible with a water percentage below 25% in the matrix/sample solution.

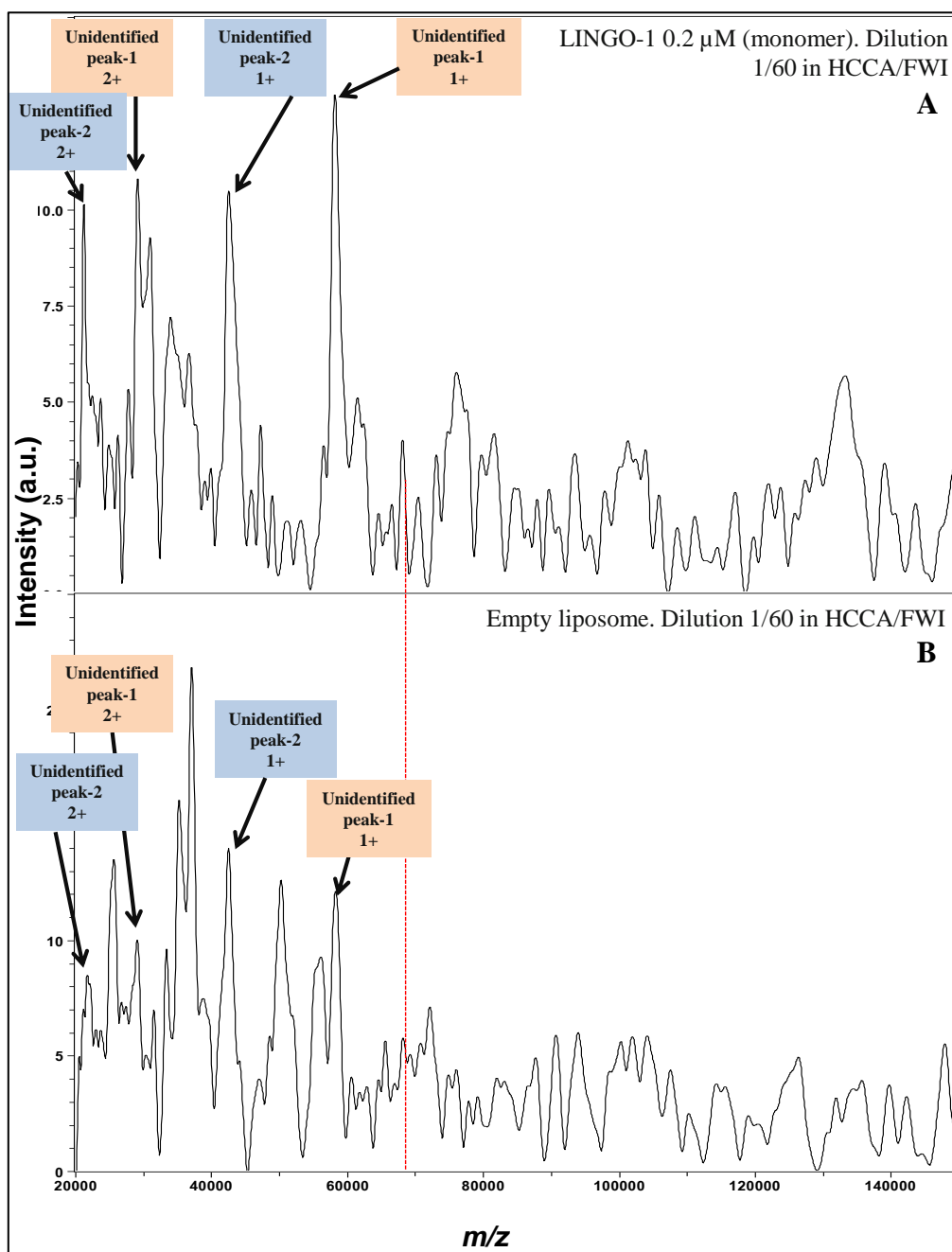
As a reference for background molecules, empty liposomes were analyzed in parallel with LINGO1-containing liposomes.





**Figure 15.** MALDI-TOF MS analysis of LINGO1 in HAG with NALIM method. The matrix/sample solution was HAG 75%, water 25%. Spectra were accumulated over 400 shots. **A.** LINGO1 in liposome 0.6  $\mu\text{M}$  (tetramer) **B.** Empty liposomes. **Red dotted line** shows the position of the theoretical mass of Lingo1 tetramer in spectra.

As seen in Figure 15.A, none of the peaks correspond to LINGO1 as either a dimer or tetramer. When compared to empty liposomes (Figure 15.B), neither do these peaks correspond to liposome complexes. Thus, it was not possible to identify a peak related to LINGO1 stoichiometry. To figure out why LINGO1 species are not observable at a 2.4  $\mu\text{M}$  monomer-equivalent concentration, we performed a quality control experiment in denaturing conditions using the UTL method (Cadene & Chait, 2000; Fenyo et al., 2007; Gabant & Cadene, 2008).



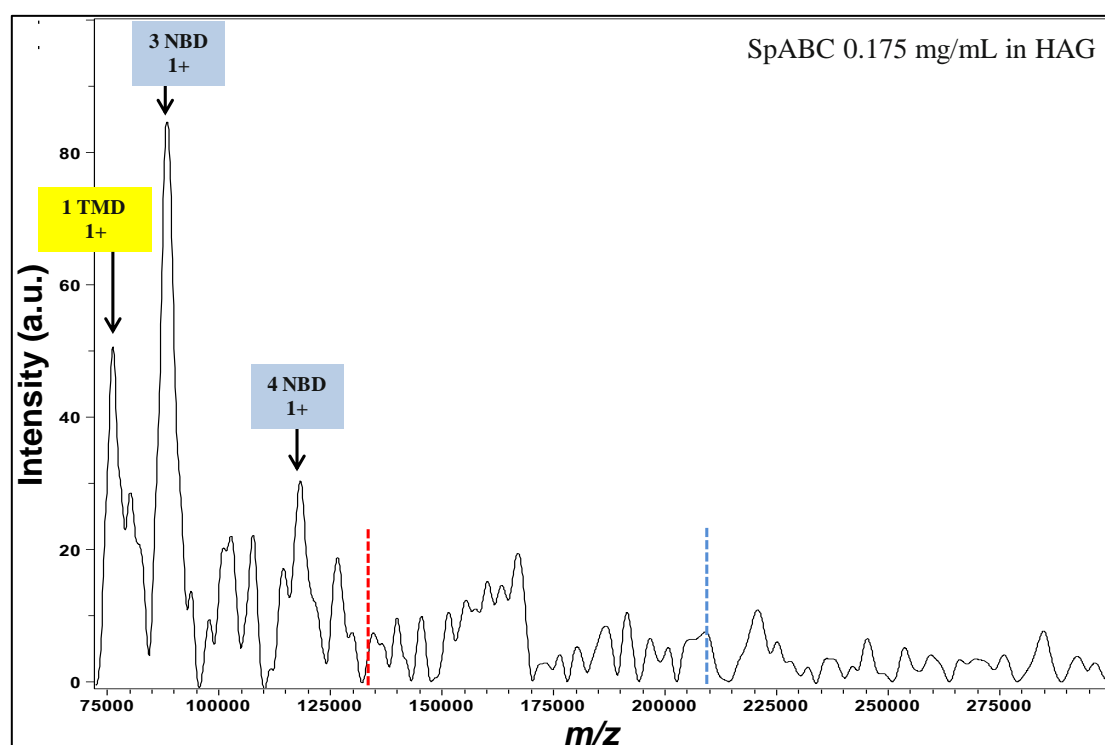
**Figure 16.** MALDI-TOF MS analysis of LINGO1 in denaturing conditions in HCCA/FWI using the UTL method. Spectra were accumulated over 200 shots. **A.** LINGO1 in liposomes at a concentration of 0.2  $\mu$ M (monomer equivalent). **B.** Empty liposomes. **Red dotted line** shows the position of the theoretical mass of Lingo1 monomer in spectra.

As can be seen in Figure 16, the sample contains other species. A search of average masses calculated from  $m/z$  values using  $z = 1$  or  $z = 2$  as likely charge states did not allow us to identify sequences related to the intact LINGO1 protein. Moreover, the main peaks in the LINGO1 analysis are found in the liposome control. This can be explained by the low LINGO1 concentration, an important limitation when analyzing this MP in NALIM.

### SpABC in cell membranes vesicles

SpABC is a BceAB-type ABC transporter present in *S. pneumoniae* and is expressed as two separate protein subunits, including a transmembrane domain (TMD) and a nucleotide binding domain (NBD) (Diagne et al., 2022). We set out to determine the stoichiometry of the complex. The two proposed hypotheses for the composition of the SpABC complex from TMD and NBD subunits were 1:2 or 2:2 (TMD:NBD), with corresponding average masses for monocharged species of 133 933 Da and 209 539 Da, respectively.

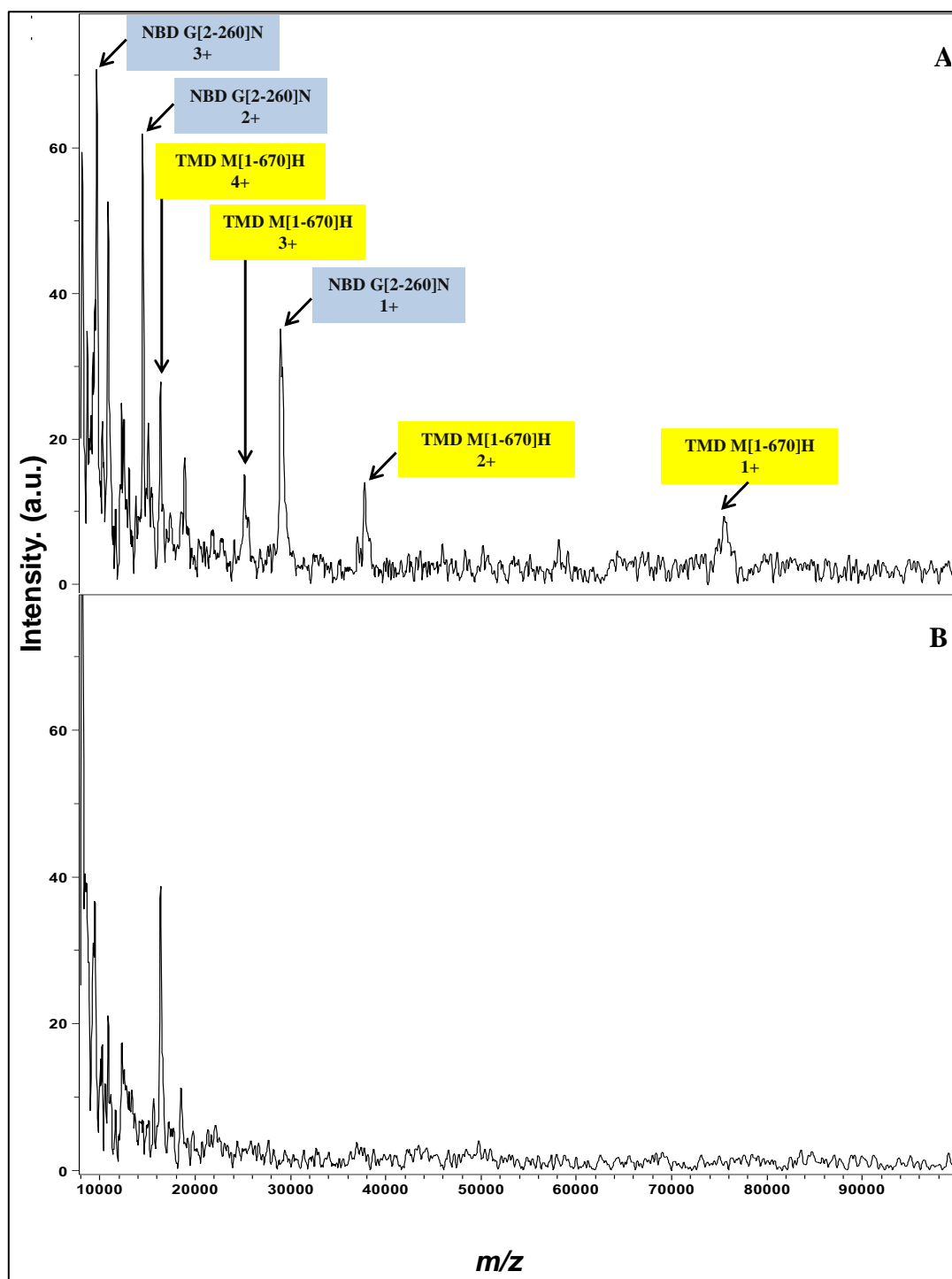
A preparation of SpABC solubilized in LMNG detergent was analyzed in NALIM in a preliminary experiment. In these conditions peaks related to the NBD or TMD proteins that form the SpABC complex can be identified (Figure 17).



**Figure 17.** NALIM MALDI-TOF analysis of SpABC (in LMNG) at 0.175 mg/mL in HAG using the NALIM method. The matrix/sample solution was HAG 50%, water 20%, glycerol 30%. The spectrum was accumulated over 400 shots. The red dotted line shows the position in the spectrum of the theoretical  $m/z$  for the monocharged 1:2 (TMD:NBD) complex. The blue dotted line shows the position in the spectrum of the theoretical  $m/z$  for the monocharged 2:2 (TMD:NBD) complex.

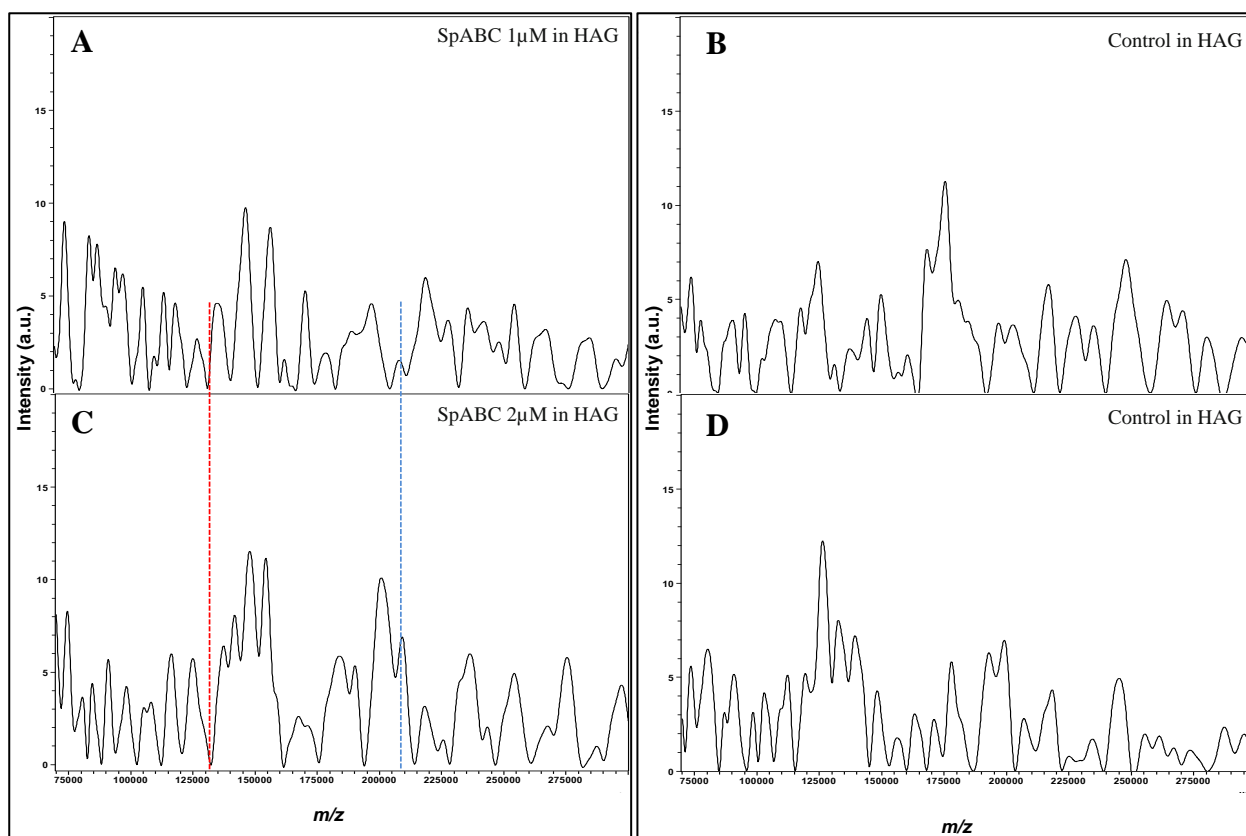
At 1.675  $\mu\text{M}$  (based on 2:2 stoichiometry), aggregates of the NBD subunits were observed. However, the hypothesized 1:2 or 2:2 (TMD:NBD) SpABC complexes were not observed. At concentrations higher than 1.675  $\mu\text{M}$  SpABC the ionization is suppressed (data not shown), while at concentrations lower than 0.42  $\mu\text{M}$  we observe mainly TMD and no NBD (data not shown).

Since the concentrations that are compatible with the observation of both subunits already produce aggregates of individual domain subunit, we hypothesized that conditions closer to those found *in vivo* could increase the stability of the complex, therefore making it easier to observe in NALIM at lower concentrations. Thus, we decided to attempt direct analysis of SpABC proteins embedded in membrane lipids by working with membrane vesicles (MVs). MVs were made from bacterial cell membranes using the protocol established by Chorev *et al.* (2020), using a membrane preparation from a strain of *E. coli* which overexpresses SpABC and localizes it to the membrane (Di Cesare *et al.*, 2022). As a control, a membrane preparation from a strain containing an empty plasmid was processed in parallel. The quality of the protein in MVs was assessed by analyzing a solution of MVs in denaturing conditions using the UTL method (Figure 18).



**Figure 18.** MALDI-TOF analysis of SpABC in denaturing conditions in HCCA/FWI using the UTL method. Spectra were accumulated over 200 shots. **A.** SpABC 0.6  $\mu$ M in MV **B.** Negative control: MV from bacteria transformed with an empty plasmid.

As seen in Figure 18.A, the spectrum of the SpABC sample shows peaks that are not present in the control (Figure 18.B). The average masses calculated from  $m/z$  values correspond to the sequence of intact TMD and NBD subunits, both without the initiator methionine, confirming that SpABC is present in the membrane vesicles. Thus, these samples were deemed appropriate for testing in NALIM.



**Figure 19.** MALDI-TOF MS analysis of SpABC in HAG using the NALIM method. The spectra were accumulated over 400 shots. **A.** SpABC in MV (1  $\mu\text{M}$ ) and the matrix/sample solution was HAG 75%, water 12.5%, glycerol 12.5%. **B.** Negative control. Empty MV at the same dilution and matrix/sample solution as A. **C.** SpABC in MV (2  $\mu\text{M}$ ) and the matrix/sample solution was HAG 75%, water 6%, glycerol 19%. **D.** Negative control. Empty MV at the same dilution and matrix/sample solution as C. The red dotted line shows the position in the spectrum of the theoretical  $m/z$  for the monocharged 1:2 (TMD:NBD) complex. The blue dotted line shows the position in the spectrum of the theoretical  $m/z$  for the monocharged 2:2 (TMD:NBD) complex.

The protein concentration was purposefully kept low in this experiment to prevent aggregation. As seen in Figures 19.A and 19.B, it is not possible to identify SpABC complexes either in a 1:2 or a 2:2 (TMD:NBD) stoichiometry from the SpABC membrane vesicles.

### Cell membrane vesicles (Spiking experiment)

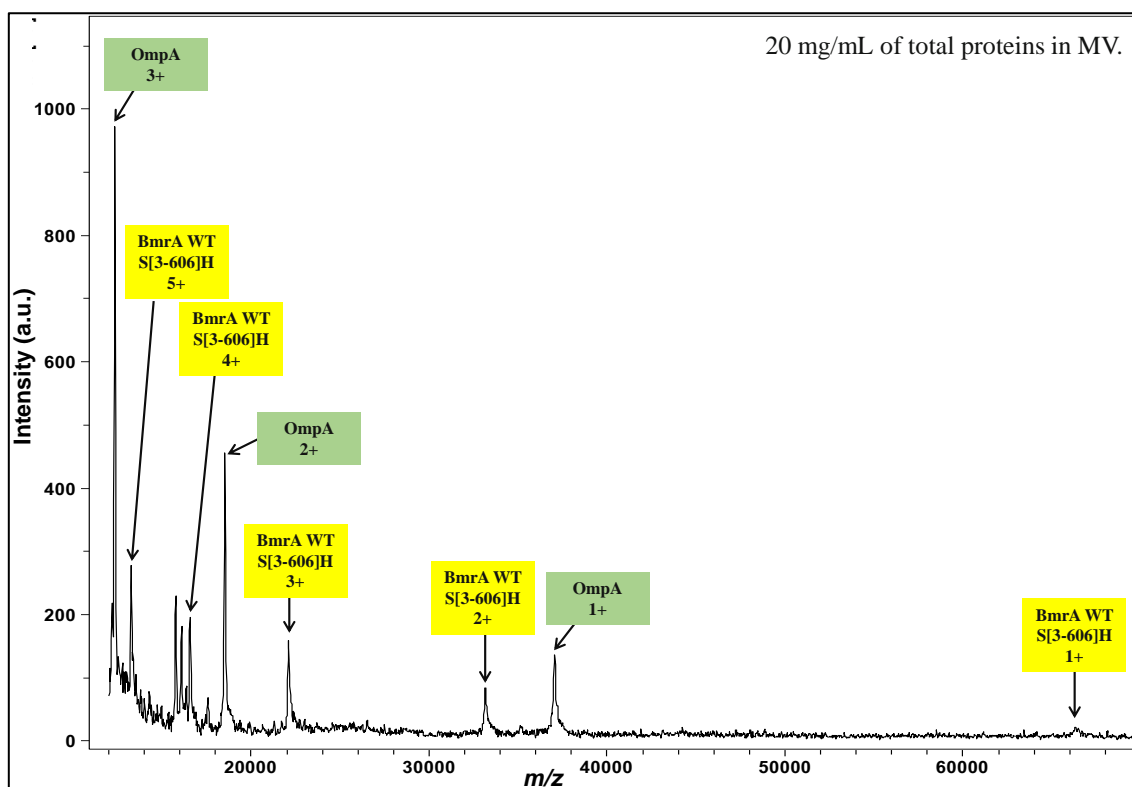
In the previous experiment, no signal for proteins could be detected from MVs. Several hypotheses can be made to explain the absence of detection of the protein of interest: (i) the protein embedded in the MV is not ejected during the ionization stage of NALIM; (ii) ion suppression due to contaminants prevents the detection of the protein of interest; (iii) contrary to initial assumptions, the membrane protein of interest is not stable enough in MVs to be detected. In order to test the aforementioned hypotheses, we designed the following experiment. Spiking consists in adding a small aliquot of a protein to the sample, to test both the effect the sample has on the spiked protein, and the effect the

spiked protein has on the protein of interest. By spiking the protein-containing MV with a protein that is known to give a good signal in NALIM, we can determine if the problem is due to ion suppression by a component of the vesicle. If the spiked protein is detected, it shows that ion suppression is not the source of the problem. In favorable cases, adding a spiking protein to the MV can even act as a kind of “sponge” for possible contaminants, so that both proteins can be ionized. If only the spiked protein is detected, we can conclude that ejection of the protein is the difficult part.

In addition to spiking with a protein, spiking with a detergent allows us to test if membrane lipids are too closely associated with the membrane protein, which could prevent ejection. The detergent could then facilitate ejection by substituting partially or completely for membrane lipids, and/or destabilizing the MV without compromising the protein complex stability.

Initially the spiking experiment was designed to improve signal on SpABC. However, since NALIM was well optimized for BmrA and optimum dimer complex stability was achieved in LMNG detergent for this protein, the spiking experiment was designed using BmrA instead. Thus, BmrA in MVs was analyzed by itself as a control, or in the presence of a second protein, or in the presence of detergent. As a positive control for detection, an analysis of the second “spiking” protein by itself was also performed. Spiking was thus performed with a soluble protein previously tested in NALIM, in this case an IgG which produces peaks that are not expected to overlap with BmrA-associated peaks. Detergent spiking was performed with LMNG, as the detergent allowed for detection of detergent-solubilized, purified BmrA (Figures 3, 5 and 6).

***Evidence of BmrA in cell membrane vesicles.*** Membrane vesicles from bacterial membranes containing WT BmrA were prepared using the protocol of Chorev *et al.* (2020). A quality control (QC) of our sample was performed in denaturing conditions using the UTL method (Figure 20).



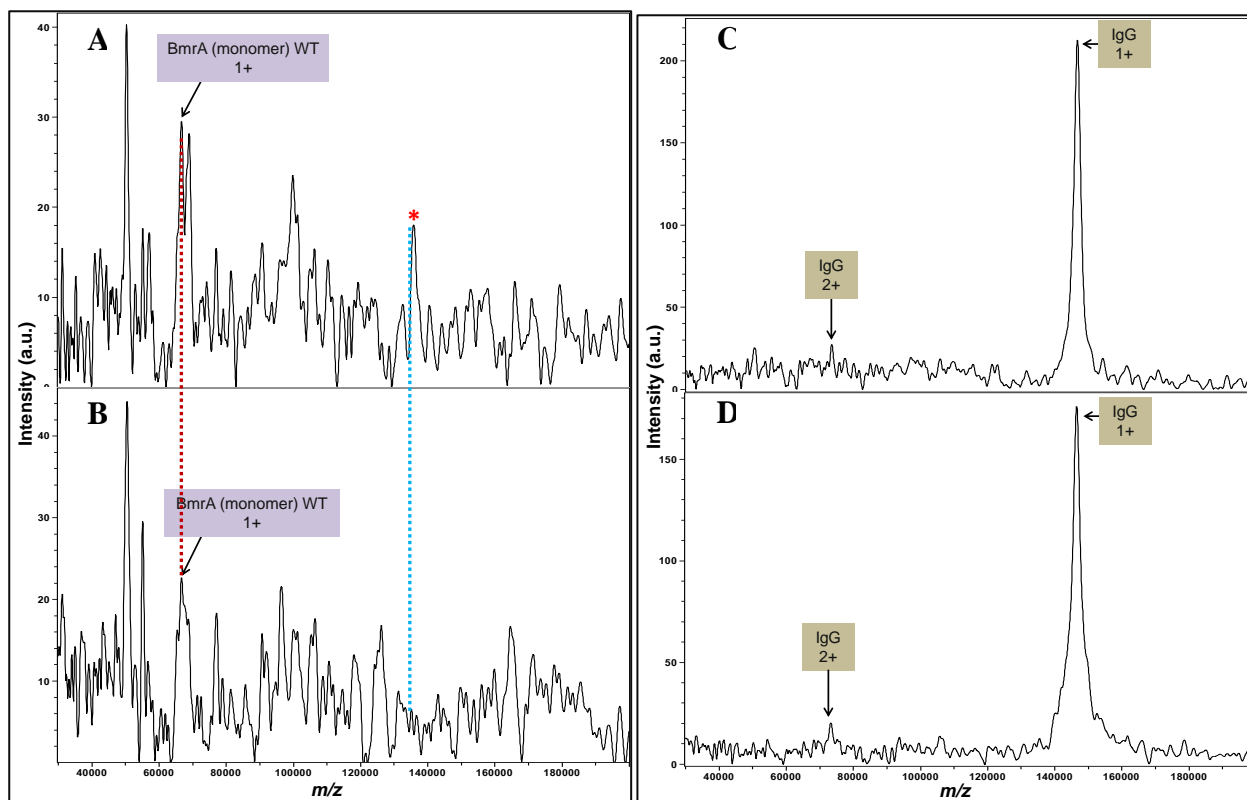
**Figure 20.** MALDI-TOF MS analysis of BmrA WT in denaturing conditions in HCCA/FWI using the UTL method. The MV sample was diluted 1/50 in HCCA/FWI to reach *circa* 0.4 mg/mL of total proteins. Spectra were accumulated over 200 shots.

Figure 20 shows that the BmrA protein is globally intact in MV. The average mass of the protein determined in the QC is indeed 66 342.5 Da, which corresponds to the theoretical mass calculated from the protein sequence without the first two amino acids Met and Ala (mass error of +2.5 Da, or 37 ppm).

Additionally, we observe another protein in MVs. Based on the average mass of this protein (37 064 Da), we identify OmpA as a possible match (mass error of -5.4 Da, or 145 ppm).

After sample preparation, the total protein concentration calculated from the absorbance measured at 280 nm was 20 mg/mL. Based on the UTL result, we roughly estimated that BmrA represents 20% of the total protein concentration. Thus, BmrA concentration would be close to 4 mg/mL, which represents a concentration of *circa* 51  $\mu$ M of BmrA dimer in MVs.





**Figure 21.** MALDI-TOF MS analysis of the spiking experiment in HAG with the NALIM method. Spectra were accumulated over 200 shots. **A.** BmrA in MV at a concentration of 1.5  $\mu\text{M}$ . The red asterisk shows a peak that may correspond to a BmrA dimer species **B.** BmrA in MV at a concentration of 1.5  $\mu\text{M}$  with LMNG at 0.02 mM. **C.** BmrA at 1.5  $\mu\text{M}$  with IgG at 0.2  $\mu\text{M}$ . **D.** IgG 0.2  $\mu\text{M}$ .

The red dotted line shows an unidentified peak at 50 376  $m/z$ . The blue dotted line shows peaks at 99 896 and 101 265  $m/z$ . For all tested conditions, the matrix/sample solution contained 75% HAG, 4% water, 21% glycerol.

As seen in Figure 21.A, a peak close to the mass of the BmrA monomer is observed in MVs. The mass found for this peak is 383 Da higher than the mass found in the QC, which itself corresponds to the sequence missing the first 2 amino acids (Table 6), with a satellite peak with a delta mass of +2 257 Da with respect to the monomer peak.

**Table 6:** Interpretation of masses found in BmrA expressing vesicles.

Protein	Observed <i>m/z</i>	Hypothetical subunits		$M_{\text{obs}} - M_{\text{theo}}$ (Da)	Delta mass
		Channel monomers	OmpA		
Unidentified peak	50376	1		-15965	
BmrA monomer	66724	1		383	
	68981	1		2640	2257
Unidentified peaks	99896			99895	
	101265			101264	
BmrA dimer	135836	2		3155	

\* Theoretical mass ( $M_{\text{theo}}$ ) of BmrA S[3-607]H : monomer = 66 340 Da, and dimer = 132 680 Da.

Interestingly, the peak at  $m/z$  135 836, labelled with a red asterisk, is shifted by +3 155 Da with respect to the [3-607] protein sequence. One possible explanation for this peak is that it corresponds to a BmrA dimer species with a ligand or ligands. In this case, the extra mass would suggest that a molecule or set of molecules, possibly lipids, is closely associated to the dimer or even trapped by within the dimer interface and forms a stable enough complex with BmrA to be ejected with it and detected in NALIM analysis. No peak related to any OmpA complex or monomer was detected. However, three additional peaks at  $m/z$  50 376, 99 896, and 101 265 do not have a direct correspondence with proteins found in the QC. They are also present in the detergent spiking condition (Figure 21.B), yet they are not observed in the SpABC-containing MVs, suggesting that they are specific to the BmrA expressing bacterial membrane. One possible interpretation is that during the sonication stage of MV preparation, shearing forces cleave off a peptide bond in BmrA so that shorter monomer and dimer forms are observed. Indeed, in the quality control, a cluster of peaks was observed from 15 700 to 16 300 Da, which could correspond to complementary fragments of the shorter BmrA dimer forms observed at  $m/z$  99 896, and 101 265. The peak at  $m/z$  50 376 could correspond to the monomer of these shorter BmrA dimers generated by gas phase dissociation.

As seen in Figure 21.B, when analyzing MVs in the presence of detergent, the only observed peak seem to correspond to the monomer BmrA. The average mass found for this protein is 66 700 Da, which is 363 Da higher than the mass found in the QC. This mass shift is coherent with the shift found in the absence of detergent, which could indicate the presence of adducts on the protein. The lack of dimer observation may be due to partial replacement of native lipids by LMNG, which may produce enough destabilization to push the dimer signal below the limit of sensitivity.

As seen in Figure 21.C, in the protein spiking condition, we only observe the IgG peak. Although we took care to leave BmrA concentration in a 5-fold excess over IgG concentration, the strong IgG signal apparently overwhelms BmrA-related signals. On the other hand, the IgG peak has a similar mass and S/N ratio as in the IgG control (Figure 21.D), suggesting that the MV itself does not bring any significant ionization-suppressing contaminants or adduct-forming salts.

## DISCUSSION

### Observation of specific membrane protein oligomers in NALIM

Previously, we conceived the NALIM approach and showed for the first time that it was possible to observe noncovalent complexes of soluble proteins directly from a liquid phase (Beaufour et al., 2018). Soluble proteins differ from membrane proteins by their hydrophilic surfaces, which makes them more amenable to being analyzed directly from the sample solution. In contrast, MPs require different molecules (e.g., lipids, detergents) to solubilize them and maintain their ternary and quaternary structures in solution. Attempting to extend NALIM to MPs requires the stability of the protein and their complexes to be ensured from the sample preparation to the detection inside the instrument.

In the present work, oligomers of different MPs were observed in the native state in NALIM. In detergent micelles, we were able to identify monocharged peaks as, respectively, the dimer of BmrA and the tetramer of KcsA (Figure 3 and 9). To ensure that any observed oligomer species are not the product of gas-phase aggregation, the final protein concentration inside the instrument was kept in the low micromolar or high nanomolar range ( $< 2 \mu\text{M}$ ). In fact, we do not observe any oligomeric state of membrane proteins higher than the ones expected from known information about these complexes. Moreover, the percentages and intensities of the BmrA dimer and the KcsA tetramer peaks are higher than their monomer and trimer peaks, respectively, whereas gas-phase aggregation usually results in a typical exponential decay (Daubenfeld et al., 2006). As a consequence, the lower oligomeric states observed in NALIM are best explained as a product of gas-phase dissociation of the complex, and the oligomers of MPs observed in NALIM are already present in solution prior to analysis. Interestingly, as noted above, KcsA dissociation seems to favor the dimer form over the trimer. Since no dimer was observed in solution, this observation suggests that in the tetramer, the interaction between dimers is stronger than between dimers and monomers, so that dimers stabilize each other, which leads to less trimers than dimers and tetramers. If true, this observation reinforces the notion that the lower-number oligomers we observe here is the result of gas-phase dissociation of the solution-phase complex rather than gas-phase aggregation of monomers, as was the case with the soluble HU protein heterodimer (Beaufour et al., 2018).

### Stability of MP complexes in NALIM

Complex stability is one of the cornerstones of native MS methods, since stability both ensures limited gas-phase dissociation and prevents spurious rearrangement of subunits. When evaluating which factors in NALIM can affect the stability of MP complexes for the purpose of maximizing it,

we identified possible areas of improvement as: (i) instrumental parameters, (ii) the matrix/sample ratio, and (iii) sample composition.

When examining the effect of instrumental parameters, we found that the laser power has an effect on oligomer stability. In NALIM, to ionize noncovalent complexes, a threshold of ionization has to be reached through adequate laser power (a proxy for laser fluence). A threshold of ionization is likewise observed for individual analytes in non-native MALDI methods. This parameter has to be meticulously tuned because, as reviewed by Bolbach (2005), the ionization of nonvolatile and labile molecules as proteins can only be achieved in a narrow range of “energy densities”. The energy in excess of what is needed for ionization could thus be detrimental for the sample, as was seen with the KcsA tetramer (Figure 9). This can be related to the energy absorbed by the matrix. The absorbed energy ( $E_a$ ) per volume unit ( $V$ ) is directly related to laser fluence ( $H$ ) and the matrix absorption coefficient ( $\alpha$ ), through the equation  $(1) E_a/V = H * \alpha$  (Hillenkamp & Karas, 2007). In MALDI, the energy absorbed by the matrix is dissipated mostly through the subsurface sputtering/explosions that occur during the desorption process. However, part of the absorbed energy can be dissipated by transfer to the analyte. Thus, if the absorbed energy is in excess of what can be dissipated through desorption, this excess will be passed on to the analyte, which can lead to prompt fragmentation and, in the case of complexes, disrupt their stability.

Incidentally, this can also explain why a higher matrix/sample ratio better preserves noncovalent complexes. A higher matrix concentration means more matrix molecules per volume unit, which potentially makes the absorption of energy more efficient, and consequently, decreases the amount of excess energy available to perturb noncovalent complexes.

At high matrix concentration, the stability of the MP complexes was found for BmrA and KcsA to also depend on the water percentage in the matrix/sample solution (Figure 4, Figure 8). MPs orient their hydrophilic domains and surfaces to be in contact with water, while their hydrophobic domains are sheltered within their structural core and the lipid bilayer. MP oligomer stability is thought to be based on hydrophobic interactions between the hydrophobic domains of each monomer, with potential stabilization by strategically situated salt bridges and other electrostatic interactions (Otzen, 2014; Stowell & Rees, 1995). Hence, a water concentration should favor the stabilization of MPs. However, BmrA dimer stability was improved by lowering the water percentage in the matrix/sample solution as seen in Figure 4. Since water and glycerol have an inverse relationship in the matrix/sample solution, the higher dimer stability at lower water percentage could be explained by a glycerol stabilization effect of the dimer, with glycerol acting as a surrogate for water with a higher stabilization efficiency. Upon transfer to the gas phase, most water is supposed to have

evaporated, removing one of the main forces driving hydrophobic interactions, and leaving other noncovalent interactions to maintain the structure. However, in MALDI-TOF MS, the timescale of the analysis is in the microsecond range, which is too short for gas-phase unfolding to have progressed to any significant degree, thus offering a distinct advantage over analyzers usually coupled to ESI.

Since NALIM, as all native MS methods, is sensitive to the intrinsic stability of protein complexes, it can reveal how different factors affect the stability of MPs (or other proteins) in native conditions. For example, a difference in stability between WT BmrA and the K380A mutant was observed. While gas-phase disruption involves different mechanisms than heat-induced melting of a protein structure, this effect can be compared with the similar but less noticeable effect observed when measuring thermal stability by nanoDSF (Javed et al., 2022), where the mutant protein was found to be slightly more stable than the WT. This difference in stability may be related to the fact that the K380A mutation prevents the hydrolysis of the ATP ligand. One explanation is that the mutant does not undergo conformational changes associated with ligand processing. This could be reflected as an increased stability within the detergent micelles, whereas the WT must adapt its conformation to carry out ATP hydrolysis, with a concomitant decrease in stability.

An inverse phenomenon is observed with KcsA. The WT protein is slightly more stable than the chimeric protein. The mutations made in the K<sup>+</sup> channel pore entrance in order to mimic the eukaryotic Shaker channel could have an effect on its structure.

In summary, while a point mutation affecting ligand hydrolysis can be expected to lead to stabilization compared to hydrolysis-competent variants, ligand binding itself can have a stabilizing effect. On the other hand, extensive mutation of a segment of protein sequence to produce a variant with a new binding ability can be expected to destabilize the protein oligomer.

### **Protein-protein, and protein-ligand interaction**

By studying the interactions of proteins with other biomolecules, it is possible to obtain information that helps us understand how the quaternary structure, i.e., stoichiometry and spatial arrangement of subunits and ligands, underlie the function of MPs. Here we used NALIM to characterize protein-protein and protein-ligand complexes to get key structural information about BmrA and KcsA, respectively.

The interaction between WT BmrA and its ligands (ATP/Mg<sup>2+</sup>/V<sub>i</sub>) can be summarized in three steps: (i) the binding of ATP on BmrA with Mg<sup>2+</sup> as cofactor, (ii) the hydrolysis of ATP in ADP, (iii) the trapping of the ADP-BmrA complex by V<sub>i</sub>. Ligand binding was found to have a stabilizing effect on

the WT homodimer stability in native conditions, as indicated by the homodimer percentage (Figure 6).

These results are consistent with the effect observed by measuring the melting temperature ( $T_m$ ) of BmrA to assess protein stability (Javed et al., 2022). The  $T_m$  of WT BmrA increases after incubation with ATP/Mg<sup>2+</sup>/V<sub>i</sub>, which indicates better stability, while for the K380A mutant which does not hydrolyze ATP, the  $T_m$  remains more or less unchanged in the presence of ligand.

Additionally, in the presence of ligand, the  $m/z$  value of the homodimer shifted higher than in its absence. One obvious explanation is that the shift in the  $m/z$  value corresponds to mass added by the ligand, in this case the trapped ADP product. It has been established that the stoichiometry of the protein-ligand complex is two ATP molecules per homodimer (Chaptal et al., 2022). However, Mg<sup>2+</sup> and Na<sup>+</sup> ions brought in by the sample buffer can also generate adduct peaks which cannot be resolved by NALIM, so that the number of ATP molecules bound to the homodimer cannot be precisely determined in NALIM (see table 2 and 3).

The experimental control with the K380A mutant (Figure 7) also helps with interpretation of the dimer stabilization. As noted above, the K380A mutant cannot hydrolyze ATP, so that the observed shift in  $m/z$  values in the presence of ligand can be attributed to bound ATP and/or salts. The homodimer percentages are roughly similar between the K380A negative control (no ligand) and other tested conditions, confirming that there is no stabilization in the presence of ligand. This clearly shows that, regardless of adduct formation, the WT dimer stability improvement observed upon ligand binding is directly related to the protein-ligand interaction.

Interestingly we observe an increase in the WT homodimer percentage at  $t_i = 60$  (Figure 6). This effect remains to be explained.

One of the most informative experiments of this work sought to determine if the original, activating Tx7335 toxin binds to the pore vestibule the same way that charybdotoxin binds to eukaryotic channels (MacKinnon et al., 1998; Yu et al., 2005). In this experiment (Figure 12), analyzing the PPI also produces information about the stoichiometry of the complex, in addition to the binding site of Tx7335. The stoichiometry of the ShKcsA-Tx7335 complex was found to be exclusively 1:1, as there are no peaks indicating a different stoichiometry. The ChTx-ShKcsA complex likewise produces a 1:1 stoichiometry, as expected (Yu et al., 2005).

Using a chimeric ShKcsA K<sup>+</sup> channel, we were able to set up a competition experiment for pore binding. The competition experiment also show evidence that molecules or ions are bound to the channel in addition to the toxins. K<sup>+</sup> ions are brought in by the KCl-containing buffer, so some K<sup>+</sup> adduction is highly probable. Although we took precautions to desalt the ChTx sample twice before incubation, there may be enough Na<sup>+</sup> to produce adducts as well. Moreover, some ions may be

necessary for the interaction or may be strongly bound to the protein. Within the pore for example, three  $K^+$  ions could be trapped by ChTx. This would account for 117 Da of mass shift. A 546 to 549 Da mass shift for the tetramer was observed in all control conditions in this experiment. Since adducts are not resolved in these experiments, a conclusive attribution of the mass shifts to ligands and/or adducts is not possible. A better desalting of the samples before the experiment can decrease the probability of adduct formation, thus increasing the resolution and S/N ratio and simplifying the interpretation of the results. However, it may also affect complex stability if some ions and bound molecules are integral to the complex.

In the presence of both ChTx and Tx7335 toxins, peaks with a delta mass of 11 669 Da or 11 625 Da relative to adducted ShKcsA tetramer and monomer peaks, respectively, was found. Given the presence of adducts and the limited resolution in these experiments, it is reasonable to assume that these agree with the sum of the theoretical masses of both toxins (11 631). Thus, ChTx and Tx7335 can concomitantly bind to the chimeric ShKcsA  $K^+$  channel. As ChTx binds the ShKcsA in the pore vestibule, it follows that Tx7335 does not.

One intriguing observation is the presence of peaks corresponding to oligomeric states lower than the tetramer, bound to either one or both toxins. As it is unlikely that the monomer, dimer, or trimer are present in the sample solution, these peaks probably originate from the gas-phase dissociation of the complex formed between the tetramer and either or both toxins. While this phenomenon also decreases the signal for the tetramer bound to the toxins, it shows that the toxin-channel interactions are robust enough to survive gas-phase dissociation.

### **Direct analysis from membrane vesicles**

In the cell, the lipids sheltering a MP from the solvent also contribute to the stability of homo-oligomer complexes as well as complexes with other biomolecules (Marty, 2020). While detergent molecules are supposed to provide a surrogate for membrane lipids, the sheltering and stabilizing effects may be imperfect. This is one of the main reasons why ideally, the conditions for MS analysis should be as close as possible to the *in vivo* conditions. To explore the possibility of using NALIM in more "*in vivo*"-like conditions, we considered two other types of samples: 1) proteoliposomes, which mimic the membrane by incorporating the protein into a bilayer micelle made of canonical membrane lipids during the *in vitro* translation process, and 2) membrane vesicles (MV), which natively contain the membrane protein expressed by the cell in its cytoplasmic membrane, or in an organelle's membrane.



Analysis of MVs from BmrA-expressing bacteria shows that it is possible to eject the BmrA monomer from MVs (Figure 21). One interesting point is the observation of a satellite peak with a delta mass of +2 257 Da with respect to the monomer peak. Based on this interpretation, it is possible to identify a peak that could correspond to the BmrA dimer, with a mass shift of +3 155 Da. The extra mass could correspond to the binding of endogenous lipids to the BmrA dimer, which was observed for other ABC transporters. For instance, Bechara *et al.* (2015) report endogenous lipids bound to a heterodimeric ABC half-transporter found in *Thermus thermophilus* (TmrAB), which are not displaced by detergent or lost in the gas phase, some of which remain associated even after extensive delipidation.

Moreover, it was not possible to identify a peak corresponding to the BmrA dimer without lipids, as in the case of the monomer. Since there is evidence of gas phase dissociation of the dimer into monomers, the lack of lipid-free dimer would suggest that the lipid-monomer or lipid-dimer interaction is stronger than the inter-dimer interaction, and/or that lipids are necessary for dimer stabilization. If dimerization is conditioned by the presence of lipids, it would produce the same apparent saturation in lipid binding. The monomer peak with a delta mass of +2 257 Da could correspond to gas phase dissociation of the dimer-lipid complex with partial loss of lipid.

In contrast, no signal corresponding to any of the proteins which form the SpABC complex was detected (Figure 19). A possible explanation is that, contrary to BmrA, the SpABC complex is not ejected from the MV. SpABC could be anchored to the MV more strongly than BmrA. Based on the BmrA spiking experiment (Figure 21), contaminants in the sample do not appear to suppress ionization in NALIM as long as samples are prepared according to the same protocol. More likely, and despite the fact that efforts were made to have comparable protein concentration in both cases, it is possible that SpABC represents a lower proportion of the total membrane protein pool than BmrA. Practically speaking, this decreases the sensitivity of the analysis, so that the SpABC complex may be disrupted by the desorption process at laser values that are lower than the threshold of desorption for its monomer and oligomers.

The masses found for the BmrA dimer show evidence that other molecules are bound or adducted to the protein. In native ESI, experiments on ATP synthase showed that lipids can be strongly bound to MPs and resist the desolvation step (Barrera *et al.*, 2013; Chorev *et al.*, 2018). In cases where these lipids are essential for the MP structure and activity (Bao *et al.*, 2013; Dowhan & Bogdanov, 2011), desolvation steps actually need to be precisely adjusted to preserve some of these bound lipids. It should be noted that in a typical MALDI-TOF instrument configuration, there are limited possibilities to tune molecular desolvation. This gives the advantage to ESI-sourced instruments, although to fully leverage native ESI analysis and to achieve great results, specialized instruments with improved

transmission had to be developed (Rose et al., 2012; Sobott et al., 2002). In the future, ad-hoc MALDI-TOF instruments may be developed to facilitate desolvation, so that the extremely large mass range (tens of MDa) accessible to the TOF geometry can be leveraged as well for native MALDI-TOF MS methods.

As is generally the case with native MS methods, gas phase dissociation (GPD) was observed to varying degrees. Decreasing GPD could improve the observation of complexes in the high  $m/z$  range and facilitate data interpretation. Although instrumental parameters such as the IS1 and IS2 accelerating voltage, as well as the lens voltage were optimized on soluble proteins for the mass ranges explored in the present NALIM work (data to be published), these parameters may be further tailored for MPs in liposomes or in MVs to improve ejection from the lipid bilayer and/or stabilize the complex. Based on the successful stabilization of BmrA by ATP binding and orthovanadate in detergent, ligand-associated stabilization could similarly improve the detection in MVs not only of BmrA homodimer, but also of MPs more generally.

## CONCLUSION

The present work highlights the potential of NALIM as a promising method to characterize membrane protein complexes in the native state, quickly providing information that is complementary to structural biology tools and other biophysical methods. Here we give examples of instances where it can be used to gain information about MP complexes stoichiometry, binding site location, and stabilization by a ligand. Moreover, we show the stabilization of the BmrA dimer by a ligand binding effect. By applying NALIM in more *in vivo*-like conditions, we showed that it is possible to eject and observe BmrA dimer from MVs, revealing at the same time the likely presence of endogenous lipids and potential stabilizing effect on BmrA dimer stability.

The small amount of sample required and the fast data acquisition make the NALIM method *de facto* attractive for MP analysis. Moreover, data interpretation is straightforward when proteins are detected mainly as monocharged peaks. When a mixture of protein species is present in a sample, this characteristic presents an advantage over ESI, as it circumvents issues related to deconvolution, and potential biases due to the shifting of the charge state distribution upon ligand binding. Another potential advantage of NALIM over ESI is the use of a glycerol-based matrix which is biocompatible, opening up a whole new range of applications to native MS. Glycerol and other viscous solvents can emulate the crowded intracellular medium and be stepping stones towards mimicking the extracellular medium.

Additionally, NALIM share with ESI-based nMS the ability to directly analyze MPs and their complexes without the need for chemical modifications, which can be laborious, time consuming, and/or have an impact on the stability of the complex, as in the case of crosslinking-MALDI MS.

Overall, these results provide an encouraging basis to further develop NALIM for direct analysis of membrane protein complexes from membrane vesicles, and potentially even from whole cells. Given the role of membrane proteins in the resistance of bacterial and eukaryotic cells to respectively, antibiotics and chemotherapeutic agents, the range of applications for direct live cells analysis of membrane protein complexes associated with drugs or drug candidates can be expected to be quite large.

# **GENERAL CONCLUSION AND PERSPECTIVES**

The main objective of this work was to obtain structural and functional information about membrane proteins and their complexes in their native state. To achieve this goal, the optimization of instrumental and sample preparation parameters played a key role in the detection and preservation of the MPs' quaternary structure. Once experimental conditions were established, different types of membrane transporter proteins, including an ABC transporter and an ion channel were characterized. For instance, through NALIM, the stabilization of the dimer of the *Bacillus* multidrug resistance ATP (BmrA) protein by a ligand-binding effect was demonstrated. Furthermore, evidence that the binding site of the Tx7335 actitoxin in the potassium ion channel KcsA is distinct from the pore channel was found. Thanks to this, we can infer that this toxin activates the potassium channel through a possible allosteric effect.

In order to get closer to the conditions of the membrane protein in its native membrane, we applied NALIM to complexes in proteoliposomes and in membrane vesicles. Although proteoliposomes are a challenge in terms of sensitivity, ejection of BmrA dimer from native membrane vesicles giving signal in NALIM analysis was achieved. NALIM hence has the power to characterize membrane proteins in their native lipid bilayer. These results open the way for further analysis of membrane protein complexes in more *in vivo*-like conditions, such as directly from a live bacterial cell.

In summary, NALIM provides information about stoichiometry, stability, and ligand binding for different types of membrane protein complexes, either detergent solubilized or from native membranes.

Different proteins were explored during the development of this method in order to find a suitable calibrant, which allows us to obtain interpretable and reliable data in the high mass range (i.e.,  $m/z$  higher than 150 000). The use of a mouse immunoglobulin A (IgA), either found as a monomer or dimer, initially seemed promising. However, its conformation prevented complete deglycosylation of dimeric species. Through a collaboration, alpha1-antitrypsin ( $\alpha$ 1AT), a serine-protease inhibitor was made available in a recombinant non-glycosylated form that produces stable oligomers of high molecular weight. A first test of this protein in NALIM showed a distribution of oligomers with a mass higher than 250 000 Da, which can be used as molecular ladder for calibration. Through optimization of instrumental parameters, the detection of oligomers with a mass higher than 500 000 Da was achieved. Thus, instrumental optimizations improved the resolution and sensitivity of the method in a relatively high mass range. The validation of  $\alpha$ 1AT as a calibrant was achieved through the accurate mass determination of a deglycosylated IgG.

The results obtained on A1AT in turn opened the door for the study of large soluble molecular complexes (e.g., higher than 200,000 Da). We looked for candidate systems forming oligomers that are biologically relevant, such as a helicase complex and a transcription factor.

NALIM, once again, enabled rapid access to structural information such as the degree of oligomerization of ZBTB8A, which had not been previously obtained with other structural techniques such as analytical SEC or electron microscopy. Furthermore, a stabilizing effect of NusG on the Rho hexamer was discovered thanks in part to the observation of an increase in the intensity ratio between higher and lower oligomeric states in the presence of NusG.

On the other hand, limitations were also identified, such as a high percentage of dissociation in the gas phase or the current low resolution, which has so far prevented the identification of the binding stoichiometry of low-mass ligand molecules such as ATP in a complex. Thus, these limitations provide opportunities for future research aimed at improving the technique and its application to a wide range of high molecular mass systems. For instance, since the matrix significantly influences sample ionization in MALDI experiments (with potential effects on sensitivity and resolution), finer optimizations of the matrix composition could be performed to enhance the quality of obtained spectra. These optimizations have the potential to lead to a more accurate and reliable identification of MP complexes.

Additionally, the potential role of ligands as stabilizing agents of molecular complexes could be explored as a means to significantly decrease the gas phase dissociation during NALIM analysis. Hence, better preservation of the structures of these molecular complexes will facilitate their comprehensive characterization.

Finally, instrumental modifications represent another potential strategy to improve the characterization of large molecular complexes by NALIM. For instance, the integration of a specific high-mass detector can increase sensitivity in the high-mass range, thereby enabling the detailed structural analysis of complexes up to 2 MDa. This improvement can lead to valuable insights into these complexes, contributing to a deeper understanding of their structure and function.

In this work, no statistical analyses (e.g., repetitions, standard deviation) are presented, as experiments were not designed for this purpose. As mentioned earlier, this study establishes the basis for the implementation of NALIM and its feasibility to characterize and obtain structural insights of complex samples. However, the repetitive character of the signal and complexes stability values obtained from the proteins studied during the different experiments - under variable conditions - allowed us to be confident in the results obtained and the interpretations made. Thus, experiments specifically focused

on statistical studies can be conducted in the future as well as NALIM implementation in other systems.

In summary, this work demonstrates that NALIM provides valuable information regarding the structure and function of a variety of biological systems, such as stoichiometry, stability, and ligand-binding effect, in a fast and straightforward manner. Moreover, all these findings contribute to the understanding of membrane proteins and large soluble molecular complexes, which have significant implications in the fields of biology and pharmacology, with the potential to open new possibilities for drug design and therapies targeted at these proteins.

## REFERENCES

- Abrahams, J. P., Leslie, A. G. W., Lutter, R., & Walker, J. E. (1994). Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. *Nature*, *370*(6491), 621-628. <https://doi.org/10.1038/370621a0>
- Agúndez, J. A., Jiménez-Jimenez, F. J., Alonso-Navarro, H., & García-Martín, E. (2015). The potential of LINGO-1 as a therapeutic target for essential tremor. *Expert Opinion on Therapeutic Targets*, *19*(8), 1139-1148. <https://doi.org/10.1517/14728222.2015.1028360>
- Alam, A., Kowal, J., Broude, E., Roninson, I., & Locher, K. P. (2019). Structural insight into substrate and inhibitor discrimination by human P-glycoprotein. *Science*, *363*(6428), 753-756. <https://doi.org/10.1126/science.aav7102>
- Alam, S. R., Mahadevan, M. S., & Periasamy, A. (2023). Detecting RNA-Protein Interactions With EGFP-Cy3 FRET by Acceptor Photobleaching. *Current Protocols*, *3*(2), e689. <https://doi.org/10.1002/cpz1.689>
- Alberts, B. (2013). *Essential cell biology* (Fourth edition). Garland Science.
- Alexander, S., Mathie, A., & Peters, J. (2011). LIGAND-GATED ION CHANNELS: Ligand-gated ion channels. *British Journal of Pharmacology*, *164*, S115-S135. [https://doi.org/10.1111/j.1476-5381.2011.01649\\_4.x](https://doi.org/10.1111/j.1476-5381.2011.01649_4.x)
- Ali, M. H., & Imperiali, B. (2005). Protein oligomerization: How and why. *Bioorganic & Medicinal Chemistry*, *13*(17), 5013-5020. <https://doi.org/10.1016/j.bmc.2005.05.037>
- Ambrose, S., Housden, N. G., Gupta, K., Fan, J., White, P., Yen, H.-Y., Marcoux, J., Kleanthous, C., Hopper, J. T. S., & Robinson, C. V. (2017). Native Desorption Electrospray Ionization Liberates Soluble and Membrane Protein Complexes from Surfaces. *Angewandte Chemie International Edition*, *56*(46), 14463-14468. <https://doi.org/10.1002/anie.201704849>
- Amzel, L. M., & Poljak, R. J. (1979). *Three-Dimensional Structure of Immunoglobulins*. *48*, 961-997. <https://doi.org/10.1146/annurev.bi.48.070179.004525>
- Banerjee, S., Huber, T., & Sakmar, T. P. (2008). Rapid Incorporation of Functional Rhodopsin into Nanoscale Apolipoprotein Bound Bilayer (NABB) Particles. *Journal of Molecular Biology*, *377*(4), 1067-1081. <https://doi.org/10.1016/j.jmb.2008.01.066>
- Bao, H., Dalal, K., Wang, V., Rouiller, I., & Duong, F. (2013). The maltose ABC transporter: Action of membrane lipids on the transporter stability, coupling and ATPase activity. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1828*(8), 1723-1730. <https://doi.org/10.1016/j.bbamem.2013.03.024>
- Bardwell, V. J., & Treisman, R. (1994). The POZ domain: A conserved protein-protein interaction motif. *Genes & Development*, *8*(14), 1664-1677. <https://doi.org/10.1101/gad.8.14.1664>
- Barrera, N. P., Di Bartolo, N., Booth, P. J., & Robinson, C. V. (2008). Micelles Protect Membrane Complexes from Solution to Vacuum. *Science*, *321*(5886), 243-246. <https://doi.org/10.1126/science.1159292>
- Barrera, N. P., Isaacson, S. C., Zhou, M., Bavro, V. N., Welch, A., Schaedler, T. A., Seeger, M. A., Miguel, R. N., Korkhov, V. M., Van Veen, H. W., Venter, H., Walmsley, A. R., Tate, C. G., & Robinson, C. V. (2009). Mass spectrometry of membrane transporters reveals subunit stoichiometry and interactions. *Nature Methods*, *6*(8), 585-587. <https://doi.org/10.1038/nmeth.1347>
- Barrera, N. P., Zhou, M., & Robinson, C. V. (2013). The role of lipids in defining membrane protein interactions: Insights from mass spectrometry. *Trends in Cell Biology*, *23*(1), 1-8. <https://doi.org/10.1016/j.tcb.2012.08.007>
- Beaufour, M., Ginguené, D., Le Meur, R., Castaing, B., & Cadene, M. (2018). Liquid Native MALDI Mass Spectrometry for the Detection of Protein-Protein Complexes. *Journal of the American Society for Mass Spectrometry*, *29*(10), 1981-1994. <https://doi.org/10.1007/s13361-018-2015-x>



- Bechara, C., Nöll, A., Morgner, N., Degiacomi, M. T., Tampé, R., & Robinson, C. V. (2015). A subset of annular lipids is linked to the flippase activity of an ABC transporter. *Nature Chemistry*, 7(3), 255-262. <https://doi.org/10.1038/nchem.2172>
- Behnke, J.-S., & Urner, L. H. (2023). Emergence of mass spectrometry detergents for membrane proteomics. *Analytical and Bioanalytical Chemistry*. <https://doi.org/10.1007/s00216-023-04584-z>
- Berg, J. M., Tymoczko, J. L., & Stryer, L. (2012). *Biochemistry* (7th ed). W.H. Freeman.
- Berthaud, A., Manzi, J., Pérez, J., & Mangenot, S. (2012). Modeling Detergent Organization around Aquaporin-0 Using Small-Angle X-ray Scattering. *Journal of the American Chemical Society*, 134(24), 10080-10088. <https://doi.org/10.1021/ja301667n>
- Bianchet, M., Ysern, X., Hullihen, J., Pedersen, P. L., & Amzel, L. M. (1991). Mitochondrial ATP synthase. Quaternary structure of the F1 moiety at 3.6 Å determined by x-ray diffraction analysis. *Journal of Biological Chemistry*, 266(31), 21197-21201. [https://doi.org/10.1016/S0021-9258\(18\)54840-7](https://doi.org/10.1016/S0021-9258(18)54840-7)
- Boeri Erba, E., & Petosa, C. (2015). The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes: The Emerging Role of Native Mass Spectrometry. *Protein Science*, 24(8), 1176-1192. <https://doi.org/10.1002/pro.2661>
- Bolbach, G. (2005). Matrix-Assisted Laser Desorption/Ionization Analysis of Non-Covalent Complexes: Fundamentals and Applications. *Current Pharmaceutical Design*, 11(20), 2535-2557. <https://doi.org/10.2174/1381612054546923>
- Bolla, J. R., Agasid, M. T., Mehmood, S., & Robinson, C. V. (2019). Membrane Protein–Lipid Interactions Probed Using Mass Spectrometry. *Annual Review of Biochemistry*, 88(1), 85-111. <https://doi.org/10.1146/annurev-biochem-013118-111508>
- Bolla, J. R., Sauer, J. B., Wu, D., Mehmood, S., Allison, T. M., & Robinson, C. V. (2018). Direct observation of the influence of cardiolipin and antibiotics on lipid II binding to MurJ. *Nature Chemistry*, 10(3), 363-371. <https://doi.org/10.1038/nchem.2919>
- Bollen, A., Herzog, A., Cravador, A., Héron, P., Chuchana, P., Vander Straten, A., Loriau, R., Jacobs, P., & Van Elsen, A. (1983). Cloning and Expression in *Escherichia coli* of Full-Length Complementary DNA Coding for Human  $\alpha_1$ -Antitrypsin. *DNA*, 2(4), 255-264. <https://doi.org/10.1089/dna.1983.2.255>
- Bonchuk, A., Balagurov, K., & Georgiev, P. (2023). BTB domains: A structural view of evolution, multimerization, and protein–protein interactions. *BioEssays*, 45(2), 2200179. <https://doi.org/10.1002/bies.202200179>
- Borchers, C., & Tomer, K. B. (1999). Characterization of the Noncovalent Complex of Human Immunodeficiency Virus Glycoprotein 120 with Its Cellular Receptor CD4 by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Biochemistry*, 38(36), 11734-11740. <https://doi.org/10.1021/bi990935w>
- Botamanenko, D. Y., & Jarrold, M. F. (2019). Ion-Ion Interactions in Charge Detection Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, 30(12), 2741-2749. <https://doi.org/10.1007/s13361-019-02343-y>
- Bragg, W. H., & Bragg, W. L. (1913). The structure of the diamond. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 89(610), 277-291. <https://doi.org/10.1098/rspa.1913.0084>
- Bragg, W. L. (1913). The structure of some crystals as indicated by their diffraction of X-rays. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 89(610), 248-277. <https://doi.org/10.1098/rspa.1913.0083>

- Burmann, B. M., Schweimer, K., Luo, X., Wahl, M. C., Stitt, B. L., Gottesman, M. E., & Rösch, P. (2010). A NusE:NusG Complex Links Transcription and Translation. *Science*, 328(5977), 501-504. <https://doi.org/10.1126/science.1184953>
- Burova, E., Hung, S. C., Sagitov, V., Stitt, B. L., & Gottesman, M. E. (1995). Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro. *Journal of Bacteriology*, 177(5), 1388-1392. <https://doi.org/10.1128/jb.177.5.1388-1392.1995>
- Busenlehner, L., Codreanu, S., Holm, P., Bhakat, P., Hebert, H., Morgenstern, R., & Armstrong, R. (2004). Stress Sensor Triggers Conformational Response of the Integral Membrane Protein Microsomal Glutathione Transferase 1. *Biochemistry*, 43(35), 11145-11152. <https://doi.org/10.1021/bi048716k>
- Büyükköroğlu, G., Dora, D. D., Özdemir, F., & Hızal, C. (2018). Techniques for Protein Analysis. En *Omics Technologies and Bio-Engineering* (pp. 317-351). Elsevier. <https://doi.org/10.1016/B978-0-12-804659-3.00015-4>
- Cadene, M., & Chait, B. T. (2000). A Robust, Detergent-Friendly Method for Mass Spectrometric Analysis of Integral Membrane Proteins. *Analytical Chemistry*, 72(22), 5655-5658. <https://doi.org/10.1021/ac0008111>
- Cala, O., Guillière, F., & Krimm, I. (2014). NMR-based analysis of protein–ligand interactions. *Analytical and Bioanalytical Chemistry*, 406(4), 943-956. <https://doi.org/10.1007/s00216-013-6931-0>
- Chakraburty, K. (2001). Translational regulation by ABC systems. *Research in Microbiology*, 152(3-4), 391-399. [https://doi.org/10.1016/S0923-2508\(01\)01210-4](https://doi.org/10.1016/S0923-2508(01)01210-4)
- Chaptal, V., Zampieri, V., Wiseman, B., Orelle, C., Martin, J., Nguyen, K.-A., Gobet, A., Di Cesare, M., Magnard, S., Javed, W., Eid, J., Kilburg, A., Peuchmaur, M., Marcoux, J., Monticelli, L., Hogbom, M., Schoehn, G., Jault, J.-M., Boumendjel, A., & Falson, P. (2022). Substrate-bound and substrate-free outward-facing structures of a multidrug ABC exporter. *Science Advances*, 8(4), eabg9215. <https://doi.org/10.1126/sciadv.abg9215>
- Chintalacharuvu, S. R., & Emancipator, S. N. (1997). The glycosylation of IgA produced by murine B cells is altered by Th2 cytokines. *The Journal of Immunology*, 159(5), 2327-2333. <https://doi.org/10.4049/jimmunol.159.5.2327>
- Chorev, D. S., Baker, L. A., Wu, D., Beilsten-Edmands, V., Rouse, S. L., Zeev-Ben-Mordehai, T., Jiko, C., Samsudin, F., Gerle, C., Khalid, S., Stewart, A. G., Matthews, S. J., Grünewald, K., & Robinson, C. V. (2018). Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. *Science*, 362(6416), 829-834. <https://doi.org/10.1126/science.aau0976>
- Chorev, D. S., Tang, H., Rouse, S. L., Bolla, J. R., von Kügelgen, A., Baker, L. A., Wu, D., Gault, J., Grünewald, K., Bharat, T. A. M., Matthews, S. J., & Robinson, C. V. (2020). The use of sonicated lipid vesicles for mass spectrometry of membrane protein complexes. *Nature Protocols*, 15(5), 1690-1706. <https://doi.org/10.1038/s41596-020-0303-y>
- Chowdhury, S. K., Katta, V., & Chait, B. T. (1990). *ELECTROSPRAY IONIZATION MASS SPECTROMETRIC PEPTIDE MAPPING: A RAPID, SENSITIVE TECHNIQUE FOR PROTEIN STRUCTURE ANALYSIS*. 167(2).
- Clegg, R. M. (1995). *Fluorescence resonance energy transfer*. 6, 103-110.
- Clement, K., Reynaud, A., Defoort, M., Vysotskyi, B., Fortin, T., Lai, S.-H., Çumaku, V., Dominguez-Medina, S., Hentz, S., & Masselon, C. (2021). Requirements and attributes of nano-resonator mass spectrometry for the analysis of intact viral particles. *Analytical and Bioanalytical Chemistry*, 413(29), 7147-7156. <https://doi.org/10.1007/s00216-021-03511-4>
- Cole, D., Young, G., Weigel, A., Sebesta, A., & Kukura, P. (2017). Label-Free Single-Molecule Imaging with Numerical-Aperture-Shaped Interferometric Scattering Microscopy. *ACS Photonics*, 4(2), 211-216. <https://doi.org/10.1021/acsp Photonics.6b00912>

- Corte, E. D., & Parkhouse, R. M. E. (1973). Biosynthesis of immunoglobulin A (IgA). Secretion and addition of carbohydrate to monomer and polymer forms of a mouse myeloma protein. *Biochemical Journal*, *136*(3), 589-596. <https://doi.org/10.1042/bj1360589>
- Da Vela, S., & Svergun, D. I. (2020). Methods, development and applications of small-angle X-ray scattering to characterize biological macromolecules in solution. *Current Research in Structural Biology*, *2*, 164-170. <https://doi.org/10.1016/j.crstbi.2020.08.004>
- Dafun, A. S., & Marcoux, J. (2022). Structural mass spectrometry of membrane proteins. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1870*(8), 140813. <https://doi.org/10.1016/j.bbapap.2022.140813>
- Daubenfeld, T., Bouin, A.-P., & Van Der Rest, G. (2006). A deconvolution method for the separation of specific versus nonspecific interactions in noncovalent protein-ligand complexes analyzed by ESI-FT-ICR mass spectrometry. *Journal of the American Society for Mass Spectrometry*, *17*(9), 1239-1248. <https://doi.org/10.1016/j.jasms.2006.05.005>
- Dauter, Z., & Wlodawer, A. (2016). Progress in protein crystallography. *Protein & Peptide Letters*, *23*(3), 201-210. <https://doi.org/10.2174/0929866523666160106153524>
- Davidson, A. L., Dassa, E., Orelle, C., & Chen, J. (2008). Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiology and Molecular Biology Reviews*, *72*(2), 317-364. <https://doi.org/10.1128/MMBR.00031-07>
- De Hoffmann, E., & Stroobant, V. (2011). *Mass spectrometry: Principles and applications* (3. ed., reprinted). Wiley.
- De Oliveira, T. M., Van Beek, L., Shilliday, F., Debreczeni, J. É., & Phillips, C. (2021). Cryo-EM: The Resolution Revolution and Drug Discovery. *SLAS Discovery*, *26*(1), 17-31. <https://doi.org/10.1177/2472555220960401>
- De Vera, I. M. S., Zheng, J., Novick, S., Shang, J., Hughes, T. S., Brust, R., Munoz-Tello, P., Gardner, W. J., Marciano, D. P., Kong, X., Griffin, P. R., & Kojetin, D. J. (2017). Synergistic Regulation of Coregulator/Nuclear Receptor Interaction by Ligand and DNA. *Structure*, *25*(10), 1506-1518.e4. <https://doi.org/10.1016/j.str.2017.07.019>
- Denisov, I. G., Grinkova, Y. V., Lazarides, A. A., & Sligar, S. G. (2004). Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *Journal of the American Chemical Society*, *126*(11), 3477-3487. <https://doi.org/10.1021/ja0393574>
- Depping, P., Román Lara, M. M., Kesidis, A., Bill, R. M., Rothnie, A. J., Browning, D. F., & Goddard, A. D. (2022). Heterologous Expression of Membrane Proteins in *E. coli*. En I. Mus-Veteau (Ed.), *Heterologous Expression of Membrane Proteins* (Vol. 2507, pp. 59-78). Springer US. [https://doi.org/10.1007/978-1-0716-2368-8\\_4](https://doi.org/10.1007/978-1-0716-2368-8_4)
- Di Cesare, M., Diagne, A. M., Bourgey, B., Jault, J.-M., & Orelle, C. (2022). Functional Overexpression of Membrane Proteins in *E. coli*: The Good, the Bad, and the Ugly. En I. Mus-Veteau (Ed.), *Heterologous Expression of Membrane Proteins* (Vol. 2507, pp. 41-58). Springer US. [https://doi.org/10.1007/978-1-0716-2368-8\\_3](https://doi.org/10.1007/978-1-0716-2368-8_3)
- Diagne, A. M., Pelletier, A., Durmort, C., Faure, A., Kanonenberg, K., Freton, C., Page, A., Delolme, F., Vorac, J., Vallet, S., Bellard, L., Vivès, C., Fieschi, F., Vernet, T., Rousselle, P., Guiral, S., Grangeasse, C., Jault, J.-M., & Orelle, C. (2022). Identification of a two-component regulatory system involved in antimicrobial peptide resistance in *Streptococcus pneumoniae*. *PLOS Pathogens*, *18*(4), e1010458. <https://doi.org/10.1371/journal.ppat.1010458>
- Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D., & Alice, M. B. (1968). Molecular Beams of Macroions. *The Journal of Chemical Physics*, *49*(5), 2240-2249. <https://doi.org/10.1063/1.1670391>
- Dörr, J. M., Koorengevel, M. C., Schäfer, M., Prokofyev, A. V., Scheidelaar, S., Van Der Crujnsen, E. A. W., Dafforn, T. R., Baldus, M., & Killian, J. A. (2014). Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K<sup>+</sup> channel: The power of native nanodiscs. *Proceedings of the National Academy of Sciences*, *111*(52), 18607-18612. <https://doi.org/10.1073/pnas.1416205112>

- Dowhan, W., & Bogdanov, M. (2011). Lipid–protein interactions as determinants of membrane protein structure and function. *Biochemical Society Transactions*, 39(3), 767-774. <https://doi.org/10.1042/BST0390767>
- Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., & MacKinnon, R. (1998). The Structure of the Potassium Channel: Molecular Basis of  $K^+$  Conduction and Selectivity. *Science*, 280(5360), 69-77. <https://doi.org/10.1126/science.280.5360.69>
- Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T., & MacKinnon, R. (2002). *X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity*. 415.
- Edwards, W. F., Jin, L., & Thies, M. C. (2003). *MALDI-TOF mass spectrometry: Obtaining reliable mass spectra for insoluble carbonaceous pitches*.
- EDYP LAB. (2019). *NEMS*. NEMS. <https://www.edyp.fr/web/2019/11/14/nems/>
- Electron Microscopy Data Bank [EMDB]. (2023). <https://www.ebi.ac.uk/emdb/>
- Elmlund, D., & Elmlund, H. (2015). Cryogenic Electron Microscopy and Single-Particle Analysis. *Annual Review of Biochemistry*, 84(1), 499-517. <https://doi.org/10.1146/annurev-biochem-060614-034226>
- Emwas, A.-H. M. (2015). The Strengths and Weaknesses of NMR Spectroscopy and Mass Spectrometry with Particular Focus on Metabolomics Research. En J. T. Bjerrum (Ed.), *Metabonomics* (Vol. 1277, pp. 161-193). Springer New York. [https://doi.org/10.1007/978-1-4939-2377-9\\_13](https://doi.org/10.1007/978-1-4939-2377-9_13)
- Espino, J. A., Mali, V. S., & Jones, L. M. (2015). In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. *Analytical Chemistry*, 87(15), 7971-7978. <https://doi.org/10.1021/acs.analchem.5b01888>
- Evans, J. T., & Rohrmann, G. F. (1997). The baculovirus single-stranded DNA binding protein, LEF-3, forms a homotrimer in solution. *Journal of Virology*, 71(5), 3574-3579. <https://doi.org/10.1128/jvi.71.5.3574-3579.1997>
- Fan, X., Wang, J., Zhang, X., Yang, Z., Zhang, J.-C., Zhao, L., Peng, H.-L., Lei, J., & Wang, H.-W. (2019). Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution. *Nature Communications*, 10(1), 2386. <https://doi.org/10.1038/s41467-019-10368-w>
- Fang, C., Huang, Y., & Zhao, Y. (2023). *Review of FRET biosensing and its application in biomolecular detection*. 15(2), 694-709.
- Fantin, S. M., Parson, K. F., Yadav, P., Juliano, B., Li, G. C., Sanders, C. R., Ohi, M. D., & Ruotolo, B. T. (2021). Ion mobility–mass spectrometry reveals the role of peripheral myelin protein dimers in peripheral neuropathy. *Proceedings of the National Academy of Sciences*, 118(17), e2015331118. <https://doi.org/10.1073/pnas.2015331118>
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. (1989). Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science*, 246(4926), 64-71. <https://doi.org/10.1126/science.2675315>
- Fenyo, D., Wang, Q., DeGrasse, J. A., Padovan, J. C., Cadene, M., & Chait, B. T. (2007). MALDI Sample Preparation: The Ultra Thin Layer Method. *Journal of Visualized Experiments*, 3, 192. <https://doi.org/10.3791/192>
- Fitzgerald, M. C., Chernushevich, I., Standing, K. G., Whitman, C. P., & Kent, S. B. (1996). Probing the oligomeric structure of an enzyme by electrospray ionization time-of-flight mass spectrometry. *Proceedings of the National Academy of Sciences*, 93(14), 6851-6856. <https://doi.org/10.1073/pnas.93.14.6851>
- Foulds, J., & Barrett, C. (1973). Characterization of *Escherichia coli* Mutants Tolerant to Bacteriocin JF246: Two New Classes of Tolerant Mutants. *Journal of Bacteriology*, 116(2), 885-892. <https://doi.org/10.1128/jb.116.2.885-892.1973>
- Frank, M., Mears, C. A., Labov, S. E., Benner, W. H., Horn, D., Jaklevic, J. M., & Barfknecht, A. T. (1996). High-efficiency Detection of 66 000 Da Protein Molecules Using a Cryogenic Detector in a Matrix-assisted Laser

- Desorption/Ionization Time-of-flight Mass Spectrometer. *Rapid Communications in Mass Spectrometry*, 10(15), 1946-1950. [https://doi.org/10.1002/\(SICI\)1097-0231\(199612\)10:15<1946::AID-RCM774>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-0231(199612)10:15<1946::AID-RCM774>3.0.CO;2-3)
- Friedrich, W., Knipping, P., & Laue, M. (1913). Interferenzerscheinungen bei Röntgenstrahlen. *Annalen der Physik*, 346(10), 971-988. <https://doi.org/10.1002/andp.19133461004>
- Fuller, F. D., Loukianov, A., Takanashi, T., You, D., Li, Y., Ueda, K., Fransson, T., Yabashi, M., Katayama, T., Weng, T.-C., Alonso-Mori, R., Bergmann, U., Jan Kern, Yachandra, V. K., Wernet, P., & Yano, J. (2021). Resonant X-ray emission spectroscopy from broadband stochastic pulses at an X-ray free electron laser. *Communications Chemistry*, 4(1), 84. <https://doi.org/10.1038/s42004-021-00512-3>
- Gabant, G., & Cadene, M. (2008). Mass spectrometry of full-length integral membrane proteins to define functionally relevant structural features. *Methods*, 46(2), 54-61. <https://doi.org/10.1016/j.ymeth.2008.10.021>
- Gaddy, J. A., Tomaras, A. P., & Actis, L. A. (2009). The *Acinetobacter baumannii* 19606 OmpA Protein Plays a Role in Biofilm Formation on Abiotic Surfaces and in the Interaction of This Pathogen with Eukaryotic Cells. *Infection and Immunity*, 77(8), 3150-3160. <https://doi.org/10.1128/IAI.00096-09>
- Gan, J., Ben-Nissan, G., Arkind, G., Tarnavsky, M., Trudeau, D., Noda Garcia, L., Tawfik, D. S., & Sharon, M. (2017). Native Mass Spectrometry of Recombinant Proteins from Crude Cell Lysates. *Analytical Chemistry*, 89(8), 4398-4404. <https://doi.org/10.1021/acs.analchem.7b00398>
- Gertz, J., Reddy, T. E., Varley, K. E., Garabedian, M. J., & Myers, R. M. (2012). Genistein and bisphenol A exposure cause estrogen receptor 1 to bind thousands of sites in a cell type-specific manner. *Genome Research*, 22(11), 2153-2162. <https://doi.org/10.1101/gr.135681.111>
- Gogol, E. P., Lucken, U., Bork, T., & Capaldi, R. A. (1989). *Molecular Architecture of Escherichia coli F1Adenosinetriphosphatase*. 28(11), 4709-4716.
- Goldstein, S. (1994). The charybdotoxin receptor of a Shaker K<sup>+</sup> channel: Peptide and channel residues mediating molecular recognition. *Neuron*, 12(6), 1377-1388. [https://doi.org/10.1016/0896-6273\(94\)90452-9](https://doi.org/10.1016/0896-6273(94)90452-9)
- Goodsell, D. S., & Olson, A. J. (2000). Structural Symmetry and Protein Function. *Annual Review of Biophysics and Biomolecular Structure*, 29(1), 105-153. <https://doi.org/10.1146/annurev.biophys.29.1.105>
- Goosen, N., & Moolenaar, G. F. (2001). Role of ATP hydrolysis by UvrA and UvrB during nucleotide excision repair. *Research in Microbiology*, 152(3-4), 401-409. [https://doi.org/10.1016/S0923-2508\(01\)01211-6](https://doi.org/10.1016/S0923-2508(01)01211-6)
- Goren, M. A., & Fox, B. G. (2008). Wheat germ cell-free translation, purification, and assembly of a functional human stearoyl-CoA desaturase complex. *Protein Expression and Purification*, 62(2), 171-178. <https://doi.org/10.1016/j.pep.2008.08.002>
- Goren, M. A., Nozawa, A., Makino, S., Wrobel, R. L., & Fox, B. G. (2009). Chapter 37 Cell-Free Translation of Integral Membrane Proteins into Unilamellar Liposomes. En *Methods in Enzymology* (Vol. 463, pp. 647-673). Elsevier. [https://doi.org/10.1016/S0076-6879\(09\)63037-8](https://doi.org/10.1016/S0076-6879(09)63037-8)
- Gossert, A. D., & Jahnke, W. (2016). NMR in drug discovery: A practical guide to identification and validation of ligands interacting with biological macromolecules. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 97, 82-125. <https://doi.org/10.1016/j.pnmrs.2016.09.001>
- Grand View Research. (2023). *Monoclonal Antibodies Market Size, Share & Trends Analysis Report By Source Type (Chimeric, Murine, Humanized, Human), By Production Type (In Vivo, In Vitro), By Application, By End-use, By Region, And Segment Forecasts, 2023—2030*. <https://www.grandviewresearch.com/industry-analysis/monoclonal-antibodies-market#>
- Greenfield, N. J. (2006). Analysis of the kinetics of folding of proteins and peptides using circular dichroism. *Nature Protocols*, 1(6), 2891-2899. <https://doi.org/10.1038/nprot.2006.244>

- Greig, M. J., Gaus, H., Cummins, L. L., Sasmor, H., & Griffey, R. H. (1995). Measurement of Macromolecular Binding Using Electrospray Mass Spectrometry. Determination of Dissociation Constants for Oligonucleotide: Serum Albumin Complexes. *Journal of the American Chemical Society*, *117*(43), 10765-10766. <https://doi.org/10.1021/ja00148a028>
- Griffiths, R. L., Dexter, A., Creese, A. J., & Cooper, H. J. (2015). Liquid extraction surface analysis field asymmetric waveform ion mobility spectrometry mass spectrometry for the analysis of dried blood spots. *The Analyst*, *140*(20), 6879-6885. <https://doi.org/10.1039/C5AN00933B>
- Guillemain, A., Laouarem, Y., Cobret, L., Štefok, D., Chen, W., Bloch, S., Zahaf, A., Blot, L., Reverchon, F., Normand, T., Decoville, M., Grillon, C., Traiffort, E., & Morisset-Lopez, S. (2020). LINGO family receptors are differentially expressed in the mouse brain and form native multimeric complexes. *The FASEB Journal*, *34*(10), 13641-13653. <https://doi.org/10.1096/fj.202000826R>
- Gupta, K., Donlan, J. A. C., Hopper, J. T. S., Uzdavinyis, P., Landreh, M., Struwe, W. B., Drew, D., Baldwin, A. J., Stansfeld, P. J., & Robinson, C. V. (2017). The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature*, *541*(7637), 421-424. <https://doi.org/10.1038/nature20820>
- Hale, O. J., & Cooper, H. J. (2021). Native Mass Spectrometry Imaging of Proteins and Protein Complexes by Nano-DESI. *Analytical Chemistry*, *93*(10), 4619-4627. <https://doi.org/10.1021/acs.analchem.0c05277>
- Hale, O. J., Sisley, E. K., Griffiths, R. L., Styles, I. B., & Cooper, H. J. (2020). Native LESA TWIMS-MSI: Spatial, Conformational, and Mass Analysis of Proteins and Protein Complexes. *Journal of the American Society for Mass Spectrometry*, *31*(4), 873-879. <https://doi.org/10.1021/jasms.9b00122>
- Hambly, D. M., & Gross, M. L. (2005). Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *Journal of the American Society for Mass Spectrometry*, *16*(12), 2057-2063. <https://doi.org/10.1016/j.jasms.2005.09.008>
- Hamdi, K., Salladini, E., O'Brien, D. P., Brier, S., Chenal, A., Yacoubi, I., & Longhi, S. (2017). Structural disorder and induced folding within two cereal, ABA stress and ripening (ASR) proteins. *Scientific Reports*, *7*(1), 15544. <https://doi.org/10.1038/s41598-017-15299-4>
- Haq, I., Irving, J. A., Faull, S. V., Dickens, J. A., Ordóñez, A., Belorgey, D., Gooptu, B., & Lomas, D. A. (2013). Reactive centre loop mutants of  $\alpha$ -1-antitrypsin reveal position-specific effects on intermediate formation along the polymerization pathway. *Bioscience Reports*, *33*(3), e00046. <https://doi.org/10.1042/BSR20130038>
- Hashimoto, K., & Panchenko, A. R. (2010). Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states. *Proceedings of the National Academy of Sciences*, *107*(47), 20352-20357. <https://doi.org/10.1073/pnas.1012999107>
- Hellwig, N., Martin, J., & Morgner, N. (2022). LILBID-MS: Using lasers to shed light on biomolecular architectures. *Biochemical Society Transactions*, *50*(3), 1057-1067. <https://doi.org/10.1042/BST20190881>
- Hellwig, N., Peetz, O., Ahdash, Z., Tascón, I., Booth, P. J., Mikusevic, V., Diskowski, M., Politis, A., Hellmich, Y., Hänel, I., Reading, E., & Morgner, N. (2018). Native mass spectrometry goes more native: Investigation of membrane protein complexes directly from SMALPs. *Chemical Communications*, *54*(97), 13702-13705. <https://doi.org/10.1039/C8CC06284F>
- Henderson, R., Baldwin, J. M., Ceskat, T. A., Zemlin, F., Beckmann, E., & Downing, K. H. (1990). *Model for the Structure of Bacteriorhodopsin Based on High-resolution Electron Cryo-microscopy*. *213*(4), 899-929. [https://doi.org/doi:10.1016/s0022-2836\(05\)80271-2](https://doi.org/doi:10.1016/s0022-2836(05)80271-2)

- Hendriks, I. A., D'Souza, R. C. J., Yang, B., Verlaan-de Vries, M., Mann, M., & Vertegaal, A. C. O. (2014). Uncovering global SUMOylation signaling networks in a site-specific manner. *Nature Structural & Molecular Biology*, *21*(10), 927-936. <https://doi.org/10.1038/nsmb.2890>
- Henrich, E., Löhr, F., Pawlik, G., Peetz, O., Dötsch, V., Morgner, N., De Kroon, A. I., & Bernhard, F. (2018). Lipid Conversion by Cell-Free Synthesized Phospholipid Methyltransferase Opi3 in Defined Nanodisc Membranes Supports an *in Trans* Mechanism. *Biochemistry*, *57*(40), 5780-5784. <https://doi.org/10.1021/acs.biochem.8b00807>
- Henrich, E., Sörmann, J., Eberhardt, P., Peetz, O., Mezhyrova, J., Morgner, N., Fendler, K., Dötsch, V., Wachtveitl, J., Bernhard, F., & Bamann, C. (2017). From Gene to Function: Cell-Free Electrophysiological and Optical Analysis of Ion Pumps in Nanodiscs. *Biophysical Journal*, *113*(6), 1331-1341. <https://doi.org/10.1016/j.bpj.2017.03.026>
- Hernández, H., & Robinson, C. V. (2007). Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature Protocols*, *2*(3), 715-726. <https://doi.org/10.1038/nprot.2007.73>
- Hillenkamp, F., & Karas, M. (2007). The MALDI Process and Method. En F. Hillenkamp & J. Peter-Katalini (Eds.), *MALDI MS* (pp. 1-28). Wiley-VCH Verlag GmbH & Co. KGaA. <https://doi.org/10.1002/9783527610464.ch1>
- Hofstadler, S. A., & Sannes-Lowery, K. A. (2006). Applications of ESI-MS in drug discovery: Interrogation of noncovalent complexes. *Nature Reviews Drug Discovery*, *5*(7), 585-595. <https://doi.org/10.1038/nrd2083>
- Hu, Y., Cheng, K., He, L., Zhang, X., Jiang, B., Jiang, L., Li, C., Wang, G., Yang, Y., & Liu, M. (2021). NMR-Based Methods for Protein Analysis. *Analytical Chemistry*, *93*(4), 1866-1879. <https://doi.org/10.1021/acs.analchem.0c03830>
- Huang, N., Deng, H., Liu, B., Wang, D., & Zhao, Z. (2021). Features and futures of X-ray free-electron lasers. *The Innovation*, *2*(2), 100097. <https://doi.org/10.1016/j.xinn.2021.100097>
- Huang, Y., & Orlando, R. (2017). Kinetics of N-Glycan Release from Human Immunoglobulin G (IgG) by PNGase F: All Glycans Are Not Created Equal. *Journal of Biomolecular Techniques: JBT*, *28*(4), 150-157. <https://doi.org/10.7171/jbt.17-2804-002>
- Irving, J. A., Ekeowa, U. I., Belorgey, D., Haq, I., Gooptu, B., Miranda, E., Pérez, J., Roussel, B. D., Ordóñez, A., Dalton, L. E., Thomas, S. E., Marciniak, S. J., Parfrey, H., Chilvers, E. R., Teckman, J. H., Alam, S., Mahadeva, R., Rashid, S. T., Vallier, L., & Lomas, D. A. (2011). The Serpinopathies. En *Methods in Enzymology* (Vol. 501, pp. 421-466). Elsevier. <https://doi.org/10.1016/B978-0-12-385950-1.00018-3>
- Jamshad, M., Grimard, V., Idini, I., Knowles, T. J., Dowle, M. R., Schofield, N., Sridhar, P., Lin, Y., Finka, R., Wheatley, M., Thomas, O. R. T., Palmer, R. E., Overduin, M., Govaerts, C., Ruyschaert, J.-M., Edler, K. J., & Dafforn, T. R. (2015). Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins. *Nano Research*, *8*(3), 774-789. <https://doi.org/10.1007/s12274-014-0560-6>
- Jan, Y. N., Jan, L. Y., & Dennis, M. J. (1977). *Two mutations of synaptic transmission in Drosophila*. *198*, 87-108.
- Jaquillard, L., Saab, F., Schoentgen, F., & Cadene, M. (2012). Improved Accuracy of Low Affinity Protein-Ligand Equilibrium Dissociation Constants Directly Determined by Electrospray Ionization Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, *23*(5), 908-922. <https://doi.org/10.1007/s13361-011-0305-7>
- Javed, W., Vallet, S., Clement, M.-P., Le Roy, A., Moulin, M., Härtle, M., Breyton, C., Burlet-Schiltz, O., Marcoux, J., Orelle, C., Ebel, C., Martel, A., & Jault, J.-M. (2022). Structural Insights into the Catalytic Cycle of a Bacterial Multidrug ABC Efflux Pump. *Journal of Molecular Biology*, *434*(9), 167541. <https://doi.org/10.1016/j.jmb.2022.167541>

- Jiang, L., Xiong, Z., Song, Y., Lu, Y., Chen, Y., Schultz, J. S., Li, J., & Liao, J. (2019). Protein–Protein Affinity Determination by Quantitative FRET Quenching. *Scientific Reports*, 9(1), 2050. <https://doi.org/10.1038/s41598-018-35535-9>
- Jiang, Y., Lee, A., Chen, J., Cadene, M., Chait, B. T., & MacKinnon, R. (2002). Crystal structure and mechanism of a calcium-gated potassium channel. *Nature*, 417(6888), 515-522. <https://doi.org/10.1038/417515a>
- Johnson, D. T., Di Stefano, L. H., & Jones, L. M. (2019). Fast photochemical oxidation of proteins (FPOP): A powerful mass spectrometry–based structural proteomics tool. *Journal of Biological Chemistry*, 294(32), 11969-11979. <https://doi.org/10.1074/jbc.REV119.006218>
- Kamilar, E., Bariwal, J., Zheng, W., Ma, H., & Liang, H. (2023). SMALPs Are Not Simply Nanodiscs: The Polymer-to-Lipid Ratios of Fractionated SMALPs Underline Their Heterogeneous Nature. *Biomacromolecules*, 24(4), 1819-1838. <https://doi.org/10.1021/acs.biomac.3c00034>
- Karas, M., Bahr, U., & Hillenkamp, F. (1989). UV laser matrix desorption/ionization mass spectrometry of proteins in the 100 000 dalton range. *International Journal of Mass Spectrometry and Ion Processes*, 92, 231-242. [https://doi.org/10.1016/0168-1176\(89\)83030-7](https://doi.org/10.1016/0168-1176(89)83030-7)
- Karas, Michael., Bachmann, Doris., & Hillenkamp, Franz. (1985). Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Analytical Chemistry*, 57(14), 2935-2939. <https://doi.org/10.1021/ac00291a042>
- Karas, Michael., & Hillenkamp, Franz. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry*, 60(20), 2299-2301. <https://doi.org/10.1021/ac00171a028>
- Katta, V., & Chait, B. T. (1991). Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry. *Journal of the American Chemical Society*, 113(22), 8534-8535. <https://doi.org/10.1021/ja00022a058>
- Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C., & Shore, V. C. (1960). Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. *Nature*, 185(4711), 422-427. <https://doi.org/10.1038/185422a0>
- Kertesz, V., & Van Berkel, G. J. (2010). Fully automated liquid extraction-based surface sampling and ionization using a chip-based robotic nanoelectrospray platform. *Journal of Mass Spectrometry*, 45(3), 252-260. <https://doi.org/10.1002/jms.1709>
- Khanal, A., Pan, Y., Brown, L. S., & Konermann, L. (2012). Pulsed hydrogen/deuterium exchange mass spectrometry for time-resolved membrane protein folding studies: Membrane protein folding studied by pulsed HDX. *Journal of Mass Spectrometry*, 47(12), 1620-1626. <https://doi.org/10.1002/jms.3127>
- Kikhney, A. G., & Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters*, 589(19PartA), 2570-2577. <https://doi.org/10.1016/j.febslet.2015.08.027>
- Kim, M.-K., Jeon, B.-N., Koh, D.-I., Kim, K.-S., Park, S.-Y., Yun, C.-O., & Hur, M.-W. (2013). Regulation of the Cyclin-dependent Kinase Inhibitor 1A Gene (CDKN1A) by the Repressor BOZF1 through Inhibition of p53 Acetylation and Transcription Factor Sp1 Binding. *Journal of Biological Chemistry*, 288(10), 7053-7064. <https://doi.org/10.1074/jbc.M112.416297>
- King, L. S., Kozono, D., & Agre, P. (2004). From structure to disease: The evolving tale of aquaporin biology. *Nature Reviews Molecular Cell Biology*, 5(9), 687-698. <https://doi.org/10.1038/nrm1469>
- Klare, J. P. (2017). Electron Paramagnetic Resonance of Membrane Proteins. En *Encyclopedia of Spectroscopy and Spectrometry* (pp. 442-446). Elsevier. <https://doi.org/10.1016/B978-0-12-409547-2.12118-3>



- Köhler, M., Neff, C., Perez, C., Brunner, C., Pardon, E., Steyaert, J., Schneider, G., Locher, K. P., & Zenobi, R. (2018). Binding Specificities of Nanobody•Membrane Protein Complexes Obtained from Chemical Cross-Linking and High-Mass MALDI Mass Spectrometry. *Analytical Chemistry*, 90(8), 5306-5313. <https://doi.org/10.1021/acs.analchem.8b00236>
- Krainer, G., Keller, S., & Schlierf, M. (2019). Structural dynamics of membrane-protein folding from single-molecule FRET. *Current Opinion in Structural Biology*, 58, 124-137. <https://doi.org/10.1016/j.sbi.2019.05.025>
- Krügel, H., Licht, A., Biedermann, G., Petzold, A., Lassak, J., Hupfer, Y., Schlott, B., Hertweck, C., Platzer, M., Brantl, S., & Saluz, H.-P. (2010). Cervimycin C resistance in *Bacillus subtilis* is due to a promoter up-mutation and increased mRNA stability of the constitutive ABC-transporter gene *bmrA*: *bmrA* promoter up-mutation and cervimycin C resistance. *FEMS Microbiology Letters*, 313(2), 155-163. <https://doi.org/10.1111/j.1574-6968.2010.02143.x>
- Kumar Bharathkar, S., Parker, B. W., Malyutin, A. G., Haloi, N., Huey-Tubman, K. E., Tajkhorshid, E., & Stadtmueller, B. M. (2020). The structures of secretory and dimeric immunoglobulin A. *eLife*, 9, e56098. <https://doi.org/10.7554/eLife.56098>
- Kumari, K. (2023). Functionally active cross-linked protein oligomers formed by homocysteine thiolactone. *Scientific Reports*.
- Lacabanne, D., Orelle, C., Lecoq, L., Kunert, B., Chuilon, C., Wiegand, T., Ravaud, S., Jault, J.-M., Meier, B. H., & Böckmann, A. (2019). Flexible-to-rigid transition is central for substrate transport in the ABC transporter BmrA from *Bacillus subtilis*. *Communications Biology*, 2(1), 149. <https://doi.org/10.1038/s42003-019-0390-x>
- Laganowsky, A., Reading, E., Allison, T. M., Ulmschneider, M. B., Degiacomi, M. T., Baldwin, A. J., & Robinson, C. V. (2014). Membrane proteins bind lipids selectively to modulate their structure and function. *Nature*, 510(7503), 172-175. <https://doi.org/10.1038/nature13419>
- Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T. R., & Weirauch, M. T. (2018). The Human Transcription Factors. *Cell*, 172(4), 650-665. <https://doi.org/10.1016/j.cell.2018.01.029>
- Lawson, M. R., Ma, W., Bellecourt, M. J., Artsimovitch, I., Martin, A., Landick, R., Schulten, K., & Berger, J. M. (2018). Mechanism for the Regulated Control of Bacterial Transcription Termination by a Universal Adaptor Protein. *Molecular Cell*, 71(6), 911-922.e4. <https://doi.org/10.1016/j.molcel.2018.07.014>
- Lee, S. C., Knowles, T. J., Postis, V. L. G., Jamshad, M., Parslow, R. A., Lin, Y., Goldman, A., Sridhar, P., Overduin, M., Muench, S. P., & Dafforn, T. R. (2016). A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nature Protocols*, 11(7), 1149-1162. <https://doi.org/10.1038/nprot.2016.070>
- Lemoine, D., Jiang, R., Taly, A., Chataigneau, T., Specht, A., & Grutter, T. (2012). Ligand-Gated Ion Channels: New Insights into Neurological Disorders and Ligand Recognition. *Chemical Reviews*, 112(12), 6285-6318. <https://doi.org/10.1021/cr3000829>
- Leney, A. C., & Heck, A. J. R. (2017). Native Mass Spectrometry: What is in the Name? *Journal of the American Society for Mass Spectrometry*, 28(1), 5-13. <https://doi.org/10.1007/s13361-016-1545-3>
- Leney, A. C., McMorran, L. M., Radford, S. E., & Ashcroft, A. E. (2012). Amphipathic Polymers Enable the Study of Functional Membrane Proteins in the Gas Phase. *Analytical Chemistry*, 84(22), 9841-9847. <https://doi.org/10.1021/ac302223s>
- LibreTexts.org. (2022). *Host-Guest Chemistry and  $\pi$ - $\pi$  Stacking Interactions*. <https://chem.libretexts.org/@go/page/162915>

- Liddell, E. (2013). Antibodies. En D. Wild, R. Jhon, C. Sheehan, S. Binder, & J. He (Eds.), *The immunoassay Handbook* (Fourth, pp. 245-265). Elsevier. <https://doi.org/10.1016/B978-0-08-097037-0.00008-7>
- Lu, M. (2021). Single-Molecule FRET Imaging of Virus Spike–Host Interactions. *Viruses*, *13*(2), 332. <https://doi.org/10.3390/v13020332>
- Lu, Y., Zhang, H., Niedzwiedzki, D. M., Jiang, J., Blankenship, R. E., & Gross, M. L. (2016). Fast Photochemical Oxidation of Proteins Maps the Topology of Intrinsic Membrane Proteins: Light-Harvesting Complex 2 in a Nanodisc. *Analytical Chemistry*, *88*(17), 8827-8834. <https://doi.org/10.1021/acs.analchem.6b01945>
- Luecke, H., Schobert, B., Richter, H.-T., Cartailier, J.-P., & Lanyi, J. K. (1999). Structure of bacteriorhodopsin at 1.55 Å resolution. *Journal of Molecular Biology*, *291*(4), 899-911. <https://doi.org/10.1006/jmbi.1999.3027>
- M. Hunt, J., & Tuder, R. (2012). Alpha 1 Anti-Trypsin: One Protein, Many Functions. *Current Molecular Medicine*, *12*(7), 827-835. <https://doi.org/10.2174/156652412801318755>
- Ma, J., Lu, Y., Wu, D., Peng, Y., Loa-Kum-Cheung, W., Peng, C., Quinn, R. J., Shui, W., & Liu, Z.-J. (2017). Ligand identification of the adenosine A<sub>2A</sub> receptor in self-assembled nanodiscs by affinity mass spectrometry. *Anal. Methods*, *9*(40), 5851-5858. <https://doi.org/10.1039/C7AY01891F>
- Ma, Y., Ding, Y., Song, X., Ma, X., Li, X., Zhang, N., Song, Y., Sun, Y., Shen, Y., Zhong, W., Hu, L. A., Ma, Y., & Zhang, M.-Y. (2020). Structure-guided discovery of a single-domain antibody agonist against human apelin receptor. *Science Advances*, *6*(3), eaax7379. <https://doi.org/10.1126/sciadv.aax7379>
- MacKinnon, R., Cohen, S. L., Kuo, A., Lee, A., & Chait, B. T. (1998). Structural Conservation in Prokaryotic and Eukaryotic Potassium Channels. *Science*, *280*(5360), 106-109. <https://doi.org/10.1126/science.280.5360.106>
- Marchand, J., Martineau, E., Guitton, Y., Dervilly-Pinel, G., & Giraudeau, P. (2017). Multidimensional NMR approaches towards highly resolved, sensitive and high-throughput quantitative metabolomics. *Current Opinion in Biotechnology*, *43*, 49-55. <https://doi.org/10.1016/j.copbio.2016.08.004>
- Marty, M. T. (2020). Nanodiscs and mass spectrometry: Making membranes fly. *International Journal of Mass Spectrometry*, *458*, 116436. <https://doi.org/10.1016/j.ijms.2020.116436>
- Marty, M. T., Das, A., & Sligar, S. G. (2012). Ultra-thin layer MALDI mass spectrometry of membrane proteins in nanodiscs. *Analytical and Bioanalytical Chemistry*, *402*(2), 721-729. <https://doi.org/10.1007/s00216-011-5512-3>
- Matic, I., Schimmel, J., Hendriks, I. A., van Santen, M. A., van de Rijke, F., van Dam, H., Gnad, F., Mann, M., & Vertegaal, A. C. O. (2010). Site-Specific Identification of SUMO-2 Targets in Cells Reveals an Inverted SUMOylation Motif and a Hydrophobic Cluster SUMOylation Motif. *Molecular Cell*, *39*(4), 641-652. <https://doi.org/10.1016/j.molcel.2010.07.026>
- Maveyraud, L., & Mourey, L. (2020). Protein X-ray Crystallography and Drug Discovery. *Molecules*, *25*(5), 1030. <https://doi.org/10.3390/molecules25051030>
- Mehmood, S., Allison, T. M., & Robinson, C. V. (2015). Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annual Review of Physical Chemistry*, *66*(1), 453-474. <https://doi.org/10.1146/annurev-physchem-040214-121732>
- Mikhailov, V. A., Griffiths, R. L., & Cooper, H. J. (2017). Liquid extraction surface analysis for native mass spectrometry: Protein complexes and ligand binding. *International Journal of Mass Spectrometry*, *420*, 43-50. <https://doi.org/10.1016/j.ijms.2016.09.011>
- Miles, A. J., & Wallace, B. A. (2016). Circular dichroism spectroscopy of membrane proteins. *Chemical Society Reviews*, *45*(18), 4859-4872. <https://doi.org/10.1039/C5CS00084J>

- Miller, C., Moczydlowski, E., Latorre, R., & Phillips, M. (1985). *Charybdotoxin, a protein inhibitor of single Ca<sup>2+</sup>-activated K<sup>+</sup> channels from mammalian skeletal muscle*. 313, 316-318. <https://doi.org/10.1038/313316a0>
- Molodenskiy, D. S., Mertens, H. D. T., & Svergun, D. I. (2020). An automated data processing and analysis pipeline for transmembrane proteins in detergent solutions. *Scientific Reports*, 10(1), 8081. <https://doi.org/10.1038/s41598-020-64933-1>
- Morgner, N., Barth, H.-D., & Brutschy, B. (2006). A New Way To Detect Noncovalently Bonded Complexes of Biomolecules from Liquid Micro-Droplets by Laser Mass Spectrometry. *Australian Journal of Chemistry*, 59(2), 109. <https://doi.org/10.1071/CH05285>
- Morgner, N., Kleinschroth, T., Barth, H.-D., Ludwig, B., & Brutschy, B. (2007). A novel approach to analyze membrane proteins by laser mass spectrometry: From protein subunits to the integral complex. *Journal of the American Society for Mass Spectrometry*, 18(8), 1429-1438. <https://doi.org/10.1016/j.jasms.2007.04.013>
- Morona, R., Klose, M., & Henning, U. (1984). Escherichia coli K-12 outer membrane protein (OmpA) as a bacteriophage receptor: Analysis of mutant genes expressing altered proteins. *Journal of Bacteriology*, 159(2), 570-578. <https://doi.org/10.1128/jb.159.2.570-578.1984>
- Nadiras, C., Schwartz, A., Delaleau, M., & Boudvillain, M. (2018). Evaluating the Effect of Small RNAs and Associated Chaperones on Rho-Dependent Termination of Transcription In Vitro. En V. Arluison & C. Valverde (Eds.), *Bacterial Regulatory RNA* (Vol. 1737, pp. 99-118). Springer New York. [https://doi.org/10.1007/978-1-4939-7634-8\\_7](https://doi.org/10.1007/978-1-4939-7634-8_7)
- Neutze, R., Huldtt, G., Hajdu, J., & Van Der Spoel, D. (2004). Potential impact of an X-ray free electron laser on structural biology. *Radiation Physics and Chemistry*, 71(3-4), 905-916. <https://doi.org/10.1016/j.radphyschem.2004.04.121>
- Nevoltris, D., & Chames, P. (Eds.). (2018). *Antibody Engineering: Methods and Protocols* (Vol. 1827). Springer New York. <https://doi.org/10.1007/978-1-4939-8648-4>
- Nie, D., Hu, Y., Chen, Z., Li, M., Hou, Z., Luo, X., Mao, X., & Xue, X. (2020). Outer membrane protein A (OmpA) as a potential therapeutic target for Acinetobacter baumannii infection. *Journal of Biomedical Science*, 27(1), 26. <https://doi.org/10.1186/s12929-020-0617-7>
- Olerinyova, A., Sonn-Segev, A., Gault, J., Eichmann, C., Schimpf, J., Kopf, A. H., Rudden, L. S. P., Ashkinadze, D., Bomba, R., Frey, L., Greenwald, J., Degiacomi, M. T., Steinhilper, R., Killian, J. A., Friedrich, T., Riek, R., Struwe, W. B., & Kukura, P. (2021). Mass Photometry of Membrane Proteins. *Chem*, 7(1), 224-236. <https://doi.org/10.1016/j.chempr.2020.11.011>
- Orelle, C., Gubellini, F., Durand, A., Marco, S., Lévy, D., Gros, P., Di Pietro, A., & Jault, J.-M. (2008). Conformational Change Induced by ATP Binding in the Multidrug ATP-Binding Cassette Transporter BmrA. *Biochemistry*, 47(8), 2404-2412. <https://doi.org/10.1021/bi702303s>
- Orelle, C., Mathieu, K., & Jault, J.-M. (2019). Multidrug ABC transporters in bacteria. *Research in Microbiology*, 170(8), 381-391. <https://doi.org/10.1016/j.resmic.2019.06.001>
- Orwick-Rydmark, M., Lovett, J. E., Graziadei, A., Lindholm, L., Hicks, M. R., & Watts, A. (2012). Detergent-Free Incorporation of a Seven-Transmembrane Receptor Protein into Nanosized Bilayer Lipodisq Particles for Functional and Biophysical Studies. *Nano Letters*, 12(9), 4687-4692. <https://doi.org/10.1021/nl3020395>
- Otzen, D. (2014). Membrane protein folding and stability. *Archives of Biochemistry and Biophysics*, 564, 262-264. <https://doi.org/10.1016/j.abb.2014.10.014>

- Pan, Y., Stocks, B. B., Brown, L., & Konermann, L. (2009). Structural Characterization of an Integral Membrane Protein in Its Natural Lipid Environment by Oxidative Methionine Labeling and Mass Spectrometry. *Analytical Chemistry*, 81(1), 28-35. <https://doi.org/10.1021/ac8020449>
- Panda, A., Giska, F., Duncan, A. L., Welch, A. J., Brown, C., McAllister, R., Hariharan, P., Goder, J. N. D., Coleman, J., Ramakrishnan, S., Pincet, F., Guan, L., Krishnakumar, S., Rothman, J. E., & Gupta, K. (2023). Direct determination of oligomeric organization of integral membrane proteins and lipids from intact customizable bilayer. *Nature Methods*, 20(6), 891-897. <https://doi.org/10.1038/s41592-023-01864-5>
- Park, S.-H., Ko, W., Lee, H. S., & Shin, I. (2019). Analysis of Protein-Protein Interaction in a Single Live Cell by Using a FRET System Based on Genetic Code Expansion Technology. *Journal of the American Chemical Society*, 141(10), 4273-4281. <https://doi.org/10.1021/jacs.8b10098>
- Pasman, Z., & von Hippel, P. H. (2000). Regulation of Rho-Dependent Transcription Termination by NusG Is Specific to the *Escherichia coli* Elongation Complex. *Biochemistry*, 39(18), 5573-5585. <https://doi.org/10.1021/bi992658z>
- Pellegrini, C. (2020). The development of XFELs. *Nature Reviews Physics*, 2(7), 330-331. <https://doi.org/10.1038/s42254-020-0197-1>
- Pepinsky, R. B., Arndt, J. W., Quan, C., Gao, Y., Quintero-Monzon, O., Lee, X., & Mi, S. (2014). Structure of the LINGO-1–Anti-LINGO-1 Li81 Antibody Complex Provides Insights into the Biology of LINGO-1 and the Mechanism of Action of the Antibody Therapy. *Journal of Pharmacology and Experimental Therapeutics*, 350(1), 110-123. <https://doi.org/10.1124/jpet.113.211771>
- Perez-Torrado, R., Yamada, D., & Defossez, P.-A. (2006). Born to bind: The BTB protein-protein interaction domain. *BioEssays*, 28(12), 1194-1202. <https://doi.org/10.1002/bies.20500>
- Peters, J. M., Mooney, R. A., Grass, J. A., Jessen, E. D., Tran, F., & Landick, R. (2012). Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. *Genes & Development*, 26(23), 2621-2633. <https://doi.org/10.1101/gad.196741.112>
- Petroff, J. T., Tong, A., Chen, L. J., Dekoster, G. T., Khan, F., Abramson, J., Frieden, C., & Cheng, W. W. L. (2020). Charge Reduction of Membrane Proteins in Native Mass Spectrometry Using Alkali Metal Acetate Salts. *Analytical Chemistry*, 92(9), 6622-6630. <https://doi.org/10.1021/acs.analchem.0c00454>
- Petsko, G., & Ringe, D. (2004). *Protein structure and function (Primers in Biology)*. New Science Press.
- Pollack, L., & Doniach, S. (2009). Time-Resolved X-ray Scattering and RNA Folding. En *Methods in Enzymology* (Vol. 469, pp. 253-268). Elsevier. [https://doi.org/10.1016/S0076-6879\(09\)69012-1](https://doi.org/10.1016/S0076-6879(09)69012-1)
- Postis, V., Rawson, S., Mitchell, J. K., Lee, S. C., Parslow, R. A., Dafforn, T. R., Baldwin, S. A., & Muench, S. P. (2015). The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1848(2), 496-501. <https://doi.org/10.1016/j.bbamem.2014.10.018>
- Qian, J., Liu, T., Yang, L., Daus, A., Crowley, R., & Zhou, Q. (2007). Structural characterization of N-linked oligosaccharides on monoclonal antibody cetuximab by the combination of orthogonal matrix-assisted laser desorption/ionization hybrid quadrupole-quadrupole time-of-flight tandem mass spectrometry and sequential enzymatic digestion. *Analytical Biochemistry*, 364(1), 8-18. <https://doi.org/10.1016/j.ab.2007.01.023>
- Raja, P., & Barron, A. (2023). *NMR Spectroscopy*. Libretexts.Org. [https://chem.libretexts.org/Bookshelves/Analytical\\_Chemistry/Physical\\_Methods\\_in\\_Chemistry\\_and\\_Nano\\_Science\\_\(Barron\)/04%3A\\_Chemical\\_Speciation/4.07%3A\\_NMR\\_Spectroscopy](https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Physical_Methods_in_Chemistry_and_Nano_Science_(Barron)/04%3A_Chemical_Speciation/4.07%3A_NMR_Spectroscopy)

- Rapoport, T. A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature*, *450*(7170), 663-669. <https://doi.org/10.1038/nature06384>
- Rapoport, T. A., Goder, V., Heinrich, S. U., & Matlack, K. E. S. (2004). Membrane-protein integration and the role of the translocation channel. *Trends in Cell Biology*, *14*(10), 568-575. <https://doi.org/10.1016/j.tcb.2004.09.002>
- Rapoport, T. A., Jungnickel, B., & Kutay, U. (1996). *Protein Transport Across the Eukaryotic Endoplasmic Reticulum and Bacterial Inner Membranes*. *65*, 271-303. <https://doi.org/10.1146/annurev.bi.65.070196.001415>
- RCSB.org. (2023). *RESEARCH COLLABORATORY FOR STRUCTURAL BIOINFORMATICS (RCSB Protein Data Bank)*. <https://www.rcsb.org/>
- Reading, E. (2018). Structural Mass Spectrometry of Membrane Proteins within Their Native Lipid Environments. *Chemistry - A European Journal*, *24*(51), 13391-13398. <https://doi.org/10.1002/chem.201801556>
- Refeyn Ltd. (2022). *Mass photometry A new way of characterising biomolecules*. <https://www.refeyn.com/about-mass-photometry>
- Rehkamp, A., Tänzler, D., Tüting, C., Kastritis, P. L., Iacobucci, C., Ihling, C. H., Kipping, M., Koch, K.-W., & Sinz, A. (2021). First 3D-Structural Data of Full-Length Guanylyl Cyclase 1 in Rod-Outer-Segment Preparations of Bovine Retina by Cross-Linking/Mass Spectrometry. *Journal of Molecular Biology*, *433*(10), 166947. <https://doi.org/10.1016/j.jmb.2021.166947>
- Rigaud, J.-L., & Lévy, D. (2003). Reconstitution of Membrane Proteins into Liposomes. En *Methods in Enzymology* (Vol. 372, pp. 65-86). Elsevier. [https://doi.org/10.1016/S0076-6879\(03\)72004-7](https://doi.org/10.1016/S0076-6879(03)72004-7)
- Ritchie, T. K., Grinkova, Y. V., Bayburt, T. H., Denisov, I. G., Zolnerciks, J. K., Atkins, W. M., & Sligar, S. G. (2009). Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs. En *Methods in Enzymology* (Vol. 464, pp. 211-231). Elsevier. [https://doi.org/10.1016/S0076-6879\(09\)64011-8](https://doi.org/10.1016/S0076-6879(09)64011-8)
- Rivera-Torres, I. O., Jin, T. B., Cadene, M., Chait, B. T., & Poget, S. F. (2016). Discovery and characterisation of a novel toxin from *Dendroaspis angusticeps*, named Tx7335, that activates the potassium channel KcsA. *Scientific Reports*, *6*(1), 23904. <https://doi.org/10.1038/srep23904>
- Roach, P. J., Laskin, J., & Laskin, A. (2010). Nanospray desorption electrospray ionization: An ambient method for liquid-extraction surface sampling in mass spectrometry. *The Analyst*, *135*(9), 2233. <https://doi.org/10.1039/c0an00312c>
- Roberts, J. W. (1969). Termination Factor for RNA Synthesis. *Nature*, *224*(5225), 1168-1174. <https://doi.org/10.1038/2241168a0>
- Robinson, C. V. (2019). Mass spectrometry: From plasma proteins to mitochondrial membranes. *Proceedings of the National Academy of Sciences*, *116*(8), 2814-2820. <https://doi.org/10.1073/pnas.1820450116>
- Rodwell, V. W., Bender, D. A., Botham, K. M., Kennelly, P. J., & Weil, P. A. (2018). *Harper bioquímica ilustrada* (Trigésima primera edición). McGraw-Hill Intereamericana.
- Rose, R. J., Damoc, E., Denisov, E., Makarov, A., & Heck, A. J. R. (2012). High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nature Methods*, *9*(11), 1084-1086. <https://doi.org/10.1038/nmeth.2208>
- Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J., & Robinson, C. V. (2008). Ion mobility-mass spectrometry analysis of large protein complexes. *Nature Protocols*, *3*(7), 1139-1152. <https://doi.org/10.1038/nprot.2008.78>
- Sahu, I. D., Craig, A. F., Dunagan, M. M., Troxel, K. R., Zhang, R., Meiberg, A. G., Harmon, C. N., McCarrick, R. M., Kroncke, B. M., Sanders, C. R., & Lorigan, G. A. (2015). Probing Structural Dynamics and Topology of the KCNE1 Membrane Protein in Lipid Bilayers via Site-Directed Spin Labeling and Electron Paramagnetic Resonance Spectroscopy. *Biochemistry*, *54*(41), 6402-6412. <https://doi.org/10.1021/acs.biochem.5b00505>

- Sannes-Lowery, K. A., Griffey, R. H., & Hofstadler, S. A. (2000). Measuring Dissociation Constants of RNA and Aminoglycoside Antibiotics by Electrospray Ionization Mass Spectrometry. *Analytical Biochemistry*, 280(2), 264-271. <https://doi.org/10.1006/abio.2000.4550>
- Scalise, M., Pochini, L., Giangregorio, N., Tonazzi, A., & Indiveri, C. (2013). Proteoliposomes as Tool for Assaying Membrane Transporter Functions and Interactions with Xenobiotics. *Pharmaceutics*, 5(4), 472-497. <https://doi.org/10.3390/pharmaceutics5030472>
- Scheibel, T., & Buchner, J. (2006). Protein Aggregation as a Cause for Disease. En K. Starke & M. Gaestel (Eds.), *Molecular Chaperones in Health and Disease* (Vol. 172, pp. 199-219). Springer-Verlag. [https://doi.org/10.1007/3-540-29717-0\\_9](https://doi.org/10.1007/3-540-29717-0_9)
- Sejwal, K., Chami, M., Baumgartner, P., Kowal, J., Müller, S. A., & Stahlberg, H. (2017). Proteoliposomes – a system to study membrane proteins under buffer gradients by cryo-EM. *Nanotechnology Reviews*, 6(1), 57-74. <https://doi.org/10.1515/ntrev-2016-0081>
- Shih, A. Y., Freddolino, P. L., Sligar, S. G., & Schulten, K. (2007). Disassembly of Nanodiscs with Cholate. *Nano Letters*, 7(6), 1692-1696. <https://doi.org/10.1021/nl0706906>
- Shimadzu Corporation. (2022). *Detection of Ultra High-Mass Proteins Using the OmegaToF Mass Spectrometer*. <https://www.ssi.shimadzu.com/products/maldi/ms/omegatof/applications.html>
- Shoemaker, G. K., Van Duijn, E., Crawford, S. E., Uetrecht, C., Baclayon, M., Roos, W. H., Wuite, G. J. L., Estes, M. K., Prasad, B. V. V., & Heck, A. J. R. (2010). Norwalk Virus Assembly and Stability Monitored by Mass Spectrometry. *Molecular & Cellular Proteomics*, 9(8), 1742-1751. <https://doi.org/10.1074/mcp.M900620-MCP200>
- Shukla, A. K., Westfield, G. H., Xiao, K., Reis, R. I., Huang, L.-Y., Tripathi-Shukla, P., Qian, J., Li, S., Blanc, A., Oleskie, A. N., Dosey, A. M., Su, M., Liang, C.-R., Gu, L.-L., Shan, J.-M., Chen, X., Hanna, R., Choi, M., Yao, X. J., ... Lefkowitz, R. J. (2014). Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature*, 512(7513), 218-222. <https://doi.org/10.1038/nature13430>
- Sigworth, F. J. (2016). Principles of cryo-EM single-particle image processing. *Microscopy*, 65(1), 57-67. <https://doi.org/10.1093/jmicro/dfv370>
- Siligardi, G., Hussain, R., Patching, S. G., & Phillips-Jones, M. K. (2014). Ligand- and drug-binding studies of membrane proteins revealed through circular dichroism spectroscopy. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(1), 34-42. <https://doi.org/10.1016/j.bbamem.2013.06.019>
- Sinz, A., Arlt, C., Chorev, D., & Sharon, M. (2015). Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology: Chemical Crosslinking and Native MS. *Protein Science*, 24(8), 1193-1209. <https://doi.org/10.1002/pro.2696>
- Snijder, J., Rose, R. J., Veesler, D., Johnson, J. E., & Heck, A. J. R. (2013). Studying 18 MDa Virus Assemblies with Native Mass Spectrometry. *Angewandte Chemie International Edition*, 52(14), 4020-4023. <https://doi.org/10.1002/anie.201210197>
- Sobott, F., Hernández, H., McCammon, M. G., Tito, M. A., & Robinson, C. V. (2002). A Tandem Mass Spectrometer for Improved Transmission and Analysis of Large Macromolecular Assemblies. *Analytical Chemistry*, 74(6), 1402-1407. <https://doi.org/10.1021/ac0110552>
- Soltermann, F., Foley, E. D. B., Pagnoni, V., Galpin, M., Benesch, J. L. P., Kukura, P., & Struwe, W. B. (2020). Quantifying Protein-Protein Interactions by Molecular Counting with Mass Photometry. *Angewandte Chemie International Edition*, 59(27), 10774-10779. <https://doi.org/10.1002/anie.202001578>

- Sonn-Segev, A., Belacic, K., Bodrug, T., Young, G., VanderLinden, R. T., Schulman, B. A., Schimpf, J., Friedrich, T., Dip, P. V., Schwartz, T. U., Bauer, B., Peters, J.-M., Struwe, W. B., Benesch, J. L. P., Brown, N. G., Haselbach, D., & Kukura, P. (2020). Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nature Communications*, *11*(1), 1772. <https://doi.org/10.1038/s41467-020-15642-w>
- Stead, M. A., Trinh, C. H., Garnett, J. A., Carr, S. B., Baron, A. J., Edwards, T. A., & Wright, S. C. (2007). A Beta-Sheet Interaction Interface Directs the Tetramerisation of the Miz-1 POZ Domain. *Journal of Molecular Biology*, *373*(4), 820-826. <https://doi.org/10.1016/j.jmb.2007.08.026>
- Stein, R. A., Beth, A. H., & Hustedt, E. J. (2015). A Straightforward Approach to the Analysis of Double Electron–Electron Resonance Data. En *Methods in Enzymology* (Vol. 563, pp. 531-567). Elsevier. <https://doi.org/10.1016/bs.mie.2015.07.031>
- Steinfels, E., Orelle, C., Fantino, J.-R., Dalmas, O., Rigaud, J.-L., Denizot, F., Di Pietro, A., & Jault, J.-M. (2004). Characterization of YvcC (BmrA), a Multidrug ABC Transporter Constitutively Expressed in *Bacillus subtilis*. *Biochemistry*, *43*(23), 7491-7502. <https://doi.org/10.1021/bi0362018>
- Stetsenko, A., & Guskov, A. (2017). An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization. *Crystals*, *7*(7), 197. <https://doi.org/10.3390/cryst7070197>
- Stogios, P. J., Downs, G. S., Jauhal, J. J., Nandra, S. K., & Privé, G. G. (2005). [No title found]. *Genome Biology*, *6*(10), R82. <https://doi.org/10.1186/gb-2005-6-10-r82>
- Stowell, M., & Rees, D. (1995). Structure and Stability of Membrane Proteins. En *Advances in Protein Chemistry* (Vol. 46, pp. 279-311). Elsevier. [https://doi.org/10.1016/S0065-3233\(08\)60338-1](https://doi.org/10.1016/S0065-3233(08)60338-1)
- Strnad, P., McElvaney, N. G., & Lomas, D. A. (2020). Alpha<sub>1</sub>-Antitrypsin Deficiency. *New England Journal of Medicine*, *382*(15), 1443-1455. <https://doi.org/10.1056/NEJMra1910234>
- Sun, P., Tropea, J. E., & Waugh, D. S. (2011). Enhancing the Solubility of Recombinant Proteins in Escherichia coli by Using Hexahistidine-Tagged Maltose-Binding Protein as a Fusion Partner. En T. C. Evans, & M.-Q. Xu (Eds.), *Heterologous Gene Expression in E.coli* (Vol. 705, pp. 259-274). Humana Press. [https://doi.org/10.1007/978-1-61737-967-3\\_16](https://doi.org/10.1007/978-1-61737-967-3_16)
- Susa, A. C., Xia, Z., & Williams, E. R. (2017). Native Mass Spectrometry from Common Buffers with Salts That Mimic the Extracellular Environment. *Angewandte Chemie - International Edition*, *56*(27), 7912-7915. <https://doi.org/10.1002/anie.201702330>
- Takáts, Z., Wiseman, J. M., Gologan, B., & Cooks, R. G. (2004). Mass Spectrometry Sampling Under Ambient Conditions with Desorption Electrospray Ionization. *Science*, *306*(5695), 471-473. <https://doi.org/10.1126/science.1104404>
- Talens, S., Malfliet, J. J. M. C., Van Hal, P. Th. W., Leebeek, F. W. G., & Rijken, D. C. (2013). Identification and characterization of  $\alpha$ 1-antitrypsin in fibrin clots. *Journal of Thrombosis and Haemostasis*, *11*(7), 1319-1328. <https://doi.org/10.1111/jth.12288>
- Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., & Matsuo, T. (1988). Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, *2*(8), 151-153. <https://doi.org/10.1002/rcm.1290020802>
- Tarentino, A. L., Gomez, C. M., & Plummer, T. H. (1985). Deglycosylation of asparagine-linked glycans by peptide:N-glycosidase F. *Biochemistry*, *24*(17), 4665-4671. <https://doi.org/10.1021/bi00338a028>
- Tempel, B. L., Papazian, D. M., Schwarz, T. L., Jan, Y. N., & Jan, L. Y. (1987). Sequence of a Probable Potassium Channel Component Encoded at *Shaker* Locus of *Drosophila*. *Science*, *237*(4816), 770-775. <https://doi.org/10.1126/science.2441471>

- Tholey, A., & Heinzle, E. (2006). Ionic (liquid) matrices for matrix-assisted laser desorption/ionization mass spectrometry—Applications and perspectives. *Analytical and Bioanalytical Chemistry*, 386(1), 24-37. <https://doi.org/10.1007/s00216-006-0600-5>
- Torchilin, V. (2006). Multifunctional nanocarriers☆. *Advanced Drug Delivery Reviews*, 58(14), 1532-1555. <https://doi.org/10.1016/j.addr.2006.09.009>
- Twerenbold, D. (1996). Cryogenic particle detectors. *Reports on Progress in Physics*, 59(3), 349-426. <https://doi.org/10.1088/0034-4885/59/3/002>
- Wang, Y. (2002). The Function of OmpA in Escherichia coli. *Biochemical and Biophysical Research Communications*, 292(2), 396-401. <https://doi.org/10.1006/bbrc.2002.6657>
- Weidmann, S., Barylyuk, K., Nespovityaya, N., Mädler, S., & Zenobi, R. (2013). A New, Modular Mass Calibrant for High-Mass MALDI-MS. *Analytical Chemistry*, 85(6), 3425-3432. <https://doi.org/10.1021/ac400129h>
- Wessels, H. P., & Spiess, M. (1988). Insertion of a multispinning membrane protein occurs sequentially and requires only one signal sequence. *Cell*, 55(1), 61-70. [https://doi.org/10.1016/0092-8674\(88\)90009-8](https://doi.org/10.1016/0092-8674(88)90009-8)
- Wikipedia. (2023). *Hydrogen bond*. [https://en.wikipedia.org/wiki/Hydrogen\\_bond](https://en.wikipedia.org/wiki/Hydrogen_bond)
- Wilkins, S., & Capaldi, R. A. (1998). Solution Structure of the  $\epsilon$  Subunit of the F1-ATPase from Escherichia coli and Interactions of This Subunit with  $\beta$  Subunits in the Complex. *Journal of Biological Chemistry*, 273(41), 26645-26651. <https://doi.org/10.1074/jbc.273.41.26645>
- Wilm, M. (2011). Principles of Electrospray Ionization. *Molecular & Cellular Proteomics*, 10(7), M111.009407. <https://doi.org/10.1074/mcp.M111.009407>
- Xu, Y., Johnson, J., Kohn, H., & Widger, W. R. (2003). ATP Binding to Rho Transcription Termination Factor. *Journal of Biological Chemistry*, 278(16), 13719-13727. <https://doi.org/10.1074/jbc.M212979200>
- Yanes, O., Nazabal, A., Wenzel, R., Zenobi, R., & Aviles, F. X. (2006). Detection of Noncovalent Complexes in Biological Samples by Intensity Fading and High-Mass Detection MALDI-TOF Mass Spectrometry. *Journal of Proteome Research*, 5(10), 2711-2719. <https://doi.org/10.1021/pr060202f>
- Yu, L., Sun, C., Song, D., Shen, J., Xu, N., Gunasekera, A., Hajduk, P. J., & Olejniczak, E. T. (2005). Nuclear Magnetic Resonance Structural Studies of a Potassium Channel–Charybdotoxin Complex. *Biochemistry*, 44(48), 15834-15841. <https://doi.org/10.1021/bi051656d>
- Zhang, Z., Pan, H., & Chen, X. (2009). Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrometry Reviews*, 28(1), 147-176. <https://doi.org/10.1002/mas.20190>
- Zheng, C., Yang, L., Hoopmann, M. R., Eng, J. K., Tang, X., Weisbrod, C. R., & Bruce, J. E. (2011). Cross-linking Measurements of In Vivo Protein Complex Topologies. *Molecular & Cellular Proteomics*, 10(10), M110.006841. <https://doi.org/10.1074/mcp.M110.006841>
- Zou, Y., Weis, W. I., & Kobilka, B. K. (2012). N-Terminal T4 Lysozyme Fusion Facilitates Crystallization of a G Protein Coupled Receptor. *PLoS ONE*, 7(10), e46039. <https://doi.org/10.1371/journal.pone.0046039>



# **Résumé de la thèse :**

## **Caractérisation structurale et fonctionnelle de complexes de protéines membranaires et d'oligomères solubles d'intérêt thérapeutique directement par NALIM (Native Liquid MALDI) – TOF MS, une approche originale de Spectrométrie de Masse**

par Edison Fabian Zhamungui Sanchez

dirigée par Martine Cadène

co-encadrée par Martine Beaufour

*Date de soutenance : 5 Juillet 2023.*

La thèse comprend 5 sections principales: (1) une introduction générale sur les protéines membranaires et les oligomères solubles et une comparaison des méthodes d'analyses de la structure des complexes de protéines, en particulier membranaires, par spectrométrie de masse et d'autres techniques, (2) les objectifs du travail, (3) un chapitre sur le développement de la méthode NALIM pour les oligomères solubles, (4) un chapitre sur le NALIM des protéines membranaires, et (5) une conclusion générale avec les perspectives de ce travail. Ces sections sont résumées ci-dessous.

## INTRODUCTION GENERALE

Les protéines membranaires (MP) sont liées à d'importants processus biologiques à l'intérieur de la cellule. Elles agissent en tant que récepteurs de signalisation, transporteurs de molécules et/ou enzymes (Depping et al., 2022 ; Marty, 2020). En effet, les MPs constituent plus de 50 % des cibles potentielles de médicaments. Cependant, notre compréhension de la structure et de la fonction des MPs est encore limitée par rapport aux protéines solubles en raison de la difficulté de les caractériser. Par exemple, les études biochimiques et structurales nécessitent de grandes quantités de MPs pures et stables. Malheureusement, de nombreuses MPs sont présentes en faible quantité dans la membrane cellulaire. Dans certains cas, la surexpression induite d'une MP recombinante peut déclencher des problèmes de toxicité pour l'organisme ou la cellule qui l'exprime (Depping et al., 2022 ; Di Cesare et al., 2022). De plus, la cristallographie aux rayons X des MPs nécessite leur solubilisation à l'aide de détergents ou de nanodisques soigneusement choisis, avec des taux de succès variables. Cela explique pourquoi seuls 2 ou 3 % de toutes les structures à haute résolution disponibles sont des MPs, bien que près de 30 % du protéome soit composé de MPs (Di Cesare et al., 2022 ; Marty, 2020).

L'analyse par spectrométrie de masse (MS, pour ses sigles en anglais) de biomolécules a été rendue possible grâce à l'invention de méthodes d'ionisation douce, c'est-à-dire l'ESI et le MALDI. Grâce à elles, différentes approches utilisant des conditions *in vitro* contrôlées et non dénaturantes, afin de mieux imiter l'environnement biologique des protéines, ont été développées pour caractériser des complexes moléculaires dans son état natif. Ces méthodes sont collectivement connues sous le nom de spectrométrie de masse native (nMS) et ont largement contribué à la compréhension des protéines et de leurs complexes du point de vue biologique et biophysique (Barrera et al., 2008 ; Chorev et al., 2020 ; Fitzgerald et al., 1996 ; Hernández & Robinson, 2007). Par exemple, l'analyse directe de complexes biomoléculaires intacts peut fournir des informations sur leur stœchiométrie, l'identité des partenaires de liaison, la dynamique des complexes, voire les affinités de liaison (Boeri Erba & Petosa, 2015 ; Gan et al., 2017 ; Jaquillard et al., 2012 ; Leney & Heck, 2017). Donc, les méthodes de spectrométrie de masse native (nMS) ont désormais atteint le statut d'outils puissants pour l'analyse et la caractérisation des complexes biomoléculaires, fournissant des informations clés en soutien aux méthodes à haute résolution telles que la cristallographie aux rayons X, la RMN et la cryo-EM.

Jusqu'à présent, la méthode privilégiée pour l'étude des assemblages macromoléculaires natifs (nMS) repose sur l'ionisation par électrospray (ESI). La première méthode basée sur ESI pour la nMS a été développée au début des années 1990 (Katta & Chait, 1991). Au cours des dernières décennies, plusieurs études révolutionnaires sur les complexes membranaires ont été publiées. La première réussite avec l'ESI-MS natif a été réalisée en 2008 en conservant le transporteur à cassette de liaison

à l'ATP BtuC2D2 en phase gazeuse (Barrera et al., 2008). Depuis lors, des approches de nMS ont été utilisées pour déterminer, par exemple, la stœchiométrie des assemblages contenant des protéines transmembranaires (Barrera et al., 2009) ou pour étudier les interactions des protéines membranaires avec les lipides (pour une revue, voir Bolla et al., 2019). Afin de mieux imiter les membranes cellulaires et/ou faciliter les études structurales, différentes méthodes alternatives de solubilisation telles que les polymères amphipathiques (Leney et al., 2012), les nanodisques (Marty et al., 2012), et les vésicules membranaires soniquées (Chorev et al., 2020) ont été développées et utilisées en MS. Cependant, l'analyse par ESI-MS natif peut encore être difficile en raison du besoin d'additifs et de sels pour stabiliser les protéines dans des tampons aqueux. Ces additifs ne peuvent pas toujours être éliminés sans altérer la structure et la fonction complexes, tandis que les laisser dans le milieu d'analyse peut déclencher la suppression d'ions, la formation d'adduits et le dépôt de résidus sur les surfaces internes des instruments. Ceci est particulièrement vrai pour l'analyse de nMS des protéines membranaires.

En revanche, la technique MALDI est plus tolérante aux contaminants et consomme moins d'échantillons que l'ESI. Cependant, la préparation classique des échantillons MALDI nécessite la dissolution d'un petit acide organique ("matrice") dans un solvant organique dénaturant, suivie de la co-cristallisation des analytes et de la matrice. Cela signifie de plus d'importants obstacles à surmonter pour rendre les conditions MALDI non dénaturantes. Donc, pour surmonter les limitations de la méthode MALDI avec dépôt solide pour la nMS, une nouvelle méthode appelée NALIM pour « Native LIquid MALDI » a été conçue par l'équipe de M. Cadène (Beaufour *et al.*, 2018), bien que l'acronyme ait été conçu après cette publication. NALIM exploite les propriétés uniques du MALDI dans une nouvelle méthode de nMS basée sur un dépôt liquide. Ainsi, contrairement au MALDI conventionnel, le NALIM évite la transition par un état solide.

Le succès de la méthode NALIM dans des conditions natives repose sur une matrice liquide à base de HCCA/3-AQ/glycérol (1:4:6 w/w) avec un pH proche des conditions physiologiques, contournant ainsi le pH acide des matrices traditionnelles. De plus, le glycérol confère une stabilité sous vide à l'analyte (Beaufour et al., 2018). Le fait que le MALDI soit plus tolérant aux détergents, combiné à la capacité du NALIM à maintenir les protéines dans leur état natif, ouvre de nouvelles possibilités et permet de caractériser les protéines membranaires et leurs complexes. En plus, puisque le NALIM est couplé à un analyseur TOF, il n'y a pas de limite théorique à la taille de l'analyte. Ainsi, NALIM ouvre également la voie à l'analyse de grands complexes moléculaires.

Les grands complexes moléculaires représentent une autre cible importante en biologie parce qu'afin de remplir sa ou ses fonctions, une protéine (soluble ou membranaire) doit interagir avec de petites

molécules organiques et de plus grandes biomolécules (par exemple, les acides nucléiques, d'autres protéines), que ce soit à l'intérieur ou à l'extérieur de la cellule. La fonction d'une protéine est généralement assurée par sa participation à des complexes dynamiques non covalents de taille variable, allant des petites aux grandes.

Étant donné que les grands complexes biomoléculaires sont fréquemment hétérogènes, présents en faibles quantités dans la cellule, et que leurs interactions peuvent être transitoires, ils ne sont pas facilement caractérisés par les méthodes biophysiques classiques telles que la cristallographie aux rayons X, la spectroscopie de résonance magnétique nucléaire (RMN) ou la cryo-microscopie électronique (cryo-EM) (Ruotolo et al., 2008). De plus, ces méthodes nécessitent des quantités d'échantillon plus importantes que le NALIM.

De plus, la nMS basée sur l'ESI est généralement couplée à des analyseurs ayant une gamme de masse limitée par conception. Par exemple, les quadripôles balayent à l'aide d'ensembles de tensions RF/DC qui maintiennent les ions stables dans le chemin du quadripôle uniquement dans des limites de  $m/z$  définies. En revanche, la gamme de masse du NALIM couplé à un analyseur de temps de vol (TOF) est en principe illimitée, de sorte que les complexes biomoléculaires peuvent être analysés simultanément sur une large gamme de masse. Donc, grâce à la sensibilité du NALIM et le fait qu'il est couplé à des analyseurs TOF peut présenter des avantages pour caractériser de grands complexes biomoléculaires d'une manière beaucoup plus accessible que les méthodes basées sur l'ESI et d'autres méthodes biophysiques.

Étant donné les avantages potentiels de la méthode NALIM exposés ci-dessus, ce travail de thèse présente les différentes optimisations instrumentales et de préparation des échantillons qui ont été nécessaires pour caractériser les protéines membranaires et les grands complexes moléculaires solubles. De même, il démontre comment le NALIM a été utilisé pour obtenir des informations d'un grand intérêt du point de vue biologique, avec des perspectives pour son utilisation future en tant que méthode complémentaire aux techniques de biologie structurale classiques, par exemple pour le développement de médicaments ciblant ces biomolécules.

# OBJECTIFS

L'objectif général de ce travail est d'obtenir des informations structurales sur les complexes de protéines membranaires qui ne sont pas facilement, dynamiquement ou rapidement accessibles par d'autres méthodes biophysiques en biologie structurale, ce qui peut contribuer à compléter l'image obtenue par ces autres méthodes. En accédant à la structure de ces protéines et de leurs complexes, nous pouvons mieux comprendre les mécanismes d'action qui sous-tendent leur rôle dans des processus biologiques clés, tels que la signalisation cellulaire, l'absorption des nutriments et le transport des médicaments, et contribuer au développement de médicaments pour traiter diverses maladies.

À cette fin, des développements spécifiques d'une méthode exploitant les avantages de la spectrométrie de masse native (native MS) et de la source d'ionisation MALDI étaient nécessaires. A l'objectif 1 défini au début de la thèse et affiné au fil du temps, s'est ajouté l'objectif 2 qui s'est naturellement imposé en cours de thèse:

## **1. Obtenir des informations structurales et fonctionnelles sur les protéines membranaires et leurs complexes dans leur état natif, telles que la stœchiométrie, la stabilité et la liaison aux ligands.**

Dans cette optique, il est nécessaire de développer une méthode de nMS (spectrométrie de masse native) basée sur MALDI préétablie appelée NALIM spécifiquement pour les protéines membranaires et leurs partenaires interactifs.

De plus, les objectifs secondaires suivants nous aideront à atteindre notre objectif principal :

- Évaluer la potentialité d'application de cette méthode à différents types de complexes de protéines membranaires, tels que les transporteurs, les canaux ioniques et les protéines récepteurs.
- Étendre l'utilisation de NALIM à la caractérisation de complexes de protéines membranaires dans un environnement lipidique *in vivo*, car les mimétiques de membrane ne peuvent pas fournir toutes les caractéristiques de la bicouche lipidique native.
- Optimiser la méthode pour une résolution et une sensibilité dans une gamme de masse relativement élevée.
- Établir un calibrant fiable basé sur un système soluble, plus facile à manipuler qu'une protéine membranaire, et couvrant une gamme de masse adéquate allant jusqu'à 100 000 Da et plus.

*En poursuivant ces objectifs, la recherche d'un calibrant de masse élevée a conduit à la découverte d'une gamme d'applicabilité plus étendue que prévu initialement, ouvrant la voie à un nouveau champ d'application pour la méthode NALIM. Comme les grands oligomères solubles constituent une classe essentielle de complexes en biologie, nous avons décidé de développer et de tirer parti de la méthode NALIM pour obtenir des informations structurales sur ces systèmes. Ainsi, mon deuxième objectif majeur était :*

**2. Caractériser de grands complexes moléculaires solubles afin de déterminer leur degré d'oligomérisation et l'effet de la liaison de ligands.**

Les oligomères sont des structures essentielles par lesquelles un certain nombre de protéines ayant une pertinence thérapeutique exercent leurs fonctions à l'intérieur de la cellule. Notre objectif était de tirer parti de l'analyseur TOF, qui en principe a une gamme de masse illimitée, et du NALIM, pour fournir des informations sur la structure/fonction de ces complexes moléculaires malgré leur grande taille, leur hétérogénéité et leur faible quantité disponible.

# **CHAPITRE 1: ANALYSE DE GRANDS OLIGOMÈRES DE PROTÉINES SOLUBLES PAR NALIM**

Ce chapitre est l'ébauche d'un manuscrit dont les auteurs sont :

**Zhamungui E<sup>1</sup>, Mance L<sup>1</sup>, Suskiewicz MJ<sup>1</sup>, Boudvillain M<sup>1</sup>, Irving JA<sup>2</sup>, Beaufour M<sup>1</sup>, Cadène M<sup>1</sup>**

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR4301, affilié à l'Université d'Orléans, Orléans 45071, France.

<sup>2</sup> Cambridge Institute for Medical Research, Department of Medicine, University of Cambridge, Welcome Trust/MRC Building, Hills Road, Cambridge CB2 0XY, U.K.

Dans ce chapitre, nous nous concentrons sur de grands complexes biomoléculaires impliqués dans la régulation de la transcription, tels que les facteurs de transcription (TF, pour ses sigles en anglais) et une hélicase ADN.

Les TF peuvent former des complexes dynamiques avec d'autres biomolécules pour contrôler l'expression génique, ce qui influencera la différenciation, le cycle cellulaire et le destin de la cellule (Lambert et al., 2018). Cependant, malgré leur importance principale pour la cellule et les connaissances structurales et fonctionnelles approfondies les concernant, notre compréhension des aspects fondamentaux de leur assemblage en structures plus grandes et du rôle de ces structures est encore limitée. Ainsi, les facteurs de transcription ZBTB8A ont été étudiés par NALIM pour obtenir des informations structurales, en particulier sur la distribution d'oligomères que forme leurs dimères. En parallèle, nous nous sommes également concentrés sur le complexe Rho-NusG formé par l'association du facteur de transcription NusG, avec l'hélicase Rho hexamérique. Rho est impliquée dans la terminaison de la transcription chez *Escherichia coli* et de nombreuses autres bactéries, processus qui pourrait être ciblé à des fins antibiotiques, et NusG est soupçonné de jouer un rôle dans la modulation de ce processus. Nous avons donc envisagé d'appliquer la méthode NALIM afin de caractériser la liaison Rho/NusG, en particulier sa stœchiométrie et sa stabilité.

Les protéines NusG/Spt5 appartiennent à une famille de facteurs de transcription conservés chez tous les organismes. Alors que NusG se trouve chez les bactéries, Spt5 est son homologue correspondant chez les archées et les eucaryotes (Lawson et al., 2018). NusG peut jouer deux rôles opposés dans la transcription. Son domaine N-terminal (NTD) peut se lier à l'ARN polymérase (RNAP) afin d'améliorer l'activité de transcription de la RNAP (c'est-à-dire l'antitermination). En revanche, le domaine C-terminal (CTD) peut recruter des facteurs accessoires pour favoriser la fin de la transcription (c'est-à-dire la terminaison) (Burmam et al., 2010).

D'autre part, Rho est une hélicase dépendante de l'ATP associée à un processus de terminaison de la transcription trouvé exclusivement chez les bactéries. Rho reconnaît et se lie à des séquences d'environ 80 nucléotides connues sous le nom de sites d'utilisation de Rho (RUT) (Lawson et al., 2018 ; Peters et al., 2012). En l'absence d'ATP et avant la reconnaissance de RUT, Rho est un homohexamère avec une conformation en anneau ouverte. Dans cette conformation ouverte, Rho se lie aux sites RUT et fait pénétrer l'ARN naissant à l'intérieur de l'anneau. En présence d'ATP, Rho passe à une conformation en anneau fermé, piégeant l'ARN à l'intérieur du pore central de l'anneau. Ensuite, Rho se déplace en aval, utilisant l'énergie de l'hydrolyse de l'ATP, jusqu'à ce qu'il rencontre la RNAP, ce qui arrêtera la transcription (Lawson et al., 2018).

Il a été suggéré que NusG joue un rôle clé en permettant la terminaison de la transcription dépendante de Rho de manière indépendante de la séquence.

Le facteur de transcription ZBTB8A, également connu sous le nom de BOZF1, appartient à la famille des facteurs de transcription ZBTB, définie par un domaine BTB (broad-complex, tramtrack and bric-à-brac) dans la zone N-terminale et des doigts de zinc à l'extrémité C-terminale. Selon Kim *et al.* (2013), ZBTB8A est un répresseur transcriptionnel qui agit sur les promoteurs proximaux du gène de la cycline-dépendante kinase, CDKN1A, inhibant sa transcription. Cela conduit à la répression de la transcription en aval de certaines protéines régulatrices liées à l'apoptose, à la différenciation cellulaire, à la progression du cycle cellulaire et à l'oncogenèse, suggérant un lien possible avec les tumeurs et les métastases. De plus, les protéines ZBTB sont fréquemment des substrats pour la SUMOylation, une modification post-traductionnelle (PTM) réversible directement associée à des processus cellulaires clés tels que la réplication et la réparation de l'ADN, la régulation transcriptionnelle, la progression du cycle cellulaire, la transduction des signaux, et plus encore (Hendriks et al., 2014 ; Matic et al., 2010).

Chez certaines protéines ZBTB, en particulier celles des insectes, une interaction entre dimères a été suggérée. Les découvertes récentes de Mance et ses collaborateurs (publication en préparation) suggèrent que le domaine BTB de ZBTB8A peut générer des filaments. La dynamique d'assemblage de ces filaments n'est pas connue, ni l'abondance relative de chaque oligomère dans la distribution des oligomères formés. Les travaux effectués sur le calibrant A1AT (décrit ci-dessous) ont montré que le NALIM peut s'avérer utile pour commencer à répondre à ces questions.

### **Calibrant NALIM pour une gamme de $m/z$ élevée**

Si on considère une stœchiométrie d'une protéine NusG par hexamère de Rho, la masse de ce complexe moléculaire à l'état natif est proche de 305 000 Da. Afin de développer la méthode NALIM de manière à aider les biologistes à mieux comprendre les grands complexes moléculaires solubles,



la première exigence est donc de disposer de calibrants appropriés dans une gamme de  $m/z$  plus élevée que ce qui est couramment observé en Spectrométrie de Masse MALDI-TOF. À notre connaissance, de tels étalons de calibration n'étaient pas disponibles commercialement pour la spectrométrie de masse native lorsque nous avons initiés ces travaux. Ainsi, étendre les applications de NALIM à une gamme de  $m/z$  élevée nécessitait la découverte de nouveaux calibrants de haute masse et l'optimisation des paramètres instrumentaux pour leur mise en œuvre.

Nous avons envisagé différentes protéines comme calibrants potentiels de haute masse. Dans un premier temps l'IgA, en tant que dimère d'une immunoglobuline (Ig), était susceptible d'atteindre facilement cette gamme. Cependant, les IgA commerciales disponibles à ce moment-là étaient glycosylées, ce qui produit une microhétérogénéité et donc des pics larges en spectrométrie de masse. LA calibration nécessitant des pics les plus fins possibles, nous avons tenté de déglycosyler la protéine avant les expériences d'optimisation.

L'un de nos premiers résultats a été d'observer que l'IgA de souris se trouve sous forme de monomère ou de dimère. De plus, contrairement aux IgA humaines, les chaînes lourdes et légères de l'IgA de souris peuvent être liées de manière covalente ou non covalente. Cela explique les différentes espèces d'IgA trouvées lors du contrôle de qualité dans des conditions dénaturantes.

Par ailleurs, les pics des dimères d'IgA restent plus larges que les pics des monomères d'IgA. Ainsi, bien que le dimère ait deux fois plus de glycanes, la déglycosylation résulte en une perte de masse moitié moindre que sur le monomère. De plus, la diminution de masse après déglycosylation montre que plus le nombre de chaînes légères est élevé, moins le nombre de glycanes éliminés est important. Par conséquent, la stœchiométrie des chaînes légères et lourdes est directement liée au degré de déglycosylation de l'IgA, ce qui suggère une déglycosylation incomplète des espèces dimères.

Nous avons tenté d'améliorer le rapport signal/bruit (S/N, par ces sigles en anglais) et la résolution des pics à  $m/z$  élevé en nous concentrant sur l'optimisation du délai d'application du voltage d'accélération, afin d'obtenir des pics homogènes et bien résolus qui garantissant une calibration correcte.

Pour cet objectif, nous avons testé l'effet sur le rapport S/N et la résolution pour cinq valeurs de délai d'extraction retardée allant de 850 à 2 550 ns par incréments de 425 ns. Quatre espèces d'IgA ont été sélectionnées pour l'optimisation afin de couvrir la gamme de  $m/z$  de 100 000 à 320 000.

Les résultats montrent que plus la valeur de délai est élevée, meilleure est la sensibilité pour les complexes moléculaires de grande taille (c'est-à-dire les espèces de dimères d'IgA), comme le montre l'augmentation de la valeur du rapport signal/bruit (S/N). De plus, à 1 275 ns, les ions de moins de 150 000  $m/z$  atteignent la valeur de rapport S/N la plus élevée. Les délais supérieurs à 1 275 ns sont

préjudiciables à leur rapport S/N, alors que pour les ions de plus de 150 000  $m/z$ , un délai supérieur à 1 700 ns est toujours nécessaire. Malheureusement, comme mentionné précédemment, la déglycosylation de l'IgA de souris était incomplète. Ce fait a limité son utilisation en tant que calibrant pour une gamme de masse supérieure à 150 000 Da et nous a poussés à rechercher un calibrant alternatif.

Sur la base des résultats précédents avec l'IgA, nous avons axé notre travail sur des protéines recombinantes exprimées dans des bactéries afin de pouvoir travailler sur des protéines non-glycosylées. L'alpha-1-antitrypsine ( $\alpha$ 1AT), un inhibiteur de protéases à sérine exprimé dans *E. coli* forme dans certaines conditions des oligomères de manière non covalente. Des données de SDS-PAGE ont montré que cette distribution d'oligomères pouvait s'avérer intéressante pour la calibration sur des centaines de kDa, et l' $\alpha$ 1AT a donc été évalué comme calibrant pour le NALIM. Comme notre objectif était de détecter des ions avec des rapports  $m/z$  supérieurs à 150 000, le premier délai testé avec  $\alpha$ 1AT était de 1 700 ns.

Nous avons évalué l'effet sur le rapport signal/bruit (S/N) et la résolution de quatorze valeurs de délai allant de 1700 à 4500 ns (valeur maximale atteignable) par incréments de 200 ns.

Les résultats ont montré que l'augmentation du délai améliore la sensibilité pour les ions de  $m/z$  supérieur à 150 000. À 3 100 ns, l'undécamère devient le plus grand état oligomérique observable, ce qui représente un gain de 100 000 dans la gamme de  $m/z$  par rapport aux ions observés à 1 700 ns (c'est-à-dire l'octamère). Un fait intéressant est que l'effet positif du délai sur le rapport S/N semble atteindre une limite à 3 100 ns. Cette observation est cohérente avec le comportement trouvé avec l'IgA. Cela confirme que l'optimisation du délai - qui a un impact positif sur le rapport S/N dans une gamme de  $m/z$  élevée - est limitée à une gamme de masse définie.

Après avoir effectué les optimisations de délai, nous avons testé l' $\alpha$ 1AT en tant que calibrant pour les ions de grande masse ( $m/z > 100\,000$ ). L'objectif de cette expérience était de vérifier l'exactitude de la détermination de la masse obtenue en utilisant l' $\alpha$ 1AT comme calibrant en NALIM. À cette fin, une IgG déglycosylée a été utilisée comme protéine de test. Étant donné que la séquence de l'IgG n'était pas connue, sa masse déterminée par HRMS a été utilisée comme proxy de la masse théorique. La masse déterminée en NALIM a ensuite été comparée à la masse déterminée par HRMS.

En utilisant l' $\alpha$ 1AT comme calibrant, la masse moyenne déterminée en NALIM est de 6 Da ou 41 ppm de plus que la masse déterminée par HRMS (146 448 Da). Compte tenu de la résolution limitée obtenue en conditions natives, cette erreur peut être considérée comme faible et acceptable pour des objectifs tels que la détermination de la stœchiométrie et la mise en évidence d'interactions protéine-

protéine ou protéine-ligand, où des erreurs de quelques daltons n'affectent pas sérieusement l'interprétation des résultats. En formant une distribution d'homo-oligomères qui constitue une échelle moléculaire régulière, la protéine  $\alpha$ 1AT est à notre connaissance, le premier exemple d'un calibrant pouvant être utilisé en spectrométrie de masse native des grands assemblages moléculaires.

## **Aperçu structural de grands complexes protéiques solubles par NALIM**

### ***Le complexe NusG-Rho en NALIM***

La stœchiométrie du complexe Rho-NusG d'*E. coli* a été déterminée par titrage en solution. Pour confirmer cette stœchiométrie et caractériser davantage ce complexe, nous avons conçu une expérience pour évaluer si NusG affecte la stabilité de l'hexamère de Rho en utilisant la méthode NALIM.

Cette expérience a permis d'identifier le complexe hexamérique de Rho-NusG en NALIM ainsi que des monomères jusqu'à des pentamères de Rho associés à une sous-unité NusG. Étant donné qu'il est peu probable que des stœchiométries inférieures à l'hexamère de Rho soient présentes en solution, l'observation de pics correspondant à des espèces allant des monomères aux pentamères suggère que de la dissociation en phase gazeuse (GPD) se produit. La fraction de l'hexamère de Rho liée à NusG, calculée comme  $f_{\text{bound}} = \text{Rho NusG} / (\text{Rho} + \text{Rho NusG})$ , est de 30%. Comme l'expérience a été conçue pour atteindre une saturation de 100 %, cette fraction de liaison inférieure suggère que le complexe NusG-Rho subit une forte GPD.

Un fait intéressant est que la résolution mesurée sur les pics du complexe Rho/NusG est plus élevée que celle des pics de Rho libre. L'élargissement des pics de Rho libre par rapport aux pics du complexe Rho/NusG pourrait indiquer que le Rho est enclin à une décomposition métastable. Si c'est le cas, les pics de complexe relativement étroits suggéreraient que NusG peut prévenir la décomposition métastable de Rho, ce qui peut être interprété comme l'effet stabilisateur de NusG sur la protéine Rho.

### ***Oligomérisation du facteur de transcription ZBTB8A***

L'idée d'une oligomérisation des protéines ZBTB8A a été suggérée pour la première fois par l'identification, par cristallographie aux rayons X, d'une interface de contact entre les dimères à travers des surfaces hydrophobes (Mance *et al.*, publication en préparation). Cependant, le degré d'oligomérisation n'a pas pu être déterminé précisément à l'aide de cette technique. Par conséquent, nous avons entrepris d'obtenir des informations structurales sur l'oligomérisation de ZBTB8A en utilisant la méthode NALIM. Pour cela, la protéine ZBTB8A WT (résidus 1-146) de *Xenopus laevis* a été produite par nos collaborateurs. Sur la base de la structure mentionnée ci-dessus, la sérine 103

a été identifiée comme un acide aminé clé pour la capacité à former des interfaces dimère-dimère et des oligomères de haut ordre. Une protéine mutée ZBTB8A (S103R), dans laquelle la sérine 103 est remplacée par une arginine, a donc également été produite comme témoin négatif pour la détection d'oligomères dans les expériences.

Les résultats montrent que des oligomères de ZBTB8A WT sont observés en NALIM. Le plus haut état oligomérique identifié est le tétradécamère (14-mère), qui a une masse moléculaire d'environ 246 kDa. Comme aucun oligomère supérieur au dimère n'est observé avec la mutation ZBTB8A S103, les oligomères WT observés peuvent être attribués à une association spécifique, plutôt qu'à une agrégation en phase gazeuse. La formation d'oligomères a été confirmée par l'observation de filaments en microscopie électronique avec la protéine WT et l'absence de formation de filaments avec la protéine mutante.

Dans l'ensemble, ces résultats confirment que ZBTB8A WT forme des oligomères spécifiques de haut ordre basés sur des interactions non covalentes de monomères et/ou de dimères.

## **Conclusion**

Les résultats décrits dans ce chapitre montrent pour la première fois qu'en exploitant la gamme de masse de l'analyseur TOF, la méthode NALIM peut rendre accessibles à l'analyse en nMS les grands complexes non covalents allant de 100 000 à 500 000 Da. Notre travail a permis de montrer que le facteur de transcription ZBTB8A forme des oligomères en solution, suite à des observations d'un motif polymère infini par cristallographie aux rayons X. La formation de filaments de ZBTB8A a ensuite été confirmée par microscopie électronique. L'avantage de la méthode NALIM est d'accéder à la distribution des oligomères avec des nombres de sous-unités clairement définis, sur une large de gamme de masse, ce qui est difficilement faisable avec les méthodes de biologie structurale qu'elle vient compléter.

La méthode NALIM a permis de fournir une image de l'interaction entre le facteur de transcription NusG et l'hélicase Rho hexamérique. Ici, nous avons constaté une augmentation de la stabilité à la fois de l'hexamère de Rho et du complexe Rho-ATP suite à la liaison de NusG.

Comme le montre cette étude, l'optimisation du délai est essentielle pour l'étude des complexes de haute masse, car elle améliore la sensibilité pour les ions de haute masse.

Contrairement aux immunoglobulines, l' $\alpha$ 1AT génère des pics homogènes et résolus qui forment une échelle moléculaire de 150 000 Da à au moins 500 000 Da, ce qui en fait un calibrant approprié pour

les masses élevées. Bien que le NALIM produise des pics moins résolus que le MALDI en conditions dénaturantes, il est adapté à la détermination de la stœchiométrie des complexes et à la caractérisation des interactions protéine-protéine.

## **CHAPITRE 2: ANALYSE NALIM DES COMPLEXES DE PROTEINES MEMBRANAIRES (MP)**

Ce chapitre est l'ébauche d'un manuscrit dont les auteurs sont :

**Zhamungui E<sup>1</sup>, Orelle C<sup>2</sup>, Jault JM<sup>2</sup>, Arizala JD<sup>3</sup>, Poget SF<sup>3</sup>, Morisset-Lopez S<sup>1</sup>, Beaufour M<sup>1</sup>, Cadène M<sup>1</sup>**

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR4301, affilié à l'Université d'Orléans, Orléans 45071, France.

<sup>2</sup>Molecular Microbiology & Structural Biochemistry (MMSB) UMR 5086, CNRS/University of Lyon, Lyon 69367, France

<sup>3</sup>Department of Chemistry, College of Staten Island, City University of New York, 2800 Victory Boulevard, Staten Island, New York 10314, United States.

L'objectif du travail dans ce chapitre est d'étendre NALIM à l'analyse de protéines membranaires (MPs) multimériques et des complexes qu'elles forment avec des ligands. À cette fin, différentes protéines membranaires solubilisées dans des détergents ont été étudiées pour couvrir un large éventail de fonctions biologiques. Cela comprenait d'une part KcsA, un archétype des canaux ioniques au potassium (K<sup>+</sup>) provenant de *Streptomyces lividans*, et d'autre part BmrA et SpABC, deux transporteurs de xénobiotiques de la famille ABC provenant de *Bacillus subtilis* et de *Streptococcus pneumoniae*, respectivement. La protéine LINGO humaine a été abordée comme représentante de protéine eucaryote. De plus, dans le but de se rapprocher des conditions de solubilisation *in vivo* des protéines membranaires, des analyses NALIM ont été réalisées à partir de liposomes, ainsi que de nano-vésicules membranaires (MV) obtenus directement à partir de membranes cellulaires.

KcsA est un archétype de canal K<sup>+</sup> produit par la bactérie Gram+ *S. lividans*. KcsA est une petite protéine solidement repliée, dont la structure quaternaire est formée de quatre monomères identiques positionnés en un cône inversé. Bien que tous les canaux potassiques partagent une séquence d'acides aminés critique (séquence caractéristique) dans leur filtre de sélectivité au K<sup>+</sup>, les résidus bordant l'entrée du pore du canal sont assez variables. Pour comprendre le mode d'interaction des toxines de venin de différents animaux avec l'entrée du pore des canaux potassiques eucaryotes, le canal K<sup>+</sup> Shaker (Sh) sert généralement de modèle. Ainsi, par des expériences de mutagenèse dirigée effectuées de longue date, les acides aminés clés directement liés à l'interaction entre le canal K<sup>+</sup> Shaker et des toxines telles que l'agitoxine et la charybdotoxine ont été déterminés.

Afin de mimer la séquence de l'entrée du pore du canal K<sup>+</sup> Shaker eucaryote dans celle du canal bactérien KcsA, ces mutations ponctuelles clés (Q58A, T61S et R64D) ont été introduites dans la séquence de KcsA (MacKinnon et al., 1998), permettant ainsi d'exprimer un canal K<sup>+</sup> chimérique Sh-KcsA sensible à ces toxines. Une toxine capable d'activer KcsA, c'est à dire une actitoxine, nommée

TX7335, a été découverte à partir de venin de serpent mamba au cours d'une collaboration avec S. Poget. Son site de liaison au canal et sa stœchiométrie de liaison restaient à déterminer. La méthode NALIM a été mise à profit pour tenter d'obtenir ces informations, en réalisant une expérience de compétition entre TX7335 et la charybdotoxine.

BmrA est un transporteur ABC que l'on trouve dans *Bacillus subtilis* et qui fonctionne comme une pompe à efflux pour les molécules susceptibles d'endommager les cellules. Il s'agit d'un système modèle pour les transporteurs associés à la 'multi-drug resistance' (MDR), c'est-à-dire impliqués dans la résistance des bactéries pathogènes aux antibiotiques. BmrA est également pertinent en raison de son homologie avec la P-glycoprotéine que l'on trouve chez les mammifères, une protéine membranaire liée à la résistance aux agents chimiothérapeutiques dans le traitement des cancers.

Sur sa structure quaternaire, BmrA est un homodimère, où chaque monomère comporte un domaine de transport transmembranaire (TMD) fusionné à un domaine de liaison à l'ATP (NBD). Les TMD sont responsables de la reconnaissance et du transport des substrats à travers la membrane, tandis que les NBD hydrolysent l'ATP et fournissent l'énergie nécessaire au transport (Chaptal et al., 2022 ; Javed et al., 2022 ; Lacabanne et al., 2019).

### **Optimisations pour augmenter la stabilité des complexes membranaires**

Afin d'obtenir des informations structurales sur les complexes membranaires, il était important d'avoir un pourcentage du complexe membranaire qui permet d'évaluer les résultats de nos expériences. C'est pourquoi l'optimisation des paramètres instrumentaux, du rapport matrice/échantillon et de la composition de l'échantillon a été entreprise pour augmenter la stabilité de ces complexes. Pour cela, le transporteur BmrA a été choisi comme modèle et le pourcentage de dimère comme indicateur de stabilité en phase gazeuse.

D'abord, il a été observé que dans la méthode NALIM, la puissance du laser influence la stabilité des complexes membranaires. Par ailleurs, passé le seuil d'ionisation, la stabilité maximale du dimère de BmrA est atteinte à 200 tirs lasers.

Ensuite, étant donné l'importance des matrices dans les expériences MALDI, des efforts ont été concentrés sur le rapport matrice/échantillon pour améliorer la stabilité du dimère BmrA. Les résultats indiquent une meilleure préservation du dimère de BmrA à un rapport matrice/échantillon plus élevé. Finalement, afin de préserver l'interaction à l'interface entre les monomères du dimère, des modifications des paramètres affectant les interactions microscopiques (liaisons hydrophobes, liaisons ioniques), tels que le pourcentage d'eau et de glycérol, et la concentration en sel(s), ont été explorés. Les résultats montrent que le pourcentage de dimères augmente avec la diminution du

pourcentage d'eau. De plus, le pourcentage d'eau a un impact sur le rapport signal/bruit (S/N), avec une diminution des rapports S/N du monomère et du dimère à des concentrations d'eau plus élevées. Enfin, une concentration en sel supérieure à 15 mM, selon les résultats obtenus par analyse NALIM de BmrA, conduit à des pics plus larges et à une augmentation du bruit de fond, suggérant la formation probable d'adduits de Na<sup>+</sup>.

Après avoir défini les conditions optimales pour l'analyse de complexes de protéines membranaires, différentes expériences ont été conçues en NALIM pour obtenir des informations sur la structure, la stœchiométrie, la stabilité, et/ou la localisation de sites de liaison de complexes protéiques membranaires de BmrA, SpABC ou de KcsA.

### **Stabilisation du dimère de BmrA par liaison au ligand ATP en présence de Mg<sup>2+</sup> et V<sub>i</sub>**

La forme sauvage (WT) de BmrA peut être piégé dans la conformation « outward facing » (OF), ce qui résulte en une conformation du dimère plus stable que l'apoprotéine. Cette conformation est induite par la liaison d'ATP et d'ions magnésium (Mg<sup>2+</sup>) en présence de vanadate (V<sub>i</sub>). Dans ce cas, le Mg<sup>2+</sup> agit comme cofacteur pour la liaison de l'ATP, tandis que le vanadate agit comme agent de piégeage - dans la conformation OF du complexe BmrA-ADP après l'hydrolyse de l'ATP (Javed et al., 2022).

Afin de pouvoir observer en NALIM une éventuelle stabilisation du dimère de BmrA par l'ATP, le piégeage du complexe BmrA-ADP par le V<sub>i</sub> a été mis à profit et quatre rapports Protéine-Ligand (P-L) différents ont été testés.

En NALIM, on observe que la stabilité du dimère augmente lorsque le ligand est ajouté dans un rapport BmrA/ATP/Mg<sup>2+</sup>/V<sub>i</sub> de 1:10:10:1. Afin de confirmer que l'augmentation de la stabilité du dimère de type sauvage (WT) est due à la liaison du ligand, l'hydrolyse de l'ATP a été contrôlée en utilisant une forme mutante de BmrA qui ne possède pas d'activité ATPase, de sorte qu'il ne se forme pas de complexe BmrA-ADP pouvant être piégé par le V<sub>i</sub>. Dans la mutation K380A, le remplacement de la lysine invariante du motif Walker-A dans la protéine de type sauvage (WT) par une alanine empêche l'hydrolyse de l'ATP (Javed et al., 2022). Ainsi, l'ajout de ligand et d'agent de piégeage à l'incubation ne devrait pas améliorer la stabilité du dimère K380A comme cela a été le cas pour le WT.

Les résultats de cette expérience ont montré que le pourcentage de dimères du mutant BmrA K380A reste à peu près constant en fonction de la concentration de ligand. Ceci confirme qu'en présence de V<sub>i</sub>, le dimère de type sauvage (WT) de BmrA est stabilisé par la liaison de l'ATP hydrolysé en ADP, tandis que le dimère K380A ne peut pas bénéficier de cet effet.



### **Localisation par compétition du site de liaison Tx7335 sur KcsA**

Dans cette section, le but était de déterminer quel(s) site(s) de liaison est (sont) ciblé(s) par la toxine Tx7335 sur le canal K<sup>+</sup>, ainsi que sa stœchiométrie de liaison. Cette expérience repose sur le fait que Tx7335 se lie à la protéine WT, car cette toxine a été isolée grâce à un test de pull-down en utilisant KcsA WT comme "appât" (Rivera-Torres et al., 2016). La charybdotoxine (ChTx) se lie à Sh-KcsA, une protéine KcsA chimérique qui mime la structure du « pore » du canal Shaker eucaryote de manière à pouvoir se lier à cette toxine (MacKinnon et al., 1998). Tx7335 se lie également au canal Sh-KcsA sensible à la charybdotoxine (Rivera-Torres et al., 2016). Si les toxines ne partagent pas le même site de liaison sur KcsA, nous devrions pouvoir observer la liaison concomitante des toxines à la protéine chimérique.

Dans l'expérience de compétition, une série de pics correspondant potentiellement à ShKcsA ayant formé des adduits de K<sup>+</sup> a été observée. Des shifts de masse correspondant à l'addition d'une seule molécule de Tx7335 ou de ChTx sont observés lorsque l'on ajoute l'une ou l'autre toxine. En présence des deux toxines, un pic à 87 504 *m/z*, avec un shift de masse de 11 669 Da par rapport à l'un des pics d'adduits du tétramère est observé. Cette masse dépasse de 40 Da la somme de la masse théorique des deux toxines (11 631 Da), un ajout de masse qui compte tenu des arrondis, pourrait correspondre à un ion K<sup>+</sup>.

En se concentrant sur la gamme de *m/z* plus basse, des formes trimères, dimères et monomères de ShKcsA sont observées, comme cela a été le cas pour les expériences sur la protéine KcsA WT. Étant donné que ces formes ne sont pas présentes en solution, elles sont vraisemblablement produites par dissociation en phase gazeuse de la forme de tétramère. Ces formes dissociées montrent davantage de preuves de l'interaction entre la protéine et la toxine, et potentiellement le piégeage de 3 ions K<sup>+</sup> dans le filtre de sélectivité du canal.

Dans l'ensemble, ces données montrent la liaison concomitante des deux toxines à la protéine chimérique ShKcsA, ce qui indique que Tx7335 et ChTx ne partagent pas le même site de liaison. La stœchiométrie du complexe ShKcsA-Tx7335 semble être de 1:1. Ceci pourrait être confirmé par une expérience de titration.

### **Preuve de BmrA dans les vésicules de la membrane cellulaire.**

Des vésicules membranaires provenant de membranes bactériennes contenant la forme sauvage (WT) de BmrA ont été préparées en utilisant le protocole de Chorev et al. (2020). Un contrôle de qualité (QC) de notre échantillon a été effectué en utilisant la méthode UTL. Le QC a montré que la protéine

BmrA est globalement intacte dans les vésicules membranaires (VM). La masse moyenne de la protéine déterminée dans le contrôle de qualité (QC) est en effet de 66 342,5 Da, ce qui correspond à la masse théorique calculée à partir de la séquence protéique sans les deux premiers acides aminés Met et Ala, avec une erreur de masse de +2,5 Da, soit 37 ppm. De plus, une autre protéine est observée à partir des VMs. En se basant sur la masse moyenne de cette protéine (37 064 Da), OmpA est identifiée comme une correspondance possible avec une erreur de masse de -5,4 Da, soit 145 ppm.

Les expériences en NALIM ont montré un pic à  $m/z$  135 836, décalé de +3 155 Da par rapport à la masse théorique du dimère basée dans la séquence protéique [3-607]. Une explication possible pour ce pic est qu'il correspond à une espèce dimérique de BmrA avec un ou plusieurs ligands. Dans ce cas, la masse supplémentaire suggère qu'une molécule ou un ensemble de molécules, peut-être des lipides, est étroitement associé au dimère ou même piégé à l'intérieur de l'interface du dimère, formant un complexe suffisamment stable avec BmrA pour être éjecté avec lui et détecté dans l'analyse NALIM. Un pic proche de la masse du monomère BmrA est observé avec une masse de 383 Da supérieure à la masse trouvée dans le QC, qui correspond elle-même à la séquence manquant les 2 premiers acides aminés. Un pic satellite de +2 257 Da par rapport au pic du monomère est également trouvé, suggérant que le monomère lie moins bien le(s) ligand(s) trouvés associés au dimère. Le dimère n'étant pas observé seul, alors que du monomère libre est présent, semble indiquer que le(s) ligand(s) est capable de stabiliser le dimère face à la dissociation en phase gazeuse.

Aucun pic lié à un complexe ou à un monomère d'OmpA n'a été détecté. Cependant, trois pics supplémentaires à  $m/z$  50 376, 99 896 et 101 265 n'ont pas de correspondance directe avec les protéines trouvées dans le contrôle de qualité. Ils sont également présents dans la condition de "spiking" avec le détergent, mais ils ne sont pas observés dans les VMs contenant SpABC, ce qui suggère qu'ils sont spécifiques à la membrane bactérienne exprimant BmrA. Une interprétation possible est que lors de l'étape de sonication de la préparation des VMs, des forces de cisaillement rompent une liaison peptidique dans BmrA, de sorte que des formes monomères et dimères plus courtes sont observées. En effet, dans le contrôle de qualité, un groupe de pics a été observé de 15 700 à 16 300 Da, ce qui pourrait correspondre à des segments complémentaires des formes dimères plus courtes de BmrA observées à  $m/z$  99 896 et 101 265. Le pic à  $m/z$  50 376 pourrait correspondre au monomère de ces dimères plus courts de BmrA généré par dissociation en phase gazeuse, bien que celui-ci ne soit pas observé dans le contrôle de qualité.

## Conclusion

Ce travail met en évidence le potentiel de la méthode NALIM comme une méthode prometteuse pour caractériser les complexes de protéines membranaires dans leur état natif, fournissant rapidement des

informations complémentaires aux outils classiques de biologie structurale et à d'autres méthodes biophysiques comme par exemple la stœchiométrie des complexes, l'emplacement des sites de liaison et la stabilisation par un ligand.

En appliquant le NALIM dans des conditions plus proches de l'*in vivo*, nous avons montré qu'il était possible lors de l'ionisation d'expulser le dimère de BmrA à partir de vésicules membranaires, révélant en même temps sa probable association à des lipides endogènes avec un potentiel effet de stabilisation du dimère par ces lipides.

Dans cette étude, nous avons élaboré et appliqué des optimisations des paramètres instrumentaux et de la préparation d'échantillon pour améliorer le signal et la stabilité des complexes de MPs dans NALIM. De plus, nous montrons que le NALIM peut être utilisé pour suivre les interactions protéine-protéine (PPI) et les interactions protéine-ligand (PLi) et fournir des informations structurales telles que la stœchiométrie des complexes, l'effet de différents facteurs sur leur stabilité et leurs interactions avec les partenaires de liaison à l'état natif.

## CONCLUSION GÉNÉRALE ET PERSPECTIVES

L'objectif principal de ce travail était d'obtenir des informations structurales et fonctionnelles sur les protéines membranaires et leurs complexes dans leur état natif. Pour atteindre cet objectif, l'optimisation des paramètres instrumentaux et de la préparation des échantillons a joué un rôle clé dans la détection et la préservation de la structure tertiaire et quaternaire des protéines membranaires (MPs). Une fois les conditions expérimentales établies, différents types de protéines transporteurs membranaires, notamment un transporteur ABC et un canal ionique, ont été caractérisés. Par exemple, BmrA est une protéine de *Bacillus* modèle des pompes à efflux MDR responsables de résistances multiples aux antibiotiques. Grâce à la méthode NALIM, la stabilisation du dimère de la protéine par un effet de liaison à un ligand a été démontrée. De plus, une expérience de compétition a donné des éléments de preuve que le site de liaison de la toxine actitoxine Tx7335 au canal potassique KcsA est distinct du site de liaison de la charybdotoxine à l'entrée du pore. Grâce à cela, nous pouvons émettre l'hypothèse que cette toxine active le canal potassique par un possible effet allostérique.

Afin de se rapprocher des conditions de la protéine membranaire dans sa membrane native, la méthode NALIM a été appliquée à des complexes dans des protéoliposomes et des vésicules membranaires. Bien que les protéoliposomes présentent des défis en termes de sensibilité, le dimère de BmrA a pu être éjecté à partir de nano-vésicules membranaires natives, et donner un signal dans l'analyse NALIM. Ainsi, le NALIM a le pouvoir de caractériser des complexes de protéines membranaires dans leur bicouche lipidique native. Ces résultats ouvrent la voie à une analyse plus approfondie des complexes de protéines membranaires dans des conditions plus proches de l'*in vivo*, telles que directement à partir d'une cellule bactérienne vivante.

Par ailleurs, différentes protéines ont été explorées lors du développement de cette méthode afin de trouver un calibrant approprié, permettant d'obtenir des données interprétables et fiables dans la gamme de masse élevée, c'est-à-dire, pour des rapports  $m/z$  supérieurs à 150 000. L'utilisation d'une immunoglobuline A (IgA) de souris commerciale, présente soit sous forme monomère soit sous forme dimère, semblait initialement prometteuse. Cependant, sa conformation a empêché une déglycosylation complète des espèces dimériques. Grâce à une collaboration, l'alpha1-antitrypsine ( $\alpha$ 1AT), un inhibiteur de sérine-protéase, a été rendue disponible sous une forme recombinante non glycosylée produisant des oligomères stables de poids moléculaire élevé. Une première analyse de cette protéine en NALIM a montré une distribution d'oligomères de masse supérieure à 250 000 Da, qui peut être utilisée comme une échelle moléculaire pour l'étalonnage. Grâce à l'optimisation des paramètres instrumentaux, la détection d'oligomères de masse supérieure à 500 000 Da a été obtenue. Ainsi, les optimisations instrumentales ont amélioré la résolution et la sensibilité de la méthode dans

une gamme de masse relativement élevée. La validation de l' $\alpha$ 1AT en tant que calibrant a été réalisée grâce à la détermination précise de la masse d'une IgG déglycosylée.

Les résultats obtenus avec l' $\alpha$ 1AT ont ouvert la voie à l'étude de grands complexes moléculaires solubles (par exemple, d'une masse supérieure à 200 000 Da). Des systèmes candidats formant des oligomères qui sont biologiquement pertinents, tels que la distribution d'oligomères d'un facteur de transcription, ou un complexe entre une hélicase et un autre facteur de transcription, ont été caractérisés en NALIM. Cette méthode, une fois de plus, a permis un accès rapide à des informations structurales telles que le degré d'oligomérisation de ZBTB8A, qui n'avait pas été précédemment obtenu avec d'autres techniques structurales telles que la chromatographie d'exclusion stérique analytique (SEC) ou la microscopie électronique. De plus, un effet de stabilisation de l'hexamère Rho par NusG a été découvert en partie grâce à la plus grande finesse des pics pour les pics du complexe Rho/NusG par comparaison aux pics de Rho seul.

Finalement, des limitations ont également été identifiées, telles qu'un pourcentage élevé de dissociation en phase gazeuse (GPD), ce qui confirme les informations initiales obtenus pour cette méthode (Beaufour et al.). En général, la GDP a pu être minimisée en partie par des optimisations de paramètres instrumentaux et de préparation de l'échantillon. Pour les espèces non-glycosylées, la résolution reste faible par rapport à la nMS basée sur l'ESI, ce qui a jusqu'à présent empêché l'identification de la stœchiométrie de liaison de molécules de ligands de faible masse tels que l'ATP à un complexe. Ainsi, ces limitations offrent des opportunités pour des recherches futures visant à améliorer la technique et son application à une large gamme de systèmes de masse moléculaire élevée. Par exemple, puisque la matrice influence significativement l'ionisation des échantillons dans les expériences MALDI des optimisations plus fines de la composition de la matrice pourraient être réalisées pour améliorer la qualité des spectres obtenus en termes de sensibilité et de résolution. Ces optimisations ont le potentiel de favoriser une identification plus précise et fiable des complexes de protéines membranaires.

De plus, l'exploration du rôle potentiel des ligands en tant qu'agents de stabilisation des complexes moléculaires pourrait contribuer à fortement réduire la dissociation en phase gazeuse lors de l'analyse NALIM. Ainsi, une meilleure préservation des structures de ces complexes moléculaires facilitera leur caractérisation complète. Enfin, des modifications instrumentales représentent une autre stratégie potentielle pour améliorer la caractérisation des grands complexes moléculaires par NALIM. Par exemple, l'intégration d'un détecteur spécifique de haute masse peut augmenter la sensibilité dans la gamme de masse élevée, permettant ainsi l'analyse structurale détaillée de complexes allant jusqu'à

2 MDa et plus. Cette amélioration peut fournir des informations précieuses sur ces complexes, contribuant à une meilleure compréhension de leur structure et de leur fonction.

En résumé, ce travail démontre que la méthode NALIM fournit des informations précieuses sur la structure et la fonction de divers systèmes biologiques, telles que la stœchiométrie, la stabilité et l'effet de liaison des ligands, de manière rapide et directe. C'est une bonne méthode complémentaire des méthodes de biologie structurale classiques ainsi que des méthodes existantes pour l'étude de la structure quaternaire des protéines. De plus, les phénomènes découverts au cours de ce travail contribuent à la compréhension des protéines membranaires et des grands complexes moléculaires solubles, et ont des implications importantes dans les domaines de la biologie et de la pharmacologie, avec le potentiel d'ouvrir de nouvelles voies pour la conception de médicaments et les thérapies ciblant ces protéines.

## Caractérisation structurale et fonctionnelle de complexes de protéines membranaires et d'oligomères solubles d'intérêt thérapeutique directement par NALIM (Native Liquid MALDI) – TOF MS, une approche originale de Spectrométrie de Masse

### Résumé :

Bien que les protéines membranaires (PMs) représentent les deux tiers des cibles thérapeutiques potentielles, seulement 10 % de ces protéines ont été exploitées comme cibles jusqu'à présent. Ceci est dû à un manque d'information structurale, attribué aux défis liés à la production des PMs et au besoin d'agents de solubilisation. De plus, des fonctions biologiques clés dépendent de la formation de grands complexes par le biais d'interactions avec d'autres biomolécules telles que les acides nucléiques, les peptides ou d'autres protéines. Ainsi, la taille de ces complexes (c'est-à-dire, 100 000 Da ou plus) et leur grande hétérogénéité imposent des limitations à leur caractérisation structurale.

Au cours de la dernière décennie, la caractérisation de complexes non covalents par spectrométrie de masse native (nMS) est apparue comme un bon complément aux techniques traditionnelles de biologie structurale. Nous montrons ici comment une nouvelle méthode appelée « NAtive LIquid MALDI » (NALIM), qui exploite les avantages intrinsèques de MALDI-TOF MS – une grande tolérance aux contaminants, une faible consommation d'échantillons et une gamme de masse d'analyse théoriquement illimitée - peut répondre aux questions sur la structure des PMs et des complexes biomoléculaires de grande taille. De plus, pour se rapprocher des conditions *in vivo* de la bicouche lipidique de la membrane native, la méthode NALIM a également été appliquée à la caractérisation de complexes de protéines membranaires dans les proteoliposomes et les vésicules membranaires.

À cette fin, des conditions non dénaturantes soigneusement contrôlées, des composants de solution d'échantillon et des paramètres instrumentaux ont été optimisés pour améliorer la stabilité des complexes, ainsi que la résolution et la sensibilité dans la gamme des hautes masses. Grâce à ces optimisations, la méthode NALIM a facilité la caractérisation de divers transporteurs membranaires (par exemple, un transporteur ABC et un canal ionique), de grandes distributions d'oligomères (par exemple, le facteur de transcription ZBTB8A) et de grands complexes moléculaires (complexe Rho-NusG). De plus, le développement de la méthode NALIM a impliqué l'exploration de différentes protéines pour trouver un calibrant approprié pour l'analyse dans la gamme des hautes masses. Une préparation d'alpha1-antitrypsine ( $\alpha$ 1AT) formant une échelle moléculaire sur une large gamme de masse a été validée comme calibrant.

En résumé, l'analyse par NALIM a permis un accès rapide et direct à la caractérisation des PMs et des grands complexes biomoléculaires. Les informations fournies par NALIM ont une valeur significative en biologie et en pharmacologie. Ainsi, la méthode NALIM promet d'être une alternative pour de nouvelles stratégies de découverte de médicaments.

Mots-clés: Spectrométrie de masse native, nMS, NALIM, MALDI-TOF, biologie structurale, protéine membranaire, oligomères, grands complexes, échelle moléculaire, interactions protéine-protéine, stœchiométrie, ligand binding.

## Direct structural/functional characterization of therapeutically relevant membrane protein complexes and soluble oligomers by NALIM (Native Liquid MALDI) – TOF MS, an original Mass Spectrometry approach

### Summary:

Although membrane proteins (MPs) represent two-thirds of potential therapeutic targets, only 10% of these proteins have been exploited thus far. This is due to a lack of structural information, attributed to the challenges associated with producing MPs and the need for solubilizing agents. Additionally, key biological functions depend on the formation of large complexes through interactions with other biomolecules such as nucleic acids, peptides, or other proteins. Thus, the size of these complexes (i.e., 100,000 Da or higher) and their high heterogeneity impose additional limitations on their structural characterization.

Over the past decade, the characterization of noncovalent complexes through native mass spectrometry (nMS) has emerged as a valuable complement to traditional Structural Biology techniques. Here we show how a novel method called Native Liquid MALDI (NALIM), which leverages the inherent strengths of MALDI-TOF MS - high tolerance to contaminants, low sample consumption, and a theoretically limitless mass range of analysis – to answer questions about the structure of MPs and large biomolecular complexes. Additionally, to get closer to the *in vivo* conditions of the native membrane's lipid bilayer, NALIM was also applied to the characterization of membrane proteins in proteoliposomes and membrane vesicles.

To this end, carefully controlled non-denaturing conditions, sample solution components, and instrumental parameters were fine-tuned to enhance complex stability, as well as resolution and sensitivity in the high mass range. Thanks to these optimizations, NALIM facilitated the characterization of diverse membrane transporters (e.g., an ABC transporter and an ion channel), large oligomer distributions (e.g., ZBTB8A transcription factor), and large molecular complexes (Rho-NusG complex). Furthermore, the development of the NALIM method involved exploring different proteins to find a suitable calibrant for high mass range analysis. A preparation of alpha1-antitrypsin ( $\alpha$ 1AT) that forms a molecular ladder over a wide mass range was validated as a calibrant.

In summary, NALIM enabled rapid and straightforward access to characterizing MPs and large biomolecular complexes. The information provided by NALIM has significant value in biology and pharmacology. Thus, NALIM promises to be an alternative for new strategies of drug discovery.

Keywords: Native mass spectrometry, nMS, NALIM, MALDI-TOF, structural biology, membrane protein, oligomers, large complexes, molecular ladder, protein-protein interactions, stoichiometry, ligand binding.

