

---

# **Native Mass Spectrometry for the Structural Characterization of Membrane-Related Protein Complexes**

Dissertation with the aim of achieving a doctoral degree at the Faculty of  
Mathematics, Informatics and Natural Sciences

Department of Chemistry  
of Universität Hamburg

submitted by

**Johannes Heidemann**

born in Lübeck

Hamburg, 2018

---

## Reviewers

---

**Dr. Charlotte Uetrecht**

Heinrich-Pette-Institute,  
Leibniz Institute for Experimental Virology, Hamburg, Germany  
  
European XFEL GmbH, Schenefeld, Germany

---

**Prof. Dr. Hartmut Schlüter**

Institute for Clinical Chemistry and Laboratory Medicine,  
Mass Spectrometric Proteomics Group,  
University Medical Center Hamburg-Eppendorf, Hamburg, Germany  
  
Department of Chemistry, Universität Hamburg, Germany

Submission of Thesis: 31.08.2018

Date of the Oral Defense: 09.11.2018

Approval of the Dissertation for Publication: 09.11.2018

---

This thesis was prepared from July 2014 to August 2018 under the supervision of Dr. Charlotte Uetrecht at the Heinrich-Pette-Institute, Leibniz Institute for Experimental Virology in the working group ‘Dynamics of Viral Structures’. Second supervisor was Prof. Dr. Wolfram Brune, located at the Heinrich Pette Institute and associated to the Department of Chemistry of the University of Hamburg.



---

# Table of Contents

	Page
Table of Contents	I
Published Data	IV
Abstract	V
Zusammenfassung	VII
<b>1 Introduction: Native Mass Spectrometry</b>	<b>1</b>
1.1 Studying Non-Covalent Interactions in Structural Biology	1
1.2 General Overview	2
1.3 Sample Requirements	3
1.3.1 Protein Production	3
1.3.2 Native MS-compatible Buffers	4
1.4 Nano Electrospray Ionization	6
1.5 Mass Analysers	10
1.5.1 Time-of-Flight Mass Analysers	11
1.5.2 Quadrupole Mass Analysers	12
1.6 Studying Non-Covalent Assemblies in the Gas Phase	13
1.6.1 Solution Structures versus Gas-phase Structures	13
1.6.2 Protein-Ligand interactions	14
1.6.3 Protein-Protein Interactions	17
1.7 Gas-phase Dissociation and Fragmentation Techniques	20
1.8 Ion Mobility Mass Spectrometry	21
1.9 High-mass Nano ESI QToF Instruments	24
<b>2 Objective and Problem Definition of the Thesis</b>	<b>25</b>
<b>3 Binding of Two Calmodulins to the Plasma-membrane Ca<sup>2+</sup>-ATPase ACA8</b>	<b>27</b>
3.1 Introduction	27
3.2 Results and Discussion	29
3.3 Methods	33
3.3.1 Protein Production and Purification	33
3.3.2 Native Mass Spectrometry	34
3.4 Outlook	35
<b>4 Epsin and Sla2 Form Assemblies through Phospholipid Interfaces</b>	<b>37</b>
4.1 Introduction	37
4.1.1 Clathrin-mediated Endocytosis	37
4.1.2 Phosphatidylinositol-4,5-bisphosphate in Clathrin-Mediated Endocytosis	39
4.1.3 Clathrin Adaptor Proteins: Epsins and Sla2/Hip1R	40

<b>4.2</b>	<b>Results and Discussion</b>	<b>44</b>
4.2.1	Native MS Reveals Cooperative Binding of 2 PI(4,5)P <sub>2</sub> to Epsin ENTH Domains	44
4.2.2	ANTH Domains Bind Two PI(4,5)P <sub>2</sub> Molecules	48
4.2.3	Fungal ENTH and Sla2 ANTH Domains Form Ordered Assemblies	50
4.2.4	ENTH:ANTH:PI(4,5)P <sub>2</sub> Complex Formation Dynamics	56
4.2.5	IMMS Analysis of <i>S. cerevisiae</i> ENTH:ANTH:PI(4,5)P <sub>2</sub>	57
4.2.6	SID MS of <i>S. cerevisiae</i> ENTH:ANTH:PI(4,5)P <sub>2</sub> Complexes	60
4.2.7	Human ENTH Domains Form Hexamers in Presence of PI(4,5)P <sub>2</sub>	64
4.2.8	Structural Analysis of Human ENTH Hexamers Using IMMS and SAXS	66
4.2.9	Hip1R ANTH Domains Bind to ENTH-6mers	71
4.2.10	An ANTH Subfamily Does Not Form ENTH:ANTH:PI(4,5)P <sub>2</sub> Complexes	73
4.2.11	Cross-Species <i>in Vitro</i> Assembly of ENTH:ANTH:PI(4,5)P <sub>2</sub> Complexes	74
4.2.12	General Assembly Model of ENTH:ANTH:PI(4,5)P <sub>2</sub> Complexes	76
<b>4.3</b>	<b>Methods</b>	<b>79</b>
4.3.1	Cloning	79
4.3.2	Protein Production and Purification	79
4.3.3	Native Mass Spectrometry: Sample Preparations	80
4.3.4	Native Mass Spectrometry: Protein-PI(4,5)P <sub>2</sub> Interactions	80
4.3.5	Native Mass Spectrometry: ENTH:ANTH:PI(4,5)P <sub>2</sub> Complexes	81
4.3.6	Crystal Structure Determination	81
4.3.7	Small-angle X-ray Scattering	82
4.3.8	Dynamic Light Scattering	83
4.3.9	Isothermal Titration Calorimetry	83
4.3.10	Ion Mobility Mass Spectrometry	83
4.3.11	Surface-induced Dissociation Mass Spectrometry	84
<b>4.4</b>	<b>Outlook</b>	<b>86</b>
<b>5</b>	<b>Molecular Organization of the <i>Salmonella</i> T3SS Sorting Platform</b>	<b>89</b>
<b>5.1</b>	<b>Introduction</b>	<b>89</b>
5.1.1	<i>Salmonella</i> Typhimurium Type III Secretion System	89
5.1.2	The T3SS Sorting Platform	90
<b>5.2</b>	<b>Results and Discussion</b>	<b>94</b>
5.2.1	SpaO-Isoforms Assemble to SpaO-2SpaOc Complexes	94
5.2.2	SpaO Domain Interactions	98
5.2.3	OrgB Dimers Induce Dimerization of SpaO-2SpaOc	101
5.2.4	Soluble Building Blocks Result from InvC Binding to SpaO/SpaOc/OrgB	103
5.2.5	Structural Modelling of Building Blocks against SEC-SAXS Data	107
5.2.6	C-ring Proteins Do Not Interact with Chaperone-Effector Complexes <i>in Vitro</i>	109
5.2.7	T3SS Sorting Platform Assembly Model	111
<b>5.3</b>	<b>Methods</b>	<b>113</b>
5.3.1	Cloning and Mutagenesis of <i>Salmonella</i> Genes	113
5.3.2	Detection of SpaO and SpaOc and Protein Secretion in <i>Salmonella</i> Cells	113
5.3.3	Protein Synthesis and Purification	114
5.3.4	Native Mass Spectrometry	115
5.3.5	Multi-angle Light Scattering	116
5.3.6	Small-angle X-ray Scattering	116
5.3.7	Structural Modelling against SAXS Data	117
<b>5.4</b>	<b>Outlook</b>	<b>119</b>

---

<b>Supplement</b>	<b>121</b>
<b>S1 Supplementary Figures</b>	<b>121</b>
<b>S1.1 Binding of Two Calmodulins to the Plasma membrane Ca<sup>2+</sup>-ATPase ACA8</b>	<b>121</b>
<b>S1.2 Epsin and Sla2 Form Assemblies through Phospholipid Interfaces</b>	<b>123</b>
<b>S1.3 Molecular Organization of the <i>Salmonella</i> T3SS Sorting Platform</b>	<b>131</b>
<b>S2 Supplementary Tables</b>	<b>139</b>
<b>S2.1 Binding of Two Calmodulins to the Plasma-membrane Ca<sup>2+</sup>-ATPase ACA8</b>	<b>139</b>
<b>S2.2 Epsin and Sla2 Form Assemblies through Phospholipid Interfaces</b>	<b>141</b>
<b>S2.3 Molecular Organization of the <i>Salmonella</i> T3SS Sorting Platform</b>	<b>148</b>
<b>S3 Material</b>	<b>154</b>
<b>S3.1 Chemicals and Biomaterials</b>	<b>154</b>
<b>S3.2 Consumables</b>	<b>155</b>
<b>S3.3 Instruments</b>	<b>156</b>
<b>S3.4 Software</b>	<b>156</b>
<b>S4 Used Hazardous Substances According to GHS</b>	<b>157</b>
<b>S5 Contributions</b>	<b>159</b>
<b>S6 List of Figures</b>	<b>160</b>
<b>S7 List of Tables</b>	<b>162</b>
<b>S8 List of Abbreviations</b>	<b>163</b>
<b>S9 References</b>	<b>165</b>
<b>S10 Acknowledgement</b>	<b>183</b>
<b>S11 Declaration of Authorship</b>	<b>185</b>

---

## Published Data

Parts of this thesis are published or were submitted for publishing:

### Chapter 1

Heidemann J, Krichel B, Uetrecht C (2018) Native Massenspektrometrie für die Proteinstrukturanalytik. *BIOspektrum* 24:2 164-7

### Chapter 3

Nitsche J, Josts I\*, Heidemann J\*, Mertens HD\*, Maric S, Moulin M, Haertlein M, Busch S, Forsyth VT, Svergun DI, Uetrecht C, Tidow H (2018) Structural basis for activation of plasma-membrane Ca<sup>2+</sup>-ATPase by calmodulin, *Communications Biology*, *in revision*

\*equal contribution

### Chapter 4

Garcia-Alai MM\*, Heidemann J\*, Skruzny M, Gieras A, Mertens HDT, Svergun DI, Kaksonen M, Uetrecht C, Meijers R (2018) Epsin and Sla2 form assemblies through phospholipid interfaces. *Nature Communications* 9: 328

\*equal contribution

### Chapter 5

Bernal I\*, Börnicke J\*, Heidemann J\*, Svergun DI, Tuukkanen A, Uetrecht C, Kolbe M (2018) Molecular organization of the *Salmonella* type III secretion system sorting platform, *in preparation*

\*equal contribution

---

## Abstract

Non-covalent interactions play a lead role in virtually all biological processes. In a living cell, well-working, often dynamic interactions orchestrate a complex variety of cellular pathways and thereby ensure the cell's healthy state. In order to understand biological processes and potentially prevent or cure diseases, structural studies are performed with a variety of biophysical techniques. Almost three decades after the first pioneering experiments, native mass spectrometry (MS) is still a comparatively new technique that provides insights into stoichiometry, topology and dynamics of interactions. Protein-protein interactions of membranous systems are inherently difficult to characterize in structural studies, as membrane proteins usually require extraction from the protective lipid environment prior to analysis. In this thesis, three protein complexes that interact in different manners with the membrane were studied by native MS.

In a first project, binding of the regulatory protein calmodulin to the plasma membrane  $\text{Ca}^{2+}$ -ATPase ACA8 from *Arabidopsis thaliana* was studied. Native MS revealed binding of two calmodulins to the integral membrane protein ACA8.

In the second project, native MS provided novel insights into the formation of membrane-associated protein-phospholipid complexes involved in clathrin-mediated endocytosis. The specific binding of the phospholipids  $\text{PI}(4,5)\text{P}_2$  to single membrane-binding domains (ENTH, ANTH) of clathrin adaptor proteins was investigated. ENTH and ANTH domains from three different species, including *Homo sapiens*, were demonstrated to assemble to defined complexes with  $\text{PI}(4,5)\text{P}_2$  mediating the interactions. Topology, conformation and assembly dynamics were investigated by native MS and associated structural gas-phase techniques and further complemented with results from other biophysical techniques, e.g. X-ray crystallography and small-angle X-ray scattering (SAXS), to create a comprehensive image of  $\text{PI}(4,5)\text{P}_2$ -dependent assembly of membrane-binding domains of clathrin-adaptor proteins.

The third project exemplifies how studying soluble subcomplexes from large membrane-spanning molecular machines by native MS can produce valuable results on complex composition, topology and assembly pathway. In the present case, the sorting platform of the type III secretion system from *Salmonella* Typhimurium SPI-1 was investigated. Four different proteins (SpaO, SpaOc, OrgB, InvC) and derived truncated protein constructs were studied by native MS to generate a conclusive picture of protein-protein interactions and further map interacting protein domains. The largest complexes comprising of all four mentioned proteins

---

were further analysed by SAXS and a computational model of the molecular architecture of the complex was generated, respecting the restrictions from native MS data.

Summing up, protein complexes from three different membranous settings were studied in this thesis. Results from these studies clearly illustrate that integral membrane protein complexes are most challenging to study. In contrast, sample handling of membrane-associated proteins is markedly facilitated and native MS further proved to be an excellent tool for the investigation of protein-lipid interactions. Studying soluble subcomplexes of membrane proteins is obviously a drastic simplification of the biological system and allows the usage of experimental procedures as for soluble protein complexes. Here, the obtained data revealed novel insights into complex stoichiometry, topology and assembly pathway, illustrating that the simplification can be worthwhile doing.

---

## Zusammenfassung

Nicht-kovalente Wechselwirkungen spielen bei praktisch allen biologischen Vorgängen eine zentrale Rolle. In einer lebenden Zelle arrangieren diese häufig dynamischen Wechselwirkungen eine komplexe Vielfalt an zellulären Funktionen und sichern damit die Gesundheit und Überlebensfähigkeit der Zelle. Strukturbiologische Analysemethoden ermöglichen Einblicke in derartige biologische Abläufe und können damit letztlich auch zur Erforschung von Krankheiten und der Entwicklung neuer Präventionsmethoden und Therapien beitragen. Obwohl erste Experimente bereits vor knapp drei Jahrzehnten durchgeführt wurden, ist die native Massenspektrometrie (MS) noch eine vergleichsweise neue Methode. Sie ermöglicht es, Erkenntnisse über die Stöchiometrie, Topologie und Dynamik von Interaktionen zu erhalten. Protein-Protein Wechselwirkungen im Umfeld der Membran sind besonders schwierig zu untersuchen, da Membranproteine für die Analyse gewöhnlich aus der schützenden Lipidmembran extrahiert werden müssen. In dieser Arbeit wurden drei Proteinkomplexe, die auf verschiedene Weise mit der Membran in Kontakt stehen, analysiert.

Im ersten Projekt wurde die Bindung des regulatorischen Proteins Calmodulin an die Plasmamembran- $\text{Ca}^{2+}$ -ATPase ACA8 von *Arabidopsis thaliana* untersucht. Dabei konnte die Bindung von zwei Calmodulinen an das integrale Membranprotein ACA8 nachgewiesen werden.

Im zweiten Projekt wurden mittels nativer MS neue Einblicke in die Bildung von membranassoziierten Protein-Phospholipid-Komplexen erlangt, die für die Clathrin-vermittelte Endozytose von Bedeutung sind. Die spezifische Bindung des Phospholipids  $\text{PI}(4,5)\text{P}_2$  an membranbindende Domänen (ENTH, ANTH) von Clathrin-Adapterproteinen wurde untersucht. Es konnte für drei verschiedene Spezies, unter anderem *Homo sapiens*, gezeigt werden, dass ENTH- und ANTH-Domänen zu definierten Komplexen assemblieren, in denen  $\text{PI}(4,5)\text{P}_2$  die Interaktion zwischen den Proteinen vermittelt. Topologie, Konformation und Assemblierungsdynamik wurden mithilfe von nativer MS und verwandten Gasphasen-Analysemethoden untersucht. Die Daten wurden durch Ergebnisse anderer biophysikalischer Methoden, beispielsweise Röntgenkristallographie oder Kleinwinkelröntgenstreuung (SAXS) weiter ergänzt, um ein aufschlussreiches Abbild der  $\text{PI}(4,5)\text{P}_2$ -abhängigen Assemblierung membranbindender Domänen der Clathrin-Adapterproteine zu erhalten.

Im dritten Projekt wurde beispielhaft gezeigt, wie die Untersuchung löslicher Teilkomplexe großer membranverankerter Proteinkomplexe aufschlussreiche Erkenntnisse über den

---

Komplexaufbau, die Topologie und den Assemblierungsweg liefern kann. Im vorliegenden Fall wurde die „Sortierungsplattform“ des Type III Sekretionssystems von *Salmonella* Typhimurium SPI-1 untersucht. Vier verschiedene Proteine (SpaO, SpaOc, OrgB, InvC) und verkürzte Proteinvarianten wurden mittels nativer MS untersucht, um ein informatives Abbild der Protein-Protein Interaktionen zu erstellen und weiterhin die für die Interaktionen verantwortlichen Domänen zu identifizieren. Der größte Komplex, bestehend aus allen vier genannten Proteinen, wurde zudem in SAXS Experimenten untersucht. Unter Berücksichtigung der Erkenntnisse aus den nativen MS Messungen wurde ferner ein computerbasiertes Modell der molekularen Struktur des Komplexes erstellt.

Zusammenfassend wurden im Zuge dieser Arbeit drei verschiedene Arten von Proteinkomplexen, die mit Membranen in Verbindung stehen, untersucht. Die Ergebnisse veranschaulichen, dass integrale Membranproteinkomplexe am schwierigsten zu untersuchen sind. Im Vergleich dazu ist die Handhabung von membranassoziierten Komplexen deutlich einfacher, hier erwies sich die native MS auch als geeignete Methode um Protein-Lipid-Wechselwirkungen zu untersuchen. Lösliche Teilkomplexe von membranverankerten Komplexen zu untersuchen ist offensichtlich eine starke Vereinfachung des biologischen Systems, erlaubt es jedoch, experimentelle Techniken wie bei löslichen Proteinkomplexen anzuwenden. Das vorliegende Beispiel veranschaulicht, dass diese Reduktion lohnenswert sein kann, da auf diese Weise neue Einblicke in die Stöchiometrie, Topologie und den Assemblierungsweg des löslichen Teilkomplexes erlangt wurden.

---

# 1 Introduction: Native Mass Spectrometry

## 1.1 Studying Non-Covalent Interactions in Structural Biology

The information about the location of biomolecules at a given time is often not sufficient for a profound understanding of biological processes on a molecular level. It is the interactions and further the dynamics of interactions that determine for instance if a protein is translocated across a membrane or if a ligand is binding to a receptor and thereby triggering a signalling cascade. In infection biology, these interactions determine the capability of a pathogen to infect a new host productively.

Such protein-protein and protein-ligand interactions are mostly non-covalent, meaning that there is no sharing of electrons but electrostatic interactions between molecules or different domains of one molecule. Compared to covalent bonds, these interactions, such as hydrogen bonds, ionic bonds, van-der-Waals forces and hydrophobic effects are all weak in nature (Frieden, 1975). Also, for a functioning process often transient interactions are required that strongly depend on the surrounding environment (Li, Wang et al., 2016).

Research on these non-covalent interactions of soluble and membrane-bound complexes increases the understanding of the respective biological mechanisms and potentially opens routes for new therapies. Nevertheless, studying molecular interactions in an *in vitro* environment is always a simplification of the biological system and results need to be interpreted with caution. Especially when studies target membrane protein complexes, the use of membrane-mimicking systems is a major modification to the physiological state (Shen, Lithgow et al., 2013). However, these artificial systems allow the handling of purified proteins for biophysical characterization and aim at depicting the native state as closely as possible. Soluble and membrane protein complexes are routinely studied by a variety of methods, such as surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), X-ray crystallography, nuclear magnetic resonance spectroscopy (NMR), or fluorescence spectroscopy. Moreover, *in silico* studies gain more importance as more and more reliable prediction tools.

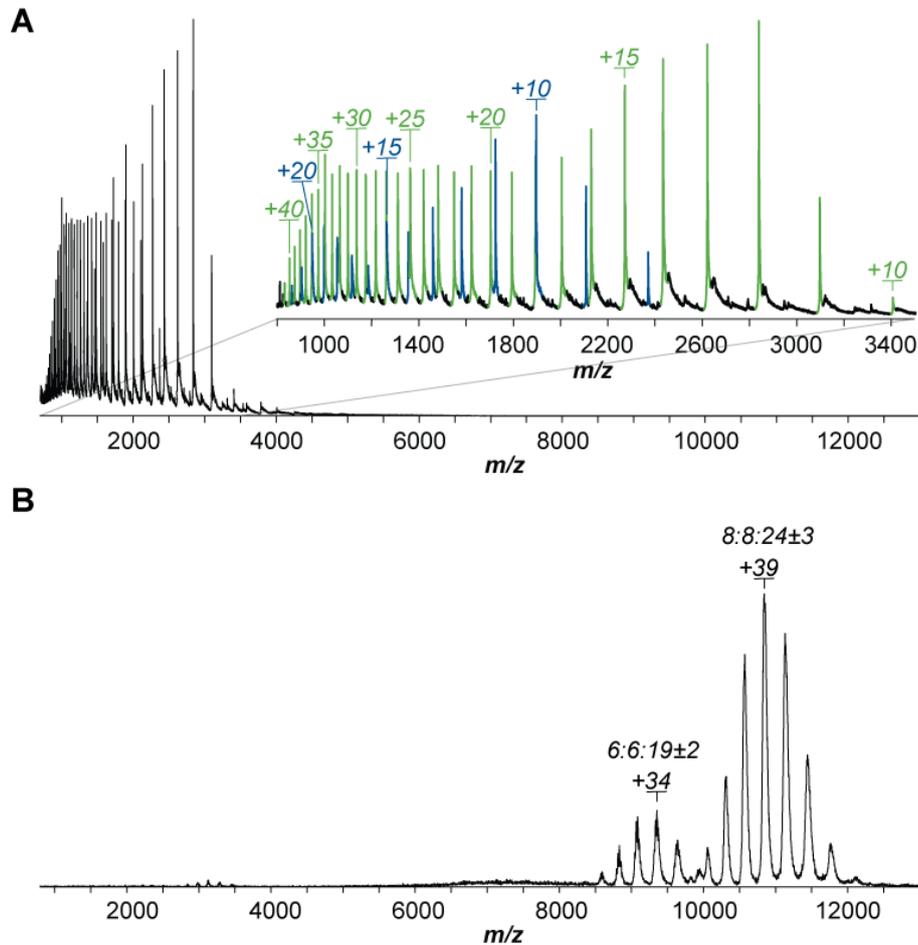
Another method well-suited method for such interaction studies is mass spectrometry (MS). The general MS working principle is remarkably easy: if two molecules interact non-covalently, the nascent complex has the summed mass of the individual components, what can be probed by a mass spectrometer.

## 1.2 General Overview

MS analyses of biological systems are not straightforward and several pitfalls have to be considered (Hernandez & Robinson, 2007). After pioneering research in the early 1990s, almost three decades of instrument and protocol development helped to establish MS in the context of structural biology research. In particular, integrating well-established structural biological methods, such as single particle cryo-electron microscopy, NMR spectroscopy, X-ray crystallography or small-angle X-ray scattering (SAXS) with MS and structural modelling can yield comprehensive characterizations of complex biological systems (Liko, Allison et al., 2016).

Conventional MS studies are often performed under harsh chemical conditions using organic solvents, heat and acids to optimize measuring conditions. These conditions typically allow the most precise analysis of monomeric analytes in MS (**Figure 1A**). In analogy to native polyacrylamide gel electrophoresis, the term *native mass spectrometry* was introduced in 2004 for MS studies that aim at understanding non-covalent interactions (van den Heuvel & Heck, 2004). It is of prime importance in native MS analyses to preserve the folding and the quaternary structure of proteins during sample preparation and measurements. By applying particular protocols and using special instrumental settings, it is possible to retain non-covalent interactions during the measurements and to draw conclusions about the physiological situation from the generated spectra (**Figure 1B**).

The MS measurement itself can be subdivided into three major steps: the ionization of a sample, the separation of the ions according to their mass-to-charge ( $m/z$ ) ratio in the gas phase and finally the detection of the ions. Ideally, the generated spectrum is a snapshot of the solution-phase situation, representing all individual non-covalent complexes and their relative intensities at a given time. This feature is noteworthy, since the output of many other techniques for the study of non-covalent interactions is an averaged signal of all present species in the sample (Sharon, 2010). As the measurement itself is fast and performed on the ms-timescale, also the study of dynamics of processes is possible. Furthermore, the sample consumption is low (< 1 nmol) and a large dynamic mass range is accessible: from a few Da up to the analysis of intact virus capsids (Utrecht, Versluis et al., 2008).



**Figure 1: Mass spectra of membrane-associated domains of epsin-2 and Sla2 from *Saccharomyces cerevisiae*.** **A)** In the presence of 0.5 % formic acid, epsin-2 (blue) and Sla2 (green) were both detected in the low  $m/z$ -range, allowing a precise analysis of the intact mass (epsin-2:  $18.958 \pm 4$  Da, Sla2:  $30.408 \pm 10$  Da). **B)** Native MS measurements in presence of the phospholipid PI(4,5)P<sub>2</sub> revealed two different complexes in the high  $m/z$ -range with the indicated epsin-2:Sla2:PI(4,5)P<sub>2</sub> stoichiometry. Adapted with permission from Springer Customer Service Centre GmbH: Springer Nature, BIOSpektrum, “Native Massenspektrometrie für die Proteinstrukturanalytik” (Heidemann, Krichel et al.), Copyright 2018.

## 1.3 Sample Requirements

### 1.3.1 Protein Production

A high sample quality is of major importance for a successful native MS analysis. In routine applications, proteins are recombinantly produced in organisms that are established in structural biology research. *Escherichia coli* (*E. coli*) lab strains optimized for high protein yields are still the most popular gene expression platforms (Rosano & Ceccarelli, 2014). Specific research questions call for different expression platforms like yeast (*Saccharomyces cerevisiae*,

*Pichia pastoris*) or baculovirus-insect cell systems (Demain & Vaishnav, 2009). These systems return lower protein yields but are capable of adding post-translational modifications (PTMs) like glycosylations to the protein of interest, which might be crucial for protein solubility or relevant in interaction studies. But these modification patterns are less complex than mammalian PTMs (Dell, Galadari et al., 2011). Therefore, it remains desirable to aim for mammalian expression systems like Chinese hamster ovary cells (CHO) or human embryonic kidney cells (HEK)-293 when studying mammalian proteins or proteins that are produced in mammalian cells, like viral proteins. As glycosylation patterns are often species or cell-type specific, expression systems that are as close as possible to the physiological system are preferably used (Goh & Ng, 2017). Disadvantages of mammalian expression systems are lower yields, more complicated handling and higher costs for consumables.

Normally, proteins are purified from the cell lysate using well-established methods (Structural Genomics, China Structural Genomics et al., 2008). Proteins that were produced with a purification tag can be specifically bound to a chromatography resin and subsequently be eluted. For some applications, it is favourable to study the protein of interest without a purification tag, which is an alteration to the physiological system. Especially when protein interactions are studied that are known or suspected to involve the tagged protein terminus, the production of tag-free proteins is required. Highly specific proteases that cleave particular amino acid sequences, like the Tobacco Etch Virus (TEV) protease, Human Rhinovirus 3C protease (PreScission), Factor Xa or thrombin can be utilized to cleave off the affinity tag prior to native MS analysis (Waugh, 2011). Alternatively, the protein of interest is produced without an affinity tag and then separated from other proteins based on its physicochemical properties, for example by ion exchange- or reversed-phase chromatography. Size-exclusion chromatography is routinely used to further purify the protein of interest.

In a recent study it was shown that overexpression of a recombinant gene in *E. coli* and subsequent lysis in highly concentrated ammonium acetate solution is sufficient to analyse proteins by native MS without further purification steps (Gan, Ben-Nissan et al., 2017). This methodology markedly facilitates sample preparation for native MS and is currently being transferred to different expression platforms.

### 1.3.2 Native MS-compatible Buffers

Once the protein of interest has been produced, non-volatile buffer components need to be removed and replaced by a volatile ammonium acetate solution. Although ammonium acetate is often referred to as a buffer, it is mostly used in the physiologically relevant pH regime that is outside of the buffered pH ranges of acetic acid (pH  $4.75 \pm 1$ ) and ammonia (pH  $9.25 \pm 1$ )

(Koneremann, 2017). Nevertheless, the change to ammonium acetate solutions is typically unproblematic and ionic strength and pH can be chosen in a wide range. For proteins that are not suited to be measured from ammonium acetate solutions, ammonium bicarbonate solutions represent an alternative. Although the buffer capacity of ammonium bicarbonate at pH 7 is generally advantageous, care must be taken of CO<sub>2</sub> outgassing that causes foaming and bubble-induced protein unfolding (Hedges, Vahidi et al., 2013).

There are several established methods for the replacement of the protein storage buffer with a native MS-compatible ammonium acetate solution. Ideally, proteins are changed to ammonium acetate solutions during protein purification, for example during size exclusion chromatography. In some cases, for instance when a long-term storage in ammonium acetate is not possible due to protein instability, proteins need to be changed to ammonium acetate solutions directly before the native MS measurement.

Centrifugal filter units are often used for that purpose. The units contain filter membranes that allow small molecules to pass with the buffer solution while large molecules like proteins are retained. As a consequence, non-volatile buffer components are washed out in several iterations of concentration and dilution. Similarly, protein solutions can be dialysed against ammonium acetate solutions until the non-volatile buffer is replaced by the ammonium acetate solution.

Additionally, there are several desalting columns commercially available that either work by centrifugation or gravity flow. In all cases, the column needs to be equilibrated with an ammonium acetate solution before the protein sample of interest is applied. The desalting process itself uses the principle of size exclusion chromatography. Small molecules can enter pores of the resin and are retained, whereas proteins are eluted after the void volume.

Although detrimental for the resolution of native MS analysis, salts and other essential buffer components can be added to the ammonium acetate solution in low amounts. The more non-volatile substances are added, the more adducts are found in the mass spectrum. Yet the addition of up to 5 mM of a non-volatile salt is generally possible. Recent studies link the number of salt adducts in the mass spectrum to the size of the electrospray emitter tip that was used during the measurement. Extra small tips with diameters below 1 µm allow the use of physiological salt concentrations (Susa, Xia et al., 2017).

For the study of membrane proteins in aqueous solutions, detergents are routinely used to solubilize the membrane proteins by covering hydrophobic protein regions. In 2008, the feasibility of transferring detergent-solubilized proteins into the gas-phase was demonstrated,

where collisions with inert gas molecules result in a release of the proteins from the detergent micelles (Barrera, Di Bartolo et al., 2008). The mass of these proteins or even protein complexes can then be determined by native MS. Several solution conditions, such as choice of detergent, detergent concentration and ammonium acetate concentration, but also instrument parameters of the mass spectrometer need to be optimized specifically for membrane proteins (Laganowsky, Reading et al., 2013). Alternatively, several detergent-free methods were shown to be applicable for native MS of membrane proteins, such as nanodiscs, bicelles and amphipols (Hopper, Yu et al., 2013).

Ammonium acetate solutions can further be supplemented with additives that modify the electrospray ionization (ESI) process. While some agents lead to decreased charge states of the analytes (e.g. triethylammonium acetate), others have a supercharging effect (e.g. *m*-nitrobenzyl alcohol) (Iavarone & Williams, 2003, Pagel, Hyung et al., 2010). The addition of acids (e.g. formic acid, acetic acid) and organic solvents (e.g. acetonitrile) to the measuring solution causes protein unfolding and allows a more precise mass determination than for folded proteins (Figure 1A).

Prior to native MS analysis, the protein concentration is usually adjusted to the range of 1-10  $\mu\text{M}$ , depending on the ionization efficiency of the protein. While low protein concentrations generally result in a more stable electrospray, highly concentrated samples often lead to higher signal intensities but are also prone to the formation of non-specific protein clusters (see 1.6.3 Protein-Protein Interactions, page 17).

## 1.4 Nano Electrospray Ionization

A prerequisite for a successful MS analysis is the ionization of analytes and their transfer into the gas phase. Especially for large biomolecules this was a major bottleneck until two revolutionary methods were established: matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI). The significance of these developments was valued in 2002 when Koichi Tanaka and John B. Fenn were awarded with the Nobel Prize in Chemistry.

Koichi Tanaka embedded proteins in a matrix consisting of ultra-fine metal powder and glycerol and ionized the proteins by irradiation with a laser (Tanaka, Waki et al., 1988). Due to its low sensitivity, this *soft laser desorption* technique is currently not used for the analysis of biomolecules. Yet it is very similar to the widely used MALDI method that was introduced by Franz Hillenkamp, Michael Karas and coworkers (Karas, Bachmann et al., 1987, Karas & Hillenkamp, 1988). As the name indicates, in MALDI the analyte is embedded in a crystalline matrix and then irradiated with a laser. The laser energy is absorbed by the matrix and is

suspected to be responsible for the proton transfer to the analyte. The precise mechanism is still under debate with thermal proton transfer models currently being favoured (Lu, Lee et al., 2015). Although MALDI proved to be a very robust and straightforward ionization technique that tolerates relatively high salt concentrations and allows easy interpretation of spectra thanks to the low number of charges, it is not the method of choice for native mass spectrometry. Already the contact to the matrix, which contains acids (e.g. sinapinic acid or trifluoroacetic acid,  $\text{pH} < 2$ ) and organic solvents, causes dissociation of non-covalent complexes. Even so, some studies demonstrated that measuring non-covalent complexes with MALDI MS is in principle possible (Chen, Gulbakan et al., 2016).

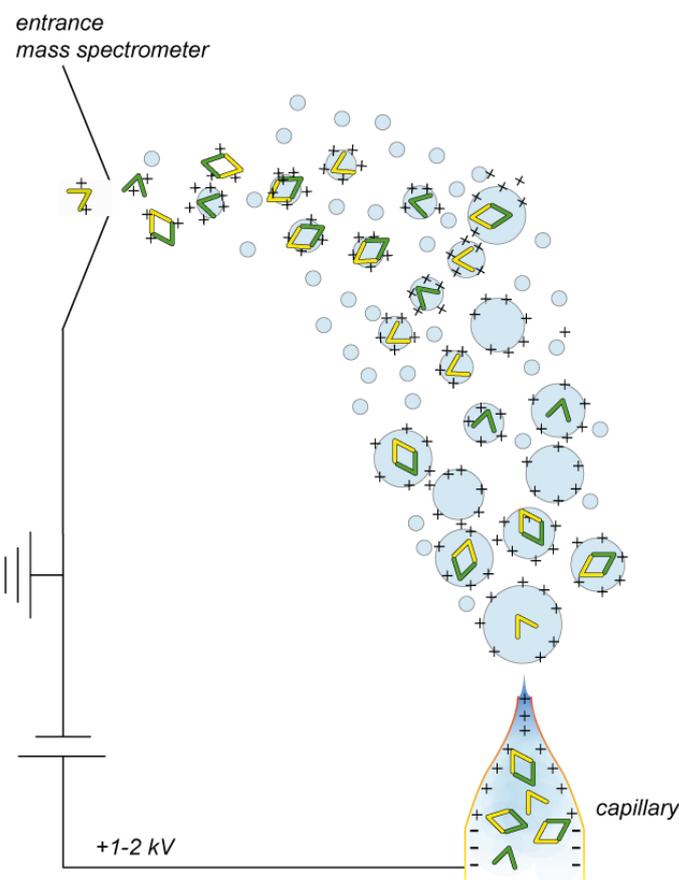
Nonetheless, ESI is generally considered a softer ionization technique. In 1968, Malcolm Dole and colleagues first coupled an electrospray to a mass spectrometer (Dole, Mack et al., 1968). But it was John B. Fenn who markedly improved the ESI source and made the technique available for the study of large biomolecules (Fenn, Mann et al., 1989, Yamashita & Fenn, 1984).

In ESI, the sample is loaded into a capillary to which a high voltage (typically 1-4 kV) is applied (Figure 2). Even though both polarities are possible, for the study of proteins normally a positive potential compared to ground potential is used. In this positive ion mode, protons are generated at the capillary tip and cations are enriched at the liquid surface (Van Berkel & Kertesz, 2007). The interplay of surface tension and coulomb forces leads to the formation of the so-called Taylor cone and the emission of charged droplets (Wilm & Mann, 1994). Here, only nano ESI will be discussed as it proved to be advantageous for the analysis of large biomolecules and is therefore most widely used in native MS (Wilm & Mann, 1994). In comparison to conventional ESI, the dimension of the inner capillary diameter is reduced from 100  $\mu\text{m}$  to the low  $\mu\text{m}$ -range. Instead of stainless steel emitters, glass capillaries coated with a conductive surface or equipped with a platinum wire that enters the sample solution are used. The size reduction does not only come along with a drastically lower sample flow rate (ESI: 1-10  $\mu\text{l}/\text{min}$ , nano ESI:  $< 20 \text{ nl}/\text{min}$ ) and sample consumption (Karas, Bahr et al., 2000). It was also shown that the ionization efficiency and the tolerance towards the presence of non-volatile buffer molecules is increased (Wilm & Mann, 1996).

During the nano ESI process, droplets with diameters of 200 nm or less are emitted from the Taylor cone (Juraschek, Dülcks et al., 1999). Solvent evaporation in the aerosol induces shrinking of droplets and increased charge densities. The Rayleigh limit describes the maximum number of charges that a droplet can contain before the Coulomb repulsion is greater than the surface tension (Rayleigh, 1882). Under this condition, the droplet undergoes a fission event that leads to the formation of smaller droplets:

$$z_R = \frac{8\pi}{e} \sqrt{\epsilon_0 \gamma R^3} \quad (1)$$

where  $z_R$  is the droplet charge of elementary charges  $e$  at a droplet radius  $R$ ,  $\epsilon_0$  is the vacuum permittivity and  $\gamma$  is the surface tension.



**Figure 2: Schematic representation of a nano electrospray ionization (ESI) source in positive ion mode.** A sample containing two proteins (green, yellow triangles) is loaded into a gold-coated glass capillary to which a high voltage is applied. From the Taylor cone at the tip of the capillary small droplets are emitted that shrink and undergo Coulomb explosions on their way to the entrance of the mass spectrometer. Positively charged protein ions, here monomers and dimers as present in solution, enter the instrument for  $m/z$  analysis. For better visibility, size proportions are not to scale. Adapted with permission from Springer Customer Service Centre GmbH: Springer Nature, BIOspektrum, “Native Massenspektrometrie für die Proteinstrukturanalytik” (Heidemann, Krichel et al.), Copyright 2018.

It is assumed that the formation of gaseous ions from these nanodroplets follows one of three different ionization mechanisms, depending on the nature of the analyte (Konermann, Ahadi et al., 2013). Small protonated analytes are suspected to be ejected from the nanodroplet (ion evaporation model). Globular folded proteins and protein complexes are proposed to become

gaseous ions by the charge residue model. Here, charges are transferred to the protein while nanodroplets containing the analyte dry out (Kebarle & Verkerk, 2009, Winger, Light-Wahl et al., 1993). Molecular dynamics simulations found a third mechanism that might apply to partially hydrophobic polymer chains like unfolded proteins. In the chain ejection model the analyte moves to the surface of the droplet, where one terminus is expelled to the surrounding air and then successively ejected from the droplet (Ahadi & Konermann, 2012).

In contrast to MALDI ionization, ionization of large analytes by ESI leads to the formation of multiply charged ions. Despite the less straightforward assignment of peaks and deduction of masses, the generation of highly charged ions has also benefits. For instance, the required  $m/z$ -range for the analysis of extremely large ions becomes considerably smaller, leading to a generally unlimited mass range of ESI MS. Moreover, an analyte of a certain mass can acquire different numbers of charges during ionization, resulting in characteristic ESI MS peak series (charge envelopes). From the relation of these ions in mass and charge, the mass of the neutral analyte  $M$  can be determined (Covey, Bonner et al., 1988). For two neighbouring peaks  $\left(\frac{m_1}{z_1}\right)$  and  $\left(\frac{m_2}{z_2}\right)$  from one charge envelope it is known that:

$$z_1 = z_2 + 1 \quad (2)$$

Assuming that all charges exclusively originate from the addition of protons with a mass  $m_H = 1$ , it follows that:

$$\left(\frac{m_2}{z_2}\right) = \frac{M + z_2}{z_2} \quad (3)$$

$$\left(\frac{m_1}{z_1}\right) = \frac{M + (z_2 + 1)}{z_2 + 1} \quad (4)$$

The number of charges is given by:

$$z_2 = \frac{\left(\frac{m_1}{z_1}\right) - 1}{\left(\frac{m_2}{z_2}\right) - \left(\frac{m_1}{z_1}\right)} \quad (5)$$

Considering equation (2), the molecular mass of the uncharged analyte is determined by:

$$M = z_2 \left( \left(\frac{m_2}{z_2}\right) - 1 \right) \quad (6)$$

If the analyte carries not only protons but also adducts of sodium, potassium or ammonia, the mass of the neutral analyte can be determined under consideration of the mass of the respective adduct. In case of sufficient spectral resolution, the number of charges and the mass of an ion can be determined from the isotope pattern of a single charge state.

## 1.5 Mass Analysers

Different methods to analyse the  $m/z$  ratio of gaseous ions were developed that all use the manoeuvrability of ions in the gas phase. In order to avoid collisions with ambient molecules that can lead to unwanted chemical reactions or ion losses, mass spectrometers are generally constructed with a high vacuum that can be as low as  $1.3 \times 10^{-10}$  mbar in modern instruments (Tolmachev, Robinson et al., 2009). However, especially for the analysis of large ions slightly increased pressures have a favourable focussing effect (Chernushevich & Thomson, 2004, Tahallah, Pinkse et al., 2001) (see 1.9 High-mass Nano ESI QToF Instruments, page 24). Furthermore, inert gases are used in a controlled manner for studies of the composition (see 1.7 Gas-phase Dissociation and Fragmentation Techniques, page 20) or conformation (see 1.8 Ion Mobility Mass Spectrometry, page 21).

An effective separation in a mass analyser enables the distinction of two ions with different  $m/z$ . One quality measure of an obtained mass spectrum is the resolution. In mass spectrometry, the resolution  $R$  is dimensionless and defined as:

$$R = \frac{M}{\Delta M} = \frac{\frac{m}{z}}{\Delta \frac{m}{z}} \quad (7)$$

It describes the capability to resolve two adjacent peaks or, if applied to a single peak, the peak width at a certain relative peak height. Throughout this work, according to IUPAC recommendations the peak width at 50 percent of the total peak height is used for  $\Delta m/z$  (Murray, Boyd et al., 2013).

Lately, due to technical developments high- and ultra-high resolution MS became applicable for native MS studies. As a common feature, these mass analysers detect image currents induced by the analytes, that are converted into a mass spectrum by Fourier transform operations (Scigelova, Hornshaw et al., 2011).

Instruments containing orbitrap mass analysers were modified for efficient transmission of high- $m/z$  ions and used for studies of large biomolecular assemblies such as ribosomal particles or an RNA-containing virus (Rose, Damoc et al., 2012, van de Waterbeemd, Fort et al., 2017).

Fourier-transform ion cyclotron resonance (FTICR) mass analysers offer highest resolving power for small analytes. Although there are concerns regarding the transmission or trapping efficiency as well as a possible degrading effect on resolution at low cyclotron frequencies during analysis of high- $m/z$  ions, the method was shown to work for protein complexes of up to 1.9 MDa (Li, Nguyen et al., 2018). In particular, the combination with protein fragmentation techniques and high-resolution analysis of the occurring fragment ions might prove useful.

Below, the two types of mass analysers used in this work, time-of-flight and quadrupole mass analysers, are discussed in more detail.

### 1.5.1 *Time-of-Flight Mass Analysers*

At the frontend of a time-of-flight (ToF) mass analyser a continuous beam of analyte ions from the ESI source is converted into ion packages. Technically, this conversion is typically achieved by a pulsed voltage that is orthogonal to the flight path of the entering ions. For all ions in one of these packages the  $m/z$  ratios are analysed simultaneously by a spatial separation according to mass and charge (Chernushevich, Loboda et al., 2001).

For that reason, ions are accelerated by an electric field, leading to the conversion of potential energy into kinetic energy:

$$E_{kin} = E_{pot} \quad (8)$$

$$zeV = \frac{1}{2}mv^2 \quad (9)$$

where  $z$  is the number of elementary charges  $e$ ,  $V$  is the applied acceleration voltage,  $m$  and  $v$  are mass and velocity of the ion, respectively. Ions then move with this kinetic energy through a field-free tube. The velocity of the ions can be determined by the time that it takes the ions to pass the drift tube with a length  $d$ :

$$v = \frac{d}{t} \quad (10)$$

Substituting (10) into (9) leads to:

$$zeV = \frac{1}{2}m\left(\frac{d}{t}\right)^2 \quad (11)$$

And rearranging the equation to express the time:

$$t = \frac{d}{\sqrt{2eV}} \sqrt{\frac{m}{z}} \quad (12)$$

Since  $d$ ,  $e$  and  $V$  are constants or instrument parameters, the time that it takes for an ion to pass the flight tube is dependent on the characteristics of the analyte ion  $\sqrt{\frac{m}{z}}$ .

Separation of ions ( $\Delta t$ ) with different  $m/z$  improves with increased path length  $d$ . The introduction of an ion mirror, called reflectron, allows elongated path lengths in compact instrument setups (Mamyrin, Karataev et al., 1973). Even more beneficial for separation and consequently spectral resolution is a compensating effect on initial energy distributions of ions of the same kind (Mamyrin, Karataev et al., 1973). Faster ions penetrate deeper into the electric field of the reflectron than slow ions and consequently have slightly increased path lengths until they reach the detector.

Multi-channel plates (MCPs) are typically used for detection. These plates are made from highly resistive material and contain a dense array of microchannels spanning from one side to the other. During the measurements a high voltage (in the range of 2 kV) is applied across the plate, creating an electron multiplier. Once an ion hits the walls of a microchannel, a signal-amplifying cascade is initiated before a time-to-digital converter records the electric signal. A very high temporal resolution, which can be in the low ps range, allows the precise determination of flight times. After conversion of flight times to  $m/z$  ratios, mass spectra are generated illustrating the relative abundance of ions as a function of the  $m/z$  ratio.

### 1.5.2 Quadrupole Mass Analysers

Quadrupole mass analysers are often used upstream of ToF mass analysers, resulting in QToF instrument setups. As the name already indicates, the quadrupole mass analyser consists of four parallel metal rods. Both of the two opposing rod pairs are electrically connected and a static voltage as well as a radiofrequency voltage are applied. The resulting electric fields have a filtering function for incoming ions. Depending on the  $m/z$  ratio, ions have either stable trajectories through the quadrupole or collide with the metal rods (Chernushevich, Loboda et al., 2001).

Quadrupoles are used either in transmission mode or in mass filter mode. In transmission mode, magnitudes of radiofrequency and static voltages are varied while their ratio is kept constant (Paul, 1990). Consequently, ions with different  $m/z$  ratios pass the quadrupole consecutively. This mode is used in standard MS analyses in QToF instruments when all the

molecules from solution are of interest. The mass determination in these setups takes place in the ToF mass analyser.

In mass filtering mode the voltages are not varied and hence only ions with a defined window of  $m/z$  ratios proceed to the succeeding mass analysis. This is particularly beneficial in MS/MS analyses when only a defined ion species, so-called precursor ions, is selected in the quadrupole and then subjected to gas-phase manipulations. Dissociation or fragmentation products, also known as daughter or product ions, can directly be associated to the selected precursor ions (see 1.7 Gas-phase Dissociation and Fragmentation Techniques, page 20).

## 1.6 Studying Non-Covalent Assemblies in the Gas Phase

### 1.6.1 *Solution Structures versus Gas-phase Structures*

Brian T. Chait and colleagues found in 1990 that the number of charge states of a protein analysed by ESI MS is dependent on the protein conformation in solution. Thus, MS was for the first time used as a probe for protein structures (Chowdhury, Katta et al., 1990). The number of charges that proteins carry after the electrospray ionization process is heavily dependent on the protein surface, whereas the number of basic amino acid residues is of less importance (Kaltashov & Mohimen, 2005). A large protein surface causes a high number of charges. The precise relationship between the protein surface and the number of charges was thoroughly investigated (Fernandez De La Mora, 2000, Kaltashov & Mohimen, 2005). It was further shown that the number of charges can be predicted from the Rayleigh limit (Equation (1)), assuming that the radius of a protein or protein complex is directly correlated to the molecular weight of the protein and that its density is similar to the density of water:

$$z_R = 0.078 \sqrt{M} \quad (13)$$

where  $M$  is the molecular weight of the protein or protein complex in Dalton (Heck & Van Den Heuvel, 2004).

But to what extent does the gas-phase structure that is probed by native MS represent the solution phase conformation of a protein? Several studies addressed this fundamental question for native MS using different techniques, like infrared and fluorescent spectroscopy, and confirmed that structural properties from solution are maintained in the gas phase (Iavarone & Parks, 2005, Oomens, Polfer et al., 2005). In ion-mobility mass spectrometry protein conformations are probed in addition to the mass of the analyte (see 1.8 Ion Mobility Mass Spectrometry, page 21). Many pieces of evidence from this technique also confirm the

structural similarity of dehydrated protein ions and proteins in solution (Scarff, Thalassinou et al., 2008, Wytenbach & Bowers, 2011). According to molecular dynamics simulations, charged amino acid side chains form salt bridges between each other within picoseconds in the gas phase and as the consequence the protein ions are even transiently stabilized (Breuker & McLafferty, 2008). Yet due to the loss of the aqueous environment during ionization, hydrophobic interactions become unstable while electrostatic interactions are strengthened. Thus, solvent-phase conformations turn out to be thermodynamically unstable in the gas phase (Breuker & McLafferty, 2008). But it was shown that even these conformations that are energetically less favourable in the gas phase are kinetically trapped over the millisecond time span of MS measurements (Silveira, Fort et al., 2013).

The first protein-ligand interactions analysed by mass spectrometry were macrolides binding to the FK binding protein and the myoglobin-heme interaction (Ganem, Li et al., 1991, Katta & Chait, 1991). Shortly afterwards, also protein-protein interactions were proven to be accessible to MS as shown by the interaction of a dimer of the HIV-1 protease with an inhibitor (Baca & Kent, 1992). Today, native MS is an established method for the investigation of protein-ligand and protein-protein interactions (Leney & Heck, 2017, Mehmood, Allison et al., 2015).

Based on the pioneering research, a variety of methods was developed for the structural analysis of proteins in the gas phase. For instance, MS proved to be a useful tool to probe the local hydrogen-deuterium exchange in proteins (Zhang & Smith, 1993). Alternative techniques for the observation of hydrogen-deuterium exchange in proteins and peptides, such as infrared spectroscopy, usually generate a global readout (Heidemann & Bernhardt, 1967). Additional techniques for the MS-based structural analysis of proteins involve for instance the cross-linking of amino acids to produce distance information or the covalent labelling of amino acids to depict the solvent accessibility (Mendoza & Vachet, 2009, Sinz, 2006).

### *1.6.2 Protein-Ligand interactions*

The verification of a specific protein-ligand interaction is often of great value. For a more detailed understanding a precise characterization is needed, such as the number of ligand binding sites and the strengths of the interactions. Nano ESI MS allows a thorough characterization of protein-ligand interactions (Kitova, El-Hawiet et al., 2012).

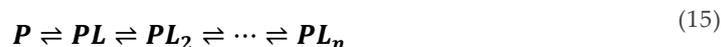
Counting the number of ligand molecules attached to a protein is generally straightforward, since every bound ligand molecule adds its mass to the mass of the protein. However, during the ESI process unbound ligands are by chance emitted in one droplet with a protein. Upon solvent evaporation, the ligand can randomly cluster to the protein surface, thereby generating

an ion that has the same mass as a specifically bound protein-ligand pair. High ligand concentrations that are required for the study of low-affinity bindings increase the chances of clustering. The usage of a reference protein was introduced to discriminate between specific protein-ligand binding events and non-specific ESI clustering effects (Sun, Kitova et al., 2006). Here, a reference protein, which is assumed to not specifically interact with the ligand, is added to the solution. The extent of non-specific clustering is identical for the reference protein and the protein of interest and consequently can be subtracted from the observed protein-ligand complexes to determine the precise ligand numbers.

For analysing the interaction strength, often expressed as the dissociation constant  $K_D$ , the assessment of the ratio between unbound protein (P), ligand (L) and protein-ligand (PL) complexes is required (Equation (14)). Conclusions from these ratios are only valid if ionization efficiencies of unbound and ligand-bound proteins are alike. But as the addition of a small ligand to a large protein barely changes the properties of the emerging particle, ionization efficiencies are typically assumed to be identical when studying protein-ligand interactions (Jecklin, Touboul et al., 2008). Hence, after correction of the unspecific clustering, the ratios found in a mass spectrum indeed depict the situation in solution. As input protein and ligand concentrations are known and the ratio of all species can be analysed simultaneously, this information can be directly fed into the law of mass action:

$$K_D = \frac{[P] \times [L]}{[PL]} \quad (14)$$

For systems covering more than one binding event, the situation is more complicated. But still, all species are observed in the mass spectrum and correction of unspecific clustering should disclose ratios as present in solution:

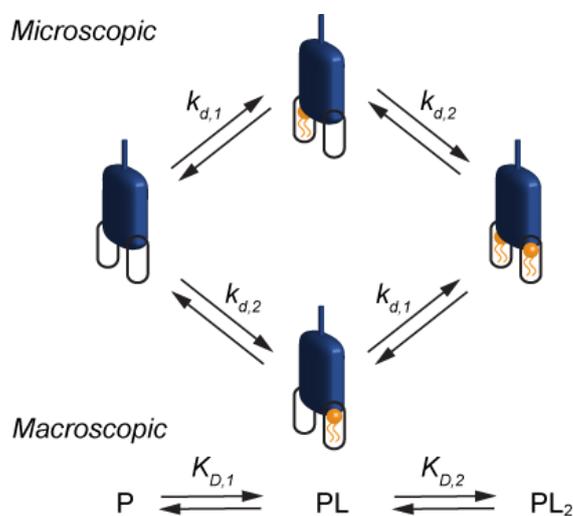


Binding affinities in these systems are determined by (El-Hawiet, Kitova et al., 2012):

$$K_{D,n} = \frac{R_{n-1} \left( [L_0] - \frac{(R_1 + 2R_2 + \dots + nR_n)[P_0]}{1 + R_1 + R_2 + \dots + R_n} \right)}{R_n} \quad (16)$$

Where  $[L_0]$  and  $[P_0]$  are initial ligand protein concentration, respectively, and  $R_n$  terms are the abundance ratios of ligand-bound to free protein ions.

If there are multiple binding sites for one ligand, native MS cannot distinguish which of the binding sites are occupied. The determined dissociation constants are macroscopic constants, summing up all microscopic binding events (Figure 3).



**Figure 3: Macroscopic and microscopic binding constants.** A protein of interest (blue) contains two binding sites (rounded rectangles) for a ligand (orange). Native MS cannot discriminate between different microscopic occupancies. Just the macroscopic states (P, PL, PL<sub>2</sub>) and macroscopic dissociation constants are accessible. Reprinted from "Epsin and Sla2 form assemblies through phospholipid interfaces" by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.

However, there is a mathematical relation of microscopic and macroscopic constants. Provided the binding sites are independent, meaning a ligand binding to one site does not alter the ligand's affinity to the other site, macroscopic and microscopic constants are related by (Klotz, 1985):

$$K_{D,1} = \frac{k_{d,1} \times k_{d,2}}{k_{d,1} + k_{d,2}} \quad (17)$$

$$K_{D,2} = k_{d,1} + k_{d,2} \quad (18)$$

The observed signal intensities of single ion species in native MS measurements are also heavily dependent on the measuring conditions, in particular on the used nano electrospray capillaries. Variable signal intensities correlate with altered ratios of free protein to ligand-bound protein. Consequently, quantitative data interpretations of ligand binding events need to be performed with caution and with data from multiple measurements.

More elaborate and more precise methods for determining ligand affinities using native MS were developed, ranging from the use of reference ligands, over ligand titrations to ligand competition assays, as reviewed by Boeri Erba & Zenobi (2011).

However, ESI MS assays are preferably used to screen small molecule libraries for binders and potential inhibitors of target proteins (Hofstadler & Sannes-Lowery, 2006, Kitova, El-Hawiet et al., 2012). For the thorough characterization of ligand binding affinity or kinetics, different methods are more widely used than native MS. In ITC measurements, a ligand is titrated into a protein containing sample cell and the temperature change upon binding is compared to the temperature change in a reference cell, where no protein is present. Multiple binding parameters, for instance stoichiometry, association constant and binding enthalpy can be derived from the data (Freyer & Lewis, 2008). Long acquisition times and a high sample consumption are generally considered to be the main disadvantages of ITC measurements, making high-throughput applications difficult (Torres, Recht et al., 2010).

Surface plasmon resonance spectroscopy is also popular for the characterization of protein-ligand interactions. Here, a ligand is immobilized on a metal-surface before the protein of interest is added. Polarized light is guided through a prism to the surface. At a critical angle of the incoming light, electron density waves, so-called surface plasmons, are generated that resonate with the light and the reflected light beam contains a characteristic intensity dip (Szabo, Stolz et al., 1995). Binding of the protein to the immobilized ligand causes a change of the refractive index and the critical angle, which is monitored by measuring the change of the resonance angle or the resonance wavelength (Patching, 2014). Expensive sensor chips and the requirement of sample immobilization are generally considered the main disadvantages of SPR studies.

Further routine methods to investigate protein-ligand interactions are for instance radiometric ligand-binding assays (Hulme & Trevethick, 2010), microscale thermophoresis (Wienken, Baaske et al., 2010), or saturation-transfer-difference NMR spectroscopy (Meyer & Peters, 2003).

### *1.6.3 Protein-Protein Interactions*

Different proteins and protein complexes have different physicochemical properties. As a consequence, they also have different ionization efficiencies and signal intensities cannot be translated into solution-phase ratios (Hermans, Ongay et al., 2017). Methods were introduced that involve thorough quantification controls or account for different ionization and MS

transmission efficiencies by considering a response factor, but their application was thus far limited to homo-oligomers (Boeri Erba, Barylyuk et al., 2011, Rose, Labrijn et al., 2011).

Due to the additional obstacles of a quantitative analysis, native MS is preferably used for a qualitative investigation of protein-protein interactions. Yet still, high protein concentrations potentially cause clustering effects during the ESI process, which can be mistaken for specific interactions.

For a time-resolved characterization of pre-steady state interactions, limited time scales are accessible. Theta-glass capillaries allow mixing of interaction partners directly before the ESI, resulting in a mixing time in the  $\mu\text{s}$ -range (Mortensen & Williams, 2014). Pre-mixing the components before loading them into the ESI capillary typically takes a few tens of seconds until data acquisition starts. The particularly interesting time span of the first milliseconds to seconds post mixing is dynamically accessible using special instrumentation (Cong, Katipamula et al., 2016, Zinck, Stark et al., 2014).

Depending on the raised research question, different biochemical or biophysical methods are available to either verify or supplement native MS results. For a qualitative proof of a suspected protein-protein interaction without any further information on stoichiometry or topology, co-immunoprecipitation coupled to western blot analysis is routinely used (Berggård, Linse et al., 2007). If putative interaction partners need to be identified first, untargeted approaches must be applied, as for instance quantitative co-affinity-purification MS or two-hybrid screenings. Especially the latter technique was reported to produce a high number of false positive and false negative results (Huang, Jedynak et al., 2007).

The use of biophysical methods allows a detailed characterization of interactions. By producing high-resolution molecular structures, X-ray crystallography, NMR spectroscopy and lately also cryo-electron microscopy are unprecedented to picture the molecular mechanism of a protein-protein interaction.

By coupling multi-angle light scattering (MALS) to size-exclusion chromatography (SEC) the molecular masses of the analytes are determined directly after elution from the chromatography column. The sample is illuminated with a monochromatic, polarized, collimated light beam that induces an oscillating dipole and anisotropic scattering in macromolecules. The angular dependence of scattering macromolecules are then used to determine the radius of gyration  $R_g$  and the molecular mass (Wyatt, 1993). In contrast, dynamic light scattering instruments measure time-dependent fluctuations of the scattered light signal

from which the hydrodynamic radius of the molecules is calculated (Stetefeld, McKenna et al., 2016).

Similar information, such as  $R_g$  and the molecular weight of the analyte is accessible by small-angle X-ray scattering analysis (Putnam, Hammel et al., 2007). As indicated by the name, samples are irradiated by X-rays and the elastic scattering in small angles is recorded. From plots that illustrate the spatially averaged scattering intensities versus the scattering vector  $q$  various sample specifications can be extracted. Scattering in the low  $q$ -range is determined by the size of the sample, the intermediate  $q$ -range by the shape of the structure and the high  $q$ -range by the surface-to-volume ratio (Boldon, Laliberte et al., 2015). Three-dimensional shapes are reconstructed by different modelling approaches, such as bead modelling or envelope function modelling (Volkov & Svergun, 2003).

As described for the analysis of protein-ligand interactions, SPR and ITC can be used for the characterization of protein-protein interactions.

Aforementioned methods qualify for an *in vitro* characterization of protein-protein interactions. Though being more difficult to obtain, ultimately *in vivo* data are wanted to highlight the biological relevance of the findings. Thanks to immense technological developments, electron cryo tomography (cryo-ET) is generally capable of studying protein complexes at near-atomic resolution (Galaz-Montoya & Ludtke, 2017). Sample preparation protocols for such *in situ* experiments do not involve purification steps and targets are analysed in their native environment (Asano, Engel et al., 2016).

Microscopic methods also provide access to *in vivo* data. In the popular FRET (Fluorescence resonance energy transfer) assay a donor fluorophore is excited by a light source. When a second chromophore with an absorbance spectrum overlapping with the donor's emission spectrum is in local proximity, a non-radiative energy transfer from one fluorophore to the other is possible (Piston & Kremers, 2007). As the efficiency of the energy is strongly dependent on the distance of the chromophores, distance information is encoded in the obtained fluorescence signals. Main drawbacks in FRET experiments are the usage of fluorescence tags that are sensitive to the local environments and low signal-to-noise ratios (Leavesley & Rich, 2016).

## 1.7 Gas-phase Dissociation and Fragmentation Techniques

The dissection of ions in the gas phase and the subsequent analysis of the emerging fragments increases the information content of native MS studies. Therefore, several methods were developed, of which only the ones relevant for this work are discussed in this section.

In collision-induced dissociation (CID) measurements, the precursor ions of interest are isolated in the quadrupole and then guided into a gas-filled chamber where collisions with inert gas molecules, such as argon or xenon take place (Benesch, 2009, Cooks, 1995). The number of collisions is controlled by the length of the collision cell and the pressure of the collision gas. The acceleration voltage that ions experience upon entering the collision cell and the number of charges of the ions define the ions' velocity and thereby the energy of the collisions. The impact of the accumulated collisions of protein complex ions on their way through the collision cell can result in different effects, depending on the total energy that has been deposited on the analyte.

If the amount of energy is comparatively low, the determined masses of the analytes are slightly reduced. This reduction is attributed to the loss of remaining non-volatile salt molecules or residual solvent molecules that attach to the analytes (Smith, Loo et al., 1990).

The application of increased activation energies to protein complexes leads to the dissociation of single protein subunits. Small complex components from the complex periphery dissociate preferentially and take a disproportionately large fraction of the protons along (Benesch, 2009, Jurchen & Williams, 2003). As a result, dissociated subunits have a low  $m/z$  ratio, while the residual complexes are found in the higher  $m/z$ -range. According to the currently prevailing opinion, the dissociation is caused by an unfolding event of a protein monomer (Ruotolo, Hyung et al., 2007). Proton migration to the unfolding protein leads to increased Coulombic repulsion between subunits until the unfolded protein subunit is ejected (Jurchen & Williams, 2003, Popa, Trecroce et al., 2016). However, in a limited number of cases an atypical dissociation behaviour without unfolding was reported (Boeri Erba, Ruotolo et al., 2010). This behaviour was later attributed to a low subunit flexibility, higher charge densities and small interfaces (Hall, Hernandez et al., 2013). Recently, Loo & Loo (2016) suggested an alternative dissociation mechanism involving the heterolytic scission and rearrangement of salt bridges.

At increased collision energies, also another energy-dependent effect occurs: fragmentations of the protein backbone. Typically, peptides that are generated in CID are cleaved at the peptide bond and named, in accordance to the Roepstorff nomenclature scheme, as  $b$ -ions (N-terminus) or  $y$ -ions (C-terminus) (Roepstorff & Fohlman, 1984). It was demonstrated that

cleavages primarily occur in protein areas that are unfolded beforehand (Han, Jin et al., 2006). A comparison of fragmentation pathways in native and denatured proteins further revealed a higher sequence specificity and a preference for surface exposed residues for native proteins (Haverland, Skinner et al., 2017). Membrane proteins tend to fragment at membrane-spanning  $\alpha$ -helices (Skinner, Catherman et al., 2014).

In surface-induced dissociation (SID) the analytes experience a collision with a surface. Hence, the deposited energy is not sequentially built up as in CID, but transferred in a single collision event, leading to subunit ejection with minimal unfolding (Zhou, Dagan et al., 2012). Charges are distributed over the dissociation products more symmetrically (Jones, Beardsley et al., 2006). Furthermore, additional structural information is preserved in the dissociation products, since complexes dissociate into logical subentities, such as a dimer of heptamers into two heptamers (Zhou, Jones et al., 2013).

## 1.8 Ion Mobility Mass Spectrometry

Structural information of proteins and protein complexes is accessible by means of ion mobility mass spectrometry (IMMS) (Uetrecht, Rose et al., 2010). In IMMS measurements, analyte ions are guided through a gas-filled chamber and are separated according to their shape, before their  $m/z$  ratio is determined. Thereby, an extra dimension of information is added to the mass analysis. In the most basic design, the gas-filled chamber consists of a linear drift tube that the protein ions pass in pulsed packages under the influence of a static electric field (Hoaglund, Valentine et al., 1998). The velocity of the ions  $v$  in the drift tube and thus also the travelling time  $t_d$  through the tube with the length  $d$ , is dependent on the strength of the electric field  $E$  and the mobility  $K$  of the ion (Mesleh, Hunter et al., 1996):

$$v = KE = \frac{d}{t_d} \quad (19)$$

To account for experimental parameters, such as temperature-dependent changes of the buffer gas pressure, the reduced mobility  $K_0$  is expressed by:

$$K_0 = \frac{d}{t_d E} \frac{P}{760 \text{ Torr}} \frac{273.2 \text{ K}}{T} \quad (20)$$

where  $P$  is the pressure (in Torr) and  $T$  is the temperature (in Kelvin). The mobility of gaseous ions in low electric fields is described by the Mason-Schamp equation (Mason & Schamp Jr, 1958). Combination with equation (20) results in:

$$K_0 = \frac{3ze}{16N\Omega} \sqrt{\frac{2\pi}{\mu k_B T}} \frac{P}{760 \text{ Torr}} \frac{273.2 \text{ K}}{T} \quad (21)$$

where  $z$  is the number of elementary charges  $e$  is the charge of the ion,  $N$  is the number density of the drift gas (number of molecules per unit volume),  $\mu$  is the reduced mass that is calculated from the ion's mass ( $M$ ) and the mass of the drift gas ( $m$ ) ( $\mu = \frac{mM}{(m+M)}$ ),  $k_B$  is the Boltzmann constant and  $\Omega$  is the average collision cross section (CCS) of the ion. The latter is the orientation-averaged cross section of the ion and hence directly depending on the analyte's conformation. Recorded drift times and derived CCSs are a valuable readout for the analysis of protein conformations and protein stabilities. For instance, unfolding of the analyte upon collisions with gas molecules (see 1.7 Gas-phase Dissociation and Fragmentation Techniques, page 20) is studied and illustrated in collision-induced unfolding plots (Dixit, Polasky et al., 2018).

A different ion mobility working principle was developed to market maturity by the company Waters (Pringle, Giles et al., 2007). In travelling wave ion mobility mass spectrometry (TWIMS) instruments of the model series *Synapt*, stacked ring-shaped electrodes generate waves of electric fields that push the ions against a counterflow of drift gas towards the exit of the ion mobility cell. Depending on the analyte's mobility and height and velocity of the waves, rollover events over the wave tops occur that lead to a prolonged transit time and consequently to a separation of ions with different mobilities (Shvartsburg & Smith, 2008). Even though these instruments offer high transmission efficiencies, the CCS is thus far not directly accessible. Instead, protein standards with known collision CCSs need to be analysed under identical measuring conditions. Comparison with drift times of the calibrant ions then allows the estimation of the CCS of the protein or protein-complexes of interest (Bush, Hall et al., 2010, Ruotolo, Benesch et al., 2008).

In short, literature CCS values are corrected for charge and mass to derive a corrected CCS':

$$CCS' = \frac{CCS \sqrt{\mu}}{z} \quad (22)$$

where  $CCS$  is the literature CCS value,  $z$  is the number of charges and  $\mu$  is the reduced mass. Moreover, measured drift times need to be corrected for  $m/z$ -dependent flight times, as they are dependent on an empirically determined instrument parameter:

$$dt' = dt - C \sqrt{\frac{m}{z}} \quad (23)$$

with  $dt$  being the drift time,  $C$  the constant that depends on the instrument's Enhanced Duty Cycle delay coefficient,  $m$  the mass and  $z$  the number of charges. Corrected CCSs are plotted against corrected drift times and fit to a logarithmic equation:

$$\ln(CCS') = x \ln(dt') + \ln A \quad (24)$$

The  $CCS'$  from the protein of interest is then estimated from the obtained fit.

Recently, a method using computational simulations for calibration was suggested that might allow direct deduction of CCSs from TWIMS data and make calibration with standard proteins dispensable (Mortensen, Susa et al., 2017).

More ion mobility setups have been introduced, such as high-field asymmetric waveform ion mobility spectrometry (FAIMS) (Kolakowski & Mester, 2007) and IMMS that uses structures for lossless ion manipulation (SLIM) (Ibrahim, Hamid et al., 2017).

If static structures from X-ray crystallography, NMR or computational modelling are available, these can be related to experimentally determined CCSs from IMMS. For the calculation of CCSs based on structural information, several calculation algorithms were developed that derive a CCS from simulated collisions of gas molecules with randomly oriented protein structures (Jurneczko & Barran, 2011, Marklund, Degiacomi et al., 2015). Although being computationally expensive, the trajectory method is considered producing most accurate results, since it includes also long-range interactions between gas molecule and target structures into the calculation (Mesleh, Hunter et al., 1996). In contrast, the simpler projection approximation method ignores such target-probe interactions that result in a scattering of gas ions. It only considers the projected area of the target, which leads to a systematic underestimation of the CCS (Jurneczko & Barran, 2011, Mack Jr, 1925). Based on the observation that the output of the two methods correlates well, the software IMPACT was developed that corrects values that were calculated based on the projection approximation method to be in agreement with results from the trajectory method (Bleiholder, Wyttenbach et al., 2011, Marklund, Degiacomi et al., 2015).

## 1.9 High-mass Nano ESI QToF Instruments

Most commercially available mass spectrometers are engineered for the analysis of low-molecular weight analytes such as metabolites or peptides. In comparison, proteins and protein complexes are considerably larger, contain different physical properties, and hence call for special instrumental setups. In 2006, a set of modifications to QToF instruments was presented that markedly improves the instrument performance in the high  $m/z$ -range by accounting for the special requirements of large analytes (van den Heuvel, van Duijn et al., 2006). On the one hand, the operating pressures of the mass spectrometer were adapted. Increased pressures in the source region and in the collision cell intensify collisional focusing of large ions and thereby improve transmission (Chernushevich & Thomson, 2004). On the other hand, electronics of the instrument need to be adapted to satisfy the requirements of high- $m/z$  ions. Decreasing the frequency of the quadrupole mass analyser increases its transmission limit at the cost of mass resolution (Sobott, Hernandez et al., 2002). A decreased ToF repetition rate enables the measurement of high- $m/z$  ions at the cost of sensitivity. Using high-transmission grids at the entry of the ToF analyser enhanced sensitivity for large analytes (van den Heuvel, van Duijn et al., 2006).

---

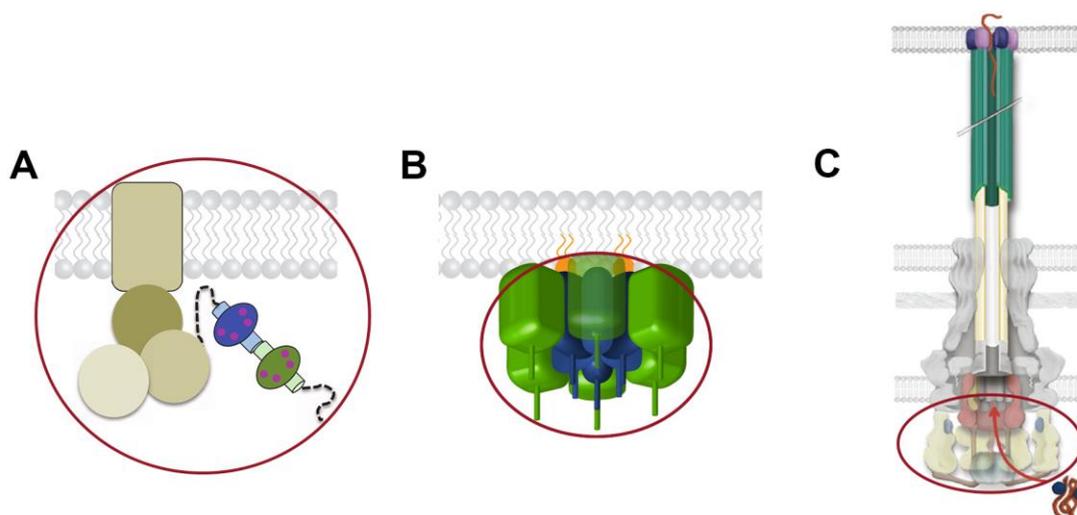
## 2 Objective and Problem Definition of the Thesis

Aim of this thesis was the characterization of protein complexes in different membranous environments by means of native MS.

Integral membrane protein complexes (**Figure 4A**) are most challenging to study, as proteins need to be solubilized in detergent or reconstituted in membrane mimics. In order to acquire conclusive MS data, the protein complex of interest is typically released from the hydrophobic environment, e.g. from the detergent micelle, by highly energetic collisions with inert gas molecules (Barrera & Robinson, 2011). A watchful dosage of collision energies is required to ensure release of the protein complex from the protecting system, while preserving informative protein-protein and protein-lipid complexes (Laganowsky, Reading et al., 2013).

In membrane-associated protein complexes (**Figure 4B**), the main protein-lipid interaction is established between the protein and hydrophilic lipid head groups (Lemmon, 2008). Only small parts of the protein might integrate into the membrane. Once a specific lipid was identified to mediate the protein-membrane interaction, the system is accessible for an in-depth *in vitro* characterization of protein-lipid interactions. In particular, when water-soluble lipids with short hydrophobic tails are used, sample handling is massively facilitated compared to integral membrane proteins. However, the low molecular weight of lipids compared to proteins poses a challenge to MS analyses. Especially very large protein-lipid complexes have high demands on sample preparation and instrumentation to obtain a mass resolution that is sufficient to identify single lipids in the complex.

Soluble subcomplexes (**Figure 4C**) of giant membrane-spanning molecular machines can be studied by native MS without any restrictions in terms of sample handling. Of course, omitting the membrane-standing complex components is a simplification of the biological system, but data on soluble subcomplexes can contain relevant information on complex structure, assembly pathway and function. Moreover, the highly dynamic nature of the subcomplex studied in this work, the sorting platform of the *S. Typhimurium* SPI-1 type 3 secretion system, presents another challenge for native MS characterization.



**Figure 4: Protein-protein interactions in membranous environments.** **A)** Integral membrane protein complexes contain large membrane-spanning domains. **B)** Membrane-associated complexes interact with membrane lipids with no or little insertion into the membrane. **C)** Soluble subcomplexes of membrane-spanning molecular machineries bind to integral membrane proteins. Panel A was adapted with permission from Springer Customer Service Centre GmbH: Springer Nature, Nature, “A bimodular mechanism of calcium control in eukaryotes” (Tidow, Poulsen et al.), Copyright 2012. Panel B is based on “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Changed orientation, added membrane association and label. Panel C was reprinted from “Type Three Secretion System in Attaching and Effacing Pathogens” by Gaytán, Martínez-Santos et al. (2016) from Frontiers in Cellular and Infection Microbiology (10.3389/fcimb.2016.00129), used under CC BY 4.0 license /Added label.

One example of each of the presented protein complexes in membranous environments was studied by means of native MS in the scope of this work. Complex stoichiometries should be identified and gas-phase dissociation techniques (CID, SID) should be utilised to unravel complex topologies (Zhou & Wysocki, 2014). From additional IMMS data, insights into complex conformations were expected (Ben-Nissan & Sharon, 2018).

By combining the MS results with data from various other biophysical techniques (for instance SAXS, molecular modelling, X-ray crystallography), acquired in the framework of highly collaborative projects, comprehensive pictures of different protein and protein-lipid complexes should be generated describing the complex structures on a molecular level (Liko, Allison et al., 2016). Based on these structures, additional conclusions and hypothesises on the assembly mechanism and function of the respective system were expected.

---

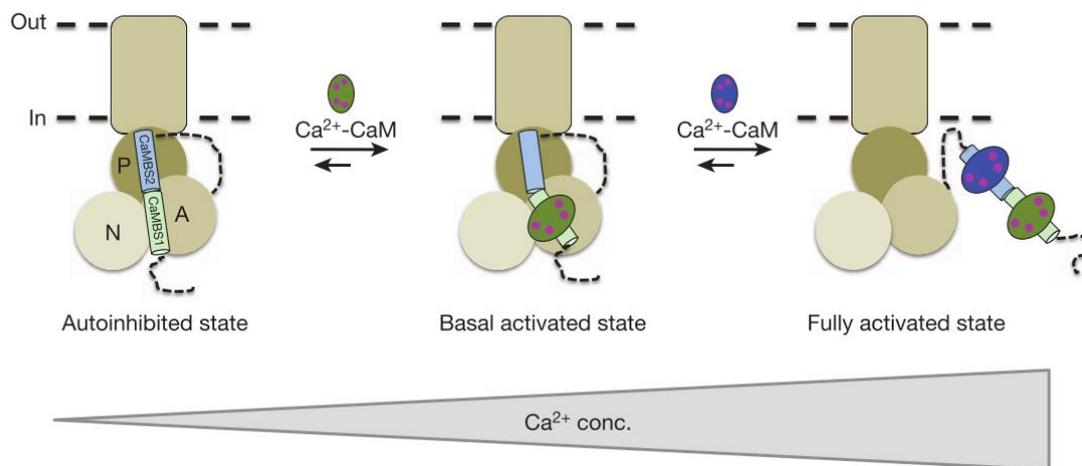
## 3 Binding of Two Calmodulins to the Plasma-membrane $\text{Ca}^{2+}$ -ATPase ACA8

### 3.1 Introduction

Many signal transduction processes involve the translation of a primary extracellular signal into a secondary intracellular messenger signal. Calcium ions ( $\text{Ca}^{2+}$ ) are ubiquitous secondary messengers and their release is for example triggered by GPCR signalling or activation of ion channels (Bootman, 2012, Clapham, 2007). A prerequisite for the signalling mechanism involving  $\text{Ca}^{2+}$  ions is a concentration gradient between low intracellular (100 nM) and high extracellular (2 mM) concentrations that sets the baseline for the signalling event, consisting of a rapidly increasing intracellular  $\text{Ca}^{2+}$  concentration (Clapham, 2007). For maintaining the concentration gradient before and re-establishing it after a signalling event, eukaryotic cells use sodium-calcium exchanger proteins and plasma-membrane  $\text{Ca}^{2+}$ -ATPases (PMCA). The latter maintain  $\text{Ca}^{2+}$  homeostasis by exporting intracellular  $\text{Ca}^{2+}$  powered by the hydrolysis of ATP (Carafoli, 1994). PMCA are tightly regulated to ensure a high pump activation only when  $\text{Ca}^{2+}$  efflux is required (Di Leva, Domi et al., 2008).

The  $\text{Ca}^{2+}$ -binding protein calmodulin (CaM) senses intracellular calcium concentrations and binds to the regulatory domain of PMCA that is located in mammalian PMCA at the C-terminus and in plant PMCA at the N-terminus of the protein (Bonza, Morandini et al., 2000). Binding of  $\text{Ca}^{2+}$ -loaded CaM to the regulatory PMCA domain abrogates autoinhibition and thus leads to an increased  $\text{Ca}^{2+}$  transport (Di Leva, Domi et al., 2008). PMCA pump rates as well as  $\text{Ca}^{2+}$  affinity are enhanced upon CaM binding (Clapham, 2007).

Although the functional role of PMCA activation by binding of  $\text{Ca}^{2+}$ -loaded CaM has been studied, the structural basis of the mechanism remains uncertain. Two CaMs bound to the regulatory domain of PMCA from *Arabidopsis thaliana* (*A. thaliana*) ACA8 were crystallized and the structure was solved to 1.95 Å resolution (Tidow, Poulsen et al., 2012). It was further hypothesized that the consecutive binding of two  $\text{Ca}^{2+}$ -loaded CaMs to the regulatory domain of ACA8 induces a two-step activation mechanism that results in the displacement of the regulatory domain from the catalytic core (Tidow, Poulsen et al., 2012). This conformational change would then allow the actuator domain of ACA8 to move freely and pump  $\text{Ca}^{2+}$  more efficiently (Tidow, Poulsen et al., 2012).

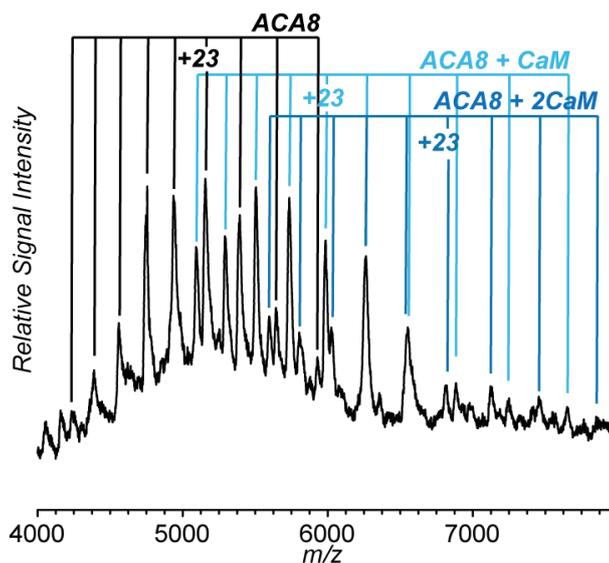


**Figure 5: Hypothesized two-step activation mechanism by binding of two  $\text{Ca}^{2+}$ -loaded CaMs to the regulatory domain of ACA8.** At low  $\text{Ca}^{2+}$  concentrations, the PMCA of *A. thaliana* ACA8 is in an autoinhibited state. The regulatory domain with two binding sites for CaMs (CaMBS1, CaMBS2) is bound to the catalytic core of the protein with its actuator (A), nucleotide-binding (N) and phosphorylation (P) domains. At increasing  $\text{Ca}^{2+}$  concentrations,  $\text{Ca}^{2+}$ -loaded CaM binds to the high-affinity CaMBS1. At even higher  $\text{Ca}^{2+}$  concentrations, also the second CaMBS2 gets occupied and consequently the regulatory domain is displaced from the catalytic core, leading to a fully activated state of the PMCA. Reprinted with permission from Springer Customer Service Centre GmbH: Springer Nature, Nature, “A bimodular mechanism of calcium control in eukaryotes” (Tidow, Poulsen et al.), Copyright 2012.

However, since crystallization trials of full-length ACA8 were not yet successful, different methods need to be applied to study the activation mechanism of the full-length PMCA. Here, the interaction of ACA8 and CaM was analysed by native mass spectrometry to validate the hypothesis that the binding of two  $\text{Ca}^{2+}$ -loaded CaMs to the isolated regulatory domain is transferable to full-length PMCAs.

### 3.2 Results and Discussion

Full-length ACA8 was solubilized in DDM, mixed with purified CaM and analysed by native mass spectrometry. As expected from previous crystallization studies (Tidow, Poulsen et al., 2012), resulting spectra revealed masses of ions that correspond to unbound ACA8, as well as ACA8 with one or two bound CaM molecules (Figure 6, theoretical and measured masses in Supplementary Table S1, protein sequences in Supplementary Table S2).



**Figure 6: Binding of two calmodulins to full-length, detergent-solubilized ACA8.** At high collision energies (200 V), proteins were released from detergent (DDM) micelles. Identified masses (Supplementary Table S1) indicate the presence of unbound ACA8 (black), ACA8 + CaM (light blue) and ACA8 + 2CaM (dark blue). Adapted from Nitsche, Josts et al. (*in revision*).

However, ratios of unbound and CaM-bound ACA8 in native mass spectra cannot be directly transferred to the ratios that were present in solution. Detergent-solubilized proteins were sprayed from ammonium acetate solution and release of the membrane proteins from their detergent micelles required high activation energies during the MS measurements. With the used instrumentation, the energy for the release of the membrane proteins needs to be applied as high collision energies, meaning that analyte ions experience fast acceleration and consequently high-energy collisions with argon molecules in the collision cell. As an unwanted side effect of such a high energy input, protein-protein interactions can be disrupted, causing collision-induced dissociation (see 1.7 Gas-phase Dissociation and Fragmentation, page 20).

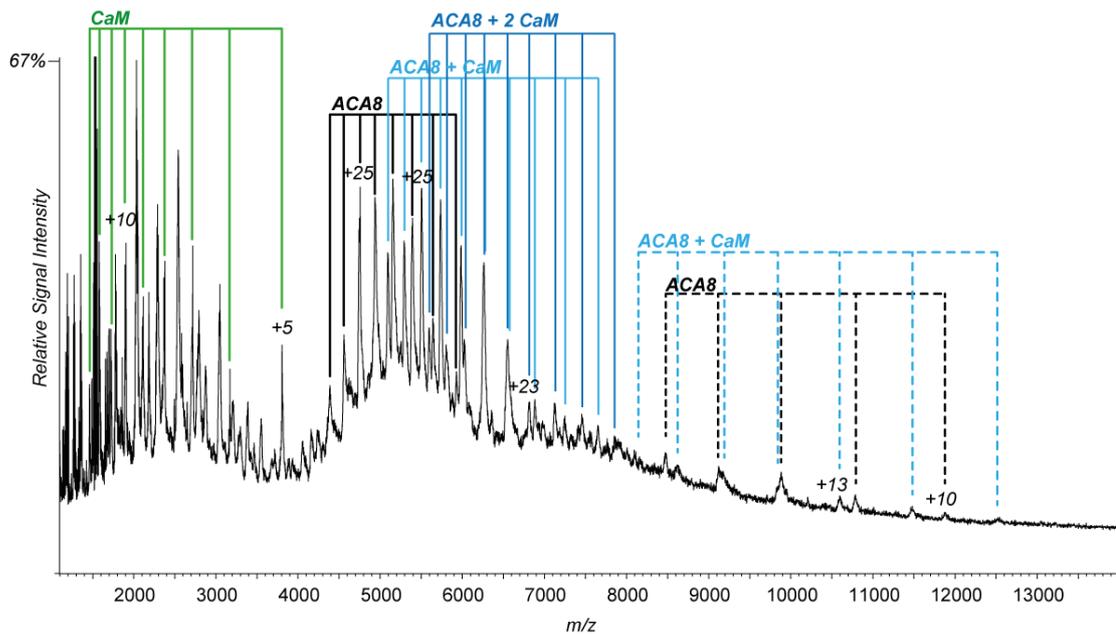
Acquired spectra show clear signs of CaM dissociation from ACA8 (Figure 7). In the high  $m/z$ -range, unbound ACA8 and ACA8 with one bound CaM molecule were identified. As expected, free CaM was found in the low  $m/z$ -range. Whether these proteins are dissociation

products, or were never bound to ACA8 can be judged based on the number of charges of CaM in the spectrum. Due to the asymmetric charge partitioning upon collision-induced dissociation of proteins, the expected number of charges for dissociated proteins is higher than for initially unbound proteins (Jurchen & Williams, 2003). CaM ions that were analysed in absence of ACA8 and DDM carry +6 to +8 charges (**Supplementary Figure S1A**). In contrast, CaM ions with the charge states +5 to +12 were observed during native MS analysis of CaM-ACA8 interaction (**Figure 7**). Another possible explanation of the high number of charges is a supercharging effect that was recently reported to be an inherent feature of non-ionic saccharide detergents (Loo, O'Brien Johnson et al., 2018). However, the lowly charged ACA8 and ACA8 + CaM ions in the mass range above 8,000 *m/z* clearly point towards a dissociation event. Possibly, gentler instrument settings can be found that ensure an efficient release of the ACA8/CaM from the DDM micelles while the protein complexes itself remain stable.

The fact that exclusively ACA8 and ACA8 + CaM were present as low-charge ions underline that a maximum number of two CaMs bind to ACA8.

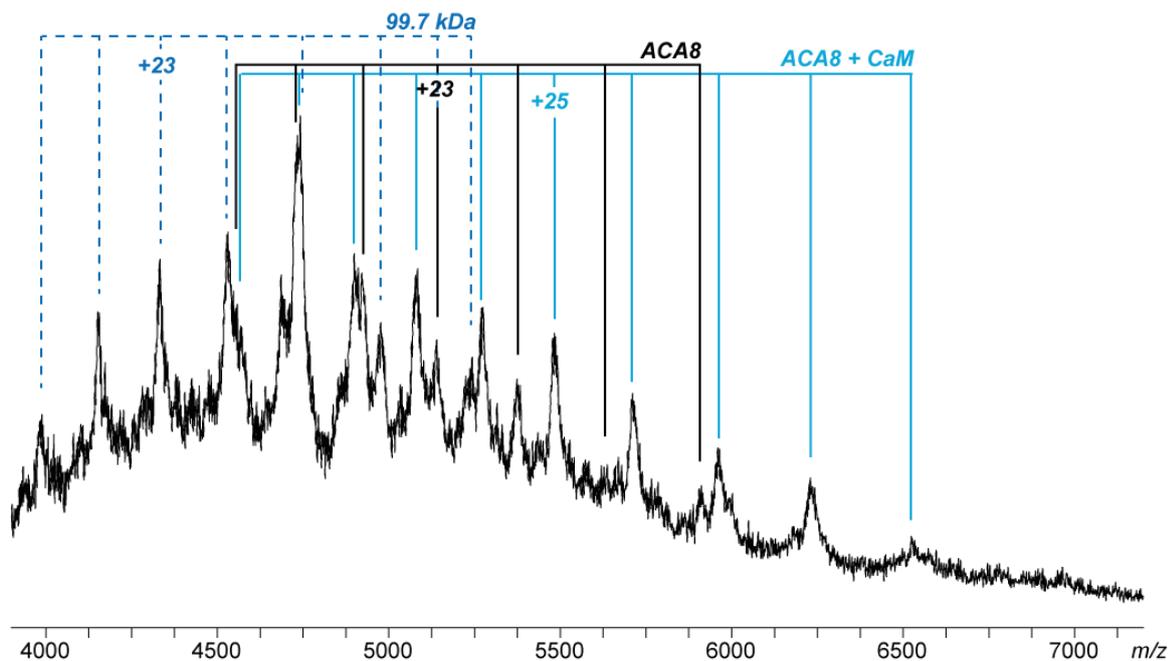
The observed fine structure of the CaM peaks further revealed information about the occupancy of the four Ca<sup>2+</sup>-binding binding sites of CaM (**Supplementary Figure S1B**). Even after washing the protein sample with EDTA-containing solution, the majority of the measured CaMs carries 2, 3 or 4 Ca<sup>2+</sup> ions. This observation is in agreement to previous studies that identified two cooperative high-affinity Ca<sup>2+</sup> binding sites at the C-terminal globular domain of CaM and additionally two low affinity cooperative binding sites at the N-terminal globular domain (Linse, Helmersson et al., 1991).

Together, these results support the previously presented structural data showing binding of two CaMs to the regulatory domain of ACA8 and do not disagree with a possible two-step activation mechanism of PMCAs.



**Figure 7: Native mass spectrum revealed binding of two CaMs to ACA8 and MS-induced dissociation.** Apart from the identified ACA8 (solid black), ACA8 + CaM (solid light blue) and ACA8 + 2CaM (solid dark blue) (see Figure 6) also lowly charged ACA8 (dotted black) and ACA8 + CaM (dotted light blue) were found in the high  $m/z$ -range. Unbound CaM (green) was present in the low  $m/z$ -range, a high number of charges indicated that some of the ions were unfolded dissociation products. Reprinted from Nitsche, Josts et al. (*in revision*).

Ion mobility mass spectrometry measurements to assess the CCSs of ACA8 and ACA8/CaM complexes were not successful. In particular, a sufficient activation of the ions to release the proteins from the detergent micelles and break large detergent clusters without simultaneously unfolding the proteins proved difficult. However, these measurements revealed an unexpected observation. In samples that contained ACA8 and CaM, either copurified or mixed after purification, unbound ACA8, ACA8 + CaM and an additional unidentified 99.7 kDa-species were found (Figure 8, Supplementary Figure S2A, B). In contrast to previous native MS measurements no ACA8 + 2CaM was found. Remarkably, the 99.7 kDa-species was only present in samples that contained ACA8 and CaM. In ACA8 samples that were measured as a reference sample prior to mixing with CaM, the 99.7 kDa-species was not observed (Supplementary Figure S2C). The ACA8 sequence was inspected for prolin residues, because CID backbone fragmentations appear preferentially N-terminal to prolin (Haverland, Skinner et al., 2017). Indeed, a cleavage N-terminal to prolin 166 was identified to yield fragments of 18.5 kDa (N-terminus) and 99.7 kDa (C-terminus). Hence, binding of the second CaM to ACA8 could induce structural rearrangements that enable fragmentation at prolin 166.



**Figure 8:** Native MS analysis of ACA8/CaM on a Synapt G2 mass spectrometer (Waters) revealed a 99.7 kDa-species (dotted lines, dark blue). Since the species was not found in ACA8 samples, a fragmentation mechanism induced by binding of two CaM molecules to ACA8 was hypothesized. The spectrum showed ACA8 (black) and ACA8 + CaM (light blue). Data were acquired in IMMS mode, the area of interest was extracted from the drift plot to reduce overlapping detergent signals (Supplementary Figure S2A, B).

Nevertheless, some arguments oppose the hypothesis of a fragmentation mechanism induced by a structural change upon binding of two CaM molecules. The hypothesized fragmentation would also produce a 56.6 kDa-species, comprised of the N-terminal regulatory domain of ACA8 and two bound CaM molecules. However, no signal corresponding to this mass was identified in the recorded spectra. Massive signal overlap with detergent signals in the range below 4,000  $m/z$  further hampered the identification of free calmodulin and possible free N-terminal ACA8 fragments in IMMS measurements. Furthermore, the relatively high number of charges of the 99.7 kDa-species argues against a preceding fragmentation event.

Therefore, the 99.7 kDa-species is possibly a contaminant that only ionized efficiently under specific spray conditions. Moreover, the massive overlap with detergent signals hampered spectra analysis (Supplementary Figure S2A, B).

Due to unstable electrospray and limited access to IMMS instruments, the obtained results give no clear indication of a conformational change, such as a detachment of the regulatory domain, upon binding of two CaMs.

### 3.3 Methods

The description of the performed experiments was adapted from the submitted manuscript Nitsche, Josts et al. (*in revision*).

#### 3.3.1 Protein Production and Purification

ACA8 was produced with an N-terminal 8xHis-tag in *Saccharomyces cerevisiae* (*S. cerevisiae*) strain BJ5460 (MATa ura352 trp1 lys2801 leu2delta1 his3delta200 pep4::HIS3 prb1delta1.6R can1 GAL) (Jones, 1991) using a pYES2 plasmid. Large-scale cultures were grown in uracil-depleted media (6.7 g/l YNB) + 0.1 % glucose at 30 °C up to an OD<sub>600</sub> of 0.6 and expression was induced by the addition of 2 % galactose. Cells were harvested 20 h post induction by centrifugation at 3,000 xg and resuspended in 30 mM Tris pH 8.0, 300 mM NaCl, 20 % (v/v) glycerol, 3 mM β-mercaptoethanol, 20 mM EDTA (buffer A) before being disrupted with glass beads. Following clearing of the lysate, cell membranes were isolated by centrifugation at 180,000 xg and membranes were solubilized in buffer A with 0.1 % undecylmaltoside (UDM) for 1.5 h with gentle stirring. To remove insoluble material the solubilized membranes were centrifuged at 100,000 xg and supernatant was incubated with Ni<sup>2+</sup> affinity resin. The resin was washed with buffer containing 20 mM Tris pH 8.3, 120 mM NaCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 1 mM β-mercaptoethanol, 0.03 % UDM and 40 mM imidazole and ACA8 was eluted with 150 mM imidazole. The purity of ACA8 was judged by SDS-PAGE and corresponding fractions were pooled and concentrated to 2 mg/ml.

CaM7 from *A. thaliana* was used as CaM ortholog throughout this study. CaM in pET42a vector was transformed into *E. coli* Bl21 Gold (DE3) and grown in LB medium at 37 °C. After induction of gene expression with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), cells were grown for another 16 h at 20 °C and harvested by centrifugation at 3,000 xg. Cells were resuspended in 30 mM Tris pH 7.5, 50 mM NaCl, 1 mM β-ME, 2 mM CaCl<sub>2</sub> (buffer B) and broken using high-pressure homogenizer (EmulsiFlex-C3, Avestin). The lysate was cleared by centrifugation at 40,000 xg and the supernatant was bound to a HiTrap Phenyl HP column, pre-equilibrated with buffer B. CaM was eluted with 5 mM EDTA. Fractions containing pure CaM were pooled, concentrated to 10 mg/ml and stored at -80 °C until further use.

Amino acid sequences of ACA8 and CaM as used for further analysis are presented as **Supplementary Table S2**.

### 3.3.2 Native Mass Spectrometry

Detergent-solubilized ACA8 was changed to 200 mM ammonium acetate solution pH 8.3, 2x CMC (0.018 % w/v) DDM using centrifugal filter units (Vivaspin 500, 100,000 MWCO, Sartorius) at 4 °C and 15,000 xg. For measurements of CaM in absence of ACA8 and DDM, CaM was washed twice with 25 mM EDTA, 400 mM ammonium acetate, and then 10 times with 400 mM ammonium acetate, in both cases using ZebaSpin desalting columns (7k MWCO, 0.5 ml, ThermoFisher Scientific).

Nano ESI capillaries were prepared in-house by processing borosilicate capillaries (1.2 mm outer diameter, 0.68 mm inner diameter, World Precision Instruments) in a two-step program with a micropipette puller (P-1000, Sutter instruments) using a squared box filament (2.5 × 2.5 mm, Sutter instruments). Subsequently, capillaries were gold-coated using a sputter coater (Q150R, Quorum Technologies) with 40 mA, 200 s, tooling factor of 2.3 and end bleed vacuum of 8×10<sup>-2</sup> mbar argon.

Native MS experiments were performed with a nano ESI source in positive ion mode on a QToF 2 (Waters, MS Vision) mass spectrometer that was modified for the analysis of high-mass ions (van den Heuvel, van Duijn et al., 2006). Capillary and sample cone voltage were used at 1700 V and 190 V, respectively. 7 mbar source pressure and 1.7×10<sup>-2</sup> mbar argon as collision gas were used. The collision energy was gradually ramped to 400 V to release protein complexes from detergent micelles. CaM spectra were acquired with 1500 V capillary voltage, 150 V sampling cone voltage, 30 V collision energy. Gas pressures were 10 mbar in the source region and 1.7×10<sup>-2</sup> mbar argon in the collision cell. Illustrated ACA8/CaM spectra (**Figure 6**, **Figure 7**) were recorded at 200 V collision energy.

CsI (25 mg/ml) spectra were acquired on the same day and used to calibrate raw data with MassLynx (Waters) software. Data were analysed using MassLynx and *Massign* (Morgner & Robinson, 2012). Average measured masses of protein complexes, standard deviations of replicate measurements and average full width at half maximum (FWHM) values as a measure of the mass resolution are listed in **Supplementary Table S1**.

Preliminary IMMS data were recorded on a Synapt G2 (Waters). Samples were measured in positive ion mode with the following settings: 1600 V capillary voltage, 200 V cone voltage, 50 V trap collision energy, 50 V trap DC bias, 35 V helium DC, 250 V transfer collision energy. As it was planned to estimate the CCS of ACA8/CaM complexes and no TWIMS calibration standard for membrane proteins is available, conditions had to be found that are harsh enough to release ACA8 from detergent micelles but gentle enough to measure drift times of soluble

protein complexes as calibration standards. Therefore, high activation energies in source and transfer cell were chosen, while the trap activation energy was relatively low for the measurement of membrane proteins.

The backing pressure was set to 8 mbar and 3 mbar nitrogen (IMS gas flow 90 ml/min) was used for ion mobility separation. Wave velocities of 400 m/s and wave heights of 40 V were used. Wave height and velocity in the trap cell was 6.0 V and 300 m/s, respectively. In the transfer cell the wave height was 2.9 V and the wave velocity was 85 m/s. Argon was used as collision gas with  $5.1 \times 10^{-2}$  mbar in the trap (flow 10 ml/min) and  $4.8 \times 10^{-2}$  mbar in the transfer cell. The pusher frequency was set to  $1900 \text{ s}^{-1}$ .

### **3.4 Outlook**

Ion mobility mass spectrometry measurements of ACA8 alone and in complex with CaM need to be repeated. A dramatically increasing CCS upon binding of the second CaM to ACA8 would be a proof of a conformational change. Since the quality of the thus far acquired data is not sufficient, measurements need to be repeated with fresh samples that have not been frozen or stored for extended periods to avoid protein degradation and an instable electrospray. Whether the 99.7 kDa protein is an ACA8 fragmentation product or a contaminant from imperfect protein purification could be tested by mutating the proline residue 166 and repeating the analysis. Alternatively, a bottom-up proteomics experiment could help to identify the protein.

Nevertheless, even if the regulatory ACA8 domain disengages from the catalytic core upon binding of two CaMs, it is not clear if this extended conformation is stable in the gas phase or if the domain collapses in the course of the measurement and the measured drift times are smaller than expected from the solution conformation. A similar behaviour was described for different non-globular proteins, such as antibodies (Devine, Fisher et al., 2017). Moreover, obtainable results from TWIMS measurements will rather be a relative change of the drift time than a change in the CCS, because finding instrument settings that allow the measurement of membrane proteins and calibration with soluble protein standards is extremely challenging. Probably a screen for detergents that require less activation energy to release proteins from micelles, e.g. Triton X-100, would be useful (Konijnenberg, Yilmaz et al., 2014).

Consequently, the use of other biophysical methods is perhaps more straightforward to study the structural basis of ACA8 activation. In the current collaborative project, structural changes of ACA8 upon CaM binding were detected in small-angle neutron and X-ray scattering studies. For these experiments, ACA8 was reconstituted in nanodiscs and not solubilized with

detergents. Attempts to release ACA8 from nanodiscs for native MS analysis were not successful.

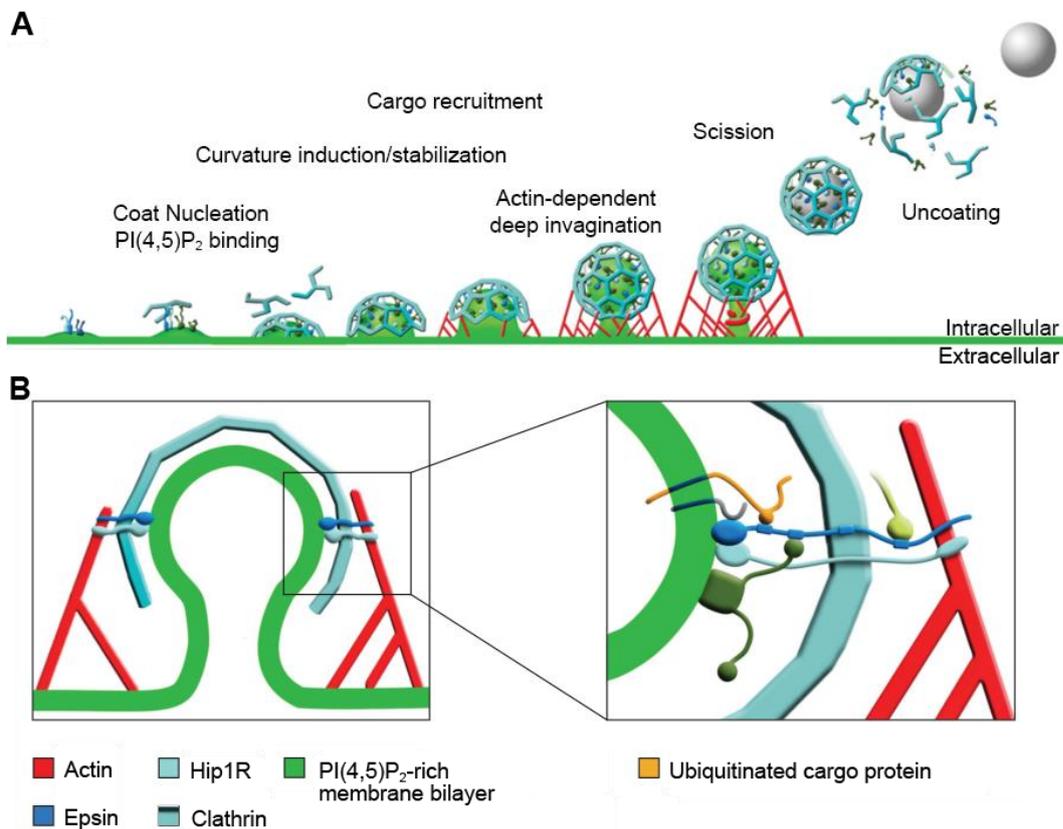
## 4 Epsin and Sla2 Form Assemblies through Phospholipid Interfaces

### 4.1 Introduction

#### 4.1.1 Clathrin-mediated Endocytosis

Clathrin-mediated endocytosis (CME) is a process in eukaryotic cells for the directed transport of cargo from the outside to the inside of the cell (Kaksonen & Roux, 2018). The nature of the cargo can be widely different, ranging from nutrients to signalling molecules that are bound extracellularly to their transmembrane receptor. The large variety of cellular pathways that rely on clathrin-mediated endocytosis results in a high frequency of events as well as evolutionary conservation. Hence, it is not surprising that many viruses exploit the pathway for cell entry, as for instance Influenza A Virus, Vesicular Stomatitis Virus or Semliki Forest Virus (Cossart & Helenius, 2014).

A complex pioneer module of different scaffold and adaptor proteins at the plasma membrane plays the central role for initiation of endocytic events (Kaksonen & Roux, 2018, Ma, Umasankar et al., 2016). Many of the adaptor proteins are recruited to the membrane by binding to the phospholipid phosphatidylinositol-4,5-bisphosphate (PI(4,5)P<sub>2</sub>) (Antonescu, Aguet et al., 2011). Clathrin coat proteins and coat-associated proteins bind to these membrane-associated modules, creating nucleation sites for coat assembly (Figure 9). In mammalian cells high local protein concentrations induce bending of the plasma membrane and lead to the formation of clathrin coated pits. Mechanistically it is still unclear, whether clathrin molecules assemble with a constant curvature or if reorganization of a preformed flat patch of clathrin coat induces membrane bending (Avinoam, Schorb et al., 2015, Saleem, Morlot et al., 2015). Different proteins of the preassembled clathrin coat bind to specific cytosolic recognition motifs of transmembrane cargo receptors (Traub, 2009). PTMs can regulate the cargo selection process. While ubiquitylation of cargo receptors is often required for interaction with clathrin adaptors, phosphorylation prevents it in some cases (Mukhopadhyay & Riezman, 2007, Shiratori, Miyatake et al., 1997). Since components of the pioneering complex are among the cargo recruiting proteins, also cargo molecules themselves might recruit components of the pioneer module and thereby promote the initiation of endocytic events by their presence (Liu, Aguet et al., 2010).



**Figure 9: Clathrin-mediated endocytosis involves interactions between plasma membrane, adaptor proteins and actin cytoskeleton.** **A)** Assembly of the clathrin coat nucleates at pioneer modules. The membrane bends and cargo is recruited. Later membrane shaping steps involve actin polymerization. After membrane scission, vesicles are uncoated and directed to early endosomes. **B)** Clathrin adaptor proteins Hip1R and epsin are major connection hubs during endocytosis. N-terminal domains establish connections with the plasma membrane and C-termini bind the actin cytoskeleton. Furthermore, the proteins interact with the clathrin coat and ubiquitin interaction motifs in epsin recruit cargo proteins. Reprinted from “Epsin deficiency impairs endocytosis by stalling the actin-dependent invagination of endocytic clathrin-coated pits” by Messa, Fernandez-Busnadiego et al. (2014) from *eLife* (10.7554/eLife.03311), used under CC BY 4.0 license /Adapted labels.

In later phases of CME, actin polymerization was shown to play a role in establishing a deep membrane invagination (Kaksonen, Toret et al., 2006). For that purpose, growing actin filaments need to be physically connected to the clathrin coated pit in order to transmit a directed force. This connection is established via interactions with the clathrin adaptor molecules epsin and Hip1R (Huntingtin-interacting protein 1-related protein in mammals, Sla2 in yeast) (Skruzny, Brach et al., 2012).

Membrane scission of vesicles is executed by the GTPase dynamin together with associated proteins (Damke, Baba et al., 1994). The coat of the resulting intracellular endocytic vesicles is afterwards disassembled by interactions with chaperones and dephosphorylation of PI(4,5)P<sub>2</sub> (McPherson, Garcia et al., 1996, Schlossman, Schmid et al., 1984). By fusion with early

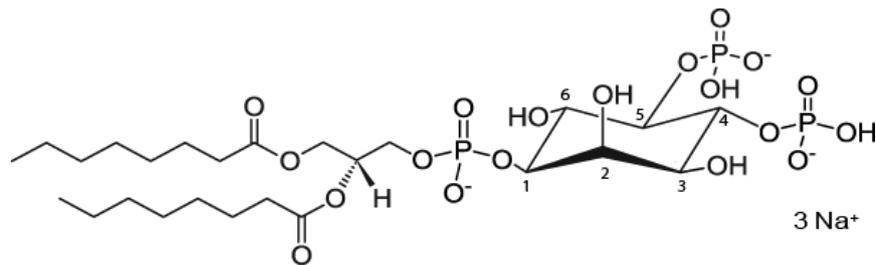
endosomes the vesicles are fed into the endosomal pathway where the vesicles' components are sorted for recycling and degradation (Maxfield & McGraw, 2004).

Although the CME process in mammals and yeast is comparable, several differences were identified and reviewed (Kaksonen & Roux, 2018, Liu, Sun et al., 2009). For instance, the membrane of yeast cells bends in much later stages and invaginations have a more tubular shape, from which vesicles are pinched off. Moreover, actin plays a more important role in yeast CME, where inhibition of actin polymerization blocks the CME event already in early stages (Kukulski, Schorb et al., 2012). Although actin was demonstrated to be not essential for mammalian endocytosis, studies indicate that actin counteracts high membrane tensions (Boulant, Kural et al., 2011, Fujimoto, Roth et al., 2000).

#### 4.1.2 *Phosphatidylinositol-4,5-bisphosphate in Clathrin-Mediated Endocytosis*

PI(4,5)P<sub>2</sub> (Figure 10) is best known for its role as substrate for the enzyme phospholipase C and hence as the precursor of the second messenger inositol-1,4,5-trisphosphate (IP<sub>3</sub>) in GPCR signalling pathways (Berridge, 1987). For its genesis, phosphatidylinositol is translocated from the ER membrane to the plasma membrane by transport proteins and then sequentially phosphorylated to PI(4,5)P<sub>2</sub> (Fruman, Meyers et al., 1998, Lev, 2012). Additionally, basal levels of phosphatidylinositol-4-phosphate are transported to the membrane and then further phosphorylated to PI(4,5)P<sub>2</sub> (Dickson, Jensen et al., 2014). Although the total PI(4,5)P<sub>2</sub> concentration was shown to be as low as 1 % of plasma membrane phospholipids or 10 μM in total cells, local concentrations are much higher (Hilgemann, 2007). Microscopy and molecular modelling both found evidence for PI(4,5)P<sub>2</sub>-dominated microdomains in the inner leaflet of the plasma membrane (Ingolfsson, Melo et al., 2014, van den Bogaart, Meyenberg et al., 2011).

PI(4,5)P<sub>2</sub>-association was found for many clathrin adaptor proteins, such as AP-2, epsin, CALM and Hip1R, as summarized by Antonescu, Aguet et al. (2011). Also, perturbations of PI(4,5)P<sub>2</sub> levels by sequestration were found to inhibit CME (Jost, Simpson et al., 1998). PI(4,5)P<sub>2</sub> is not only important as early interaction partner for proteins in CME. Its concentration is tightly regulated by the presence of kinases and phosphatases that convert PI(4,5)P<sub>2</sub> to other phospholipids, which are necessary in later steps of the process (Kaksonen & Roux, 2018, Wallroth & Haucke, 2018). Not only that the disassembly of clathrin coats is dependent on dephosphorylation (see 4.1.1 Clathrin-mediated Endocytosis, page 37), also conversion of PI(4,5)P<sub>2</sub> to PI(3,4)P<sub>2</sub> was found to be important for the recruitment of proteins that assist during fission (Posor, Eichhorn-Gruenig et al., 2013).



**Figure 10: Phosphatidylinositol-(4,5)-bisphosphate.** Throughout this study PI(4,5)P<sub>2</sub> with two octanoyl (08:0) fatty acid chains was used.

In contrast to PI(4,5)P<sub>2</sub> derivatives with longer fatty acid chains, dioctanoyl (08:0) PI(4,5)P<sub>2</sub> (**Figure 10**) is water soluble and consequently useable as additive to standard buffer systems in biophysical in-solution assays.

#### 4.1.3 Clathrin Adaptor Proteins: Epsins and Sla2/Hip1R

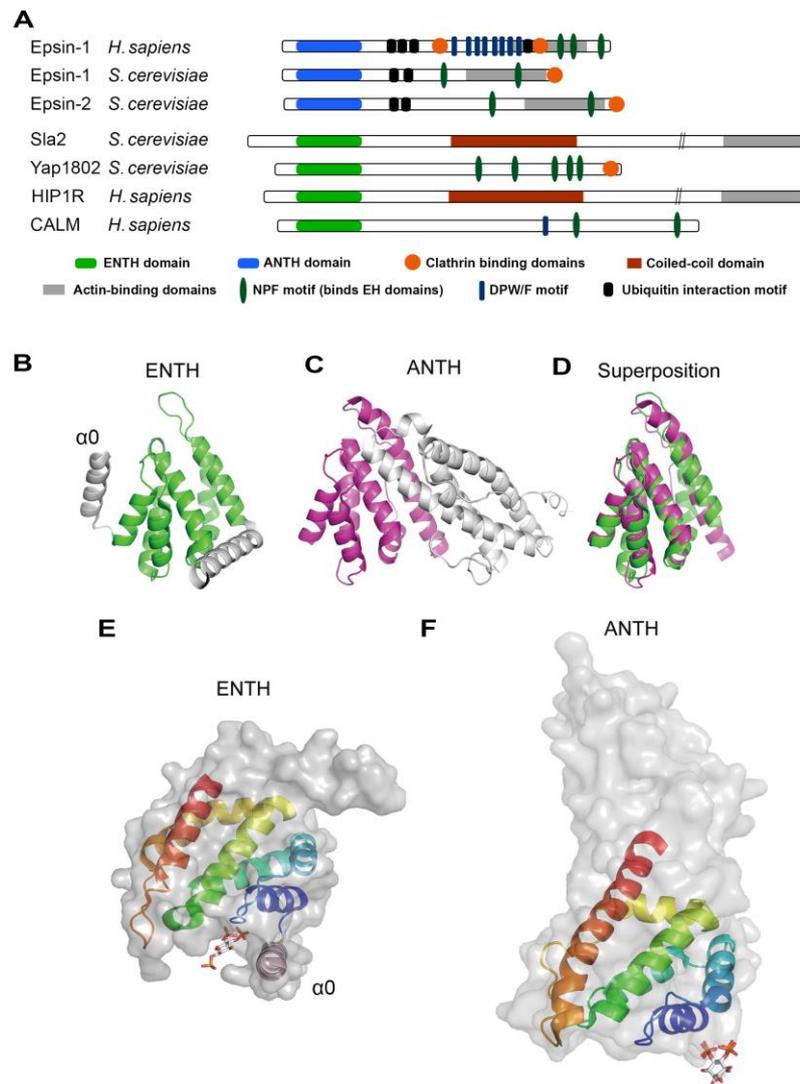
Epsins are a protein family of extensively studied clathrin adaptors. In mammals, the protein family consists of three different isoforms: epsin-1, epsin-2 and epsin-3, of which epsin-1 and epsin-2 have a broad tissue distribution and are commonly studied in the context of CME (Sen, Madhivanan et al., 2012). Since only double knock-outs of epsin-1 and epsin-2 lead to early embryonic lethality in mice, distinct roles for these isoforms remain unclear and the proteins are suspected to have redundant functions (Chen, Ko et al., 2009). In contrast, epsin-3 is mainly found in keratinocytes and was linked to wound healing (Spradling, McDaniel et al., 2001). Moreover, epsin-3 has the strongest association of the epsin isoforms to tumorigenesis (Coon, Drenzo et al., 2011). In *S. cerevisiae*, related proteins to mammalian isoforms 1 and 2 are known, as well as three epsin-related proteins that are involved in Golgi/endosome trafficking (Duncan, Costaguta et al., 2003).

As needed for their function as interaction hubs in CME, epsin structures contain several interaction domains (**Figure 11A**). At the C-terminus, clathrin binding domains and NPF-repeats are located that are recognized by proteins containing EH (eps 15 homology) domains (Confalonieri & Di Fiore, 2002). In addition, DPW/F motifs bind to the clathrin adaptor protein AP-2. *S. cerevisiae* epsin-1 contains an actin-interaction domain, which is regulated by protein phosphorylation (Skruzny, Brach et al., 2012). The direct interaction between epsin-1 and actin was also confirmed for mammals (Messa, Fernandez-Busnadiego et al., 2014). C-terminal sequences diverge between species and contain large unstructured regions (Kalthoff, Alves et al., 2002). Ubiquitin-interacting motifs in central protein regions allow selection and recruitment of ubiquitylated cargo receptors.

A common feature of all epsins is an epsin N-terminal homology (ENTH) domain that mediates the interaction with the plasma membrane. ENTH domains are the only epsin domains of which structures determined by NMR or X-ray crystallography are available (Ford, Mills et al., 2002, Itoh, Koshiba et al., 2001). ENTH domains consist of approximately 150 amino acids, forming eight  $\alpha$ -helices, which arrange to a compact structure (**Figure 11B**). ENTH domains bind PI(4,5)P<sub>2</sub> by their inositol headgroups, inducing the formation of an additional N-terminal  $\alpha$ -helix, named ‘ $\alpha$ -helix zero’ ( $\alpha$ 0). The newly formed amphipathic helix inserts into the plasma membrane, promoting membrane curvature (Ford, Mills et al., 2002, Yoon, Tong et al., 2010). However, epsins also sense the membrane curvature and preferentially assemble at deeply curved membranes (Holkar, Kamerkar et al., 2015). These observations match the general notion that epsins are not part of the pioneering module for endocytosis initiation but are involved in later stages of clathrin coated pit formation (Lu, Drubin et al., 2016).

Lipid binding was further characterized in co-sedimentation assays, ITC, and SPR and shown to be dependent on the phosphorylation of the inositol headgroups (Ford, Mills et al., 2002, Hom, Vora et al., 2007, Itoh, Koshiba et al., 2001). Binding affinities were determined to be in the low micromolar (co-sedimentation, ITC) or in the nanomolar range (SPR) and stronger in acidic environments (Hom, Vora et al., 2007, Stahelin, Long et al., 2003).

In addition to epsins, another group of clathrin adaptor proteins was found to contain a characteristic membrane domain that specifically binds PI(4,5)P<sub>2</sub>. In reference to the protein in which it was identified, it was named AP180 N-terminal homology (ANTH) domain (Ford, Pearse et al., 2001). Although ANTH domains are with approximately 260 amino acids considerably larger than ENTH domains, both domains share structural features. ANTH domains are also largely helical structures and helices 1-6 superimpose well with ENTH domains (**Figure 11D**). PI(4,5)P<sub>2</sub> binding was as well characterized and the affinity was found to be slightly lower than for ENTH domains (Stahelin, Long et al., 2003). Despite the structural similarity of ENTH and ANTH domains, the nature of the identified binding site is different. Instead of a binding pocket, in ANTH domains a set of exposed lysine and histidine side chains are found (**Figure 11E, F**) (Ford, Pearse et al., 2001).



**Figure 11: Clathrin adaptor protein structures.** **A)** Domain structures of ENTH- and ANTH-domain containing proteins. ENTH domains (blue) and ANTH domains (green) are located at the N-termini of the proteins, further interaction motifs are illustrated as annotated in the legend. The precise location of actin binding domains is yet to be identified. **B)** Crystal structure of rat ANTH domain (PDB ID: 1HG2) with the annotated helix  $\alpha 0$ . **C)** Crystal structure of rat epsin-1 ENTH domain (PDB ID: 1H0A). **D)** The superposition of rat ANTH (purple) and ENTH (green) domains reveals a similar fold of six N-terminal helices in ANTH and ENTH domains. **E)** The rat ENTH domain is depicted as transparent surface with the six N-terminal helices highlighted as rainbow cartoon representation. Binding of inositol-1,4,5-trisphosphate induced formation of helix  $\alpha 0$ . **F)** Likewise, the rat ANTH domain in surface representation with helices 1-6 highlighted in rainbow colours is shown as co-crystallized with bound inositol-4,5-bisphosphate. Note the different binding sites of substrate mimics for ENTH and ANTH domains. Panel A was prepared based on De Camilli, Chen et al. (2002) and panels B-F adapted with permission from Springer Customer Service Centre GmbH: Springer Nature, Springer eBook, “AP180 N-Terminal Homology (ANTH) and Epsin N-Terminal Homology (ENTH) Domains: Physiological Functions and Involvement in Disease” (Takatori & Tomita), Copyright 2018.

Different subfamilies of clathrin adaptor proteins contain ANTH domains. In CALM (Clathrin assembly lymphoid myeloid leukaemia, also known as AP180) the formation of an N-terminal  $\alpha$ 0-helix was described, establishing upon association to membranes in a mechanism similar to ENTH  $\alpha$ 0 helix formation (Miller, Mathiasen et al., 2015). Moreover, the overall protein structure of CALM proteins is similar to epsins, as proteins contain large unstructured regions, as well as protein interaction domains (Figure 11A).

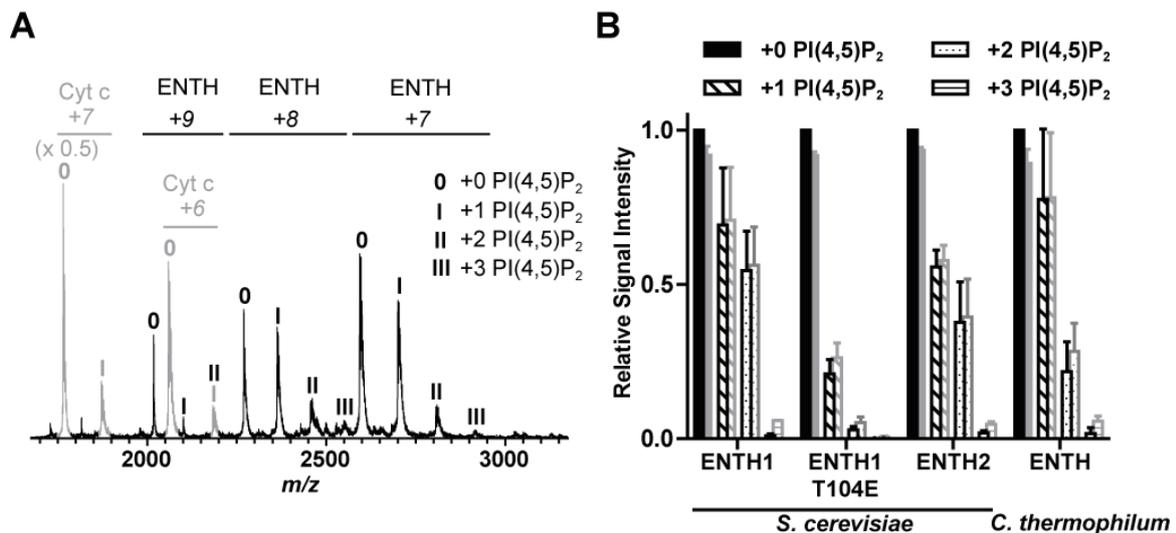
The Huntingtin interacting protein 1 related (Hip1R) and its yeast homolog Sla2 are part of another subfamily of clathrin adaptor proteins that contain an ANTH domain. Both share a similar domain structure with an N-terminal ANTH domain, followed by a coiled-coil domain for clathrin interaction and a C-terminal talin-like domain that binds to filamentous actin (Brett, Legendre-Guillemain et al., 2006) (Figure 11A). Alike epsins, Hip1R/Sla2 function as an interaction hub between plasma membrane, clathrin and actin (Engqvist-Goldstein, Warren et al., 2001, Gottfried, Ehrlich et al., 2010). Studies disclosed homodimerization via coiled-coil domains for both Hip1R and Sla2 (Wilbur, Chen et al., 2008, Yang, Cope et al., 1999).

Furthermore, several pieces of evidence display an interaction between ENTH-containing proteins and ANTH-containing proteins. For instance, recombinant yeast ENTH- and ANTH domains were co-immunoprecipitated *in vitro* (Skruzny, Brach et al., 2012). Results were later confirmed for murine full-length proteins (Messa, Fernandez-Busnadiego et al., 2014). Critically, the interaction was proven to be dependent on the presence of PI(4,5)P<sub>2</sub>. Skruzny, Desfosses et al. (2015) showed by cryo-ET the formation of large lattices of co-assembled ENTH and ANTH domains from *S. cerevisiae* on giant unilamellar vesicles. Although the physiological relevance remains uncertain, it is a clear indication that the formation of larger co-assemblies of ENTH, ANTH and PI(4,5)P<sub>2</sub> is possible. A first native MS analysis of ENTH and ANTH domains in presence of dioctanoyl PI(4,5)P<sub>2</sub> in solution revealed an ENTH:ANTH:PI(4,5)P<sub>2</sub> complex stoichiometry of 8:8:22±1 (Skruzny, Desfosses et al., 2015).

## 4.2 Results and Discussion

### 4.2.1 Native MS Reveals Cooperative Binding of 2 PI(4,5)P<sub>2</sub> to Epsin ENTH Domains

In order to produce a more complete description of the interplay of membrane-binding domains of clathrin adaptor proteins and PI(4,5)P<sub>2</sub> by native MS, in a first set of experiments the interaction of ENTH domains with PI(4,5)P<sub>2</sub> was analysed. Therefore, ENTH domains of *S. cerevisiae* epsin-1 (ENTH1) and epsin-2 (ENTH2) as well as *Chaetomium thermophilum* (*C. thermophilum*) epsin were used. Proteins were mixed with PI(4,5)P<sub>2</sub> and the reference protein cytochrome c and lipid binding was analysed in a direct MS approach (Sun, Kitova et al., 2006). In native mass spectra ENTH domains with up to 3 attached PI(4,5)P<sub>2</sub> molecules were detected, while cytochrome c showed only 1 PI(4,5)P<sub>2</sub> attachment (Figure 12A).



**Figure 12: Two PI(4,5)P<sub>2</sub> molecules bind to epsin ENTH domains.** For analysing PI(4,5)P<sub>2</sub> binding to epsin ENTH domains, proteins were measured in presence of the phospholipid PI(4,5)P<sub>2</sub> and cytochrome c as reference. Spectra showed free cytochrome c and attachment of 1 PI(4,5)P<sub>2</sub> (grey), while ENTH domains (here from epsin of *C. thermophilum*) bound 0-3 PI(4,5)P<sub>2</sub> molecules. As annotated, the +7 charge state of cytochrome c is downsized to 50 % of its original size for better visibility of the remaining signals. **B**) Signal intensities from mass spectra were summed over all charge states (back) and corrected for unspecific PI(4,5)P<sub>2</sub> attachment based on the ratio of bound/unbound reference protein (front). Data of at least three independent measurements were normalized to the corrected signal of unbound ENTH and the averages of the relative signal intensities were plotted with their standard deviations. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

Signal intensities of unbound, 1-PI(4,5)P<sub>2</sub>-bound, 2-PI(4,5)P<sub>2</sub>-bound, 3-PI(4,5)P<sub>2</sub> bound ENTH domains were summed over all charge states. The ratio of bound/unbound reference protein was considered for the correction of unspecific clustering events (**Figure 12B**). At the used concentrations (10 μM ENTH domain, 60 μM PI(4,5)P<sub>2</sub>), the largest fraction of ENTH was present in an unbound state, also after signal correction. Moreover, signals showing +1-PI(4,5)P<sub>2</sub> and +2-PI(4,5)P<sub>2</sub> attachments were found. The apparent third PI(4,5)P<sub>2</sub> binding from uncorrected spectra was within the range of the standard deviation after correction for unspecific clustering effects.

Macroscopic dissociation constants were directly derived from the spectrum as described (El-Hawiet, Kitova et al., 2012, Sun, Kitova et al., 2006) (see 1.6.2 Protein-Ligand interactions, Equation (16), page 15). Determined affinities are in the range of 50-150 μM for the first binding event and 45-240 μM for the second binding event (**Table 1**). Measurements at the MS-optimized ammonium acetate concentration of 300 mM and at the more physiological ionic strength of 160 mM ammonium acetate gave comparable results. Lower dissociation constants were determined by different biophysical techniques. Although some degree of variation between different methods is not surprising, especially the nanomolar dissociation constants (e.g. rat epsin ENTH: 330 ± 60 nM at pH 8.0) measured by SPR indicate a higher affinity than determined by native MS (Hom, Vora et al., 2007). In the SPR study, the overall binding of mammalian ENTH domains to PI(4,5)P<sub>2</sub>-enriched membranes was probed (Hom, Vora et al., 2007) while here, individual PI(4,5)P<sub>2</sub> binding events to fungal proteins were investigated. Although gentle experimental conditions were used, dissociation of a fraction of PI(4,5)P<sub>2</sub> throughout the measurement cannot be completely excluded. Moreover, the used approach for direct deduction of affinities from mass spectra is less accurate than titration methods.

**Table 1: Macroscopic  $K_D$  values (standard deviation, N=3) of PI(4,5)P<sub>2</sub> binding to different ENTH domains.** Unless stated differently, measurements were performed in 300 mM ammonium acetate, pH 8.0. The sample marked with an asterisk was measured at the more physiological ion concentration of 160 mM ammonium acetate, pH 8.0. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.

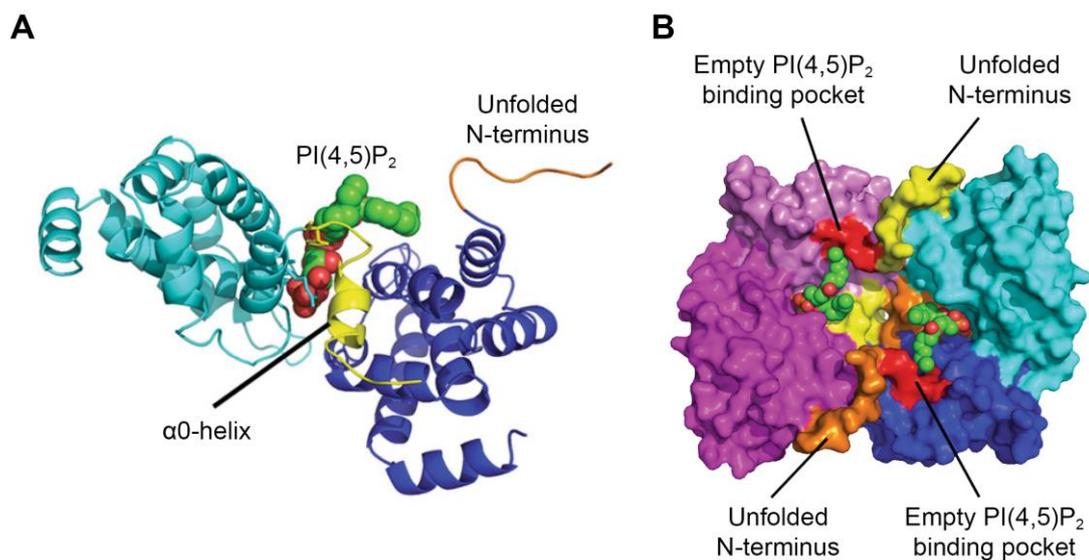
Protein	Organism	$K_{D1}$ / $\mu$ M	$K_{D2}$ / $\mu$ M
ENTH1	<i>S. cerevisiae</i>	83 (20)	65 (2)
ENTH1*	<i>S. cerevisiae</i>	100 (50)	52 (6)
ENTH1 T104E	<i>S. cerevisiae</i>	280 (80)	450 (110)
ENTH2	<i>S. cerevisiae</i>	100 (10)	82 (21)
ENTH	<i>C. thermophilum</i>	74 (23)	200 (40)

Comparing the values for  $K_{D1}$  and  $K_{D2}$  revealed a surprisingly high affinity for the second PI(4,5)P<sub>2</sub> binding event. For instance, *S. cerevisiae* ENTH1 has a higher affinity for the second binding event than for the first one. Considering that two binding sites are available for the first binding, while the second binding event is opposed by two dissociation pathways (Figure 3), solely for stochastic reasons higher  $K_{D2}$  values were expected. The low  $K_{D2}$  values point to an allosteric mechanism. This hypothesis was further assessed by insertion of the determined dissociation constants into equations (17) and (18) (see 1.6.2 Protein-Ligand interactions, page 16), which describe the relation of microscopic and macroscopic dissociation constants of two independent binding sites (Klotz, 1985). Even under consideration of the standard deviations of the determined macroscopic dissociation constants, no set of microscopic site constants was found that satisfies the equations. The numbers suggest positive cooperativity between the two PI(4,5)P<sub>2</sub> binding sites in ENTH domains, meaning that binding of one PI(4,5)P<sub>2</sub> increases the affinity for the second PI(4,5)P<sub>2</sub> binding event.

Further evidence for this cooperative binding mechanism was obtained from X-ray crystallography. A crystal structure of two ENTH domains of *S. cerevisiae* epsin-2 that surround one PI(4,5)P<sub>2</sub> molecule was determined to a resolution of 3.4 Å (Figure 13, Table S4). PI(4,5)P<sub>2</sub> is bound to the binding site that was previously identified for ENTH from rat epsin-1 in complex with the phosphorylated inositol head group IP<sub>3</sub> (Ford, Mills et al., 2002). The N-terminus is formed to an  $\alpha$ 0-helix as reported before (Ford, Mills et al., 2002). By contrast, the electron density for the N-terminus of the second ENTH domain revealed an extended conformation and the absence of an  $\alpha$ 0-helix. Instead, residue T104 is involved in PI(4,5)P<sub>2</sub> interaction that was previously reported to be crucial for the interaction of ENTH and ANTH

domains (Skruzny, Brach et al., 2012, Skruzny, Desfosses et al., 2015). The residue was described to be evolutionarily conserved and important for ENTH domain function (Aguilar, Longhi et al., 2006).

A T104E mutant of the *S. cerevisiae* epsin-1 ENTH domain was generated and analysed by native MS to characterize the role of the residue in PI(4,5)P<sub>2</sub> binding. The ratio of ENTH domains that bound two PI(4,5)P<sub>2</sub> molecules was drastically reduced and only marginally above the error range after correction for unspecific clustering (Figure 12B). Resulting from the altered ratios of PI(4,5)P<sub>2</sub>-bound and unbound ENTH domains, also the calculated dissociation constants reflect the decreased PI(4,5)P<sub>2</sub>-binding ability (Table 1).



**Figure 13: Crystal structure of two ENTH domains from *S. cerevisiae* epsin-2 binding one PI(4,5)P<sub>2</sub> molecule.** **A)** Ribbon diagram of the building block of two ENTH domains (cyan, blue) surrounding a PI(4,5)P<sub>2</sub> molecule (green and red spheres). One N-terminus is folded to an  $\alpha 0$ -helix (yellow), the other N-terminus is unfolded (orange). **B)** A surface presentation of a symmetry related ENTH tetramer (two building blocks: cyan/blue and magenta/violet) illustrates that bound PI(4,5)P<sub>2</sub> molecules (green and red spheres) are in plane with unfolded N-termini (yellow, orange). Empty PI(4,5)P<sub>2</sub> binding pockets (Y16, R24, R62, H72) are highlighted in red. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

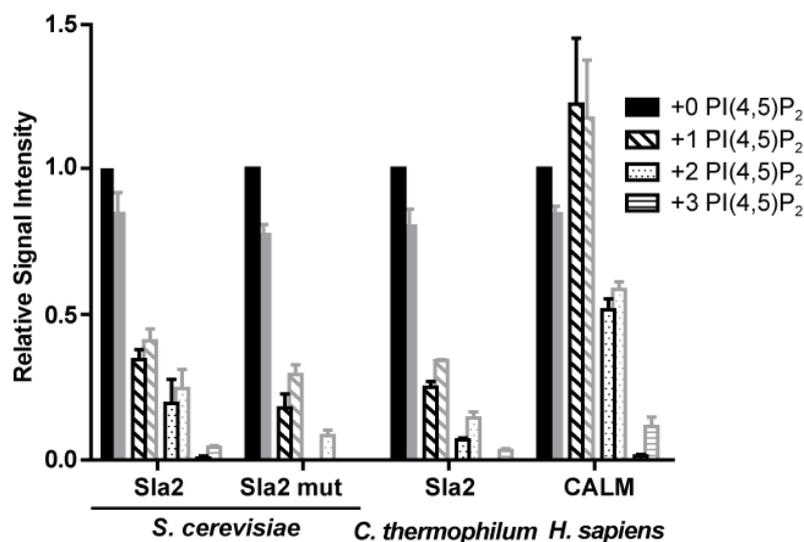
The buried surface area of PI(4,5)P<sub>2</sub> sandwiched by the two *S. cerevisiae* ENTH2 domains was analysed with the software PISA (Krissinel, 2015). The phosphatidylinositol head group has a total solvent accessible surface area of 488 Å<sup>2</sup>. The larger share of the head group (306 Å<sup>2</sup> or 63 %) is buried by the ENTH2 domain that has formed an  $\alpha 0$ -helix. Only 20 % (95 Å<sup>2</sup>) of the head group’s surface is covered by the ENTH2 domain that has not formed an  $\alpha 0$ -helix. The

$\alpha$ 0-helix itself accounts for 56 % (1020 Å<sup>2</sup>) of the total interface (1820 Å<sup>2</sup>) between the two ENTH domains.

In the crystal structure, a symmetry-related ENTH tetramer was found with two  $\alpha$ 0-helices that stack against each other (Figure 13B). The two enclosed PI(4,5)P<sub>2</sub> molecules are located in one plane so that an integration into the plasma membrane is conceivable, while the C-termini are pointing to the opposite direction. In the tetramer, also two unfolded N-termini with neighbouring positively charged patches (residues K14, R24, R62, and H72) are found are possibly set to bind two more PI(4,5)P<sub>2</sub> molecules. Thus, it can be hypothesized that the crystal structure depicts an intermediate state and that after additional PI(4,5)P<sub>2</sub> binding and  $\alpha$ 0-helix formation potentially larger ENTH clusters arise.

#### 4.2.2 ANTH Domains Bind Two PI(4,5)P<sub>2</sub> Molecules

ANTH domains from *S. cerevisiae* Sla2, *C. thermophilum* Sla2 and *H. sapiens* CALM proteins were subjected to PI(4,5)P<sub>2</sub> binding analysis as described before for ENTH domains. All wild type ANTH domains showed binding of two PI(4,5)P<sub>2</sub> molecules after correction for unspecific clustering. (Figure 14).



**Figure 14:** Two PI(4,5)P<sub>2</sub> molecules bind to wild type ANTH domains. Signal intensities from mass spectra were summed over all charge states (back) and corrected for unspecific PI(4,5)P<sub>2</sub> attachment based on the ratio of bound/unbound reference protein (front). Data of at least three independent measurements were normalized to the corrected signal of unbound ENTH and the averages of the relative signal intensities were plotted with their standard deviations. Sla2 mut refers to an ANTH domain with K13A, K24A, K26A, H27A mutations in the canonical PI(4,5)P<sub>2</sub> binding site. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

Derived dissociation constants for fungal proteins were in the range of 100-350  $\mu\text{M}$  with a slight trend towards higher affinities for the first binding event (Table 2). Yet for human CALM ANTH domains the first PI(4,5)P<sub>2</sub> binding event had a markedly higher affinity. The occupancy ratio revealed the highest abundance for CALM ANTH + 1PI(4,5)P<sub>2</sub> complexes at the used protein (10  $\mu\text{M}$ ) and ligand (60  $\mu\text{M}$ ) concentrations (Figure 14) and the derived dissociation constant was in the range of 45  $\mu\text{M}$  (Table 2). Differences in PI(4,5)P<sub>2</sub> affinity for ANTH domains from CALM and Sla2 can be explained by different phospholipid binding mechanisms. Crystal structures revealed a PI(4,5)P<sub>2</sub> binding mechanism for CALM ANTH domains that involves the formation of a helix  $\alpha 0$  as reported for ENTH domains, which was not found for ANTH domains of the Sla2 subfamily (Miller, Mathiasen et al., 2015). PI(4,5)P<sub>2</sub> binding affinities of ENTH domains were generally slightly higher than for Sla2 ANTH domains. As discussed for ENTH domains, binding constants determined by SPR generally reported a higher affinity (e.g.  $1.20 \pm 0.12 \mu\text{M}$  at pH 8.0 for human CALM ANTH) (Hom, Vora et al., 2007). Yet the overall trend with lower PI(4,5)P<sub>2</sub> binding affinities for ANTH domains than for ENTH domains was confirmed. Possible reasons for deviations between the methods apply as discussed for ENTH domains.

**Table 2: Macroscopic  $K_D$  values (standard deviation, N=3) of PI(4,5)P<sub>2</sub> binding to different ANTH domains.** Unless stated differently, measurements were performed in 300 mM ammonium acetate pH 8.0. Samples marked with an asterisk were measured at the more physiological concentration of 160 mM ammonium acetate pH 8.0. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.

Protein	Organism	$K_{D1}$ / $\mu\text{M}$	$K_{D2}$ / $\mu\text{M}$
Sla2 ANTH	<i>S. cerevisiae</i>	158 (19)	130 (90)
Sla2 ANTH*	<i>S. cerevisiae</i>	110 (2)	160 (50)
Sla2 ANTH	<i>C. thermophilum</i>	211 (7)	270 (80)
Sla2 ANTH mut	<i>S. cerevisiae</i>	300 (70)	-
CALM ANTH	<i>H. sapiens</i>	44 (8)	125 (14)
CALM ANTH*	<i>H. sapiens</i>	45 (8)	157 (20)

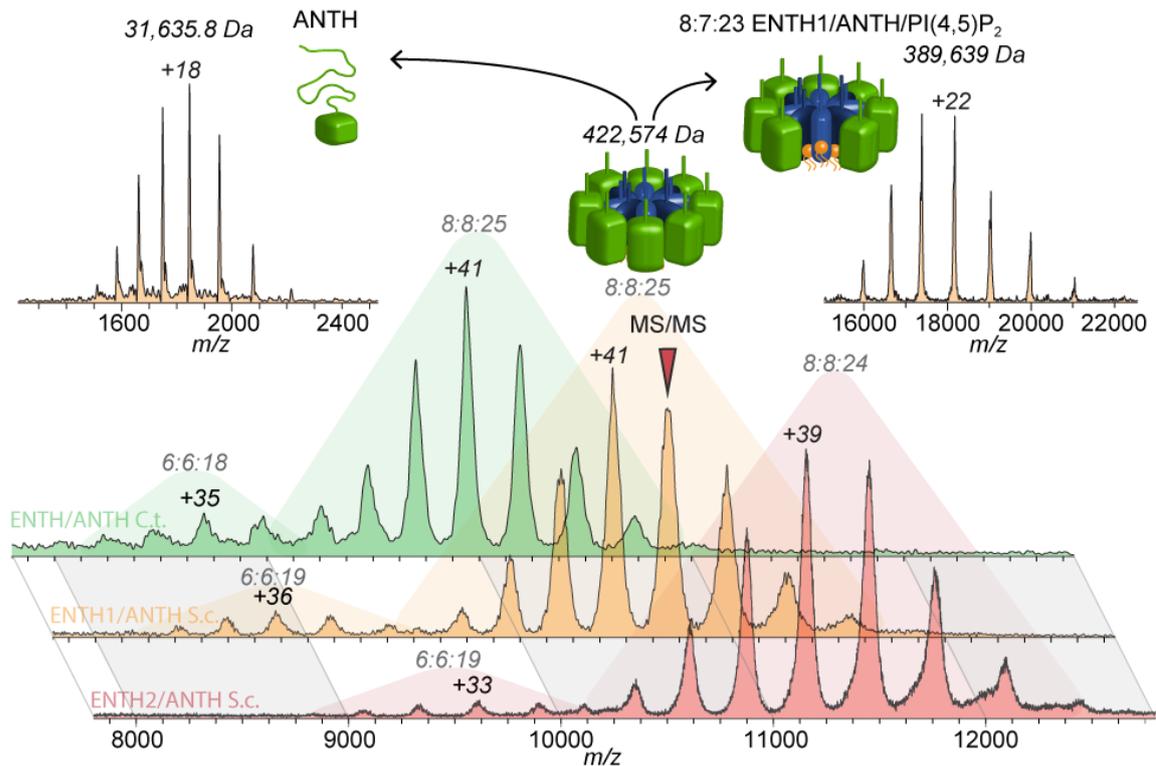
In order to precisely characterize the contributions of the two individual PI(4,5)P<sub>2</sub> binding sites, the known binding site of *S. cerevisiae* Sla2 ANTH was altered by the introduction of four mutations (K13A, K24A, K26A, H27A). As hypothesized, these mutations led to the abrogation of PI(4,5)P<sub>2</sub> binding to this site as the mutant ANTH domain unveiled only binding of one PI(4,5)P<sub>2</sub> molecule after correction for unspecific binding (Figure 14). In this case, the

determined dissociation constant of  $300 \pm 70 \mu\text{M}$  was the microscopic constant of PI(4,5)P<sub>2</sub> binding to the thus far unidentified binding site. However, as for PI(4,5)P<sub>2</sub> binding to ENTH domains, no additional site constant was found that satisfies the relation between macroscopic and microscopic dissociation constants. Hence, also a cooperative mechanism for PI(4,5)P<sub>2</sub> binding to Sla2 ANTH domains is proposed.

#### 4.2.3 Fungal ENTH and Sla2 ANTH Domains Form Ordered Assemblies

In a previous study, complex formation of ENTH and ANTH domains from *S. cerevisiae* proteins epsin-1 and Sla2 in presence of PI(4,5)P<sub>2</sub> was analysed by native MS (Skruzny, Desfosses et al., 2015). The use of tagged and untagged proteins allowed the identification of an 8:8 (ENTH:ANTH) protein stoichiometry. Here, complexes of different ENTH and ANTH domains were subjected to a more detailed native MS analysis. ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes from *C. thermophilum* proteins epsin and Sla2 were analysed, as well as complexes with domains originating from *S. cerevisiae* protein epsin-1 or epsin-2 (both ENTH) and Sla2 (ANTH). Irrespective of the proteins, all studied ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes showed the same stoichiometries, signal ratios and dissociation pathways in CID MS/MS studies. Complex masses, statistical errors, ranges of PI(4,5)P<sub>2</sub> numbers and an average FWHM value rating the spectral resolution are presented in **Supplementary Table S3**.

As reported before, the main MS signal originated from an 8:8:~24 (ENTH:ANTH:PI(4,5)P<sub>2</sub>) complex with a mass of approximately 422 kDa (**Figure 15**, Skruzny, Desfosses et al. (2015)). Spectral resolution was insufficient to directly determine the PI(4,5)P<sub>2</sub> number from the peak fine structure. For the assignment of the number of PI(4,5)P<sub>2</sub> molecules in a complex, respective multiples of the measured masses of ENTH and ANTH domains were subtracted from the measured complex mass. The residual mass was attributed to lipid binding and consequently divided by the PI(4,5)P<sub>2</sub> mass to yield the number of PI(4,5)P<sub>2</sub> molecules in the complex. The uncertainty was expressed by division of the FWHM value by the PI(4,5)P<sub>2</sub> mass. All analysed 8:8 ENTH:ANTH complexes contained 24-25 PI(4,5)P<sub>2</sub> molecules with an uncertainty of 3-4 PI(4,5)P<sub>2</sub> (**Supplementary Table S3**).



**Figure 15: ENTH:ANTH:PI(4,5)P<sub>2</sub> complex formation in fungi.** 6:6:~18 and 8:8:~24 (ENTH:ANTH:PI(4,5)P<sub>2</sub>) complex stoichiometries were identified in native MS measurements. A cartoon representation of the most prominent complex is illustrated with ANTH in green, ENTH in blue and PI(4,5)P<sub>2</sub> in orange. Complexes from *C. thermophilum* (green spectrum) and *S. cerevisiae*, the latter ones containing either ENTH from epsin-1 (yellow spectrum) or epsin-2 (red spectrum), revealed the same stoichiometries, signal ratios and dissociation pathways in CID MS/MS. The dissociation of the +40 charged 8:8:25±3 ENTH1:ANTH:PI(4,5)P<sub>2</sub> complex into a likely unfolded ANTH domain (top left spectrum, adjacent cartoon shows green ANTH domain) and a residual 8:7:23±1 complex (top right spectrum and cartoon) is depicted. Stoichiometries (ENTH:ANTH:PI(4,5)P<sub>2</sub>), charge states of main peaks, and average experimental masses are annotated. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

In native mass spectra, an additional peak series with lower signal intensities was detected in the range between 8,000 and 10,000 *m/z* (Figure 15) to which a mass of 317 kDa was assigned. Applying the same procedure as for previously described complexes led to 6:6:~19 ENTH:ANTH:PI(4,5)P<sub>2</sub> complex stoichiometries with an uncertainty of 2-3 PI(4,5)P<sub>2</sub> molecules.

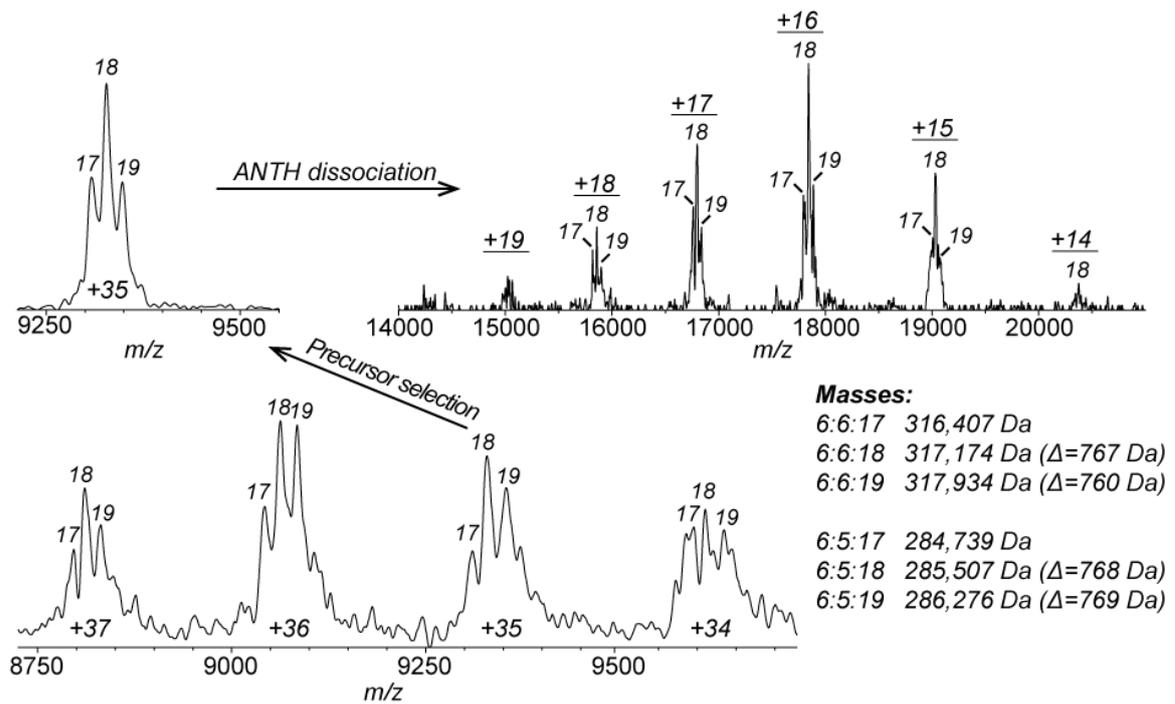
Based on the finding that ENTH and ANTH domains contain two PI(4,5)P<sub>2</sub> binding sites each, the observed number of PI(4,5)P<sub>2</sub> molecules in ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes is surprisingly low. For instance, in the 8:8:~24 complex 32 PI(4,5)P<sub>2</sub> molecules were expected. Another special feature of these complexes is a remarkably low number of charges after nano electrospray ionisation. For a protein complex with a size of 422 kDa, the expected number of

charges would be around +51 (see 1.6.1 Solution Structures versus Gas-phase Structures, Equation (13), page 13), the measured charges for the protein-phospholipid complexes is in the range of +40. A possible reason for the deviation is an exceptionally compact complex structure or alterations of the ESI process caused by the presence of phospholipids with their negatively charged head groups.

To obtain further information on complex composition and topology, complexes were subjected to CID MS/MS analysis. One ANTH domain dissociated from 8:8:~24 complexes, resulting in a residual 8:7:~23 ENTH:ANTH:PI(4,5)P<sub>2</sub> complex with a mass of approximately 390 kDa (**Figure 15**). Likewise, ANTH dissociation from 6:6:~19 complexes was detected (**Figure 16**).

Although the CID mechanism is not conclusively clarified, a general preference for the dissociation of monomeric proteins from the complex periphery was reported (Benesch, 2009). In addition, small proteins typically require lower energy for unfolding and dissociation than large proteins. In the present case, the dissociation pattern with the exclusive dissociation of ANTH domains suggests that in both complex stoichiometries, the larger ANTH domains are located in the periphery of the complex and are therefore more susceptible to the CID process than ENTH domains.

Higher resolution spectra from MS and CID MS/MS analysis of 6:6:~19 ENTH2:ANTH:PI(4,5)P<sub>2</sub> complexes from *S. cerevisiae* disclosed further details on phospholipid numbers and the role of the lipids in the investigated complex. In this case, it was possible to assign the observed fine structure in mass spectra to a specific PI(4,5)P<sub>2</sub> content (**Figure 16**). Using the same procedure as described before, 6:6 ENTH:ANTH complexes were specified to contain 17, 18 or 19 PI(4,5)P<sub>2</sub> molecules. Since the measured mass of single PI(4,5)P<sub>2</sub> adducts at a protein was roughly 744 Da, but the periodical adduct pattern reveals mass shifts of 760 to 769 Da, the presence of one additional cation adduct per PI(4,5)P<sub>2</sub> binding event can be speculated. The resolution is insufficient to clearly identify the cation among the candidates that are present in substantial amounts in the measuring solution, such as sodium and ammonia.

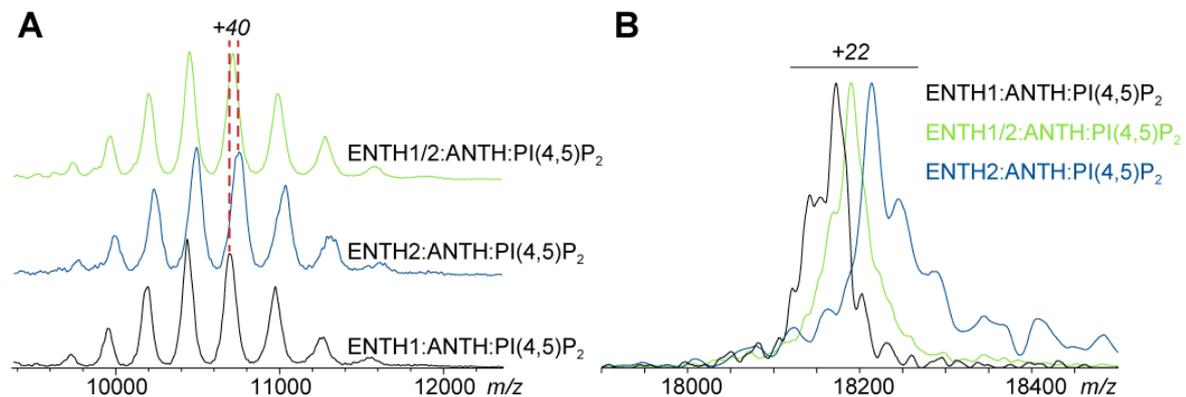


**Figure 16: Lipid stoichiometry of *S. cerevisiae* 6:6 ENTH2:ANTH complexes revealed by CID MS/MS.** Native mass spectra contained a peak fine structure that indicated varying numbers of PI(4,5)P<sub>2</sub> molecules in the complexes. By using the measured masses of *S. cerevisiae* ENTH2 and ANTH domains alone and the theoretical mass of PI(4,5)P<sub>2</sub>, the precise stoichiometry was calculated and exhibited 17 to 19 PI(4,5)P<sub>2</sub> molecules. The mass difference ( $\Delta$ ) to the complex species with one PI(4,5)P<sub>2</sub> molecule less was calculated and listed. The +35 charge state was selected as a precursor ion for CID MS/MS experiments. Increased collision voltages led to dissociation of one ANTH domain, masses of remaining 6:5 ENTH2:ANTH complexes were determined and revealed unaltered lipid stoichiometry and ratio. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.

Remarkably, a very similar pattern of PI(4,5)P<sub>2</sub> molecules was found after dissociation of an ANTH domain in CID MS/MS experiments, even with comparable signal intensity ratios. Both PI(4,5)P<sub>2</sub> molecules that interacted with the ejected ANTH domain remained bound in the complex, implying that they need to be positioned at protein-protein interfaces. In other words, one PI(4,5)P<sub>2</sub> molecule is bound by two ENTH or ANTH domains. For that reason, less PI(4,5)P<sub>2</sub> molecules were found in ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes than expected from PI(4,5)P<sub>2</sub> binding studies. Moreover, all other PI(4,5)P<sub>2</sub> molecules remain bound in the complex, despite the energy that is deposited on the complex and potentially caused protein unfolding.

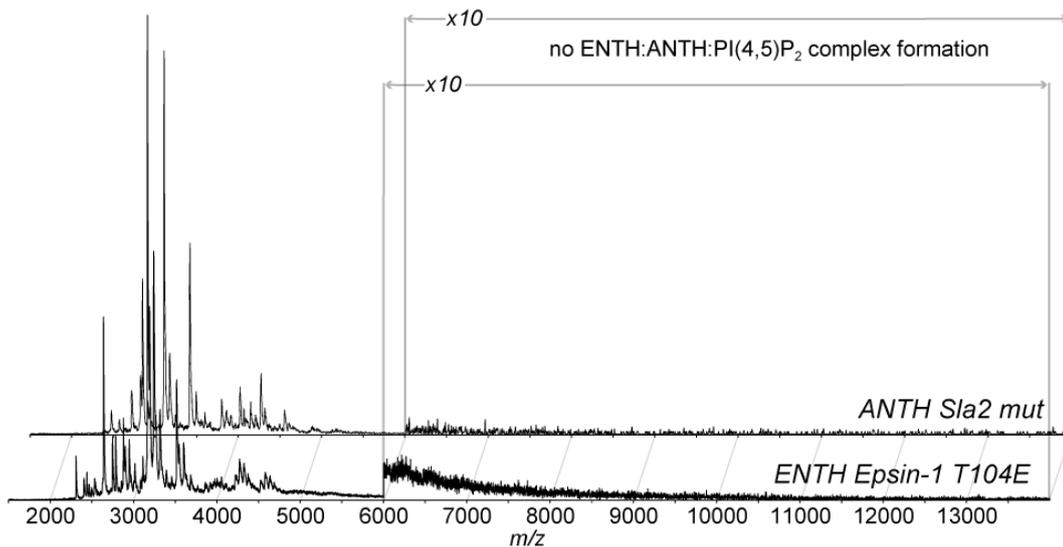
No difference in ENTH:ANTH:PI(4,5)P<sub>2</sub> complex formation was observed for ENTH domains from *S. cerevisiae* proteins epsin-1 and epsin-2. However, the used constructs differ in mass,

with ENTH2 being 111 Da heavier. To further investigate a putative difference in complex formation, ENTH1 and ENTH2 domains were mixed in an equimolar ratio prior to ENTH:ANTH:PI(4,5)P<sub>2</sub> complex formation. Complexes assembled in this manner had an intermediate mass in MS and MS/MS analyses compared to complexes that contained exclusively ENTH1 or ENTH2 (**Figure 17, Supplementary Figure S3**). Consequently, ENTH1 and ENTH2 domains from *S. cerevisiae* possess the ability to form mixed complexes *in vitro*. Furthermore, spectra showed no preference for either ENTH variant, suggesting similar binding affinities for complex formation.



**Figure 17: ENTH:ANTH:PI(4,5) assembled with mixed ENTH domains from *S. cerevisiae* epsin-1 and epsin-2. A)** Due to different molecular weights ( $\Delta m = 111$  Da) of the used ENTH1 and ENTH2 constructs, 8:8:~25 ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes with different ENTH forms have different masses. The ENTH1:ANTH:PI(4,5)P<sub>2</sub> complex (black) has a 1 kDa lower mass than the ENTH2:ANTH:PI(4,5)P<sub>2</sub> complex (blue). When ENTH1 and ENTH2 domains were mixed in equimolar ratios and used for complex assembly an intermediate mass (green) was observed. Peak tops of the +40 ions of exclusively ENTH1- or ENTH2-containing complexes are annotated in red to highlight the intermediate position of the mixed form. **B)** +40 8:8:~25 ENTH:ANTH:PI(4,5)P<sub>2</sub> ions were subjected to CID MS/MS analysis. For the product ions (here: charge state +22) an intermediate mass for mixed ENTH1/2-containing complexes in comparison to homogenous ENTH1- or ENTH2-containing complexes was observed. For the full *m/z*-range of CID MS/MS spectra, see **Supplementary Figure S3**. Panel A was reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

The importance of ENTH residue T104 for PI(4,5)P<sub>2</sub> binding was highlighted in experiments with a *S. cerevisiae* ENTH1 T104E mutant (see 4.2.1 Native MS Reveals Cooperative Binding of 2 PI(4,5)P<sub>2</sub> to Epsin ENTH Domains, page 44). Similarly, ENTH:ANTH:PI(4,5)P<sub>2</sub> complex formation was probed for the T104E mutant. Spectra revealed no complex formation after mixing the mutant ENTH domain with ANTH domains and PI(4,5)P<sub>2</sub> (Figure 18). Equally, the *S. cerevisiae* Sla2 ANTH domain with four mutations (Y16A, R24A, R62A, H72A) that abrogated PI(4,5)P<sub>2</sub> binding to the canonical PI(4,5)P<sub>2</sub> binding site showed no formation of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes upon mixing with ENTH2 domains of *S. cerevisiae* and PI(4,5)P<sub>2</sub> (Figure 18). Thus, it seems plausible that the reduced PI(4,5)P<sub>2</sub> binding competence of the mutated ENTH and ANTH domains impedes the formation of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes.

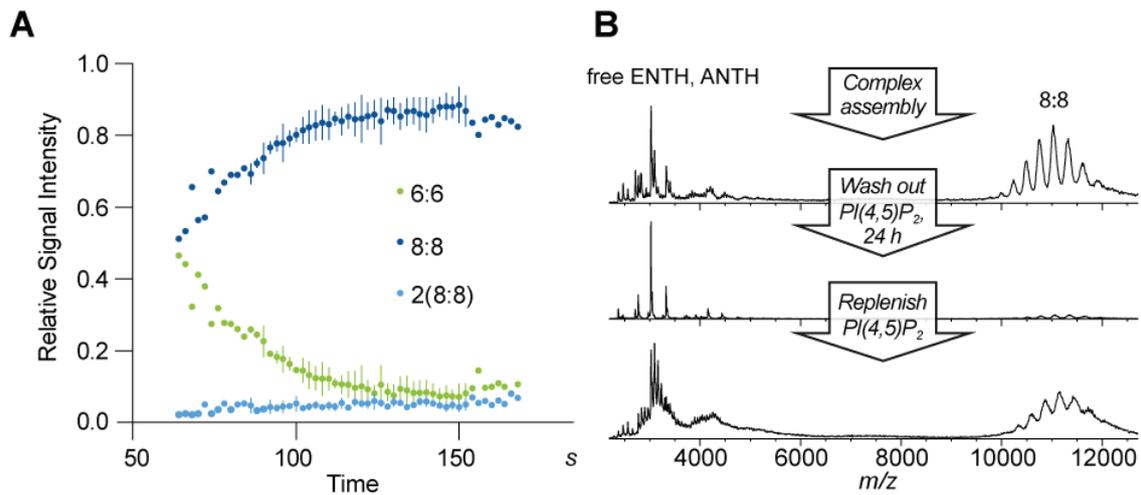


**Figure 18: Reduced PI(4,5)P<sub>2</sub> binding capacities of *S. cerevisiae* ENTH and ANTH mutant domains obviated complex formation.** No complex formation was observed in native MS after mixing of *S. cerevisiae* ENTH1 T104E domains and Sla2 ANTH domains with PI(4,5)P<sub>2</sub> (front). Likewise, *S. cerevisiae* Sla2 ANTH mutants (Y16A, R24A, R62A, H72A) did not form complexes with ENTH2 domains. The range above 6,000 *m/z* is 10x magnified to highlight the absence of large complexes. PI(4,5)P<sub>2</sub>-bound and unbound ENTH and ANTH domains of the respective domain variants are found in the lower *m/z*-range. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

#### 4.2.4 *ENTH:ANTH:PI(4,5)P<sub>2</sub> Complex Formation Dynamics*

The formation of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes was monitored over time by native MS. Therefore, data acquisition was started as quickly as possible after mixing of the components and sample loading into an electrospray capillary. Relative signal intensities of 6:6 (ENTH:ANTH) and 8:8 complexes were determined and plotted against time (**Figure 19A**). In addition, the signal intensity of the dimer of the 8:8 complex was monitored to control for unspecific protein aggregation. The first data points were acquired approximately 1 min after mixing and showed similar signal intensities for 6:6 and 8:8 complexes. Within the next minute, the signal of the 6:6 complex dropped, while the 8:8 signal increased. Since the signal of the 2(8:8) complex remained stable over the entire acquisition time, unspecific protein aggregation effects were ruled out. This time course of signal intensities suggests a distinct assembly pathway. After mixing of ENTH, ANTH domains and PI(4,5)P<sub>2</sub>, 6:6 complexes are formed and subsequently converted to 8:8 complexes. The latter complex stoichiometry is possibly the more stable form and the 6:6 complex an assembly intermediate. Because of possible differences in ionization and ion transmission efficiencies, direct conversion of depicted ratios of relative signal intensities into protein complex abundances is not possible.

ENTH:ANTH:PI(4,5)P<sub>2</sub> complex assembly was also shown to be reversible (**Figure 19B**). For that purpose, complexes were assembled and a spectrum was acquired. Then, PI(4,5)P<sub>2</sub> was depleted from the measuring solution by washing with ammonium acetate buffer using centrifugal filter units. After one day of storage, mainly signals of monomeric ENTH and ANTH domains were observed. The low signal intensity of 8:8 complexes still present after 24 hours emphasises a high stability of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes with low dissociation rates. When the measuring solution was then supplemented with PI(4,5)P<sub>2</sub>, ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes formed once more, demonstrating the reversibility of complex formation.



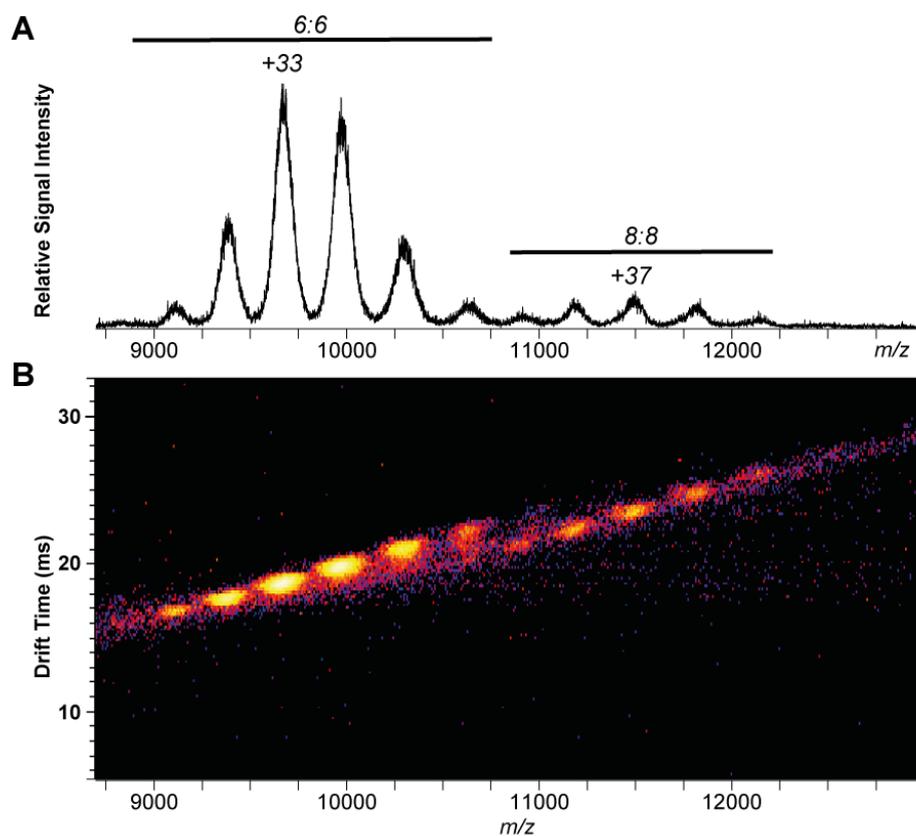
**Figure 19: Dynamics of fungal ENTH:ANTH:PI(4,5)P<sub>2</sub> assemblies.** **A)** A time course of ENTH2:ANTH:PI(4,5)P<sub>2</sub> complex formation was recorded by native MS. Components were mixed, injected into the electrospray capillary and the spectra were monitored over time. Relative signal intensities of complexes with 6:6 (green), 8:8 (dark blue) and 2(8:8) (light blue) ENTH:ANTH stoichiometries were determined and plotted against the progressed time after mixing. The signal of the 6:6 complex faded within the first 2 min after mixing, while the 8:8 signal increased. The signal intensity of the dimer of the 8:8 complex remained rather constant. Average data of three measurements, error bars illustrate the standard deviation for data points with N = 3. **B)** ENTH1:ANTH:PI(4,5)P<sub>2</sub> complex formation in *S. cerevisiae* is reversible. Complexes were assembled and a native mass spectrum was recorded (top). Then, PI(4,5)P<sub>2</sub> was washed out of the measuring solution using centrifugal filter units. Mass spectra acquired after 24 hours of storage exhibited considerably reduced signal intensities for assembled complexes (middle). Replenishing the measuring solution with PI(4,5)P<sub>2</sub> resulted in reformation of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes (bottom). Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

#### 4.2.5 IMMS Analysis of *S. cerevisiae* ENTH:ANTH:PI(4,5)P<sub>2</sub>

Complexes, built up from *S. cerevisiae* proteins ENTH2, ANTH as well as PI(4,5)P<sub>2</sub>, were further analysed by means of travelling wave ion mobility spectrometry (TWIMS). With this method, ions are separated according to their charge and conformation that both determine the transit time through a gas-filled chamber (see 1.8 Ion Mobility Mass Spectrometry, page 21). Subsequently, also the *m/z* ratio is measured, allowing the characterization of mass and conformation of analyte ions.

6:6 and 8:8 (ENTH:ANTH) complexes were identified as described for previous native MS measurements (Figure 15, Figure 20). However, the ratio of complex stoichiometries was different. In some cases, only 8:8 complexes were detected, whereas other measurements

displayed a tremendous shift towards 6:6 stoichiometries. In such cases, very high capillary voltages were necessary to generate a stable electrospray, indicating poor sample or electrospray capillary quality and normally resulting in a low spectral resolution. Another possible, yet less likely explanation is that non-ideal sample conditions impeded the transition to 8:8 complexes. In addition, transmission efficiencies of the used mass spectrometer (Synapt G2, Waters) generally favour smaller ions. Independently of the signal intensities, the signal shapes host information about the compactness or degree of unfolding of the ENTH:ANTH:PI(4,5)P<sub>2</sub> ions. Here, the signals of all charge states in the 2-dimensional IMMS plot have a rather compact shape and reveal no elongation along the y-axis, which is characteristic for narrow conformational ensembles.



**Figure 20:** Ion mobility mass spectrometry of *S. cerevisiae* ENTH2:ANTH:PI(4,5)P<sub>2</sub> complexes. **A)** 6:6 and 8:8 ENTH:ANTH complex stoichiometries were identified in mass spectra. **B)** Drift plots contain the drift time of the respective ions as additional information. Main peaks and complex stoichiometries are labelled in the mass spectrum.

Arrival time distributions of single charge states were extracted from the spectra. Peak tops (*drift times*) were used for the determination of CCSs of the respective ions (Table 3). The utilized TWIMS instrument does not allow a direct measurement of CCSs. Instead, standard calibration protein complexes were used to estimate corrected CCSs from a logarithmic fit ( $R^2 = 0.96$ , Supplementary Figure S4) as described (see 1.8 Ion Mobility Mass Spectrometry, page 23). The CCSs were in the range of 15,300 to 15,700 Å<sup>2</sup> for the 8:8 complex and in the range of 12,500 to 13,000 Å<sup>2</sup> for the smaller 6:6 complex (Table 3).

**Table 3: Measured drift times and derived CCS' for PI(4,5)P<sub>2</sub> containing 8:8 and 6:6 ENTH2:ANTH complexes from *S. cerevisiae*.** A plot showing the logarithmic fit of the calibrant ions is presented as Supplementary Figure S4.

Complex	Charge state $z$	Drift time (ms)	CCS' (Å <sup>2</sup> )
8:8	38	22.05	15,710
	37	23.15	15,570
	36	24.42	15,450
	35	25.88	15,340
6:6	35	16.58	13,040
	34	17.50	12,920
	33	18.41	12,770
	32	19.50	12,650
	31	20.78	12,540

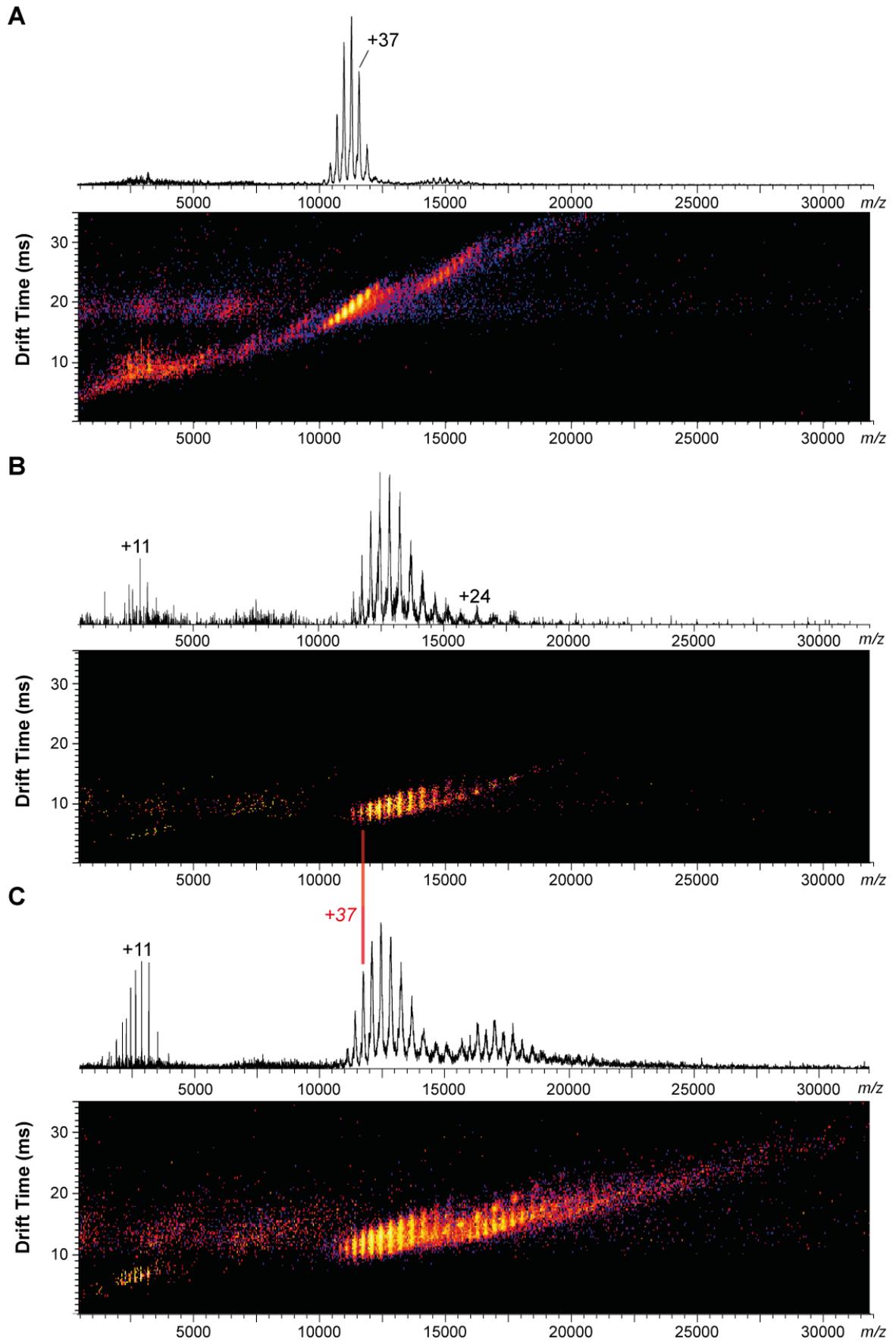
A comparison with previously determined CCSs of complexes with a comparable molecular weight allows a rough estimation of the compactness of these complexes. Glutamate dehydrogenase (GDH) hexamers from bovine liver have a mass of around 336 kDa and a CCS of 13,400 Å<sup>2</sup> (Bush, Hall et al., 2010) and are thus similar to ENTH:ANTH 6:6 complexes in respect of size and mass. Yet the observed number of charges is different for these complexes. While the +36 charge state was found to be most prominent in GDH spectra, the main charge state of ENTH:ANTH 6:6 complexes was +33. Based on the general relationship of surface area and number of charges (see 1.6.1 Solution Structures versus Gas-phase Structures, page 13), a compact complex structure of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes was expected from native MS measurements. However, these IMMS results indicate that complexes formed of clathrin adaptor proteins in presence of PI(4,5)P<sub>2</sub> are not extremely compact structures but rather gain exceptionally few charges during electrospray ionization in positive ion mode.

#### 4.2.6 SID MS of *S. cerevisiae* ENTH:ANTH:PI(4,5)P<sub>2</sub> Complexes

Colliding protein complexes with a solid surface instead of gas molecules can be of value for conformational studies, because more structural information of the precursor complex is preserved in the resulting dissociation products than after CID dissociation (Zhou, Jones et al., 2013). Therefore, ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes were assembled using ENTH1 domains of *S. cerevisiae* and ANTH domains of *S. cerevisiae* Sla2 and subjected to surface-induced dissociation (SID) MS and MS/MS analysis. The used SID-IMMS setup allowed the monitoring of drift times of SID product ions. Unfortunately, the access to the mass spectrometer equipped with a functional SID cell was very limited and thus the presented results need to be considered to be preliminary data.

In SID transmission mode, ions traverse the SID cell without being guided against the surface. The corresponding spectrum was in accordance with previously measured IMMS spectra and showed the described 6:6 (ENTH:ANTH) and 8:8 complexes, as well as a dimer of 8:8 complexes (Figure 21A).

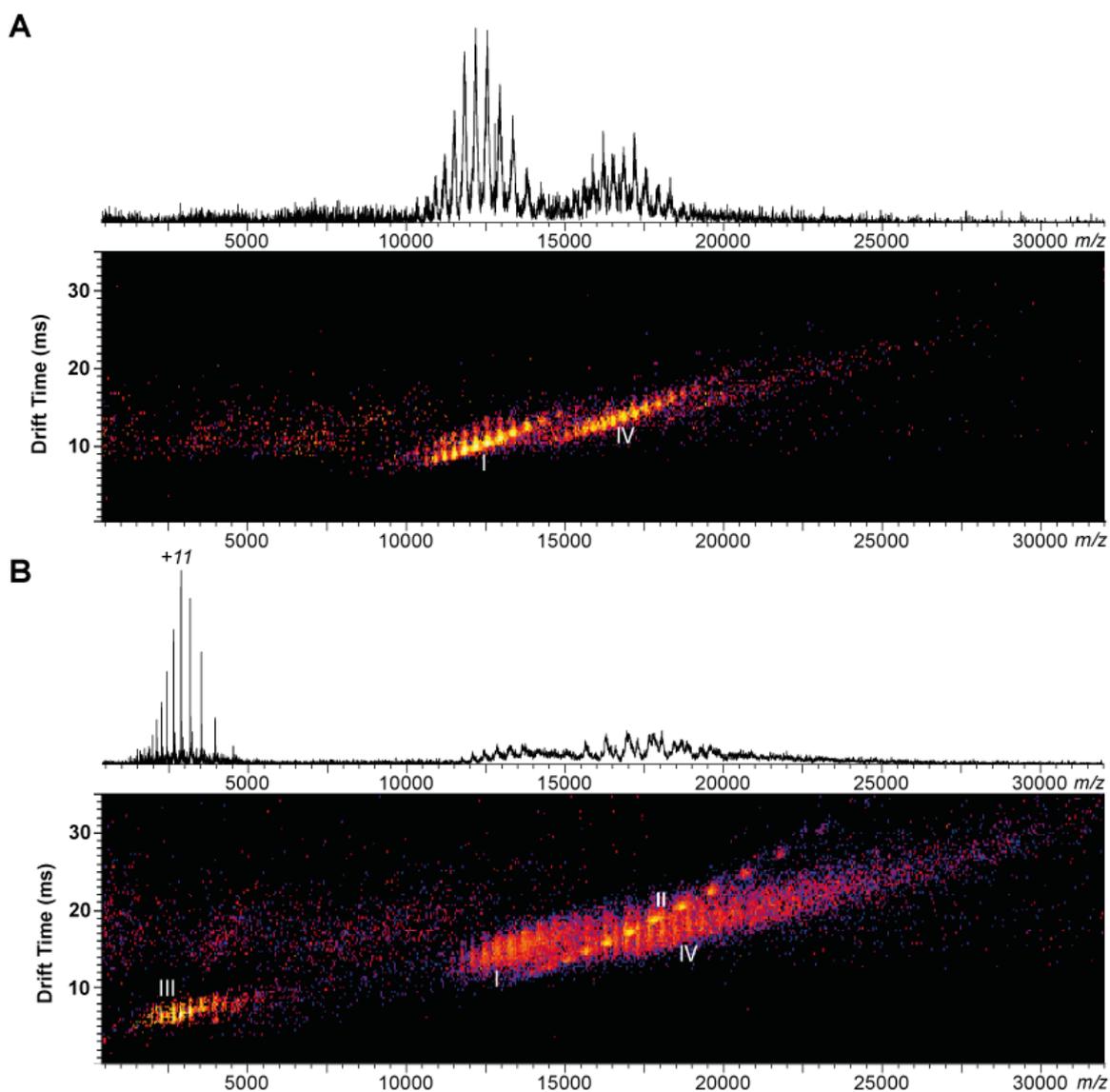
Next, charge state +37 was selected as precursor ion for SID MS/MS analysis. Ions were accelerated with 60 V SID collision energy and guided against the surface. The acquired spectrum exposed three different charge envelopes (Figure 21B). The first charge envelope corresponds to a series of charge-stripped precursor ions. The ENTH:ANTH stoichiometry of this complex is still 8:8 but the number of charges is reduced up to +29. This finding was unexpected and needs to be explained by a pronounced loss of residual buffer molecules and possibly PI(4,5)P<sub>2</sub> upon collision with the surface. Moreover, a clear reduction in mass was observed when comparing the molecular weight determined in flythrough mode (428.2 kDa, FWHM: 2.5 Da) and in 60 V SID collision energy (423.0 kDa, FWHM: 2.0 kDa) that cannot be attributed to the dissociation of a complete protein domain.



**Figure 21: Surface-induced dissociation MS of *S. cerevisiae* ENTH1:ANTH:PI(4,5)P<sub>2</sub>.** **A)** In transmission mode, no collisions with the surface occur. The resulting MS (top) and IMMS (bottom) spectra revealed the 8:8 (ENTH:ANTH) complex and low signals for the 6:6 complex and dimers of the 8:8 complex. **B)** The +37 charge state (highlighted by red line) of the 8:8 complex was selected for SID MS/MS analysis at 60 V collision energy. The spectrum exhibited a peak series of charge-stripped precursor ions (I). In addition, low intensity signals of a 8:7 complex (II) and dissociated ANTH domains (III) were found. **C)** In order to increase signal intensities, the MS profile was set to 10,000-15,000 *m/z*. These settings allowed a quadrupole transmission of ions > 10,000 *m/z*. Consequently, also 2(8:8) complexes (IV) were identified in the spectrum. Note that different IMMS settings for transmission and SID mode do not allow a direct comparison of drift times between panels A and B, C.

In the range of 15,500 to 19,000 *m/z*, a second charge envelope occurred that originated from a 8:7 complex, while the associated ANTH monomer was found in the low mass range. The total signal intensity was very low during these measurements. For that reason, the quadrupole settings were changed, such that the MS profile allowed transmission of precursor ions above 10,000 *m/z* instead of selecting only charge state +37. As expected, the resulting spectrum contained also the dimer of the 8:8 complex. However, the previously described dissociation of ANTH, resulting in a residual 8:7 complex, becomes more obvious (**Figure 21C**). It was shown that the charge partitioning in SID experiments is less asymmetric than in CID measurements, where a single dissociating protein takes a large fraction of the charges that were present in the precursor ion (Wysocki, Jones et al., 2008). Here, the main peak of the ejected ANTH domains has +11 charges. In earlier described CID MS/MS experiments (**Figure 15**), the main peak of the ejected ANTH ions was +18, clearly indicating that the underlying dissociation mechanism is SID.

SID collision energies were changed to lower (30 V) and higher (90 V) voltages. When only 30 V were applied, no dissociation of ANTH domains was observed and exclusively the previously described 8:8 and 2(8:8) complexes were identified (**Figure 22A**). Conversely, the application of 90 V SID collision energy resulted in most abundant signals for dissociated ANTH domains and residual 8:7 complexes (**Figure 22B**).



**Figure 22: Low- and high-energy SID MS of *S. cerevisiae* ENTH1:ANTH:PI(4,5)P<sub>2</sub>.** **A)** At 30 V SID collision energy the main signals originated from the 8:8 (ENTH:ANTH) complex (I) and the dimer of the 8:8 complex (IV). **B)** At 90 V SID collision energy the 8:8 complex was shifted to higher drift times and less abundant (I). Instead, dissociated ANTH domains (III) and residual 8:7 complexes (II) became the most intense signals. Since the peaks of the dimers of 8:8 complexes were less defined (IV) and increased noise in the range above 20,000 *m/z* was observed, first dissociation events of the 2(8:8) complex can be assumed.

Although no calibration with standard proteins was performed for SID measurements, relative changes of drift times upon increasing SID collision energies gave insights into the SID mechanism. At 30 V SID collision energy, drift times increased only slightly, meaning that the structure of the majority of the ions was not affected by the collision with the surface. At 60 V and 90 V SID collision energy, the signal was clearly elongated and shifted towards longer drift times, with no distinct signal maxima.

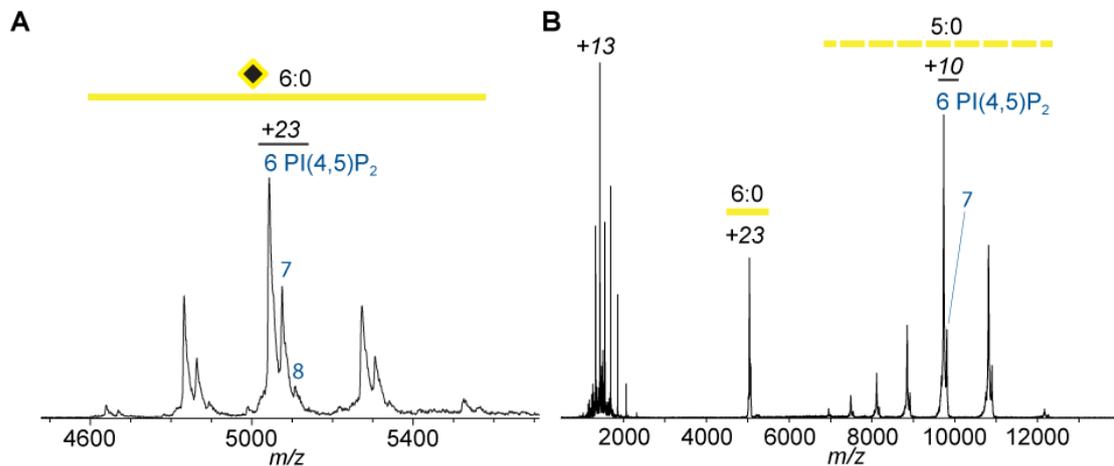
These preliminary results indicate that buffer and PI(4,5)P<sub>2</sub> molecules dissociate at low SID collision energies. Whether there are interfacial phospholipids among the dissociated molecules upon surface collision, should be investigated further, because previous studies state that protein-lipid interactions are well-retained in SID (Harvey, Liu et al., 2017). In comparison, in CID measurements no loss of interfacial lipids was observed. These preliminary results were unexpected, because in contrast to CID, the SID mechanism is assumed to take place without protein unfolding (Zhou, Dagan et al., 2012).

Oftentimes, charges of the precursor ions are reduced by the addition of triethylammonium acetate (TEAA) as this was shown to increase complex stability and to be beneficial for SID studies (Zhou, Dagan et al., 2012). Due to the intrinsically low number of charges of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes this step was omitted in the primary SID MS characterization. However, the effects of charge reduction needs to be tested in upcoming experiments.

#### 4.2.7 Human ENTH Domains Form Hexamers in Presence of PI(4,5)P<sub>2</sub>

Furthermore, complex formation of ENTH and ANTH domains in presence of PI(4,5)P<sub>2</sub> was analysed for human proteins. In a first experiment, ENTH domains of *H. sapiens* epsin-1 were mixed with PI(4,5)P<sub>2</sub>. Native mass spectra revealed not only binding of PI(4,5)P<sub>2</sub> to ENTH domains but also formation of complexes with a mass of around 116 kDa (**Figure 23A**). Comparable ENTH oligomers with fungal ENTH proteins and PI(4,5)P<sub>2</sub> were not observed by native MS.

The assignment of the stoichiometry based on the measured complex mass revealed ENTH hexamers with at least 6 PI(4,5)P<sub>2</sub> molecules bound. Based on the fine structure of the peaks, also complexes with 7 or 8 PI(4,5)P<sub>2</sub> molecules were present. An ENTH domain dissociated in CID MS/MS measurements at elevated collision energies from the ENTH:PI(4,5)P<sub>2</sub> complexes (**Figure 23B**). As shown before in CID MS/MS experiments with fungal ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes (**Figure 16**), human ENTH complexes contained the same number of PI(4,5)P<sub>2</sub> molecules before and after the protein dissociation event. Here, the human ENTH pentamer still carried 6-7 PI(4,5)P<sub>2</sub> molecules, even at high collision energies (100 V) that induced first protein backbone fragmentation events. In view of that, all PI(4,5)P<sub>2</sub> molecules need to be located at protein-protein interfaces in PI(4,5)P<sub>2</sub>-containing ENTH hexamers.



**Figure 23: ENTH domains from *H. sapiens* epsin-1 form stable hexamers in presence of PI(4,5)P<sub>2</sub>.** A) For the most abundant complexes a mass of around 116 kDa was determined in native MS. Complex stoichiometries derived from this mass yielded ENTH hexamers (6:0) with 6 to 8 PI(4,5)P<sub>2</sub> molecules bound. B) In CID MS/MS, the hexamer dissociated in ENTH monomer (1,000 – 2,300 *m/z*) and a residual pentamer (6,900 – 12,200 *m/z*). Even at high collision energies (here: 100 V), all PI(4,5)P<sub>2</sub> molecules remained bound to the ENTH pentamer. Charge states of the main peaks, signals arising from pentamers (5:0) and hexamers (6:0), and number of PI(4,5)P<sub>2</sub> molecules (blue) in the fine structure of peaks are labelled. Panel A was reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

No dissociation constants for PI(4,5)P<sub>2</sub> binding to human epsin-1 ENTH domains were determined and compared to fungal constants because of the hexamer formation that is coupled to PI(4,5)P<sub>2</sub> binding.

Compared to yeast ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes, these human ENTH hexamers proved to be remarkably stable in a thermal denaturation assay by dynamic light scattering (DLS) (Supplementary Figure S5). In this assay, ENTH1 and PI(4,5)P<sub>2</sub> from *S. cerevisiae* had a mid-aggregation temperature ( $T_{agg}$ ) of 33 °C in absence of ANTH and 37 °C in presence of ANTH. In contrast, PI(4,5)P<sub>2</sub>-containing *H. sapiens* ENTH-hexamers revealed a  $T_{agg}$  of 45 °C. With regard to the human physiological body temperature of 37 °C, these results suggest a stable formation of ENTH hexamers under physiological conditions.

#### 4.2.8 Structural Analysis of Human ENTH Hexamers Using IMMS and SAXS

Human PI(4,5)P<sub>2</sub>-containing ENTH hexamers were analysed by IMMS as described for *S. cerevisiae* clathrin adaptor complexes (see 4.2.5 IMMS Analysis of *S. cerevisiae* ENTH:ANTH:PI(4,5)P<sub>2</sub>, page 57). IMMS data acquired under gentle conditions (trap collision energy: 25 V, sampling cone voltage: 25 V) revealed the presence of ENTH hexamers, even though the obtained resolution was not sufficient to specify numbers of PI(4,5)P<sub>2</sub> molecules. The drift plot displayed a single conformation for charge states +23 to +21 (**Figure 24A**) with maxima of arrival time distributions of 11.66 to 13.30 ms (**Table 4**). In the *m/z*-range of charge states +20 and +19, another less populated species with slightly decreased drift times was found. Relevant charge states were extracted from the drift plot for further analysis (**Supplementary Figure S6**). Due to the low spectral resolution, charge states could not be identified with high certainty. Nevertheless, comparing how different eligible peak assignments match the lowly resolved peaks, revealed a shift towards higher masses. For instance, a putative assignment of the main peak as the +26 ion results in complex mass of 154.6 kDa. In a very preliminary and speculative data interpretation one could recognize this complex as a PI(4,5)P<sub>2</sub> containing ENTH octamer, which had a CCS' of approximately 8,000 Å<sup>2</sup>. Yet further analysis is required to clearly identify this complex and thereby perhaps find a link between the ENTH hexamer and larger clathrin adaptor protein complexes. Possibly, inducing CID after ion mobility separation by the application of a high collision energy in the transfer cell is a straightforward experiment to clearly identify the doubtful complex species. With this experimental setup, fragment ions can be easily assigned to their precursors as they have the same drift times.

Drift times of all charge states were extracted from the spectrum and the CCSs' were estimated from a logarithmic fit ( $R^2 = 0.92$ ) as described before (see 1.8 Ion Mobility Mass Spectrometry, page 23, **Supplementary Figure S7, Table 4**). The low  $R^2$ -value of the logarithmic fit leads to some uncertainty in the estimated CCSs' that were in the range of 6,480 to 6,650 Å<sup>2</sup> for the main peak series. A slightly decreased CCS' of around 6,120 Å<sup>2</sup> was derived for the more compact complex species.

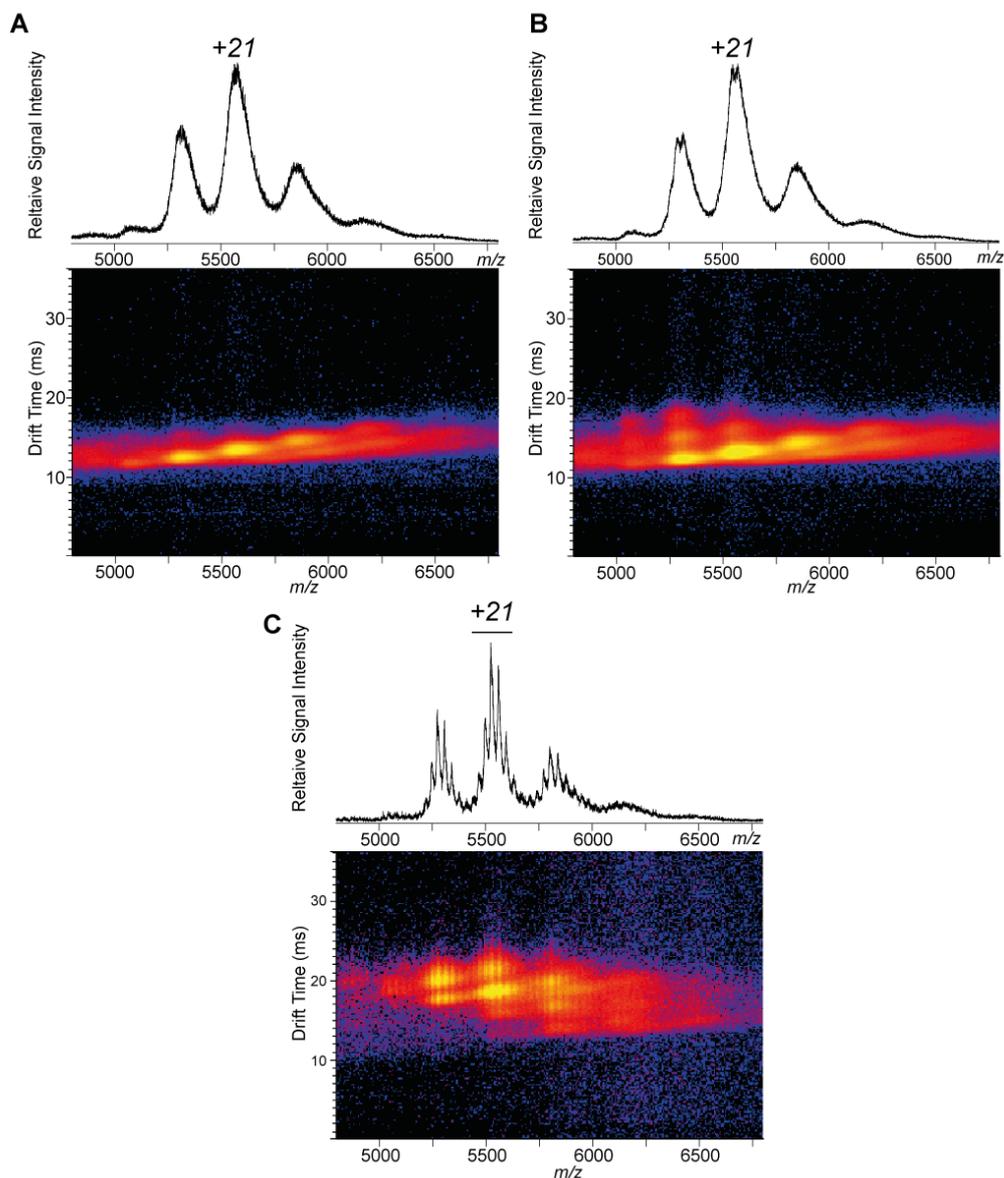
With less gentle MS measuring conditions, higher resolution spectra were obtained. When the ions were activated in the source region by the application of a high sampling cone voltage (175 V), first peak fine structures that originate from different numbers of PI(4,5)P<sub>2</sub> molecules were detected (**Figure 24B**). Nonetheless, these conditions were already too harsh to guarantee transmission of the ions in a compact state. For charge states +23 to +21, signals were elongated along the y-axis towards higher drift times with two additional maxima, for instance at

14.94 ms ( $CCS' = 6,940 \text{ \AA}^2$ ) and 17.50 ms ( $CCS' = 7,550 \text{ \AA}^2$ ) for charge state +22. As expected, higher charge states were more susceptible to unfolding because of higher energy collisions with the collision gas.

**Table 4: Measured drift times and derived  $CCSs'$  for PI(4,5) $P_2$  containing epsin-1 ENTH hexamers from *H. sapiens*.** Values for the second ion series with decreased drift times and uncertain charge states are labelled (#). An alternative interpretation of the data that designates the relevant charge states as +26 and +25 and results in  $CCSs'$  of around  $8000 \text{ \AA}^2$  is given in parentheses. A plot containing the logarithmic fit of the calibrant ions is presented as **Supplementary Figure S7**.

Charge state $z$	Drift time (ms)	$CCS' (\text{\AA}^2)$
23	11.66	6,650
22	12.39	6,570
21	13.30	6,520
20	14.40	6,480
19	15.86	6,480
#20 (26)	#12.94	#6,120 (7,952)
#19 (25)	#14.22	#6,110 (8,041)

Applying increased energies in later stages of the flight path proved to be more beneficial for increasing mass resolution. Non-volatile salt molecules or solvent molecules that remained bound to the complex during the ESI process, dissociated under the influence of 75 V trap collision energy. With the obtained resolution, different numbers of PI(4,5) $P_2$  in the ENTH hexamer were recognized (**Figure 24C**). In addition, unfolding of the complex structures became more pronounced. In particular for charge states +22 and +21, two unfolded species were observed that can be considered to be two different states of unfolding. Drift times were extracted and the corrected  $CCSs'$  were determined as described above.  $CCSs'$  were 7,800 to  $7,950 \text{ \AA}^2$  and 8,390 to  $8,590 \text{ \AA}^2$ , respectively for the different states of unfolding (**Table 5**).



**Figure 24: Ion mobility mass spectrometry of PI(4,5)P<sub>2</sub>-containing *H. sapiens* epsin-1 ENTH hexamers.** **A)** Gentle MS conditions (25 V sampling cone, 25 V trap collision energy) resulted in low resolution mass spectra (top) and drift plots (bottom) showed compact complex conformations. **B)** Exposing the ions to a high sampling cone voltage (175 V, 25 V trap collision energy) caused first unfolding of the complexes, especially for charge state +22 two new species with increased drift times were detected. Peaks in the mass spectrum (top) showed first indications of a fine structure indicating the different numbers of PI(4,5)P<sub>2</sub> molecules. **C)** When ions were instead activated by a high trap collision energy (75 V, 25 V sampling cone voltage), complex unfolding became more prominent. For instance for charge state +22, in addition to the previously observed two unfolded states, a third state with even further increased drift times was found. Spectral resolution was sufficient to distinguish different numbers of PI(4,5)P<sub>2</sub> molecules bound in the ENTH hexamer. Charge states of main peaks are annotated.

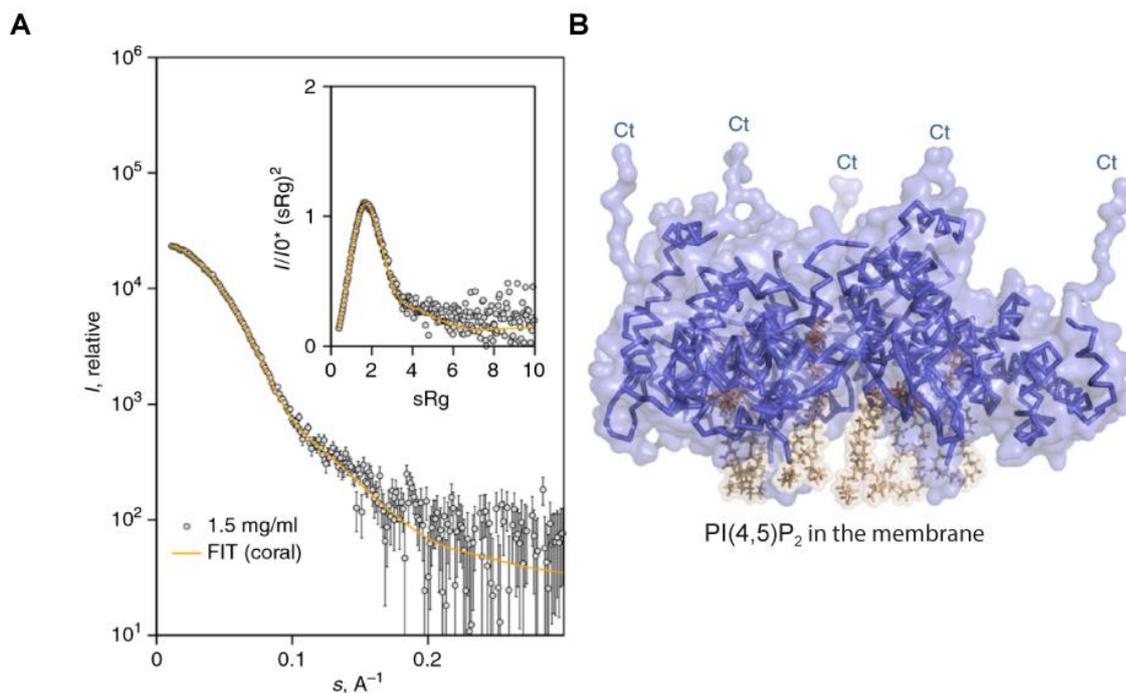
To further elucidate a possible PI(4,5)P<sub>2</sub> dependence of complex unfolding, the two unfolded states were extracted from the drift plot (**Supplementary Figure S8**). From these isolated data, PI(4,5)P<sub>2</sub> binding patterns of the two unfolded states were compared for charge states +22 and +21. Overlays of the spectra unveil no major difference in PI(4,5)P<sub>2</sub> binding and argue against a protective role of PI(4,5)P<sub>2</sub> for complex unfolding.

**Table 5: Measured drift times and derived CCSs' for PI(4,5)P<sub>2</sub> containing epsin-1 ENTH hexamers from *H. sapiens* that were exposed to 75 V trap collision energy.** Charge states are labelled for increased (+) and vastly increased (++) observed CCSs' of ion series. A plot showing the logarithmic fit of the calibrant ions is presented as **Supplementary Figure S7**.

Charge state $z$	Drift time (ms)	CCS' (Å <sup>2</sup> )
22 <sup>+</sup>	17.68	7,950
21 <sup>+</sup>	18.59	7,800
23 <sup>++</sup>	18.77	8,590
22 <sup>++</sup>	20.05	8,510
20 <sup>++</sup>	21.32	8,390

Taken together, a compact complex conformation with a CCS' of 6,570 Å<sup>2</sup> for charge state +22 was observed when mild MS conditions were applied. When ions were exposed to increased energies in the source region of the mass spectrometer or in the trap cell, distinct states with different degrees of unfolding were observed with CCSs' of 6,940, 7,950, and 8,510 Å<sup>2</sup> for charge state +22. In other words, complexes undergo defined conformational changes upon activation in the gas-phase that possibly originate from specific spatial rearrangements of the ENTH subunits. This unfolding pathway is different from the process that was observed in SID MS experiments with PI(4,5)P<sub>2</sub>-containing *S. cerevisiae* ENTH:ANTH complexes. These complexes revealed increased drift times that suggested random unfolding after colliding with a surface (**Figure 21, Figure 22**).

The formation of PI(4,5)P<sub>2</sub>-containing ENTH hexamers was also confirmed in solution by small-angle X-ray scattering (SAXS) (**Figure 23C, Supplementary Table S5**). A model of the hexamer was generated with the X-ray structure containing two *S. cerevisiae* ENTH2 domains surrounding a PI(4,5)P<sub>2</sub> molecule (**Figure 13**) as a rigid-body model with enforced p3 symmetry and dummy residue chains representing C-terminal parts missing in the crystal structure (**Figure 23D**). The model yielded a good fit to the experimental data (**Figure 23C, Supplementary Figure S9**). PI(4,5)P<sub>2</sub> molecules in this model are oriented in one direction, potentially setting up the membrane interactions. C-termini are pointing to the opposite cytoplasmic direction, where they can bind interaction partners such as clathrin.



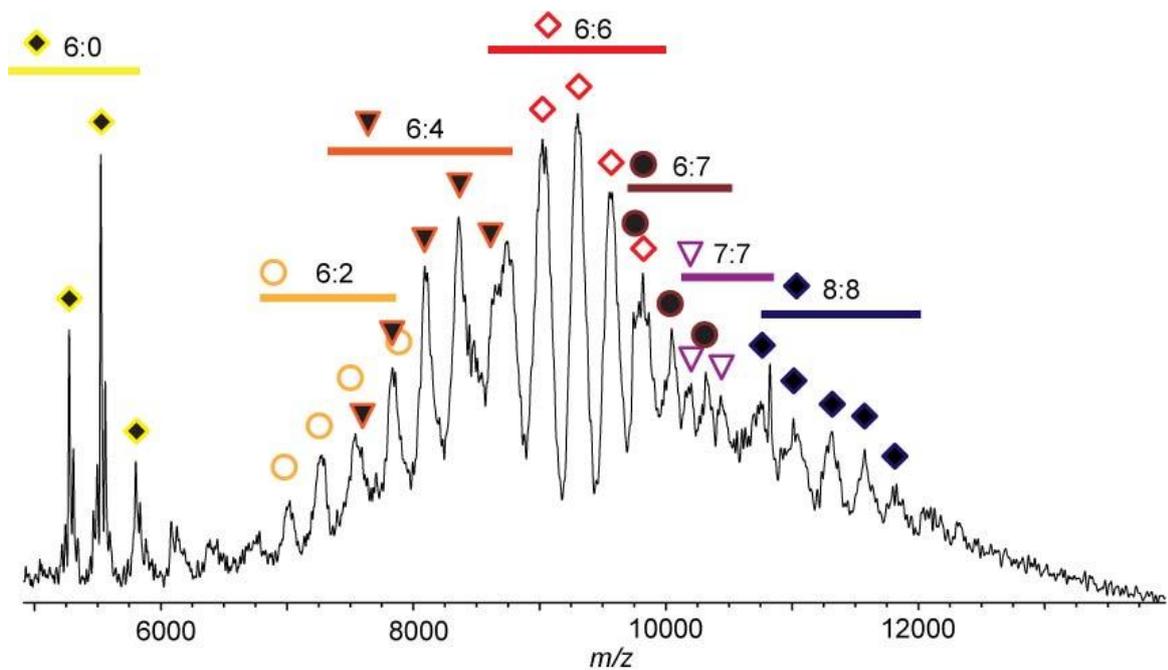
**Figure 25 SAXS modelling of human epsin-1 ENTH in the presence of PI(4,5)P<sub>2</sub>.** **A)** SAXS data recorded for human ENTH in solution are shown (grey circles) along with a fit to the rigid-body model refined against the SAXS data with p3 symmetry using CORAL (yellow line). Experimental errors are results from counting statistics on the Pilatus 2M detector and propagated through the data reduction process as standard errors in the scattering intensities. The  $\chi^2$  for the fit is 1.05. The inset shows the dimensionless Kratky plot representation of the SAXS data and the same fit. **B)** Backbone and surface representation of the SAXS refined rigid body model. C-terminal residues (Ct) that were not observed in the crystal structure of the tandem domains were modelled as dummy residues by CORAL. P3 symmetry was enforced and the tandem ENTH domains with bound PI(4,5)P<sub>2</sub> used as rigid bodies. Bound PI(4,5)P<sub>2</sub> molecules (light brown) are all located on one side of the complex, suggesting a membrane-binding interface. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

For comparison, the CCS of the SAXS model (**Figure 25B**) of the human PI(4,5)P<sub>2</sub> containing epsin-1 ENTH hexamer was determined computationally. After removal of the dummy residues at the ENTH C-termini and addition of hydrogens, the CCS was calculated using the software IMPACT (Marklund, Degiacomi et al., 2015). The resulting CCS based on the projection approximation (PA) method (5,717 Å<sup>2</sup>) was previously shown to systematically underestimate CCSs (Jurneczko & Barran, 2011). IMPACT corrects for this underestimation, producing CCSs that are in very good agreement with the more precise but computationally expensive trajectory method. This approximation yields a CCS of 7,507 Å<sup>2</sup> for the SAXS model of the human ENTH hexamer, which is about 15 % larger than the values that were determined by IMMS (**Table 4**). Judging from the structure of the SAXS model, a slight overestimation of

the CCS seems likely, as N-termini and fatty acid chains almost certainly collapse in the gas phase to the protein surface. Thus, the actual shape of the complexes during IMMS is rather cylindrical and not mushroom-like as in the SAXS model. Furthermore, a compaction of the complex structure in the gas phase, as shown for GroEL chaperones cannot be excluded (Hogan Jr, Ruotolo et al., 2011). Rather harsh settings as used in the trap cell (25 V trap CE, 45 V trap DC bias) could promote this partial collapse. In addition, the low R<sup>2</sup>-value (0.92) of the logarithmic fit for the calibration of IMMS measurements results in some uncertainty in the determined CCSs' (Supplementary Figure S7).

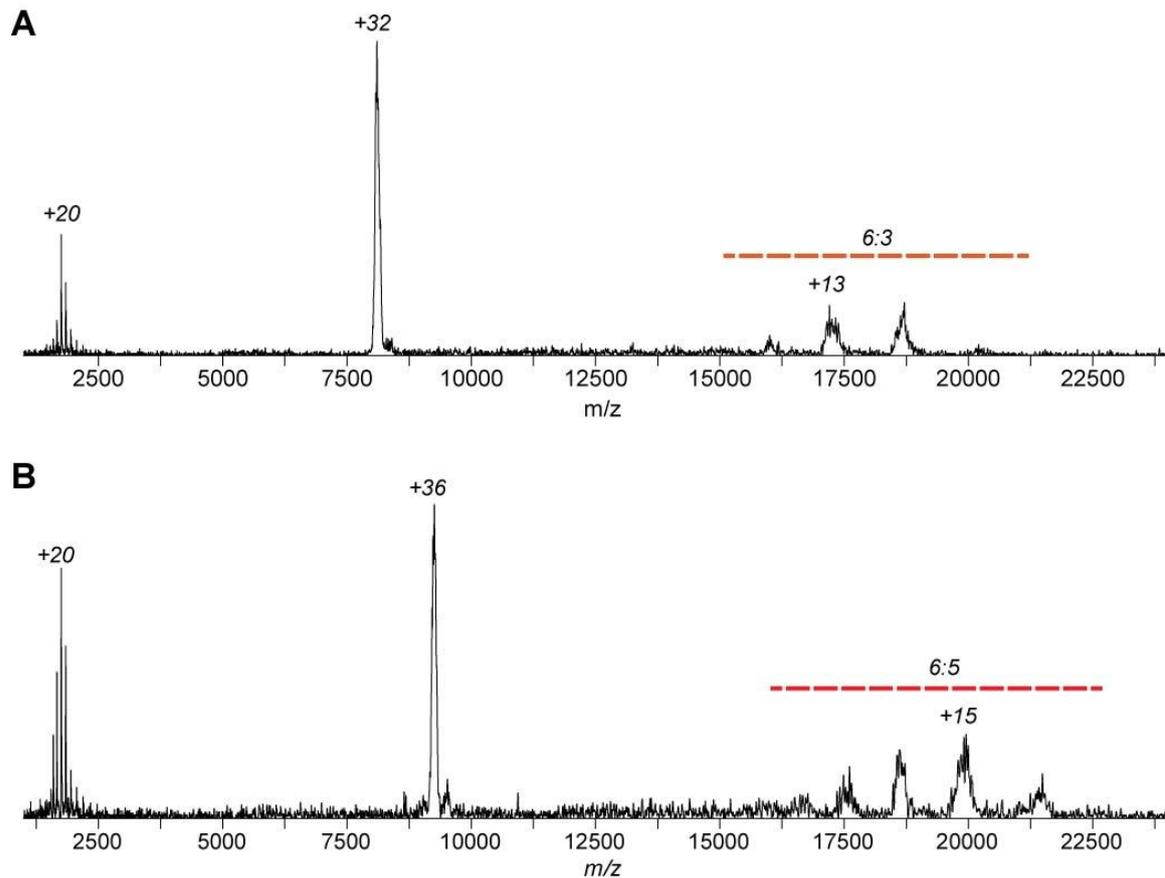
#### 4.2.9 *Hip1R ANTH Domains Bind to ENTH-6mers*

To investigate possible ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes of *H. sapiens* proteins, human ENTH domains from epsin-1, ANTH domains from the Sla2-homolog Hip1R and PI(4,5)P<sub>2</sub> were mixed and analysed for complex formation by native MS (Figure 26). As expected from previous results, hexamers of ENTH (6:0) with 6-8 PI(4,5)P<sub>2</sub> were identified. In addition, various peaks in the range of 7,000 to 13,000 *m/z* were observed. MS/MS measurements were performed to unravel complex stoichiometries. Increased collision energies led to the dissociation of Hip1R ANTH domains from all examined precursor ions. Molecular weights of the residual complexes were determined and used for the assignment of the MS spectrum. Two exemplary cases of the procedure are illustrated in Figure 27. For instance, when the main peak at 9,278 *m/z* was selected for MS/MS analysis, dissociation studies revealed a residual complex with a molecular weight of approximately 298 kDa (Figure 27B). Considering the 35 kDa from the ejected Hip1R domain and the mass of the ENTH hexamer, a complex stoichiometry of 6:6 (ENTH:ANTH) for the highest abundant peak series was assigned (Figure 26).



**Figure 26: Human Hip1R ANTH domains bind to epsin-1 ENTH hexamers.** PI(4,5)P<sub>2</sub>-containing ENTH hexamers (yellow) were mixed with Hip1R ANTH domains. ANTH binding to preformed ENTH hexamers was detected in different stoichiometries with 6:6 ENTH:ANTH complexes being the most abundant complex species. Moreover, assemblies containing more than six ENTH domains with stoichiometries of up to 8:8 were identified. ENTH:ANTH stoichiometries as assigned also by CID MS/MS (Figure 27) are labelled. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.

Due to low signal intensities and repeatedly clogging electrospray capillaries, the analysis was cumbersome. This behaviour is indicative for non-ideal sample conditions and the formation of aggregates. Compared to the spectra of fungal ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes the relative amount of observed 8:8 ENTH:ANTH protein complexes is distinctly lower. Non-ideal buffer and protein conditions can alter and distort the observed ratios. Assuming that this was not the case, the high stability of the human ENTH hexamers could obstruct the transition of 6:6 complexes to complexes with more than six ENTH entities, such as 8:8 complexes.

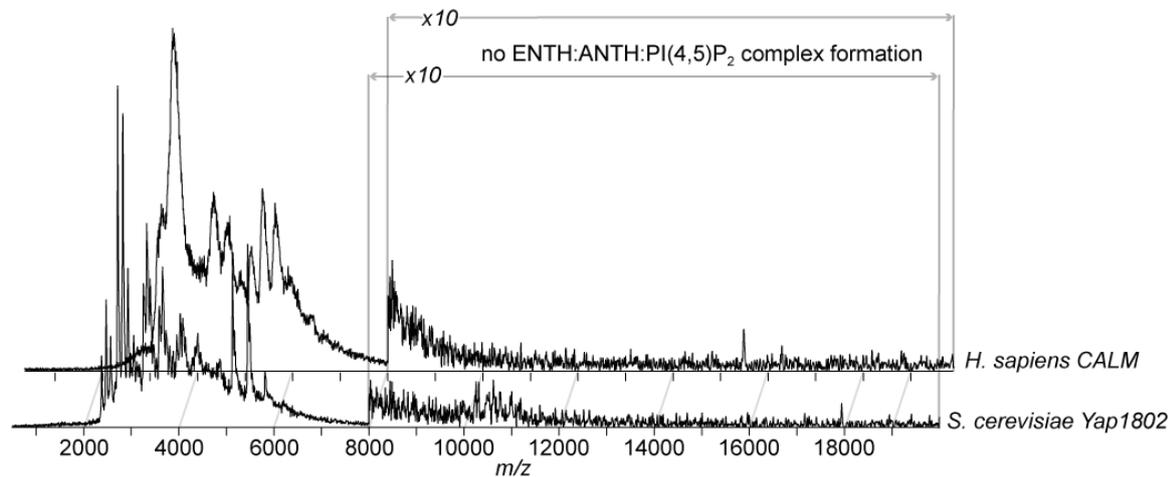


**Figure 27: Collision-induced dissociation analysis of *H. sapiens* ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes.** **A)** Precursor ions encompassing 8,140 *m/z* were selected, accelerated with 170 V collision energy and collided with a collision gas. Hip1R monomers (35 kDa) dissociated and a residual complex with a mass of 223 kDa was found. Thus, a precursor ion mass of 258 kDa was derived. **B)** Likewise, precursor ions encompassing 9,278 *m/z* were selected and dissociated into Hip1R monomers and a residual complex with molecular weight around 298 kDa. The assigned precursor ions had a mass of approximately 334 kDa. Charge states of the main peaks and determined ENTH:ANTH stoichiometries of residual complexes are annotated.

#### 4.2.10 An ANTH Subfamily Does Not Form ENTH:ANTH:PI(4,5)P<sub>2</sub> Complexes

ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes were assembled with ANTH domains of the Sla2/Hip1R subfamily for three different species. However, not all tested ENTH and ANTH domains assembled to ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes. When ANTH domains from the CALM subfamily were used instead of Sla2/Hip1R ANTH, no complex formation was observed. In mixtures containing human epsin-1 ENTH, CALM ANTH and PI(4,5)P<sub>2</sub>, only low resolution signal for free ANTH, PI(4,5)P<sub>2</sub>-bound ANTH and ENTH hexamers was detected. However, no signal was observed for the range above 8,000 *m/z* (Figure 28). Similarly, *S. cerevisiae* ENTH2 was mixed with PI(4,5)P<sub>2</sub> and ANTH domains of the CALM yeast homolog Yap1802. Apart from free and PI(4,5)P<sub>2</sub>-bound ENTH and ANTH domains an 87 kDa contaminating protein

was identified. No obvious signal appeared in the range above 8,000  $m/z$ . Only a faint and unresolved signal around 10,500  $m/z$  was noticed upon spectrum magnification. Since the origin of the signal is unclear and the signal intensity is extremely low, still the conclusion was drawn that ENTH:ANTH:PI(4,5)P<sub>2</sub> complex formation could not be verified for ANTH proteins from the CALM subfamily.



**Figure 28: No ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes with ANTH domains of the CALM subfamily were identified.** *S. cerevisiae* epsin-2 ENTH and PI(4,5)P<sub>2</sub> were mixed with ANTH domains of *S. cerevisiae* Yap1802 (front). Only a marginal increase in the signal noise was observed in the range around 10,500  $m/z$ . In the low  $m/z$ -range, PI(4,5)P<sub>2</sub>-bound ENTH and ANTH monomers and contaminant protein with approximately 87 kDa (at around 5,000  $m/z$ ) molecular weight were identified. Similarly, after mixing of human epsin-1 ENTH, PI(4,5)P<sub>2</sub> and ANTH domains of CALM no ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes were identified (back). The range above 8,000  $m/z$  is presented in 10-fold magnification to highlight the absence of signals.

#### 4.2.11 Cross-Species in Vitro Assembly of ENTH:ANTH:PI(4,5)P<sub>2</sub> Complexes

The formation of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes was demonstrated for fungal (*S. cerevisiae*, *C. thermophilum*) and mammalian (*H. sapiens*) proteins. In the next set of experiments, ENTH and ANTH domains were mixed together with PI(4,5)P<sub>2</sub> across species boundaries (Figure 29). For instance, *C. thermophilum* ENTH was mixed with *S. cerevisiae* Sla2 ANTH and PI(4,5)P<sub>2</sub> and the resulting spectra showed prominent signals for chimeric ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes with 6:6, 6:7 and 8:8 ENTH:ANTH stoichiometries. Complex assembly was much less efficient when proteins were used inversely, meaning *S. cerevisiae* ENTH2, *C. thermophilum* ANTH and PI(4,5)P<sub>2</sub>. But still, low signals in the  $m/z$ -range of the 8:8 ENTH:ANTH complexes were observed, demonstrating that the assembly of chimeric complexes is also possible.

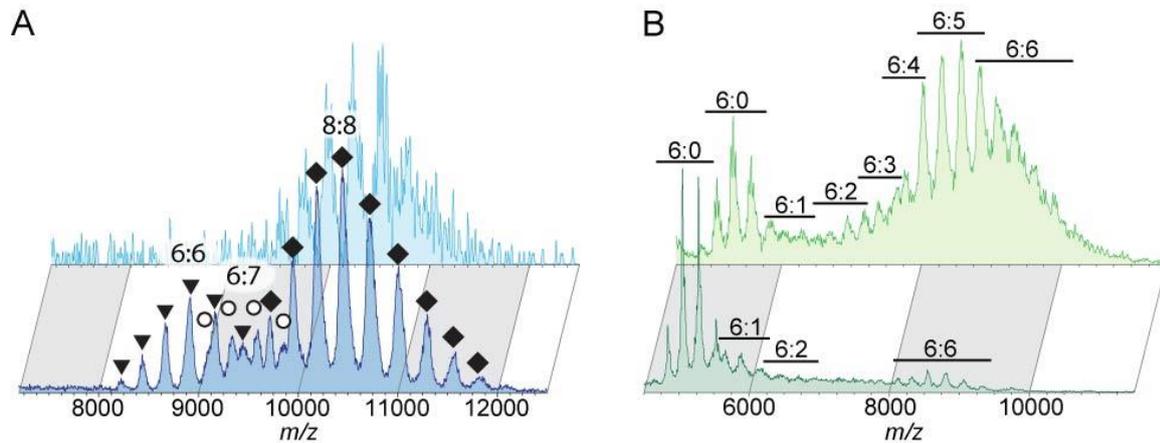
These results are not too astonishing, as these ENTH and ANTH domains are all fungal proteins with sequence identities of around 55 % (ENTH domains) and almost 60 % (ANTH

proteins), reflecting the homology between these proteins (Supplementary Table S6, Supplementary Table S7). Furthermore, ENTH:ANTH:PI(4,5)P<sub>2</sub> complex formation proved to proceed in the same manner in *S. cerevisiae* and *C. thermophilum* in previous experiments.

In contrast, sequence identity between human and fungal proteins is lower, especially for ANTH proteins. While ENTH domains share 42 % to 50 % of the same residues, for fungal Sla2 ANTH domains and human Hip1R ANTH only about 27 % of the residues are identical. Nevertheless, it was also possible to assemble chimeric complexes with human ENTH and fungal ANTH domains (Figure 29B). For both domains, ANTH from *S. cerevisiae* and *C. thermophilum* binding to human ENTH hexamers was observed. Signal intensities suggest a more efficient binding of *C. thermophilum* ANTH domains to human ENTH hexamers since signal intensities in the high *m/z*-range are higher than for ANTH domains from *S. cerevisiae*. Furthermore, for the first-mentioned mixture, all the different ENTH:ANTH stoichiometries from 6:0, over 6:1 up to 6:6 were identified. But for both cases, no complexes larger than 6:6 were formed, suggesting that a transition from a complex with the stable PI(4,5)P<sub>2</sub>-containing ENTH hexamer to complex species with up to 8 ENTH domains cannot be induced by binding of ANTH domains of foreign origin.

These MS data on the formation of cross-species ENTH:ANTH:PI(4,5)P<sub>2</sub> samples were confirmed in solution by ITC measurements. In a first experiment, an exothermic reaction was observed when *S. cerevisiae* ANTH was added to pre-mixed *S. cerevisiae* ENTH2 and PI(4,5)P<sub>2</sub>, indicating ANTH binding and ENTH:ANTH:PI(4,5)P<sub>2</sub> complex formation (Supplementary Figure S10A). Although weaker pronounced, binding was also observed when ANTH from *C. thermophilum* was used instead and added to the *S. cerevisiae* ENTH2:PI(4,5)P<sub>2</sub> mixture (Supplementary Figure S10B).

The presented *in vitro* formation of cross-species ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes highlights the conservation of the general mechanism of PI(4,5)P<sub>2</sub>-dependent assembly of clathrin adaptor proteins. Considering the rather low sequence identity between human and fungal ANTH proteins and the common competency to bind PI(4,5)P<sub>2</sub> molecules, it can be concluded that ENTH:ANTH:PI(4,5)P<sub>2</sub> complex formation is mainly dependent of the protein-PI(4,5)P<sub>2</sub> interactions and less reliant on protein-protein interactions.



**Figure 29: Cross-species complex assembly of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes.** A) Mixing of epsin ENTH (*C. thermophilum*), Sla2 ANTH (*S. cerevisiae*) and PI(4,5)P<sub>2</sub> led to the formation of 6:6 (triangles), 6:7 (circles) and 8:8 (rhombuses) ENTH:ANTH-complexes (dark blue). Mixing fungal ENTH and ANTH domains inversely, so epsin-2 ENTH (*S. cerevisiae*), Sla2 ANTH (*C. thermophilum*) and PI(4,5)P<sub>2</sub>, resulted in poor quality signal in the high *m/z*-range, demonstrating inefficient 8:8 complex formation (light blue). B) Sla2 ANTH proteins from *S. cerevisiae* (dark green) and *C. thermophilum* (light green) were added to human epsin-1 ENTH and PI(4,5)P<sub>2</sub>. In both cases, ANTH proteins bound to the human ENTH 6-mer in different stoichiometries. No complexes larger than 6:6 were observed. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.

#### 4.2.12 General Assembly Model of ENTH:ANTH:PI(4,5)P<sub>2</sub> Complexes

In-depth native MS analyses in conjunction with other biophysical techniques disclosed a general sequence of events that lead to the PI(4,5)P<sub>2</sub>-dependent clustering of clathrin adaptor proteins with ENTH and ANTH domains (Figure 30).

All tested ENTH and ANTH domains from *S. cerevisiae*, *C. thermophilum* and *H. sapiens* were demonstrated to contain two PI(4,5)P<sub>2</sub> interaction sides. The dissociation constants that were derived from native mass spectra indicate a positive allosteric mechanism of PI(4,5)P<sub>2</sub> binding. These interactions of ENTH and ANTH domains with PI(4,5)P<sub>2</sub> allow an initial recruitment of clathrin adaptor proteins to the inner plasma membrane where ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes form.

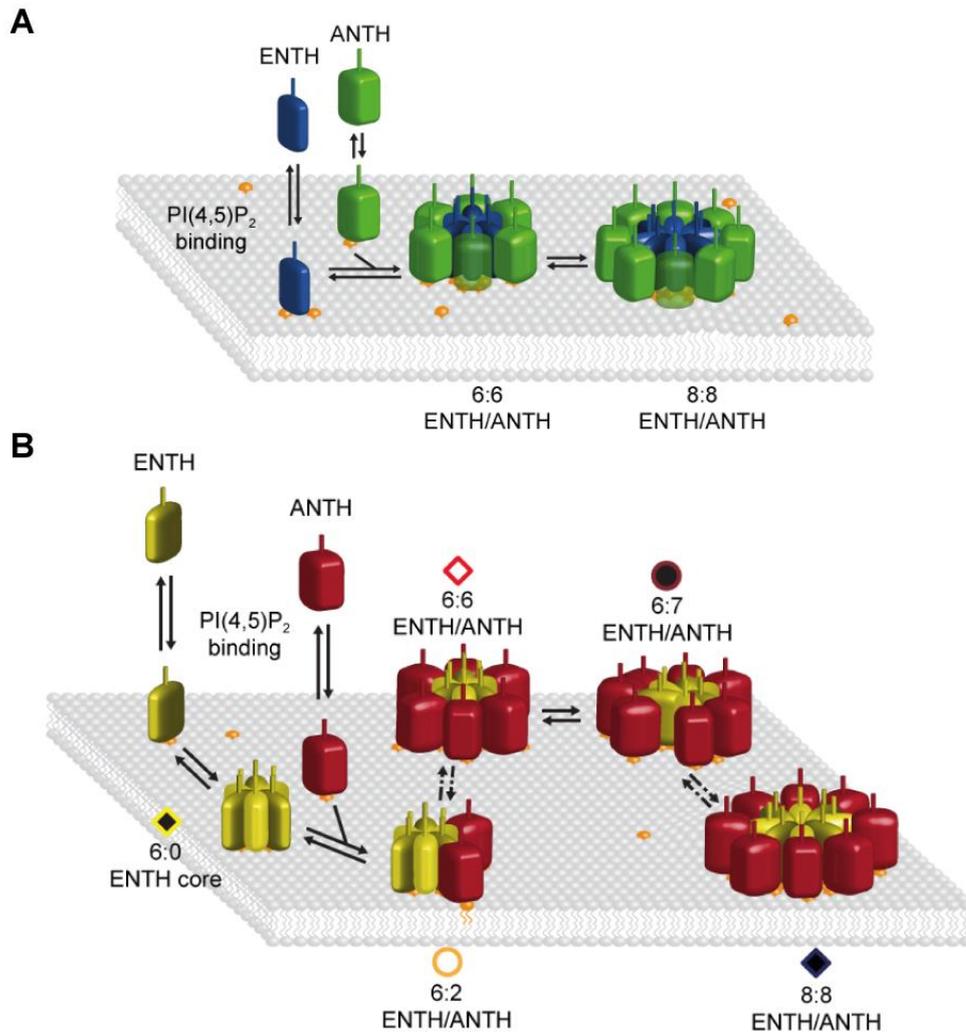
The presented data from biophysical experiments revealed a difference for the assembly of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes between fungal and human proteins. In fungal systems (*S. cerevisiae*, *C. thermophilum*) two different ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes were identified. Changes in native MS signal ratios in the course of time suggest a transformation of smaller 6:6:~19 (ENTH:ANTH:PI(4,5)P<sub>2</sub>) complexes to more stable 8:8:~25 complexes (Figure 30A). The

assembly was shown to be reversible and dependent on the availability of PI(4,5)P<sub>2</sub> in the assembly solution.

ENTH and ANTH domains of the human clathrin adaptor proteins epsin-1 and Hip1R follow a slightly different assembly pathway (**Figure 30B**). For these proteins, stable ENTH hexamers with at least 6 bound PI(4,5)P<sub>2</sub> molecules were recognized that can be decorated by ANTH domains. Several different ENTH:ANTH protein stoichiometries were identified in native MS experiments with 6:6 complexes being the most abundant species. Yet, a transition to complexes containing more than 6 ENTH domains is possible and 8:8 complexes were the largest observed complexes in this study.

The experiments also gave a detailed insight into the 3-dimensional structure of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes. Various pieces of evidence from native MS and X-ray crystallography confirm that PI(4,5)P<sub>2</sub> molecules are located between ENTH and ANTH domains and interact with two different protein domains, thus mediating the contact between the protein domains. Similar integral membrane protein complexes with interfacial lipids were recently identified by native MS (Gupta, Donlan et al., 2017). Furthermore, the stable human ENTH hexamer and CID MS/MS dissociation pathways suggest a complex superstructure with ENTH domains building the core and ANTH proteins binding in a PI(4,5)P<sub>2</sub>-dependent manner to the core's surface.

These assembly models, generated based on *in vitro* data, are in good agreement with previously published cell biological findings on epsins and Sla2/Hip1R from yeast and mammals. Sla2 ANTH and epsin ENTH were reported to work cooperatively in *S. cerevisiae* endocytosis with ANTH domains stabilizing the ENTH domains at the endocytic sites (Skruzny, Brach et al., 2012). However, for murine proteins it was shown that ENTH domains bind to the endocytic site in absence of Hip1R and are then required for Hip1R recruitment (Messa, Fernandez-Busnadiego et al., 2014). The general order of events that was described in aforementioned studies is in very good agreement with the assembly model, which is proposed in this work based on the results from native MS experiments.



**Figure 30: Schematic models of ENTH:ANTH:PI(4,5)P<sub>2</sub> complex assembly.** **A)** In *S. cerevisiae* and *C. thermophilum* ENTH (blue) and ANTH (green) domains bind PI(4,5)P<sub>2</sub> (orange) in the membrane. ENTH and ANTH domains cluster with PI(4,5)P<sub>2</sub>, mediating the interactions to 6:6 (ENTH:ANTH) complexes that are subsequently transformed to more stable 8:8 complexes. **B)** Human ENTH domains (yellow) bind PI(4,5)P<sub>2</sub> (orange) and cluster to hexamers. ANTH domains (red) bind in a PI(4,5)P<sub>2</sub> dependent manner in different stoichiometries. 6:6 complexes are the most abundant species, but a transition to larger complexes, up to 8:8 was demonstrated. Symbols annotating the different complexes are chosen as in the mass spectrum of **Figure 26**. Dashed arrows indicate that not all complex stoichiometries are represented in this model. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

## 4.3 Methods

Experimental methods described in this chapter were adapted from Garcia-Alai, Heidemann et al. (2018).

### 4.3.1 Cloning

ENTH and ANTH domains from *S. cerevisiae* were codon-optimized and cloned into pETM30 vectors as described (Skruzny, Desfosses et al., 2015). The sequence coding for the ANTH domain of *S. cerevisiae* Yap1802 (aa 1-272) was PCR amplified from gDNA and cloned into *NcoI* - *XhoI* sites of pETM30 (EMBL).

Sequences for epsin-1 ENTH and Sla2 ANTH from *C. thermophilum* were amplified from an RNA library kindly provided by Peer Bork's laboratory (EMBL Heidelberg) using the QIAGEN OneStep RT-PCR Kit and cloned into the pETM-11/LIC vector (EMBL).

*H. sapiens* epsin-1 ENTH coding sequence (aa 1-158) was PCR amplified from human cDNA and cloned into the *NcoI* - *XhoI* sites of pETM30 with the addition of a codon for glycine after the *start* codon. The coding sequence of human Hip1R ANTH domain (aa 1-300) was PCR amplified from human cDNA and cloned into *NcoI* - *XhoI* sites of the vector pETM30 (EMBL). The coding sequence for the ANTH domain of the human CALM protein was purchased from Addgene (Plasmid #27691) and cloned in pETM-11/LIC vector for expression with an N-terminal 6xHis-tag. Further experimental details are described in Garcia-Alai, Heidemann et al. (2018).

### 4.3.2 Protein Production and Purification

Recombinant human Hip1R ANTH, human epsin-1 ENTH and *S. cerevisiae* ENTH and ANTH domains were produced from the pETM30 vector in *E. coli* BL21 (DE3) (Novagen) as GST fusion constructs with an additional N-terminal 6xHis-tag and a cleavage site of the TEV protease between the His-GST tag and the protein of interest. N-terminal 6xHis-tags followed by a TEV cleavage site were also present the pETM-11/LIC constructs of recombinant human ANTH domains from CALM and *C. thermophilum* ENTH and ANTH domains during production in *E. coli* BL21 (DE3) pLYS (Novagen). Expression was induced with 0.5 mM IPTG at an optical density of 1 (OD<sub>600</sub> = 1) in LB medium supplemented with kanamycin and carried out for 4 h at 20 °C. After harvest by centrifugation, cells were lysed by sonication. The soluble fraction of the lysate was loaded on a nickel-nitrilotriacetic acid (Ni-NTA) resin and fusion proteins eluted with 250 mM imidazole. N-terminal His-tags or His-GST tags were cleaved off

using an excess of 6xHis-tagged TEV protease. Digestion was performed overnight during dialysis against buffer without imidazole at 4 °C. Proteins of interest were separated from cleaved tags and TEV protease in a second affinity chromatography step with Ni-NTA. The flow-through was concentrated to approximately 5 mg/ml and subjected to a size-exclusion chromatography using an Äkta liquid chromatography system (Amersham Biosciences) and an S75 10/300 GL (Tricorn) column (GE Healthcare).

Amino acid sequences of ENTH and ANTH domains as used in this study are presented as **Supplementary Table S8**. For more experimental details see Garcia-Alai, Heidemann et al. (2018) and Skruzny, Desfosses et al. (2015).

### 4.3.3 *Native Mass Spectrometry: Sample Preparations*

Prior to native MS analysis, purified proteins were buffer exchanged to 300 mM ammonium acetate (#431311, 99.99 % purity, Sigma-Aldrich) and 1 mM DL-dithiothreitol (Sigma-Aldrich), pH 8.0, via centrifugal filter units (Vivaspin 500, MWCO 5,000 and 10,000, Sartorius) at 13,000 xg and 4 °C.

ESI capillaries were produced in house as described before (see 3.3.2 Native Mass Spectrometry, page 34).

### 4.3.4 *Native Mass Spectrometry: Protein-PI(4,5)P<sub>2</sub> Interactions*

Lipid binding was studied on an LCT ESI-ToF system (Waters and MS Vision) using the gentlest ionization possible (capillary 1.4 kV, cone 100–120 V, extraction cone 0 V, 6.5 mbar source pressure). Samples were prepared by adding 60 μM PI(4,5)P<sub>2</sub> (Phosphatidylinositol-4,5-bisphosphate diC8, Echelon) to the protein of interest (10 μM). Thus, soluble PI(4,5)P<sub>2</sub> was used at a submicellar concentration that is in the range of its physiological concentration in the cell (Gamper & Shapiro, 2007, Hilgemann, 2007).

For ENTH domains, cytochrome c from equine heart (#129021, Sigma-Aldrich) and for ANTH domains, carbonic anhydrase isoenzyme II from bovine erythrocytes (#C2522, Sigma-Aldrich) was used as reference protein to test for unspecific clustering as described (Sun, Kitova et al., 2006) (see 1.6.2 Protein-Ligand interactions, page 14). If there was signal overlap of the ENTH or ANTH domain with signal of the PI(4,5)P<sub>2</sub>-bound reference protein, this charge state was excluded from the calculation of the ratio of unspecific binding. As an approximation the unspecific binding that was determined using the remaining charge states of the reference protein was subtracted from overlapping signals and the residual signals of the overlapping peaks were considered for further analysis. Relative peak intensities were used to determine

the ratio of lipid-bound reference protein to non-bound reference protein. This ratio was used to correct peak intensities and visualize the data using GraphPad Prism (GraphPad Software).

Cooperativity of the two binding sites was assessed by reviewing the mathematical relation of microscopic (binding site) and macroscopic (apparent) dissociation constants of independent binding sites (see 1.6.2 Protein-Ligand interactions, page 14). A lack of microscopic constants  $k_{d,1}$  and  $k_{d,2}$  fulfilling the required condition for macroscopic constants (Equations (17), (18)) is an indication for dependence of PI(4,5)P<sub>2</sub>-binding sites. Macroscopic dissociation constants in the same order of magnitude for both binding events suggest positive cooperativity. All errors given for native MS data refer to the standard deviation and were based on at least three independent measurements (Table 1, Table 2).

#### 4.3.5 Native Mass Spectrometry: ENTH:ANTH:PI(4,5)P<sub>2</sub> Complexes

Complexes were assembled after buffer exchange by mixing 10 μM ENTH, 10 μM ANTH and 60 μM PI(4,5)P<sub>2</sub> (Phosphatidylinositol-4,5 biphosphate diC8, Echelon).

Native MS on protein complexes was performed on a QToF 2 (Waters and MS Vision) modified for high mass experiments with nano ESI source in positive ion mode (van den Heuvel, van Duijn et al., 2006). The gas pressures were 10 mbar in the source region and  $1.1 \times 10^{-2}$  mbar xenon (purity 5.0) in the collision cell (Lorenzen, Versluis et al., 2007, Tahallah, Pinkse et al., 2001). Mass spectra were recorded with applied voltages for capillary, cone, and collision of 1.35 kV, 120–150 V, and 30–60 V, respectively, optimized for resolution and minimal complex dissociation. Complexes were analysed in MS/MS by ramping collision voltages from 10 to 400 V in order to eject protein subunits. Raw data were calibrated with 25 mg/ml cesium iodide spectra of the same day with the software MassLynx (Waters). MassLynx and *Massign* (Morgner & Robinson, 2012) were used to assign peak series to complex species.

Average measured masses of protein complexes, standard deviations of replicate measurements and average full width at half maximum (FWHM) values as a measure of the mass heterogeneity and resolution are listed in **Supplementary Table S3**.

#### 4.3.6 Crystal Structure Determination

For structure determination of *S. cerevisiae* epsin-2 ENTH domains in presence of PI(4,5)P<sub>2</sub>, 6 mg/ml purified protein solution was incubated with 200 μM PI(4,5)P<sub>2</sub> at 4 °C for 2 h. After filtration through 0.22 μm centrifugal filters (Millipore), crystals were obtained in a vapour diffusion hanging drop setup. 1 μl protein-PI(4,5)P<sub>2</sub> solution was mixed with 1 μl of mother

liquor containing 0.1 M MES pH 6.5, 0.1 M NaCl, 1.45 M ammonium sulfate. Three cryo-cooled crystals soaked in paraffin oil were used to acquire a data set to 3.35 Å (space group F432) at PETRA III beamline P14 (DESY, Germany). The structure was solved by molecular replacement using PHASER with a 1.8 Å structure of epsin-2 ENTH domain (PDB ID: 6ENR, **Supplementary Table S4**) as search model, refined with Phenix (Afonine, Grosse-Kunstleve et al., 2012) and manually rebuilt with Coot (Emsley, Lohkamp et al., 2010), resulting in a final  $R_{\text{factor}}$  of 27.9 % ( $R_{\text{free}} = 30.3$  %). The stereochemistry was checked with MolProbity (Chen, Arendall et al., 2010), indicating good overall geometry with only 1.8 % of the residues in disallowed regions of the Ramachandran plot. Structure diagrams were prepared with Pymol (The PyMOL Molecular Graphics System, Version 1.7.x, Schrödinger, LLC) and Chimera (Pettersen, Goddard et al., 2004). For further experimental details see Garcia-Alai, Heidemann et al. (2018).

#### 4.3.7 Small-angle X-ray Scattering

PI(4,5)P<sub>2</sub>-containing human epsin-1 ENTH hexamers were assembled by mixing 80 μM protein and 200 μM PI(4,5)P<sub>2</sub> and incubation overnight at 4 °C. ENTH without PI(4,5)P<sub>2</sub> was used as a control. Synchrotron radiation X-ray scattering data were collected at PETRA III beamline P12 (DESY, Germany) (Blanchet, Spilotros et al., 2015) (**Supplementary Table S5**) with a PILATUS 2M pixel detector (DECTRIS, Switzerland). Solutions of human ENTH (0.4-1.5 mg/ml) were measured through a capillary at 20 °C. The sample-to-detector distance was 3.1 m, covering a range of momentum transfer  $0.01 \leq s \leq 0.46 \text{ \AA}^{-1}$  ( $s = 4\pi \sin\theta/\lambda$ ). Obtained data were normalized, averaged, buffer subtracted, and placed on an absolute scale that is relative to water, according to standard procedures. Data manipulations were performed using PRIMUSqt and the ATSAS software package (Petoukhov, Franke et al., 2012). From Guinier analysis, forward scattering  $I(0)$  and radius of gyration ( $R_g$ ) were determined. With program GNOM an indirect Fourier transform method was applied to obtain the distance distribution function  $p(r)$  and the maximum particle dimensions  $D_{\text{max}}$  (Semenyuk & Svergun, 1991). Molecular masses of solutes were estimated from SAXS data by comparing the extrapolated forward scattering with the respective data of a reference solution of glucose isomerase (Hampton), the hydrated particle/Porod volume  $V_p$ , where the molecular mass is estimated as 0.588 times  $V_p$ , and from the excluded solvent volumes,  $V_{\text{ex}}$  that were obtained from DAMMIF through *ab initio* modelling (Franke & Svergun, 2009). Rigid body modelling was performed using the program CORAL (Petoukhov, Franke et al., 2012). Further experimental details are described in Garcia-Alai, Heidemann et al. (2018).

#### 4.3.8 Dynamic Light Scattering

Measurements were performed on a DynaPro Nanostar (Wyatt Technology Corporation) after mixing 30  $\mu\text{M}$  protein and 200  $\mu\text{M}$  PI(4,5)P<sub>2</sub> and overnight incubation at 4 °C. Prior to measurements, samples were filtered through 0.22  $\mu\text{m}$  centrifugal filters (Millipore). Data were acquired in 20 cycles of 3 s.

#### 4.3.9 Isothermal Titration Calorimetry

A MicroCal VP-ITC calorimeter (GE Healthcare) was used at 25 °C. 100  $\mu\text{M}$  ENTH domains (*S. cerevisiae* ANTH: 1/150 of cell volume, *C. thermophilum* ANTH: 1/100 of cell volume) was titrated into the cell with 10  $\mu\text{M}$  *S. cerevisiae* ANTH domains that was premixed with 200  $\mu\text{M}$  PI(4,5)P<sub>2</sub>. 20 mM Tris-HCl buffer pH 8.0 was used for experiments with *S. cerevisiae* ANTH, 20 mM sodium phosphate buffer pH 8.0 with 250 mM NaCl was used for *C. thermophilum* ANTH.

#### 4.3.10 Ion Mobility Mass Spectrometry

Samples were prepared as described for native MS analysis (see 4.3.3 Native Mass Spectrometry, page 80) and analysed in positive ion mode on a commercially available travelling wave ion mobility mass spectrometer (Synapt G2, Waters). Measurements were performed in ion mobility separation mode using an RF generator that produced low frequencies allowing transmission of ions with up to 32,000  $m/z$  and with a pusher frequency of 1900  $\text{s}^{-1}$ .

ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes from *S. cerevisiae* were measured using 2300 V capillary voltage, 50 V cone voltage, 0.5 V extraction cone voltage, 15 V trap collision energy, 45 V trap DC bias, 25 V helium DC, and 5 V transfer collision voltage. Backing pressure was set to 8 mbar and 2.5 mbar nitrogen (60 ml/min IMS gas flow) was used for ion mobility separation. For ion separation, wave velocities of 650 m/s and wave heights of 40 V were used. Wave velocities in trap and transfer cell were 300 m/s and 64 m/s with wave heights of 6.0 V and 2.9 V, respectively. Argon was used as collision gas with  $4.6 \times 10^{-2}$  mbar and  $4.4 \times 10^{-2}$  mbar in trap (10 ml/min gas flow) and transfer cell, respectively.

Alcohol dehydrogenase from *S. cerevisiae* (#A7011, Sigma-Aldrich), concanavalin A from *C. ensiformis* (#C2010, Sigma-Aldrich), glutamate dehydrogenase from bovine liver (#G7882, Sigma-Aldrich), pyruvate kinase from rabbit muscle (#P9136, Sigma-Aldrich), and bovine serum albumin (#P7656, Sigma-Aldrich) were measured under the same conditions to estimate

collision CCSs of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes by using a logarithmic fit procedure as described (see 1.8 Ion Mobility Mass Spectrometry, page 23) (Bush, Hall et al., 2010, Ruotolo, Benesch et al., 2008).

Human PI(4,5)P<sub>2</sub>-containing ENTH hexamers were analysed using the following conditions: 1300 V capillary voltage, 25 V sampling cone, 4 V extraction cone voltage. Pusher frequency was 1900 s<sup>-1</sup>. Trap collision energy, trap DC bias, helium DC and transfer collision energy were set to 25V, 45 V, 25 V, and 5 V, respectively. Backing pressure was 8.5 mbar and 3.1 mbar nitrogen (90 ml/min IMS gas flow) was used for ion mobility separation. Argon was used as collision gas with 3.1 × 10<sup>-2</sup> mbar in the trap (4 ml/min gas flow) and 3.2 × 10<sup>-2</sup> mbar in the transfer cell. Travelling waves with heights of 25 V and velocities of 300 m/s were used for ion mobility separation. Wave velocities in trap and transfer cell were 300 m/s and 66 m/s with wave heights of 6.0 V and 2.0 V, respectively.

Alcohol dehydrogenase from *S. cerevisiae* (#A7011, Sigma-Aldrich), concanavalin A from *C. ensiformis* (#C2010, Sigma-Aldrich), and glutamate dehydrogenase from bovine liver (#G7882, Sigma-Aldrich) were measured under the same conditions to estimate CCSs of *H. sapiens* ENTH:PI(4,5)P<sub>2</sub> complexes as described above. In order to increase mass resolution and resolve charge states with different numbers of PI(4,5)P<sub>2</sub> molecules bound, ions were either activated in the source (175 V sampling cone voltage) or in the trap region (75 V trap collision energy).

Acquired data were processed and analysed using MassLynx (Waters) and DriftScope (Waters) software.

#### 4.3.11 Surface-induced Dissociation Mass Spectrometry

Samples were prepared as described for native MS analysis (see 4.3.3 Native Mass Spectrometry, page 80) and analysed in positive ion mode on a Synapt G2 instrument with an additional SID cell located upstream of the ion mobility separation cell to allow for ion mobility separation of SID products (Zhou & Wysocki, 2014). The SID cell contained a gold surface coated with 2-(perfluorodecyl)ethanethiol (Galhena, Dagan et al., 2008). Measurements were performed in ion mobility separation mode using an RF generator that produced low frequencies that allowed transmission of ions with up to 32,000 *m/z*.

The spectrum in transmission mode, meaning without a collision with the surface, was acquired with 1600 V capillary voltage and 80 V sampling cone voltage. Trap and transfer cell collision energies were both set to 10 V. Backing pressure was 8 mbar and 2.5 mbar nitrogen

were used for ion mobility separation. Travelling waves with heights of 20 V and velocities of 300 m/s were used in the IMS cell. .

Measurements in SID mode were performed with 1800 V capillary voltage and 50 V sampling cone voltage. Trap and transfer collision energies were set to 10 V and 4 V, respectively. Backing pressure was 8 mbar and 2.2 mbar nitrogen was used for ion mobility separation of SID products in the IMS cell. The potential difference between the Trap DC bias and the surface of the SID cell defined the SID collision energy. It was ramped from 30 V to 60 V and 90 V. For 60 V SID collision energy, an MS/MS spectrum was acquired with selection of the +37 ENTH:ANTH:PI(4,5)P<sub>2</sub> ion as precursor ion. Due to the low signal intensity in MS/MS mode, for 30 V, 60 V and 90 V SID collision energy, spectra with broader precursor selection windows were acquired. For that purpose, the MS profile was set to 10,000-15,000 *m/z*.

## 4.4 Outlook

Although the generated data allowed to draw a comprehensive picture of ENTH:ANTH:PI(4,5)P<sub>2</sub> complex formation during clathrin-mediated endocytosis, several questions remained unclear and need to be addressed in future experiments.

For instance, due to limited access to the instruments, a detailed characterization of the ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes by IMMS and SID MS was not possible. Especially the analysis of human complexes by IMMS might reveal further insights into the assembly process, since many different complex stoichiometries that potentially represent intermediate forms can be observed. Repeating SID experiments, as well with charge reduced complexes, possibly unveils further details of the complex structures. In addition, a comprehensive understanding of the mechanism in the present case helps to further fathom the general SID mechanism.

Of course, investigations can be extended to more ENTH and ANTH domains to better understand the interplay between different clathrin adaptor proteins. By using longer protein constructs with additional interaction sites for ubiquitin, clathrin or actin, a variety of research questions that cover different CME related topics, such as cargo recruitment, clathrin assembly or actin anchoring can be studied by native MS.

The role of clathrin adaptor proteins in infection biology needs to be analysed more thoroughly. For example, epsin-1 was reported to be a cargo-specific adaptor for endocytosis of influenza viruses. The interaction is mediated via the ubiquitin interaction motif of epsin, but the interaction partner and the precise mechanism remain unclear (Chen & Zhuang, 2008).

Further research on the role of PI(4,5)P<sub>2</sub> microdomains for the recruitment of ENTH and ANTH domains is required. Is the existence of these microdomains a prerequisite for a successful CME event? Can membrane-associated ENTH and ANTH domains move laterally on the membrane surface prior to assembly? Answers to such questions will hardly be found by native MS but rather require the application of different biophysical methods, such as microscopy.

To emphasize the informative value of biophysical experiments as presented in this work, results need to be verified for their correctness and physiological relevance by *in vivo* experiments.

In general, this study illustrates the accessibility for a detailed structural characterization of membrane-associated protein-lipid complexes by native MS in conjunction with other

biophysical methods. Compared to integral membrane proteins (see 3 Binding of Two Calmodulins to the Plasma-membrane  $\text{Ca}^{2+}$ -ATPase ACA8, page 27), sample handling and measurements are significantly facilitated when no solubilisation or reconstitution of the membrane proteins is required.



---

## 5 Molecular Organization of the *Salmonella* T3SS Sorting Platform

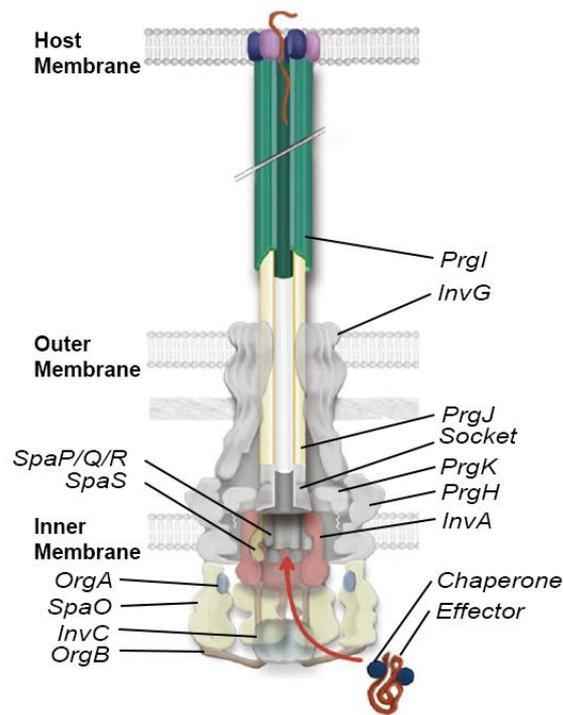
### 5.1 Introduction

#### 5.1.1 *Salmonella Typhimurium* Type III Secretion System

Many Gram-negative bacteria (e.g. *Shigella*, *Salmonella*, *E. coli*, *Vibrio*, *Yersinia*, *Chlamydia*, *Pseudomonas*) contain huge, membrane-spanning molecular syringes, called Type III Secretion System (T3SS). As indicated by their alternative name *injectisome*, these protein complexes secrete effector proteins from the bacterial cytoplasm directly into the cytoplasm of host cells. A large quantity of effector proteins has been identified that mostly induce subtle alterations in host cell physiology, such as cytoskeleton dynamics, cell cycle, transcription, protein degradation and vesicular trafficking (Dean, 2011, Galan, 2007, Gaytan, Martinez-Santos et al., 2016). Yet together, these multifunctional effectors extensively modulate the pathogen's environment in the host to be advantageous for infection. In *Salmonella enterica* serovar Typhimurium (*S. Typhimurium*) two T3SSs with specific functions were reported, encoded on the pathogenicity islands SPI-1 and SPI-2. While SPI-1 T3SSs are exclusively associated with gastrointestinal disease, T3SSs from SPI-2 are relevant for gastrointestinal and disseminated infection (Galan & Curtiss, 1989, Hensel, Shea et al., 1995). If not specified differently, it is referred to the T3SS from *S. Typhimurium* SPI-1 throughout this work.

The needle complex forms a 3.5 MDa molecular machine that structurally resembles bacterial flagella (Figure 31). Based on phylogenetic analyses it is assumed that evolutionary exaptation of the flagellar apparatus formed the injectisomes in a stepwise process (Abby & Rocha, 2012).

The secretion mechanism itself is tightly controlled and linked to T3SS assembly. To ensure complex assembly, proteins forming the needle and translocon of the complex are translocated first. Models suggest a switch to effector substrates by different mechanisms, sensing either conformational changes at the cytoplasmic side of the complex or the length of the needle (Journet, Agrain et al., 2003, Marlovits, Kubori et al., 2006). For secretion, chaperone-bound effector proteins are recognized on the cytoplasmic side of the complex.



**Figure 31: Schematic representation of Type III secretion systems.** Chaperone-bound effector proteins are selected for translocation by the sorting platform and then translocated across both bacterial membranes and the host membrane directly into the host cell. Protein names are annotated as established for *Salmonella* Typhimurium. Reprinted from “Type Three Secretion System in Attaching and Effacing Pathogens” by Gaytán, Martínez-Santos et al. (2016) from *Frontiers in Cellular and Infection Microbiology* (10.3389/fcimb.2016.00129), used under CC BY 4.0 license /Added labels.

As the inner diameter of the needle spans only 2 nm, translocation of substrates involves at least partial protein unfolding (Radics, Königsmair et al., 2014). The capability of the cytoplasmic ATPase (InvC) to unfold proteins in an ATP-dependent manner was proven in *in vitro* experiments and can probably be transferred to the *in vivo* mechanism (Akeda & Galán, 2005, Galán, Lara-Tejero et al., 2014).

The inner membrane complex, also known as export apparatus, is comprised of the transmembrane proteins SpaP/Q/R and was subject of detailed native MS studies (Dietsche, Tesfazgi Mebrhatu et al., 2016, Kuhlen, Abrusci et al., 2018).

### 5.1.2 The T3SS Sorting Platform

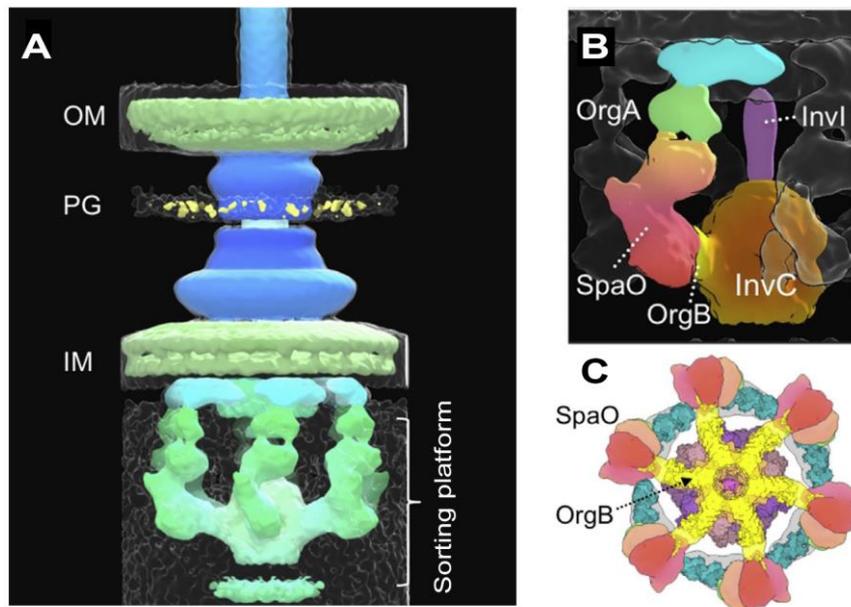
Being responsible for substrate selection, the cytoplasmic subcomplex of the T3SS was termed sorting platform (Lara-Tejero, Kato et al., 2011, Morita-Ishihara, Ogawa et al., 2006). In *S. Typhimurium*, the sorting platform of the SPI-1 injectisome consists of the ATPase InvC, its regulator protein OrgB, the cytosolic ring (C-ring) protein SpaO, and the accessory protein OrgA (Lara-Tejero, Kato et al., 2011). The most detailed structural characterization of the T3SS

sorting platform was recently presented by Hu, Lara-Tejero et al. (2017) and depicts the shape of the complex as determined by cryo-ET.

The results confirmed a previously reported overall shape of the sorting platform that deviates from related structures (**Figure 32A**) (Hu, Morado et al., 2015, Makino, Shen et al., 2016). While flagella and other related structural variants, e.g. from *Yersinia enterocolitica*, contain a continuous, wheel-like C-ring surrounding a central nave, the C-ring of the *S. Typhimurium* sorting platform is discontinuous, resulting in six pods that are linked to the central nave (Diepold, Kudryashev et al., 2015). In some instances, results from different structural studies on C-rings are not fully compatible. While cryo-ET studies visualized a discontinuous C-ring for the injectisome from *Shigella flexneri*, native mass spectra revealed the formation of higher order assemblies of C-ring proteins, thus suggesting an oligomerization and ring formation (Hu, Morado et al., 2015, McDowell, Marcoux et al., 2016). However, to resolve these discrepancies, it was suggested that the pods captured in cryo-ET are only the most stable C-ring components and that binding of additional subunits could be detrimental for substrate delivery (McDowell, Marcoux et al., 2016, Notti, Bhattacharya et al., 2015).

Moreover, the location of individual proteins in the sorting platform was assessed with injectisomes that contained tagged protein components and deletion mutant variants (Hu, Lara-Tejero et al., 2017). Based on differences in electron density maps of wild type and mutant injectisomes, the approximate position of proteins was determined (**Figure 32B**). In addition, atomic protein structures as available in the protein database or generated by *in silico* modelling, were fit into the cryo-ET maps (**Figure 32C**).

The ATPase InvC is located at the central nave of the sorting platform and assembles to hexamers (Claret, Calder et al., 2003). An atomic structure of the hexamer was modelled based on available crystal structures from related T3SS ATPases (Kato, Lefebvre et al., 2015). InvC was demonstrated to be involved in the recognition of chaperone-bound substrates for secretion, as well as in the unfolding of the substrate prior to transport (Akedo & Galan, 2005). The system is fuelled by ATP hydrolysis but also reliant on the proton motive force (Minamino & Namba, 2008). Some experimental results even point to a rotation-dependent T3SS secretion mechanism, similar to the rotation of flagella (Ohgita, Hayashi et al., 2013). The presence of the stalk protein InvI between the InvC hexamer and the nonameric InvA ring located in the inner membrane suggests a functional interplay between ATPase and inner membrane ring (Hu, Lara-Tejero et al., 2017).



**Figure 32: *Salmonella* Typhimurium T3SS structure as determined by electron cryo tomography.** A) A 3D injectisome surface structure was determined *in situ* to 17 Å resolution by cryo-ET. Blue fragments correlate well with single particle electron microscopy structures from isolated needle complexes. Locations of outer membrane (OM), peptidoglycan layer (PG) and inner membrane (IM), as well as the sorting platform are annotated. B) Locations of individual sorting platform proteins as determined using tagged protein variants and deletion mutants in side view. C) Bottom view of the injectisome with available and modelled atomic structures (PDB IDs: 5B0O, 3AJW, 4YX7). Adapted with permission from Elsevier and Copyright Clearance Center: Elsevier, Cell, “In Situ Molecular Architecture of the *Salmonella* Type III Secretion Machine” (Hu, Lara-Tejero et al.), Copyright 2017.

In the electron density map, six spokes connect the central nave of the sorting platform with the surrounding, discontinuous C-ring. The electron density was associated with the stator protein OrgB. For OrgB homologs, e.g. from *Yersinia*, an ATPase inhibiting function was reported upon protein overexpression (Blaylock, Riordan et al., 2006). Possibly, this inhibiting function ensures ATP hydrolysis only after assembly of the injectisome complex (Romo-Castillo, Andrade et al., 2014). Moreover, it was suggested that a conformational change of OrgB upon binding to membrane-embedded proteins (homologous to InvA, SpaS) of the injectisome triggers oligomerization and activation of the ATPase InvC (Ibuki, Imada et al., 2011, Romo-Castillo, Andrade et al., 2014).

Structurally, limited information is available about OrgB. Yet, the C-terminus of the flagellar homolog FliH was described to form a dimer that binds to the InvC-homolog FliI (Imada, Minamino et al., 2016). Sequence-based prediction tools identified an N-terminal disordered region in OrgB, followed by a coiled-coil domain (Notti, Bhattacharya et al., 2015). For OrgB homologs, binding of an extended stretch of the N-terminus to InvA homologs was reported

(Hara, Morimoto et al., 2012). In addition, a crystal structure of OrgB amino acid residues 1-30 in complex with the C-terminal domains of SpaO was solved (Notti, Bhattacharya et al., 2015). If both interactions to InvA and SpaO form at the same time or whether dynamic changes in interaction partners occur, is not clarified.

Of all sorting platform components, SpaO was studied in most detail. As indicated before, continuous and discontinuous C-rings consisting of proteins homologous to SpaO were reported. Recent cryo-ET studies of *S. Typhimurium* show discontinuous rings with SpaO capping the OrgB spokes forming pod structures (Hu, Lara-Tejero et al., 2017). At the top of these pods, SpaO proteins are physically linked to OrgA and thereby fixed to distal injectisome components (**Figure 32**). Structurally, SpaO is composed of an N-terminal domain and two surface presentation of antigen domains (SPOA1 and SPOA2) at the C-terminus. As shown for homologous injectisomes, SpaO contains an internal translation initiation site (Bzymek, Hamaoka et al., 2012). The short isoform that is produced in addition to the full-length protein, comprising the SPOA2 domain, will be referred to as SpaOc (Song, Sukovich et al., 2017). Structures of SpaOc homodimers and SPOA1-2 domains, the latter also in complex with OrgA residues 1-30, were solved by X-ray crystallography (Notti, Bhattacharya et al., 2015). For homologous proteins, two SpaOc and one SpaO were shown to assemble to heterotrimers (Bzymek, Hamaoka et al., 2012). Studies proposed SpaOc homologs to be essential components of T3SS sorting platforms (Bzymek, Hamaoka et al., 2012, Diepold, Kudryashev et al., 2015). However, the homolog from *Salmonella* SPI-2 was described as a stabilizing chaperone of the full-length gene product, which is not essential for secretion (Yu, Liu et al., 2011). The precise role of SpaOc in the pods of the *S. Typhimurium* SPI-1 sorting platform remains elusive.

Since type III secretion needs to be precisely controlled, a dynamic exchange of SpaO subunits was shown and proposed to have a regulatory function for selection and export of substrates (Diepold, Kudryashev et al., 2015). In conjunction with the idea that the cryo-ET pod-structures are biased for C-ring components that are fixed to the inner membrane ring of the injectisome, an attractive model of the C-ring being a dynamic hub with SpaO subunits attaching and detaching was hypothesized (McDowell, Marcoux et al., 2016, Notti, Bhattacharya et al., 2015).

## 5.2 Results and Discussion

### 5.2.1 *SpaO*-Isoforms Assemble to *SpaO*-2*SpaO*<sub>C</sub> Complexes

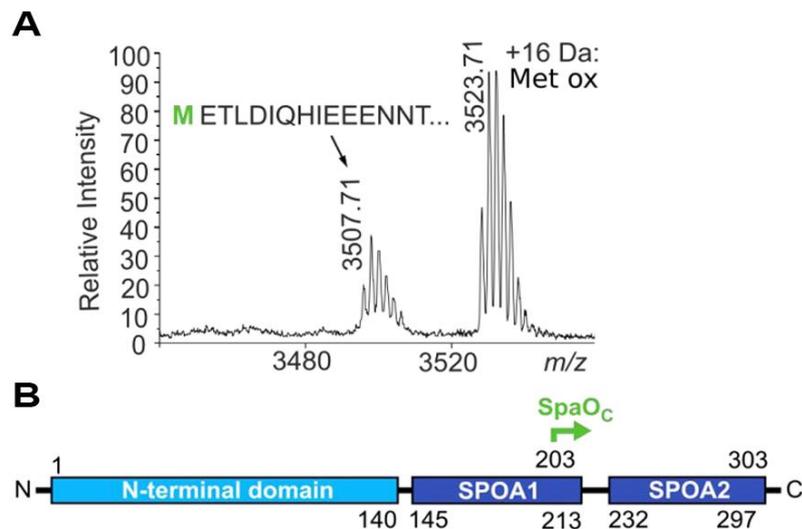
In a first set of experiments, the two isoforms of the *S. Typhimurium* T3SS C-ring protein *SpaO* were characterized. As previously reported for homologs and *Salmonella* proteins, a full-length *SpaO* protein (*SpaO*/*SpaO*<sub>FL</sub>) and a short *SpaO* isoform (*SpaO*<sub>C</sub>) comprising of the C-terminal part of the *spaO* gene were identified (Bzymek, Hamaoka et al., 2012, Diepold, Kudryashev et al., 2015, Song, Sukovich et al., 2017, Yu, Liu et al., 2011).

An in-gel tryptic digest of *SpaO*<sub>C</sub> was performed and the tryptic peptides were analysed by MALDI MS (Figure 33A, Supplementary Figure S11A). The amino acid sequence of the N-terminus of *SpaO*<sub>C</sub> was found to be initiated by a methionine residue instead of valine as encoded for amino acid 203 in the *spaO* gene (Figure 33B). The +16 Da mass shift indicated a partial oxidation of methionine residues that probably took place during sample preparation. Results were verified in MALDI MS/MS measurements by sequencing of the 3,523.71 *m/z* oxidized precursor peptide (Figure S11B). Various *b*- and *y*-fragments were assigned, thereby confirming the N-terminal amino acid sequence of *SpaO*<sub>C</sub> (Supplementary Table S9). The presence of an N-terminal methionine in *SpaO*<sub>C</sub> strongly opposes a proteolysis of full-length *SpaO* and argues that *SpaO*<sub>C</sub> is the product of a translation initiation event with a GTG start codon at position 203.

Summing up, these results map the synthesis of two *SpaO* isoforms. The full-length *SpaO* protein consists of the N-terminal domain, the SPOA1 domain, and the SPOA2 domain (Figure 33B). Internal translation initiation at codon 203 leads to the production of a short protein isoform (*SpaO*<sub>C</sub>) containing the SPOA2 domain.

The importance of *SpaO* and *SpaO*<sub>C</sub> for *S. Typhimurium* type III secretion was evaluated in secretion assays. To that end, the presence of type III secretion effector proteins (*SipB*, *SipC*, *SopB*, *SipD*) in culture supernatants of different *Salmonella* strains was detected by western blot analysis (Supplementary Figure S12A). No effector proteins were detected in the supernatant of *spaO* knockout mutants. Similarly, loss of full-length *SpaO* caused a complete shutoff of the T3SS. On the contrary, abrogation of *SpaO*<sub>C</sub> production led to a reduced secretion, as indicated by lower intensities of bands in western blot analysis. Complementation with *spaO*<sub>C</sub> partly restored secretion. Western blot analyses of *Salmonella* cell lysates confirmed that effector proteins were present in the bacteria even though the secretion process was hampered or abrogated (Supplementary Figure S12B). However, low concentrations of *SpaO*<sub>C</sub> were detected in  $\Delta$ *spaO* *Salmonella* strains, which carried a *spaO*-containing plasmid with

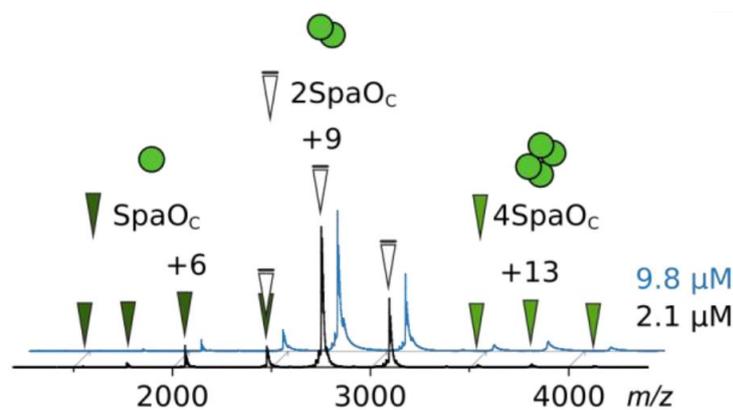
mutations in the putative internal ribosomal binding site or in the internal start codon (Supplementary Figure S12C). Consequently, these basal SpaOc protein levels could cause the observed residual activity of the  $\Delta spaOc$  *Salmonella* strain in the secretion assay. Most recent MALDI MS measurements of SpaOc produced by the *spaO*<sub>V203A</sub> mutant disclosed either a methionine or an alanine at position 203. The identification of methionine points to an incomplete shutoff of SpaOc translation in the *spaO*<sub>V203A</sub> mutant and the alanine to another mechanism for production of SpaOc, such as proteolysis.



**Figure 33: SpaOc is produced by translation from an internal initiation start site in the *spaO* gene.** A) MALDI MS of the tryptic digest of SpaOc revealed that the N-terminal peptide of SpaOc contains a methionine or an oxidized methionine in position 1. Masses of the peptides and the N-terminal sequence are annotated. Spectra of a wider *m/z*-range and MS/MS analysis are presented as **Supplementary Figure S11**. B) A schematic representation illustrates the SpaO domain structure. Sequence positions of domain boundaries and the SpaOc initiation site are labelled. Adapted from Bernal, Börnicke et al. (*in preparation*).

These results provide insights into the importance of the *spaO* gene products for type III secretion. Full-length SpaO is an essential component of the injectisome as deletion mutants showed no secretion of effector proteins. In contrast, deletion mutants of *spaOc* did not cause a shutoff but only a reduction of T3SS secretion. However, most recent results hint at a basal level of SpaOc, which is produced despite the mutations. The mechanism of SpaOc production needs to be addressed in future experiments and a confirmation of the additional SpaOc production by proteolysis would further highlight the importance of SpaOc for type III secretion. These findings imply that SpaOc is rather an integral factor for *Salmonella* SPI-1 type III secretion, similar as reported for proteins from *Yersinia* (Diepold, Kudryashev et al., 2015), and not just a chaperone for the full-length C-ring protein as proposed for *Salmonella* SPI-2 proteins (Yu, Liu et al., 2011).

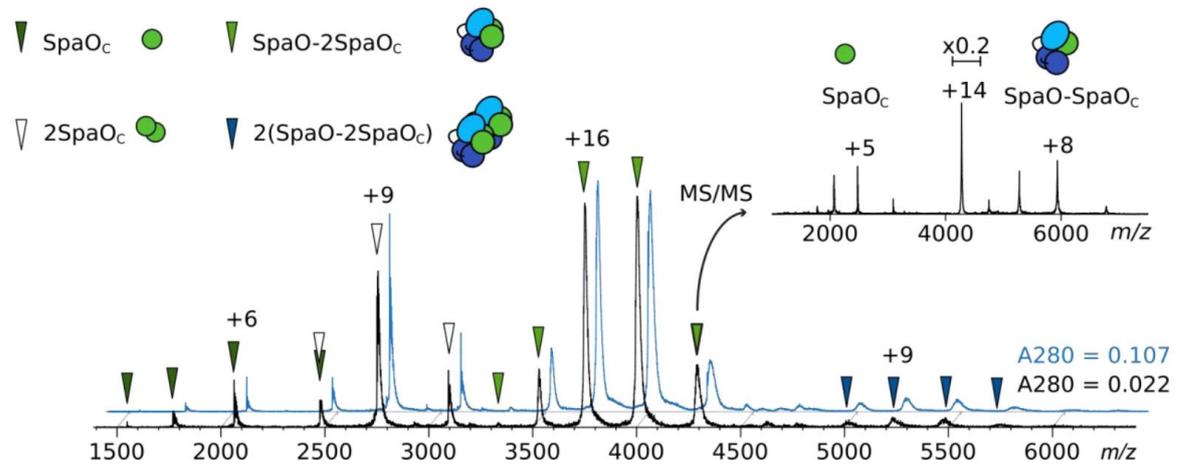
SpaOc proteins were subjected to native MS analysis and spectra revealed that the protein is prone to form dimers (Figure 34, experimental and theoretical masses, standard deviations and FWHM values for all proteins and protein complexes relevant for this chapter are listed in Supplementary Table S10). Homodimers were the predominant oligomerization state at two different protein concentrations. Monomers and homotetramers were observed in low amounts at both measured concentrations. The specificity of the interaction is reassured by rather constant signal intensity ratios for the oligomeric species at both tested protein concentrations. Moreover, based on previous results from homologous proteins (e.g. *S. Typhimurium* SPI-2 SsaQs), homodimerization of the short SpaO isoform was expected (Yu, Liu et al., 2011).



**Figure 34: Dimerization of SpaOc proteins.** A native mass spectrum of *S. Typhimurium* SpaOc revealed few monomers (dark green arrows) and homotetramers (light green arrows), while the majority of the protein was observed as homodimers (white arrows). Two different protein concentrations (blue spectrum, black spectrum) revealed comparable signal intensity ratios, thus arguing against extensive unspecific clustering during the ESI process. Adapted from Bernal, Börnicke et al. (*in preparation*).

Subsequently, co-purified C-terminally Strep-tagged SpaO/SpaOc was analysed by native MS (Figure 35). As expected from previous measurements, SpaOc was identified in low amounts as a monomer and in higher amounts as a homodimer. In addition, a signal of a 59.9 kDa complex was found. The comparison with expected theoretical masses of SpaO and SpaOc pointed to the presence of a heterotrimer with a SpaO-2SpaOc stoichiometry. CID MS/MS analysis of these ions confirmed the heterotrimer that dissociated upon high-energy collisions with argon into a SpaOc monomer and a residual SpaO-SpaOc heterodimer. The absence of monomeric SpaO in native MS measurements emphasises a high binding-affinity to SpaOc dimers and highlights the critical role of SpaOc for SpaO solubility. Moreover, a dimerization of heterotrimers to 2(SpaO-2SpaOc) heterohexamers was found in low amounts. Although the

presence of this complex was confirmed at any tested protein concentration, signal intensities were low and a specificity of this complex cannot be claimed with absolute certainty.



**Figure 35: SpaO isoforms form stable SpaO-2SpaOc complexes.** Native mass spectra disclosed monomers (dark green arrows) and dimers of SpaOc (white arrows). Furthermore, the efficient formation of SpaO-2SpaOc complexes (light green arrows), as well as further dimerization to heterohexamers (dark blue arrows) was observed. Spectra were acquired at two different protein concentrations (blue spectrum, black spectrum). Ratios of signal intensities were largely unaltered, indicating that concentrations were used that are below the levels where considerable unspecific protein clustering occurs. When the +14 SpaO-2SpaOc ion was subjected to CID MS/MS analysis, a SpaOc monomer and a residual SpaO-SpaOc complex were detected (inset). For better visibility of dissociation products, the precursor peak was scaled to 20 % of its original size. Both SpaO and SpaOc carried a C-terminal Strep-tag. Adapted from Bernal, Börnicke et al. (*in preparation*).

Further complex oligomerization was only registered at high protein concentrations that do not allow to distinguish a specific complex formation from ESI clustering products (see 1.6.3 Protein-Protein Interactions, page 17). Therefore, the reported oligomerization of *Shigella flexneri* T3SS C-ring proteins (McDowell, Marcoux et al., 2016) is not transferable to the *S. Typhimurium* SPI-1 injectisome. Results from multi-angle light scattering coupled to size exclusion chromatography (SEC-MALS) confirmed these observations. Even with protein concentrations as high as 140  $\mu$ M, no complexes larger than heterohexamers were identified (**Supplementary Figure S13A**).

Obviously, using tagged proteins to study protein-protein interactions potentially creates a biased picture of the physiological situation. To reduce the risk of sterical clashes induced by protein tags, native MS measurements were repeated with N-terminally Strep-tagged SpaO proteins (**Supplementary Figure S13**). No striking difference between spectra with Strep-tags on either side of SpaO proteins was recognized. The only distinct feature of Strep-SpaO/SpaOc complex formation was a low-abundant 2SpaO-2SpaOc species that was not present for C-

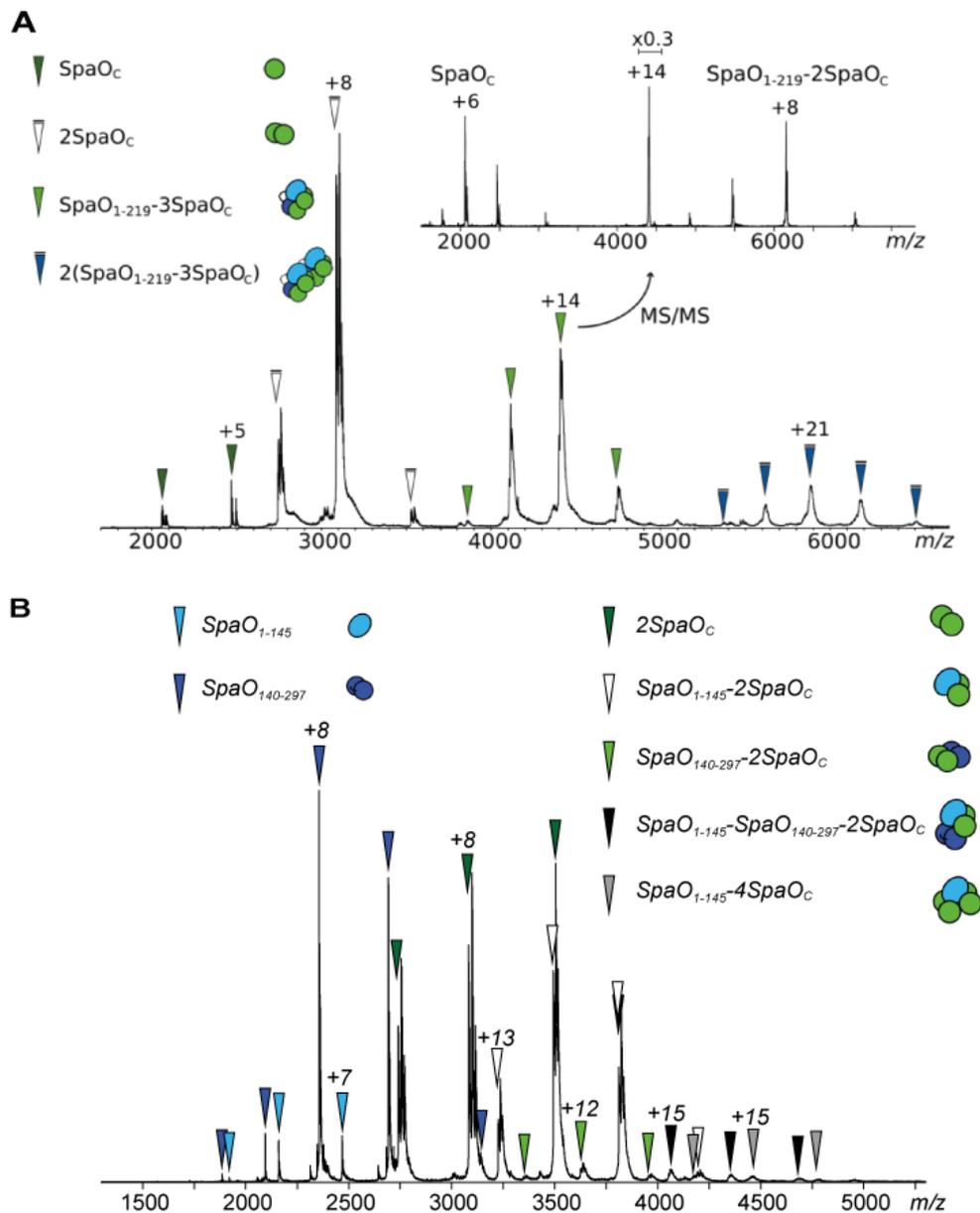
terminally Strep-tagged SpaO proteins. Possibly, the C-terminal Strep-tag forestalled a low-affinity interaction between SpaO-2SpaOc heterotrimers and an additional SpaO protein.

### 5.2.2 SpaO Domain Interactions

In order to characterize the interacting domains in SpaO-2SpaOc heterotrimers in more detail, native MS and SEC-MALS analyses of protein constructs covering different SpaO domains were performed.

For a first domain interaction experiment, SpaO was truncated after the first SPOA1 domain, giving rise to a SpaO<sub>1-219</sub> construct. Native MS measurements of SpaO<sub>1-219</sub> mixed with SpaOc revealed the effective formation of SpaO<sub>1-219</sub>-3SpaOc complexes (**Figure 36A**). Single SpaOc proteins dissociated in CID MS/MS analyses, leading to the formation of a residual SpaO<sub>1-219</sub>-2SpaOc complex. These results point out that the covalent linkage between SPOA1 and SPOA2 domains in SpaO is not essential for the formation of SpaO-2SpaOc complexes. The absence of SpaO<sub>1-219</sub>-2SpaOc in the mass spectra emphasises the high affinity of SPOA domain dimerization. The capability to form higher order dimers, in this case 2(SpaO<sub>1-219</sub>-3SpaOc), was not disturbed by shortening SpaO. In fact, the signals in the presented spectrum (**Figure 36A**) were higher than previously shown for 2(SpaO-2SpaOc). Yet, the protein concentration during measurements of the SpaO<sub>1-219</sub> complexes was around three fold higher than for SpaO-2SpaOc complexes (**Figure 35**).

Next, SpaO was split into two proteins covering either the N-terminal domain (SpaO<sub>1-145</sub>) or the C-terminal SPOA1-SPOA2 domains (SpaO<sub>140-297</sub>). Different from SpaO, SpaO<sub>1-145</sub> was stable in solution and subsequently the interactions with the SPOA-containing constructs SpaOc and SpaO<sub>140-297</sub> were investigated. Native mass spectra revealed an interaction of SpaO<sub>1-145</sub> and SpaOc homodimers (**Supplementary Figure S14A**). In contrast, mass spectra gave no indication for a stable interaction of SpaO<sub>1-145</sub> and SpaO<sub>140-297</sub>, as the very low signals corresponding to SpaO<sub>1-145</sub>-SpaO<sub>140-297</sub> complexes were likely ESI clustering artefacts (**Supplementary Figure S14B**). These findings imply that the N-terminal domain of SpaO and the C-terminal dimer of SPOA1-SPOA2 do not form a compact structure with stable non-covalent interactions between the domains.



**Figure 36: Inter- and intramolecular domain interactions of SpaO-2SpaOc complexes.** **A)** Native mass spectrum of SpaO<sub>1-219</sub>/SpaOc revealed formation of SpaO<sub>1-219</sub>-3SpaOc complexes (light green arrows) that further dimerized to a low extent (dark blue arrows). CID MS/MS of +14 SpaO<sub>1-219</sub>-3SpaOc ions showed dissociation of SpaO<sub>C</sub>, leaving a residual SpaO<sub>1-219</sub>-2SpaOc complex behind. The CID MS/MS precursor peak was scaled to 30 % peak height. **B)** A native mass spectrum of SpaO<sub>1-145</sub>/SpaO<sub>140-297</sub>/SpaOc contained signals of the individual proteins SpaO<sub>1-145</sub> (light blue arrows), SpaO<sub>140-297</sub> (dark blue arrows) and 2SpaO<sub>C</sub> (dark green arrows). In addition, SpaO<sub>1-145</sub>-2SpaO<sub>C</sub> (white arrows) complexes were detected. Signals of SpaO<sub>1-145</sub>-SpaO<sub>140-297</sub> (light green arrows), SpaO<sub>1-145</sub>-SpaO<sub>140-297</sub>-2SpaO<sub>C</sub> (black arrows) and SpaO<sub>1-145</sub>-4SpaO<sub>C</sub> had a low intensity and unspecific clustering in the ESI process cannot be excluded. Reprinted from Bernal, Börnicke et al. (*in preparation*).

Furthermore, there is a binding site for SpaOc dimers on the N-terminal SpaO domain, which is functioning independently of the SPOA1-2 domains. Remarkably, this binding site cannot be occupied by the structurally very similar SPOA1-2 dimer (SpaO<sub>140-297</sub>). Previous crystallization studies reported a root mean square deviation (RMSD) of 2.47 Å for C $\alpha$  traces of SpaOc dimers and SPOA1-2 (Notti, Bhattacharya et al., 2015). The structures of SPOA2 domains in complex with SPOA1 and in SpaOc homodimers overlap well with a C $\alpha$ -trace RMSD of 1.67 Å (Notti, Bhattacharya et al., 2015).

Mixing the partitioned SpaO protein (SpaO<sub>1-145</sub> and SpaO<sub>140-297</sub>) with SpaOc gave the same results in native MS as mixing N- or C-terminal parts of SpaO with SpaOc individually (**Figure 36B**). In addition, SpaO<sub>1-145</sub>-SpaO<sub>140-297</sub>-2SpaOc complexes were barely detectable, thereby emphasising the importance of the covalent linkage between the N-terminal SpaO and the SPOA domains. Furthermore, hardly any SpaO<sub>1-145</sub>-4SpaOc was observed, inferring that the deleted SPOA1-2 domains cannot be efficiently replenished by an additional SpaOc dimer.

In order to verify the results obtained in native MS, samples were also subjected to SEC-MALS analysis. Only stable SpaO<sub>1-145</sub>-2SpaOc complexes were identified, thereby confirming previously described native MS results (**Supplementary Figure S15**).

All domain interaction studies were performed with tagged protein variants. In particular, SpaO<sub>1-145</sub> contained a C-terminal Strep-tag while SpaO<sub>140-297</sub> had an N-terminal Strep-tag. Together with the overlap of five amino acid residues in the used constructs, these tags could interfere substantially with the formation of a complex between the N-terminal SpaO domain and the C-terminal SPOA domains. Results need to be validated with untagged protein constructs that do not have an overlap in the amino acid sequence to exclude steric hindrances.

Summing up, the covalent linkage between SPOA1 and SPOA2 domains is dispensable for the formation of SpaO-2SpaOc heterodimers. SpaOc dimers bind to the N-terminus of SpaO proteins, whereas SPOA1-2 dimers are not binding to the N-terminal domains. Hence, these results further demonstrate that SPOA1-2 dimers from the SpaO C-terminus and SPOA2 homodimers (equivalent to 2SpaOc) are not interchangeable for the formation of SpaO-2SpaOc heterotrimers, despite their structural similarity.

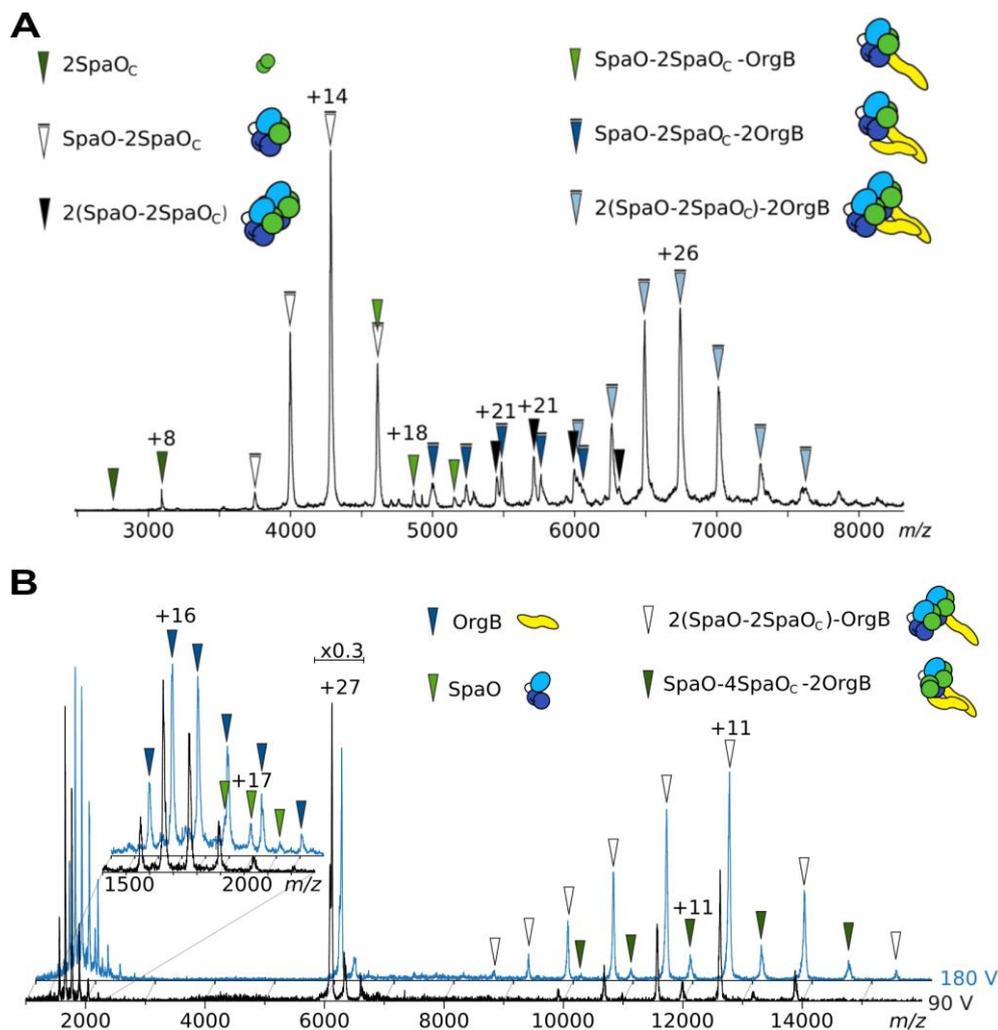
### 5.2.3 *OrgB* Dimers Induce Dimerization of SpaO-2SpaOc

Next, complex formation of the SpaO isoforms with OrgB was investigated. OrgB was insoluble when produced individually and accumulated in inclusion bodies. In contrast, co-expression with *spaO* caused a substantial increase of OrgB solubility (**Supplementary Figure S16**). The observed change in OrgB solubility hinted at SpaO/SpaOc/OrgB interactions that were analysed by native MS.

As expected from previous studies, SpaOc homodimers, SpaO-2SpaOc heterotrimers and dimerized heterotrimers with a 2(SpaO-2SpaOc) stoichiometry were observed (**Figure 37A**). No unbound OrgB was detected, confirming the observations in protein solubility during recombinant gene expression and protein purification. Instead, a strong signal of a 165 kDa complex was observed that was identified to originate from 2(SpaO-2SpaOc)-2OrgB complexes. Notably, this signal was stronger than the 2(SpaO-2SpaOc) complexes, indicating that OrgB dimers not only bind to preformed 2(SpaO-2SpaOc) complexes, but also induce dimerization of SpaO-2SpaOc heterotrimers. Only low signals of SpaO-2SpaOc-2OrgB complexes were detected, possibly representing an assembly intermediate or the product of dynamic dissociations and associations of SpaO-2SpaOc heterotrimers.

In protein complexes characterized by native MS, OrgB was mostly identified as a dimer. Only a very low-abundant SpaO-2SpaOc-OrgB complex was found. This result is in agreement with previous studies on homologous flagellar proteins that described the C-terminal fragment of the protein forming an asymmetric dimer (Imada, Minamino et al., 2016).

Complex stoichiometries were confirmed by CID MS/MS measurements on OrgB-containing complexes (**Figure 37B**). From all selected precursor ions the same dissociation pattern was recognized with OrgB dissociating predominantly from the complexes. In a second, less populated pathway, dissociation of SpaO from 2(SpaO-2SpaOc)-2OrgB complexes was observed. For both pathways, the respective residual complex after dissociation was identified.



**Figure 37: Dimerization of SpaO-2SpaOc is induced by binding of OrgB dimers.** **A**) As expected, in samples containing SpaO/SpaOc/OrgB the previously identified complexes 2SpaOc (dark green arrows), SpaO-2SpaOc (white arrows), and 2(SpaO-2SpaOc) (black arrows) were identified in native MS. In addition, 2(SpaO-2SpaOc)-2OrgB complexes (light blue arrows) were observed with a high signal intensity. Low intensity signals of SpaO-2SpaOc-OrgB (light green arrows) and SpaO-2SpaOc-2OrgB (dark blue arrows) were detected. **B**) +27 2(SpaO-2SpaOc)-2OrgB ions were subjected to CID MS/MS analysis at 90 V (black spectrum) and 180 V (blue spectrum) collision energy. At both energies, primarily OrgB (blue arrows) dissociated from the complex, giving rise to residual 2(SpaO-2SpaOc)-OrgB complexes (white arrows). SpaO (light green arrows) dissociated less frequently, the corresponding SpaO-4SpaOc-2OrgB complex (dark green arrows) was also identified. The precursor peak was scaled to 30 % signal intensity. Adapted from Bernal, Börnicke et al. (*in preparation*).

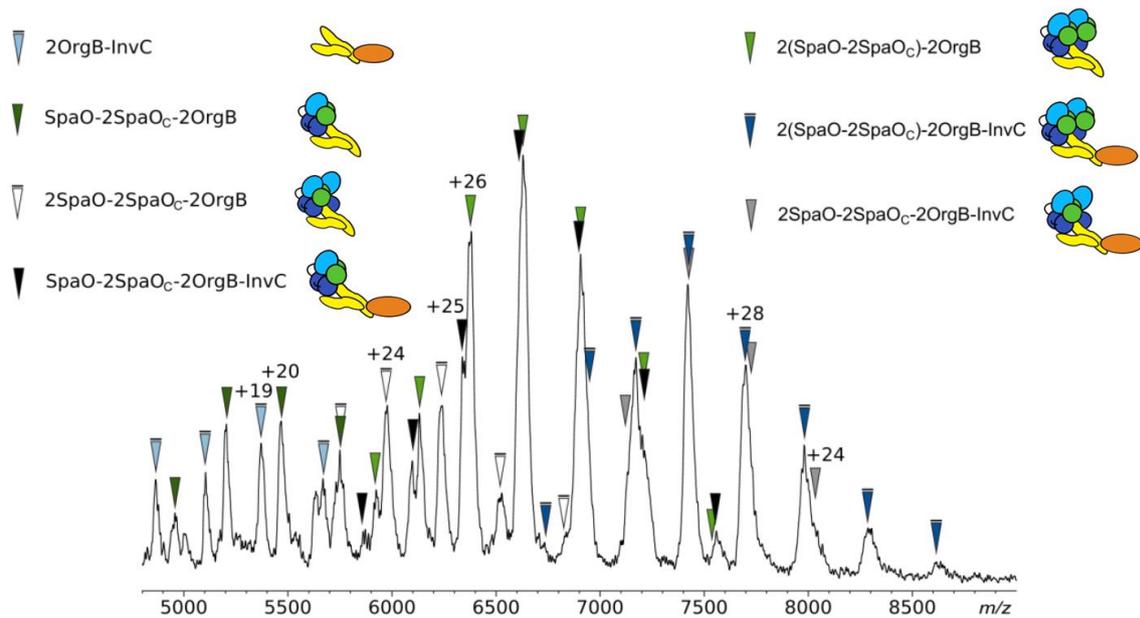
By considering the general CID mechanism, some structural features of the OrgB-containing complexes can be derived. Although the process of CID is not completely understood and exceptions have been reported, mostly monomeric proteins from the periphery of protein complexes dissociate (Benesch, 2009). As large proteins generally require more energy for unfolding, an additional general preference for small proteins seems self-evident.

In the present case, the 26.4 kDa protein OrgB or the 33.8 kDa protein SpaO dissociated from the complex and not SpaOc, which has less than half the molecular weight (12.4 kDa). In SpaO-2SpaOc heterotrimers and dimers thereof, exclusively SpaOc dissociation was observed. Therefore, OrgB dimers likely bind directly to SpaOc in the complex and thereby prevent their dissociation. The fact that SpaO dissociation occurred and the SpaOc dimer was present in the residual SpaO-4SpaOc-2OrgB complex further confirms the OrgB/2SpaOc interaction. Alternatively, OrgB binding induces structural changes in the complex that lead to SpaOc protection.

In contrast to these results, OrgB was reported to bind to the SPOA1-2 domain and the structure of the complex was solved by X-ray crystallography (Notti, Bhattacharya et al., 2015). Possibly, the presented SPOA1-2/OrgB structure illustrates a low-affinity interaction that was obtained in presence of the molecular crowding agent PEG400. Provided, the assembly of functioning injectisomes in absence of SpaOc is possible, as indicated by results from secretion assays, an interaction between OrgB and SpaO must be possible (see 5.2.1 SpaO-Isoforms Assemble to SpaO-2SpaOc Complexes, page 94). However, based on the structural similarity of SPOA1-2 and 2SpaOc an additional interaction of OrgB with 2SpaOc does not seem to be very far-fetched.

#### *5.2.4 Soluble Building Blocks Result from InvC Binding to SpaO/SpaOc/OrgB*

In the most extensive native MS experiment in this study, complexes of SpaO, SpaOc, OrgB and InvC were characterized. Although a large fraction of InvC accumulated in inclusion bodies, it was possible to express and purify the protein together with SpaO/SpaOc and OrgB (**Supplementary Figure S16**). Native mass spectra revealed that the ATPase InvC is forming different types of complexes with the other sorting platform components (**Figure 38**). Low-abundant species with 2OrgB-InvC or 2SpaO-2SpaOc-2OrgB-InvC stoichiometries were identified. Yet, the most abundant species were SpaO-2SpaOc-2OrgB-InvC and 2(SpaO-2SpaOc)-2OrgB-InvC complexes.



**Figure 38: The ATPase InvC binds to SpaO/SpaOc/OrgB complexes.** InvC-containing complexes with the stoichiometries 2OrgB-InvC (light blue arrows), SpaO-2SpaOc-2OrgB-InvC (black arrows), 2(SpaO-2SpaOc)-2OrgB-InvC (dark blue arrows) and 2SpaO-2SpaOc-2OrgB-InvC (grey arrows) were identified. In addition, SpaO/SpaOc/OrgB complexes were observed: SpaO-2SpaOc-2OrgB (dark green arrows), 2SpaO-2SpaOc-2OrgB (white arrows) and 2(SpaO-2SpaOc)-2OrgB. Combining MS with MS/MS analyses (**Figure 39**) was necessary to assign the different complex stoichiometries unambiguously from overlapping peaks. Adapted from Bernal, Börnicke et al. (*in preparation*).

Signal intensity ratios of the different complex species were dependent on the electrospray conditions and fairly instable, consequently the signal intensities of the presented spectrum (**Figure 38**) cannot be translated into complex ratios in solution. Nevertheless, based on previous SpaO/SpaOc/OrgB measurements that revealed a strong dimerization of SpaO-2SpaOc heterotrimers, species like SpaO-2SpaOc-2OrgB-InvC and 2SpaO-2SpaOc-2OrgB-InvC were not expected in such high abundances.

Importantly, InvC was exclusively detected in complexes containing OrgB dimers, demonstrating a direct interaction between these proteins as previously reported in cryo-ET maps and pull-down assays (Hu, Lara-Tejero et al., 2017, Notti, Bhattacharya et al., 2015).

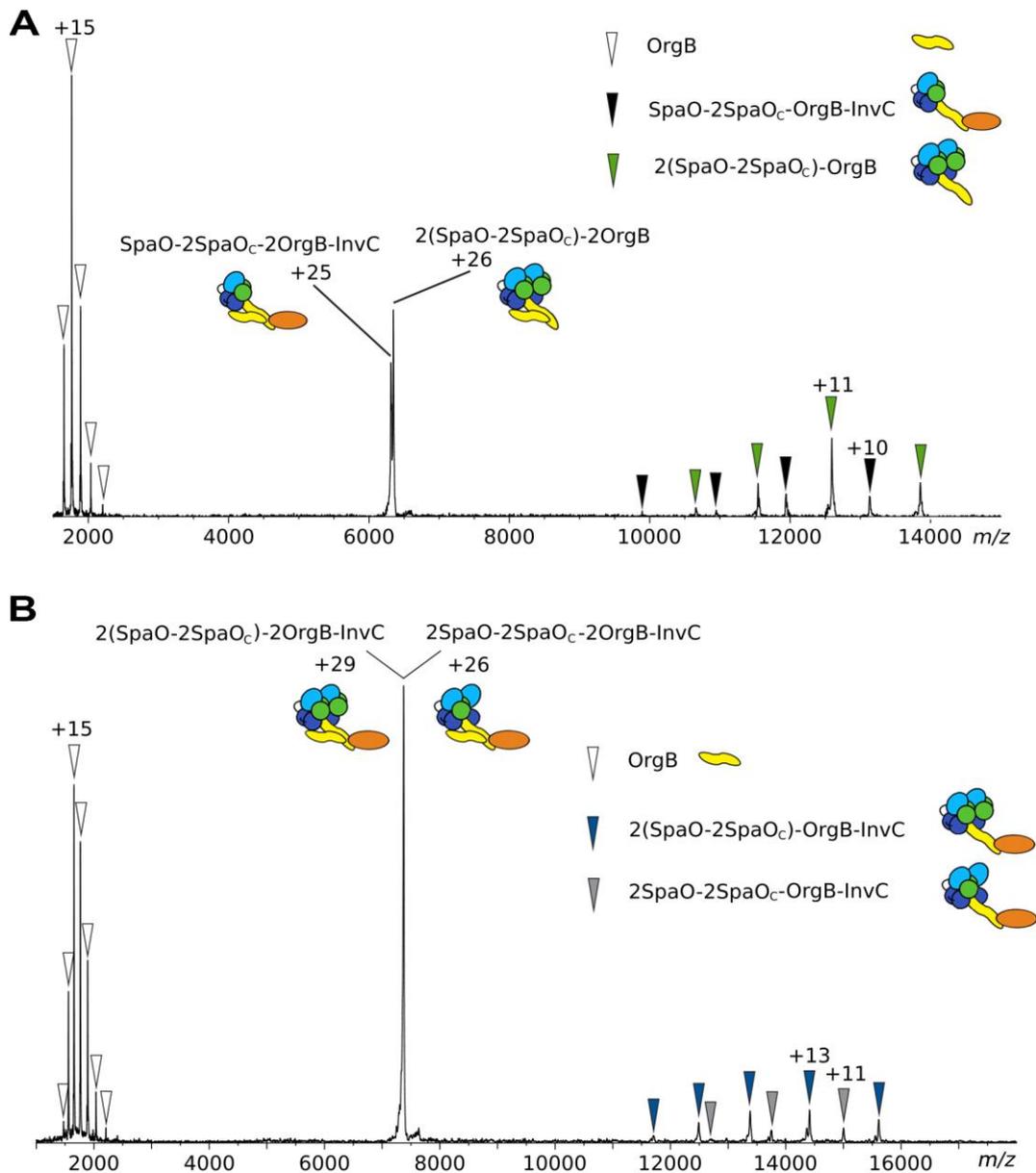
CID MS/MS measurements on all InvC-containing complexes unveiled dissociation of an OrgB monomer (**Figure 39**). The residual complexes consist of at least a SpaO-2SpaOc heterotrimer, OrgB and InvC. From this dissociation pattern, it is obvious that the same OrgB protein is bound to the SpaO-2SpaOc heterotrimer and to InvC. Possibly, one OrgB monomer establishes these interactions, while the second OrgB is less tightly integrated into the complex and primarily acts to stabilize extended helical regions of the first OrgB. This asymmetry in OrgB

dimerization is consistent with the previously determined structure of the C-terminus of a homologous protein from flagella (Imada, Minamino et al., 2016).

Several peaks in the SpaO/SpaOc/OrgB/InvC mass spectrum were massively overlapping, making peak assignment troublesome. Nonetheless, CID MS/MS product ions of overlapping peak series were undoubtedly identified and correlated with the respective precursor peaks in the spectrum (**Figure 39**).

The largest assembly identified in substantial amounts was a 2(SpaO-2SpaOc)-2OrgB-InvC complex. Dependent on the electrospray conditions, occasionally signals in the higher  $m/z$ -range were observed (**Supplementary Figure S17**). Due to the low resolution and signal intensity, the charge states could not be identified unambiguously in MS or MS/MS measurements. The mass range and peak interval suggests the presence of complexes with masses of approximately 433 kDa, possibly dimers of 2(SpaO-2SpaOc)-2OrgB-InvC.

With regard to the overall structure of the sorting platform as reported by cryo-ET, the obtained soluble complexes seem to comprise one pod, one spoke and one InvC subunit. Apart from the inefficient putative dimerization, no further oligomerization of the building blocks was observed. The addition of a non-hydrolysable ATP analog (ATP- $\gamma$ -S) to the sample did not trigger oligomerization. Possibly, further oligomerization of the building blocks is only achieved by adding additional protein subunits to the system, such as the proteins of the export apparatus of the inner membrane InvA or SpaS, that were suggested to cause hexamerization of the homologous ATPase in the flagellar system (Ibuki, Imada et al., 2011).



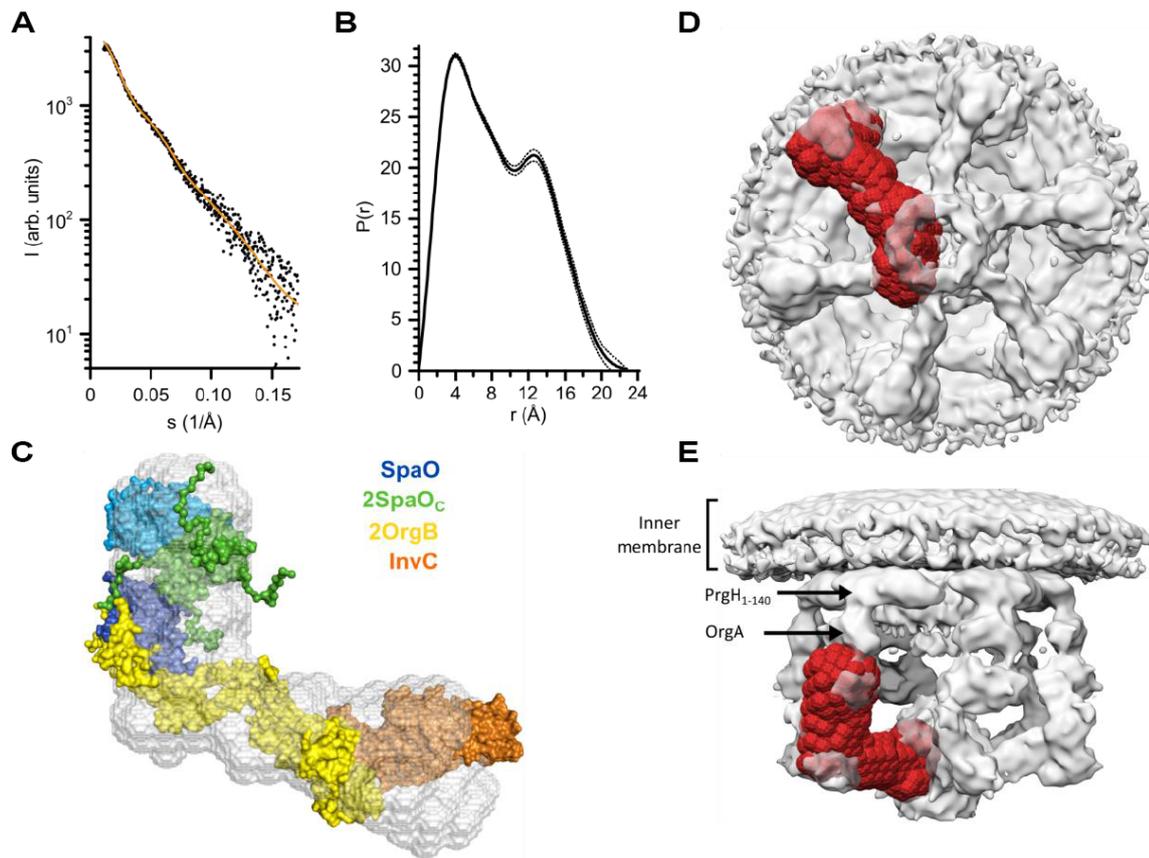
**Figure 39: Identification of SpaO/SpaOc/OrgB/InvC complex stoichiometries by CID MS/MS. A)** A precursor region (encompassing 6,350  $m/z$ ) with overlapping peaks was selected for CID MS/MS analysis with 120 V collision energy. OrgB monomers were found to dissociate from the precursor ions and the residual complexes were identified as SpaO-2SpaOc-OrgB-InvC and 2(SpaO-2SpaOc)-OrgB. From these product signals, precursor ions were determined as +25 SpaO-2SpaOc-2OrgB-InvC ions and +26 2(SpaO-2SpaOc)-2OrgB ions. **B)** Likewise, another region with overlapping signals (encompassing 7,456  $m/z$ ) was selected for CID MS/MS analysis with 150 V collision energy. Product ions were identified as OrgB (white arrows), 2(SpaO-2SpaOc)-OrgB-InvC (blue arrows) and 2SpaO-2SpaOc-OrgB-InvC (grey arrows) and allowed the determination of the precursor ions as +29 2(SpaO-2SpaOc)-2OrgB-InvC and +26 2SpaO-2SpaOc-2OrgB-InvC. Adapted from Bernal, Börnicke et al. (*in preparation*).

### 5.2.5 Structural Modelling of Building Blocks against SEC-SAXS Data

In order to obtain additional structural information on the sorting platform building blocks described by native MS, samples containing SpaO, SpaOc, OrgB and InvC were subjected to size exclusion chromatography (SEC) coupled to small-angle-X-ray scattering (SAXS).

In contrast to native MS analyses, online SAXS measurements produce an averaged signal over all complex species in solution. In agreement with the native MS data (Figure 38), SEC-MALS measurements revealed complex sizes of approximately 210 to 165 kDa, which fits well to the identified complex stoichiometries of 2(SpaO-2SpaOc)-2OrgB-InvC, 2SpaO-2SpaOc-2OrgB-InvC, 2(SpaO-2SpaOc)-2OrgB and SpaO-2SpaOc-2Org-InvC in one SEC elution peak (Supplementary Figure S18). Although only the first half of the SEC elution peak was considered for SAXS analysis, the analysed sample was not entirely homogeneous. Moreover, fast kinetics of disassembly and assembly reactions could induce further heterogeneity after separation by size exclusion chromatography. Therefore, some degree of averaging over different complex species for the obtained SAXS data must be taken into account.

The *ab initio* DAMMIF bead model, which was generated from the SAXS data revealed an extended L-shape of the complex (Figure 40A-C, Supplementary Table S11) (Franke & Svergun, 2009). A SAXS-based molecular model of the SpaO/SpaOc/OrgB/InvC complex was generated under consideration of previously published structural data and results from native MS measurements. The model includes the binding of SpaOc dimers to the N-terminal SpaO domain (SpaO<sub>1-145</sub>), the interaction of SpaO SPOA1-SPOA2 dimer with the N-terminus of OrgB (Notti, Bhattacharya et al., 2015), and an interaction of the C-terminal OrgB domain with the N-terminal InvC domain, which was modelled on the interface of structures of flagellar homologs FliH/FliI (Imada, Minamino et al., 2016). In the final model, 2OrgB-InvC form the longer leg of the L-shape, while the shorter leg is made up of a SpaO-2SpaOc heterotrimer. Consequently, the generated model has a SpaO-2SpaOc-2OrgB-InvC stoichiometry. Although the molecular weight of the complex that was estimated from the SAXS data hints at the presence of another SpaO-2SpaOc heterotrimer, additional proteins could not be fit in the SAXS envelope.



**Figure 40: SEC-SAXS analysis and computational modelling of SpaO/SpaOc/OrgB/InvC complexes.** **A)** A small-angle X-ray scattering profile was recorded SpaO/SpaOc/OrgB/InvC samples. SEC-SAXS experimental data is denoted by black dots and the fit computed from the SAXS rigid-body model from panel C is given as a solid orange line ( $\chi^2 = 1.12$ ). **B)** A pair-distance distribution function  $P(r)$  was computed from the SAXS data in panel A. **C)** A rigid-body SAXS model of SpaO-2SpaOc-2OrgB-InvC was generated. The SpaO-2SpaOc heterotrimer was modelled using a fragment-based model for SpaO<sub>1-145</sub> and the 2SpaOc homodimer structure (PDB ID: 4YX7). Fragment-based models were used for OrgB and InvC. The model was aligned with a representative *ab initio* DAMMIF bead model (grey). **D)** A superposition of the SAXS-based bead model (red) with the *in situ* cryo-ET structure of the *S. Typhimurium* sorting platform (EMDB ID: EMD-8544, grey) was generated in bottom view. **E)** The same superposition is presented in side view with the N-terminal domain of PrgH (PrgH<sub>1-140</sub>) and OrgA being labelled according to Hu, Lara-Tejero et al. (2017). Adapted from Bernal, Börnicke et al. (*in preparation*).

The SAXS L-shape of the complex is similar to the pod structures described by *in situ* cryo-ET analysis (Hu, Lara-Tejero et al., 2017). Moreover, the complex stoichiometries determined in native MS measurements are in good agreement with the structural characterization of the pods with tagged protein variants and mutant proteins (Hu, Lara-Tejero et al., 2017). Hence, the SpaO-2SpaOc-2OrgB-InvC complex characterized by SAXS conceivably represents a soluble building block and six of them further oligomerize to a fully functional T3SS sorting platform. The high correlation of the determined SAXS-based bead model and the previously

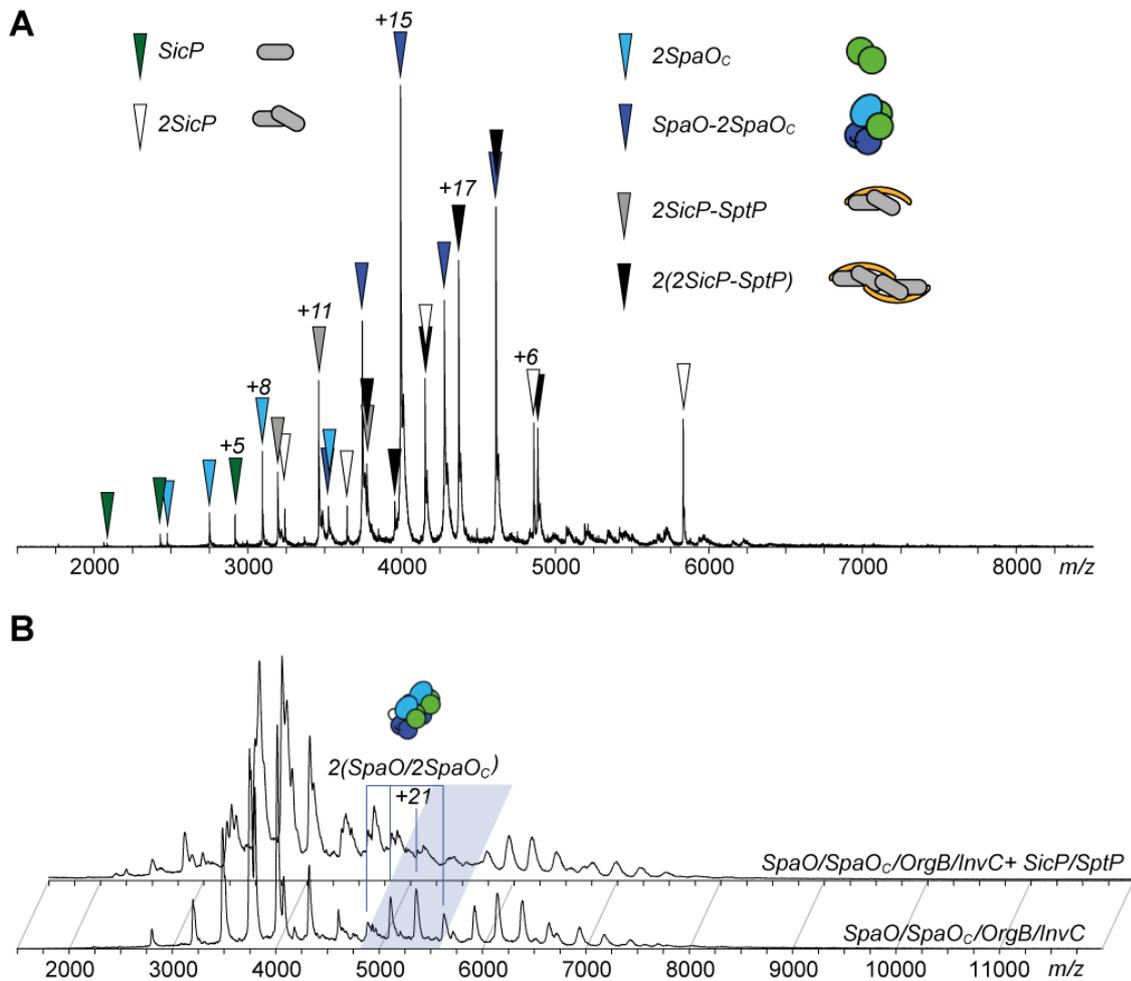
published cryo-ET map supports the hypothesis (**Figure 40D, E**) (Hu, Lara-Tejero et al., 2017). The superposition further matches the previously reported overall protein topology with InvC forming the central hub and OrgB acting as a linker to the SpaO pods (Hu, Lara-Tejero et al., 2017, Hu, Morado et al., 2015).

Nevertheless, filling the remaining five pods with additional SAXS models to generate a molecular model of the entire sorting platform created steric clashes in the central InvC region. Given the clear indications from native MS and SEC-MALS analysis for complex heterogeneity, imperfect SAXS models are likely. Alternatively, the clashes could argue for conformational changes that are required for the oligomerization of the building blocks.

### 5.2.6 *C-ring Proteins Do Not Interact with Chaperone-Effector Complexes in Vitro*

When Lara-Tejero, Kato et al. (2011) first described the sorting platform for type III secretion, they recognized SpaO proteins as key players for the control of sequential protein secretion in *Salmonella*. Therefore, it was hypothesized that SpaO/OrgA/OrgB complexes establish the interaction with chaperone/effector complexes (Stamm & Goldberg, 2011). Here, this hypothesis was verified by testing protein-protein interactions between the chaperone/effector complex SicP/SptP and T3SS sorting platform components by means of native MS.

First, a SicP/SptP sample was mixed with SpaO/SpaOc. The acquired spectra revealed dimerization of SicP chaperones and binding to SptP to form heterotrimers (**Figure 41A**). Furthermore, efficient dimerization of heterotrimers to 2(2SicP-SptP) hexamers was observed, confirming previously published data. For instance, the structure of 2(2SicP/SptP) complexes was solved by X-ray crystallography (Stebbins & Galán, 2001). Although SpaO-2SpaOc heterotrimers were present in large quantities, no interaction between SicP/SptP complexes and SpaO/SpaOc was detected. The obtained mass spectrum showed an uncommonly broad charge envelope for SicP dimers with a very low number of charges that hint at a CID event. However, no corresponding product ion was identified in the low  $m/z$ -range and rather gentle MS settings were used (1.5 kV capillary voltage, 150 V cone voltage, 10 V collision energy).



**Figure 41: SicP/SptP chaperone/effector complexes do not interact with T3SS sorting platform components SpaO/SpaOc/OrgB/InvC.** **A)** SicP/SptP was mixed with SpaO/SpaOc. Spectra revealed the formation of 2SicP/SptP heterotrimers (grey arrows) and further dimerization to 2(SicP/SptP) complexes (black arrows). SpaO-2SpaOc heterotrimers (dark blue arrows) were recognized as expected. Low signal intensities above 5,000  $m/z$  are likely a result of unspecific ESI clustering. **B)** In a preliminary experiment, SpaO/SpaOc/OrgB/InvC samples were measured (bottom) and compared to samples of the same batch that were supplemented with SicP/SptP, 200  $\mu\text{M}$  ATP- $\gamma$ -S and 1 mM  $\text{Mg}^{2+}$  (top). Putative complexes of sorting platform components bound to SicP/SptP were expected in the range above 5,000  $m/z$ . The only pronounced difference between the spectra is the absence of SpaO-2SpaOc dimers in the presence of SicP/SptP (highlighted in blue).

To get a first hint of a possible interaction between SicP/SptP and OrgB- and InvC-containing complexes, SicP/SptP was added to a sample with SpaO/SpaOc/OrgB/InvC complexes (**Figure 41B**). Additionally, ATP- $\gamma$ -S and magnesium ions were added to the sample to account for a possible ATP-dependence. A comparison of spectra of the sorting platform sample in absence and in presence of SicP/SptP revealed only one pronounced difference in the range above 5,000  $m/z$ , where putative complexes were expected. The dimer of SpaO/2SpaOc complexes that was present in absence of SicP/SptP was no longer observed when SicP/SptP was added

to the sample. Binding of SicP/SptP to 2(SpaO-2SpaOc) complexes could certainly result in the non-appearance of the latter complexes in the obtained spectrum. However, in that case the nascent larger protein complexes were expected to appear in the spectrum.

The analysis needs to be repeated to rule out dilution effects and dependencies on the electrospray. Different from the previous experiment, measurements need to be performed with fresh protein samples that had not been frozen for storage and contain higher concentrations of OrgB- and InvC-containing complexes. To avoid protein degradation, a minimal chaperone-binding construct of SptP (amino acids 35-139) was used. Although an interaction of the chaperone with the sorting platform seems likely, using truncated protein constructs involves the risk of deleting amino acids that are essential for the interaction.

### 5.2.7 T3SS Sorting Platform Assembly Model

An assembly model of the T3SS sorting platform was generated by combining the results from native MS analyses of different subcomplexes (Figure 42).

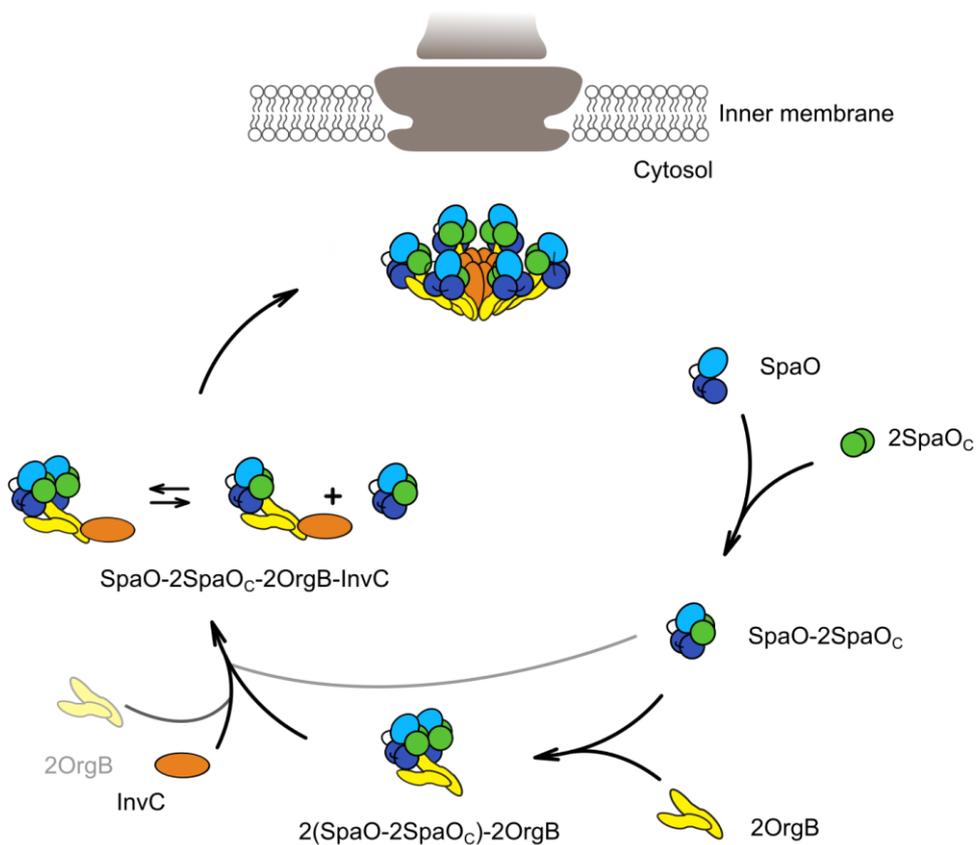
First, SpaOc dimers bind to the N-terminal domains of SpaO to form stable SpaO-2SpaOc heterotrimers that are required for SpaO solubility. Binding of OrgB dimers to SpaO-2SpaOc induces dimerization of SpaO heterotrimers resulting in 2(SpaO-2SpaOc)-2OrgB complexes. Single subunits of the ATPase InvC bind to OrgB in these complexes, thereby giving rise to 2(SpaO-2SpaOc)-2OrgB-InvC complexes. In these building blocks, one OrgB molecule binds to InvC and SpaO-2SpaOc while the other has probably a stabilizing function.

An alternative pathway, with OrgB dimers first binding to InvC before the recruitment of SpaO-2SpaOc heterotrimers cannot be excluded with the current data. The low signal intensity of 2OrgB-InvC signals in native mass spectra and problems with protein solubility of OrgB/InvC samples argue against this pathway in the *in vitro* system. Of course, *in vivo* these arguments may not hold true.

Dynamic dissociation and association of SpaO-2SpaOc heterodimers was observed for SpaO/SpaOc/OrgB/InvC blocks. Moreover, the pod structures of the electron density map only fit one SpaO-2SpaOc heterotrimer as illustrated by the overlay of SAXS model and electron density maps (Figure 40D, E). The contradiction of these observations can be solved by assuming that classification and subtomogram averaging algorithms with template structures in cryo-ET analysis introduce a bias towards pods with one SpaO-2SpaOc heterotrimer only (Yu & Frangakis, 2014). Additional SpaO-2SpaOc heterotrimers could dynamically bind and

dissociate and thus form instable assemblies that are structurally closer to the flagellar C-ring. Functional proof from *in vivo* experiments is required to validate this hypothesis.

Albeit some mass spectra disclosed indications for dimerization of these building blocks, a trigger for an efficient oligomerization to yield fully assembled T3SS sorting platforms *in vitro* was not yet identified.



**Figure 42: T3SS sorting platform assembly model.** SpaO and SpaOc form heterodimers. Binding of OrgB dimers induces oligomerization of SpaO-2SpaOc heterotrimers. The recruitment of InvC monomers gives rise to SpaO-2SpaOc-2OrgB-InvC building blocks that dynamically lose and recruit SpaO-2SpaOc heterotrimers. Alternatively, OrgB dimers bind to InvC before SpaO-2SpaOc heterotrimers are recruited (transparent cartoons and arrows). Higher-level oligomerization of SpaO/SpaOc/OrgB/InvC building blocks to the full hexameric sorting platform might involve interactions with proteins of the injectisome that are located in the inner membrane or conformational changes. Adapted from Bernal, Börnicke et al. (*in preparation*).

## 5.3 Methods

Experimental methods described in this chapter were taken with minor adaptations from Bernal, Börnicke et al. (*in preparation*).

### 5.3.1 Cloning and Mutagenesis of Salmonella Genes

Genes ligated into the expression vectors pASK-IBA (IBA GmbH), pET (Novagen), or pCDFDuet-1 (Novagen) were derived from *Salmonella* Typhimurium strain SL1344 using standard techniques. All PCRs were performed using *Phusion*<sup>®</sup> polymerase (New England Biolabs) and oligonucleotides synthesized by Sigma-Aldrich or Eurofins Genomics. Site-directed mutagenesis of the *spaO* gene was performed according to the *QuikChange* PCR site-directed mutagenesis protocol (Agilent).

Genomic *spaO* deletion in *Salmonella* was carried out by homologous recombination using the  $\lambda$  Red recombinase system (Datsenko & Wanner, 2000). Briefly, the  $\lambda$  Red recombinase plasmid pKD46 was expressed in *S. Typhimurium* SL1344 and a kanamycin cassette flanked by two 50 bp regions homologous to the *spaO* gene was subsequently transformed into the strain for homologous recombination. The  $\Delta spaO_C$ ,  $\Delta spaO_{FL}$ , *spaO*-3xFLAG,  $\Delta spaO_C$ -3xFLAG and  $\Delta spaO_{FL}$ -3xFLAG strains were generated following a similar protocol, introducing a tetracycline cassette into the *spaO* region as described above. In a second step, the tetracycline cassette was replaced by *spaO* DNA carrying mutations and colonies were selected on tetracycline-sensitivity selection media (Bochner, Huang et al., 1980, Maloy & Nunn, 1981). To generate the  $\Delta spaO_C$  strain, silent mutations at the internal putative Shine-Dalgarno region (position 593 to 600) and start codon (position 606) of *spaO* were introduced, while the  $\Delta spaO_{FL}$  strain was produced by introduction nonsense mutations shortly after the start codon of *spaO* at amino acid position 28 and 29. For the generation of the strains *spaO*-3xFLAG,  $\Delta spaO_C$ -3xFLAG and  $\Delta spaO_{FL}$ -3xFLAG, a 3xFLAG-tag was inserted at the C-terminus of *spaO* in the chromosome. PCR and DNA sequencing verified introduction of mutations.

### 5.3.2 Detection of SpaO and SpaO<sub>C</sub> and Protein Secretion in Salmonella Cells

*Salmonella spaO*-3xFLAG,  $\Delta spaO_C$ -3xFLAG and  $\Delta spaO_{FL}$ -3xFLAG strains were grown in LB medium at 37 °C to an OD<sub>600</sub> of 1. Total cell lysates were separated by SDS-PAGE and analysed by western blot using anti-FLAG M2 primary antibody (Sigma-Aldrich), horseradish peroxidase (HRP)-conjugated secondary antibodies (Jackson Immunoresearch) and ECL western blotting substrates (Thermo Fischer Scientific) for protein detection.

To measure protein secretion, strains were grown in LB medium at 37 °C for 6 h to induce SPI-1 effector protein secretion. Where appropriate, expression was induced with anhydrotetracycline (AHT, Sigma-Aldrich) at an OD<sub>600</sub> of 0.1. Proteins were precipitated from 12-13 ml of filtered culture supernatants by addition of 15 % ice-cold trichloroacetic acid (TCA) and centrifugation at 3,200 xg for 90 min. Pellets were washed with ice-cold acetone, air-dried and resuspended in 200 mM Tris-HCl (pH 8.0) containing 200 mM NaCl. Samples were loaded onto SDS-PAGE gels and analysed by Coomassie-staining and western blot. Rabbit anti-SipB, anti-SipC, anti-SipD and anti-SopB polyclonal antibodies were raised and applied for detection of T3SS-dependent substrates in western blot analysis. Anti-FliC (kindly provided by Marc Erhardt's lab) and anti-DnaK antibodies (Stress Gen) were used as loading control and lysis control, respectively. HRP-conjugated secondary antibodies and ECL western blotting substrates were used for protein detection.

### 5.3.3 Protein Synthesis and Purification

*E. coli* BL21 (DE3) cells were grown in LB medium with the appropriate antibiotics at 37 °C. At an OD<sub>600</sub> of 0.5, the temperature was reduced to 20 °C and gene expression was induced by addition of 200 µg/l AHT for pASK-IBA vectors and/or 0.5 mM IPTG for pET and pCDFDuet plasmids. Cells were grown further for 18 h and harvested by centrifugation. All purification steps were performed at 4 °C. To purify SpaO<sub>C</sub>, SpaO<sub>1-145</sub>, SpaO<sub>140-297</sub>, SpaO<sub>1-145</sub>/SpaO<sub>C</sub>, SpaO<sub>1-219</sub>/SpaO<sub>C</sub>, SpaO/SpaO<sub>C</sub>, and SpaO/SpaO<sub>C</sub>/OrgB/InvC complexes, cell pellets were resuspended in buffer B1 (100 mM Tris pH 7.5, 150 mM NaCl) supplemented with complete EDTA-free protease inhibitor cocktail (Roche), 1 mg/ml lysozyme, 10 µg/ml DNase I and 2 mM β-mercaptoethanol. Cell lysis was achieved with a French press and lysates were clarified by centrifugation at 48,000 xg for 30 min. Protein complexes were purified by *Strep*-Tactin affinity chromatography and eluted with buffer B1 supplemented with 7.5 mM desthiobiotin. Affinity-purified proteins were polished by size-exclusion chromatography (SEC) on Superdex 75 or Superdex 200 columns (GE Healthcare), equilibrated with buffer B2 (20 mM HEPES pH 7.5, 350 mM NaCl). Affinity-purified SpaO/SpaO<sub>C</sub>/OrgB/InvC complexes were dialyzed against 10 mM Tris-HCl pH 8.0, 50 mM NaCl and further purified by SEC on a Superose 6 column equilibrated with the same buffer. For SpaO/SpaO<sub>C</sub>/OrgB complex purification, cells were resuspended in buffer B3 (20 mM sodium phosphate buffer pH 7.4, 500 mM NaCl) supplemented with 40 mM imidazole, protease inhibitors, 1 mg/ml lysozyme, 10 µg/ml DNase I and 2 mM β-mercaptoethanol. The SpaO/SpaO<sub>C</sub>/OrgB complex was immobilized on HisTrap HP columns (GE Healthcare), washed with buffer B3 containing 3 mM ATP and 10 mM MgCl<sub>2</sub> and eluted with buffer B3 containing 400 mM imidazole. Eluted proteins were diluted three-fold in buffer B1, purified by *Strep*-Tactin affinity chromatography, followed by

SEC in a Superdex 200 column equilibrated with buffer B2. For solubility analysis of sorting platform proteins, cells were lysed by sonication (Sonopuls HD 2070, Bandelin). Soluble and insoluble fractions were separated by centrifugation and analysed by SDS-PAGE and western blot using anti-*Strep* (Qiagen) and anti-His (GE Healthcare) primary antibodies, HRP-conjugated secondary antibodies and SuperSignal West Dura Extended Duration Substrate for detection. For the purification of OrgB fragments with SpaO/SpaOc and InvC, cells were resuspended in buffer B1 supplemented with complete EDTA-free protease inhibitor cocktail (Roche), 1 mg/ml lysozyme, 10 µg/ml DNase I, 2 mM β-mercaptoethanol and 1 mM MgCl<sub>2</sub>. Cells were lysed by sonication, lysates clarified by centrifugation and protein from the soluble fraction purified by *Strep*-Tactin affinity purification. Eluted proteins were analysed by SDS-PAGE, followed by Coomassie-staining or western blot analysis using an anti-His primary antibody (Thermo Fisher Scientific), HRP-conjugated secondary antibody (Jackson ImmunoResearch) and ClarityMax ECL substrate (Bio-Rad).

Amino acid sequences of all protein constructs used for MS analysis are listed in **Supplementary Table S12**.

#### 5.3.4 *Native Mass Spectrometry*

Protein samples, purified by affinity chromatography and size exclusion chromatography were buffer-exchanged into 50 mM ammonium acetate pH 7.5 (SpaO constructs) or 300 mM ammonium acetate pH 7.5 (SpaOc/SpaO/OrgB) using Vivaspin 500 centrifugal filter units (Sartorius). SpaO/SpaOc/OrgB/InvC samples were exchanged to 60 mM ammonium acetate pH 8.0 during size exclusion chromatography and diluted with 300 mM ammonium acetate pH 8.0 to reach a final concentration of 180 mM. MS/MS measurements for SpaO/SpaOc/OrgB complexes were also performed with this sample. For interaction studies of SicP/SptP with SpaO/SpaOc/OrgB/InvC proteins were mixed after the change to ammonium acetate and supplemented with 2 mM ATP-γ-S (Jena Bioscience) and 1 mM magnesium acetate (Sigma-Aldrich).

Samples were loaded into homemade gold-coated nano ESI glass capillaries, which were produced as described before (see 3.3.2 Native Mass Spectrometry, page 34).

Sample-loaded capillaries were mounted into the nano electrospray ionization (ESI) source of a QToF 2 mass spectrometer (Waters, MS Vision) adapted for high-mass experiments (van den Heuvel, van Duijn et al., 2006) and operated in positive ion mode. Capillary voltages of 1.3 to 1.5 kV and cone voltages of 110 to 150 V were applied. The source pressure was set in the range of 6 to 10 mbar and argon was used as collision gas at 1.7 to 1.9 × 10<sup>-2</sup> mbar. Collision voltages

were optimized for resolution and minimal complex dissociation. CID MS/MS experiments on protein complexes were performed to confirm mass assignments and deduce topological information by selecting specific precursor peaks for dissociation and ramping collision energies up to 400 V or until the entire precursor signal disappeared.

Cesium iodide spectra (25 mg/ml) were acquired on the same day of each measurement and used to calibrate raw data using MassLynx software (Waters). Peak series were assigned with MassLynx and *Massign* (Morgner & Robinson, 2012). Average measured masses of protein complexes, standard deviations of replicate measurements and average full width at half maximum (FWHM) values as a measure of the mass heterogeneity and resolution are listed in **Supplementary Table S10**.

### 5.3.5 Multi-angle Light Scattering

The sample eluting from the SEC column was directed into a triple detector array of UV absorption, multi-angle light scattering (MALS, Wyatt MiniDawn Treos), and RI detectors (Wyatt Optilab T-rEX). The molecular masses of the separated sample components eluting from the column were estimated by combining the results from light scattering with RI and UV absorption measurements.

### 5.3.6 Small-angle X-ray Scattering

Small-angle X-ray scattering (SAXS) measurements were carried out at the beamline P12 (EMBL/DESY, Hamburg) at the PETRA III storage ring using a Pilatus 2M detector (Dectris) (Blanchet, Spilotros et al., 2015). The SAXS camera was set to a sample-detector distance of 3.1 m, covering the momentum transfer range  $0.008 \text{ \AA}^{-1} < s < 0.47 \text{ \AA}^{-1}$ ,  $s = 4\pi \sin(\theta)/\lambda$  (where  $2\theta$  is the scattering angle and  $\lambda = 1.24 \text{ \AA}$  is the X-ray wavelength). For the SAXS measurement, 75-90  $\mu\text{l}$  of affinity-purified protein sample was loaded onto a Superose 6 10/300 GL (GE Healthcare) column, previously equilibrated with 10 mM Tris-HCl pH 8.0, 50 mM NaCl and eluted at 0.5 ml/min. The scattering profile over the elution peak, collected with an exposure time of 1 s each, was separated into sample and buffer regions, appropriately averaged and the signal from the buffer was subtracted using CHROMIXS (Franke, Petoukhov et al., 2017).

The radius of gyration  $R_g$  and forward scattering intensity  $I(0)$  were determined using Guinier analysis (Guinier, 1939) and an indirect Fourier transformation approach by the program GNOM (Svergun, 1992), the latter also providing maximum particle dimensions  $D_{\text{max}}$ .

### 5.3.7 *Structural Modelling against SAXS Data*

Structural models for SpaO<sub>1-145</sub>, OrgB and InvC were obtained using I-TASSER fragment-based modelling (Yang, Yan et al., 2015). Theoretical scattering profiles of SpaO<sub>1-145</sub> models were calculated using the program CRY SOL (Svergun, Barberato et al., 1995) and used to determine the best model fitting the experimental scattering data. The model of full-length SpaO was constructed using the fragment-based model of SpaO<sub>1-145</sub> and the atomic model of the SpaO SPOA1-SPOA2 domain dimer (PDB ID: 4YX5) and subjected to nanosecond timescale all-atom molecular dynamics (MD) simulation in order to refine the model and study its flexibility. The program NAMD (Phillips, Braun et al., 2005) was employed with CHARMM27 force field for the proteins (MacKerell Jr, Bashford et al., 1998) and the TIP3P solvent model was used for water (Jorgensen, Chandrasekhar et al., 1983). Constant particle number, constant pressure and constant temperature ( $NpT$ ) ensembles were assumed. *Langevin* dynamics was used to maintain constant temperature. Pressure was controlled using a hybrid *Nose-Hoover Langevin* piston method. An in-house computational pipeline for high-throughput MD simulations and the visualization program VMD (Humphrey, Dalke et al., 1996) were used to prepare input files and to analyse the simulation trajectories.

*Ab initio* models were reconstructed from the scattering data using bead modelling program DAMMIF (Franke & Svergun, 2009). Ten independent reconstructions were averaged to generate a representative model with the program DAMAVER (Volkov & Svergun, 2003). The average DAMMIF model was also used to calculate the excluded volume of the particle,  $V_{DAM}$ , from which an independent mass weight estimate can be derived (empirically,  $MM_{DAM} \sim V_{DAM}/2$ ). Resolutions of the *ab initio* models were computed using a Fourier Shell Correlation based approach by Tuukkanen, Kleywegt et al. (2016).

The program CORAL was employed for SAXS-based rigid-body modelling and building missing structural fragments (Petoukhov, Franke et al., 2012). The rigid-body model of the SpaO<sub>1-145</sub>-2SpaOc complex was based on the crystal structure of the SpaOc homodimer (PDB ID: 4YX1, SPOA2-SPOA2) and the fragment-based model of SpaO<sub>1-145</sub>. Similarly, the rigid-body SAXS model of the SpaO-2SpaOc complex used the crystal structures of the SpaOc homodimer (PDB ID: 4YX1), the MD-optimized full-length SpaO model constructed from the SPOA1-SPOA2 domain dimer (PDB ID: 4YX5) and the fragment-based SpaO<sub>1-145</sub> model. The SpaO-2SpaOc-2OrgB-InvC complex was built from the crystal structures of the SpaOc homodimer (PDB ID: 4YX1), the SPOA1-SPOA2 domain dimer in complex with the N-terminus of OrgB (PDB ID: 4YX7), and fragment-based models for SpaO<sub>1-145</sub>, InvC, and OrgB.

For the comparison between SAXS data and the electron microscopy density map the program Chimera (Pettersen, Goddard et al., 2004) was used to superimpose a bead model based on the *ab initio* SAXS shape with the *Salmonella* T3SS cryo-ET map (EMDB ID: EMD-8544). A contour level of 2.53 was used for the cryo-ET map.

The SAXS data and the generated model is available at the Small-Angle Scattering Biological Data Bank (SASBDB) under the code: SASDCA8 (SpaO/SpaOc/OrgB/InvC).

## 5.4 Outlook

Structure and function of the T3SS were subject of intensive studies in recent years for two main reasons: First, its presence in many highly pathogenic bacteria makes it a promising target for the development of new antibiotics. In addition, a profound understanding of the T3SS potentially opens new routes to exploit the system for the highly specific delivery of therapeutic proteins into target cells.

The sorting platform on the cytoplasmic side of the injectisome was demonstrated to have a pivotal role in substrate recognition for delivery (Lara-Tejero, Kato et al., 2011). In this work, a recently reported complex shape was complemented with details on the molecular architecture of the complex that contribute to the understanding of type III secretion systems (Hu, Lara-Tejero et al., 2017). However, the structural characterization is still incomplete and various questions need to be addressed in later experiments.

First of all, the role of SpaOc for type III secretion needs to be investigated further. Preliminary findings that suggested the production of SpaOc by proteolysis in an alternative pathway need to be verified.

For the structural analysis of the complexes, SEC-SAXS measurements need to be repeated and coupled to a MALS analysis. The use of a mobile phase-flow splitter allows to split size exclusion chromatography elution fractions online for simultaneous SAXS and MALS analysis. Thereby, the SAXS data can be recorded with real-time monitoring of the molecular weights. So far, data from the first half of the SEC elution peak was used for SAXS analysis. A more restrictive cut-off would possibly decrease complex heterogeneity and increase the quality of the model. Comparisons of SAXS envelopes generated from early and late time spans of the SEC elution peak possibly allow the mapping of InvC or SpaO-2SpaOc binding sites.

Moreover, SpaO/SpaOc/OrgB/InvC complexes could be subjected to additional structural analyses to generate further constraints that increase the accuracy and reliability of the structural modelling against SAXS data. Correlating CCSs from ion mobility mass spectrometry with SAXS models could further validate the correctness of the models. MS-based methods, such as hydrogen-deuterium exchange MS, crosslinking MS or surface labelling MS could be applied to map protein-protein interaction sites in the complexes. Such studies would be severely aggravated by the existence of SpaO isoforms, complex dynamics and resulting heterogeneities and the fact that several proteins exist in multiple copies in the complexes. Therefore, the identification of peptides from digested samples would not always

allow the clear assignment to a specific protein or complex. The use of isotope labels could partially help to overcome such problems.

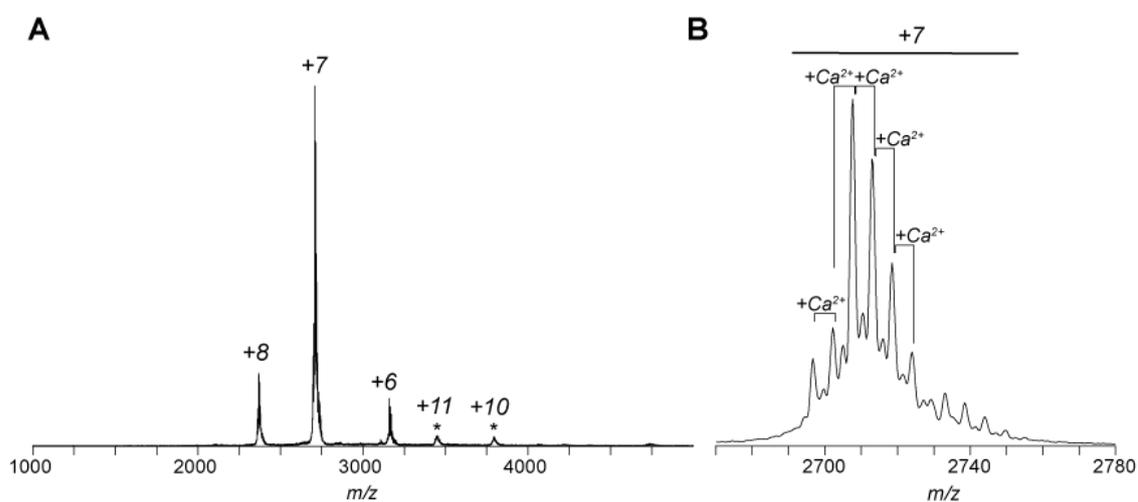
Certainly, a more comprehensive picture of the sorting platform can be created by adding more components to the system. For instance, complex formation in the presence of the proteins OrgA or InvI could reveal further insights. In this work, all efforts to produce soluble OrgA, also in presence of putative interaction partners, were unsuccessful. Maybe the presence of additional proteins from the T3SS is required for a correct OrgA folding and the production of soluble protein.

Another key interaction for type III secretion is the transfer of the substrate from the chaperone dimer to the sorting platform. As a part of this study, different chaperone-effector complexes were analysed, but no interaction with SpaO-2SpaOc heterotrimers or SpaO/SpaOc/OrgB complexes was documented. Further investigations with additional injectisome components are required to shed light on the processes of substrate recognition and transfer.

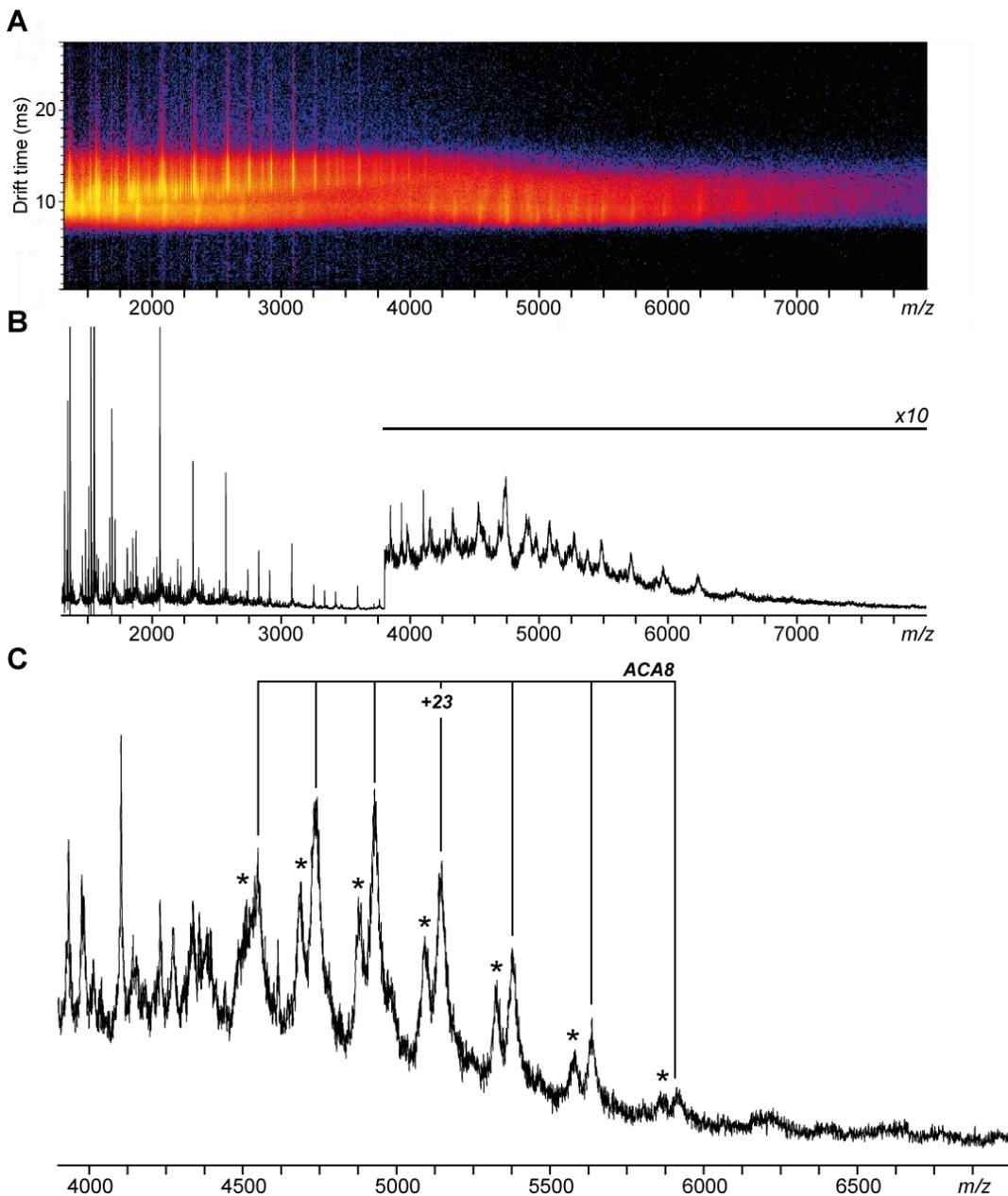
## Supplement

### S1 Supplementary Figures

#### S1.1 Binding of Two Calmodulins to the Plasma membrane $\text{Ca}^{2+}$ -ATPase ACA8



**Figure S1: Native mass spectrum of *A. thaliana* CaM in absence of ACA8 and DDM.** A) The protein was predominantly present in a monomeric state with +6 to +8 charges. Peaks labelled with an asterisk are representing CaM dimers, probably originating from unspecific clustering during the electrospray process. B)  $\text{Ca}^{2+}$ -binding of CaM was analysed from the fine structure of the peaks. Charge state +7 disclosed attachment of up to 5  $\text{Ca}^{2+}$  ions. The characteristic mass shift of +38 Da that results from the replacement of two protons by one  $\text{Ca}^{2+}$  ion is annotated (+ $\text{Ca}^{2+}$ ).



**Figure S2: Preliminary IMMS analysis of ACA8/CaM.** **A)** The drift plot showed massive detergent signals in the range below 4,000  $m/z$ . Protein signals of CaM ( $< 3,250 m/z$ ,  $< 10$  ms) were separated from the bulk detergent signal by their drift time. In the range above 4,000  $m/z$ , ACA8, ACA8-CaM complexes and an unknown 99.7 kDa species were identified. **B)** The mass spectrum showed the different proteins and protein complexes more clearly, a zoom into the 4,000-7,000  $m/z$ -range is presented as **Figure 8**. For better visibility of the protein signals, the spectrum was normalized to the highest signal  $> 2,000 m/z$  and signals in the range  $> 3,800 m/z$  are presented in 10-fold magnification. **C)** ACA8 samples did not show a signal corresponding to a 99.7 kDa species. Degradation of the ACA8, probably due to an extended storage period (4 °C, overnight) was observed, yielding a 117 kDa species. Peaks corresponding to this degradation product are highlighted with asterisks.

## S1.2 Epsin and Sla2 Form Assemblies through Phospholipid Interfaces

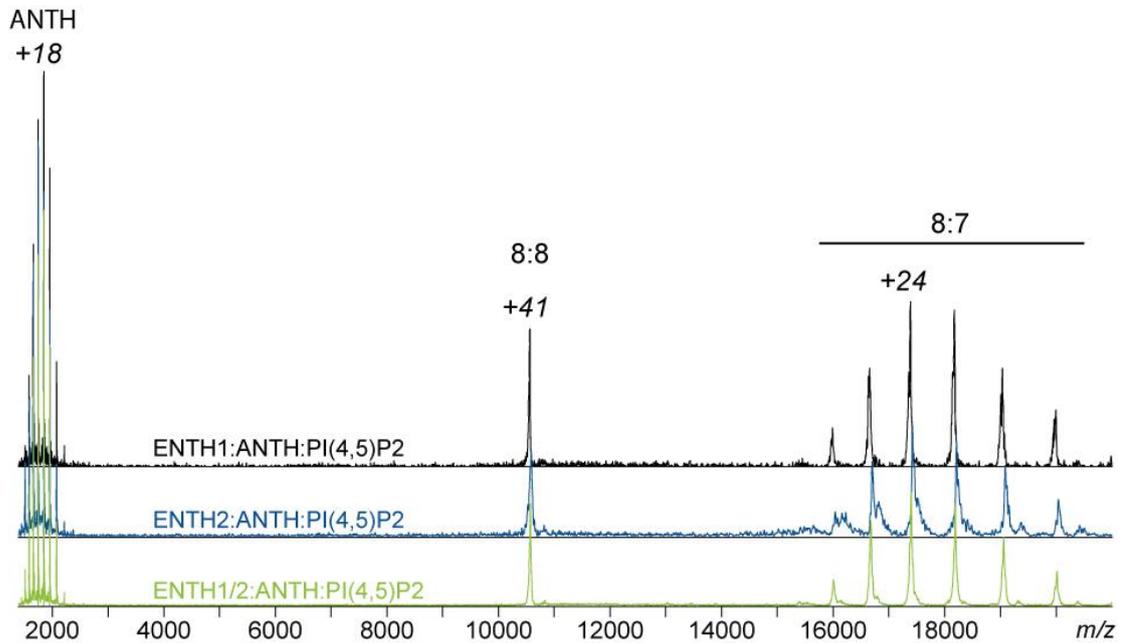
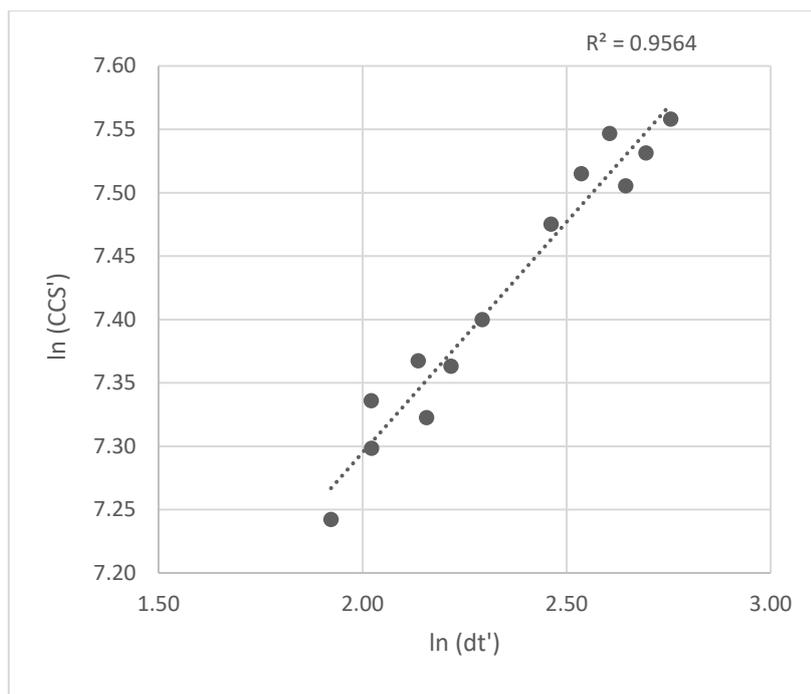
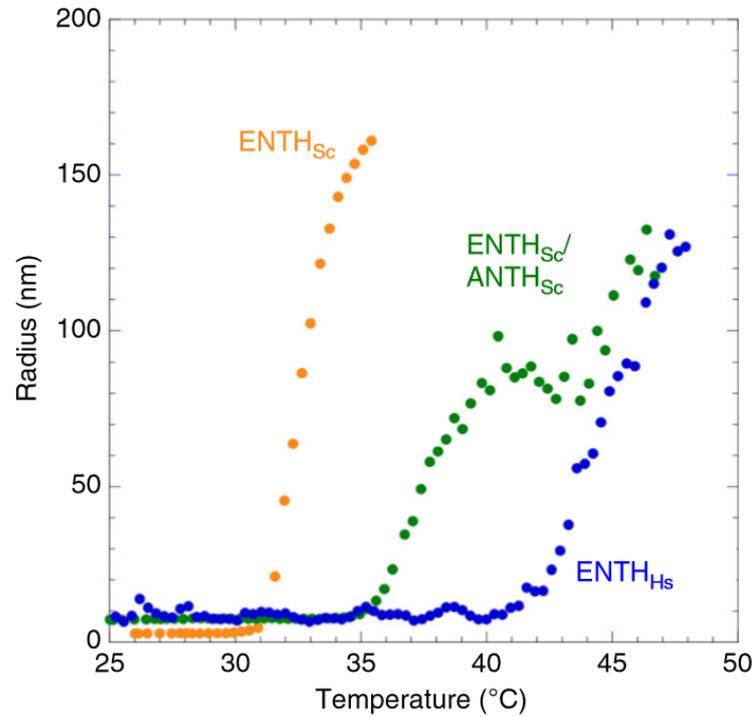


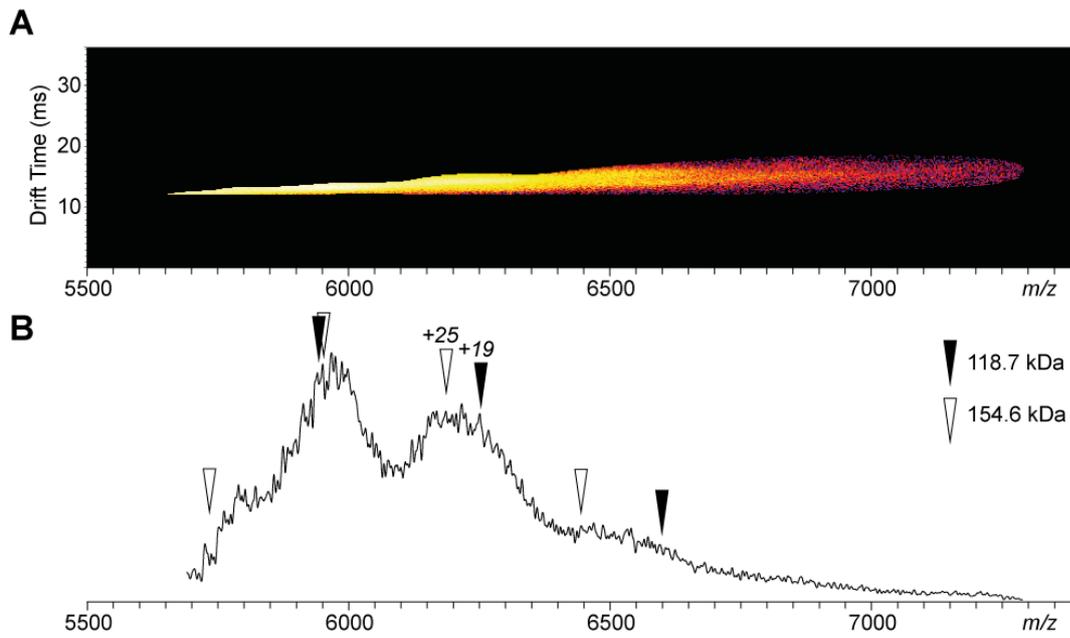
Figure S3: CID MS/MS analysis of *S. cerevisiae* ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes containing ENTH domains from epsin-1 and/or epsin-2. +41 precursor ions with 8:8 (ENTH:ANTH) stoichiometry were selected for dissociation analysis. After dissociation of an ANTH domain, 8:7 complexes with mixed ENTH domains from epsin-1 and epsin-2 (green) revealed an intermediate mass compared to uniform complexes that contained either epsin-1 ENTH (black) or epsin-2 ENTH (blue). For a magnification of charge state +22 of 8:7 complexes, see **Figure 17**.



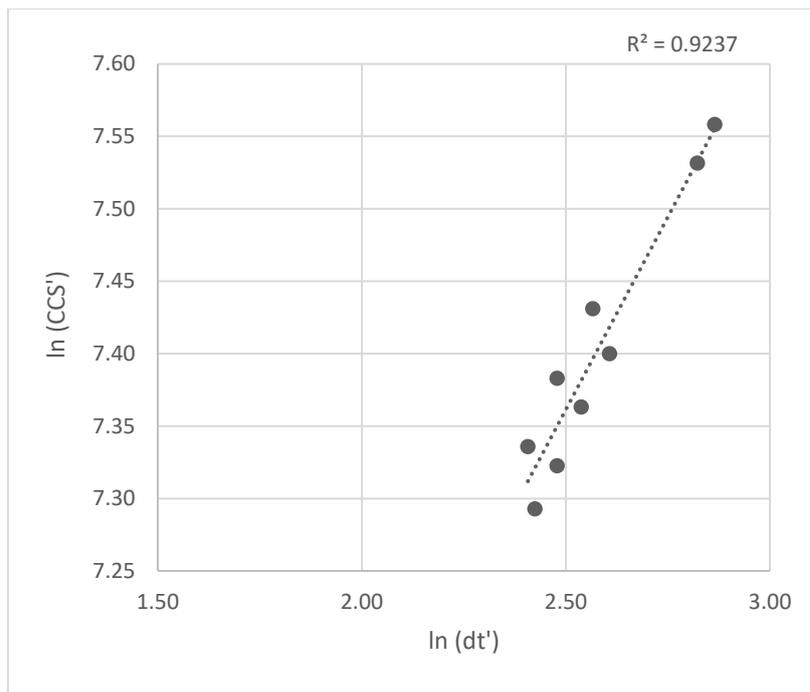
**Figure S4: Logarithmic fit calibration for the CCS' determination of *S. cerevisiae* ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes.** For TWIMS calibration, alcohol dehydrogenase (*S. cerevisiae*, charge states +24 to +27), concanavalin A (*C. ensiformis*, charge state 21), glutamate dehydrogenase (bovine liver, charge states +37 to +39), pyruvate kinase (rabbit heart, charge states +31 to +33), serum albumin (bovine, charge states +15 to +17) were used. From the logarithmic fit ( $\ln(\text{CCS}') = 0.3639 \ln(dt') + \ln 711.2581$ ,  $R^2 = 0.96$ ) the CCS' of ENTH:ANTH:PI(4,5)P<sub>2</sub> complexes was determined.



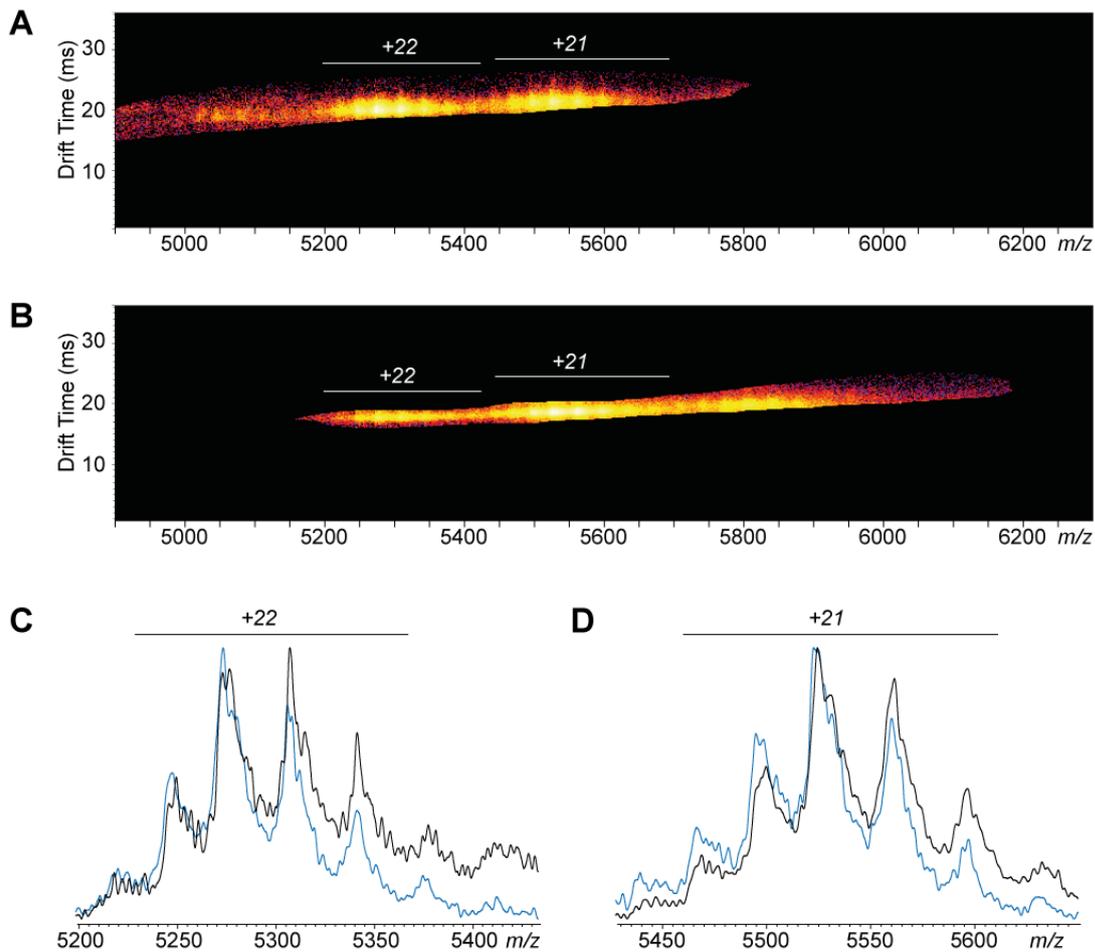
**Figure S5: Stability of clathrin adaptor protein complexes assessed in a dynamic light scattering thermal aggregation assay.** The hydrodynamic radius of the respective protein or protein-PI(4,5)P<sub>2</sub> complex is shown as a function of temperature. Depicted are ENTH1 in presence of PI(4,5)P<sub>2</sub> (orange) and ENTH1:ANTH:PI(4,5)P<sub>2</sub> complexes (green) from *S. cerevisiae* as well as human PI(4,5)P<sub>2</sub>-containing ENTH hexamers (blue). Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.



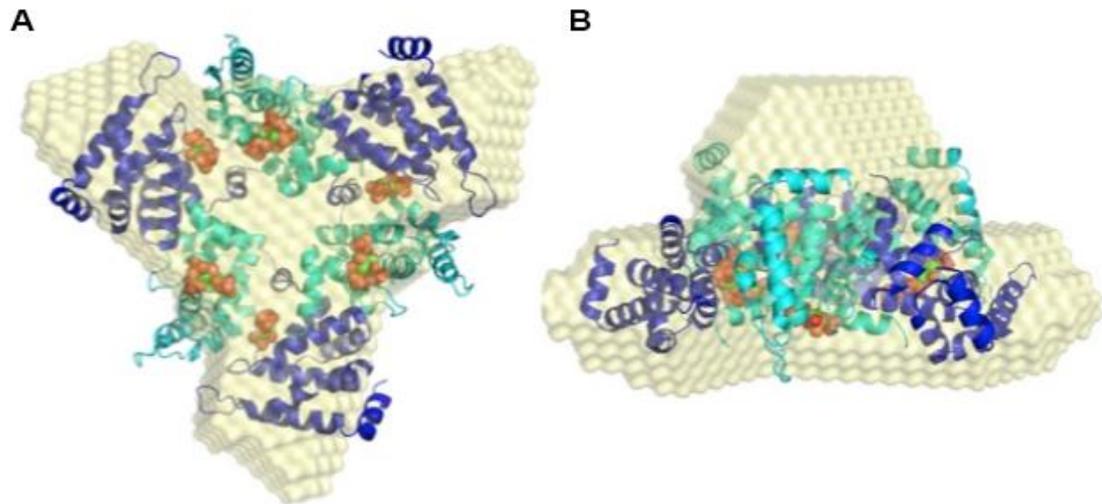
**Figure S6: Characterization of the unidentified human PI(4,5)P<sub>2</sub>-containing ENTH complexes.** **A)** Signals corresponding to the complex species were extracted from the drift plot (**Figure 24A**) to exclude overlapping signals. **B)** Low mass resolution did not allow a confident assignment of charge states. Testing how different charge state assignments match the lowly resolved peaks revealed a probable complex size larger than the ENTH hexamer (a theoretical assignment is labelled with black arrows). For comparison, also a putative assignment for a 154.6 kDa complex (white arrows) is displayed.



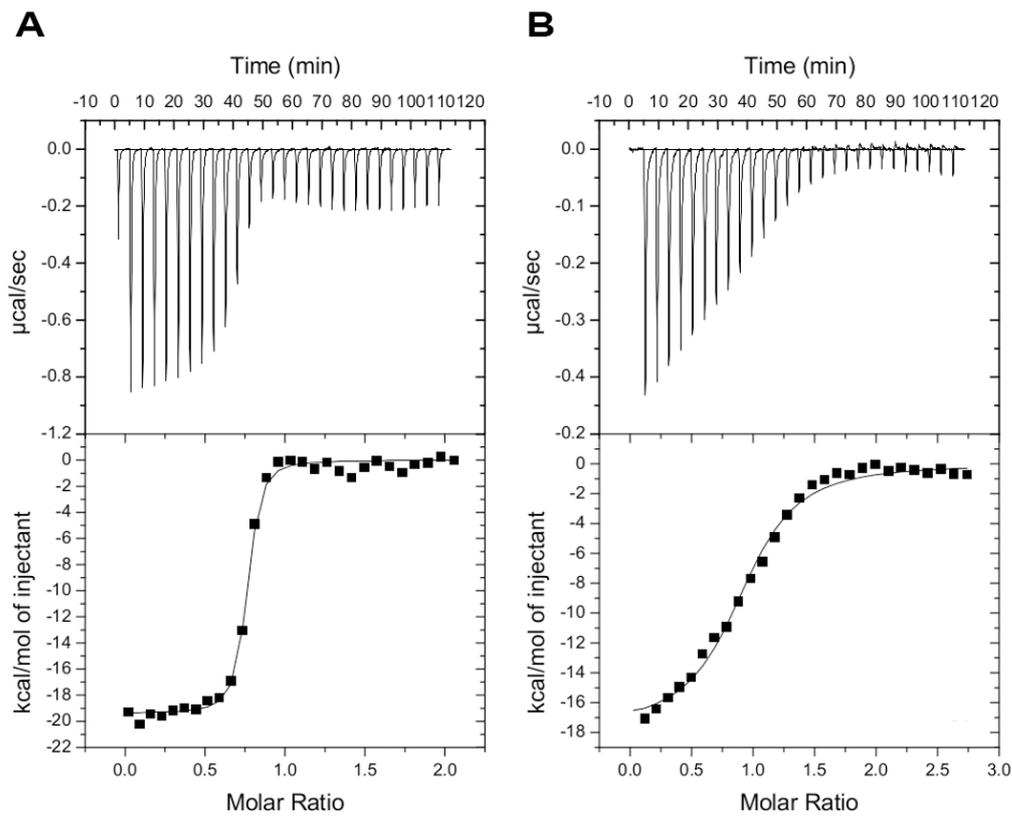
**Figure S7: Logarithmic fit calibration for the CCS' determination of PI(4,5)P<sub>2</sub>-containing *H. sapiens* ENTH hexamers.** For TWIMS calibration, protein complexes alcohol dehydrogenase (*S. cerevisiae*, charge states +24 to +27), concanavalin A (*C. ensiformis*, charge states +19 to +21), glutamate dehydrogenase (bovine liver, charge states +37, +38) were used. From a logarithmic fit ( $\ln(CCS') = 0.5319\ln(dt') + \ln 416.4156$ ,  $R^2 = 0.92$ ) the corrected CCS of PI(4,5)P<sub>2</sub> containing ENTH hexamers were determined.



**Figure S8: The phospholipid content in human PI(4,5)P<sub>2</sub>-containing ENTH hexamers is not determining complex unfolding.** **A)** Highly unfolded complex species were extracted from the drift plot (Figure 24C). **B)** Likewise, less unfolded complex species were extracted from the drift plot. **C)** A comparison of the peak fine structures of charge state +22 revealed similar PI(4,5)P<sub>2</sub> contents for highly unfolded complexes as extracted in A (blue spectrum) and less unfolded complexes as extracted in B (black spectrum). **D)** Likewise, massively unfolded +21 complex ions (blue spectrum) exhibited a PI(4,5)P<sub>2</sub> binding pattern similar to less unfolded complexes (black spectrum).

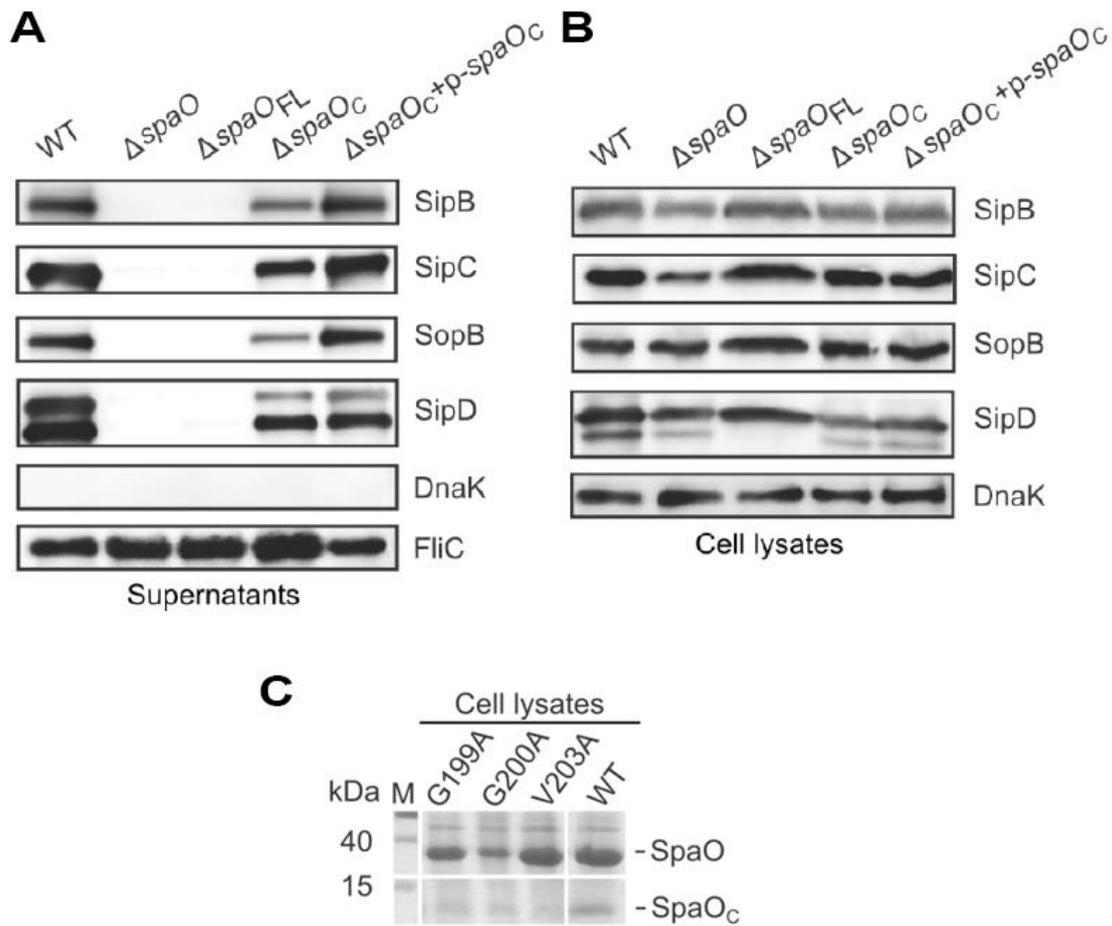


**Figure S9: Human PI(4,5)P<sub>2</sub>-containing ENTH hexamers analysed by small-angle X-ray scattering.** **A)** A superposition of SAXS refined rigid-body models (cartoon) and the reconstructed *ab initio* envelope (yellow semi-transparent surface) of the ENTH hexamer is illustrated. P3 symmetry was enforced and the tandem ENTH domains with bound PI(4,5)P<sub>2</sub> used as rigid bodies are shown in blue and cyan. **B)** Side view of the superposition. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.

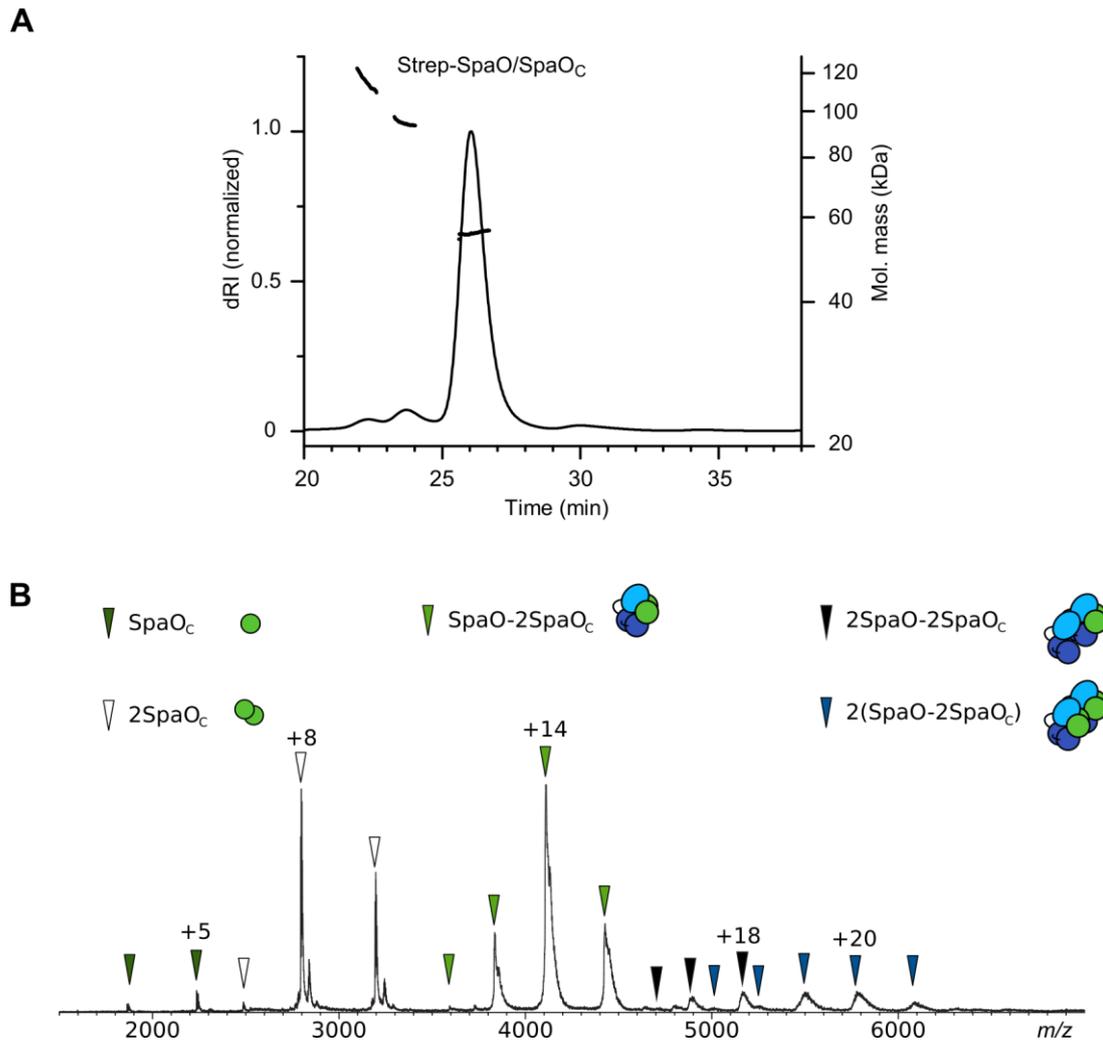


**Figure S10: Isothermal titration calorimetry revealed PI(4,5)P<sub>2</sub>-dependent binding of ANTH and ENTH domains.** **A)** Measured µcal/sec (top) and derived kcal/mol (bottom) of *S. cerevisiae* ANTH that was added to pre-mixed *S. cerevisiae* ENTH2 and PI(4,5)P<sub>2</sub>. **B)** In an analogous experiment, ANTH from *C. thermophilum* was used instead and added to the *S. cerevisiae*/PI(4,5)P<sub>2</sub> mixture. Likewise, a temperature change caused by binding was recognized. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license /Adapted labels.

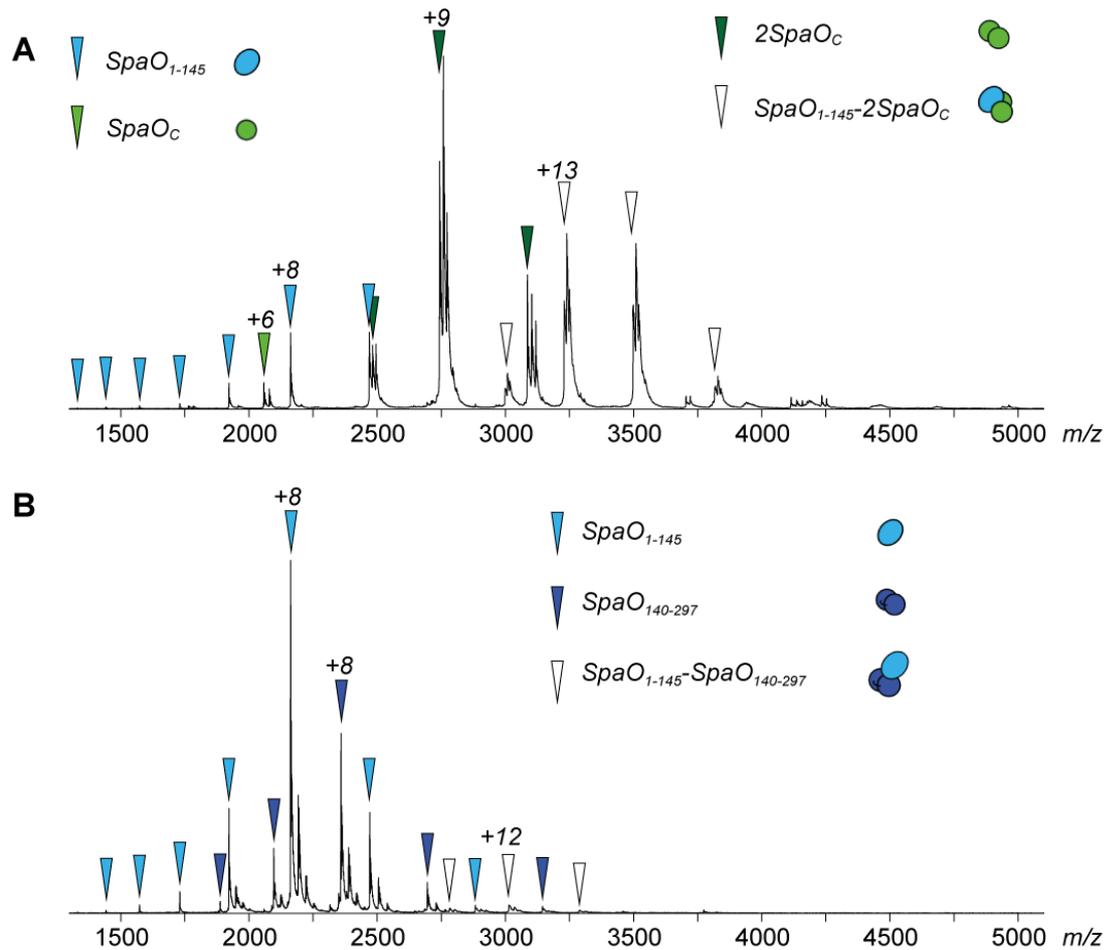




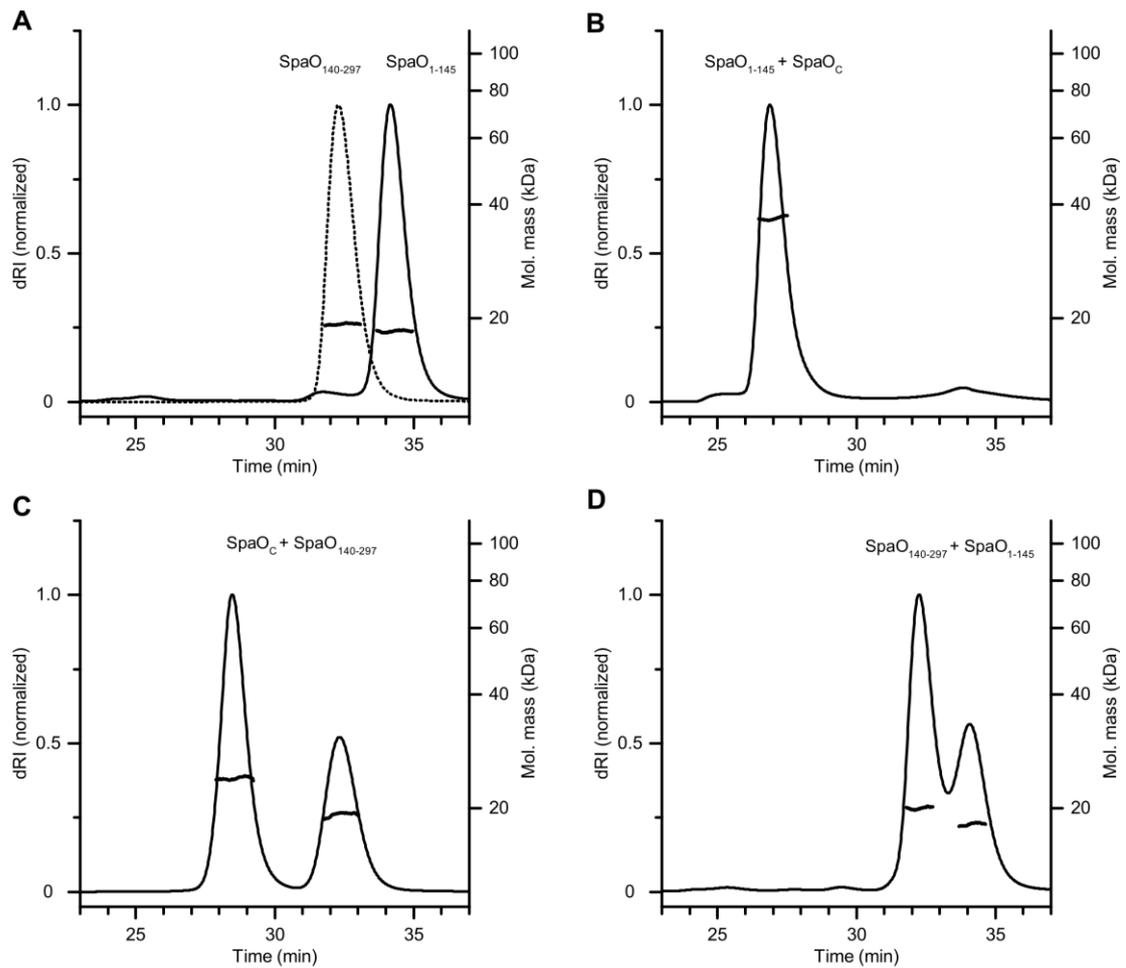
**Figure S12: Fully efficient *Salmonella* Typhimurium type III secretion is dependent on SpaO<sub>c</sub>.** **A)** Proteins secreted by *Salmonella* wild type and mutant strains were detected in culture supernatants by western blot analysis. Deletions of *spaO* and *spaO<sub>FL</sub>* abrogated secretion completely, while deletion of *spaO<sub>c</sub>* only reduced secretion. Complementation with *spaO<sub>c</sub>* partially restored type III secretion ( $\Delta spaO_c + p-spaO_c$ ). **B)** Western blot analysis of T3SS substrates in cell lysates revealed comparable substrate amounts in mutant and wild type strains. **C)** A residual SpaO<sub>c</sub> production was detected by SDS-PAGE of whole cell lysates of plasmid-encoded SpaO (WT) and SpaO mutants expressed in *Salmonella*  $\Delta spaO$ . In G199A and G200A strains the putative internal ribosomal binding site was mutated, while V203A describes a mutation in the putative internal start codon of *spaO*. Reprinted from Bernal, Börnicke et al. (*in preparation*).



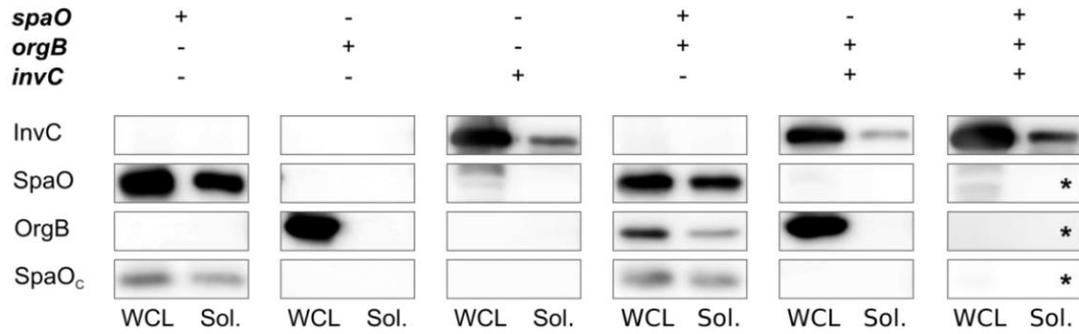
**Figure S13: Characterization of SpaO/SpaOc oligomerization.** **A)** SEC-MALS analysis of Strep-SpaO/SpaOc. The SEC elution profile (dRI trace) and the weight-averaged molar mass across the elution peaks are presented. The detected experimental mass of the main peak is in good agreement with the theoretical mass of 57 kDa. Smaller elution peaks likely correspond to  $2(\text{SpaO}-2\text{SpaOc})$  and  $2\text{SpaO}-2\text{SpaOc}$  complexes with theoretical molecular weights of 115 kDa and 92 kDa, respectively. **B)** The effect of the location of the Strep-tag on SpaO/SpaOc complex formation was assessed by native MS. Here, Strep-SpaO/SpaOc samples were analysed. The formation of SpaOc monomers (dark green arrows), SpaOc dimers (white arrows), SpaO-2SpaOc (light green arrows) and dimers of SpaO-2SpaOc (blue arrows) was identified as observed for samples containing C-terminally Strep-tagged proteins (**Figure 35**). In addition, low intensity signals for  $2\text{SpaO}-2\text{SpaOc}$  complexes were found. Reprinted from Bernal, Börnicke et al. (*in preparation*).



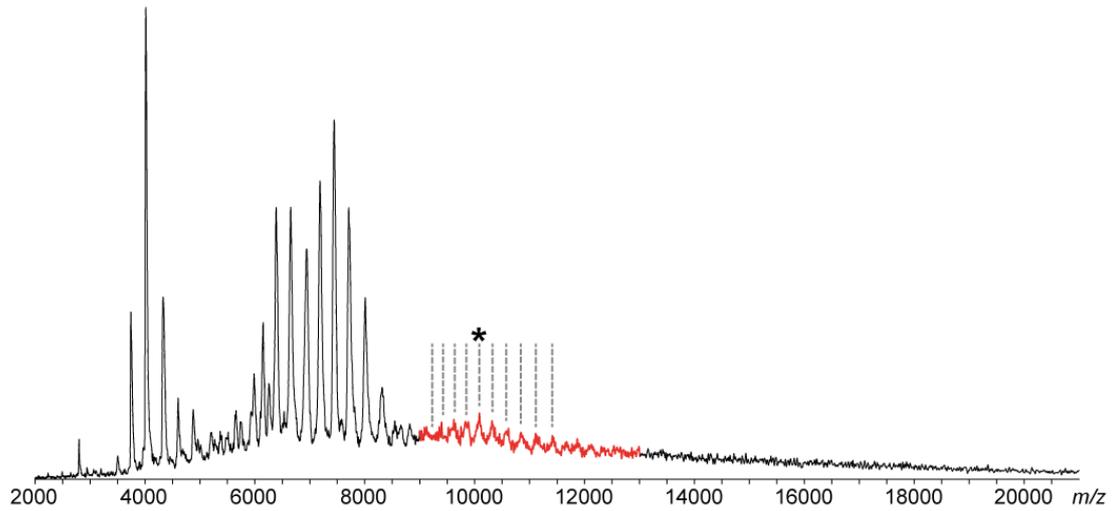
**Figure S14: Native MS analysis of SpaO-SpaOc domain interactions.** **A)** Interacting proteins were detected in samples of mixed SpaO<sub>1-145</sub>/SpaOc. Apart from SpaO<sub>1-145</sub> (blue arrow) monomers and SpaOc monomers (light green arrow) and homodimers (dark green arrow), SpaO<sub>1-145</sub>-2SpaOc heterotrimers (white arrows) were detected. **B)** Mixed SpaO<sub>1-145</sub>/SpaO<sub>140-297</sub> samples revealed no interaction. Monomeric SpaO<sub>1-145</sub> (light blue arrow) and SpaO<sub>140-297</sub> (dark blue arrow) were identified. Weak signals of SpaO<sub>1-145</sub>-SpaO<sub>140-297</sub> complexes likely originated from non-specific clustering during the ESI process. Reprinted from Bernal, Börnicke et al. (*in preparation*).



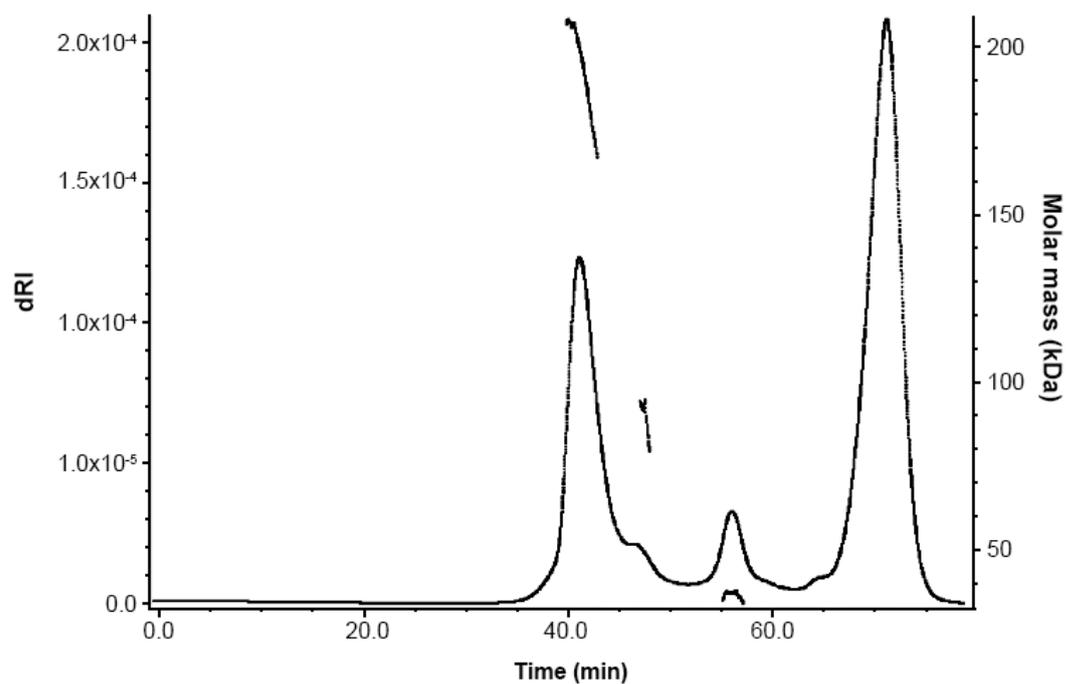
**Figure S15: SEC-MALS analysis of SpaO-SpaOc domain interaction.** **A)** Analysis of SpaO<sub>1-145</sub> (solid line) and SpaO<sub>140-297</sub> (dotted line) in individual measurements. Experimental masses are congruent with theoretical masses of monomeric SpaO<sub>1-145</sub> (17 kDa) and SpaO<sub>140-297</sub> (19 kDa), respectively. SEC elution profiles (dRI traces) and weight-averaged molecular masses across the elution peaks are shown. **B)** Analysis of co-purified SpaO<sub>1-145</sub>/SpaOc revealed a signal corresponding to a higher molecular mass. The experimental mass is suggesting a SpaO<sub>1-145</sub>-2SpaOc (41 kDa theoretical mass) complex formation. **C)** Analysis of mixed SpaO<sub>140-297</sub>/SpaOc samples revealed no interaction. Monomeric SpaO<sub>140-297</sub> (19 kDa theoretical mass) and dimeric SpaOc (24 kDa theoretical mass) were detected. **D)** Analysis of combined SpaO<sub>1-145</sub>/SpaO<sub>140-297</sub> exposed no interaction. Reprinted from Bernal, Börnicke et al. (*in preparation*).



**Figure S16: Solubility of *Salmonella* T3SS sorting platform proteins after recombinant gene (co-) expression in *E. coli*.** Genes were (co-)expressed as indicated at the top and then detected in western blots of whole cell lysates (WCL) and soluble fractions (Sol.). Primary antibodies were targeted against Strep-tags (SpaO, SpaO<sub>c</sub>, InvC) or His-tags (OrgB). Asterisks label experiments with untagged protein constructs that could not be detected in this assay. Results are representative for at least two biological replicates. Reprinted from Bernal, Börnicke et al. (*in preparation*).



**Figure S17: SpaO/SpaOc/OrgB/InvC spectra revealed formation of higher-order oligomers.** Dependent on the electrospray conditions, occasionally a signal in the range of 9,000 to 13,000  $m/z$  was observed (red). Resolution was not sufficient for an unequivocal assignment of the charge states. However,  $m/z$ -range and peak distances suggest a complex mass of approximately 430 kDa. If the main peak, highlighted with an asterisk, was charge state +43, the dotted lines indicate the expected peak positions for a 433 kDa complex. Reprinted from Bernal, Börnicke et al. (*in preparation*).



**Figure S18: SEC-MALS analysis of SpaO/SpaOc/OrgB/InvC complexes.** The SEC peak with a retention time of 39 minutes revealed a molecular mass decreasing from approximately 210 kDa at the beginning of the peak to 165 kDa at the end of the peak (red), indicating a mixture of different complexes.

---

## S2 Supplementary Tables

### S2.1 Binding of Two Calmodulins to the Plasma-membrane Ca<sup>2+</sup>-ATPase ACA8

Table S1: Theoretical and experimental molecular weights (M) of *A. thaliana* ACA8, CaM and complexes of these proteins. Experimental masses (M<sub>Exp</sub>) are listed with standard deviations and average full width half maximum (FWHM), all values in Da. Adapted from Nitsche, Josts et al. (*in revision*).

Protein/-complex	M <sub>Theo</sub>	M <sub>Exp</sub>	St. dev. (Da, N≥3)	Avg. FWHM
CaM	18,593.5	18,869.9	0.8	11
ACA8	118,194.2	118,520	60	804
ACA8 + CaM	136,787.7	137,500	110	700
ACA8 + 2 CaM	155,381.2	156,600	200	730

Table S2: Amino acid sequences of *A. thaliana* proteins ACA8 and CaM7.

Protein	Organism	Sequence
ACA8	<i>A. thaliana</i>	MHHHHHHHHENLYFQGATSLKSSPGRRRGGDVESGKSEHADSDSDTF YIPSKNASIERLQQWRKAALVLNASRRFRYTLDLKKEQETREMRQKIR SHAHALLAANRFMDMGRESGVEKTTGPATPAGDFGITPEQLVIMSKDH NSGALEQYGGTQGLANLLKTNPEKGISGDDDDLLKRKTIYGSNTYPRK KKGFLRFLWDACHDLTLIIIMVAVASLALGIKTEGIKEGWYDGGSI AFAVILVIVVTAVSDYKQSLQFQNLNDEKRNIHLEVLRGGRVEISYI DIVVGDVPLNIGNQVPADGVLISGHSALDESSTGESKIVNKDANK DPFLMSGCKVADNGSMLVTG VGVNTEWGLLMASISEDNGEETPLQVR LNGVATFIGSIGLAVAAAVLVILLTRYFTGHTKDNNGGPQFVKGKTKV GHVIDDVVKVLTVAVTIVVVAVPEGLPLAVTTLTAYSMRKMMAKALV RRLSACETMGSATTICSDKTGTLTLNQM TVVESYAGGKKTDEQLPAT ITSLVVEGISQNTTGSIFVPEGGGDLEYS GSPT EKAILGWGVKLG MNF ETARSQSSILHAFPFNSEKKRGGVAVKTADGEVHVHWKGASEIVLASC RSYIDEDGNVAPMTDDKASFFKNGINDMAGRTLRCVALAFRTYEA EKV PTGEELSKWVLPEDDLILLAI VGIKPCRPGVKDSV VLCQ NAGVKV RM VTGDNVQTARAI ALECGILSSDADLSEPTLIEGKSFREMTDAERDKIS DKISVMGRSSPNDKLLLVQSLRRQGHVVAVTGDGTNDAPALHEADIGL AMGIAGTEVAKESDIIILDDNFASVVKVVRWGRSVYANIQKFIQFQL TVNVAALVINVVAAISSGDVPLTAVQLLWVNLIMDTLGALALATEPPT DHLMGRPPVGRKEPLITNIMWRNLLIQAIYQVSVLLTLNFRGISILGL EHEVHEHATRVKNTIIFNAFVLCQAFNEFNARKPDEKNIFKGVIKNRL FMGIIVITLVLQVIIVEFLGKFASTTKLNWKQWLICVIGIVISWPLAL VGKFI PVPAAPISNKLKVLKFWGKKKNSSGEGSL
CaM7	<i>A. thaliana</i>	MHHHHHHENLYFQGAADQLTDDQISEFKEAFSLFDKGDGDCITTKELG TVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLNLMARKMKD TDS EEELKEAFRVFDKQNGFISAAELRHVMTNLGEKLTDEEVDEMIREAD VDGDGQINYE EFVKVMMAK

## S2.2 Epsin and Sla2 Form Assemblies through Phospholipid Interfaces

**Table S3: Molecular weights (M) of single ENTH and ANTH domains and ENTH/ANTH/PI(4,5)P<sub>2</sub> complexes from *S. cerevisiae*, *C. thermophilum* and *H. sapiens* were determined from at least three MS or MS/MS measurements and listed with standard deviations and the average full width half maximum (FWHM), all values in Da. Using the experimental masses of single ENTH and ANTH proteins, an expected protein mass for the complexes was determined, ENTH:ANTH stoichiometries are listed. By subtracting the expected protein mass from the measured mass, the overall lipid mass of the complex was calculated. Division of the lipid mass by the theoretical PI(4,5)P<sub>2</sub> mass (741.25 Da) led to an estimated number of PI(4,5)P<sub>2</sub> molecules. Variations of the PI(4,5)P<sub>2</sub> content were calculated based on the FWHM value. Results of 6:6 and 8:8 complexes from MS measurements showed broader peaks due to remaining buffer molecules, resulting in a larger FWHM value and a slight overestimation of the number of PI(4,5)P<sub>2</sub> molecules. For *H. sapiens* 6:0 and 5:0 ENTH complexes resolution was sufficient to directly determine the exact number of phospholipids. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.**

Protein	Species	$M_{\text{Theo}}$	$M_{\text{Exp}}$	St.dev.	FWHM
ENTH1	<i>S. cerevisiae</i>	18,847.4	18,846.8	0.6	8.8
ENTH2	<i>S. cerevisiae</i>	18,958.4	18,957.6	3.5	10.0
Sla2	<i>S. cerevisiae</i>	31,638.1	31,626.7	9.8	16.7
Sla2 *	<i>S. cerevisiae</i>	31,638.1	31,635.8	0.6	23.3
ENTH	<i>C. thermophilum</i>	18,041.6	18,168.0	5.4	10.9
Sla2	<i>C. thermophilum</i>	30,407.8	30,263.3	0.6	14.8
Sla2 *	<i>C. thermophilum</i>	30,407.8	30,263.4	4.0	27.0
ENTH	<i>H. sapiens</i>	18,311.8	18,571.4	1.0	11.2
ENTH *	<i>H. sapiens</i>	18,311.8	18,570.3	0.4	13.2
Hip1R	<i>H. sapiens</i>	34,996.3	34,997.5	1.9	16.0
Hip1R *	<i>H. sapiens</i>	34,996.3	34,994.4	0.5	23.7
CALM	<i>H. sapiens</i>	32,894.8	32,990.8	16.4	181.5

Complex	Species	N PI(4,5)P <sub>2</sub>	$M_{\text{Exp}}$	St.dev.	FWHM
6:6 ENTH1:ANTH	<i>S. cerevisiae</i>	19±2	317,332	209	1,660
8:8 ENTH1:ANTH	<i>S. cerevisiae</i>	25±3	422,574	298	2,373
8:7 ENTH1:ANTH *	<i>S. cerevisiae</i>	23±1	389,639	134	701
6:6 ENTH2:ANTH	<i>S. cerevisiae</i>	19±2	317,727	356	1,119
8:8 ENTH2:ANTH	<i>S. cerevisiae</i>	24±3	422,790	240	2,385
8:7 ENTH2:ANTH *	<i>S. cerevisiae</i>	23±1	389,897	87	935
6:6 ENTH:ANTH	<i>C. thermophilum</i>	18±3	304,232	345	2,406
8:8 ENTH:ANTH	<i>C. thermophilum</i>	25±4	405,983	247	2,720
8:7 ENTH:ANTH *	<i>C. thermophilum</i>	24±1	374,759	436	504
6:0 ENTH	<i>H. sapiens</i>	6±0	115,971	20	249
5:0 ENTH *	<i>H. sapiens</i>	6±0	97358	12	153

6:6 ENTH:ANTH	<i>H. sapiens</i>	18±7	334,702	290	5,495
6:5 ENTH:ANTH *	<i>H. sapiens</i>	17±4	298,960	81	3,332

---

\*determined in MS/MS experiments

**Table S4: *S. cerevisiae* ENTH2 X-ray crystallographic data collection and refinement statistics.** Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.

	ENTH2/PI(4,5)P <sub>2</sub>	ENTH2
<b>Data collection</b>		
Space group	F432	P2 <sub>1</sub>
Unit cell dimensions		
<i>a, b, c</i> (Å)	<i>a, b, c</i> = 211.4	24.4, 65.7, 33.7
$\alpha, \beta, \gamma$ (°)		B = 97.3
Resolution (Å)	120 – 3.35 (8.85-3.35)*	66 – 1.80 (1.84-1.980)*
<i>R</i> <sub>pim</sub>	0.053 (1.04)	0.020 (0.353)
<i>R</i> <sub>merge</sub>	0.134 (1.62)	0.048 (0.580)
<i>I</i> / $\sigma$ <i>I</i>	9.4 (0.9)	23.6 (2.1)
Completeness (%)	98.6 (97.1)	99.2 (94.6)
Redundancy	6.2 (3.0)	6.4 (3.5)
<i>CC</i> <sub>1/2</sub> *	0.34	0.76
<b>Refinement</b>		
Resolution (Å)	75.0 – 3.35 (4.22 – 3.35)	33.5 – 1.90
No. reflections	5,798	7,571
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.28/0.30	0.17/0.22
No. atoms		
Protein	2,378	1,104
Ligand/ion	36	N/A
Water	N/A	61
<i>B</i> -factors (Å <sup>2</sup> )		
Protein	155	26
Ligand/ion	165	N/A
Water	N/A	35
RMS deviations		
Bond lengths (Å)	0.004	0.018
Bond angles (°)	0.83	1.95
Ramachandran Statistics		
Favoured (%)	91.9	100.0
Disallowed (%)	1.8	0.0

\*Values in parentheses are for the highest-resolution shell.

**Table S5: SAXS Data collection and derived parameters for human ENTH.** Abbreviations:  $M_r$ : molecular mass;  $R_g$ : radius of gyration;  $D_{max}$ : maximal particle dimension;  $V_p$ : Porod volume;  $V_{ex}$ : Particle excluded volume. Reprinted from “Epsin and Sla2 form assemblies through phospholipid interfaces” by Garcia-Alai, Heidemann et al. (2018) from Nature Communications (10.1038/s41467-017-02443-x), used under CC BY 4.0 license.

	ENTH	ENTH + PI(4,5)P <sub>2</sub>
<b>Data collection parameters</b>		
Instrument	EMBL P12 beam line (PETRA-III, DESY, Hamburg)	
Beam geometry	0.2 x 0.12 mm <sup>2</sup>	
Wavelength (Å)	1.24	
$s$ range (Å <sup>-1</sup> ) <sup>a</sup>	0.01-0.46	
Exposure time (s)	1 (20x0.05 s)	
Concentration range (mg/mL)	0.4-1.5	0.4-1.5
Temperature (K)	288	288
<b>Structural parameters<sup>b</sup></b>		
$I(0)$ (cm <sup>-1</sup> ) [from $p(r)$ ]	0.012 ± 0.001	0.078 ± 0.001
$R_g$ (Å) [from $p(r)$ ]	19 ± 1	34 ± 1
$I(0)$ (cm <sup>-1</sup> ) (from Guinier)	0.012 ± 0.001	0.078 ± 0.001
$R_g$ (Å) (from Guinier)	19 ± 1	35 ± 1
$D_{max}$ (Å)	60	112
Porod volume estimate (Å <sup>3</sup> )	34,020 ± 10,000	189,000 ± 20,000
Excluded volume estimate (Å <sup>3</sup> )	17,900 ± 2,000	202,000 ± 20,000
Dry volume calculated from sequence (Å <sup>3</sup> ) <sup>c</sup>	22315/133782 (mon/hex)	
<b>Molecular-mass determination</b>		
$I(0)$ (cm <sup>-1</sup> ) Glucose Isomerase (173,000 Da)	0.118 ± 0.001	
Molecular mass $M_r$ (Da) [from $I(0)$ ]	17,593 ± 2,000	114,356 ± 10,000
Molecular mass $M_r$ (Da) [from Porod volume ( $V_p/1.6$ )]	21,263 ± 5,000	118,256 ± 10,000
Molecular mass $M_r$ (Da) [from excluded volume ( $V_{ex}/2$ )]	8,950 ± 1,000	101,000 ± 10,000
Calculated monomeric $M_r$ from sequence (Da)	~18,443	
<b>Software employed</b>		
Primary data reduction	RADAVER	
Data processing	PRIMUS/Qt	
<i>Ab initio</i> analysis	DAMMIF, DAMMIN	
Validation and averaging	DAMAVR	
Rigid-body modeling	CORAL	
Computation of model intensities	CRY SOL	
3D graphics representations	PyMOL, UCSF Chimera	

<sup>a</sup>Momentum transfer  $s = 4\pi\sin(\theta)/\lambda$ . <sup>b</sup>Values reported for 1.5 mg mL<sup>-1</sup>. Dry volume determined using the server: <http://www.basic.northwestern.edu/biotools/proteincalc.html>

**Table S6: Sequence identity matrix of investigated ENTH domains.** Sequences are numbered and the sequence identity (in percent) to corresponding numbers is given as calculated by ClustalOmega (Sievers, Wilm et al., 2011). Only half of the matrix is filled to avoid redundant entries.

No	Protein	Organism	1	2	3	4
1	Epsin-1	<i>H. sapiens</i>	100	50.32	45.57	42.41
2	Epsin-1	<i>C. thermophilum</i>	-	100	55.13	53.21
3	Epsin-1	<i>S. cerevisiae</i>	-	-	100	73.49
4	Epsin-2	<i>S. cerevisiae</i>	-	-	-	100

**Table S7: Sequence identity matrix of investigated ANTH domains.** Sequences are numbered and the sequence identity (in percent) to corresponding numbers is given as calculated by ClustalOmega (Sievers, Wilm et al., 2011). Only half of the matrix is filled to avoid redundant entries.

No	Protein	Organism	1	2	3	4	5
1	Hip1R	<i>H. sapiens</i>	100	14.83	27.86	26.64	19.76
2	CALM	<i>H. sapiens</i>	-	100	21.52	21.14	26.77
3	Sla2	<i>C. thermophilum</i>	-	-	100	58.75	16.67
4	Sla2	<i>S. cerevisiae</i>	-	-	-	100	22.73
5	Yap1802	<i>S. cerevisiae</i>	-	-	-	-	100

Table S8: Amino acid sequences of ENTH and ANTH domains.

Domain	Protein	Organism	Sequence
ENTH	Epsin-1	<i>S. cerevisiae</i>	GAMGSMKQFVRSKLNLVKGYSSSTQVLVLRN ATSNDNHQVSKDSLIELAEKSYDSADFFFEI MDMLDKRLNDKGYWRHIAKALTVIDYLIR FGSENCVLWCRENLYIIKTLKEFRHEDDEG IDQGQIVRVKAKELTALLSDDERLNEERNM NIKGRNRKGRRR
ENTH	Epsin-2	<i>S. cerevisiae</i>	GAMGSMKQFVRSKNNMMKGYSSSTQVLVLRD ATANDSRTPSIDTLDDLAQRSYDSVDFFFEI MDMLDKRLNDKGYWRHVAKSLTVLDYLVR FGSENCVLWCRENFYVIKTLREFRHENESG FDEGQIIRVKAKELVSLNDEERLREERSM NTRNRRANRAAR
ANTH	Sla2	<i>S. cerevisiae</i>	GAMGSMRIDSIDLQKALKKACSVVEETAPKR KHVRACIVYTWHDQSSKAVFTTLKTLPLAN DEVQLFKMLIVLHKIIQEGHPSALAEAIRD RDWIRSLGRVHSGGSSYSKLIREYVRYLVL KLDFFHAHHRGFNNGTFEYEEYVSLVSVSDP DEGYETILDLMISLQDSLDEFSQIIFASIQS ERRNTECKISALIPLIAESYGIYKFITSM RAHARLFEFYADCSSVKYLTTLVTIPKLPV DAPDVMHRQLNDAEGDAALQPLKERYELQ LINDVDESKEIKFKKREPSVT
ANTH	Sla2 AAAA mut	<i>S. cerevisiae</i>	NGTFEYEEYVSLVSVSDPDEGYETILDLM LQDSGAMGSMRIDSIDLQKALAKACSVVEET APARAACRACIVYTWHDQSSKAVFTTLKTL PLANDEVQLFKMLIVLHKIIQEGHPSALAE AIRDRDWIRSLGRVHSGGSSYSKLIREYVR YLVKLKLDFFHAHHRGFNLDEFSQIIFASIQS ERRNTECKISALIPLIAESYGIYKFITSM RLINDVDESKEIKFKKREPSVTTHARLFEFY ADCSSVKYLTTLVTIPKLPVDAPDVFAMHR QLNDAEGDAALQPLKERYELQ
ANTH	Yap1802	<i>S. cerevisiae</i>	GAMGSMSSLYTKLVKGATKIKMAPPKQKYV DPILSGTSSARGLQEIHALDIRLSDTAWT IVYKALIVLHLMIQQGEKDVTLRHYSNLD VFQLRKISHTTKWSSNDMRALQRYDEYLKT RCEEYGRGMDHLRDNYSKLGSKNQLSM DEELDHVESLEIQINALIRNKYSVSDLENH LLLYAFQLLVQDLLGLYNALNEGVIITLLES FFELSIHAKRTL DLYKDFVDMTEYVRYL KIGKAVGLKIPVIKHITTKLINSLEEHLRE ETKRQRG
ENTH	Epsin-1	<i>C. thermophilum</i>	GAMSKVIRSVKNVTKGYSSVQIKVREATSN DPWGPTGTQMSEIAQLTYGSSTDFYEIMDM LDKRLNDKGNWRHVVLKALKVMDYCLHEGS ELVVTWAKKNIFIIKTLREFQYIDEEGRDV GQNIRVAARELTALIQDEERLRAERNDRKM WKNRVNG
ANTH	Sla2	<i>C. thermophilum</i>	GAMATTRSLDHAKAEAEELAINIKKATSPEE TAPKRKHVRSCIVYTDHKSLSFWAGLKV QPILADEVQTFKALITIHKVLQEGHPVTLR EAMANRGWIDSLSRGMMGEGVRGYGPLIRE YVHFLAKLSFHKQHPEFNGTFEYEEYISL KAIHDPNEGYETITDMLTLQDKIDQFQKLI FSHFRHIGNNECRISALVPLVAESYGIYAQ HYRLVKFYECNLRYLTSLITIPKLFKIT SMLRAMHSSSTGDNEALEPLRQRYD

ENTH	Epsin-1	<i>H. sapiens</i>	GAMGSTSSLRQMKNIVHNYSEAEIKVREA TSNDPWFSPSSSLMSEIADLTYNVVAFSEIM SMIWKRLNDHGKNWRHVYKAMTLMEYLIKT GSERVSQQCKENMYAVQTLKDFQYVDRDGK DQGVNVREKAKQLVALLRDEDRLREERAHA LKTKEKLAQTA
ANTH	Hip1R	<i>H. sapiens</i>	GAMGSMNSIKNVPARVLSRRPGHSLEAERE QFDKTQAISISKAINTEAPVKEKHARRII LGTHHEKGAFTFWSYAIGLPLPSSSILSWK FCHVLHKVLRDGHPNVLHDCQRYRSNIREI GDLWGHLHLDRYGQLVNVYTKLLLTAKISFHL KHPQFPAGLEVTDDEVLEKAAGTDVNNIFQL TVEMFDYMDCELKLSVFRQLNTAIAVSQ MSSGQCRLAPLIQVIQDCSHLYHYTVKLLF KLHSCLPADTLQGHRDRFHEQFHS LRNFRR RASDMLYFKRLIQIPRLPEGPPNFLRASAL AEHIK
ANTH	CALM	<i>H. sapiens</i>	GAMSGQSLTDRITAAQHSVTGSAVSKTVCK ATTHEIMGPKKKHLDYLIQCTNEMNVNIPQ LADSLFERTTNSSWVVFKSLITTHHLMVY GNERFIQYLASRNTEFNLSNFLDKSGLQGY DMSTFIRRYSRYLNEKAVSYRQVAFDFTKV KRGADGVMRTMNTTEKLLKTVPIIQNQMDAL LDFNVNSNELTNGVINAAFMLLFKDAIRLF AAYNEGIINLLEKYFDMKKNQCKEGLDIYK KFLTRMTRISEFLKVAEQVGIDRGDIPDLS QAPSSLLDALEQHLASLEGKK

---

## S2.3 Molecular Organization of the *Salmonella* T3SS Sorting Platform

Table S9: Peak list of the MALDI MS/MS spectrum of the N-terminal peptide (3,523.71 *m/z*) of SpaOc. Reprinted from Bernal, Börnicke et al. (*in preparation*).

Mass	Intensity	Assigned ion	Mass	Intensity	Assigned ion
343.18668	61	<i>y</i> 3	1,811.74097	214	<i>y</i> 17
382.17477	38		1,908.74207	93	
456.24936	47		1,925.74500	309	<i>y</i> 18
510.21487	59		2,039.79761	235	<i>y</i> 19
584.29779	49	<i>y</i> 5	2,043.65857	127	
590.13623	36		2,080.66260	70	
595.33911	40		2,151.72168	88	
623.27216	115		2,168.81299	168	<i>y</i> 20
698.33301	79	<i>y</i> 6	2,171.71826	100	
736.35266	36		2,215.61000	43	<i>b</i> 19
811.37231	43	<i>y</i> 7	2,280.69385	91	
819.38165	40		2,297.82300	167	<i>y</i> 21
831.31989	32		2,344.67432	99	<i>b</i> 20
948.44794	118		2,426.84839	217	<i>y</i> 22
965.45569	738	<i>y</i> 9	2,445.79000	81	<i>b</i> 21
984.27399	59	<i>b</i> 8	2,494.78516	141	
1,033.33374	46		2,539.92090	189	<i>y</i> 23
1,078.52539	163	<i>y</i> 10	2,543.74902	205	
1,081.43127	78		2,558.78101	203	<i>b</i> 22
1,097.36902	65	<i>b</i> 9	2,561.68604	116	
1,179.55261	156	<i>y</i> 11	2,676.90894	176	<i>y</i> 24
1,210.44080	69		2,712.86000	31	<i>b</i> 24
1,226.40308	51	<i>b</i> 10	2,804.96069	71	<i>y</i> 25
1,291.44458	47		2,825.81519	62	<i>b</i> 25
1,308.61865	125	<i>y</i> 12	2,918.03223	264	<i>y</i> 26
1,339.47437	47		2,940.92212	97	<i>b</i> 26
1,355.41821	64	<i>b</i> 11	3,034.00098	82	<i>y</i> 27
1,379.62048	125	<i>y</i> 13	3,067.96000	45	<i>b</i> 27
1,480.64709	213	<i>y</i> 14	3,117.90283	85	
1,484.49255	75	<i>b</i> 12	3,119.01563	89	
1,535.44385	50		3,180.97510	53	<i>b</i> 28
1,599.46033	80	<i>b</i> 13	3,231.16504	42	
1,609.66479	166	<i>y</i> 15	3,248.09351	61	<i>y</i> 29
1,710.71301	190	<i>y</i> 16	3,375.27344	82	

**Table S10: Theoretical masses and average experimental masses of *Salmonella* SPI-1 T3SS sorting platform proteins and protein complexes as determined by native MS. (N $\geq$ 3, unless stated differently). Reprinted from Bernal, Börnicke et al. (*in preparation*).**

Protein/-complex	Theoretical mass (Da)	Exp. avg. mass (Da)	St.dev (Da)	Avg. FWHM (Da)
SpaOc MS/MS	11,176.0	11,170	50	40
SpaOc-Strep MS/MS	12,374.0	12,371	3	10
SicP MS/MS	14,580.7	14,579.3	0.1	10
SpaO <sub>1-145</sub> -Strep	17,291.8	17,292.0	0.9	12
Strep-SpaO <sub>140-297</sub>	18,863.6	18,863.7	0.6	13
2SpaOc	22,351.0	22,349	5	15
2SpaOc-Strep	24,748.0	24,746.9	1.5	14
OrgB MS/MS	26,448.4	26,459	21	100
2SicP	29,161.4	29,161.3	0.6	30
SpaO <sub>1-145</sub> -Strep/SpaOc-Strep MS/MS	29,665.8	29,640	11	130
SpaO MS/MS	33,793.7	33,800	40	250
2SicP-SptP <sub>35-139</sub>	41,523.7	41,557	50	140
SpaO <sub>1-145</sub> -Strep/2SpaOc-Strep	42,039.8	42,000	70	240
2SpaOc-Strep/Strep-SpaO <sub>140-297</sub>	43,611.3	43,663	26	410
SpaO/SpaOc MS/MS	44,838.1	44,880	50	50
SpaO-Strep/SpaOc-Strep MS/MS	47,524.0	47,504	23	70
InvC-Strep	48,808.9	48,240	110	370
SpaO/2SpaOc	56,013.6	56,050	40	190
Strep-SpaO/2SpaOc	57,471.2	57,552	19	270
SpaO-Strep/2SpaOc-Strep	59,897.9	59,930	40	250
SpaO <sub>1-145</sub> -Strep/2SpaOc-Strep/Strep-SpaO <sub>140-297</sub>	60,903.0	60,980	50	460
SpaO <sub>1-219</sub> /3SpaOc-Strep	61,786.4	61,780	70	450
SpaO <sub>1-145</sub> -Strep/4SpaOc-Strep	66,787.1	66,920	70	710
2(2SicP-SptP <sub>35-139</sub> )	83,047.5	83,112	60	190
2OrgB/InvC-Strep	101,705.6	102,220	190	820
2SpaO-Strep/3SpaOc-Strep MS/MS	107,422.0	107,390	50	190
SpaO/2SpaOc/2OrgB	109,041.5	109,230	100	890
2(SpaO/2SpaOc)	112,027.2	112,480	170	530
2SpaO/2SpaOc/OrgB MS/MS	116,386.9	116,150	40	320
2(SpaO-Strep/2SpaOc-Strep)	119,795.7	120,020	180	750
SpaO/2SpaOc/1OrgB/InvC-Strep MS/MS	131,402.1	131,350	90	240
2SpaO/4SpaOc/OrgB MS/MS	138,475.6	138,850	290	250
2SpaO/2SpaOc/2OrgB	142,835.2	143,400	240	1,300
2SpaO-Strep/4SpaOc-Strep/OrgB-His MS/MS	147,511.5	147,380	50	330
SpaO/2SpaOc/2OrgB/InvC-Strep	157,850.4	158,800	400	1,300
2SpaO/4SpaOc/2OrgB	164,923.9	165,370	160	910
2SpaO/2SpaOc/1OrgB/InvC-Strep MS/MS*	165,195.8	165,019	28	260
2SpaO-Strep/4SpaOc-Strep/2OrgB-His	175,227.2	175,500	500	890

SpaO/4SpaOc/2OrgB/InvC-Strep MS/MS	180,070.2	180,400	400	1,400
2SpaO/4SpaOc/OrgB/InvC-Strep MS/MS	187,284.5	187,620	170	490
2SpaO/2SpaOc/2OrgB/InvC-Strep	191,644.2	193,200	220	1,400
2SpaO/4SpaOc/2OrgB/InvC-Strep	213,732.8	214,500	500	1,800

---

\*N=2

**Table S11: SAXS data collection and structural parameters for SpaO/SpaOc/OrgB/InvC complex characterization.** Reprinted from Bernal, Börnicke et al. (*in preparation*).*Data collection parameters*

Instrument	P12 (EMBL/DESY, storage ring PETRA III, Germany)
Beam geometry	0.2 x 0.12 mm <sup>2</sup>
Wavelength (Å)	1.24
$q$ -range (Å <sup>-1</sup> )	0.008 – 0.47
Exposure time (s)	3,600 x 1
Temperature (K)	283

*Structural parameters*

Sample	$R_{g,Guinier}$ (nm)	$R_h$ (nm)	$D_{max}$ (nm)	$MM_{theor}$ (kDa)*	$MM_{DAM}$ (kDa)	SASBDB ID
SpaO/SpaOc/OrgB/InvC	5.7	7.3	19.0	169	208	SASDCA8

\*based on stoichiometry from native MS/MALS

**Table S12: Protein sequences of the used constructs of *S. Typhimurium* SPI-1 sorting platform proteins and chaperone-effector proteins.** Reprinted from Bernal, Börnicke et al. (*in preparation*).

<b>Protein</b>	<b>Sequence</b>
<i>InvC-Strep</i>	MKTPRLLQYLLAYPQKITGPIIEAELRDVAIGELCEIRRGWHQKQVVARAQVVG LQRETRVLSLIGNAQGLSRDVLVLYPTGRALSAWVGYSVLGAVLDPTGKIVERF TPEVAPISEERVIDVAPPSYASRVGVREPLITGVRAIDGLLTCGVGQRMGIFA SAGCGKTMMLMHMLIEQTEADVFIIGLIGERGREVTEFVDMRLRASHKKEKCVLV FATSDFPSVDRCNAAQLATTVAEYFRDQGKRVVLFIDSMTRYARALRDVALAS GERPARRGYPASVFDNLPRLLEP GATSEGSITAFYTVLLESEEEADPMADI RSILDGHLVLSRKLQGGHYPAIDVLKSVSRVFGQVTTPTTHAEQASAVRKLMT RLEELQLFIDLGEYRPGENIDNDRAMQMRDSLKAWLCQVPAQYSSFFDDTLSGM NAFADQNSAWSHPPQFEK
<i>OrgB</i>	MVKNIPSPSPVPEGILIKRKTLERYFSIERLEQQAHQRAKRILREAESEEAK TLRMYAYQEGYEQGMIDALQQVAAAYLTDNQTMAWKWMEKIQIYARELFSAAVD HPETLLTVLDEWLRDFDKPEGQLFLTLTPVNAKKDHQKLMVLLMENWPGTFNLK YHQEQRFIMSCGDQIAEFSPEQFVETAVGVIKHHLDELDPQDCRTISDNAINAL IDEWKTKTQAEVIR
<i>OrgB-His</i>	MGMLKNIPSPSPVPEGILIKRKTLERYFSIERLEQQAHQRAKRILREAESEE AKTLRMYAYQEGYEQGMIDALQQVAAAYLTDNQTMAWKWMEKIQIYARELFSAA VDHPETLLTVLDEWLRDFDKPEGQLFLTLTPVNAKKDHQKLMVLLMENWPGTFN LKYHQEQRFIMSCGDQIAEFSPEQFVETAVGVIKHHLDELDPQDCRTISDNAIN ALIDEWKTKTQAEVIRLEHHHHHH
<i>SicP</i>	MLQAHQDIIANIGEKLGPLTFDDNNQCLLLLDSDIFTSIEAKDDIWLNGMI IPLSPVCGDSIWRQIMVINGELAEANNEGTLAYIDAAETLLLIHAITDLTNTYH IISQLESFVNQQEALKNILQEYAKV
<i>SpaO</i>	MSLRVRQIDRREWLLAQTATECQRHGREATLEYPTRQGMWVRLSDAEKRWSAW IKPGDWLEHVSPALAGAAVSAGAEHLVVPWLAATERPFELPVPHLSCRRLCVE NPVPGSALPEGKLLHIMSDRGGLWFEHLPELPAVGGGRPKMLRWPLRFVIGSS DTQRSLGRIIGIDVLLIRTSRAEVYCYAKKLGHFNRVEGGIIVETLDIQHIE EENNTTETAETLPGLNQLPVKLEFVLYRKNVTLAELEAMGQQQLLSLPTNAEL NVEIMANGVLLGNGELVQMNDTLGVEIHEWLSESGNGE
<i>SpaO-Strep</i>	MASMSLRVRQIDRREWLLAQTATECQRHGREATLEYPTRQGMWVRLSDAEKRW SAWIKPGDWLEHVSPALAGAAVSAGAEHLVVPWLAATERPFELPVPHLSCRRL CVENPVPGSALPEGKLLHIMSDRGGLWFEHLPELPAVGGGRPKMLRWPLRFVI GSSDTQRSLGRIIGIDVLLIRTSRAEVYCYAKKLGHFNRVEGGIIVETLDIQ HIEEENNTTETAETLPGLNQLPVKLEFVLYRKNVTLAELEAMGQQQLLSLPTN AELNVEIMANGVLLGNGELVQMNDTLGVEIHEWLSESGNGESAWSHPPQFEK
<i>Strep-SpaO</i>	MASWSHPQFEKMGMSLRVRQIDRREWLLAQTATECQRHGREATLEYPTRQGMW VRLSDAEKRWSAWIKPGDWLEHVSPALAGAAVSAGAEHLVVPWLAATERPFEL PVPHLSCRRLCVENPVPGSALPEGKLLHIMSDRGGLWFEHLPELPAVGGGRPK MLRWPLRFVIGSSDTQRSLGRIIGIDVLLIRTSRAEVYCYAKKLGHFNRVEG GIIVETLDIQHIEEENNTTETAETLPGLNQLPVKLEFVLYRKNVTLAELEAMG QQQLLSLPTNAELNVEIMANGVLLGNGELVQMNDTLGVEIHEWLSESGNGE
<i>SpaOc</i>	METLDIQHIEEENNTTETAETLPGLNQLPVKLEFVLYRKNVTLAELEAMGQQQ LLSLPTNAELNVEIMANGVLLGNGELVQMNDTLGVEIHEWLSESGNGE
<i>SpaOc-Strep</i>	METLDIQHIEEENNTTETAETLPGLNQLPVKLEFVLYRKNVTLAELEAMGQQQ LLSLPTNAELNVEIMANGVLLGNGELVQMNDTLGVEIHEWLSESGNGESAWSH PPQFEK
<i>SpaO<sub>1-145</sub>-Strep</i>	MSLRVRQIDRREWLLAQTATECQRHGREATLEYPTRQGMWVRLSDAEKRWSAW IKPGDWLEHVSPALAGAAVSAGAEHLVVPWLAATERPFELPVPHLSCRRLCVE NPVPGSALPEGKLLHIMSDRGGLWFEHLPELPAVGGGRPSAWSHPPQFEK
<i>SpaO<sub>1-219</sub></i>	MSLRVRQIDRREWLLAQTATECQRHGREATLEYPTRQGMWVRLSDAEKRWSAW IKPGDWLEHVSPALAGAAVSAGAEHLVVPWLAATERPFELPVPHLSCRRLCVE NPVPGSALPEGKLLHIMSDRGGLWFEHLPELPAVGGGRPKMLRWPLRFVIGSS DTQRSLGRIIGIDVLLIRTSRAEVYCYAKKLGHFNRVEGGIIVETLDIQHIE EENNTTE

*SpaO*<sub>140-297</sub> MASWSHPQFEKGAVGGGRPKMLRWPLRFVIGSSDTQRSLLGRIGIGDVLLIRT  
SRAEVYCYAKKLGHFNRVEGGIIVETLDIQHIEEENNTTETAETLPGLNQLPV  
KLEFVLYRKNVTLAELEAMGQQQLLSLPTNAELNVEIMANGVLLGNGELVQMN  
DTLGVEIHEWLS

*SptP*<sub>35-139</sub> GSHMTDKAYVAPEKFSSKVLTWLGKMPLFKNTEVVQKHTENIRVQDQKILQTF  
LHALTEKYGETAVNDALLMSRINMNKPLTQRLAVQITECVKAADEGFINLIKS  
KDN

---

---

## S3 Material

### S3.1 Chemicals and Biomaterials

---

Chemical	CAS	Distributor
Acetic acid	64-19-7	Th. Geyer (Renningen, Germany)
Adenosine 5'-gamma-thio-triphosphate lithium salt (ATP- $\gamma$ -S)	93839-89-5	Jena Bioscience (Jena, Germany)
Alcohol Dehydrogenase (ADH) ( <i>S. cerevisiae</i> )	9031-72-5	Sigma-Aldrich (St. Louis, USA)
Ammonium acetate	631-61-8	Sigma-Aldrich (St. Louis, USA) Honeywell (Morristown, USA)
Ammonium hydroxide	1336-21-6	Sigma-Aldrich (St. Louis, USA)
Argon	7440-37-1	SOL Germany (Gersthofen, Germany)
Bovine serum albumin	9048-46-8	Sigma-Aldrich (St. Louis, USA)
Carbonic anhydrase II from human erythrocytes	9001-03-0	Sigma-Aldrich (St. Louis, USA)
Cesium Iodide	7789-17-5	Sigma-Aldrich (St. Louis, USA)
Concanavalin A ( <i>Canavalia ensiformis</i> )	11028-71-0	Sigma-Aldrich (St. Louis, USA)
Cytochrome c (equine heart)	9007-43-6	Sigma-Aldrich (St. Louis, USA)
Dithiothreitol,1,4- (DTT)	3483-12-3	Sigma-Aldrich (St. Louis, USA)
Dodecyl- $\beta$ -D-maltoside	69227-93-6	Sigma-Aldrich (St. Louis, USA)
Ethylenediaminetetraacetic acid (EDTA)	60-00-4	Sigma-Aldrich (St. Louis, USA)
Formic acid	64-18-6	Sigma-Aldrich (St. Louis, USA)
L-Glutamic Dehydrogenase (GDH) from bovine liver	9029-12-3	Sigma-Aldrich (St. Louis, USA)
Magnesium acetate tetrahydrate	16674-78-5	Sigma-Aldrich (St. Louis, USA)
Phosphatidylinositol 4,5-bisphosphate diC8	204858-53-7	Echelon Biosciences Inc (Salt Lake City, USA)
Pyruvate Kinase from rabbit muscle	9001-59-6	Sigma-Aldrich (St. Louis, USA)
Triethylammonium acetate	5204-74-0	Sigma-Aldrich (St. Louis, USA)
Xenon	7440-63-3	Linde (München, Germany)

---

## S3.2 Consumables

Consumable	Manufacturer
Centrifugal filter units Vivaspin 500 (5,000; 10,000; 30,000; 100,000 MWCO)	Sartorius (Göttingen, Germany)
Desalting spin columns Micro Bio-Spin™ 6	Bio-Rad Laboratories, Inc. (Hercules, USA)
Desalting spin columns Zeba™, 7K and 40K MWCO	Thermo Fisher Scientific (Waltham, USA)
Fused Silica tubing, flexible	Polymicro Technologies (Phoenix, USA)
Glass capillaries, 100 mm, 1.2 mm outer diameter, 0.68 mm inner diameter	World Precision Instruments (Sarasota, USA)
Gold Target (57mm Ø x 0.5mm)	LOT-QuantumDesign (Darmstadt, Germany)
Reaction tube 0.5 ml	VWR collection (Radnor, USA)
Reaction tube 1.5 ml SuperSpin™	VWR collection (Radnor, USA)
Reaction vessel, 15 ml	Greiner Bio-One (Kremsmünster, Austria)
Reaction vessel, 50 ml	Greiner Bio-One (Kremsmünster, Austria)

### S3.3 Instruments

<b>Instrument</b>	<b>Manufacturer</b>
Analytical balance ED 224S	Sartorius (Göttingen, Germany)
Analytical balance PB 303	Mettler Toledo (Columbus, USA)
Analytical balance PM 4600	Mettler Toledo (Columbus, USA)
Centrifuge 5417R	Eppendorf (Hamburg, Germany)
Centrifuge 5430R	Eppendorf (Hamburg, Germany)
Centrifuge Biofuge fresco	Thermo Fisher Scientific (Waltham, USA)
Ion mobility mass spectrometer Synapt G2	Waters Corporation (Milford, USA)
Mass spectrometer LCT	Waters Corporation (Milford, USA)
	MS Vision (Almere, The Netherlands)
Mass spectrometer QToF2	Waters Corporation (Milford, USA)
	MS Vision (Almere, The Netherlands)
pH-meter 913	Metrohm (Herisau, Switzerland)
Spectrophotometer DS-11 FX	DeNovix (Wilmington, USA)
Sputter coater Q150R	Quorum Technologies (Lewes, UK)

### S3.4 Software

<b>Software</b>	<b>Version</b>	<b>Company</b>
DriftScope	2.3	Waters Corporation(Milford, USA)
GraphPad Prism	5.03	GraphPad Software, Inc. (San Diego, USA)
IMPACT	0.9.1	Marklund, Degiacomi et al. (2015)
Massign	As published 02/2012	(Morgner & Robinson, 2012)
MassLynx	4.0/4.1	Waters Corporation(Milford, USA)

## S4 Used Hazardous Substances According to GHS

Name (abbreviation)	Symbols	Hazard Statements (H) Precautionary Statements (P)
Acetic acid	 	<p><b>H226</b> Flammable liquid and vapour  <b>H290</b> May be corrosive to metals  <b>H314</b> Causes severe skin burns and eye damage  <b>P210</b> Keep away from heat, hot surfaces, sparks, open flames and other ignition sources. No smoking.  <b>P280</b> Wear protective gloves/protective clothing/eye protection/face protection.  <b>P303+P361+P353</b> IF ON SKIN (or hair): take off immediately all contaminated clothing. Rinse skin with water/shower.  <b>P305+P351+P338</b> IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing.  <b>P310</b> Immediately call a POISON CENTER/doctor.</p>
Adenosine 5'-gamma-thio-triphosphate lithium salt (ATP-γ-S)		<p><b>H315</b> Causes skin irritation.  <b>H319</b> Causes serious eye irritation.  <b>H335</b> May cause respiratory irritation  <b>P312</b> Call a POISON CENTER or doctor/ physician if you feel unwell.  <b>P305+P351+P338</b> IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing.  <b>P304+P340</b> IF INHALED: Remove to fresh air and keep at rest in a position comfortable for breathing.  <b>P264</b> Wash face, hands and any exposed skin thoroughly after handling.  <b>P280</b> Wear protective gloves/ eye protection/ face protection.  <b>P362</b> Take off contaminated clothing and wash before reuse.  <b>P261</b> Avoid breathing dust/ fume/ gas/ mist/ vapors/ spray.  <b>P271</b> Use only outdoors or in a well-ventilated area.  <b>P403+P233</b> Store in a well-ventilated place. Keep container tightly closed.  <b>P405</b> Store locked up.  <b>P501</b> Dispose of contents/ container to an approved waste disposal plant.</p>
Ammonium hydroxide	  	<p><b>H290</b> May be corrosive to metals.  <b>H314</b> Causes severe skin burns and eye damage.  <b>H335</b> May cause respiratory irritation.  <b>H400</b> Very toxic to aquatic life.  <b>P273</b> Avoid release to the environment.  <b>P280</b> Wear protective gloves/protective clothing/eye protection/face protection.  <b>P303+P361+P353</b> IF ON SKIN (or hair): take off immediately all contaminated clothing. Rinse skin with water/shower.  <b>P304+P340</b> IF INHALED: Remove person to fresh air and keep comfortable for breathing.  <b>P305+P351+P338</b> IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing.  <b>P310</b> Immediately call a POISON CENTER/doctor.</p>

Argon		<b>H280</b> Contains gas under pressure; may explode if heated. <b>P403</b> Store in a well-ventilated place.
Cesium Iodide	 	<b>H315</b> Causes skin irritation. <b>H317</b> May cause an allergic skin reaction. <b>H319</b> Causes serious eye irritation. <b>H335</b> May cause respiratory irritation. <b>H410</b> Very toxic to aquatic life with long lasting effects. <b>P280</b> Wear protective gloves. <b>P305+P351+P338</b> IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing.
Dithiothreitol,1,4- (DTT)		<b>H302</b> Harmful if swallowed <b>H315</b> Causes skin irritation <b>H319</b> Causes serious eye irritation <b>H335</b> May cause respiratory irritation <b>H412</b> Harmful to aquatic life with long lasting effects <b>P260</b> Do not breathe dust. <b>P270</b> Do not eat, drink or smoke when using this product. <b>P305+P351+P338</b> IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. <b>P337+P313</b> If eye irritation persists: Get medical advice/attention.
Ethylenediaminetetraacetic acid (EDTA)		<b>H319</b> Causes serious eye irritation. <b>P305+P351+P338</b> IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing.
Formic acid	  	<b>H226</b> Flammable liquid and vapour <b>H290</b> May be corrosive to metals <b>H302</b> Harmful if swallowed <b>H314</b> Causes severe skin burns and eye damage <b>H331</b> Toxic if inhaled <b>P210</b> Keep away from heat. No smoking. <b>P280</b> Wear protective gloves/protective clothing/eye protection/face protection. <b>P303+P361+P353</b> IF ON SKIN (or hair): take off immediately all contaminated clothing. Rinse skin with water/shower. <b>P304+P340</b> IF INHALED: Remove person to fresh air and keep comfortable for breathing. <b>P305+P351+P338</b> IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. <b>P310</b> Immediately call a POISON CENTER/doctor.
Helium		<b>H280</b> Contains gas under pressure; may explode if heated. <b>P403</b> Store in a well-ventilated place.
Nitrogen		<b>H280</b> Contains gas under pressure; may explode if heated. <b>P403</b> Store in a well-ventilated place.
Xenon		<b>H280</b> Contains gas under pressure; may explode if heated. <b>P403</b> Store in a well-ventilated place.

---

## S5 Contributions

Experimental chapters 3-5 are results from collaborative projects that involved valuable contributions from several people. The here presented data are mass spectrometry-centred views on the projects. Supportive results from other techniques were added to this thesis to underline and accentuate the MS findings. Unless specified differently, all work has been performed by me, Johannes Heidemann (JH). The following list is an overview over other author's contributions:

**Chapter 3** JN produced and purified the proteins. JH performed native MS measurements and data analysis. IMMS data were acquired by JH with assistance of AK and JvD, IMMS results were analysed by JH. JH wrote the chapter, subsection 3.3.1 (Methods) was taken with minor adaptations by JH from the submitted manuscript.

**Chapter 4** MGA produced and purified the proteins for native MS analysis and performed biophysical experiments (ITC, DLS). HM performed the SAXS analysis and together with DS the structural modelling. AG crystallized the ENTH domains in presence of PI(4,5)P<sub>2</sub>. JH performed native MS measurements. IMMS data were acquired by JH with assistance of AK and JvD. SID MS experiments were performed by JH and AK. All MS data were analysed and interpreted by JH. JH wrote the chapter, parts of the Methods section (4.3) are summarized by JH based on the methods section of a manuscript that was originally written by RM, MGA and JH (Garcia-Alai, Heidemann et al., 2018).

**Chapter 5** IB and JB cloned the genes, produced and purified proteins, performed secretion assays and biochemical experiments. AT performed SAXS experiments and together with DS the structural modelling. PJ and MS performed the MALDI MS measurements. Chapters 5.3.1-3, 5.3.5-7 (Methods) are taken with minor adaptations by JH from the manuscript, which is in preparation by all authors. JH performed native MS experiments, analysed the data and wrote the chapter.

### Initials and Names:

AG	Anna Gieras	JN	Julius Nitsche
AK	Albert Konijnenberg	JvD	Jeroen van Dyck
AT	Anne Tuukkanen	MGA	Maria Garcia Alai
DS	Dmitri Svergun	MS	Monika Schmidt
HM	Haydyn Mertens	PJ	Peter Jungblut
JH	Johannes Heidemann	RM	Rob Meijers

## S6 List of Figures

		Page
Figure 1	Mass spectra of membrane-associated domains of epsin-2 and Sla2 from <i>Saccharomyces cerevisiae</i> .	3
Figure 2	Schematic representation of a nano electrospray ionization (ESI) source in positive ion mode.	8
Figure 3	Macroscopic and microscopic binding constants.	16
Figure 4	Protein-protein interactions in membranous environments.	26
Figure 5	Hypothesized two-step activation mechanism by binding of two Ca <sup>2+</sup> -loaded CaMs to the regulatory domain of ACA8.	28
Figure 6	Binding of two calmodulins to full-length, detergent-solubilized ACA8.	29
Figure 7	Native mass spectrum revealed binding of two CaMs to ACA8 and MS-induced dissociation.	31
Figure 8	Native MS analysis of ACA8/CaM on a Synapt G2 mass spectrometer (Waters) revealed a 99.7 kDa-species.	32
Figure 9	Clathrin-mediated endocytosis involves interactions between plasma membrane, adaptor proteins and actin cytoskeleton.	38
Figure 10	Phosphatidylinositol-(4,5)-bisphosphate.	40
Figure 11	Clathrin adaptor protein structures.	42
Figure 12	Two PI(4,5)P <sub>2</sub> molecules bind to epsin ENTH domains.	44
Figure 13	Crystal structure of two ENTH domains from <i>S. cerevisiae</i> epsin-2 binding one PI(4,5)P <sub>2</sub> molecule.	47
Figure 14	Two PI(4,5)P <sub>2</sub> molecules bind to wild type ANTH domains.	48
Figure 15	ENTH:ANTH:PI(4,5)P <sub>2</sub> complex formation in fungi.	51
Figure 16	Lipid stoichiometry of <i>S. cerevisiae</i> 6:6 ENTH2:ANTH complexes revealed by CID MS/MS.	53
Figure 17	ENTH:ANTH:PI(4,5) assembled with mixed ENTH domains from <i>S. cerevisiae</i> epsin-1 and epsin-2.	54
Figure 18	Reduced PI(4,5)P <sub>2</sub> binding capacities of <i>S. cerevisiae</i> ENTH and ANTH mutant domains obviated complex formation.	55
Figure 19	Dynamics of fungal ENTH:ANTH:PI(4,5)P <sub>2</sub> assemblies.	57
Figure 20	Ion mobility mass spectrometry of <i>S. cerevisiae</i> ENTH2:ANTH:PI(4,5)P <sub>2</sub> complexes.	58
Figure 21	Surface-induced dissociation MS of <i>S. cerevisiae</i> ENTH1:ANTH:PI(4,5)P <sub>2</sub> .	62
Figure 22	Low- and high-energy SID MS of <i>S. cerevisiae</i> ENTH1:ANTH:PI(4,5)P <sub>2</sub> .	63
Figure 23	ENTH domains from <i>H. sapiens</i> epsin-1 form stable hexamers in presence of PI(4,5)P <sub>2</sub> .	65
Figure 24	Ion mobility mass spectrometry of PI(4,5)P <sub>2</sub> -containing <i>H. sapiens</i> epsin-1 ENTH hexamers.	68
Figure 25	SAXS modelling of human epsin-1 ENTH in the presence of PI(4,5)P <sub>2</sub> .	70
Figure 26	Human Hip1R ANTH domains bind to epsin-1 ENTH hexamers.	72
Figure 27	Collision-induced dissociation analysis of <i>H. sapiens</i> ENTH:ANTH:PI(4,5)P <sub>2</sub> complexes.	73
Figure 28	No ENTH:ANTH:PI(4,5)P <sub>2</sub> complexes with ANTH domains of the CALM subfamily were identified.	74
Figure 29	Cross-species complex assembly of ENTH:ANTH:PI(4,5)P <sub>2</sub> complexes.	76
Figure 30	Schematic models of ENTH:ANTH:PI(4,5)P <sub>2</sub> complex assembly.	78
Figure 31	Schematic representation of Type III secretion systems.	90
Figure 32	<i>Salmonella</i> Typhimurium T3SS structure as determined by electron cryo tomography.	92
Figure 33	SpaOc is produced by translation from an internal initiation start site in the <i>spaO</i> gene.	95
Figure 34	Dimerization of SpaOc proteins.	96

Figure 35	SpaO isoforms form stable SpaO-2SpaOc complexes.	97
Figure 36	Inter- and intramolecular domain interactions of SpaO-2SpaOc complexes	99
Figure 37	Dimerization of SpaO-2SpaOc is induced by binding of OrgB dimers.	102
Figure 38	The ATPase InvC binds to SpaO/SpaOc/OrgB complexes.	104
Figure 39	Identification of SpaO/SpaOc/OrgB/InvC complex stoichiometries by CID MS/MS.	106
Figure 40	SEC-SAXS analysis and computational modelling of SpaO/SpaOc/OrgB/InvC complexes.	108
Figure 41	SicP/SptP chaperone/effector complexes do not interact with T3SS sorting platform components SpaO/SpaOc/OrgB/InvC.	110
Figure 42	T3SS sorting platform assembly model.	112
<b>Supplement</b>		
Figure S1	Native mass spectrum of <i>A. thaliana</i> CaM in absence of ACA8 and DDM.	121
Figure S2	Preliminary IMMS analysis of ACA8/CaM.	122
Figure S3	CID MS/MS analysis of <i>S. cerevisiae</i> ENTH:ANTH:PI(4,5)P <sub>2</sub> complexes containing ENTH domains from epsin-1 and/or epsin-2.	123
Figure S4	Logarithmic fit calibration for the CCS' determination of <i>S. cerevisiae</i> ENTH:ANTH:PI(4,5)P <sub>2</sub> complexes.	124
Figure S5	Stability of clathrin adaptor protein complexes assessed in a dynamic light scattering thermal aggregation assay.	125
Figure S6	Characterization of the unidentified human PI(4,5)P <sub>2</sub> -containing ENTH complexes.	126
Figure S7	Logarithmic fit calibration for the CCS' determination of PI(4,5)P <sub>2</sub> -containing <i>H. sapiens</i> ENTH hexamers.	127
Figure S8	The phospholipid content in human PI(4,5)P <sub>2</sub> -containing ENTH hexamers is not determining complex unfolding.	128
Figure S9	Human PI(4,5)P <sub>2</sub> -containing ENTH hexamers analysed by small-angle X-ray scattering.	129
Figure S10	Isothermal titration calorimetry revealed PI(4,5)P <sub>2</sub> -dependent binding of ANTH and ENTH domains.	130
Figure S11	MALDI MS analysis revealed the sequence of the N-terminal SpaOc peptide.	131
Figure S12	Fully efficient <i>Salmonella</i> Typhimurium type III secretion is dependent on SpaOc.	132
Figure S13	Characterization of SpaO/SpaOc oligomerization.	133
Figure S14	Native MS analysis of SpaO-SpaOc domain interactions.	134
Figure S15	SEC-MALS analysis of SpaO-SpaOc domain interaction.	135
Figure S17	SpaO/SpaOc/OrgB/InvC spectra revealed formation of higher-order oligomers.	137
Figure S18	SEC-MALS analysis of SpaO/SpaOc/OrgB/InvC complexes.	138

---

## S7 List of Tables

		Page
Table 1	Macroscopic $K_D$ values (standard deviation, N=3) of PI(4,5)P <sub>2</sub> binding to different ENTH domains.	46
Table 2	Macroscopic $K_D$ values (standard deviation, N=3) of PI(4,5)P <sub>2</sub> binding to different ANTH domains.	49
Table 3	Measured drift times and derived CCS' for PI(4,5)P <sub>2</sub> containing 8:8 and 6:6 ENTH2:ANTH complexes from <i>S. cerevisiae</i> .	59
Table 4	Measured drift times and derived CCSs' for PI(4,5)P <sub>2</sub> containing epsin-1 ENTH hexamers from <i>H. sapiens</i> .	67
Table 5	Measured drift times and derived CCSs' for PI(4,5)P <sub>2</sub> containing epsin-1 ENTH hexamers from <i>H. sapiens</i> that were exposed to 75 V trap collision energy.	69
Table S1	Theoretical and experimental molecular weights (M) of <i>A. thaliana</i> ACA8, CaM and complexes of these proteins.	139
Table S2	Amino acid sequences of <i>A. thaliana</i> proteins ACA8 and CaM7.	140
Table S3	Molecular weights (M) of single ENTH and ANTH domains and ENTH/ANTH/PI(4,5)P <sub>2</sub> complexes from <i>S. cerevisiae</i> , <i>C. thermophilum</i> and <i>H. sapiens</i> .	141
Table S4	<i>S. cerevisiae</i> ENTH2 X-ray crystallographic data collection and refinement statistics.	143
Table S5	SAXS Data collection and derived parameters for human ENTH.	144
Table S6	Sequence identity matrix of investigated ENTH domains.	145
Table S7	Sequence identity matrix of investigated ANTH domains.	145
Table S8	Amino acid sequences of ENTH and ANTH domains.	146
Table S9	Peak list of the MALDI MS/MS spectrum of the N-terminal peptide (3,523.71 <i>m/z</i> ) of SpaOc.	148
Table S10	Theoretical masses and average experimental masses of <i>Salmonella</i> SPI-1 T3SS sorting platform proteins and protein complexes as determined by native MS.	149
Table S11	SAXS data collection and structural parameters for SpaO/SpaOc/OrgB/InvC complex characterization.	151
Table S12	Protein sequences of the used constructs of <i>S. Typhimurium</i> SPI-1 sorting platform proteins and chaperone-effector proteins.	152

---

## S8 List of Abbreviations

aa	Amino acids
CC BY	Creative Commons Attribution License
CCS	Collision cross section
CID	Collision-induced dissociation
CMC	Critical micelle concentration
CME	Clathrin mediated endocytosis
C-ring	Cytoplasmic ring
Cryo-ET	Electron cryo tomography
DC	Direct current
DLS	Dynamic light scattering
dRI	Differential refractive index
EMBL	The European Molecular Biology Laboratory
ER	Endoplasmatic reticulum
ESI	Electrospray ionization
FWHM	Full width at half maximum
GHS	Globally Harmonized System of Classification and Labelling of Chemicals
GPCR	G protein-coupled receptor
HIV	Human immunodeficiency virus
IMMS	Ion mobility mass spectrometry
IMS	Ion mobility separation
ITC	Isothermal titration calorimetry
IUPAC	International Union of Pure and Applied Chemistry
MALDI	Matrix-assisted laser desorption/ionization
MALS	Multi-angle light scattering
MD	Molecular dynamics
MS	Mass spectrometry
MWCO	Molecular weight cut-off
NMR	Nuclear magnetic resonance
PMCA	Plasma membrane Ca <sup>2+</sup> ATPase
PTM	Post-translational modification
QToF	Quadrupole time-of-flight
RF	Radio frequency
RI	Refractive index
RMSD	Root mean square deviation
RNA	Ribonucleic acid
SAXS	Small-angle X-ray scattering
SEC	Size-exclusion chromatography
SID	Surface-induced dissociation
SPI	<i>Salmonella</i> pathogenicity island

SPR	Surface plasmon resonance
St. dev.	Standard deviation
T3SS	Type 3 secretion system
TEV	Tobacco Etch Virus
ToF	Time-of-flight
TWIMS	Travelling wave ion mobility mass spectrometry

---

## S9 References

Abby SS, Rocha EP (2012) The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. *PLoS Genet* 8: e1002983

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68: 352-67

Aguilar RC, Longhi SA, Shaw JD, Yeh LY, Kim S, Schon A, Freire E, Hsu A, McCormick WK, Watson HA, Wendland B (2006) Epsin N-terminal homology domains perform an essential function regulating Cdc42 through binding Cdc42 GTPase-activating proteins. *P Natl Acad Sci USA* 103: 4116-21

Ahadi E, Konermann L (2012) Modeling the behavior of coarse-grained polymer chains in charged water droplets: implications for the mechanism of electrospray ionization. *J Phys Chem B* 116: 104-12

Akeda Y, Galan JE (2005) Chaperone release and unfolding of substrates in type III secretion. *Nature* 437: 911-5

Antonescu CN, Aguet F, Danuser G, Schmid SL (2011) Phosphatidylinositol-(4,5)-bisphosphate regulates clathrin-coated pit initiation, stabilization, and size. *Mol Biol Cell* 22: 2588-600

Asano S, Engel BD, Baumeister W (2016) In Situ Cryo-Electron Tomography: A Post-Reductionist Approach to Structural Biology. *J Mol Biol* 428: 332-43

Avinoam O, Schorb M, Beese CJ, Briggs JA, Kaksonen M (2015) Endocytic sites mature by continuous bending and remodeling of the clathrin coat. *Science* 348: 1369-72

Baca M, Kent SBH (1992) Direct Observation of a Ternary Complex between the Dimeric Enzyme Hiv-1 Protease and a Substrate-Based Inhibitor. *J Am Chem Soc* 114: 3992-3

Barrera NP, Di Bartolo N, Booth PJ, Robinson CV (2008) Micelles protect membrane complexes from solution to vacuum. *Science* 321: 243-6

Barrera NP, Robinson CV (2011) Advances in the mass spectrometry of membrane proteins: from individual proteins to intact complexes. *Annu Rev Biochem* 80: 247-71

Ben-Nissan G, Sharon M (2018) The application of ion-mobility mass spectrometry for structure/function investigation of protein complexes. *Curr Opin Chem Biol* 42: 25-33

Benesch JL (2009) Collisional activation of protein complexes: picking up the pieces. *J Am Soc Mass Spectrom* 20: 341-8

Berggård T, Linse S, James P (2007) Methods for the detection and analysis of protein-protein interactions. *Proteomics* 7: 2833-42

Bernal I, Börnicke J, Heidemann J, Svergun D, Tuukkanen A, Uetrecht C, Kolbe M (*in preparation*) Molecular organization of the Salmonella type III secretion system sorting platform.

- Berridge MJ (1987) Inositol trisphosphate and diacylglycerol: two interacting second messengers. *Annu Rev Biochem* 56: 159-93
- Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, Franke D, Mark D, Zengerle R, Cipriani F, Fiedler S, Roessle M, Svergun DI (2015) Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J Appl Crystallogr* 48: 431-43
- Blaylock B, Riordan KE, Missiakas DM, Schneewind O (2006) Characterization of the *Yersinia enterocolitica* type III secretion ATPase YscN and its regulator, YscL. *J Bacteriol* 188: 3525-34
- Bleiholder C, Wyttenbach T, Bowers MT (2011) A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). *Method. Int J Mass Spectrom* 308: 1-10
- Bochner BR, Huang HC, Schieven GL, Ames BN (1980) Positive selection for loss of tetracycline resistance. *J Bacteriol* 143: 926-33
- Boeri Erba E, Barylyuk K, Yang Y, Zenobi R (2011) Quantifying protein-protein interactions within noncovalent complexes using electrospray ionization mass spectrometry. *Anal Chem* 83: 9251-9
- Boeri Erba E, Ruotolo BT, Barsky D, Robinson CV (2010) Ion mobility-mass spectrometry reveals the influence of subunit packing and charge on the dissociation of multiprotein complexes. *Anal Chem* 82: 9702-10
- Boeri Erba E, Zenobi R (2011) Mass spectrometric studies of dissociation constants of noncovalent complexes. *Annu Rep Prog Chem C* 107: 199-228
- Boldon L, Laliberte F, Liu L (2015) Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application. *Nano Rev* 6: 25661
- Bonza MC, Morandini P, Luoni L, Geisler M, Palmgren MG, De Michelis MI (2000) At-ACA8 encodes a plasma membrane-localized calcium-ATPase of *Arabidopsis* with a calmodulin-binding domain at the N terminus. *Plant Physiol* 123: 1495-505
- Bootman MD (2012) Calcium signaling. *Cold Spring Harb Perspect Biol* 4: a011171
- Boulant S, Kural C, Zeeh JC, Ubelmann F, Kirchhausen T (2011) Actin dynamics counteract membrane tension during clathrin-mediated endocytosis. *Nat Cell Biol* 13: 1124-31
- Brett TJ, Legendre-Guillemain V, McPherson PS, Fremont DH (2006) Structural definition of the F-actin-binding THATCH domain from HIP1R. *Nat Struct Mol Biol* 13: 121-30
- Breuker K, McLafferty FW (2008) Stepwise evolution of protein native structure with electrospray into the gas phase,  $10^{-12}$  to  $10^2$  s. *P Natl Acad Sci USA* 105: 18145-52
- Bush MF, Hall Z, Giles K, Hoyes J, Robinson CV, Ruotolo BT (2010) Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. *Anal Chem* 82: 9557-65
- Bzymek KP, Hamaoka BY, Ghosh P (2012) Two translation products of *Yersinia* yscQ assemble to form a complex essential to type III secretion. *Biochemistry* 51: 1669-77

- Carafoli E (1994) Biogenesis: plasma membrane calcium ATPase: 15 years of work on the purified enzyme. *FASEB J* 8: 993-1002
- Chen C, Zhuang X (2008) Epsin 1 is a cargo-specific adaptor for the clathrin-mediated endocytosis of the influenza virus. *P Natl Acad Sci USA* 105: 11790-5
- Chen F, Gulbakan B, Weidmann S, Fagerer SR, Ibanez AJ, Zenobi R (2016) Applying mass spectrometry to study non-covalent biomolecule complexes. *Mass Spectrom Rev* 35: 48-70
- Chen H, Ko G, Zatti A, Di Giacomo G, Liu L, Raiteri E, Perucco E, Collesi C, Min W, Zeiss C, De Camilli P, Cremona O (2009) Embryonic arrest at midgestation and disruption of Notch signaling produced by the absence of both epsin 1 and epsin 2 in mice. *P Natl Acad Sci USA* 106: 13838-43
- Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66: 12-21
- Chernushevich IV, Loboda AV, Thomson BA (2001) An introduction to quadrupole-time-of-flight mass spectrometry. *J Mass Spectrom* 36: 849-65
- Chernushevich IV, Thomson BA (2004) Collisional cooling of large ions in electrospray mass spectrometry. *Anal Chem* 76: 1754-60
- Chowdhury SK, Katta V, Chait BT (1990) Probing Conformational-Changes in Proteins by Mass-Spectrometry. *J Am Chem Soc* 112: 9012-3
- Clapham DE (2007) Calcium signaling. *Cell* 131: 1047-58
- Claret L, Calder SR, Higgins M, Hughes C (2003) Oligomerization and activation of the FliI ATPase central to bacterial flagellum assembly. *Mol Microbiol* 48: 1349-55
- Confalonieri S, Di Fiore PP (2002) The Eps15 homology (EH) domain. *FEBS Lett* 513: 24-9
- Cong Y, Katipamula S, Trader CD, Orton DJ, Geng T, Baker ES, Kelly RT (2016) Mass spectrometry-based monitoring of millisecond protein-ligand binding dynamics using an automated microfluidic platform. *Lab on a Chip* 16: 1544-8
- Cooks RG (1995) Special feature: Historical. Collision-induced dissociation: Readings and commentary. *J Mass Spectrom* 30: 1215-21
- Coon BG, D'Amico DM, Konieczny SF, Aguilar RC (2011) Epsins' novel role in cancer cell invasion. *Commun Integr Biol* 4: 95-7
- Cossart P, Helenius A (2014) Endocytosis of viruses and bacteria. *Cold Spring Harb Perspect Biol* 6: a016972
- Covey TR, Bonner RF, Shushan BI, Henion J (1988) The determination of protein, oligonucleotide and peptide molecular weights by ion-spray mass spectrometry. *Rapid Commun Mass Spectrom* 2: 249-56
- Damke H, Baba T, Warnock DE, Schmid SL (1994) Induction of mutant dynamin specifically blocks endocytic coated vesicle formation. *J Cell Biol* 127: 915-34

- Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *P Natl Acad Sci USA* 97: 6640-5
- De Camilli P, Chen H, Hyman J, Panepucci E, Bateman A, Brunger AT (2002) The ENTH domain. *FEBS Lett* 513: 11-8
- Dean P (2011) Functional domains and motifs of bacterial type III effector proteins and their roles in infection. *FEMS Microbiol Rev* 35: 1100-25
- Dell A, Galadari A, Sastre F, Hitchen P (2011) Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *Int J Microbiol* 2010
- Demain AL, Vaishnav P (2009) Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv* 27: 297-306
- Devine PWA, Fisher HC, Calabrese AN, Whelan F, Higazi DR, Potts JR, Lowe DC, Radford SE, Ashcroft AE (2017) Investigating the Structural Compaction of Biomolecules Upon Transition to the Gas-Phase Using ESI-TWIMS-MS. *J Am Soc Mass Spectrom* 28: 1855-62
- Di Leva F, Domi T, Fedrizzi L, Lim D, Carafoli E (2008) The plasma membrane Ca<sup>2+</sup> ATPase of animal cells: structure, function and regulation. *Arch Biochem Biophys* 476: 65-74
- Dickson EJ, Jensen JB, Hille B (2014) Golgi and plasma membrane pools of PI(4)P contribute to plasma membrane PI(4,5)P<sub>2</sub> and maintenance of KCNQ2/3 ion channel current. *P Natl Acad Sci USA* 111: E2281-90
- Diepold A, Kudryashev M, Delalez NJ, Berry RM, Armitage JP (2015) Composition, formation, and regulation of the cytosolic c-ring, a dynamic component of the type III secretion injectisome. *PLoS Biol* 13: e1002039
- Dietsche T, Tesfazgi Mebrhatu M, Brunner MJ, Abrusci P, Yan J, Franz-Wachtel M, Scharfe C, Zilkenat S, Grin I, Galan JE, Kohlbacher O, Lea S, Macek B, Marlovits TC, Robinson CV, Wagner S (2016) Structural and Functional Characterization of the Bacterial Type III Secretion Export Apparatus. *PLoS Pathog* 12: e1006071
- Dixit SM, Polasky DA, Ruotolo BT (2018) Collision induced unfolding of isolated proteins in the gas phase: past, present, and future. *Curr Opin Chem Biol* 42: 93-100
- Dole M, Mack LL, Hines RL (1968) Molecular Beams of Macroions. *J Chem Phys* 49: 2240-9
- Duncan MC, Costaguta G, Payne GS (2003) Yeast epsin-related proteins required for Golgi-endosome traffic define a  $\gamma$ -adaptin ear-binding motif. *Nat Cell Biol* 5: 77
- El-Hawiet A, Kitova EN, Klassen JS (2012) Quantifying carbohydrate-protein interactions by electrospray ionization mass spectrometry analysis. *Biochemistry* 51: 4244-53
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66: 486-501
- Engqvist-Goldstein AE, Warren RA, Kessels MM, Keen JH, Heuser J, Drubin DG (2001) The actin-binding protein Hip1R associates with clathrin during early stages of endocytosis and promotes clathrin assembly in vitro. *J Cell Biol* 154: 1209-23

- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246: 64-71
- Fernandez De La Mora J (2000) Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal Chim Acta* 406: 93-104
- Ford MG, Mills IG, Peter BJ, Vallis Y, Praefcke GJ, Evans PR, McMahon HT (2002) Curvature of clathrin-coated pits driven by epsin. *Nature* 419: 361-6
- Ford MG, Pearse BM, Higgins MK, Vallis Y, Owen DJ, Gibson A, Hopkins CR, Evans PR, McMahon HT (2001) Simultaneous binding of PtdIns(4,5)P<sub>2</sub> and clathrin by AP180 in the nucleation of clathrin lattices on membranes. *Science* 291: 1051-5
- Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM, Svergun DI (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr* 50: 1212-25
- Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* 42: 342-6
- Freyer MW, Lewis EA (2008) Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods Cell Biol* 84: 79-113
- Frieden E (1975) Non-covalent interactions: Key to biological flexibility and specificity. *J Chem Educ* 52: 754-61
- Fruman DA, Meyers RE, Cantley LC (1998) Phosphoinositide kinases. *Annu Rev Biochem* 67: 481-507
- Fujimoto LM, Roth R, Heuser JE, Schmid SL (2000) Actin assembly plays a variable, but not obligatory role in receptor-mediated endocytosis in mammalian cells. *Traffic* 1: 161-71
- Galan JE (2007) SnapShot: effector proteins of type III secretion systems. *Cell* 130: 192
- Galan JE, Curtiss R, 3rd (1989) Cloning and molecular characterization of genes whose products allow *Salmonella typhimurium* to penetrate tissue culture cells. *P Natl Acad Sci USA* 86: 6383-7
- Galán JE, Lara-Tejero M, Marlovits TC, Wagner S (2014) Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells. *Annu Rev Microbiol* 68: 415-38
- Galaz-Montoya JG, Ludtke SJ (2017) The advent of structural biology in situ by single particle cryo-electron tomography. *Biophys Rep* 3: 17-35
- Galhena AS, Dagan S, Jones CM, Beardsley RL, Wysocki VH (2008) Surface-induced dissociation of peptides and protein complexes in a quadrupole/time-of-flight mass spectrometer. *Anal Chem* 80: 1425-36
- Gamper N, Shapiro MS (2007) Target-specific PIP<sub>2</sub> signalling: how might it work? *J Physiol* 582: 967-75

- Gan J, Ben-Nissan G, Arkind G, Tarnavsky M, Trudeau D, Noda Garcia L, Tawfik DS, Sharon M (2017) Native Mass Spectrometry of Recombinant Proteins from Crude Cell Lysates. *Anal Chem* 89: 4398-404
- Ganem B, Li YT, Henion JD (1991) Detection of Noncovalent Receptor Ligand Complexes by Mass-Spectrometry. *J Am Chem Soc* 113: 6294-6
- Garcia-Alai MM, Heidemann J, Skruzny M, Gieras A, Mertens HDT, Svergun DI, Kaksonen M, Uetrecht C, Meijers R (2018) Epsin and Sla2 form assemblies through phospholipid interfaces. *Nat Commun* 9: 328
- Gaytan MO, Martinez-Santos VI, Soto E, Gonzalez-Pedrajo B (2016) Type Three Secretion System in Attaching and Effacing Pathogens. *Front Cell Infect Microbiol* 6: 129
- Goh JB, Ng SK (2017) Impact of host cell line choice on glycan profile. *Crit Rev Biotechnol*: 1-17
- Gottfried I, Ehrlich M, Ashery U (2010) The Sla2p/HIP1/HIP1R family: similar structure, similar function in endocytosis? *Biochem Soc Trans* 38: 187-91
- Guinier A (1939) La diffraction des rayons X aux très petits angles: application à l'étude de phénomènes ultramicroscopiques. *Ann Phys (Paris)* 11: 161-237
- Gupta K, Donlan JA, Hopper JT, Uzdavinyas P, Landreh M, Struwe WB, Drew D, Baldwin AJ, Stansfeld PJ, Robinson CV (2017) The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* 541: 421-4
- Hall Z, Hernandez H, Marsh JA, Teichmann SA, Robinson CV (2013) The role of salt bridges, charge density, and subunit flexibility in determining disassembly routes of protein complexes. *Structure* 21: 1325-37
- Han XM, Jin M, Breuker K, McLafferty FW (2006) Extending top-down mass spectrometry to proteins with masses greater than 200 kilodaltons. *Science* 314: 109-12
- Hara N, Morimoto YV, Kawamoto A, Namba K, Minamino T (2012) Interaction of the extreme N-terminal region of FliH with FlhA is required for efficient bacterial flagellar protein export. *J Bacteriol* 194: 5353-60
- Harvey SR, Liu Y, Liu W, Wysocki VH, Laganowsky A (2017) Surface induced dissociation as a tool to study membrane protein complexes. *Chemical Communications* 53: 3106-9
- Haverland NA, Skinner OS, Fellers RT, Tariq AA, Early BP, LeDuc RD, Fornelli L, Compton PD, Kelleher NL (2017) Defining gas-phase fragmentation propensities of intact proteins during native top-down mass spectrometry. *J Am Soc Mass Spectrom* 28: 1203-15
- Heck AJ, Van Den Heuvel RH (2004) Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom Rev* 23: 368-89
- Hedges JB, Vahidi S, Yue X, Konermann L (2013) Effects of ammonium bicarbonate on the electrospray mass spectra of proteins: evidence for bubble-induced unfolding. *Anal Chem* 85: 6469-76
- Heidemann E, Bernhardt H (1967) Deuterium-Hydrogen Exchange of Collagen Like Synthetic Polypeptides. *Nature* 216: 263-4

- Heidemann J, Krichel B, Uetrecht C (2018) Native Massenspektrometrie für die Proteinstrukturanalytik. *BIOspektrum* 24: 164-7
- Hensel M, Shea JE, Gleeson C, Jones MD, Dalton E, Holden DW (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science* 269: 400-3
- Hermans J, Ongay S, Markov V, Bischoff R (2017) Physicochemical Parameters Affecting the Electrospray Ionization Efficiency of Amino Acids after Acylation. *Anal Chem* 89: 9159-66
- Hernandez H, Robinson CV (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc* 2: 715-26
- Hilgemann DW (2007) Local PIP2 signals: when, where, and how? *Pflügers Arch* 455: 55-67
- Hoaglund CS, Valentine SJ, Sporleder CR, Reilly JP, Clemmer DE (1998) Three-dimensional ion mobility/TOFMS analysis of electrosprayed biomolecules. *Anal Chem* 70: 2236-42
- Hofstadler SA, Sannes-Lowery KA (2006) Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes. *Nat Rev Drug Discov* 5: 585-95
- Hogan Jr CJ, Ruotolo BT, Robinson CV, Fernandez de la Mora J (2011) Tandem differential mobility analysis-mass spectrometry reveals partial gas-phase collapse of the GroEL complex. *J Phys Chem B* 115: 3614-21
- Holkar SS, Kamerkar SC, Pucadyil TJ (2015) Spatial Control of Epsin-induced Clathrin Assembly by Membrane Curvature. *J Biol Chem* 290: 14267-76
- Hom RA, Vora M, Regner M, Subach OM, Cho W, Verkhusha VV, Stahelin RV, Kutateladze TG (2007) pH-dependent binding of the Epsin ENTH domain and the AP180 ANTH domain to PI(4,5)P2-containing bilayers. *J Mol Biol* 373: 412-23
- Hopper JT, Yu YT-C, Li D, Raymond A, Bostock M, Liko I, Mikhailov V, Laganowsky A, Benesch JL, Caffrey M (2013) Detergent-free mass spectrometry of membrane protein complexes. *Nat Methods* 10: 1206-8
- Hu B, Lara-Tejero M, Kong Q, Galán JE, Liu J (2017) In situ molecular architecture of the Salmonella type III secretion machine. *Cell* 168: 1065-74. e10
- Hu B, Morado DR, Margolin W, Rohde JR, Arizmendi O, Picking WL, Picking WD, Liu J (2015) Visualization of the type III secretion sorting platform of *Shigella flexneri*. *P Natl Acad Sci USA* 112: 1047-52
- Huang H, Jedynak BM, Bader JS (2007) Where have all the interactions gone? Estimating the coverage of two-hybrid protein interaction maps. *PLoS Comp Biol* 3: e214
- Hulme EC, Trevethick MA (2010) Ligand binding assays at equilibrium: validation and interpretation. *Br J Pharmacol* 161: 1219-37
- Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14: 33-8, 27-8
- Iavarone AT, Parks JH (2005) Conformational change in unsolvated Trp-cage protein probed by fluorescence. *J Am Chem Soc* 127: 8606-7

- Iavarone AT, Williams ER (2003) Mechanism of charging and supercharging molecules in electrospray ionization. *J Am Chem Soc* 125: 2319-27
- Ibrahim YM, Hamid AM, Deng L, Garimella SV, Webb IK, Baker ES, Smith RD (2017) New frontiers for mass spectrometry based upon structures for lossless ion manipulations. *Analyst* 142: 1010-21
- Ibuki T, Imada K, Minamino T, Kato T, Miyata T, Namba K (2011) Common architecture of the flagellar type III protein export apparatus and F- and V-type ATPases. *Nat Struct Mol Biol* 18: 277-82
- Imada K, Minamino T, Uchida Y, Kinoshita M, Namba K (2016) Insight into the flagella type III export revealed by the complex structure of the type III ATPase and its regulator. *P Natl Acad Sci USA* 113: 3633-8
- Ingolfsson HI, Melo MN, van Eerden FJ, Arnarez C, Lopez CA, Wassenaar TA, Periole X, de Vries AH, Tieleman DP, Marrink SJ (2014) Lipid organization of the plasma membrane. *J Am Chem Soc* 136: 14554-9
- Itoh T, Koshiba S, Kigawa T, Kikuchi A, Yokoyama S, Takenawa T (2001) Role of the ENTH domain in phosphatidylinositol-4,5-bisphosphate binding and endocytosis. *Science* 291: 1047-51
- Jecklin MC, Touboul D, Bovet C, Wortmann A, Zenobi R (2008) Which electrospray-based ionization method best reflects protein-ligand interactions found in solution? A comparison of ESI, nanoESI, and ESSI for the determination of dissociation constants with mass spectrometry. *J Am Soc Mass Spectrom* 19: 332-43
- Jones CM, Beardsley RL, Galhena AS, Dagan S, Cheng G, Wysocki VH (2006) Symmetrical gas-phase dissociation of noncovalent protein complexes via surface collisions. *J Am Chem Soc* 128: 15044-5
- Jones EW (1991) Tackling the protease problem in *Saccharomyces cerevisiae*. In *Methods Enzymol*, pp 428-53. Elsevier
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* 79: 926-35
- Jost M, Simpson F, Kavran JM, Lemmon MA, Schmid SL (1998) Phosphatidylinositol-4,5-bisphosphate is required for endocytic coated vesicle formation. *Curr Biol* 8: 1399-402
- Journet L, Agrain C, Broz P, Cornelis GR (2003) The needle length of bacterial injectisomes is determined by a molecular ruler. *Science* 302: 1757-60
- Juraschek R, Dülcks T, Karas M (1999) Nanoelectrospray—more than just a minimized-flow electrospray ionization source. *J Am Soc Mass Spectrom* 10: 300-8
- Jurchen JC, Williams ER (2003) Origin of asymmetric charge partitioning in the dissociation of gas-phase protein homodimers. *J Am Chem Soc* 125: 2817-26
- Jurneczko E, Barran PE (2011) How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *Analyst* 136: 20-8

- Kaksonen M, Roux A (2018) Mechanisms of clathrin-mediated endocytosis. *Nat Rev Mol Cell Biol* 19: 313-26
- Kaksonen M, Toret CP, Drubin DG (2006) Harnessing actin dynamics for clathrin-mediated endocytosis. *Nat Rev Mol Cell Biol* 7: 404-14
- Kaltashov IA, Mohimen A (2005) Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Anal Chem* 77: 5370-9
- Kalthoff C, Alves J, Urbanke C, Knorr R, Ungewickell EJ (2002) Unusual structural organization of the endocytic proteins AP180 and epsin 1. *J Biol Chem* 277: 8209-16
- Karas M, Bachmann D, Bahr U, Hillenkamp F (1987) Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int J Mass Spectrom Ion Process* 78: 53-68
- Karas M, Bahr U, Dulcks T (2000) Nano-electrospray ionization mass spectrometry: addressing analytical problems beyond routine. *Fresenius J Anal Chem* 366: 669-76
- Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60: 2299-301
- Kato J, Lefebvre M, Galan JE (2015) Structural Features Reminiscent of ATP-Driven Protein Translocases Are Essential for the Function of a Type III Secretion-Associated ATPase. *J Bacteriol* 197: 3007-14
- Katta V, Chait BT (1991) Observation of the Heme Globin Complex in Native Myoglobin by Electrospray-Ionization Mass-Spectrometry. *J Am Chem Soc* 113: 8534-5
- Kebarle P, Verkerk UH (2009) Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom Rev* 28: 898-917
- Kitova EN, El-Hawiet A, Schnier PD, Klassen JS (2012) Reliable determinations of protein–ligand interactions by direct ESI-MS measurements. Are we there yet? *J Am Soc Mass Spectrom* 23: 431-41
- Klotz IM (1985) Ligand–receptor interactions: facts and fantasies. *Q Rev Biophys* 18: 227-59
- Kolakowski BM, Mester Z (2007) Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS). *Analyst* 132: 842-64
- Konermann L (2017) Addressing a Common Misconception: Ammonium Acetate as Neutral pH “Buffer” for Native Electrospray Mass Spectrometry. *J Am Soc Mass Spectrom*: 1-9
- Konermann L, Ahadi E, Rodriguez AD, Vahidi S (2013) Unraveling the mechanism of electrospray ionization. *Anal Chem* 85: 2-9
- Konijnenberg A, Yilmaz D, Ingolfsson HI, Dimitrova A, Marrink SJ, Li Z, Venien-Bryan C, Sobott F, Kocer A (2014) Global structural changes of an ion channel during its gating are followed by ion mobility mass spectrometry. *P Natl Acad Sci USA* 111: 17170-5
- Krissinel E (2015) Stock-based detection of protein oligomeric states in jsPISA. *Nucleic Acids Res* 43: W314-9

- Kuhlen L, Abrusci P, Johnson S, Gault J, Deme J, Caesar J, Dietsche T, Mebrhatu MT, Ganief T, Macek B, Wagner S, Robinson CV, Lea SM (2018) Structure of the core of the type III secretion system export apparatus. *Nat Struct Mol Biol* 25: 583-90
- Kukulski W, Schorb M, Kaksonen M, Briggs JAG (2012) Plasma Membrane Reshaping during Endocytosis Is Revealed by Time-Resolved Electron Tomography. *Cell* 150: 508-20
- Laganowsky A, Reading E, Hopper JT, Robinson CV (2013) Mass spectrometry of intact membrane protein complexes. *Nat Protoc* 8: 639-51
- Lara-Tejero M, Kato J, Wagner S, Liu X, Galan JE (2011) A sorting platform determines the order of protein secretion in bacterial type III systems. *Science* 331: 1188-91
- Leavesley SJ, Rich TC (2016) Overcoming limitations of FRET measurements. *Cytometry A* 89: 325-7
- Lemmon MA (2008) Membrane recognition by phospholipid-binding domains. *Nat Rev Mol Cell Biol* 9: 99-111
- Leney AC, Heck AJ (2017) Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectrom* 28: 5-13
- Lev S (2012) Nonvesicular lipid transfer from the endoplasmic reticulum. *Cold Spring Harb Perspect Biol* 4: a013300
- Li HL, Nguyen HH, Loo RRO, Campuzano IDG, Loo JA (2018) An integrated native mass spectrometry and top-down proteomics method that connects sequence to structure and function of macromolecular complexes. *Nat Chem* 10: 139-48
- Li Y, Wang Z, Wei Q, Luo M, Huang G, Sumer BD, Gao J (2016) Non-covalent interactions in controlling pH-responsive behaviors of self-assembled nanosystems. *Polym Chem* 7: 5949-56
- Liko I, Allison TM, Hopper JT, Robinson CV (2016) Mass spectrometry guided structural biology. *Curr Opin Struct Biol* 40: 136-44
- Linse S, Helmersson A, Forsen S (1991) Calcium binding to calmodulin and its globular domains. *J Biol Chem* 266: 8050-4
- Liu AP, Aguet F, Danuser G, Schmid SL (2010) Local clustering of transferrin receptors promotes clathrin-coated pit initiation. *J Cell Biol* 191: 1381-93
- Liu J, Sun Y, Drubin DG, Oster GF (2009) The mechanochemistry of endocytosis. *PLoS Biol* 7: e1000204
- Loo RR, Loo JA (2016) Salt Bridge Rearrangement (SaBRe) Explains the Dissociation Behavior of Noncovalent Complexes. *J Am Soc Mass Spectrom* 27: 975-90
- Loo RRO, O'Brien Johnson R, Nshanian M, Loo JA (2018) Predicting Agents that "Supercharge", ID 295501, TOA 09:30 am. Proceedings of the 66th ASMS Conference on Mass Spectrometry and Allied Topics, San Diego, CA, USA

- Lorenzen K, Versluis C, van Duijn E, van den Heuvel RHH, Heck AJR (2007) Optimizing macromolecular tandem mass spectrometry of large non-covalent complexes using heavy collision gases. *Int J Mass Spectrom* 268: 198-206
- Lu IC, Lee C, Lee YT, Ni CK (2015) Ionization Mechanism of Matrix-Assisted Laser Desorption/Ionization. *Annu Rev Anal Chem (Palo Alto Calif)* 8: 21-39
- Lu R, Drubin DG, Sun YD (2016) Clathrin-mediated endocytosis in budding yeast at a glance. *J Cell Sci* 129: 1531-6
- Ma L, Umasankar PK, Wrobel AG, Lyman A, McCoy AJ, Holkar SS, Jha A, Pradhan-Sundt T, Watkins SC, Owen DJ (2016) Transient Fcho1/2· Eps15/R· AP-2 nanoclusters prime the AP-2 clathrin adaptor for cargo binding. *Dev Cell* 37: 428-43
- Mack Jr E (1925) Average cross-sectional areas of molecules by gaseous diffusion methods. *J Am Chem Soc* 47: 2468-82
- MacKerell Jr AD, Bashford D, Bellott M, Dunbrack Jr RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102: 3586-616
- Makino F, Shen D, Kajimura N, Kawamoto A, Pissaridou P, Oswin H, Pain M, Murillo I, Namba K, Blocker AJ (2016) The Architecture of the Cytoplasmic Region of Type III Secretion Systems. *Sci Rep* 6: 33341
- Maloy SR, Nunn WD (1981) Selection for loss of tetracycline resistance by *Escherichia coli*. *J Bacteriol* 145: 1110-1
- Mamyrin B, Karataev V, Shmikk D, Zagulin V (1973) The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Zh Eksp Teor Fiz* 64: 82-9
- Marklund EG, Degiacomi MT, Robinson CV, Baldwin AJ, Benesch JL (2015) Collision cross sections for structural proteomics. *Structure* 23: 791-9
- Marlovits TC, Kubori T, Lara-Tejero M, Thomas D, Unger VM, Galan JE (2006) Assembly of the inner rod determines needle length in the type III secretion injectisome. *Nature* 441: 637-40
- Mason EA, Schamp Jr HW (1958) Mobility of gaseous ions in weak electric fields. *Ann Phys* 4: 233-70
- Maxfield FR, McGraw TE (2004) Endocytic recycling. *Nat Rev Mol Cell Biol* 5: 121-32
- McDowell MA, Marcoux J, McVicker G, Johnson S, Fong YH, Stevens R, Bowman LA, Degiacomi MT, Yan J, Wise A, Friede ME, Benesch JL, Deane JE, Tang CM, Robinson CV, Lea SM (2016) Characterisation of *Shigella* Spa33 and *Thermotoga* FliM/N reveals a new model for C-ring assembly in T3SS. *Mol Microbiol* 99: 749-66
- McPherson PS, Garcia EP, Slepnev VI, David C, Zhang X, Grabs D, Sossin WS, Bauerfeind R, Nemoto Y, De Camilli P (1996) A presynaptic inositol-5-phosphatase. *Nature* 379: 353-7
- Mehmood S, Allison TM, Robinson CV (2015) Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annu Rev Phys Chem* 66: 453-74

- Mendoza VL, Vachet RW (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom Rev* 28: 785-815
- Mesleh MF, Hunter JM, Shvartsburg AA, Schatz GC, Jarrold MF (1996) Structural information from ion mobility measurements: Effects of the long-range potential. *J Phys Chem-U* 100: 16082-6
- Messa M, Fernandez-Busnadiego R, Sun EW, Chen H, Czapla H, Wrasman K, Wu Y, Ko G, Ross T, Wendland B, De Camilli P (2014) Epsin deficiency impairs endocytosis by stalling the actin-dependent invagination of endocytic clathrin-coated pits. *Elife* 3: e03311
- Meyer B, Peters T (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Edit* 42: 864-90
- Miller SE, Mathiasen S, Bright NA, Pierre F, Kelly BT, Kladt N, Schauss A, Merrifield CJ, Stamou D, Honing S, Owen DJ (2015) CALM regulates clathrin-coated vesicle size and maturation by directly sensing and driving membrane curvature. *Dev Cell* 33: 163-75
- Minamino T, Namba K (2008) Distinct roles of the FliI ATPase and proton motive force in bacterial flagellar protein export. *Nature* 451: 485-8
- Morgner N, Robinson CV (2012) Massign: an assignment strategy for maximizing information from the mass spectra of heterogeneous protein assemblies. *Anal Chem* 84: 2939-48
- Morita-Ishihara T, Ogawa M, Sagara H, Yoshida M, Katayama E, Sasakawa C (2006) Shigella Spa33 is an essential C-ring component of type III secretion machinery. *J Biol Chem* 281: 599-607
- Mortensen DN, Susa AC, Williams ER (2017) Collisional Cross-Sections with T-Wave Ion Mobility Spectrometry without Experimental Calibration. *J Am Soc Mass Spectrom*: 1-11
- Mortensen DN, Williams ER (2014) Theta-glass capillaries in electrospray ionization: rapid mixing and short droplet lifetimes. *Anal Chem* 86: 9315-21
- Mukhopadhyay D, Riezman H (2007) Proteasome-independent functions of ubiquitin in endocytosis and signaling. *Science* 315: 201-5
- Murray KK, Boyd RK, Eberlin MN, Langley GJ, Li L, Naito Y (2013) Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013). *Pure Appl Chem* 85: 1515-609
- Nitsche J, Josts I, Heidemann J, Mertens H, Maric S, Moulin M, Haertlein M, Busch S, Forsyth V, Svergun D, Uetrecht C, Tidow H (*in revision*) Structural basis for activation of plasma-membrane Ca<sup>2+</sup>-ATPase by calmodulin. *Commun Biol*
- Notti RQ, Bhattacharya S, Lilic M, Stebbins CE (2015) A common assembly module in injectisome and flagellar type III secretion sorting platforms. *Nat Commun* 6: 7125
- Ohgita T, Hayashi N, Hama S, Tsuchiya H, Gotoh N, Kogure K (2013) A novel effector secretion mechanism based on proton-motive force-dependent type III secretion apparatus rotation. *FASEB J* 27: 2862-72
- Oomens J, Polfer N, Moore DT, van der Meer L, Marshall AG, Eyler JR, Meijer G, von Helden G (2005) Charge-state resolved mid-infrared spectroscopy of a gas-phase protein. *Phys Chem Chem Phys* 7: 1345-8

- Pagel K, Hyung SJ, Ruotolo BT, Robinson CV (2010) Alternate dissociation pathways identified in charge-reduced protein complex ions. *Anal Chem* 82: 5363-72
- Patching SG (2014) Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. *Biochim Biophys Acta* 1838: 43-55
- Paul W (1990) Electromagnetic traps for charged and neutral particles (Nobel lecture). *Angew Chem Int Ed* 29: 739-48
- Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HD, Konarev PV, Svergun DI (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* 45: 342-50
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605-12
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26: 1781-802
- Piston DW, Kremers GJ (2007) Fluorescent protein FRET: the good, the bad and the ugly. *Trends Biochem Sci* 32: 407-14
- Popa V, Trecroce DA, McAllister RG, Konermann L (2016) Collision-Induced Dissociation of Electrosprayed Protein Complexes: An All-Atom Molecular Dynamics Model with Mobile Protons. *J Phys Chem B* 120: 5114-24
- Posor Y, Eichhorn-Gruenig M, Puchkov D, Schoneberg J, Ullrich A, Lampe A, Muller R, Zarbakhsh S, Gulluni F, Hirsch E, Krauss M, Schultz C, Schmoranzler J, Noe F, Haucke V (2013) Spatiotemporal control of endocytosis by phosphatidylinositol-3,4-bisphosphate. *Nature* 499: 233-7
- Pringle SD, Giles K, Wildgoose JL, Williams JP, Slade SE, Thalassinou K, Bateman RH, Bowers MT, Scrivens JH (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int J Mass Spectrom* 261: 1-12
- Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* 40: 191-285
- Radics J, Konigsmair L, Marlovits TC (2014) Structure of a pathogenic type 3 secretion system in action. *Nat Struct Mol Biol* 21: 82-7
- Rayleigh L (1882) XX. On the equilibrium of liquid conducting masses charged with electricity. *The London, Edinburgh and Dublin philosophical magazine and journal of science* 14: 184-6
- Roepstorff P, Fohlman J (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* 11: 601
- Romo-Castillo M, Andrade A, Espinosa N, Feria JM, Soto E, Diaz-Guerrero M, Gonzalez-Pedrajo B (2014) EscO, a Functional and Structural Analog of the Flagellar FliJ Protein, Is a Positive Regulator of EscN ATPase Activity of the Enteropathogenic *Escherichia coli* Injectisome. *J Bacteriol* 196: 2227-41

- Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol* 5: 172
- Rose RJ, Damoc E, Denisov E, Makarov A, Heck AJ (2012) High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat Methods* 9: 1084-6
- Rose RJ, Labrijn AF, van den Bremer ET, Loverix S, Lasters I, van Berkel PH, van de Winkel JG, Schuurman J, Parren PW, Heck AJ (2011) Quantitative analysis of the interaction strength and dynamics of human IgG4 half molecules by native mass spectrometry. *Structure* 19: 1274-82
- Ruotolo BT, Benesch JL, Sandercock AM, Hyung SJ, Robinson CV (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nat Protoc* 3: 1139-52
- Ruotolo BT, Hyung SJ, Robinson PM, Giles K, Bateman RH, Robinson CV (2007) Ion Mobility–Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes. *Angew Chem Int Ed* 46: 8001-4
- Saleem M, Morlot S, Hohendahl A, Manzi J, Lenz M, Roux A (2015) A balance between membrane elasticity and polymerization energy sets the shape of spherical clathrin coats. *Nat Commun* 6: 6249
- Scarff CA, Thalassinos K, Hilton GR, Scrivens JH (2008) Travelling wave ion mobility mass spectrometry studies of protein structure: biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements. *Rapid Commun Mass Spectrom* 22: 3297-304
- Schlossman DM, Schmid SL, Braell WA, Rothman JE (1984) An Enzyme That Removes Clathrin Coats - Purification of an Uncoating Atpase. *J Cell Biol* 99: 723-33
- Scigelova M, Hornshaw M, Giannakopoulos A, Makarov A (2011) Fourier transform mass spectrometry. *Mol Cell Proteomics* 10: M111 009431
- Semenyuk A, Svergun D (1991) GNOM—a program package for small-angle scattering data processing. *J Appl Crystallogr* 24: 537-40
- Sen A, Madhivanan K, Mukherjee D, Aguilar RC (2012) The epsin protein family: coordinators of endocytosis and signaling. *Biomol Concepts* 3: 117-26
- Sharon M (2010) How far can we go with structural mass spectrometry of protein complexes? *J Am Soc Mass Spectrom* 21: 487-500
- Shen HH, Lithgow T, Martin L (2013) Reconstitution of membrane proteins into model membranes: seeking better ways to retain protein activities. *Int J Mol Sci* 14: 1589-607
- Shiratori T, Miyatake S, Ohno H, Nakaseko C, Isono K, Bonifacino JS, Saito T (1997) Tyrosine phosphorylation controls internalization of CTLA-4 by regulating its interaction with clathrin-associated adaptor complex AP-2. *Immunity* 6: 583-9
- Shvartsburg AA, Smith RD (2008) Fundamentals of traveling wave ion mobility spectrometry. *Anal Chem* 80: 9689-99
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7: 539

- Silveira JA, Fort KL, Kim D, Servage KA, Pierson NA, Clemmer DE, Russell DH (2013) From solution to the gas phase: stepwise dehydration and kinetic trapping of substance P reveals the origin of peptide conformations. *J Am Chem Soc* 135: 19147-53
- Sinz A (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom Rev* 25: 663-82
- Skinner OS, Catherman AD, Early BP, Thomas PM, Compton PD, Kelleher NL (2014) Fragmentation of integral membrane proteins in the gas phase. *Anal Chem* 86: 4627-34
- Skruzny M, Brach T, Ciuffa R, Rybina S, Wachsmuth M, Kaksonen M (2012) Molecular basis for coupling the plasma membrane to the actin cytoskeleton during clathrin-mediated endocytosis. *P Natl Acad Sci USA* 109: E2533-42
- Skruzny M, Desfosses A, Prinz S, Dodonova Svetlana O, Gieras A, Uetrecht C, Jakobi Arjen J, Abella M, Hagen Wim JH, Schulz J, Meijers R, Rybin V, Briggs John AG, Sachse C, Kaksonen M (2015) An Organized Co-assembly of Clathrin Adaptors Is Essential for Endocytosis. *Dev Cell* 33: 150-62
- Smith RD, Loo JA, Barinaga CJ, Edmonds CG, Udseth HR (1990) Collisional activation and collision-activated dissociation of large multiply charged polypeptides and proteins produced by electrospray ionization. *J Am Soc Mass Spectrom* 1: 53-65
- Sobott F, Hernandez H, McCammon MG, Tito MA, Robinson CV (2002) A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal Chem* 74: 1402-7
- Song M, Sukovich DJ, Ciccarelli L, Mayr J, Fernandez-Rodriguez J, Mirsky EA, Tucker AC, Gordon DB, Marlovits TC, Voigt CA (2017) Control of type III protein secretion using a minimal genetic system. *Nat Commun* 8: 14737
- Spradling KD, McDaniel AE, Lohi J, Pilcher BK (2001) Epsin 3 is a novel extracellular matrix-induced transcript specific to wounded epithelia. *J Biol Chem* 276: 29257-67
- Stahelin RV, Long F, Peter BJ, Murray D, De Camilli P, McMahon HT, Cho W (2003) Contrasting membrane interaction mechanisms of AP180 N-terminal homology (ANTH) and epsin N-terminal homology (ENTH) domains. *J Biol Chem* 278: 28993-9
- Stamm LM, Goldberg MB (2011) Microbiology. Establishing the secretion hierarchy. *Science* 331: 1147-8
- Stebbins CE, Galán JE (2001) Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion. *Nature* 414: 77-81
- Stetefeld J, McKenna SA, Patel TR (2016) Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys Rev* 8: 409-27
- Structural Genomics C, China Structural Genomics C, Northeast Structural Genomics C, Graslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, Gileadi O, Knapp S, Oppermann U, Arrowsmith C, Hui R, Ming J, dhe-Paganon S, Park HW, Savchenko A, Yee A, Edwards A, Vincentelli R et al. (2008) Protein production and purification. *Nat Methods* 5: 135-46

- Sun J, Kitova EN, Wang W, Klassen JS (2006) Method for distinguishing specific from nonspecific protein-ligand complexes in nanoelectrospray ionization mass spectrometry. *Anal Chem* 78: 3010-8
- Susa AC, Xia Z, Williams ER (2017) Small Emitter Tips for Native Mass Spectrometry of Proteins and Protein Complexes from Nonvolatile Buffers That Mimic the Intracellular Environment. *Anal Chem* 89: 3116-22
- Svergun D (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25: 495-503
- Svergun D, Barberato C, Koch MHJ (1995) CRY SOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* 28: 768-73
- Szabo A, Stolz L, Granzow R (1995) Surface-Plasmon Resonance and Its Use in Biomolecular Interaction Analysis (Bia). *Curr Opin Struct Biol* 5: 699-705
- Tahallah N, Pinkse M, Maier CS, Heck AJ (2001) The effect of the source pressure on the abundance of ions of noncovalent protein assemblies in an electrospray ionization orthogonal time-of-flight instrument. *Rapid Commun Mass Spectrom* 15: 596-601
- Takatori S, Tomita T (2018) AP180 N-Terminal Homology (ANTH) and Epsin N-Terminal Homology (ENTH) Domains: Physiological Functions and Involvement in Disease. *Adv Exp Med Biol*
- 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 Commun Mass Spectrom* 2: 151-3
- Tidow H, Poulsen LR, Andreeva A, Knudsen M, Hein KL, Wiuf C, Palmgren MG, Nissen P (2012) A bimodular mechanism of calcium control in eukaryotes. *Nature* 491: 468-72
- Tolmachev AV, Robinson EW, Wu S, Pasa-Tolic L, Smith RD (2009) FT-ICR MS optimization for the analysis of intact proteins. *Int J Mass Spectrom* 281: 32-8
- Torres FE, Recht MI, Coyle JE, Bruce RH, Williams G (2010) Higher throughput calorimetry: opportunities, approaches and challenges. *Curr Opin Struct Biol* 20: 598-605
- Traub LM (2009) Tickets to ride: selecting cargo for clathrin-regulated internalization. *Nat Rev Mol Cell Biol* 10: 583-96
- Tuukkanen AT, Kleywegt GJ, Svergun DI (2016) Resolution of ab initio shapes determined from small-angle scattering. *IUCrJ* 3: 440-7
- Utrecht C, Rose RJ, van Duijn E, Lorenzen K, Heck AJ (2010) Ion mobility mass spectrometry of proteins and protein assemblies. *Chem Soc Rev* 39: 1633-55
- Utrecht C, Versluis C, Watts NR, Roos WH, Wuite GJ, Wingfield PT, Steven AC, Heck AJ (2008) High-resolution mass spectrometry of viral assemblies: molecular composition and stability of dimorphic hepatitis B virus capsids. *P Natl Acad Sci USA* 105: 9216-20
- Van Berkel GJ, Kertesz V (2007) Using the electrochemistry of the electrospray ion source. *Anal Chem* 79: 5510-20

van de Waterbeemd M, Fort KL, Boll D, Reinhardt-Szyba M, Routh A, Makarov A, Heck AJ (2017) High-fidelity mass analysis unveils heterogeneity in intact ribosomal particles. *Nat Methods* 14: 283-6

van den Bogaart G, Meyenberg K, Risselada HJ, Amin H, Willig KI, Hubrich BE, Dier M, Hell SW, Grubmüller H, Diederichsen U (2011) Membrane protein sequestering by ionic protein–lipid interactions. *Nature* 479: 552

van den Heuvel RH, Heck AJ (2004) Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr Opin Chem Biol* 8: 519-26

van den Heuvel RH, van Duijn E, Mazon H, Synowsky SA, Lorenzen K, Versluis C, Brouns SJ, Langridge D, van der Oost J, Hoyes J, Heck AJ (2006) Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal Chem* 78: 7473-83

Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr* 36: 860-4

Wallroth A, Hauke V (2018) Phosphoinositide conversion in endocytosis and the endolysosomal system. *J Biol Chem* 293: 1526-35

Waugh DS (2011) An overview of enzymatic reagents for the removal of affinity tags. *Protein Expr Purif* 80: 283-93

Wienken CJ, Baaske P, Rothbauer U, Braun D, Duhr S (2010) Protein-binding assays in biological liquids using microscale thermophoresis. *Nat Commun* 1: 100

Wilbur JD, Chen CY, Manalo V, Hwang PK, Fletterick RJ, Brodsky FM (2008) Actin Binding by Hip1 (Huntingtin-interacting Protein 1) and Hip1R (Hip1-related Protein) Is Regulated by Clathrin Light Chain. *J Biol Chem* 283: 32870-9

Wilm M, Mann M (1996) Analytical properties of the nanoelectrospray ion source. *Anal Chem* 68: 1-8

Wilm MS, Mann M (1994) Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *Int J Mass Spectrom Ion Process* 136: 167-80

Winger BE, Light-Wahl KJ, Ogorzalek Loo RR, Udseth HR, Smith RD (1993) Observation and implications of high mass-to-charge ratio ions from electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom* 4: 536-45

Wyatt PJ (1993) Light-Scattering and the Absolute Characterization of Macromolecules. *Anal Chim Acta* 272: 1-40

Wysocki VH, Jones CM, Galhena AS, Blackwell AE (2008) Surface-induced dissociation shows potential to be more informative than collision-induced dissociation for structural studies of large systems. *J Am Soc Mass Spectrom* 19: 903-13

Wytttenbach T, Bowers MT (2011) Structural stability from solution to the gas phase: native solution structure of ubiquitin survives analysis in a solvent-free ion mobility–mass spectrometry environment. *J Phys Chem B* 115: 12266-75

- Yamashita M, Fenn JB (1984) Electrospray Ion-Source - Another Variation on the Free-Jet Theme. *J Phys Chem-Us* 88: 4451-9
- Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 12: 7-8
- Yang S, Cope MJ, Drubin DG (1999) Sla2p is associated with the yeast cortical actin cytoskeleton via redundant localization signals. *Mol Biol Cell* 10: 2265-83
- Yoon Y, Tong J, Lee PJ, Albanese A, Bhardwaj N, Kallberg M, Digman MA, Lu H, Gratton E, Shin YK, Cho W (2010) Molecular basis of the potent membrane-remodeling activity of the epsin 1 N-terminal homology domain. *J Biol Chem* 285: 531-40
- Yu X-J, Liu M, Matthews S, Holden DW (2011) Tandem translation generates a chaperone for the Salmonella type III secretion system protein SsaQ. *J Biol Chem* 286: 36098-107
- Yu Z, Frangakis AS (2014) M-free: scoring the reference bias in sub-tomogram averaging and template matching. *J Struct Biol* 187: 10-9
- Zhang Z, Smith DL (1993) Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci* 2: 522-31
- Zhou M, Dagan S, Wysocki VH (2012) Protein subunits released by surface collisions of noncovalent complexes: natively compact structures revealed by ion mobility mass spectrometry. *Angew Chem Int Edit* 51: 4336-9
- Zhou M, Jones CM, Wysocki VH (2013) Dissecting the large noncovalent protein complex GroEL with surface-induced dissociation and ion mobility-mass spectrometry. *Anal Chem* 85: 8262-7
- Zhou M, Wysocki VH (2014) Surface induced dissociation: dissecting noncovalent protein complexes in the gas phase. *Acc Chem Res* 47: 1010-8
- Zinck N, Stark AK, Wilson DJ, Sharon M (2014) An improved rapid mixing device for time-resolved electrospray mass spectrometry measurements. *ChemistryOpen* 3: 109-14

---

## S10 Acknowledgement

Throughout the preparation of this thesis, I received input of many people and I would like to use this opportunity to express my gratitude.

First and foremost, I would like to thank my supervisor Charlotte Uetrecht for the supervision of this thesis. I am especially thankful for the arrangement of all the exciting projects and the detailed discussions of the results. I truly appreciated the combination of latitude on one hand and promotion and support on the other.

I would also like to thank Wolfram Brune for the co-supervision of the thesis.

In the course of this highly collaborative thesis, I worked with many different people who made valuable contributions, mainly subject-specifically but sometimes also personally. For instance, it was a pleasure to work with Maria Garcia-Alai, who provided me with high-quality proteins and taught me straight pragmatism – in science and other situations. Together also with Rob Meijers it was a pleasure to work on the project.

ACA8 proved us, how hard and sometimes frustrating it is to work with integral membrane proteins. Nevertheless, I enjoyed working on this project, thanks to Henning Tidow and Julius Nitsche who always created a pleasant, motivating and optimistic atmosphere.

I would further like to thank Michael Kolbe, Ivonne Bernal and Jonathan Börnicke for the productive collaboration. Although – or rather because we had some challenging situations during data interpretation, I learned a lot during this project.

My stays in Frank Sobott's lab in Antwerp certainly were a highlight during my thesis. Therefore, I would like to thank Frank for hosting me in his group and for making my trips equally educating and entertaining. A special thanks goes to Albert Konijnenberg and Jeroen van Dyck, who were always extremely welcoming and supportive during my visits. Unfortunately, the Synapt was repeatedly misbehaving and the samples that I brought were quite challenging. But I am very thankful for all the troubleshooting together with Albert that taught me a lot about the basics of mass spectrometry.

In this regard, I would also like to thank Jerre van der Horst and Jan Commandeur from MS Vision for all their educative explanations in the lab and during calzone lunch sessions during their support stays.

The pleasure in the daily work is heavily dependent on your colleagues. Therefore, I am very grateful to my whole working group at the Heinrich-Pette-Institute. Four years in one office and one lab with my 'cousin' Boris had a huge impact on me, and I am equally thankful for the scientific discussions and the fun that we had. I also want to thank Julia for creating such a nice working atmosphere in our office and I wish her all the best for the future. Many thanks also to the remaining group members: Panda, Jasmin, Lara, Janine, Ronja, Alan, Knut, Yinfei, Jürgen and Marc. It was a pleasure working in this group.

Moreover, I would like to thank Sarah, Annika, Anja, Ingo, Lisi and Vicky for spending nice coffee breaks with me, for enjoyable and fun conversations and some friendships that developed in these years.

Last but not least, I want to thank my family, especially my parents, for all the support during my studies and in the past four years. The certitude that I can always rely on them, is priceless. Along these lines, I also want to thank Lale for suddenly and lastingly coming into my life, for always supporting me and giving me confidence.

---

## **S11 Declaration of Authorship**

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids. The submitted written document corresponds to the file on the storage medium. I further declare that this thesis has not been presented previously to another examination board.

### **“Eidesstattliche Versicherung”**

Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.

29 August 2018, Johannes Heidemann