In vitro Reconstitution and Characterization of Soluble Complexes of the *Salmonella* Type III Secretion System Sorting Platform

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

Department of Chemistry
University of Hamburg

Submitted by

Carl Jonathan Börnicke

born in Berlin

Hamburg, 2020

This dissertation was prepared from September 2013 to December 2019 in the group of Prof. Dr. Michael Kolbe at the Max Planck Institute for Infection Biology, Berlin, and the Centre for Structural Systems Biology, Hamburg.

Reviewers

1. **Prof. Dr. Michael Kolbe**

   Institute for Biochemistry and Molecular Biology, Department of Chemistry, University of Hamburg, Germany

2. **PD Dr. habil. Markus Perbandt**

   Institute for Biochemistry and Molecular Biology, Department of Chemistry, University of Hamburg, Germany

Examination Commission

1.	Chair:	Prof. Dr. Michael Kolbe
2.	Vice Chair:	Prof. Dr. Tim Gilberger
3.	Member:	Dr. Jan Kosinski

Date of oral defense and approval for publication: 08.05.2020

Part of this work was published in the peer-reviewed Journal of Molecular Biology:

Bernal I, Börnicke J, Heidemann J, Svergun D, Horstmann JA, Erhardt M, Tuukkanen A, Uetrecht C, Kolbe M. 2019. “Molecular Organization of Soluble Type III Secretion System Sorting Platform Complexes.” *Journal of Molecular Biology* 431 (19): 3787–3803.

Table of Contents

Summary

During an infection, many pathogenic Gram-negative bacteria manipulate host cells by injecting them with virulence effectors using transmembrane multiprotein complexes such as type III secretion systems (T3SS). These sophisticated nanomachines are complex structures that consist of a basal body embedded in the bacterial envelope and an extracellular needle. This needle inserts into host cell membranes to form a continuous channel for the delivery of proteins into the host cell cytoplasm. The assembly and function of the T3SS are critically dependent on the hierarchical secretion of structural and effector proteins, and it has been proposed that a multi-protein complex attached to the cytosolic side of the T3SS acts as a sorting platform by selecting the appropriate substrates for secretion through the system. However, although this complex plays an essential role in type III secretion, its precise molecular organization and mechanism of function are still incompletely understood.

This work aimed to analyze the *Salmonella* Typhimurium SPI-1 T3SS sorting platform by *in vitro* reconstitution and to uncover its molecular architecture and mechanism of assembly in solution. Co-expression of sorting platform components yielded soluble complexes containing the protein SpaO, its shorter isoform $SpaO_C$, the T3SS ATPase InvC and its regulator OrgB. These complexes were analyzed by size-exclusion chromatography and multi-angle light scattering to obtain mass and shape information, and by native mass spectrometry to determine the precise stoichiometry of subunits within each complex. In addition, structural analysis by small-angle X-ray scattering revealed that the largest of the isolated complexes, SpaO/$SpaO_C$/OrgB/InvC, adopts an elongated L-shaped configuration in solution. Together with topology information from tandem mass spectrometry and protein domain pull-down assays, these data could be combined into a model of the architecture of the complex. Importantly, this model is in good agreement with the sorting platform pod densities observed in a previous *in situ* cryo-electron tomography structure, suggesting that the SpaO/$SpaO_C$/OrgB/InvC complex identified in this study is the soluble core building block from which the complete sorting platform may be assembled. Together, these findings grant insights into the formation and architecture of the T3SS sorting platform, and the purified soluble complexes provide a starting point for *in vitro* interaction studies between sorting platform complexes and T3SS substrates.

Zusammenfassung

Viele Gram-negative Bakterien manipulieren Wirtszellen während einer Infektion, indem sie ihnen Virulenzeffektoren mit Hilfe von Transmembrankomplexen wie dem Typ-III-Sekretionssystem (T3SS) injizieren. Diese raffinierten Nanomaschinen sind komplexe Strukturen, die aus einem in die Bakterienmembranen integrierten Basalkörper und einer extrazellulären Nadel bestehen. Diese Nadel durchdringt die Wirtszellenmembran und bildet einen durchgängigen Kanal für den Transport von Proteinen in das Zytoplasma der Wirtszelle. Für die Konstruktion und die Funktion des T3SS ist es zwingend erforderlich, dass Struktur- und Effektorproteine in der richtigen Reihenfolge sekretiert werden und frühere Studien deuten darauf hin, dass ein an die zytosolische Seite des T3SS gebundener Komplex aus mehreren Proteinen als eine „Sortierplattform" fungiert und die korrekten Substrate für die Sekretion durch das System auswählt. Trotz der essenziellen Rolle dieses Komplexes im Typ-III-Sekretionsprozess sind seine genaue molekulare Organisation und sein Funktionsmechanismus noch immer nicht vollständig erforscht.

Diese Arbeit hatte zum Ziel, die Sortierplattform des *Salmonella* Typhimurium SPI-1 T3SS *in vitro* zu rekonstituieren und ihre molekulare Architektur und ihren Aufbaumechanismus zu ergründen. Durch die Koexpression von Sortierplattformkomponenten konnten lösliche Komplexe aus SpaO, seiner kurzen Isoform $SpaO_C$, der T3SS-ATPase InvC und ihrem Regulatorprotein OrgB aufgereinigt werden. Diese Komplexe wurden mit Hilfe von Größenausschlusschromatographie und Mehrwinkel-Lichtstreuung hinsichtlich ihrer Masse und Form untersucht und native Massenspektrometrie wurde genutzt, um die genaue Stöchiometrie der Untereinheiten in den einzelnen Komplexen zu analysieren. Darüber hinaus zeigte die Strukturanalyse mittels Kleinwinkel-Röntgenstreuung, dass der größte der isolierbaren Komplexe, SpaO/$SpaO_C$/OrgB/InvC, eine ausgestreckte L-förmige Konformation besitzt. Tandem-Massenspektrometrie und Pulldown-Experimente mit Proteindomänen lieferten zusätzliche topologische Informationen, wodurch ein Modell der Architektur des Komplexes erstellt werden konnte. Dieses Modell zeigt hohe Übereinstimmung mit den durch Kryoelektronentomographie ermittelten Säulenstrukturen der Sortierplatform, was darauf hindeutet, dass der hier identifizierte SpaO/$SpaO_C$/OrgB/InvC-Komplex den löslichen Baustein darstellt, aus welchem sich die vollständige Sortierplatform zusammensetzt. Zusammengefasst liefern diese Ergebnisse Einsicht in den Aufbauprozess und die Architektur der T3SS-

Sortierplatform und die aufgereinigten löslichen Komplexe dienten als Ausgangspunkt für *in-vitro*-Interaktionsstudien zwischen Sortierplatformkomplexen und T3SS-Substraten.

1. Introduction

1.1. Type III secretion as a central mechanism of *Salmonella* pathogenicity

The secretion of proteins is one of the most important ways for prokaryotes to interact with their environment and at least six different mechanisms are utilized by Gram-negative bacteria to transport pathogenicity factors across their inner and outer membranes into the extracellular space (Green & Mecsas, 2016). Some of these systems further enable the perforation of a third membrane and a variety of species use a type III secretion system (T3SS) to interact with eukaryotic hosts cells by directly injecting proteins into their cytoplasm. While such trans-kingdom interactions can in some cases be mutually beneficial, like in the symbiotic relationships between *Rhizobia* and legumes, T3SS-using bacteria are more often associated with diseases of both plants and animals (Hueck, 1998). Importantly, the T3SS represent an essential virulence determinant in a variety of human pathogens like *Yersinia, Shigella, Salmonella, Chlamydia, Pseudomonas* and *E. coli*, which cause diseases such as plague, enterocolitis, typhoid fever, sexually transmitted infections or pneumonia (Coburn et al., 2007).

An important group of T3SS-utilizing bacteria is the *Salmonella enterica* subspecies *enterica*, which contains over 1500 serovars that are distinguishable by their antigenic profile and are responsible for a large number of infections in both humans and livestock animals (Crump & Wain, 2017). In humans, the serovars *S.* Typhi and Paratyphi are the causative agents of enteric (typhoid) fever with an estimated 14 million cases and over 135,000 deaths annually (Roth et al., 2018; James et al., 2018). While these two serovars are specialized pathogens that are restricted to humans, others like *S.* Tyhimurium can infect a broader host range that also includes cattle and poultry, and cause an estimated 93 million cases of human gastroenteritis each year (Majowicz et al., 2010). Because of their threat to global health and increasing multi-drug resistance, *Salmonella* species have been classified by the World Health Organization as high-priority pathogens for which the research and development of novel antimicrobial treatment options are urgently needed (Tacconelli et al., 2018).

S. Typhimurium has been studied extensively as a model of *Salmonella* infection and the function of type III secretion systems due to its ability to induce a typhoid-like disease in mice (Santos et al., 2001). Unless otherwise indicated, "*Salmonella*" in this work will refer to *S.* Typhimurium. During infection, *S.* Typhimurium uses two type III secretion systems that were named SPI-1 and SPI-2 after the genomic *Salmonella* pathogenicity islands they are encoded in. Of these systems, SPI-1 is mostly active during the initial invasion phase, while SPI-2 plays an important role in survival and persistence within the host (Haraga et al., 2008; Larock et al., 2015). The major cause of *Salmonella* infections is the ingestion of contaminated water or animal products. Upon reaching the intestine, bacteria mainly cross the otherwise impenetrable epithelial layer through M cells (Sansonetti & Phalipon, 1999; Jones et al., 1994), specialized cells that take up antigens from the gut lumen to pass them on to sub-epithelial immune cells (Figure 1.1). At the basolateral side of the intestinal mucosa, *Salmonella* are captured by phagocytes like macrophages and dendritic cells. Using their SPI-1 T3SS, bacteria either escape from these cells by inducing their cell death and causing inflammation (Monack et al., 2000; van der Velden et al., 2003), or evade destruction by preventing the recruitment of hydrolytic lysosomal enzymes

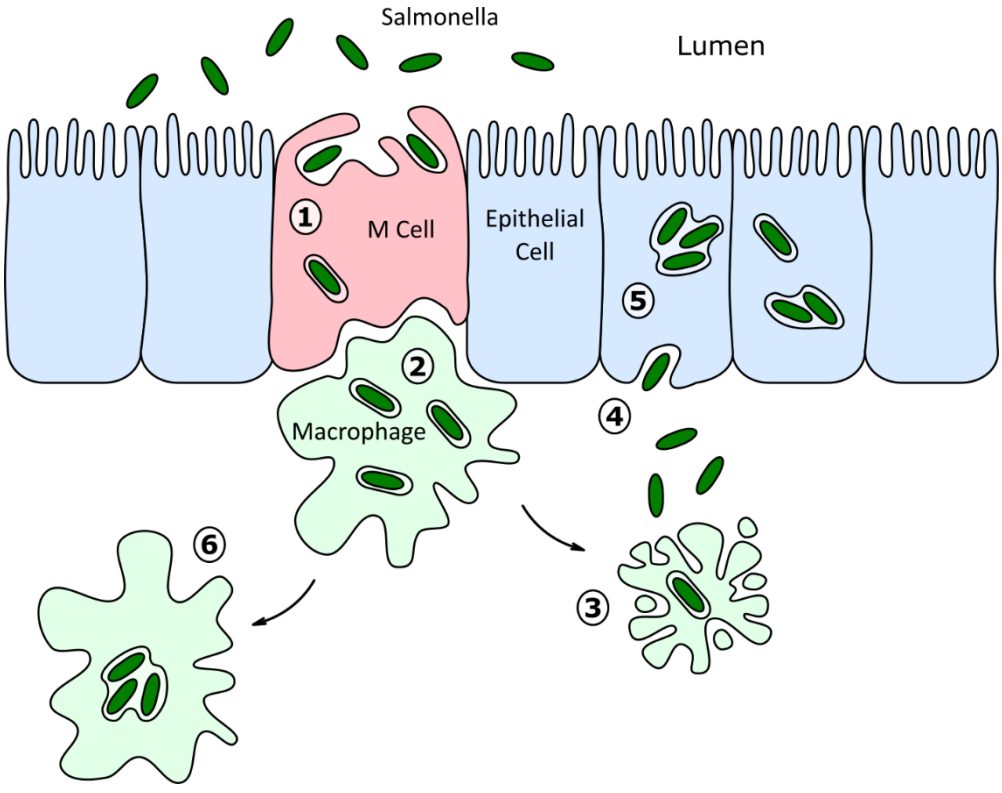


Figure 1.1. *Salmonella* infection of the intestine. *Salmonella* in the gut lumen are transported across the epithelial layer by M cells (1) and taken up by tissue-resident macrophages (2). Inducing their cell death, they escape destruction by macrophages (3) and infect epithelial cells from the baso-lateral side (4). Inside host cells, *Salmonella* modulate the phagosome to transform it into the *Salmonella*-containing vacuole for their intracellular survival and replication (5). Stably infected phagocytotic cells that leave the site of primary infection can cause systemic dissemination of *Salmonella* (6).

to the phagosome (Ishibashi & Arai, 1990; Buchmeier & Heffron, 1991; Tuli & Sharma, 2019). In the case of human infections with *S*. Typhi and murine infections with *S*. Typhimurium, such stably infected phagocytes may leave the site of initial infection and cause systemic dissemination of bacteria and the colonization of other organs (Carter & Collins, 1974; Vazquez-Torres et al., 1999). Following their escape from phagocytes, *Salmonella* can invade epithelial cells from the basolateral side by injecting them with SPI-1 effectors that result in the uptake of bacteria into these otherwise non-phagocytic cells by macropinocytosis (Criss & Casanova, 2003; Jones et al., 1994; Hume et al., 2017). After successful internalization, the SPI-2 T3SS begins to secrete proteins across the vacuolar membrane and SPI-1 and SPI-2 effectors together cause the transformation of the phagosome into the "*Salmonella*-containing vacuole" (SCV), a distinct compartment that is permissive to *Salmonella* replication (Steele-Mortimer, 2008). Thus, by using two type III secretion systems, *Salmonella* establishes itself a niche for survival within epithelial and phagocytic cells.

1.2. Structure and assembly of the *Salmonella* SPI-1 T3SS

The structure of the T3SS has been extensively studied by cryo-electron microscopy (cryo-EM) and cryo-electron tomography (CET), which revealed that it consists of a membrane-embedded basal body, an extracellular needle of approximately 50 nm length, and a cytoplasmic complex (Figure 1.2) (B. Hu et al., 2017; Schraidt & Marlovits, 2011; Worrall et al., 2016; Hodgkinson et al., 2009). At the center of this structure is a hollow channel that acts as a conduit for the transport of effector proteins from the bacterial to the host cell cytosol (Radics et al., 2014; Park et al., 2018). The T3SS shares many features with the flagellar apparatus (Diepold & Armitage, 2015), and phylogenetic analysis indicates that this self-assembling rotary engine for cell motility evolved into an ancestral protein secretion system that was then adapted by different bacteria for interactions with their specific eukaryotic hosts (Abby & Rocha, 2012). Thus, T3SSs are classed into seven different families that show a high degree of conservation within their structural core components and architecture, with greater diversity in the extracellular parts of the needle that are involved in host cell interactions (Troisfontaines & Cornelis, 2005; Abby & Rocha, 2012). The structural and functional conservation of the T3SS has made it an attractive target for the development of "anti-infective" drugs that aim to inhibit pathogenicity without otherwise harming the bacterium. Because such drugs do not directly affect the survival of the

pathogen, they are believed to exert a lower selection pressure on bacteria to become resistant (McShan & De Guzman, 2015; Lyons & Strynadka, 2019).

The assembly of the T3SS involves over 20 different proteins and proceeds in two broad phases (Deng et al., 2017; Portaliou et al., 2016). The first, the formation of the basal body, is dependent on the canonical Sec pathway for the insertion of proteins into the bacterial membranes (Tsirigotaki et al., 2016), while in the second phase subunits are secreted through the basal body to assemble into the extracellular needle (Sukhan et al., 2001). Although this chapter describes the structure and assembly of the *Salmonella* SPI-1 T3SS, findings from several different species have contributed to our understanding of this remarkable nanomachine.

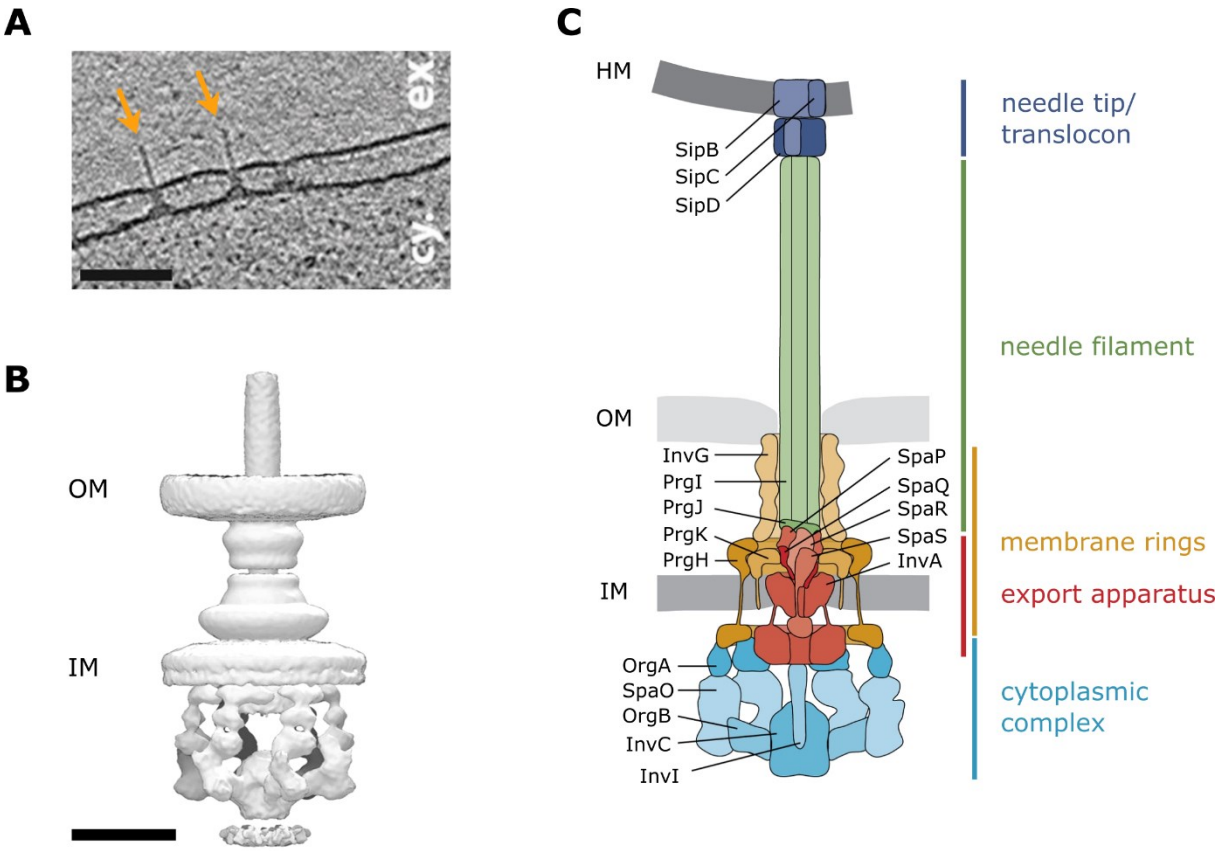


Figure 1.2. Structure of the T3SS. A) Slice of a cryo-electron tomogram of a *Salmonella* Typhimurium cell. Membrane-embedded T3SS are indicated by orange arrows. Scale bar = 100 nm. Cy.: cytosol; ex.: extracellular space. Adapted from Radics et al., 2014. **B)** In-situ cryo-electron tomography map with sub-tomogram averaging of the T3SS from *Salmonella* embedded in the bacterial membranes. Scale bar = 20 nm. Adapted from B. Hu et al., 2017. **C)** Schematic structure indicating individual components. The membrane rings and the export apparatus together constitute the T3SS basal body. Adapted from Wagner et al., 2018. OM: outer membrane; IM: inner membrane; HM: host cell membrane.

1.2.1. Phase I: Sec-dependent assembly of the basal body and cytoplasmic complex

The T3SS basal body consists of a central structure known as the export apparatus and an outer shell of stacked rings that span the bacterial membranes and the periplasmic space. Its assembly is initiated by the construction of the export apparatus, which consists of the five membrane proteins SpaP, SpaQ, SpaR, SpaS, and InvA (Figure 1.3A). Its formation begins with the assembly of a SpaP pentamer, which binds one SpaR and four SpaQ molecules to form a helical array that subsequently associates with a single SpaS (Johnson et al., 2019; Kuhlen et al., 2018; Dietsche et al., 2016). Nine InvA are then recruited to this core structure (Diepold et al., 2011; Wagner et al., 2010), forming a ring of transmembrane domains that is believed to cause the SpaPRQS complex to lift into its final position above the membrane (Johnson et al., 2019; Abrusci et al., 2013; Zilkenat et al., 2016). Both SpaS and InvA possess large cytoplasmic domains that in the case of InvA form a torus in the cytosol below the export apparatus (Abrusci et al., 2013; Minamino & Macnab, 2000).

After completion of the export apparatus, basal body assembly proceeds with the addition of PrgK (Diepold et al., 2011; Wagner et al., 2010), which forms a ring of 24 units around the export apparatus and is itself enveloped by a second ring of 24 PrgH (Schraidt et al., 2010; Kimbrough & Miller, 2000). Both of these proteins are located inside the periplasm and are anchored to the inner membrane by transmembrane helices and in the case of PrgK an additional lipid anchor (Allaoui et al., 1992). Furthermore, PrgH possesses an N-terminal domain that reaches into the cytoplasm, where it forms the innermost ring of the T3SS basal body (Bergeron et al., 2013; Schraidt et al., 2010). The ring of PrgHK connects with a second ring complex that independently assembles in the outer membrane and reaches deep into the periplasm. This outer ring is formed by InvG, a protein of the secretin family that relies on a pilotin for targeting to the outer membrane and ring assembly (Crago & Koronakis, 1998; Daefler & Russel, 1998). While in *Yersinia* and *Shigella* this ring is made up of 12 units (Hodgkinson et al., 2009; Kowal et al., 2013), in the Salmonella SPI-1 system it contains 15 InvG in the outermost region with a 16th unit inserted into the periplasmic region to ensure a symmetry match with the 24 units of the inner PrgHK ring (J. Hu et al., 2019; Worrall et al., 2016).

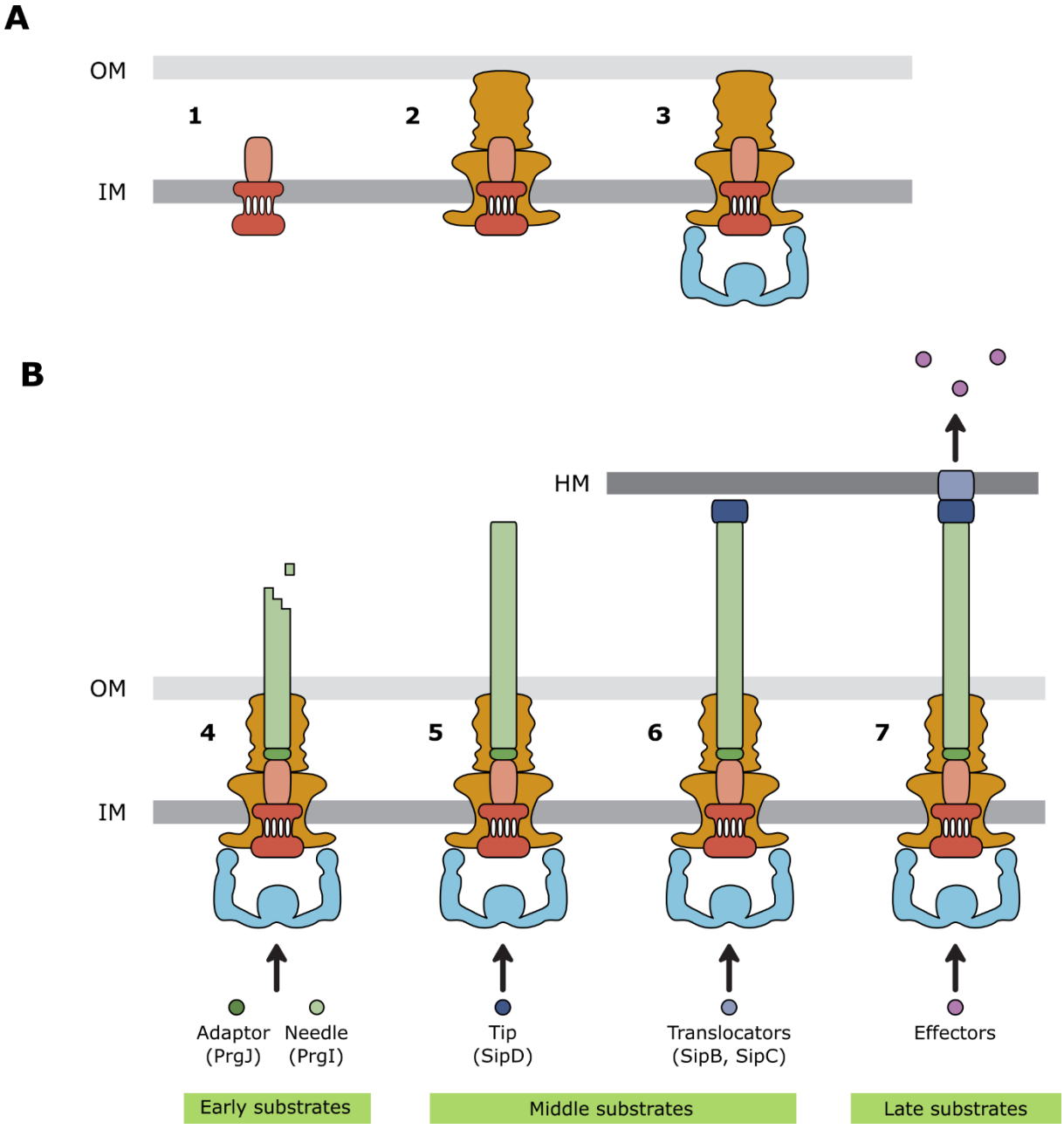


Figure 1.3. Assembly of the T3SS. Structures are colored according to Figure 1.2C. **A)** The Sec-dependent assembly of the basal body is initiated by the formation of the export apparatus. The inner ring proteins assemble around the export apparatus and associate with the pre-assembled outer ring. The completed basal body associates with the cytoplasmic complex, and together these structures constitute the core secretion system through which all further structural and effector proteins are secreted. **B)** The type III export machinery secretes early substrates that form the extracellular needle. Upon reaching its final length, the system switches to the secretion of middle substrates that form the needle tip and the translocation pore in the host cell membrane. After a second substrate switch, effector proteins are injected through the T3SS into the host cell cytosol. OM: outer membrane; IM: inner membrane; HM: host cell membrane.

Once the basal body has been formed, it enables the assembly of the cytoplasmic complex (Zhang et al., 2017; Diepold et al., 2010), which due to its role in selecting substrates for secretion through the T3SS is also called the "sorting platform" (Lara-Tejero et al., 2011). The cytoplasmic complex was visualized by CET and adopts the structure of six pods containing SpaO, which are connected to the basal body through OrgA. Capping the pods at the opposite end is a wheel-like structure of six spokes made up of OrgB connected to a central hub formed by the T3SS ATPase InvC (B. Hu et al., 2015; 2017; Makino et al., 2016). The formation of the sorting platform is incompletely understood but appears to be a highly cooperative process, since in most species each of the four proteins OrgA, SpaO, OrgB and InvC is required for the presence of the other components in the structure, although InvC in the *Salmonella* SPI-1 system seems dispensable (Lara-Tejero et al., 2011; B. Hu et al., 2017; Diepold et al., 2010; 2017; Zhang et al., 2017). An additional protein, InvI, is believed to form a connection between the ATPase hub and the export apparatus, but is not required for the formation of the sorting platform *in vivo* (B. Hu et al., 2017; Ibuki et al., 2013).

1.2.2. Phase II: Type-III-secretion-dependent formation of the extracellular needle

Together, the basal body and the sorting platform constitute the secretion-competent core of the type III secretion system and allow for secretion of subunits that assemble into the extracellular needle structure (Figure 1.3B). First, at the top of the export apparatus a ring-shaped adapter of six PrgJ is formed, which acts to anchor the needle inside the basal body (J. Hu et al., 2019; Zilkenat et al., 2016; Marlovits et al., 2006). The needle itself consists of over 100 PrgI subunits that form a hollow tube with an inner diameter of 2 nm and assembles by addition of secreted PrgI subunits to the distal end of the growing needle (Poyraz et al., 2010). Once the needle reaches its final length, a substrate switch occurs in which the secretion of the "early substrates" PrgJ and PrgI is stopped in a process that involves the protein InvJ and the cytoplasmic domain of the export apparatus protein SpaS (Sorg et al., 2007; Monjarás Feria et al., 2015; Zarivach et al., 2008). Although the precise mechanisms of length determination and substrate switching are still under debate, it is likely that InvJ acts as a molecular ruler that is secreted through the growing needle, measures the distance between the needle base and tip and signals needle completion through an interaction with SpaS (Journet et al., 2003; Wee & Hughes, 2015; Ho et al., 2017; Bergeron et al., 2016).

After formation of the needle, the first substrate switch stops the secretion of the early substrates and permits the secretion of the middle substrates, components that assemble into the needle tip and the translocation pore (Sorg et al., 2007). The needle tip of the SPI-1 T3SS is made up of a pentamer of SipD (Rathinavelan et al., 2014) and its formation marks the completion of the T3SS. The T3SS now enters an inactive state until the detection of host cells triggers the secretion of the translocators SipC and SipB. These hydrophobic proteins assemble on the SipD tip complex to form a translocation pore in the host cell membrane to establish a continuous conduit for the translocation of effector proteins from the bacterial to the host cytosol (Park et al., 2018; Lara-Tejero & Galán, 2009; Myeni et al., 2013; Miki et al., 2004). The precise stoichiometry of SipC and SipB in the translocation pore is currently unclear, although the formation of an 8:8 complex has been shown *in vitro* for *P. aeruginosa* (Romano et al., 2016).

Insertion of the translocation pore into the host cell membrane triggers a second substrate switch. This involves the inactivation of the protein InvE, which is associated with the export apparatus and acts as a gatekeeper that blocks the secretion of effectors (Kubori & Galán, 2002; Portaliou et al., 2017; Botteaux et al., 2009). Conformational changes that occur within the translocator proteins upon insertion into the membrane are believed to act as the switching signal (Armentrout & Rietsch, 2016; Roehrich et al., 2013; Veenendaal et al., 2007), which is transmitted through the needle to the basal body (Kenjale et al., 2005; Torruellas et al., 2005; Guo et al., 2019). By an unknown mechanism, this causes the release of the gatekeeper and its subsequent removal by secretion or proteolysis, triggering the secretion of effector protein (Botteaux et al., 2009; Cheng et al., 2001; Yu et al., 2010).

1.2.3. Secretion of proteins through the T3SS

Most T3SS substrates are targeted to the type III secretion pathway by two signals located within their first ~100 amino acids: a signal sequence comprising the N-terminal 20-30 amino acids and a chaperone-binding domain downstream of it. Signal sequences are not conserved at the primary structure level but share certain biases in their amino acid composition and are structurally disordered (Samudrala et al., 2009; McDermott et al., 2011; Arnold et al., 2009; Buchko et al., 2010). In addition to the signal sequence, the secretion of most substrates requires specialized T3SS chaperones that bind to the chaperone-binding domains of their dedicated substrate (Lee & Galán, 2004; Cheng et al., 1997; Wattiau et al., 1994). This keeps the chaperone-bound domain in a partially unfolded state that is believed to prime the substrate for its insertion

into the narrow secretion channel of the T3SS (Stebbins & Galán, 2003; 2001a). The chaperones act to deliver their substrates to the base of the T3SS (Lara-Tejero et al., 2011), where they interact with both the cytoplasmic complex and the export apparatus (Chen et al., 2013; Allison et al., 2014; Akeda & Galán, 2005; Spaeth et al., 2009; Xing et al., 2018). Here, the substrates are released from their chaperones and unfolded with the help of the ATPase InvC (Akeda & Galán, 2005) and subsequently traverse in an unfolded state through the export apparatus, needle and translocation pore into the host cell cytoplasm (Radics et al., 2014; Dohlich et al., 2014). How the transport through the T3SS is energized is incompletely understood, but it involves both ATP hydrolysis by InvC and the proton-motive force, both of which have been shown to be able to support secretion even in the absence of the other (Paul et al., 2008; Minamino & Namba, 2008; Terashima et al., 2018; Erhardt et al., 2014).

While the structural components of the T3SS are well conserved, the effector proteins exhibit great variability between species, reflecting the specific adaptation of each pathogen to its particular eukaryotic host or hosts. Once inside the host cell, effector proteins act to create an environment that is beneficial to the pathogen by interfering with a number of cellular processes. Although the specific points of action differ between pathogens, processes that are commonly manipulated include the cytoskeletal organization to promote or prevent internalization, vesicle trafficking to avoid the recruitment of anti-bacterial factors to the bacteria-containing phagosome, and intracellular signaling cascades that enhance survival of infected cells, induce the death of immune cells, or modulate the host inflammatory response (Pinaud et al., 2018; Galán, 2009). To achieve these effects, effectors often functionally and structurally mimic host factors and possess enzymatic activities like kinase and phosphatase, protease or ubiquitin ligase activity (Dean, 2011). While some effectors appear to have been acquired from hosts by horizontal gene transfer, others developed by convergent evolution and exhibit no homology in sequence or fold to their eukaryotic counterparts, except for remarkable structural similarity at the surfaces involved in interaction with their target proteins (Stebbins & Galán, 2001b).

1.3. The *Salmonella* SPI-1 T3SS sorting platform

The correct assembly and function of the T3SS depend on the order of the secreted proteins and the cytoplasmic complex has been identified as a critical factor that acts as a “sorting platform” in determining the hierarchy of secreted substrates (Lara-Tejero et al., 2011). This platform

consist of the five proteins OrgA, OrgB, InvC, InvI and SpaO, all of which are essential for the formation of the T3SS needle and the ability of the bacterium to invade host cells (Sukhan et al., 2001). As described above, cryo-electron tomography determined that these proteins assemble into six pods that are capped by a wheel-like structure with a central hub, and the approximate positions of the individual components could be delineated by fusing large protein tags to each of the components and observing the appearance of additional densities in cryo-electron tomography (Figure 1.4) (B. Hu et al., 2017; Makino et al., 2016).

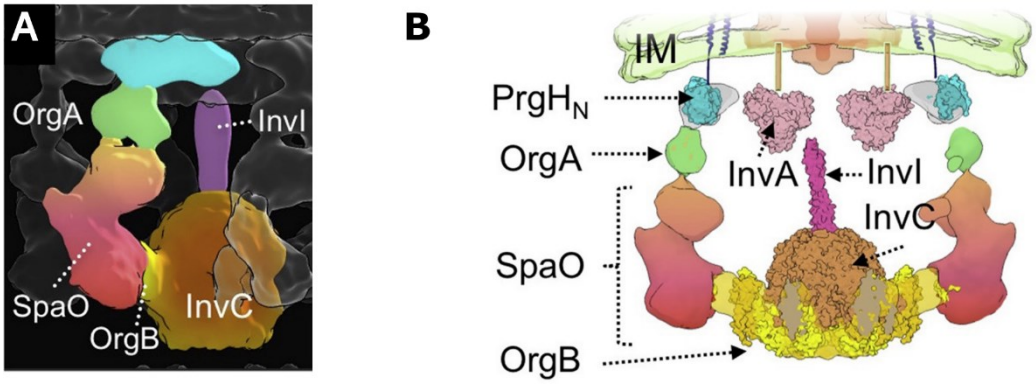


Figure 1.4. The *Salmonella* cytoplasmic complex in detail. A) Cytoplasmic region of the *Salmonella* T3SS cryo-electron tomography map (see Figure 1.2B) indicating the position of each component as determined using tagged proteins. **B)** Cut-through of the cytoplasmic complex structure fitted with the crystal structure of the N-terminal domain of PrgH and atomic models based on available structures of homologous proteins. Adapted from B. Hu et al., 2017.

The sorting platform is anchored to the T3SS basal body by OrgA, which interacts with the inner ring protein PrgH as seen in CET images (B. Hu et al., 2017). On its membrane-distal side OrgA is connected to SpaO, an interaction that has also been observed in a yeast two-hybrid assay involving the homologous proteins of *E. coli* (Soto et al., 2017). SpaO is present in the sorting platform as 24 copies (Diepold et al., 2015; Zhang et al., 2017) and is often also called the C-ring protein, due to its sequence similarity with the flagellar proteins FliM and FliN that do not form discrete pod structures but assemble into a continuous cytosolic ring below the basal body of the flagellar T3SS (Makino et al., 2016). SpaO contains a poorly characterized N-terminal domain and two C-terminal SPOA (surface presentation of antigen) domains that fold into a SPOA1-SPOA2 domain dimer that forms the binding site for OrgB (Notti et al., 2015). Like many of its homologs, the *spaO* gene possesses an internal translation initiation site, which produces a second protein that encompasses the C-terminal SPOA2 domain and is hence referred to as SpaO$_C$ (Yu et al., 2011; Song et al., 2017; Bernal, 2019; Bernal, Börnicke, et al., 2019). Dimers of

this shorter product associate with full-length SpaO to form heterotrimers (McDowell et al., 2016; Bzymek et al., 2012). Crystal structures have been solved for the $SpaO_C$ dimer (SPOA2-SPOA2) and the SPOA1-SPOA2 dimer of full-length SpaO both alone and in complex with the N-terminus of OrgB (Notti et al., 2015; Bzymek et al., 2012; McDowell et al., 2016). However, the precise role of $SpaO_C$ in type III secretion remains controversial, with contradicting results indicating that it is essential for T3SS function in both *Shigella* and *Yersinia*, but dispensable in *Salmonella* (McDowell et al., 2016; Lara-Tejero et al., 2019; Yu et al., 2011; Bzymek et al., 2012). Similarly, it is unclear whether $SpaO_C$ forms part of the assembled sorting platform (Diepold et al., 2015; Lara-Tejero et al., 2019).

The C-terminal domain of SpaO interacts with OrgB, which forms the spokes that connect the ATPase InvC to the pods of the cytoplasmic complex (B. Hu et al., 2017; Notti et al., 2015). The interaction of OrgB with InvC inhibits the ATPase activity in solution, probably to ensure ATP hydrolysis only occurs upon assembly of the cytoplasmic complex at the T3SS base (Blaylock et al., 2006; Gonzalez-Pedrajo et al., 2002; Minamino & MacNab, 2000). Like for SpaO, the structure of OrgB has only partially been solved and data are available for its N-terminal 30 residues in complex with SpaO (Notti et al., 2015), as well as for the C-terminal domain of the flagellar homolog FliH interacting with the ATPase FliI (Imada et al., 2016). While no structural data exist for the N-terminal domain following the SpaO interaction site, it is predicted to form a coiled-coil involved in homo-dimerization (Minamino, González-Pedrajo, et al., 2002).

At the center of the sorting platform and held in place by OrgB, the T3SS ATPase InvC forms a hexameric ring with an inner diameter of 2.5-3 nm that aligns with the entry pore of the export apparatus (B. Hu et al., 2017; Claret et al., 2003). The activity of this ATPase is greatly enhanced through homo-oligomerization mediated by its N-terminal domain (Claret et al., 2003; Minamino et al., 2006; Burgess, Burgess, et al., 2016), suggesting that it only becomes fully active in the assembled cytoplasmic complex. InvC plays a central role in secretion and has been shown to be able to de-chaperone and unfold T3SS substrates prior to their insertion into the T3SS secretion channel (Akeda & Galán, 2005; Yoshida et al., 2014). However, how these processes are linked to its ATPase activity on a mechanistic level has not yet been elucidated. Crystal structures of InvC and homologs from several species have been solved and revealed extensive similarities to the F_1 ATPase, which is part of the F_OF_1 synthase complex involved in proton-motive force-driven ATP synthesis (Imada et al., 2007; Zarivach et al., 2007; Allison et al., 2014; Burgess, Burgess, et al., 2016; Bernal, Römermann, et al., 2019). Like the F_1 ATPase, the central hole of the InvC hexamer binds to InvI, a protein with structural similarity to the central stalk

of the ATP synthase (Majewski et al., 2019; Ibuki et al., 2011). This protein forms a second interaction with the cytoplasmic domain of InvA, linking the cytoplasmic complex to the export apparatus (Ibuki et al., 2013). However, the role of InvI in type III secretion is currently poorly understood.

In addition to their presence in the T3SS-associated sorting platform, OrgA, SpaO, OrgB and InvC all have also been observed to exist in a cytosolic state, and SpaO has been shown to undergo a dynamic exchange between the T3SS-bound state and cytosolic pool (Diepold et al., 2015; Rocha et al., 2018; Zhang et al., 2017; Diepold et al., 2017). Activation of type III secretion leads to an increase in this subunit exchange and affects the interactions formed between sorting platform components in the cytosolic pool, indicating that the sorting platform is a highly dynamic structure and that cycling of subunit between the cytosol and the T3SS may be an important factor in its function.

1.4. Aims of the study

Despite the central role of the cytoplasmic sorting platform complex in type III secretion, there are still several open questions regarding its structure and function. For example, even though great progress has been made in the visualization of this structure in recent years, the precise stoichiometry and spatial arrangement of its components, as well as their assembly process are still uncertain. Moreover, although the complex is known to play a role in selecting substrates for secretion, little is known about the mechanism of this sorting process and the precise site of interaction between the sorting platform and the chaperone-substrate complexes. Finally, recent findings indicate that soluble complexes play a role in the function of the sorting platform and consequently the T3SS, but these complexes have only partially been characterized until now.

Therefore, this study aimed to reconstitute the sorting platform *in vitro* in order to obtain soluble complexes or even complete sorting platforms for precise molecular characterization. In particular, purified complexes would be amenable to methods like multi-angle light scattering and native mass spectrometry to provide detailed information about their mass and precise molecular composition, and might prove suitable for structural analysis by X-ray crystallography or small-angle X-ray scattering. Furthermore, using these purified complexes in quantitative interaction studies with T3SS substrates may help elucidate the mechanism of interaction and provide a path to understanding the sorting function of the T3SS cytoplasmic complex.

2. Results

2.1. Reconstitution of soluble sorting platform complexes *in vitro*

Formation of the *Salmonella* Typhimurium SPI-1 T3SS sorting platform *in vivo* has been shown to involve the six proteins OrgA, OrgB, SpaO and its short variant SpaO$_C$, InvC and InvI, which are proposed to connect to the T3SS needle base by interactions with the cytoplasmic domains of the inner ring protein PrgH and the export apparatus protein InvA (PrgH$_{1-140}$ and InvA$_{357-685}$, respectively). Therefore, in order to reconstitute the sorting platform *in vitro*, pairs of the genes encoding these components were cloned into bicistronic Duet vectors that allow for the simultaneous expression of two genes, and used in conjunction with other plasmids to assess different combinations of sorting platform components (Table 4.7). Following co-expression in *E. coli*, proteins were tested for their solubility and whether different components co-purified as complexes in Strep-Tactin affinity chromatography.

First, expression of the genes *spaO* and *orgB* showed that *spaO* produces both the C-ring protein SpaO and the shorter variant SpaO$_C$, which form a complex that is soluble and can be purified with high yields (Figure 2.1). In contrast, the ATPase regulator protein OrgB by itself is insoluble, but its solubility is greatly increased by co-expression with SpaO/SpaO$_C$ and high yields of

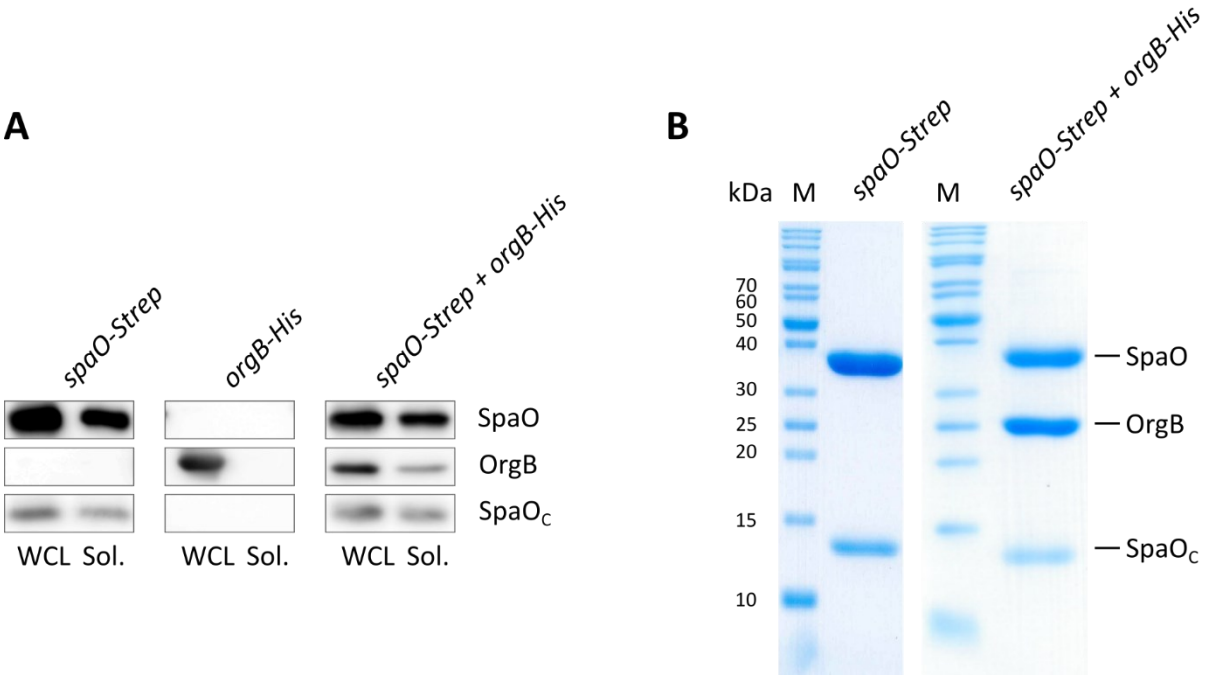


Figure 2.1. Analysis of complexes formed by SpaO/SpaO$_C$ and OrgB. A) Western blot of whole cell lysates (WCL) and soluble fractions (Sol.). **B)** Coomassie-stained SDS-PAGE of SpaO/SpaO$_C$ and SpaO/SpaO$_C$/OrgB complexes purified by Strep-Tactin affinity and size-exclusion chromatography.

complexes containing SpaO, $SpaO_C$ and OrgB could be purified. It should be noted that due to its small size and low Coomassie binding, $SpaO_C$ is often not visible in SDS-PAGE analysis, especially when only low levels of protein were isolated in small-scale pull-down assays. However, later analyses showed that in virtually all soluble complexes involving SpaO, $SpaO_C$ is also present (e.g. Figure 2.7).

Next, possible interactions between the basal body inner ring protein PrgH and SpaO or OrgA were tested, both of which have been proposed to act as proteins that anchor the sorting platform to the T3SS basal body in different species (Morita-Ishihara et al., 2006; B. Hu et al., 2017; Jackson & Plano, 2000; Soto et al., 2017). However, in pull-down experiments of SpaO/$SpaO_C$ and $PrgH_{1-140}$ no co-purification could be observed, indicating that SpaO might not directly connect to the T3SS basal body (Figure 2.2A). Similarly, combinations of OrgA with $PrgH_{1-140}$ showed that the protein OrgA not only is of low solubility and can be purified in very small amounts, but also does not form stable complexes with $PrgH_{1-140}$ (Figure 2.2B). On the other hand, when *orgA* and *$prgH_{1-140}$* were co-expressed with *spaO* and *orgB*, small amounts of complexes containing OrgA, SpaO and OrgB were purified (Figure 2.2C). Thus, while these pull-downs demonstrate an interaction between OrgA, SpaO and OrgB in *Salmonella*, no interaction of these components with $PrgH_{1-140}$ could be detected. Whether SpaO or OrgB is responsible for the binding of OrgA in forming the OrgA/SpaO/OrgB complex could not be determined from these pull-down assays, because untagged SpaO non-specifically bound to the Strep-Tactin resin and thus precluded the detection of possible OrgA/SpaO complexes.

One of the most important components of the sorting platform is the ATPase InvC. Co-expression and purification with other sorting platform proteins showed that it forms soluble complexes with both OrgB and SpaO/OrgB, but not with OrgA (Figure 2.3). SpaO by itself was also apparently pulled down by InvC, but non-specific interactions of SpaO with the Strep-Tactin resin were also detected, questioning the possible formation of SpaO/InvC complexes. Moreover, previous pull-down experiments between InvC and SpaO using nickel affinity chromatography could not detect an interaction between these proteins (Notti et al., 2015).

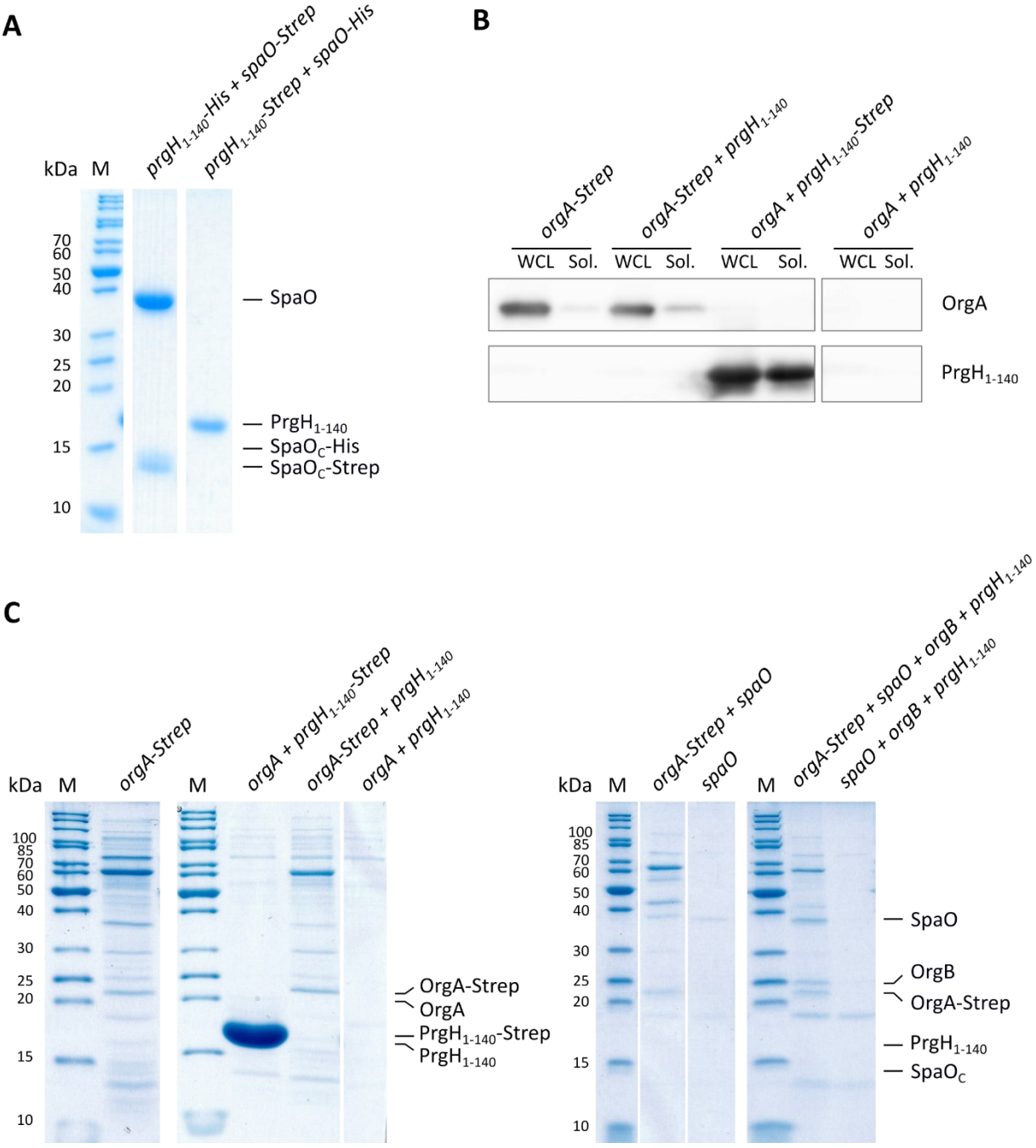


Figure 2.2. Analysis of complexes formed by PrgH$_{1-140}$, SpaO and OrgA. A) Coomassie-stained SDS-PAGE of Strep-Tactin-purified PrgH$_{1-140}$ and SpaO/SpaO$_C$ after co-expression. **B)** Western blots of whole cell lysates (WCL) and soluble fractions (Sol.) of *E. coli* expressing *orgA* and *prgH$_{1-140}$*. Proteins were detected through their C-terminal Strep-tag using an anti-Strep antibody. **C)** Coomassie-stained SDS-PAGE of Strep-Tactin-purified OrgA and PrgH$_{1-140}$ following co-expression with other sorting platform components. Image contrast was adjusted in panel C) to improve visibility of fainter bands.

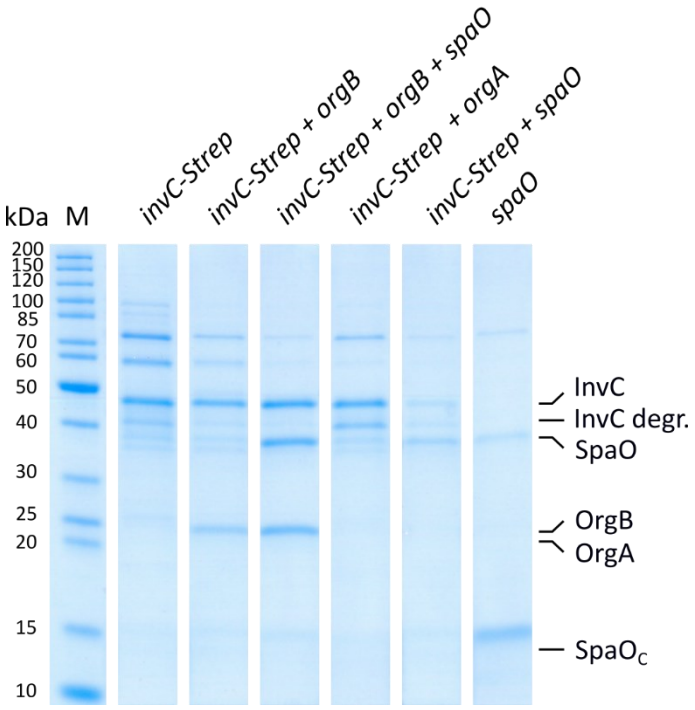


Figure 2.3. Complexes formed by InvC with other sorting platform components. Coomassie-stained SDS-PAGE of sorting platform components co-purified with Strep-Tactin-purified InvC-Strep after co-expression. The protein band running just below full-length InvC represents an InvC degradation product (InvC degr.) truncated at the N-terminus (see Figure 2.4 and Figure 2.14).

Next, combinations of the proteins InvC, InvI and the cytoplasmic domain of the export apparatus protein InvA, $InvA_{357-685}$, were tested. While the solubility of $InvA_{357-685}$ was demonstrated previously (Worrall et al., 2010), western blot analysis showed that InvI is mostly insoluble, even though low levels could be affinity-purified (Figure 2.4). When co-expressed with InvC, low levels of InvI co-purified with InvC and vice versa, demonstrating an interaction between the two proteins. In addition, InvI was capable of pulling down $InvA_{357-685}$, and both $InvA_{357-685}$ and InvI could be precipitated with InvC. Together, these results show that in *Salmonella* the ATPase InvC interacts with InvI, which in turn also binds to the C-terminal domain of InvA. However, these complexes could only be purified with very low yields that are difficult to detect by Coomassie staining, probably due to the low solubility of InvI even in the presence of its binding partners.

Collectively, these results provide biochemical evidence for a number of interactions between the proposed sorting platform components of the SPI-1 T3SS of *Salmonella* Typhimurium and allow for the construction of a linear chain of interactions in the order of SpaO(+$SpaO_C$)-OrgB-InvC-InvI-InvA. OrgA connects to this chain at the point of SpaO/OrgB, even though it could

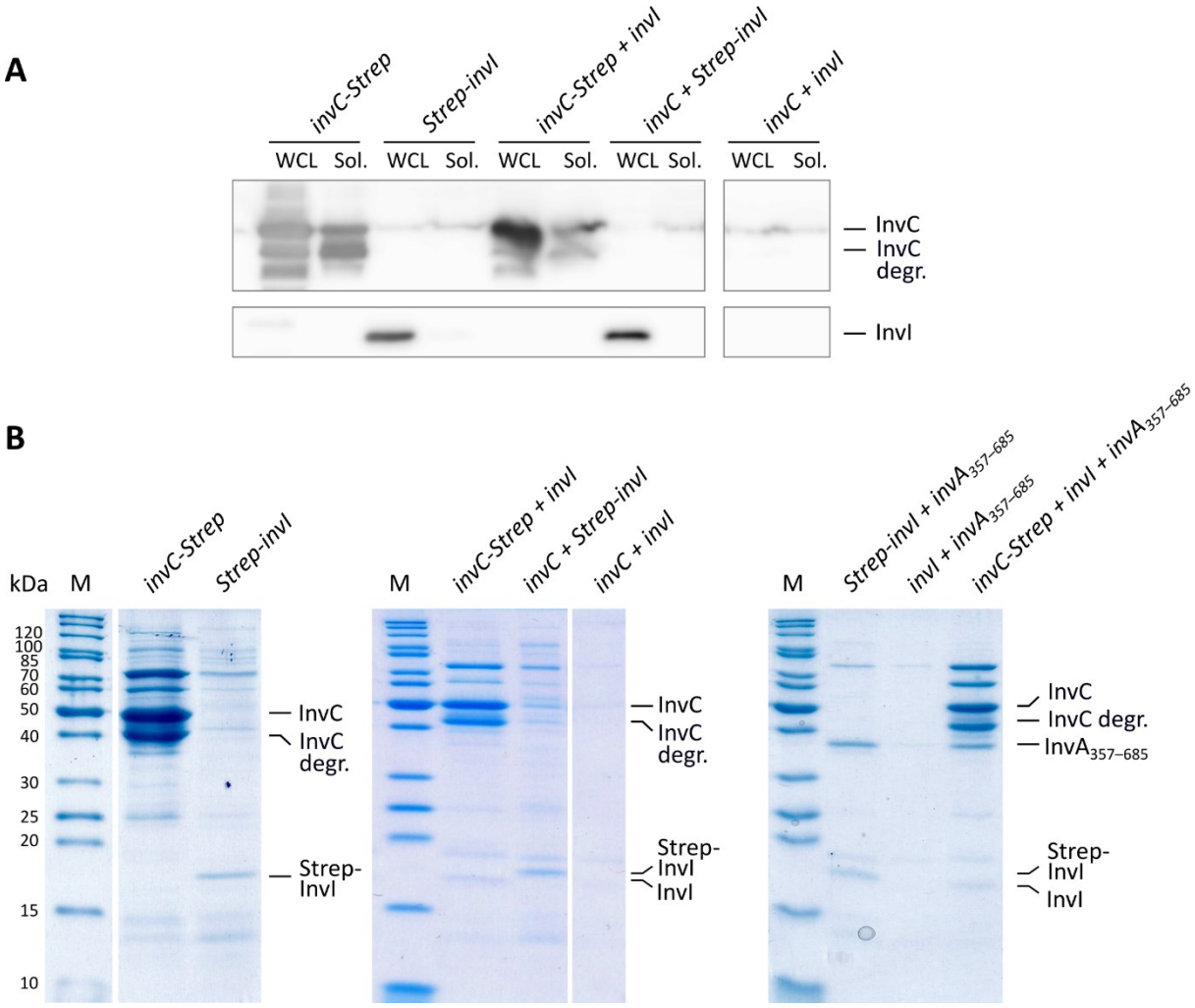


Figure 2.4. Complexes formed by InvC, InvI and InvA$_{357-685}$. A) Western blots of whole cell lysates (WCL) and soluble fractions (Sol.) of *E. coli* expressing *invC* and *invI*. Proteins were detected through their C-terminal Strep-tag using an anti-Strep antibody. **B)** Coomassie-stained SDS-PAGE of Strep-Tactin-purified combinations of InvC, InvI and InvA$_{357-685}$. The protein band running just below full-length InvC is an InvC degradation product (InvC degr.) truncated at the N-terminus, as indicated by the detection of the C-terminal Strep-tag in western blot analysis in A) and Figure 2.14B. Image contrast in B) was adjusted to enhance visibility of faint bands.

not be determined which of the two proteins forms the point of interaction. However, based on the reported interaction of the *E. coli* homologous proteins and CET imaging of the *Salmonella* T3SS, the interaction most likely occurs between OrgA and SpaO (Soto et al., 2017; B. Hu et al., 2017). While this arrangement links the sorting platform to the T3SS basal body by the simultaneous interaction of InvI with the ATPase InvC and the export apparatus protein InvA, no interactions between the inner ring protein PrgH and its proposed interaction partners SpaO or OrgA could be detected in solution.

2.1.1. Duet vectors encoding SpaO suppress the heterologous co-expression of sorting platform components from other Duet vectors

In order to reconstitute the complete sorting platform *in vitro*, the three bicistronic Duet vectors pCOLADuet-1 *orgA+prgH*$_{1-140}$, pCDFDuet-1 *orgB+spaO* and pACYCDuet-1 *invC+invI* were introduced into *E. coli* for the co-expression of the sorting platform components and subsequent purification. Despite the supposed inter-compatibility of these plasmids (Novy et al., 2002; Held et al., 2003; 2004), SDS-PAGE showed that only *spaO* and *orgB* were over-expressed in all tested combinations, with no or only very faint bands discernable for PrgH$_{1-140}$, OrgA, InvC and InvI (Figure 2.5, samples 1-7). The reasons for this could not be conclusively determined but appear to be related to the gene *spaO* rather than the pCDFDuet-1 vector it was expressed from. This is indicated by the fact that when pCDFDuet-1 *orgB* was combined with pACYCDuet-1 *invC-Strep,* both OrgB and InvC were produced (Figure 2.5, sample 8), showing that there is no general incompatibility between these two types of vectors. On the other hand, only SpaO/SpaO$_C$ was

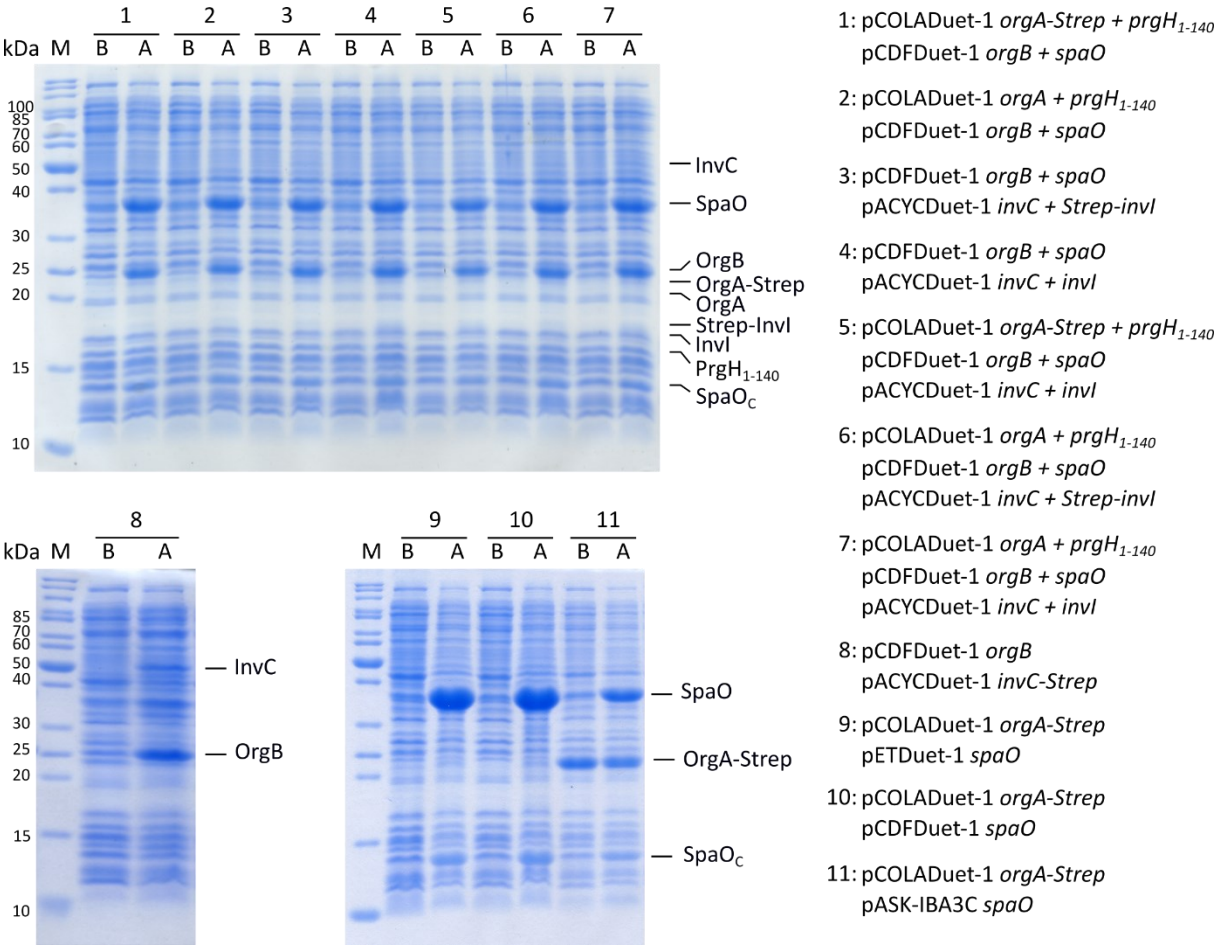


Figure 2.5. Co-expression of sorting platform components from different Duet vector combinations. SDS-PAGE of lysates of *E. coli* carrying different combinations of Duet vectors before (B) and after (A) induction of expression.

produced when *spaO* in either pCDFDuet-1 or pETDuet-1 was combined with pCOLADuet-1 *orgA* (Figure 2.5, samples 9-10), while both OrgA and SpaO/SpaO$_C$ were produced when pCOLADuet-1 *orgA-Strep* was combined with *spaO* on a different class of vector, pASK-IBA3C (Figure 2.5, sample 11). These results indicate that the expression of *spaO* from a Duet vector suppresses the expression of genes carried by other Duet vectors, making the use of this vector system unsuitable for the reconstitution of the complete sorting platform.

2.1.2. Sorting platform components do not form stable complexes with OrgC in solution

The *Salmonella* sorting platform genes *orgA* and *orgB* are transcribed as part of a larger operon that also encodes for the structural T3SS components PrgH, -I, -J and –K, as well as other regulators or effector proteins (Figure 2.6A) (Klein et al., 2000). Immediately downstream of *orgA* and *orgB* lies the gene *orgC*, whose positioning suggested that it might also be involved in the formation of the sorting platform. Therefore, different sorting platform components were co-expressed with Strep-tagged *orgC* in *E. coli* and tested for complex formation using a Strep-Tactin pull-down assay. SDS-PAGE of the purified proteins showed that no sorting platform proteins are co-purified with OrgC above levels of non-specific adhesion to the resin, with the possible exception of OrgB (Figure 2.6B), which showed a band of slightly stronger intensity when pulled down by OrgC than when expressed alone. However, since similar levels of co-purification with OrgC where not observed when OrgB was present in combination with SpaO or InvC-InvI, it is unclear whether this apparent interaction is of biological relevance. In fact, a recent study showed that OrgC is secreted through the T3SS and acts extracellularly to enhance the polymerization of PrgI into the needle structure (Kato et al., 2018), which is consistent with these results that OrgC is unlikely to be a structural component of the sorting platform.

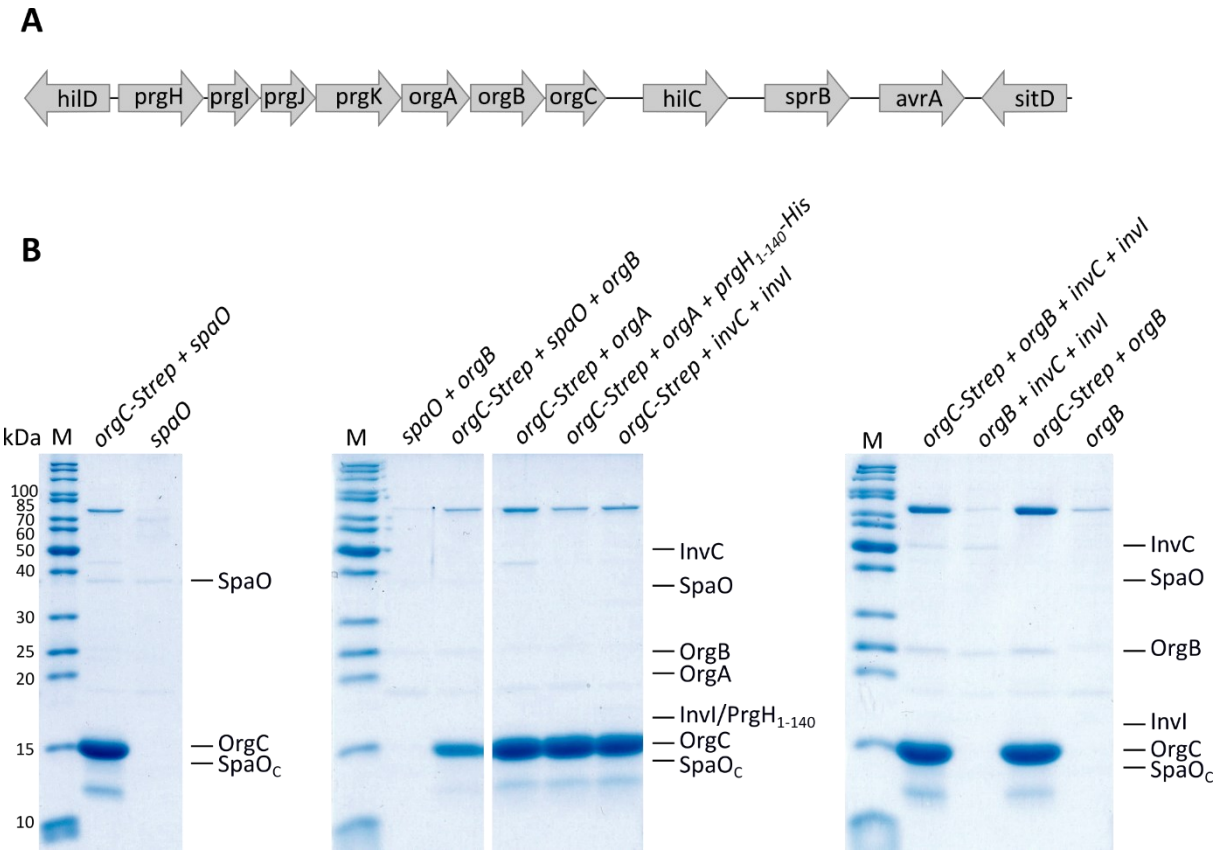


Figure 2.6. Analysis of pull-downs of OrgC with other sorting platform components. A) Schematic organization of the operon harboring *orgA*, *orgB* and *orgC*. **B)** Coomassie-stained SDS-PAGE of different sorting platform components co-purified with OrgC-Strep after co-expression in *E. coli*. Image contrast was adjusted to enhance visibility of faint bands.

2.2. Purifyable sorting platform sub-complexes

The *in vitro* reconstitution and characterization of the complete *Salmonella* sorting platform was precluded by issues with the Duet vector system and the low solubility of both OrgA and InvI. However, sub-complexes of $SpaO_C$ and $SpaO/SpaO_C$, as well as $SpaO/SpaO_C$/OrgB and $SpaO/SpaO_C$/OrgB/InvC were soluble and could be purified in sufficient yields to allow for further characterization (Figure 2.7). Of these complexes, $SpaO/SpaO_C$/OrgB and $SpaO/SpaO_C$/OrgB/InvC were characterized further in this work, while $SpaO_C$ and $SpaO/SpaO_C$ were investigated in an associated collaborative study (Bernal, 2019; Bernal, Börnicke, et al., 2019).

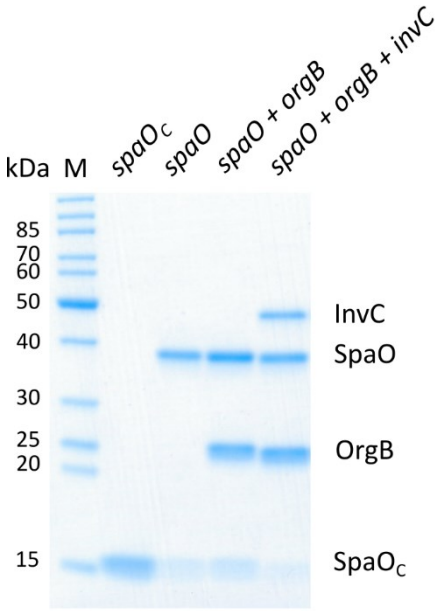


Figure 2.7. Soluble sorting platform sub-complexes. Coomassie-stained SDS-PAGE of purifiable sorting platform sub-complexes after affinity and size-exclusion chromatography.

2.3. Characterization of SpaO/SpaO$_C$/OrgB complexes

As shown in Figure 2.1, the ATPase regulator protein OrgB and the C-ring proteins SpaO/SpaO$_C$ form stable complexes that were amenable to *in vitro* characterization by analytical size-exclusion chromatography (SEC), multi-angle light scattering (MALS) and native mass spectrometry (MS). In contrast, SpaO/SpaO$_C$/OrgB complexes were not suitable for detailed structural determination by X-ray crystallography, due to a propensity to aggregate and precipitate at protein concentrations above 3 mg/ml (data not shown).

2.3.1. Analytical size-exclusion chromatography and multi-angle light scattering of SpaO/SpaO$_C$/OrgB

In order to assess the homogeneity and molecular mass of SpaO/SpaO$_C$/OrgB complexes, analytical size-exclusion chromatography and online multi-angle light scattering (SEC-MALS) was performed (Figure 2.8, left). The elution profile showed the presence of a large peak that contains all three components SpaO, SpaO$_C$ and OrgB. A second, smaller elution peak containing much lower levels of OrgB eluted at a volume very similar to that of purified SpaO/SpaO$_C$ (13.6 ml, data not shown), indicating that this peak consists mostly of SpaO/SpaO$_C$

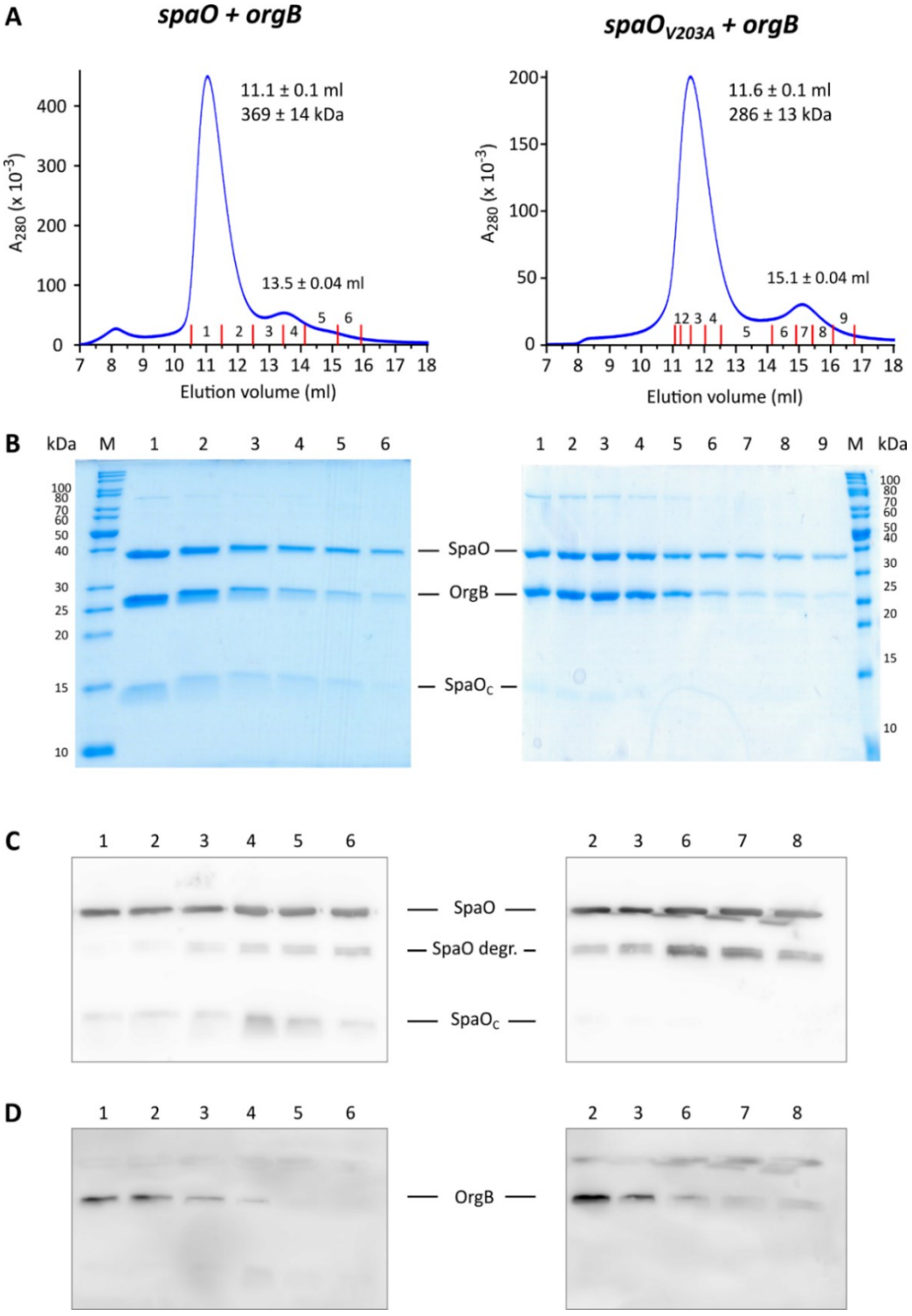


Figure 2.8. SEC analysis of SpaO/SpaO$_C$/OrgB complexes. On the left-hand side, data for complexes derived from co-expressing wild type *spaO+orgB* are shown, on the right-hand side data for *spaO$_{V203A}$+orgB*. SEC was performed on a Superdex 200 10/30 GL column. **A)** Representative SEC elution profiles of affinity-purified complexes. Mean elution volumes and their corresponding molecular masses (± standard deviations) from five measurements are indicated. Collected elution fractions are denoted by red lines. The SEC calibration curve for molecular mass determination can be found in Supplementary Figure 1. **B)** Coomassie-stained SDS-PAGE of elution fractions collected in A). **C)** Anti-Strep western blots of elution fractions, detecting the C-terminal Strep-tag on SpaO and SpaO$_C$, as well as a likely degradation product of SpaO (SpaO degr.). **D)** Blots in C) were stripped and re-probed with anti-His antibody detecting the C-terminal His-tag on OrgB.

complexes and possibly low levels of dissociation products from the full SpaO/SpaO$_C$/OrgB complex. Interestingly, while the elution volume of the major peak indicated a molecular mass of 369 ± 14 kDa for the complex, a molecular mass of 165 ± 8 kDa was determined by SEC-MALS (Figure 2.9A). Because molecular masses found by SEC are also dependent on the shape of a molecule and are only accurate for globular proteins, the higher apparent mass in SEC indicates that the complex deviates strongly from a globular shape and instead assumes a more extended configuration.

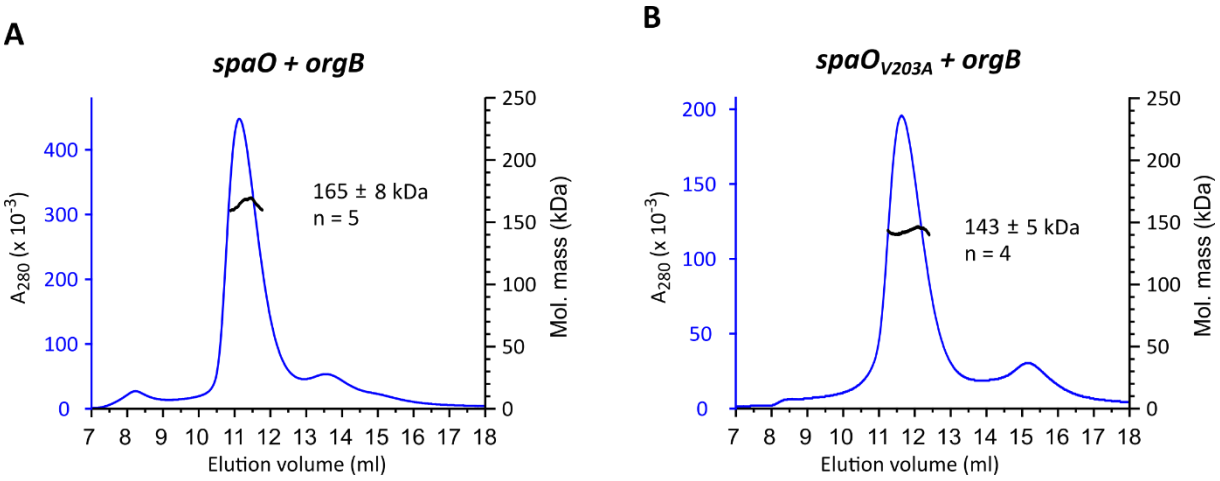


Figure 2.9. SEC-MALS analysis of SpaO/SpaO$_C$/OrgB complexes using wildtype *spaO* (A) and the *spaO$_{V203}$* mutant (B). Representative SEC elution profiles (A_{280}, blue) and weight-averaged molecular masses across the elution peaks (black) are shown. Mean molecular masses and standard deviations based on "n" measurements are indicated. Theoretical masses of individual proteins: SpaO-Strep = 35 kDa; OrgB-His = 28 kDa; SpaO$_C$-Strep = 12 kDa.

In order to investigate the role of SpaO$_C$ in these complexes, o*rgB* was co-expressed with *spaO$_{V203A}$,* a *spaO* variant carrying a mutation in the internal start codon of SpaO$_C$ that greatly diminishes the levels of SpaO$_C$ (Bernal, Börnicke, et al., 2019). SEC with subsequent SDS-PAGE and western blot analysis showed that the resulting complexes have greatly reduced levels of SpaO$_C$ and a decreased molecular mass when compared to wildtype complexes (Figure 2.8, right side). In fact, MALS analysis indicates a molecular mass of 143 ± 5 kDa for the major SpaO$_{V203A}$/OrgB species (Figure 2.9B), which represents a decrease of approximately 22 kDa and could indicate the loss of a SpaO$_C$ dimer (2 x 12 kDa for Strep-tagged SpaO$_C$) from the wildtype SpaO/SpaO$_C$/OrgB complexes. In addition, the loss of SpaO$_C$ resulted in an approximately 3-fold decrease in yield of isolated complexes (data not shown) and led to an overall increase in the levels of SpaO degradation products (Figure 2.8C). These products run at an estimated mass of

24 kDa in SDS-PAGE and are still detected through their C-terminal Strep-tags in western blot analysis, indicating that in the absence of $SpaO_C$ the N-terminal domain of SpaO becomes more susceptible to degradation. Interestingly, when the $SpaO_{V203A}$/OrgB complex was mixed with purified $SpaO_C$ prior to SEC, only the minor elution peak shifted to a position that suggests the formation of SpaO/$SpaO_C$ complexes, while the major peak of the $SpaO_{V203A}$/OrgB complex was largely unaffected (Figure 2.10). This apparent inability to bind $SpaO_C$ and form wildtype-like complexes suggests that the $SpaO_{V203A}$/OrgB complex is rather tightly connected, most likely involving aberrant interactions between regions that would otherwise be shielded by $SpaO_C$.

Together, these observations indicate that $SpaO_C$ is not required for the formation of complexes between full-length SpaO and OrgB, but acts to enhance their stability in solution. This stabilization is likely caused by the association of $SpaO_C$ with the N-terminal domain of SpaO, which has been shown to stabilize also SpaO in the absence of OrgB (Bernal, Börnicke, et al., 2019).

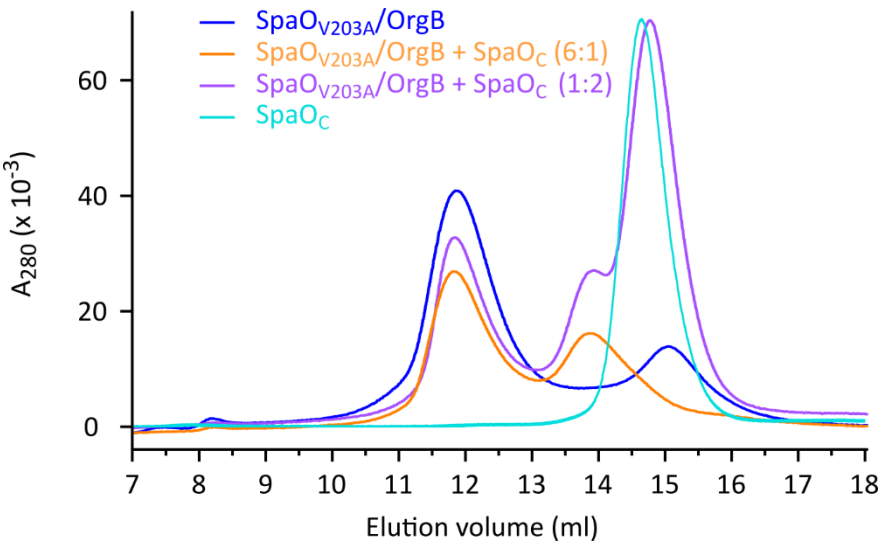


Figure 2.10. SEC analysis of $SpaO_{V203A}$/OrgB complexes mixed with purified $SpaO_C$. Mixing ratios (mass:mass) are given in brackets. The curve of $SpaO_C$ was scaled down to 50% of its original size for better comparability.

2.3.2. Native mass spectrometry of SpaO/SpaO$_C$/OrgB complexes

In order to determine the exact stoichiometry of subunits within the SpaO/SpaO$_C$/OrgB complex, complexes purified by affinity chromatography were analyzed by native MS. In native MS, proteins are ionized and transferred to the gas phase by electro-spray ionization (ESI). In this technique, samples are transferred to a volatile buffer (e.g. ammonium acetate) and loaded into a capillary to which a high voltage is applied. This voltage causes the formation of protons at the solution-capillary interface, which accumulate within the sample and lead to the expulsion of droplets from the capillary tip due to charge repulsion (Kebarle & Verkerk, 2009). Solvent evaporation from these droplets increases their charge density and causes them to break into ever smaller nanodroplets that only contain a single protein complex. Eventually, the complete evaporation of solvent molecules from the nanodroplets leaves behind protonated and thus multi-charged protein complexes in the gas phase, which can be analyzed by a mass spectrometer (Konermann et al., 2013). Compared to other methods, ESI is a "soft" ionization process that does not break non-covalent interactions, making native MS a sensitive technique for the analysis of biological complexes.

Applying native MS to SpaO/SpaO$_C$/OrgB showed that the major molecular species of OrgB-containing complexes possesses two molecules of OrgB bound to two SpaO-2SpaO$_C$ heterotrimers, forming complexes of 176 kDa and 2(SpaO-2SpaO$_C$)-2OrgB stoichiometry (Figure 2.11, Table 2.1). Additional, low-abundance species lacking either one SpaO-2SpaO$_C$ heterotrimer or SpaO-2SpaO$_C$-OrgB were observed, which might represent assembly/disassembly intermediates of the full complex. Interestingly, two molecules of OrgB are present in the great majority of OrgB-containing complexes, indicating that OrgB mainly exists as a dimer. Together with the previously reported interaction of the SpaO SPOA1-SPOA2 domain dimer and the N-terminus of OrgB (Notti et al., 2015), this suggests an architecture of the complex in which each monomer of the OrgB dimer is bound to one SpaO-2SpaO$_C$ heterotrimer.

In addition to OrgB-containing complexes, mass spectra also show the presence SpaO$_C$ dimers, SpaO-2SpaO$_C$ heterotrimers and 2(SpaO-2SpaO$_C$) heterohexamers in the lower m/z range. This region of the spectrum strongly resembles mass spectra of purified SpaO/SpaO$_C$ (Bernal, Börnicke, et al., 2019; Bernal, 2019), indicating that these species are likely stable building blocks of the larger OrgB-containing complexes and are the result of dissociation from OrgB during purification or sample preparation for native MS. It should be noted that the high intensity of SpaO/SpaO$_C$ peaks does not necessarily reflect their relative abundance in solution, because in

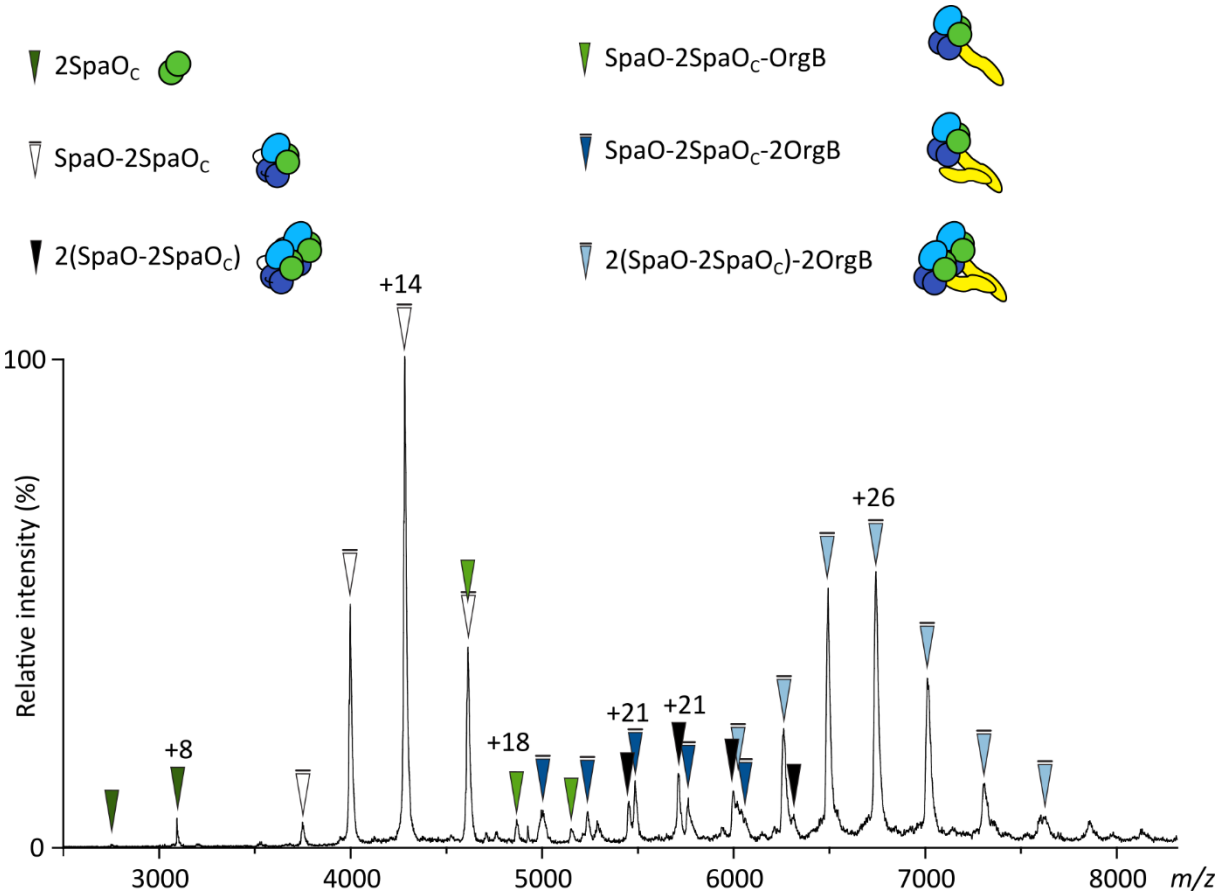


Figure 2.11. Native mass spectrum of purified SpaO/SpaO$_C$/OrgB complexes. The main charge state of each protein or protein complex is labeled. In the schematic representations of complexes SpaO$_C$ and OrgB are depicted in green and yellow, respectively, while SpaO is shown with its N-terminal domain in light blue and its C-terminal SPOA domains in dark blue. Theoretical and experimentally determined molecular masses are given in Table 2.1 . Both SpaO and SpaO$_C$ carry a Strep-Tag, a His-tag is present on OrgB. MS analysis was performed by Johannes Heidemann.

native MS different molecular species have different efficiencies of ionization and transfer to the gas phase, meaning that there is no direct correlation between the levels of a complex in solution and the intensity of its peaks in a mass spectrum.

Native MS was also applied to the SpaO$_{V203A}$/OrgB complex, but only species of 2SpaO$_C$, SpaO-2SpaO$_C$ and contaminating DnaK were detected (data not shown). This indicates that SpaO$_{V203A}$/OrgB does not remain stable through the native MS steps of buffer exchange and ionization, highlighting that SpaO$_C$ is not only important for the stability of SpaO (Bernal, Börnicke, et al., 2019), but also for that of OrgB-containing sorting platform complexes.

Table 2.1. Theoretical masses and average experimental masses of SpaO/SpaO$_C$/OrgB proteins and protein complexes as determined by native MS. n≥3; STDEV: standard deviation; Avg. FWHM: average full-width at half-maximum.

Protein/-complex	Theoretical mass (Da)	Experimental avg. mass (Da)	STDEV (Da)	Avg. FWHM (Da)
2SpaO$_C$-Strep	24,748.0	24,746.9	1.5	14
OrgB MS/MS	26,448.4	26,459	21	100
SpaO MS/MS	33,793.7	33,800	40	250
SpaO-Strep/2SpaO$_C$-Strep	59,897.9	59,930	40	250
2(SpaO-Strep/2SpaO$_C$-Strep)	119,795.7	120,020	180	750
2(SpaO/2SpaO$_C$)/OrgB MS/MS	138,475.6	138,850	290	250
2(SpaO-Strep/2SpaO$_C$-Strep)/2OrgB-His	175,227.2	175,500	500	890

2.3.3. MS/MS indicates the stabilization of interactions between SpaO-2SpaO$_C$ heterotrimers within SpaO/SpaO$_C$/OrgB complexes

Next, selected ions of the 2(SpaO-2SpaO$_C$)-2OrgB complex were subjected to collision-induced dissociation tandem MS (CID MS/MS) experiments. In these experiments selected species are collided with argon gas, which causes them to accumulate internal energy and charges, leading to the unfolding of a subunit and its subsequent ejection from the complex (Benesch, 2009). Although the process of CID is not completely understood, it has been found that the dissociating proteins are usually small monomeric proteins from the complex periphery, which means that the dissociation pathways observed in CID experiments can provide insights into the topology of a protein complex (Benesch, 2009). In the case of 2(SpaO-2SpaO$_C$)-2OrgB mainly the dissociation of an OrgB monomer was observed, while in a second, less prominent pathway SpaO was ejected from the complex (Figure 2.12). This is unexpected given that SpaO$_C$ is the smallest subunit within the complex and that monomeric SpaO$_C$ was ejected in CID MS/MS experiments on SpaO-2SpaO$_C$ complexes in the absence of OrgB (Bernal, Börnicke, et al., 2019). In addition, OrgB was able to dissociate without the simultaneous loss of other subunits, indicating that each of the SpaO-2SpaO$_C$ heterotrimers must be connected to both its respective OrgB and the second SpaO-2SpaO$_C$ heterotrimer so as not to dissociate from the complex together with its associated OrgB. Similarly, in order for SpaO to dissociate without the simultaneous loss of 2SpaO$_C$, the SpaO$_C$ dimer requires a second point of attachment within the complex (Figure 2.13).

Together, these results show that the presence of OrgB not only leads to a stabilization of SpaO$_C$ within the complex, but also promotes direct interactions between SpaO-2SpaO$_C$ heterotrimers within 2(SpaO-2SpaO$_C$)-2OrgB.

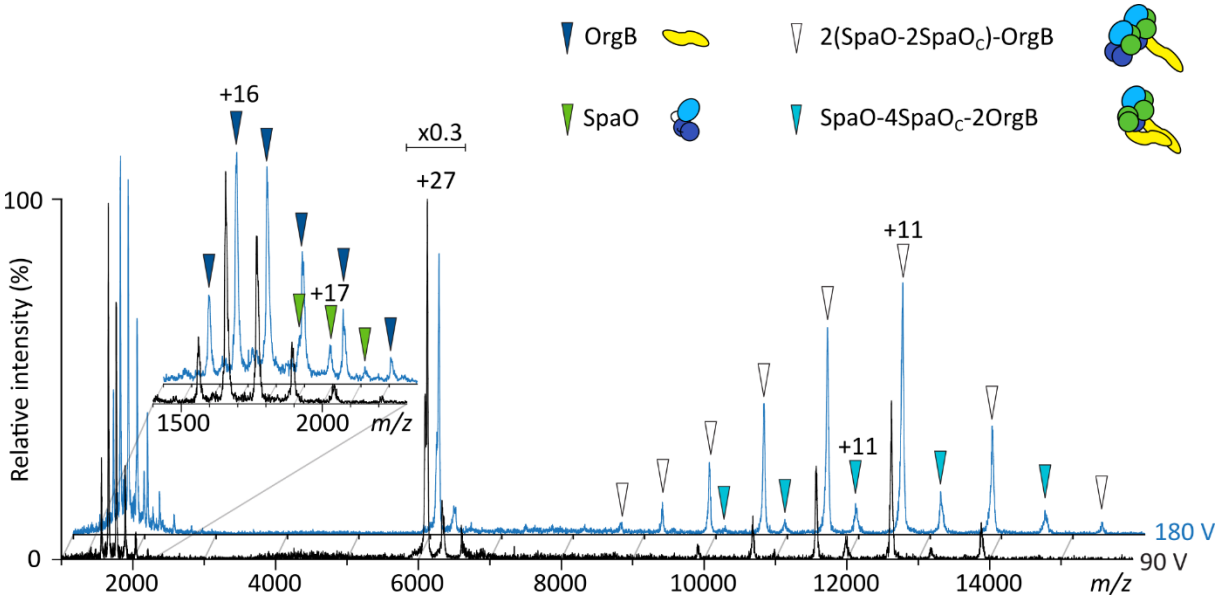


Figure 2.12. MS/MS analysis of the 2(SpaO-2SpaO$_C$)-2OrgB complex. The +27 charged complex in the spectrum of SpaO/SpaO$_C$/OrgB/InvC (Figure 2.24) was subjected to MS/MS analysis at acceleration voltages of 90 V (black spectrum) and 180 V (blue spectrum). The precursor peak in the 90 V spectrum was scaled down to 30% of its original size. Theoretical and experimentally determined molecular masses are given in Table 2.1. No affinity-tags were present on the proteins. MS analysis was performed by Johannes Heidemann.

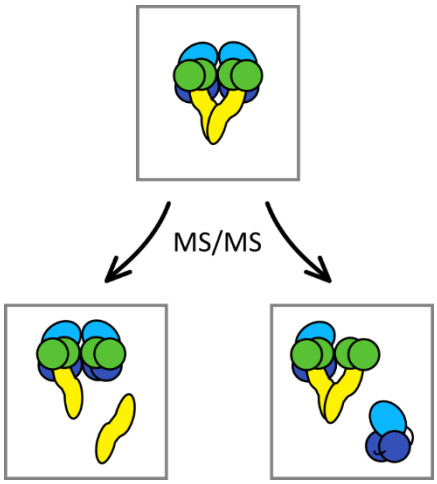


Figure 2.13. Schematic drawing of the architecture of the 2(SpaO-2SpaO$_C$)-2OrgB complex based on MS/MS. Direct interactions between SpaO-2SpaO$_C$ heterotrimers allow for the dissociation of OrgB or SpaO from the complex without the simultaneous loss of other subunits (Figure 2.12).

2.4. Characterization of SpaO/SpaO$_C$/OrgB/InvC complexes

2.4.1. OrgB and InvC domains involved in SpaO/SpaO$_C$/OrgB/InvC complex formation

The largest soluble sorting platform sub-complex purifiable in amounts suitable for *in vitro* characterization contained the proteins SpaO, SpaO$_C$, OrgB and InvC (Figure 2.7). Because a previous study reported that the N-terminus of OrgB is responsible for its interaction with SpaO/SpaO$_C$ (Notti et al., 2015), it was hypothesized that its C-terminus is involved in interactions with InvC. To test this theory, His-tagged constructs encompassing the N-terminal or C-terminal halves of OrgB (amino acids 1-105 and 106-226, respectively) were tested for their ability to co-purify with Strep-tagged SpaO/SpaO$_C$ or InvC in Strep-Tactin affinity purification following co-expression in *E. coli*. Analysis of purified complexes showed that the N-terminal fragment OrgB$_{1-105}$ co-purified with SpaO/SpaO$_C$, while InvC pulled down minute amounts that were only detectable by western blotting (Figure 2.14A). In contrast, OrgB$_{106-226}$ co-purified with InvC, and while low levels of this construct were also pulled down by SpaO/SpaO$_C$, the relative amounts of SpaO/SpaO$_C$ and OrgB$_{106-226}$ indicate that the interaction between these proteins is weak in comparison.

The region of InvC that interacts with OrgB was identified using a similar strategy. First, it was observed that when complexes of Strep-tagged InvC and His-tagged OrgB were purified using Strep-Tactin, bands for both full-length InvC and a smaller variant appeared in western blots detecting the C-terminal Strep-Tag on InvC. In contrast, only full-length InvC was detected after purification by nickel immobilized metal ion affinity chromatography (IMAC) using the His-tag on OrgB, indicating that the shorter, N-terminally truncated form of InvC had lost its ability to interact with OrgB (Figure 2.14B). Based on this observation, InvC was dissected after its proposed N-terminal domain, resulting in the fragments InvC$_{1-79}$ and InvC$_{80-431}$. Pull-down assays using Strep-tagged forms of these constructs showed that OrgB co-purified with InvC$_{1-79}$ but not with InvC$_{80-431}$ (Figure 2.14C), demonstrating that the interaction with OrgB is mediated by the N-terminal domain of InvC.

Together, these results show that the ATPase InvC interacts through its N-terminal domain with the C-terminus of OrgB and confirm the interaction of the OrgB N-terminus with SpaO/SpaO$_C$ (Notti et al., 2015).

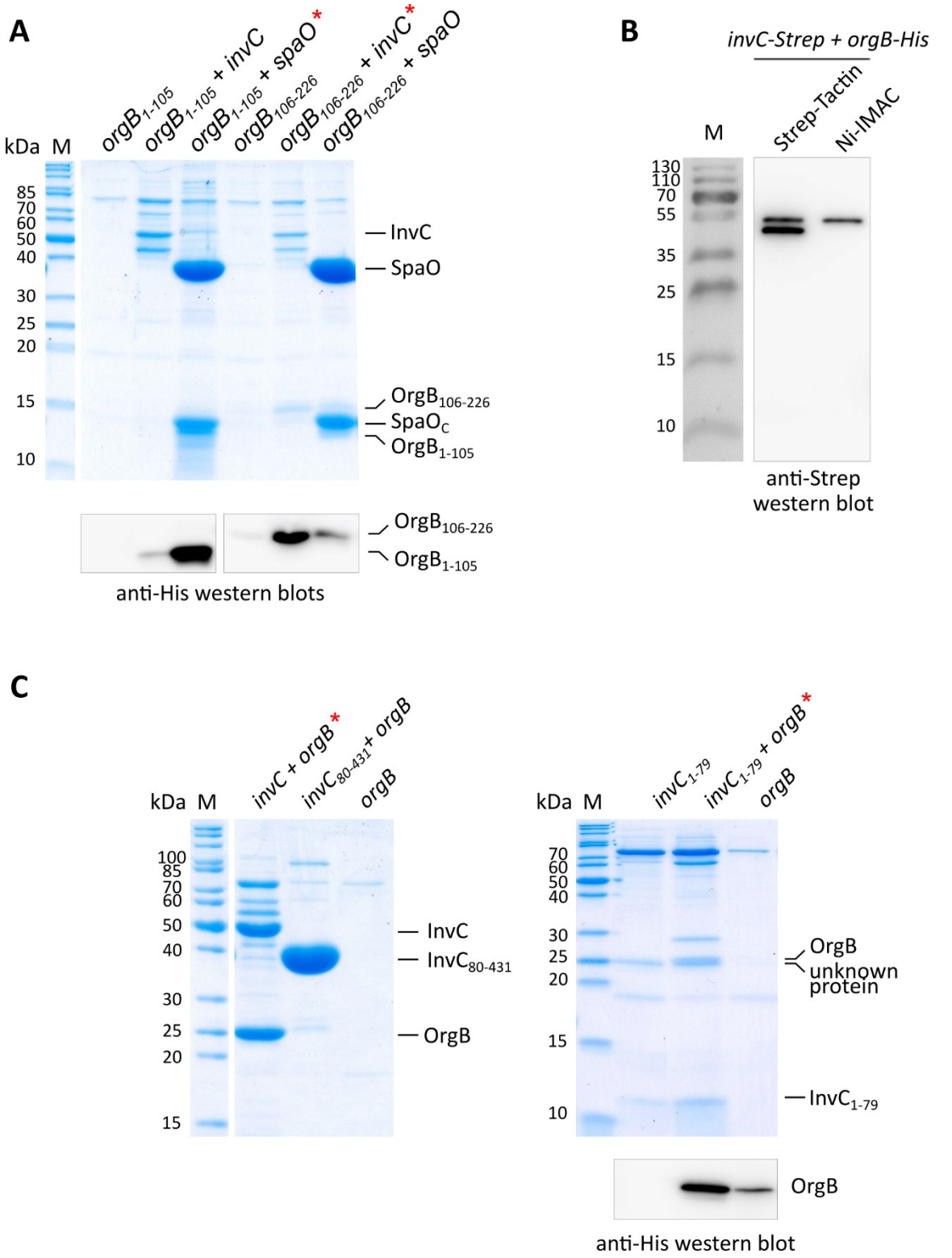


Figure 2.14. Protein domains involved in the formation of soluble complexes of SpaO, SpaO$_C$, OrgB and InvC. A) Coomassie-stained SDS-PAGE (top) and western blots (bottom) of His-tagged OrgB fragments pulled down by Strep-tagged SpaO and InvC. Combinations showing significant co-purification are indicated by red asterisks. **B)** Anti-Strep western blot of InvC-Strep/OrgB-His complexes purified by Strep-Tactin or nickel IMAC. **C)** Coomassie-stained SDS-PAGE (top) and western blot (bottom) of His-tagged OrgB pulled down by different Strep-tagged InvC fragments.

2.4.2. Development of a purification strategy for SpaO/SpaO$_C$/OrgB/InvC complexes

SpaO/SpaO$_C$/OrgB/InvC complexes were purified with the aim of their biophysical characterization and crystallization for structural studies, which require high amounts of pure and homogenous protein. However, this complex showed some unusual behavior that posed a challenge to the development of a purification strategy.

2.4.2.1. SpaO/SpaO$_C$/OrgB/InvC co-purifies with DnaK

For the purification of SpaO/SpaO$_C$/OrgB/InvC complexes, the genes *invC*-Strep, *orgB* and *spaO* were co-expressed in *E. coli* in 1.5-2.5-liter cultures and purified by Strep-Tactin affinity purification. This yielded a protein concentration of approximately 1 mg/ml at the highest point of the elution peak and ~4-5 mg of total eluting protein (both estimated from A_{280}). SDS-PAGE showed that the desired proteins co-purify with the expected InvC degradation product (see Figure 2.4, Figure 2.14), as well as two proteins that were identified by matrix-assisted laser desorption ionization MS (data not shown) as the chaperone DnaK and the biotin carboxyl carrier BCCP (Figure 2.15, first sample lane). While the proteins of interest could be separated from the smaller contaminations by Superose 6 size-exclusion chromatography in a low-salt buffer (10 mM Tris-HCL pH 8.0, 50 mM NaCl), DnaK appeared to be tightly bound and co-eluted with SpaO/SpaO$_C$/OrgB/InvC in a peak at ~14.1-14.2 ml (Figure 2.16). In addition, a peak with a high A_{280} was present at 19.77 ml, indicating that in Strep-Tactin purification small, unidentified molecules can co-purify with SpaO/SpaO$_C$/OrgB/InvC.

Because the contamination with DnaK precluded further analysis of the complex of interest, several strategies for DnaK removal during Strep-Tactin affinity purification were attempted. Since DnaK is an ATP-dependent chaperone, column-bound proteins were washed with buffers containing ATP-$MgCl_2$ or ATP-$MgCl_2$ and denatured proteins, and the inclusion of 10% glycerol in purification buffers was tested (Figure 2.15). While the addition of wash steps with ATP-$MgCl_2$ or denatured proteins did not have an effect on sample purity, the addition of 10% glycerol to purification buffers lead to a significant reduction in co-purifying DnaK. However, treatment of the samples with glycerol also reduced the yield of purified complexes by ~70% (data not shown), making it unsuitable for large-scale purification.

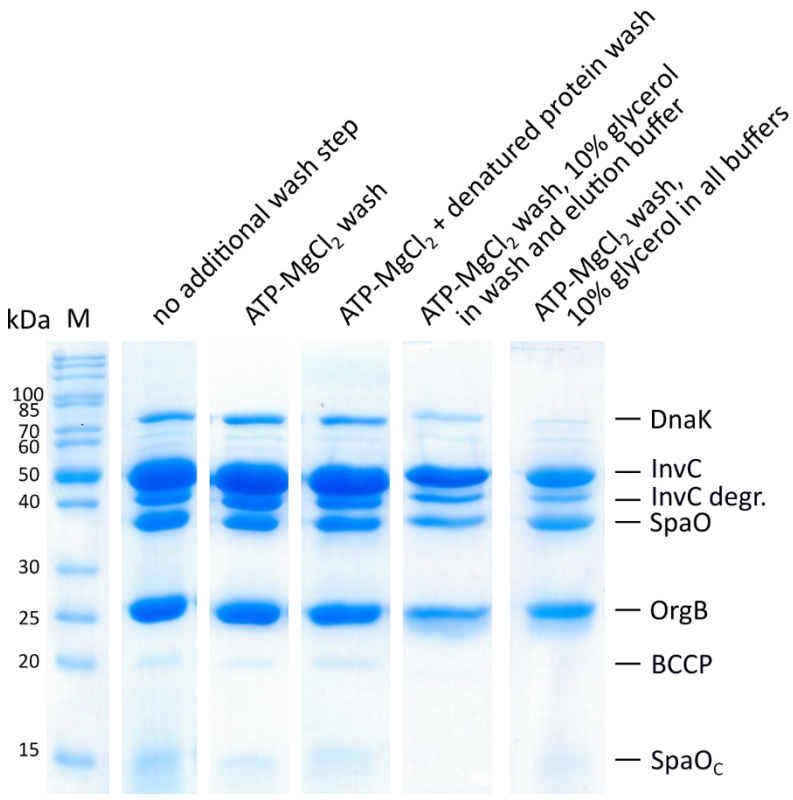


Figure 2.15. Effects of additional wash steps during Strep-Tactin affinity purification of SpaO/SpaOc/OrgB/InvC complexes. Coomassie-stained SDS-PAGE of complexes purified by Strep-Tactin affinity chromatography using a C-terminal Strep-tag on InvC. Additional wash steps and the additives included in different buffers are indicated above the lanes. "InvC degr." denotes the InvC degradation product.

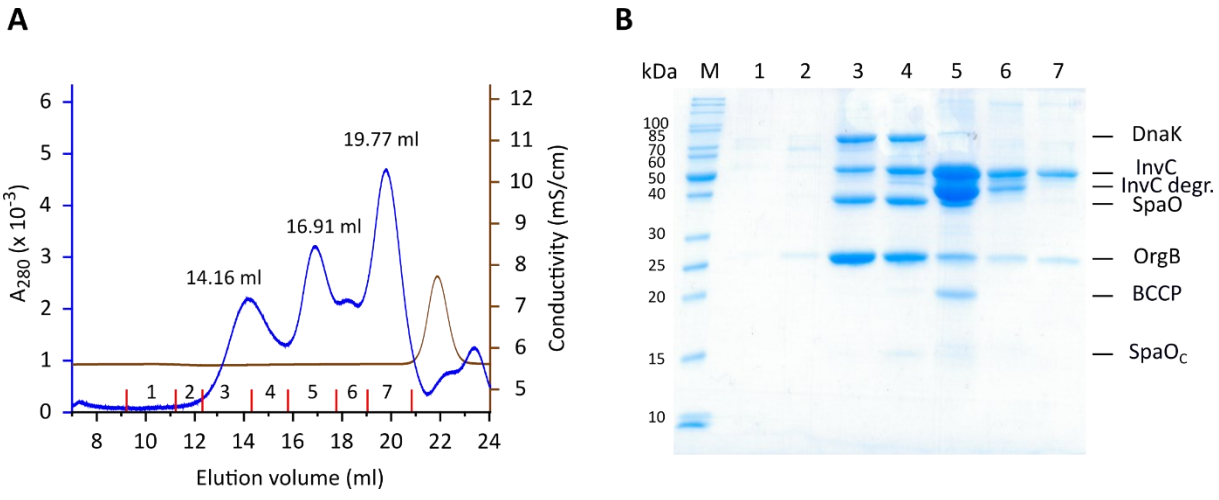


Figure 2.16. SEC analysis of SpaO/SpaOc/OrgB/InvC complexes. A) SEC elution profile of affinity-purified complexes. Collected elution fractions are denoted by red lines and the elution volumes of peaks are indicated. The small peak at approx. 18.25 ml appears to be an artifact in this particular experiment and was not found in other preparations. **B)** Coomassie-stained SDS-PAGE of elution fractions collected in A) and concentrated by TCA precipitation. Sample volumes were normalized by A_{280} of the elution fraction.

2.4.2.2. Dialysis of SpaO/SpaO$_C$/OrgB/InvC samples causes the formation of high-molecular-mass species

A possible strategy for the purification of DnaK-free complexes presented itself when affinity-purified samples were dialyzed against the low-salt buffer (10 mM Tris pH 8.0, 50 mM NaCl) prior to SEC. This resulted in the generation of high-molecular-mass complexes that eluted at ~11 ml in SEC and contained the proteins SpaO, OrgB and InvC without any associated DnaK (Figure 2.17). SpaO$_C$ appears to be missing from these complexes, although this is difficult to conclude with certainty since low levels of SpaO$_C$ are not always detectable by Coomassie-staining due to its small size and low dye binding capability. Furthermore, the larger width and the elution volume of the peak indicate the presence of a mixture of species with molecular masses in the low megadalton range, raising the possibility that these complexes could represent a mixture of different higher-order SpaO/(SpaO$_C$)/OrgB/InvC complexes such as could be expected to appear during the assembly of a full sorting platform. Similarly, SpaO/OrgB/InvC complexes with greatly reduced levels of DnaK could be obtained from dialyzed samples in ion-exchange chromatography (IEX), (Figure 2.18, fractions 9-11). Here, the shape of the elution peak indicates the presence of at least two molecular species eluting very close together, with SDS-PAGE suggesting that one difference between them could be the inclusion or absence of SpaO$_C$.

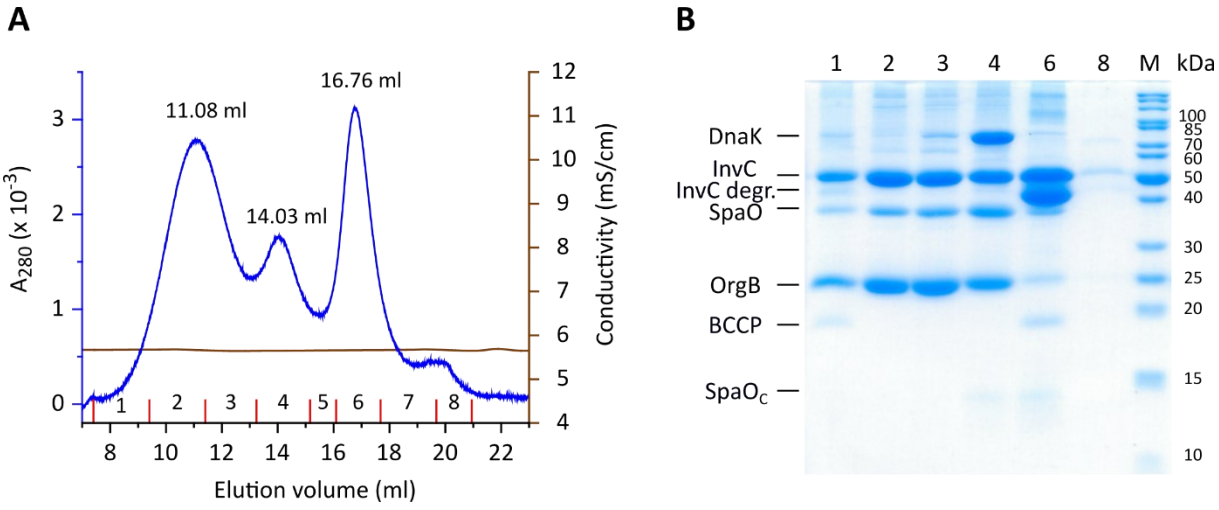


Figure 2.17. SEC analysis of SpaO/SpaO$_C$/OrgB/InvC complexes after dialysis against low-salt buffer. A) SEC elution profile of affinity-purified complexes dialyzed against 10 mM Tris-HCl pH 8.0, 50 mM NaCl. Collected elution fractions are denoted by red lines and the elution volumes of peaks are indicated. **B)** Coomassie-stained SDS-PAGE of elution fractions collected in A) and concentrated by TCA precipitation.

However, even though dialysis caused the appearance of DnaK-free SpaO/SpaO$_C$/OrgB/InvC complexes in the samples, the generated molecular species appeared to be only stable in buffers of low ionic strength (10 mM Tris-HCl pH 8.0, 50 mM NaCl). For example, in IEX the complexes of interest eluted at a NaCl concentration of 300-400 mM (Figure 2.18), and subsequent SEC analysis showed that at these salt concentrations a significant portion of the protein complex had aggregated and eluted in the void volume of the SEC column (Figure 2.19A, B). Similarly,

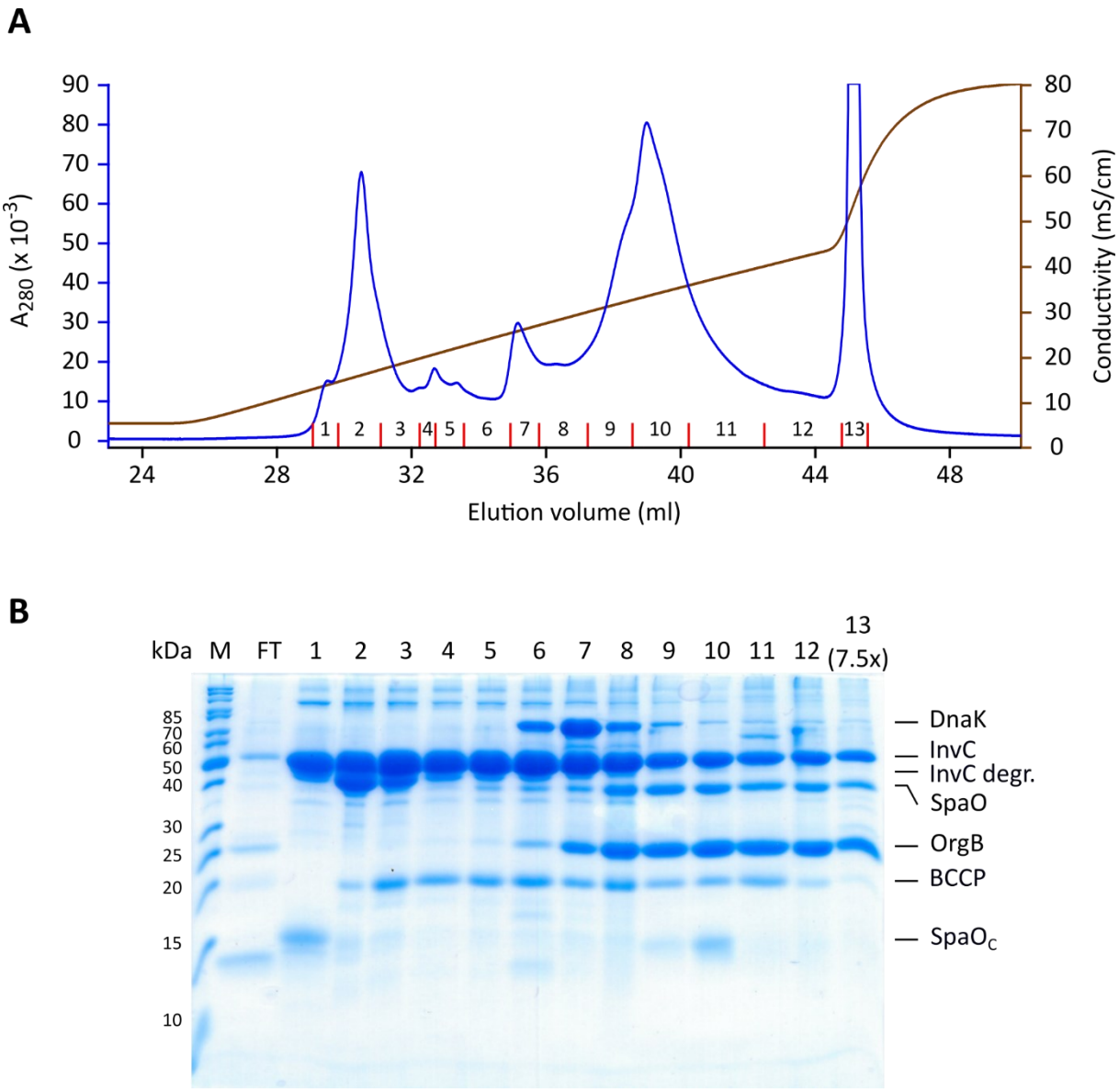


Figure 2.18. Anion-exchange chromatography of SpaO/SpaO$_C$/OrgB/InvC complexes after dialysis against low-salt buffer. A) IEX elution profile of affinity-purified complexes dialyzed against 10 mM Tris-HCL pH 8.0, 50 mM NaCl. Proteins were eluted with a gradient of NaCl, collected elution fractions are indicated by red lines. **B)** Coomassie-stained SDS-PAGE of elution fractions collected in A) and concentrated by TCA precipitation. Loaded sample volumes were normalized by A_{280} and a 7.5x excess was loaded for fraction 13, whose A_{280} signal was mostly due to a large overlapping absorption peak at 260 nm. FT = Flow-through.

when affinity-purified samples where dialyzed into a buffer of higher ionic strength (100 mM Tris-HCl pH 8.0, 150 mM NaCl; also used during Strep-Tactin purification), SEC revealed a much higher heterogeneity among the higher-mass-complexes (i.e. no bell-shaped peak), as well as significant protein aggregation (Figure 2.19C, D). Both the heterogeneity and aggregation upon exposure to higher salt concentrations meant that the DnaK-free complexes generated by dialysis were unsuitable for many types of analysis, especially structural analysis by X-ray crystallography.

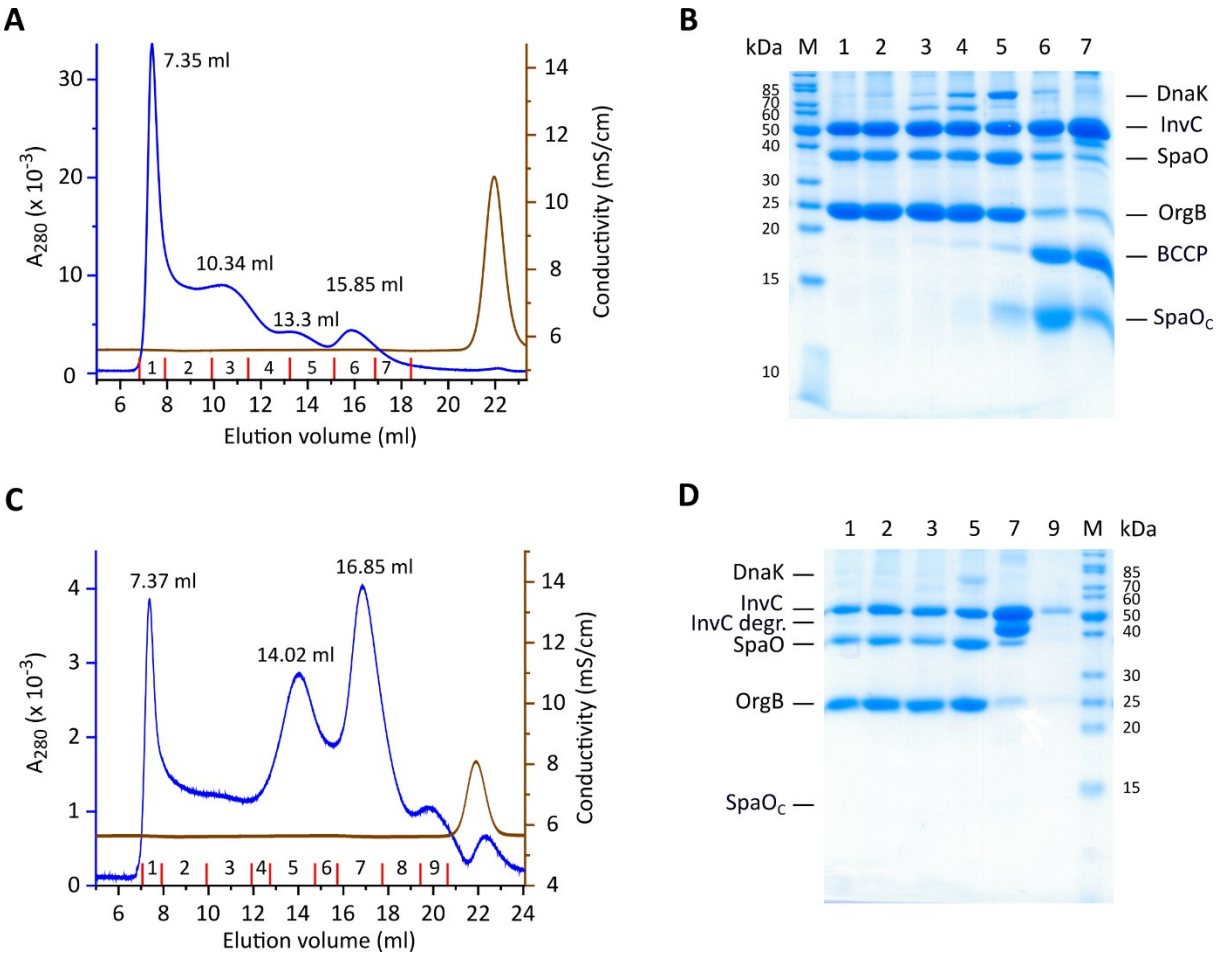


Figure 2.19. SEC analysis of SpaO/SpaO$_C$/OrgB/InvC complexes subjected to dialysis and higher salt concentrations. A) SEC elution profile of affinity-purified complexes dialyzed against 10 mM Tris-HCl pH 8.0, 50 mM NaCl and purified by IEX chromatography. Elution fractions 9, 10 and 11 of the IEX depicted in Figure 2.18A were pooled and concentrated prior to SEC in 10 mM Tris-HCl pH 8.0, 50 mM NaCl. Elution fractions are indicated by red lines. **B)** Coomassie-stained SDS-PAGE of elution fractions collected in A) and concentrated by TCA precipitation. **C)** SEC elution profile of affinity-purified complexes dialyzed against 100 mM Tris-HCl pH 8.0, 150 mM NaCl. **D)** Coomassie-stained SDS-PAGE of elution fractions collected in C) and concentrated by TCA precipitation.

2.4.2.3. SpaO/SpaO$_C$/OrgB/InvC complexes form at higher protein concentrations

Finally, it was determined that the apparent complex of SpaO/SpaO$_C$/OrgB/InvC/DnaK observed in the SEC experiments was in fact a mixture of smaller sub-complexes formed by these five proteins and that SpaO/SpaO$_C$/OrgB/InvC complexes only assemble at higher protein concentrations. Importantly, these complexes do not include DnaK and thus could be partially separated from DnaK by SEC after sample concentration. By increasing the starting material to cells from 8 liters of expression culture, the yield of Strep-Tactin purification increased approximately 4-fold to ~15-25 mg, and the concentration of eluting proteins rose to ~8 mg/ml at the highest point of the elution peak (data not shown). In SEC this increase in protein

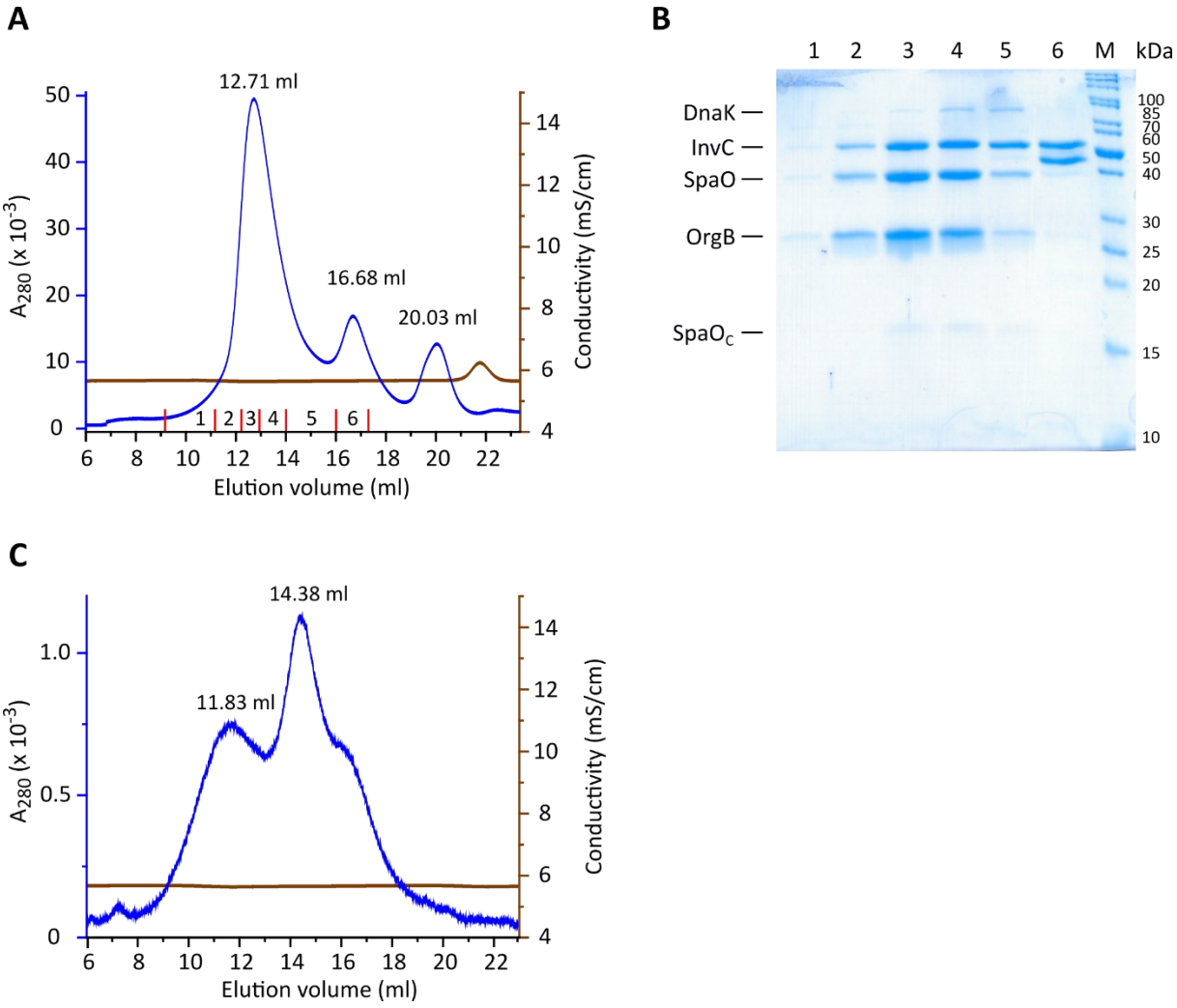


Figure 2.20. SEC analysis of SpaO/SpaO$_C$/OrgB/InvC complexes at higher protein concentration. **A)** SEC elution profile of affinity-purified complexes after concentrating to A_{280} of 21. Elution fractions are denoted by red lines and elution volumes of peaks are indicated. **B)** Coomassie-stained SDS-PAGE of elution fractions collected in A). **C)** SEC profile of proteins eluting in fraction 3 in A) re-subjected to SEC analysis without prior concentrating.

concentration was accompanied by a shift of the SpaO/SpaO$_C$/OrgB/InvC-containing peak to a lower elution volume, and by raising the concentration to 20-30 mg/ml using centrifugal concentrators, elution volumes of 12.5-12.7 ml were reached (Figure 2.20A-B, Figure 2.21A-B; see Figure 2.16 and Figure 2.17 for comparison). This shift was also accompanied by a relative decrease in height of the peak containing monomeric InvC (elution volume 16.7-16.9 ml) and a disappearance of the discrete peak at ~14 ml, suggesting that with increasing protein concentration the proteins contained in these peaks assemble into larger complexes that elute at ~12.6 ml. In fact, this conclusion is supported by a later multi-angle light scattering experiment indicating a molecular mass of approximately 100 kDa for species eluting at a volume

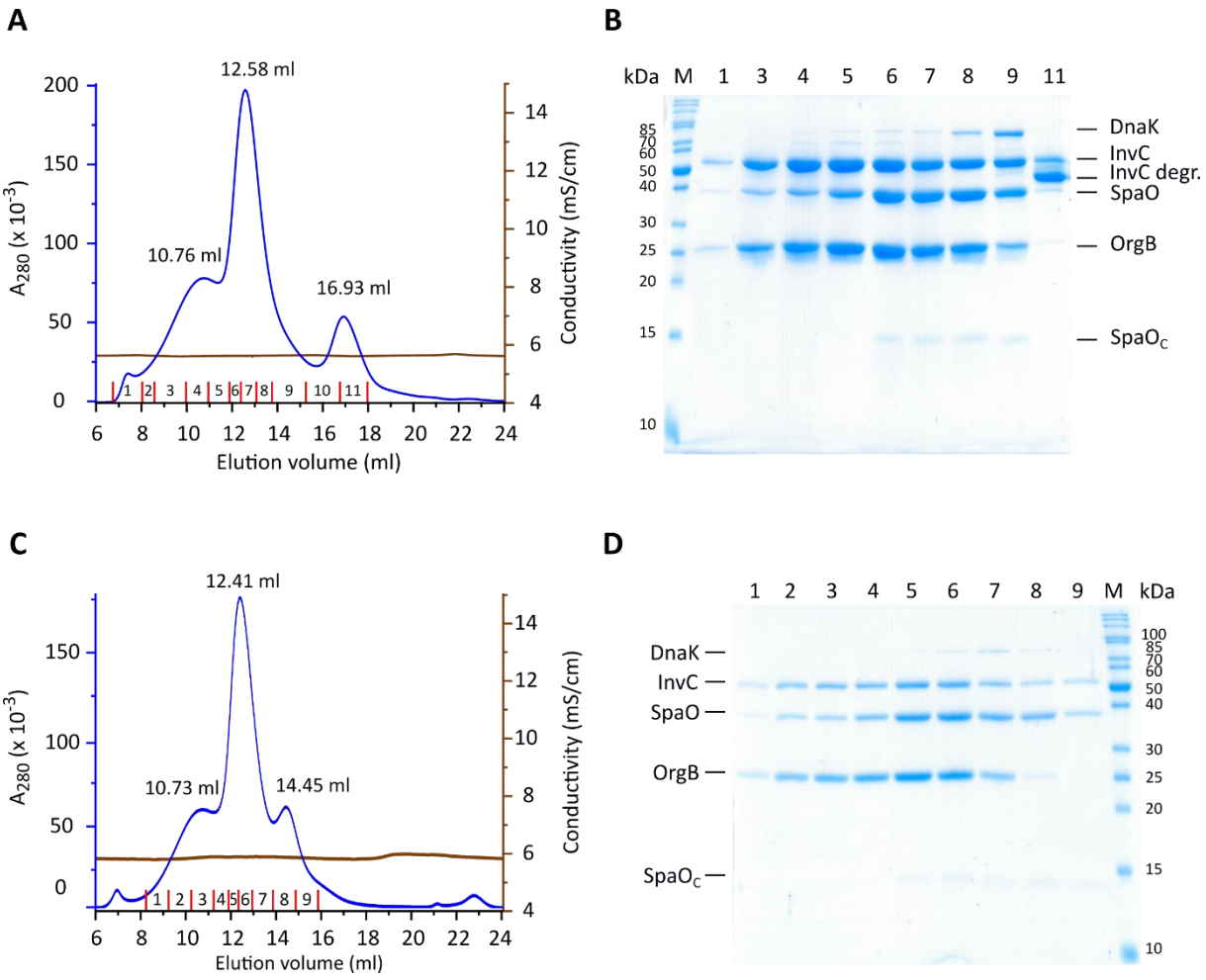


Figure 2.21. SEC analysis of SpaO/SpaO$_C$/OrgB/InvC complexes at higher protein concentration after dialysis against low-salt buffer. A) SEC elution profile of affinity-purified complexes after dialysis against 10 mM Tris-HCl pH 8.0, 50 mM NaCl and concentrating to A_{280} of 30. Elution fractions are denoted by red lines and elution volumes of peaks are indicated. **B)** Coomassie-stained SDS-PAGE of elution fractions collected in A). **C)** SEC profile of proteins collected in elution fractions 6, 7 and 8 in A). Elution fractions were pooled and concentrated to A_{280} of 3.4 before SEC. The chromatogram was obtained on an HPLC system with a more sensitive UV detector, meaning that for comparability with other chromatograms the measured A_{280} values need to be divided by a factor of ~3. **D)** Coomassie-stained SDS-PAGE of elution fractions collected in C).

of ~14 ml (Figure 2.23), confirming that a peak at this position cannot contain a single large complex of SpaO, SpaO$_C$, OrgB, InvC and DnaK, which have a combined mass of 189 kDa. Importantly, while the proteins of interest formed a larger complex and shifted to a lower elution volume, DnaK still eluted at 13-15 ml, meaning that the first half of the shifted SpaO/SpaO$_C$/OrgB/InvC peak was almost free of contaminations and thus appeared suitable for further characterization. However, subjecting these purified SpaO/SpaO$_C$/OrgB/InvC complexes to a second round of SEC revealed that after the first SEC lower-mass species of InvC and SpaO had been formed, as well as high-molecular mass species similar to those generated by dialysis (Figure 2.20C, Figure 2.21C; compare with Figure 2.17A). This indicates that such larger complexes can result not only from dialysis but from any form of buffer exchange, which in turn suggests that their formation is due to the loss of an unknown factor that is co-purified with the proteins during Strep-Tactin purification but removed by later purification steps. In fact, the dialysis of affinity-purified samples generally caused the disappearance of an SEC peak eluting at ~19.5-20.0 ml (Figure 2.17, Figure 2.19), indicating that the factor in question could be a small molecule with significant absorption at 280 nm, a description that fits ATP. However, the addition of ATP-$MgCl_2$ did not prevent the formation of larger complex species during dialysis (Figure 2.22).

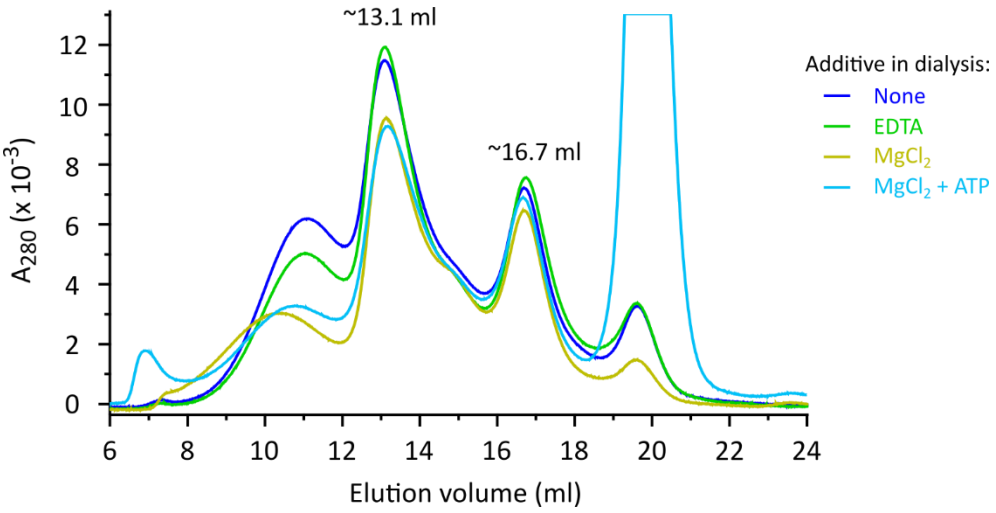


Figure 2.22. SEC elution profiles of SpaO/SpaO$_C$/OrgB/InvC complexes dialyzed against low-salt buffer supplemented with different additives. Affinity-purified complexes were dialyzed against 10 mM Tris-HCl pH 8.0, 50 mM NaCl buffer supplemented with 1 mM EDTA, 2 mM $MgCl_2$, or 2 mM $MgCl_2$ + 0.5 mM ATP and analyzed by SEC.

In conclusion, even though SpaO/SpaO$_C$/OrgB/InvC complexes of sufficient purity could be obtained by affinity purification and SEC, the formation of larger molecular species with each buffer exchange meant that no homogenous samples could be obtained by chromatographic methods. Combined with this heterogeneity, the instability and aggregation of the larger species in the presence of higher salt concentrations limited the techniques available for characterization of the complex, preventing structural analysis by X-ray crystallography in particular. However, it was possible to analyze the heterogenous mixture by native MS, and since the kinetics of the described sample changes appeared slow enough for the SpaO/SpaO$_C$/OrgB/InvC complex to remain intact during SEC, it could be characterized using SEC-coupled methods like SEC-SAXS and SEC-MALS.

2.4.3. Molecular mass analysis of SpaO/SpaO$_C$/OrgB/InvC by MALS

SEC-MALS analysis of purified and concentrated SpaO/SpaO$_C$/OrgB/InvC complexes revealed a rather uniform molecular mass of approximately 207 kDa in the first half of the major elution peak, which decreased to approximately 180 kDa in the later regions of the peak (Figure 2.23). This range in molecular mass shows that the SEC peak contains a mixture of molecular species and thus indicates that SpaO/SpaO$_C$/OrgB/InvC is a dynamic system, in which subunits may associate and dissociate to form different sub-complexes.

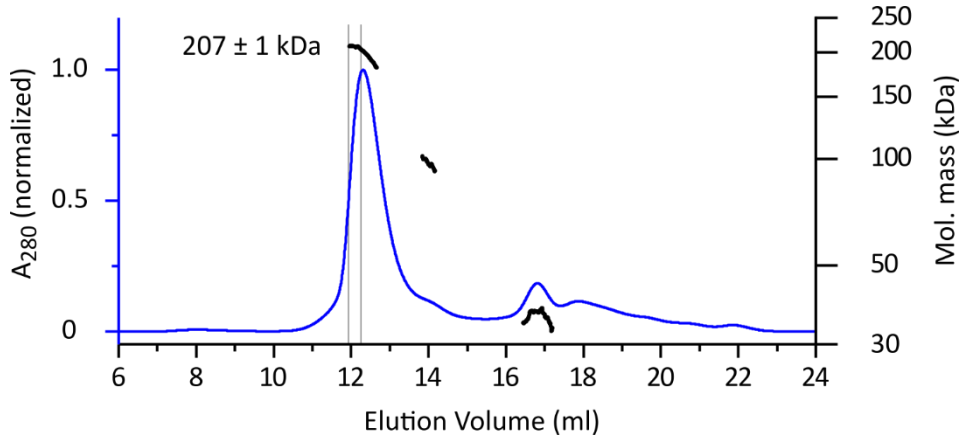


Figure 2.23. SEC-MALS analysis of SpaO/SpaO$_C$/OrgB/InvC complexes. Affinity-purified complexes were concentrated to A_{280} = 22 prior to SEC. The SEC elution profile (A_{280}, blue) and weight-averaged molecular masses across the elution peaks (black) are shown. The mean molecular mass and standard deviation of data points within the region marked by grey vertical lines are indicated. Molecular masses were calculated by Cy Jeffries.

2.4.4. Native mass spectrometry of SpaO/SpaO$_C$/OrgB/InvC complexes

In order to obtain precise information about the different SpaO/SpaO$_C$/OrgB/InvC complexes and their subunit stoichiometries, samples were subjected to analysis by native MS. The complexes were purified by affinity chromatography, followed by SEC directly into native MS-compatible ammonium acetate buffer to minimize the number of buffer exchanges experienced by the proteins and preserve sample quality.

This analysis showed that InvC is present in a number of different complexes (Figure 2.24). The most prominent of these are of 2(SpaO-2SpaO$_C$)-2OrgB-InvC (214 kDa) and SpaO-2SpaO$_C$-2OrgB-InvC (159 kDa) stoichiometry, while less intense peaks for species of 2SpaO-2SpaO$_C$-2OrgB-InvC (192 kDa) and 2OrgB-InvC (102 kDa) stoichiometry were also observed. These findings are not only in good agreement with the molecular masses determined by MALS, but also highlight the dynamic nature of these complexes and suggest that SpaO-2SpaO$_C$ heterotrimers in particular may be prone to exchange. Notably, this analysis detected InvC only in complexes containing OrgB dimers, supporting the earlier observation of an interaction between the OrgB C-terminus and InvC N-terminus (Figure 2.14). Interestingly, while the two OrgB molecules seem to be able to interact with one SpaO-2SpaO$_C$ heterotrimer each, only a single InvC was associated with OrgB dimers.

The ratio of larger and smaller complexes observed in the mass spectra of SpaO/SpaO$_C$/OrgB/InvC was dependent on the quality and stability of the electrospray, parameters that are influenced by e.g. small differences in the electrospray capillary used in a particular experiment. In general, conditions that produced the highest-resolution spectra also resulted in lower signal intensities for higher-mass complexes (Figure 2.24, Figure 2.25A). On the other hand, spectra showing a high abundance of larger complexes were of lower resolution, but in some cases additional signals emerged in the higher m/z range (Figure 2.25B). Even though it was not possible to clearly assign charge states to these peaks because of their low signal intensity and resolution, their m/z range and spacing suggest that they are produced by complexes with an approximate mass of 433 kDa, possibly representing dimers of 2(SpaO-2SpaO$_C$)-2OrgB-InvC.

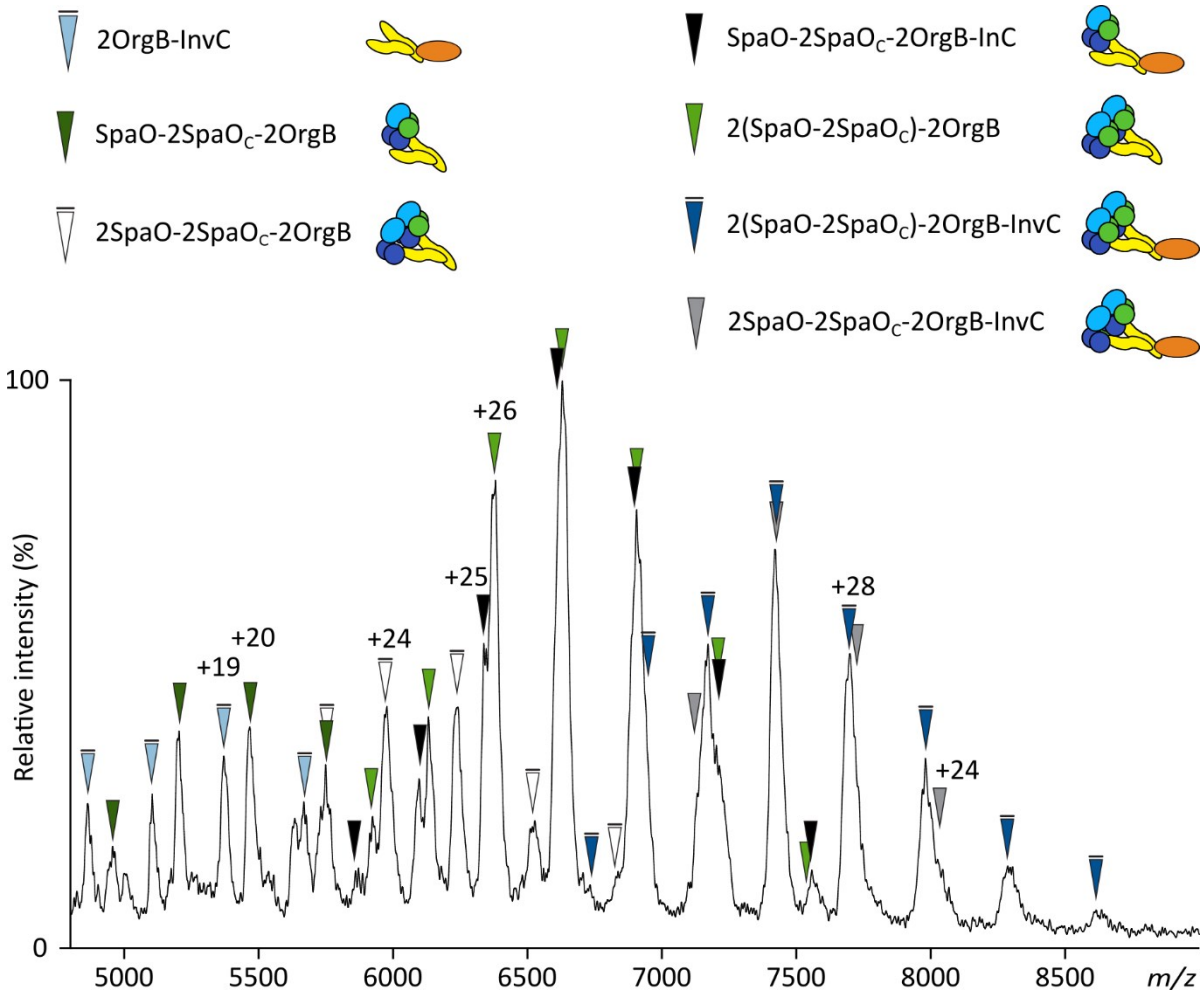


Figure 2.24. Native mass spectrometry analysis of SpaO/SpaO$_C$/OrgB/InvC complexes. SpaO/SpaO$_C$/OrgB/InvC complexes were purified by Strep-Tactin affinity purification and SEC in native-MS-compatible ammonium acetate buffer. The region of the spectrum showing the InvC-containing higher-mass complexes is depicted, the complete spectrum can be found in Figure 2.25A. Theoretical and experimentally determined molecular masses are given in Table 2.2. Species contained within overlapping peak series were identified and assigned with the help of MS/MS (Figure 2.28). InvC carries a C-terminal Strep-tag. MS analysis was performed by Johannes Heidemann.

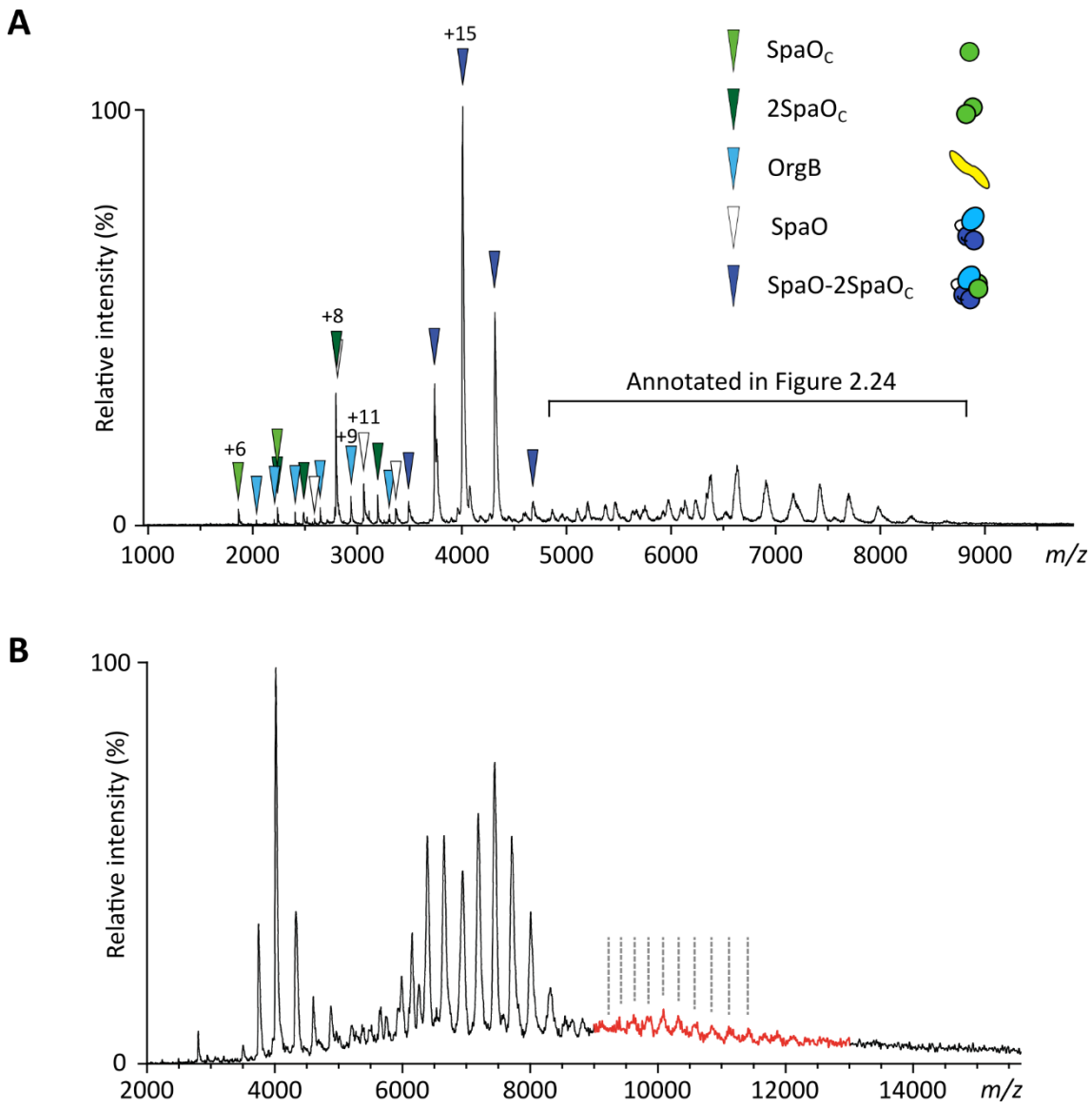


Figure 2.25. Variability in native mass spectra of SpaO/SpaO$_C$/OrgB/InvC complexes. A) Complete native mass spectrum of the partial SpaO/SpaO$_C$/OrgB/InvC spectrum shown in Figure 2.24**. B)** Depending on the conditions of the electrospray ionization, some SpaO/SpaO$_C$/OrgB/InvC spectra showed a higher abundance of high-molecular-mass complexes, as well as an additional high-*m*/*z* peak series (indicated in red). Dashed lines indicate peak positions. MS analysis was performed by Johannes Heidemann.

Next, it was tested whether oligomerization could be induced by the addition of a non-hydrolyzable ATP-analog, as previously reported for the hexamerization of the flagellar ATPase FliI (Claret et al., 2003; Kazetani et al., 2009). However, inclusion of ATPγS in the sample did not result in the appearance of higher-order oligomers (Figure 2.26), although it should be noted that the SpaO/SpaO$_C$/OrgB/InvC sample used in this experiment had previously been frozen for storage and contained a lower proportion of InvC-containing complexes. Therefore, this result needs to be confirmed using freshly prepared protein.

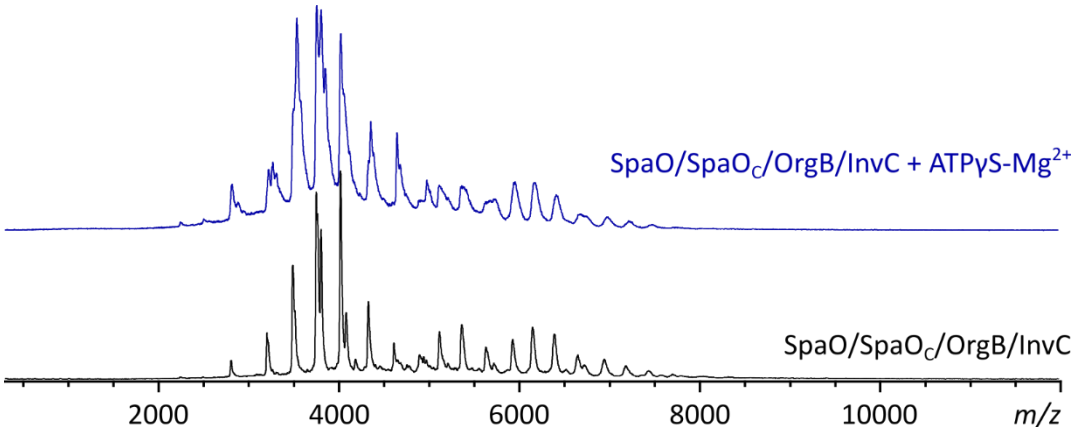


Figure 2.26. Native mass spectrometry of SpaO/SpaO$_C$/OrgB/InvC in the presence of a non-hydrolyzable ATP analog. Offset mass spectra of SpaO/SpaO$_C$/OrgB/InvC alone (black, bottom) and in the presence of 0.2 mM ATPγS + 1 mM Mg^{2+} (blue, top). Native MS was performed by Johannes Heidemann.

Finally, the possibility remained that the large complex species that formed during dialysis of SpaO/SpaO$_C$/OrgB/InvC and eluted at ~10-11 ml in SEC (Figure 2.21A) could be the result of higher-order oligomerization. However, native mass spectra showed no defined and interpretable peak series (Figure 2.27). The only signals detected were of a ~800-900 kDa complex from which a 57 kDa subunit dissociated in CID MS/MS (data not shown), probably corresponding to GroEL (van Duijn et al., 2006). Since not even single subunits of SpaO/SpaO$_C$/OrgB/InvC could be observed, this highlights the instability of these larger species and suggests that they are the product of partial protein aggregation and not of biologically relevant oligomerization.

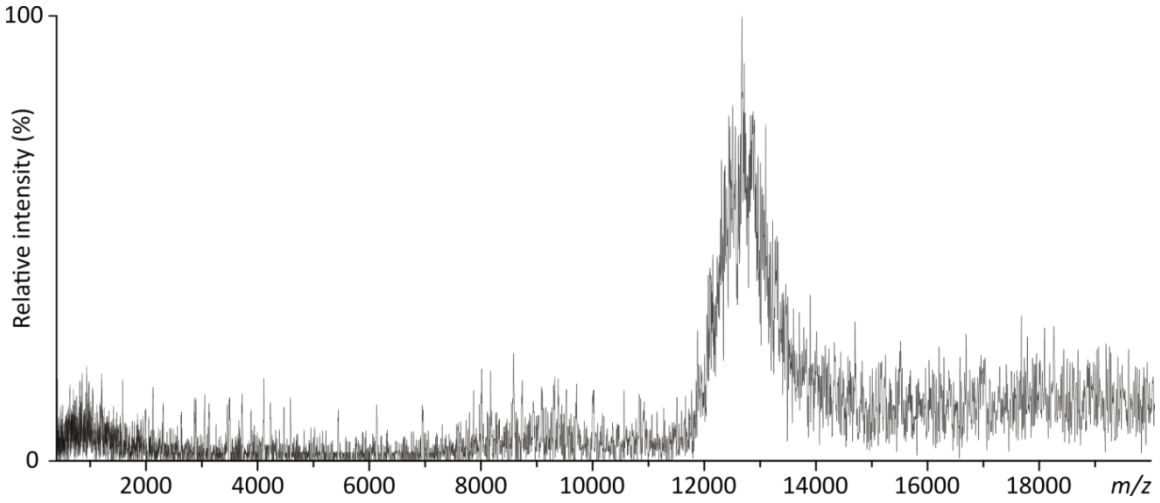


Figure 2.27. Native mass spectrum of SpaO/SpaO$_C$/OrgB/InvC complexes formed by dialysis and isolated by SEC. The spectrum was obtained from elution fraction 4 of the SEC depicted in Figure 2.21A. The signal in the 12,000-14,000 *m/z*-range is indicative of a species of 800-900 kDa. Native MS was performed by Johannes Heidemann.

Table 2.2. Theoretical masses and average experimental masses of SpaO/SpaO$_C$/OrgB/InvC proteins and protein complexes as determined by native MS. n≥3, unless otherwise indicated; STDEV: Standard deviation; Avg. FWHM: Average full-width at half-maximum.

Protein/-complex	Theoretical mass (Da)	Experimental avg. mass (Da)	STDEV (Da)	Avg. FWHM (Da)
2SpaO$_C$	22,351.0	22,349	5	15
OrgB MS/MS	26,448.4	26,459	21	100
InvC-Strep	48,808.9	48,240	110	370
SpaO/2SpaO$_C$	56,013.6	56,050	40	190
2OrgB/InvC-Strep	101,705.6	102,220	190	820
SpaO/2SpaO$_C$/2OrgB	109,041.5	109,230	100	890
2(SpaO/2SpaO$_C$)	112,027.2	112,480	170	530
SpaO/2SpaO$_C$/OrgB/InvC-Strep MS/MS	131,402.1	131,350	90	240
2(SpaO/2SpaO$_C$)/OrgB MS/MS	138,475.6	138,850	290	250
2SpaO/2SpaO$_C$/2OrgB	142,835.2	143,400	240	1,300
SpaO/2SpaO$_C$/2OrgB/InvC-Strep	157,850.4	158,800	400	1,300
2(SpaO/2SpaO$_C$)/2OrgB	164,923.9	165,370	160	910
2SpaO/2SpaO$_C$/OrgB/InvC-Strep MS/MS *	165,195.8	165,019	28	260
2(SpaO/2SpaO$_C$)/OrgB/InvC-Strep MS/MS	187,284.5	187,620	170	490
2SpaO/2SpaO$_C$/2OrgB/InvC-Strep	191,644.2	193,200	220	1,400
2(SpaO/2SpaO$_C$)/2OrgB/InvC-Strep	213,732.8	214,500	500	1,800

*n=2

2.4.5. Tandem mass spectrometry of SpaO/SpaO$_C$/OrgB/InvC complexes

In order to gain further information about the architecture of InvC-containing complexes, different SpaO/SpaO$_C$/OrgB/InvC species were subjected to CID MS/MS, which showed the dissociation of a single OrgB monomer in all cases (Figure 2.28). This dissociation pattern indicates an asymmetry within the OrgB dimer, with one OrgB monomer acting as a connector between SpaO/SpaO$_C$ and InvC, while the other is more loosely associated and can leave the complex without the simultaneous loss of other subunits. Furthermore, similar to the CID of 2(SpaO-2SpaO$_C$)-2OrgB (Figure 2.12), no dissociation of SpaO$_C$ was observed, showing that SpaO$_C$ is also protected in InvC-containing complexes. Interestingly, this was also observed for complexes containing only a single SpaO-2SpaO$_C$ heterotrimer (SpaO-2SpaO$_C$-2OrgB-InvC), indicating that the stabilization of SpaO$_C$ within OrgB-containing complexes is likely due to an interaction between SpaO$_C$ and OrgB, rather than to the dimerization of SpaO-2SpaO$_C$ heterotrimers (see section 2.3.3).

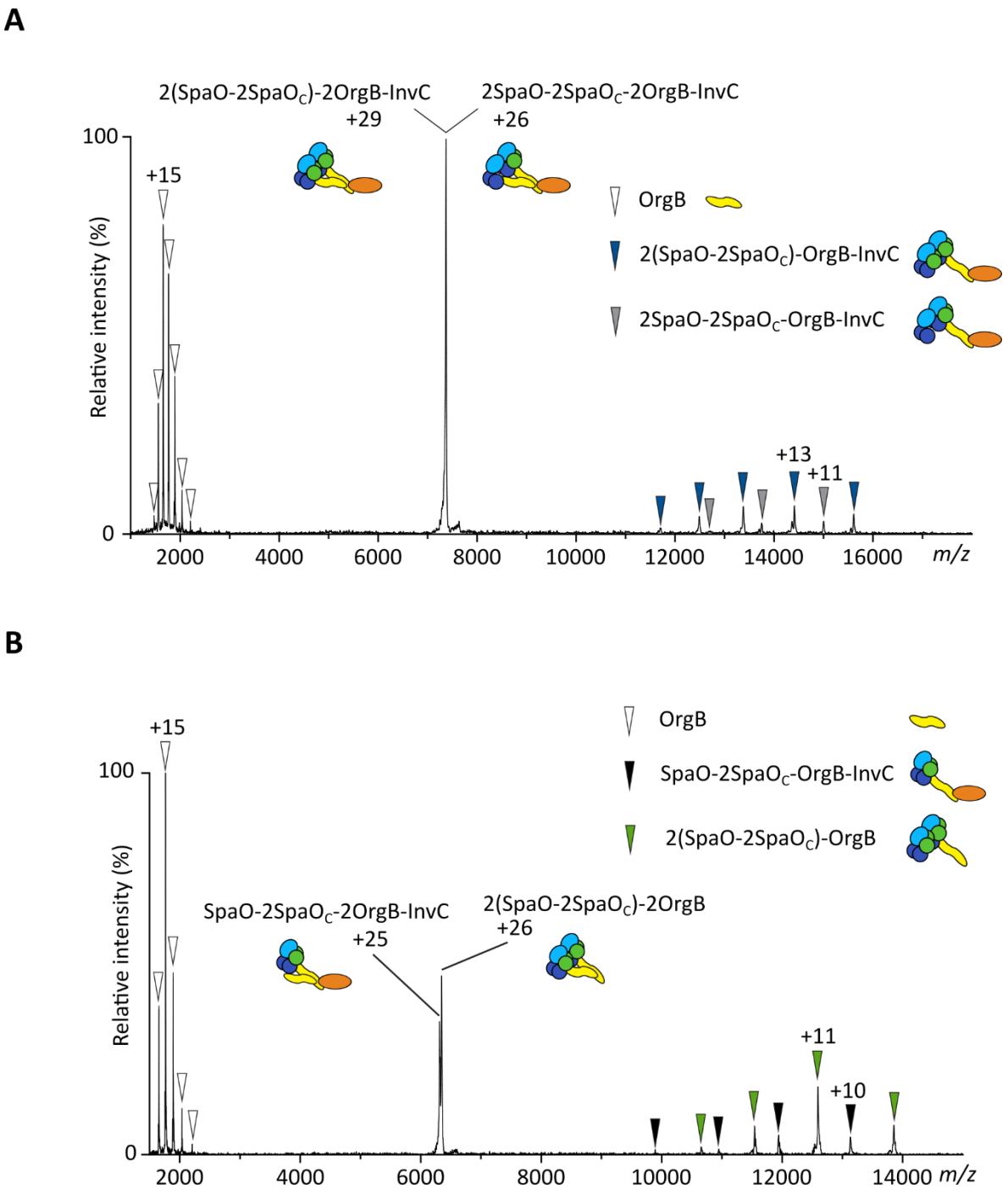


Figure 2.28. MS/MS of InvC-containing sorting platform sub-complexes. A) MS/MS of the +26 charge state of 2SpaO-2SpaO$_C$-2OrgB-InvC and the +29 charge state of 2(SpaO-2SpaO$_C$)-2OrgB-InvC. Both species were analyzed together due to peak overlap (Figure 2.24). Experimental and theoretical molecular masses are given in Table 2.2. **B)** MS/MS of the +25 charge state of SpaO-2SpaO$_C$-2OrgB-InvC complexes. The selected species was analyzed together with 2(SpaO-2SpaO$_C$)-2OrgB complexes due to peak overlap (Figure 2.24). MS/MS analysis was performed by Johannes Heidemann.

2.4.6. Small-angle X-ray scattering of SpaO/SpaO$_C$/OrgB/InvC

In order to further characterize SpaO/SpaO$_C$/OrgB/InvC, SEC-coupled small-angle X-ray scattering (SAXS) was employed to obtain structural information about these complexes. Given the heterogeneity of molecular species revealed by MALS in the later regions of the elution peak (Figure 2.23), only data from the first half of the peak was used for SAXS analysis to ensure a largely homogenous sample was measured. Nevertheless, it cannot be excluded that the resulting structure might be an average of different molecular species. An independent SEC of the sample used to acquire the SAXS data was run as a means of sample quality control and showed an elution profile very similar to that obtained in the MALS experiment (Figure 2.29).

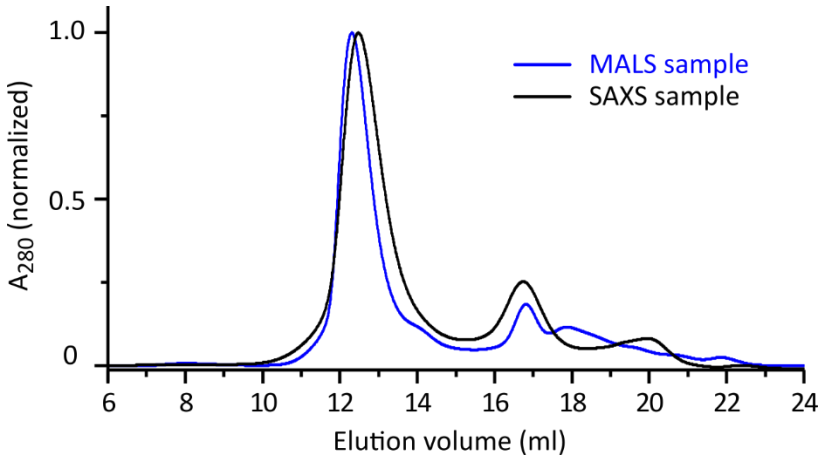


Figure 2.29. Comparison of SEC profiles of SpaO/SpaO$_C$/OrgB/InvC in MALS and SAXS analysis. Overlay of the chromatograms obtained in SEC-MALS analysis (blue, Figure 2.23) and a quality control SEC of the sample analyzed by SEC-SAXS (black, "SAXS sample"). Small differences between the chromatograms are most likely caused by differences in SEC conditions: SEC-MALS analysis was performed at room temperature and a flow-rate of 0.3 ml/min using an Agilent 1260 Infinity HPLC system, while the quality control SEC of the SAXS sample was performed at 4 °C using a flow-rate of 0.35 ml/min on an Äkta Purifier. For the acquisition of SAXS data, SEC was performed under the same conditions as for SEC-MALS.

Using the recorded SAXS data, an approach of simulated annealing of dummy atoms was used to generate an *ab initio* bead model whose theoretical scattering pattern best fits the experimental data. This model shows that the complex adopts an extended L-shape in solution (Figure 2.30A-C, Supplementary Table 1). Because scattering data was also available for the SpaO/SpaO$_C$ complex in the laboratory of Prof. Dr. Michael Kolbe, multi-phase modeling could be used to determine the position of SpaO/SpaO$_C$ within the larger SpaO/SpaO$_C$/OrgB/InvC complex and indicated that SpaO/SpaO$_C$ is located in the shorter arm of the L-shaped structure (Figure 2.30D).

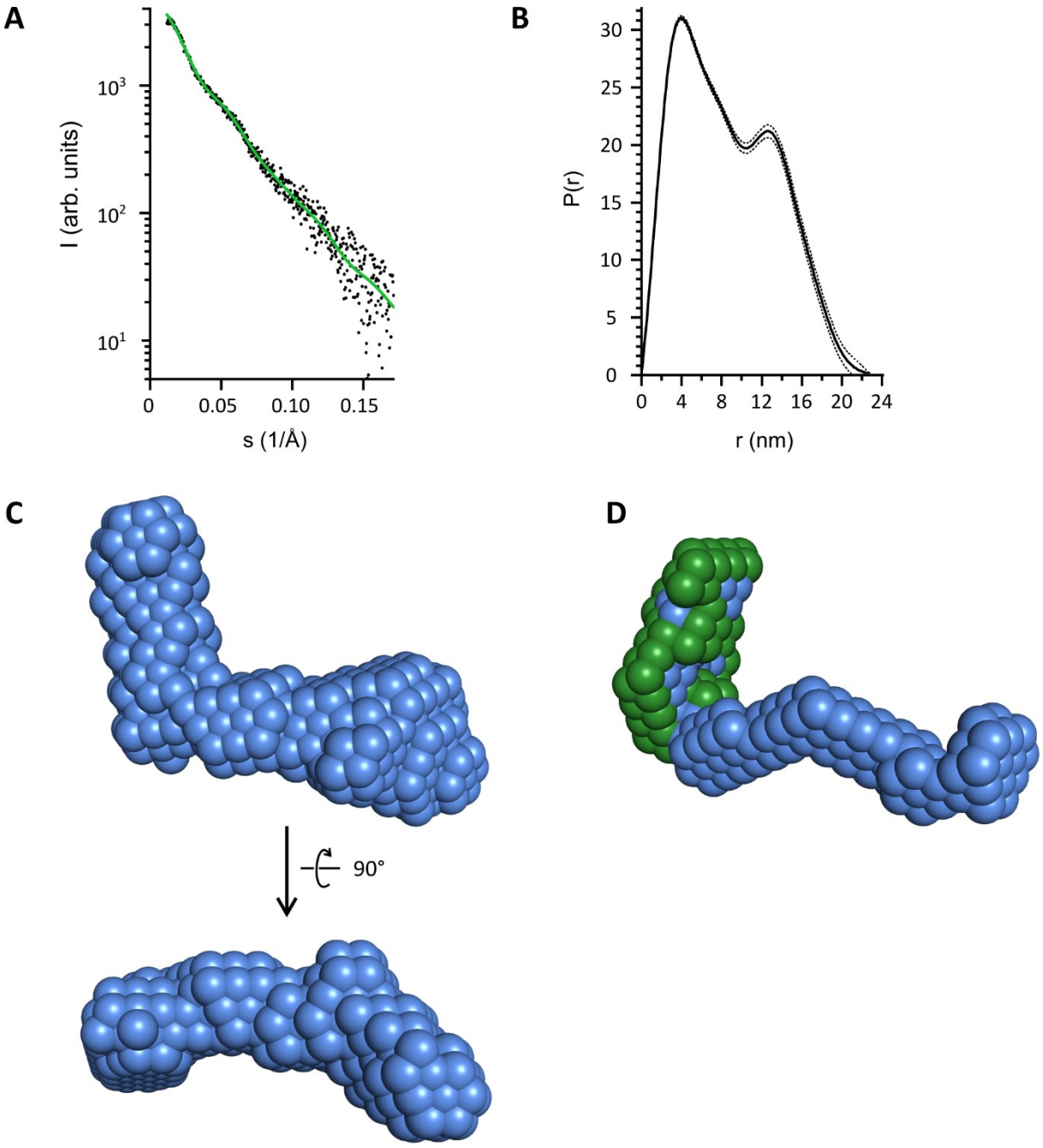


Figure 2.30. Small-angle X-ray scattering of SpaO/SpaOC/OrgB/InvC. A) Scattering profile of SpaO/SpaOC/OrgB/InvC. Experimental SEC-SAXS data is depicted as black dots, the green line represents the fit calculated from the SAXS hybrid model shown in Figure 2.31 (χ^2 = 1.12). **B)** Pair-distance distribution function P(r) computed from the SAXS data in A). **C)** Representative *ab initio* bead model in side view (top) and bottom view (bottom). **D)** Multiphase model using SAXS data of both SpaO/SpaOC and SpaO/SpaOC/OrgB/InvC. The phase corresponding to SpaO/SpaOC is colored in green. Details regarding data acquisition, analysis and modeling can be found in Supplementary Table 1. SAXS analysis was performed by Anne Tuukkanen.

The SAXS data was further combined with previously determined structural information to generate a SAXS hybrid model. In this approach, the crystal structures of $SpaO_C$ and the SpaO C-terminal domain dimer were combined with fragment-based models of OrgB, InvC and the SpaO N-terminal domain and arranged into an assembly that best fits the experimental SAXS data and takes into account known interactions between the components (see Materials & Methods section 4.6.4 for details). Similar to the *ab initio* and multi-phase bead models (Figure 2.30C, D), the rigid-body hybrid model shows an extended L-shape of the complex, with 2OrgB-InvC positioned in the longer and SpaO-2$SpaO_C$ in the shorter arm (Figure 2.31). Although MALS and native MS indicated that the measured complex was of 2(SpaO-2$SpaO_C$)-2OrgB-InvC stoichiometry, it was only possible to fit a complex containing a single SpaO-2$SpaO_C$ heterotrimer (i.e. SpaO-2$SpaO_C$-2OrgB-InvC) to the SAXS data. The reason for this discrepancy could lie in ambiguities in the SAXS structures resulting from polydispersity in the system, as indicated by the Kratky plot of the SAXS data (Figure 2.32). Such polydispersity is often the result of conformational flexibility, meaning that the final SAXS structure would be an average of all the conformations of the complex in solution, which could especially affect the analysis of a large extended complex. Furthermore, as noted above, the possibility remains that the region of the SEC peak used for SAXS analysis may not have been entirely homogenous and that the final SAXS structure therefore might be an average structure with contributions from smaller complexes.

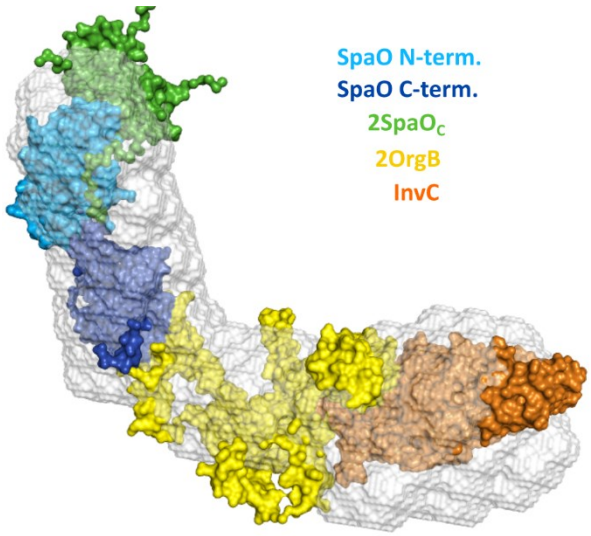


Figure 2.31. SAXS-based hybrid model of SpaO-2$SpaO_C$-2OrgB-InvC. The model was aligned with a representative *ab initio* bead model (grey). SAXS modeling was performed by Anne Tuukkanen.

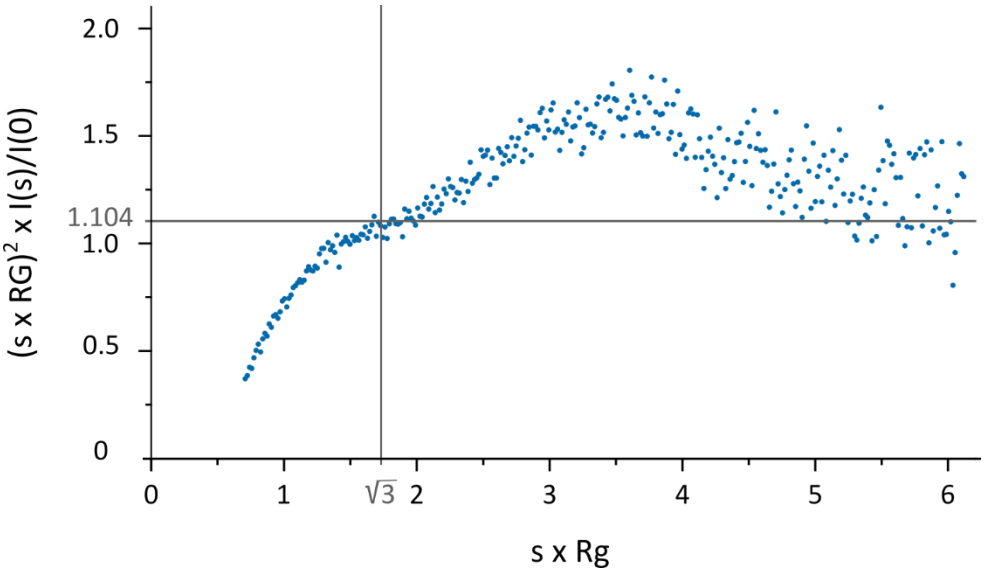


Figure 2.32. Normalized Kratky plot of SpaO/SpaO$_C$/OrgB/InvC SAXS data.

2.5. The SpaO/SpaO$_C$/OrgB/InvC structure is in good agreement with the cryo-electron tomography structure of the sorting platform

Because of the apparent resemblance between the extended L-shape of the SpaO/SpaO$_C$/OrgB/InvC complex in solution and the pod-like densities seen in the CET structure of the *Salmonella* sorting platform (B. Hu et al., 2017), the *ab initio* SAXS structure was superimposed with the CET map to assess whether the complete sorting platform could possibly be assembled from soluble SpaO/SpaO$_C$/OrgB/InvC building blocks. Although it should be noted that in this approach two low-resolution structures associated with their own uncertainties are compared, this superposition shows a good correspondence between the two structures with the short arm of the SAXS model aligning with the outer pod densities and the longer arm with the spokes and central hub region of the CET map (Figure 2.33A). Taking also into account the subunit positions from the SAXS multiphase analysis and rigid-body hybrid model (Figure 2.30D, Figure 2.31), this orientation places SpaO-2SpaO$_C$ in the outer pods, InvC inside the central hub and 2OrgB in the spokes connecting them. These positions correspond well with the assignments made in CET using deletion mutants and the localization of protein tags fused to sorting platform components (B. Hu et al., 2017). When the SAXS *ab initio* model is aligned with each of the legs of the CET map, overlaps between the bead models and thus steric clashes arise in the central hub region (Figure 2.33B). However, the Kratky analysis of the SAXS data (Figure 2.32) indicates that the SpaO/SpaO$_C$/OrgB/InvC complex may be a flexible

structure capable of undergoing conformational changes upon assembly into the complete sorting platform. In fact, a straightening of the OrgB dimer in the rigid-body model and an upward rotation of InvC by 90° around its interaction site with OrgB would result in a

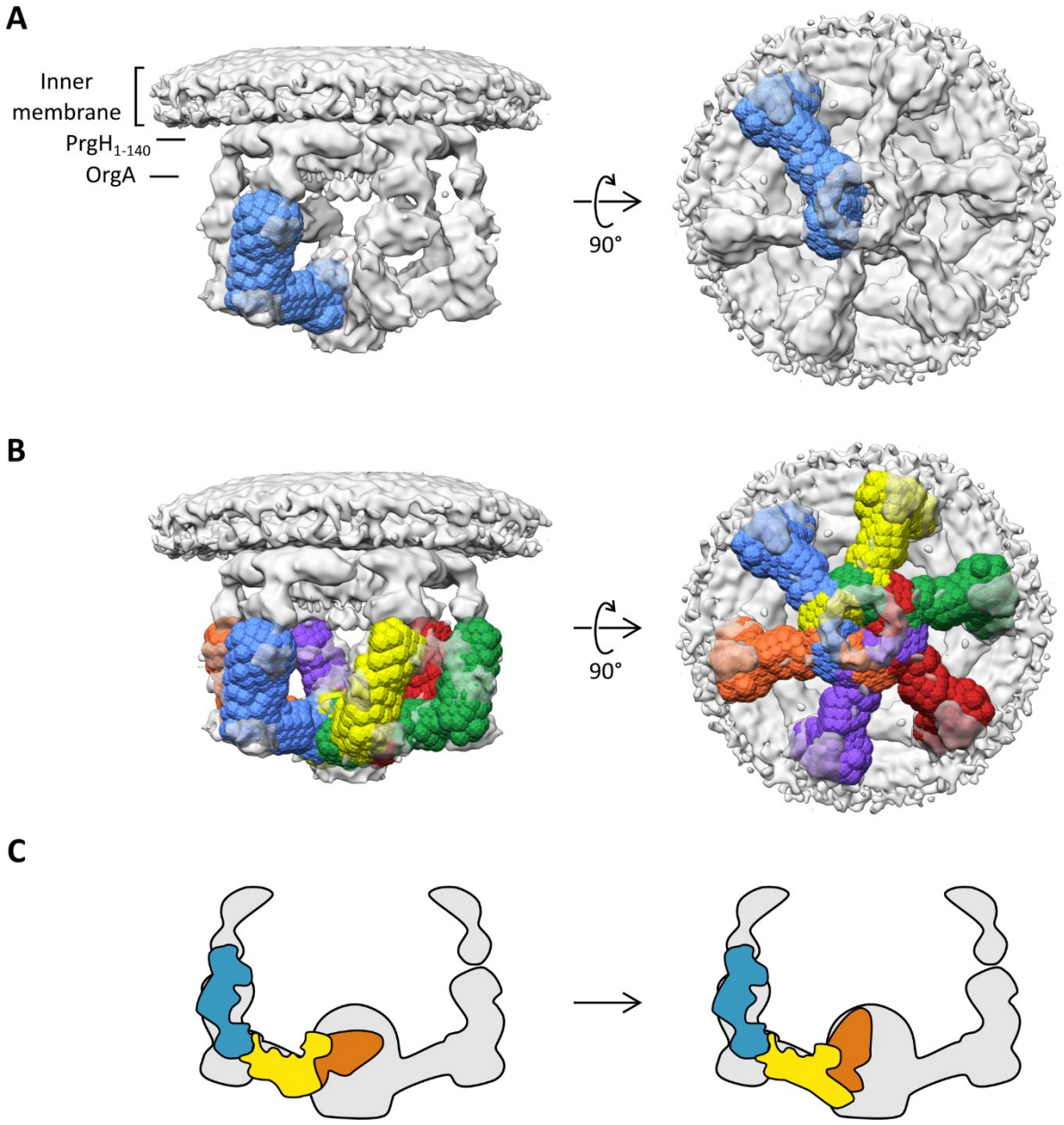


Figure 2.33. Comparison of the SpaO/SpaOC/OrgB/InvC SAXS structure and the *in situ* cryo-electron tomography map of the *Salmonella* sorting platform. A) Superposition of the *ab initio* bead model (blue) with the CET map of the *S.* Typhimurium sorting platform (EMDB ID: EMD-8544, grey) in side view (left) and bottom view (right). PrgH1-140 and OrgA are labeled according to B. Hu et al., 2017. **B)** Superposition of six copies of the SAXS bead model with the CET map in side and bottom view. **C)** Schematic model of proposed conformational changes within the SpaO/SpaOC/OrgB/InvC complex upon assembly at the T3SS base. Left: Schematic representation of a central slice through the CET map (grey) aligned with a representation of the SAXS hybrid model (Figure 2.31) indicating the conformation of the complex in solution. Right: Hypothetical conformation of the complex after assembly. (SpaO/SpaOC: blue; OrgB: yellow; InvC: orange)

conformation in which InvC is oriented toward the T3SS export apparatus, parallel to SpaO/SpaO$_C$ (Figure 2.33C). Adopting this configuration would not only resolve the steric clashes, but also perfectly position InvC for the formation of a hexameric ATPase complex in the central hub of the CET map. Together, these results suggest that the SpaO/SpaO$_C$/OrgB/InvC complex represents the major soluble building block from which the complete sorting platform could be assembled.

2.6. Sorting platform sub-complexes do not interact with a T3SS substrate *in vitro*

2.6.1. Microscale thermophoresis appears unsuitable for *in vitro* interaction analysis of dynamic or unstable binding partners

The T3SS sorting platform is a critical player in the export of T3SS substrates and has been shown to associate *in vivo* with both translocator and effector proteins in complex with their respective chaperones (Lara-Tejero et al., 2011). Therefore, binding of different sorting platform sub-complexes was also tested *in vitro* using microscale thermophoresis (MST) (Seidel et al., 2013). This technique measures the migration of a fluorescently labeled molecule across a thermal gradient, a process that is influenced by the size, charge and hydration shell of the particle (Duhr & Braun, 2006). Because these factors change upon binding to another protein, changes in thermophoretic mobility can be used to determine binding between the fluorescently labeled protein and another, unlabeled molecule. Here, a complex of the dimeric T3SS chaperone SicP with a fragment of the effector SptP (residues 1-158), which contains the T3SS signal sequence and chaperone-binding domain (Fu & Galan, 1998), was fluorescently labelled and mixed with different concentrations of both SpaO/SpaO$_C$ and SpaO/SpaO$_C$/OrgB ligands. SpaO/SpaO$_C$/OrgB/InvC complexes were excluded from this analysis due to their post-purification inhomogeneity (see section 2.4.2).

Testing the binding of SpaO/SpaO$_C$/OrgB to SicP/SptP$_{1\text{-}158}$, no clear sigmoidal binding curve could be obtained (Figure 2.34A, B). Instead, MST signals are seen to first decrease and then increase with increasing SpaO/SpaO$_C$/OrgB concentration. Because thermophoresis is influenced by several molecular parameters whose changes may lead to either increases or decreases in MST signals, the distribution of data points suggests binding events at lower and higher ligand concentration, possibly the interaction of SicP/SptP$_{1\text{-}158}$ with a single 2(SpaO-

2SpaO$_C$)-2OrgB complex and higher-order oligomers, respectively. When MST experiments were performed using SpaO/SpaO$_C$ as the unlabeled ligand, one experiment showed an almost random distribution of points (Experiment C, Figure 2.34C), while in a second experiment a shape more similar to the SpaO/SpaO$_C$/OrgB experiments was obtained (Experiment D, Figure 2.34D). Because the possible binding events are not completely covered by the concentration range tested in these preliminary experiments, dissociation constants (Kd) can only be estimated as Kd <50 nM and Kd >10^5 nM (100 µM) for the apparent high-affinity and low-affinity interaction between SicP/SptP$_{1-158}$ and the sorting platform complexes (Figure 2.35).

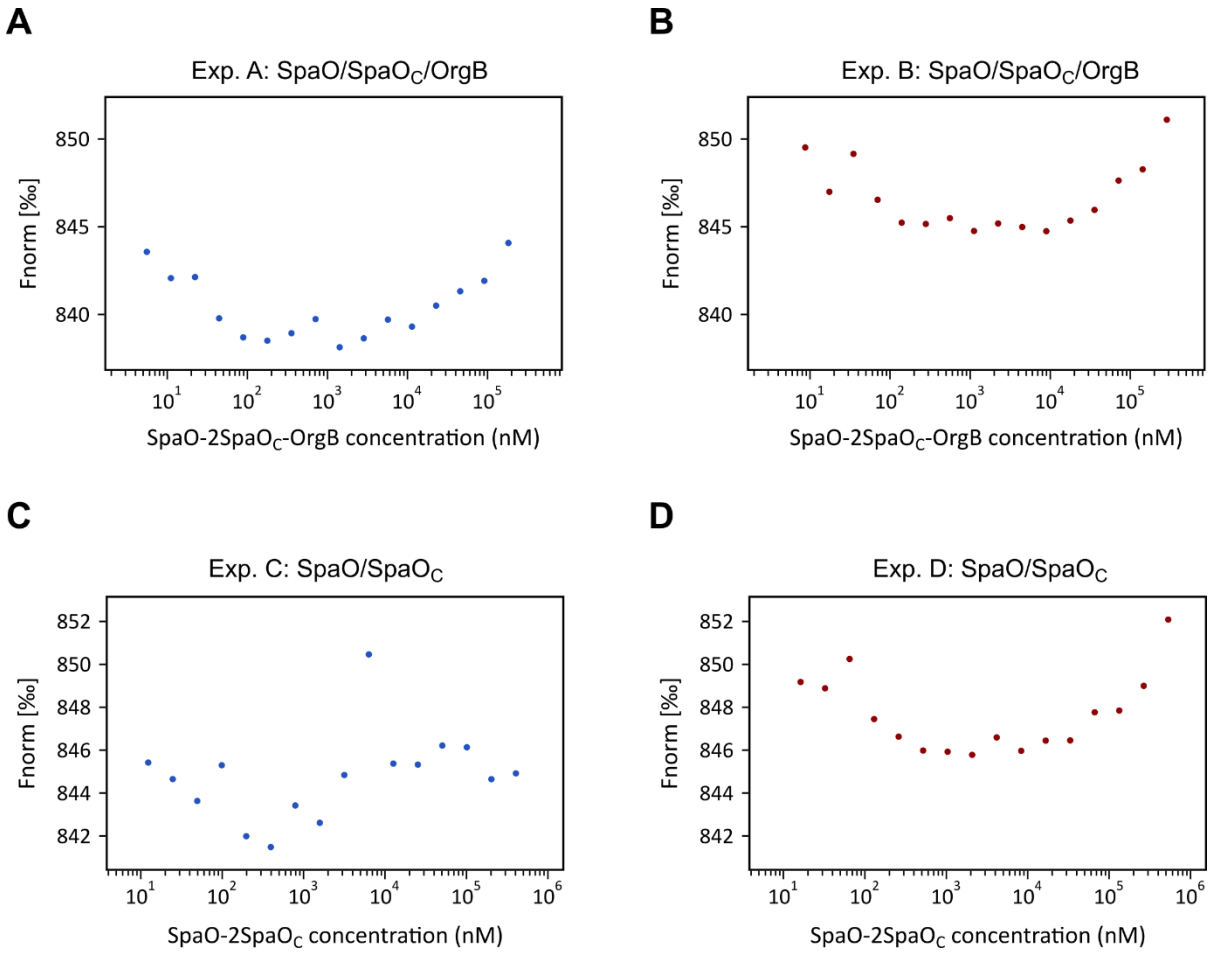
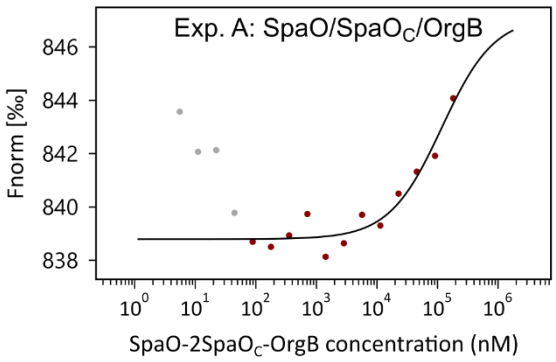


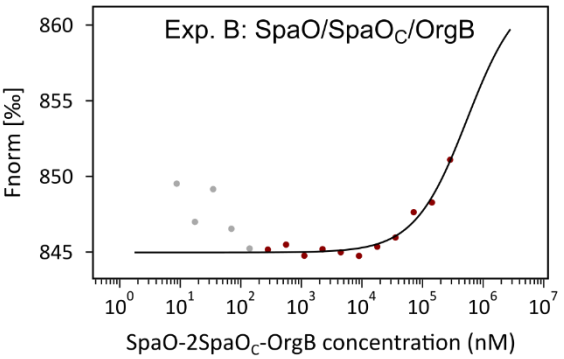
Figure 2.34. Binding analysis of sorting platform sub-complexes to a T3SS substrate-chaperone complex by microscale thermophoresis. 100 nM fluorescently labeled SicP/SptP$_{1-158}$ were mixed with different concentrations of SpaO/SpaO$_C$/OrgB or SpaO/SpaO$_C$ and analysed by MST. Exp: Experiment. **A and B)** Interaction between SpaO/SpaO$_C$/OrgB and SicP/SptP$_{1-158}$ from two independent experiments. SpaO/SpaO$_C$/OrgB concentrations are given in terms of complexes containing a single SpaO-2SpaO$_C$ (i.e. half of the 2SpaO-4SpaO$_C$-2OrgB complexes found by native MS) for easier comparability to measurements involving SpaO-2SpaO$_C$. **C and D)** Interaction between SpaO/SpaO$_C$ and SicP/SptP$_{1-158}$. MST binding curves from two independent experiments are shown.

A

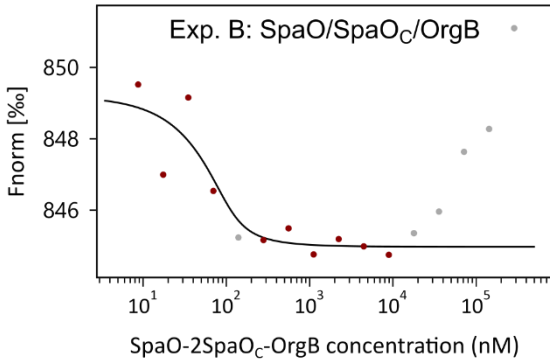


Kd	117 µM
Kd Confidence (68%)	± 70 µM

B

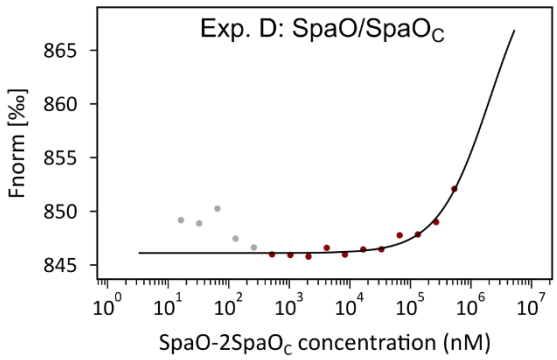


Kd	544 µM
Kd Confidence (68%)	± 369 µM

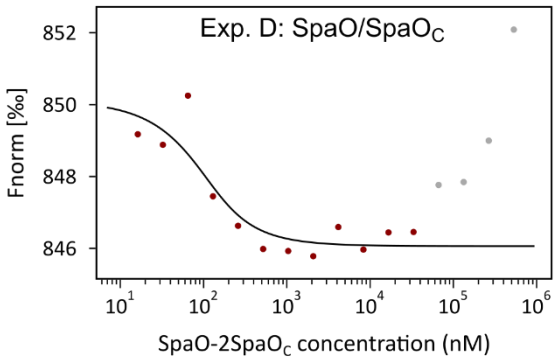


Kd	12 nM
Kd Confidence (68%)	± 27 nM

C



Kd	2.1 mM
Kd Confidence (68%)	± 2.5 mM



Kd	49 nM
Kd Confidence (68%)	± 44 nM

Figure 2.35. Possible binding curves calculated from MST experiments in Figure 2.34. The calculated dissociation constants Kd and the 68%-confidence interval are given below each curve. A, B and C show possible curves from experiments A, B and D, respectively. Data points not included in the curve calculations are shaded grey.

However, these interactions could not be confirmed by other biophysical methods (see section 2.6.2 below) and several observations suggest that they are false positive results caused by non-specific interactions. First, SEC repeatedly showed that concentrating SpaO/SpaO$_C$/OrgB samples to the high concentrations required for MST leads to a loss of sample quality through complex dissociation and aggregation (Figure 2.36A). Although SpaO/SpaO$_C$ was generally stable at high concentrations, the particular sample used in experiment D, in which possible interactions were detected, also showed dissociation (Figure 2.36B). Therefore, it is possible that non-specific interactions involving aggregates or dissociation products of sorting platform complexes could have led to aberrant MST results, especially at high ligand concentrations. Second, an MST experiment that tested the suggested dimerization of the SicP/SptP$_{1-158}$ complex (Stebbins & Galán, 2001a) indicated that the labeled SicP/SptP$_{1-158}$ is partially dissociated in the binding assays with sorting platform complexes, raising the possibility of further non-specific interactions. In this experiment, a dissociation constant of approximately 100 nM for a homo-interaction between SicP/SptP$_{1-158}$ complexes was determined (Figure 2.37), and although this suggests that at micromolar protein concentrations virtually all complexes should be present as 2(2SicP-SptP$_{1-158}$) heterohexamers, native MS and SEC-MALS performed at such concentrations mostly detected 2:1 heterotrimers (Supplementary Figure 2 and Figure 2.38). Therefore, the Kd value of the homo-interaction instead suggests that at the concentration of 100 nM used in MST experiments 50% of the labeled SicP/SptP$_{1-158}$ complex exists as dissociated subunits that can specifically interact with added SicP/SptP$_{1-158}$ to re-assemble into full complexes and cause a

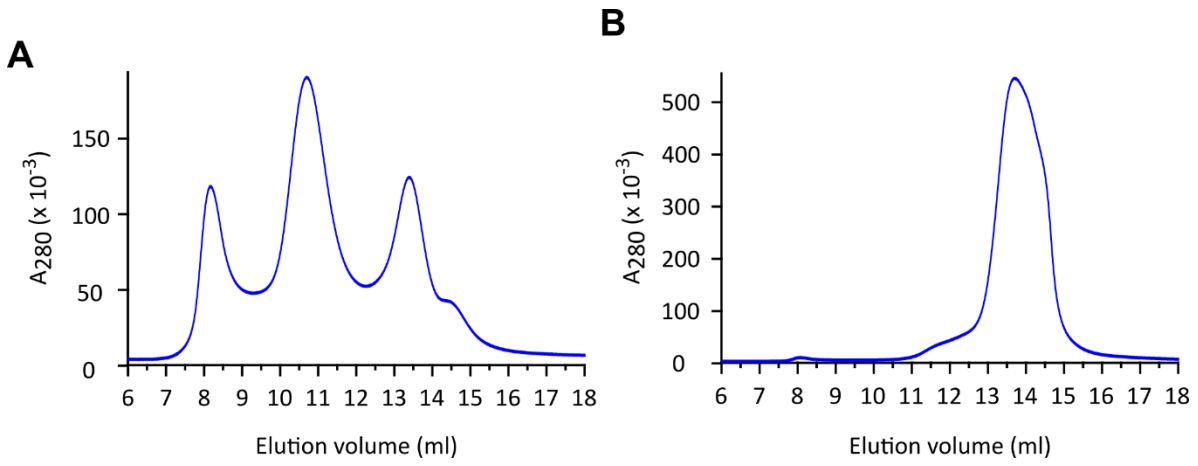
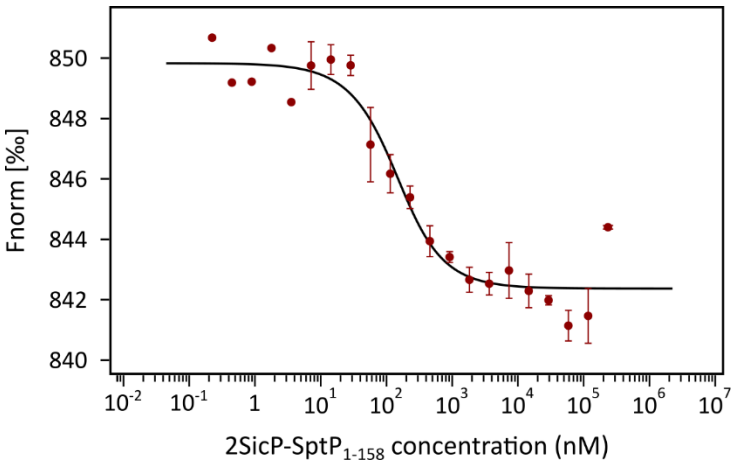


Figure 2.36. SEC of concentrated proteins after MST. SEC was performed on a Superdex 200 10/300 (GL) in MST buffer. **A)** SEC of SpaO/SpaO$_C$/OrgB. Aggregated protein elutes in the void volume of the column (8.2 ml), dissociation products at 13.4 ml. The small peak at 14.5 ml elution volume is of a small amount of SicP/SptP$_{1-158}$ accidentally included in the injected sample. **B)** SEC of SpaO/SpaO$_C$ used in MST experiment D. SpaO-2SpaO$_C$ complexes elute at a volume of 13.6 ml, proteins at higher elution volumes indicate complex dissociation.

detectable change in overall MST signal (Figure 2.37). Because such free subunits will expose extensive hydrophobic surfaces that are otherwise buried in the SicP/SptP$_{1-158}$ complex (Stebbins & Galán, 2001a), it is likely that in other MST experiments they also engaged in non-specific interactions with sorting platform complexes, their aggregates or dissociation products and led to Kd values similar to that of formation of the specific 2SicP-SptP$_{1-158}$ complex.

Together, these experiments indicate that the used MST procedure is not a reliable strategy to obtain quantitative data on possible interactions between dynamic T3SS substrates and sorting platform sub-complexes. However, optimizations of the assay could be tested, such as using a higher concentration of labeled SicP/SptP$_{1-158}$ to eliminate artifacts that appear as high-affinity interactions. Similarly, it might be viable to change to a strategy in which the sorting platform complexes are fluorescently labeled and used at low concentration, because SicP/SptP$_{1-158}$ might prove more stable during the concentration procedure.



Kd	94 nM
Kd Confidence (68%)	± 36 nM
Standard Error (RMSE)	0.831

Figure 2.37. MST of labeled SicP/SptP$_{1-158}$ with increasing concentrations of unlabeled SicP/SptP$_{1-158}$. Error bars indicate standard deviations. Results of the binding curve analysis are displayed on the right. RMSE: Root mean square error.

2.6.2. Native MS and SEC-MALS indicate no interactions between sorting platform components and the T3SS substrate

Given the unreliability of the MST results, interactions between sorting platform sub-complexes and the chaperone-substrate complex were also tested using other biophysical methods. First, SpaO/SpaO$_C$/OrgB and SicP/SptP$_{1-158}$ were mixed and analyzed by SEC-MALS, but despite the high affinity suggested by the low apparent Kd in MST, only elution peaks with masses corresponding to the individual complexes were detected without the formation of larger interaction products (Figure 2.38). Next, in order to also detect possible lower-affinity interactions, native MS was employed. In native MS, mixtures of SpaO/SpaO$_C$ and SicP/SptP$_{1-158}$ produced complex spectra showing SpaO$_C$, SpaO-2SpaO$_C$, 2(SpaO-2SpaO$_C$) and a variety of 2:1 and 4:2 complexes of SicP with both full-length SptP$_{1-158}$ and a smaller degradation product, but no interaction products between SpaO/SpaO$_C$ and SicP/SptP$_{1-158}$ could be detected (Figure 2.39A). Similarly, SicP/SptP$_{1-158}$ failed to interact with the larger sorting platform complexes present in purified SpaO/SpaO$_C$/OrgB/InvC (Figure 2.39B). However, while this sample was rich in SpaO-2SpaO$_C$ and 2(SpaO-2SpaO$_C$)-2OrgB, it showed reduced levels of InvC-containing complexes due to prior freezing. Therefore, this result should be confirmed using freshly prepared protein. Together, SEC-MALS and native MS did not detect high-affinity interactions between sorting platform complexes and the substrate-chaperone complex and thus support the conclusions that the observed interactions in MST are likely experimental artifacts.

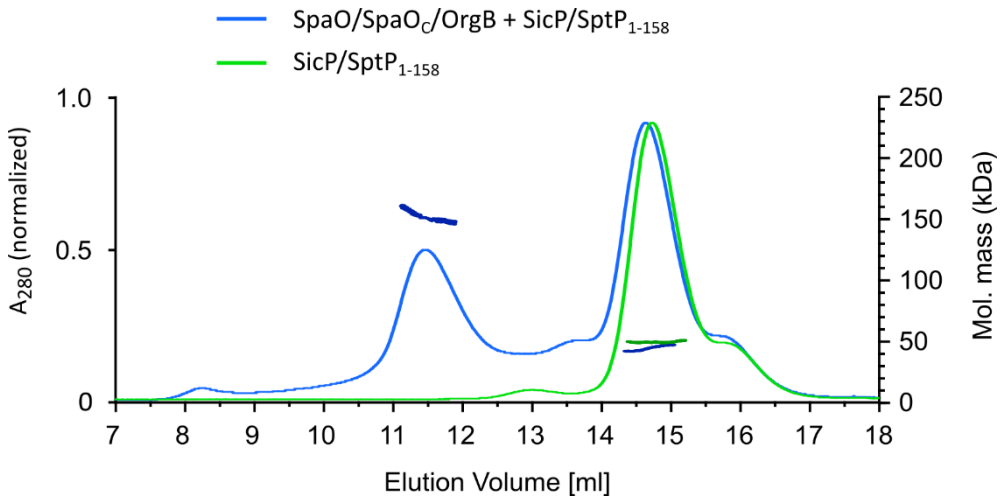


Figure 2.38. MALS analysis of SicP/SptP$_{1-158}$ and SpaO/SpaO$_C$/OrgB. SEC elution profiles (A$_{280}$) of SicP/SptP$_{1-158}$ alone (light green trace) and mixed 1:1 (mass:mass) with SpaO/SpaO$_C$/OrgB (light blue trace). Weight-averaged molecular masses across the elution peaks (dark green and dark blue, respectively) are shown. In both elution profiles the SicP/SptP$_{1-158}$ peak is followed by a smaller peak of SicP without bound SptP$_{1-158}$ and the peak at 13.5 ml in the light blue profile is of SpaO/SpaO$_C$/OrgB dissociation products (see Figure 2.8).

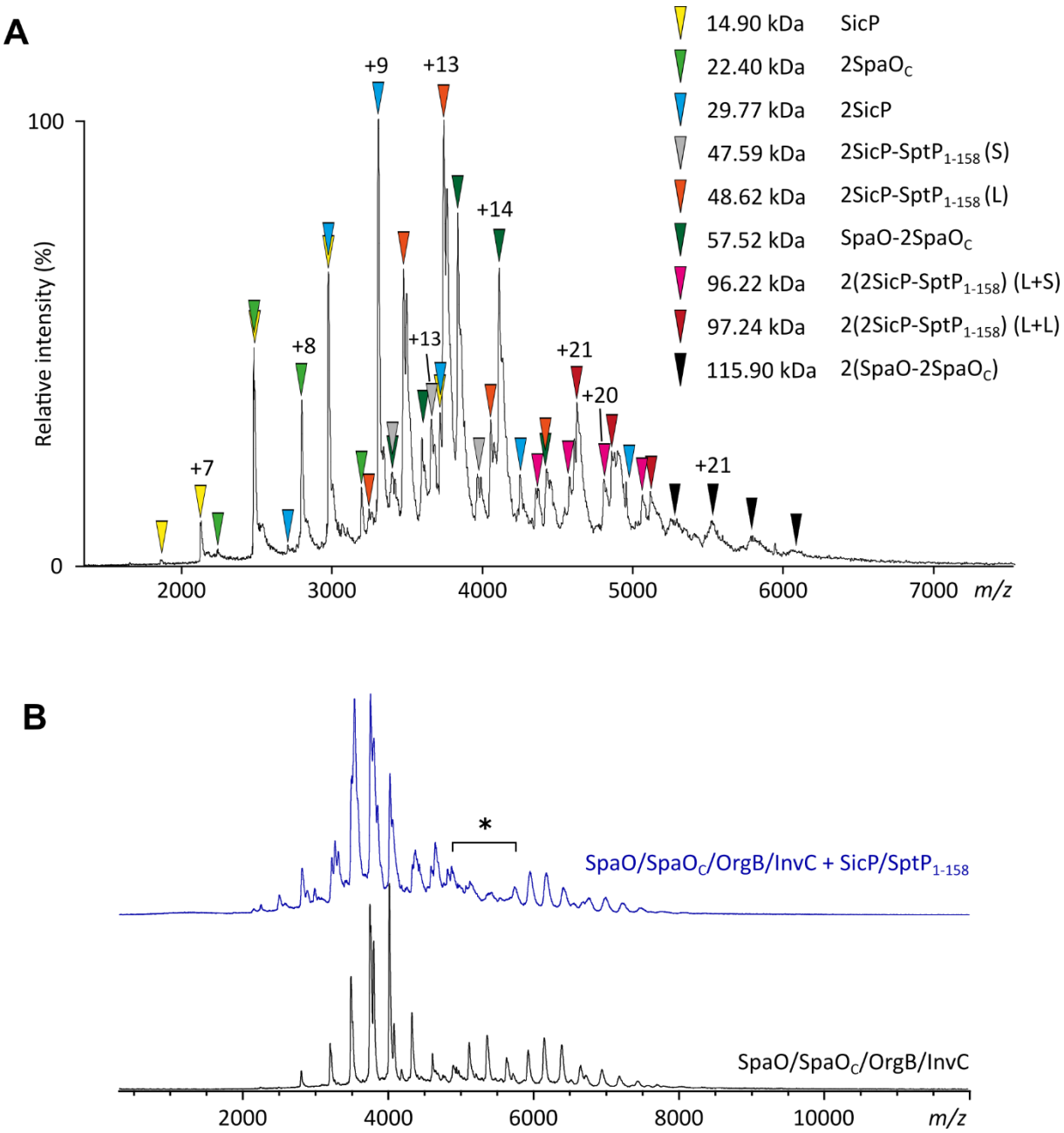


Figure 2.39. Native mass spectrometry of sorting platform complexes with a T3SS chaperone-substrate complex. A) Native mass spectrum of SicP/SptP$_{1-158}$ mixed with SpaO/SpaO$_C$. SptP$_{1-158}$ is present as a longer form (18.8 kDa, indicated as "L") and a shorter degradation product (17.8 kDa, indicated as "S"). **B)** Offset mass spectra of SpaO/SpaO$_C$/OrgB/InvC alone (black, bottom) and in combination with SicP/SptP$_{1-158}$ + 0.2 mM ATPγS + 1 mM Mg^{2+} (blue, top). The addition of SicP/SptP$_{1-158}$ led to an apparent decrease in 2(SpaO-2SpaO$_C$) complexes (marked by an asterisk), but no interaction products between sorting platform complexes and SicP/SptP$_{1-158}$ were detected. Native MS was performed by Johannes Heidemann.

3. Discussion & Perspectives

The type III secretion system is a complex nanomachine used by many Gram-negative bacteria to inject effector proteins into host cells in order to establish and maintain an infection. A critical aspect of its assembly and function is the hierarchical secretion of substrates, with proteins that form the inner rod and needle structures secreted in the early phase, tip proteins and translocators in the middle phase, and effector proteins in the late phase after host cell contact (Deane et al., 2010; Büttner, 2012). The selection of the correct substrates at different stages is believed to involve a "sorting platform", a cytoplasmic complex associated with the T3SS basal body (Lara-Tejero et al., 2011). However, despite its vital role in type III secretion the precise molecular makeup and organization of the sorting platform, as well as its mode of action in substrate sorting are incompletely understood. Because purification of the sorting platform from T3SS-expressing bacteria has so far been unsuccessful, this work aimed to reconstitute the complex to make it amenable to methods of *in vitro* analysis.

3.1. Interactions between sorting platform components

In order to reconstitute the sorting platform of the *Salmonella* SPI-1 T3SS, different components were co-expressed in *E. coli* and tested for stable interactions by co-purification assays. This showed a chain of interactions in the order of SpaO/SpaO$_C$-OrgB-InvC-InvI (Figure 2.1 to Figure 2.4), consistent with previous findings in both *Salmonella* and other bacteria (Notti et al., 2015; Jackson & Plano, 2000; Ibuki et al., 2011). OrgA connects to this chain at the point of SpaO/SpaO$_C$-OrgB, and although the exact binding partner could not be determined here, other findings suggest that the interaction occurs between OrgA and SpaO (Soto et al., 2017; B. Hu et al., 2017). Additionally, an interaction between InvI and the cytoplasmic domain of the export apparatus protein InvA provides a point of contact between the sorting platform and the T3SS needle base similar to that seen in the flagellar system (Fraser et al., 2003). On the other hand, although an interaction between the cytoplasmic domain of the inner membrane ring protein PrgH and OrgA was proposed as a major point of attachment (B. Hu et al., 2017), no complexes between these two proteins could be purified. A possible reason for this is suggested by CET, which showed that the binding of the sorting platform causes the cytoplasmic domains of PrgH to undergo a re-arrangement from a 24-unit ring into 6 patches that align with the pods of the

sorting platform (B. Hu et al., 2017). This indicates cooperativity between multiple PrgH molecules in the interaction with OrgA and thus suggests that the isolated cytoplasmic domain of PrgH, which is monomeric in solution (Bergeron et al., 2013), would not be able to interact with OrgA in co-purification experiments.

3.2. Soluble sub-complexes of the T3SS sorting platform

Individual sorting platform components could be reconstituted into different soluble complexes, the largest of which contained the proteins SpaO, $SpaO_C$, OrgB and InvC. This composition is similar to that observed *in vivo* for cytosolic sorting platform complexes in *Yersinia* and *Salmonella*, although OrgA and its homolog YscK were also detected in those complexes (Diepold et al., 2017; Zhang et al., 2017). However, compared to OrgB and InvC, a lower fraction of OrgA was in a cytosolic state and not associated with needle complexes (Zhang et al., 2017), which probably reflects the low solubility exhibited by OrgA-containing complexes in co-purification assays (Figure 2.2). Similarly, low solubility was also observed for InvI, and complexes containing InvI or OrgA could not be purified in sufficient amounts for biophysical analysis. However, methods like the fusion to solubility tags might in the future render these proteins soluble and make them amendable to *in vitro* characterization (Paraskevopoulou & Falcone, 2018; Nallamsetty & Waugh, 2006).

In addition to SpaO/$SpaO_C$/OrgB/InvC, smaller sub-complexes of $SpaO_C$, SpaO/$SpaO_C$ and SpaO/$SpaO_C$/OrgB could be obtained, which likely represent intermediates in the assembly of the complete sorting platform (see section 3.6 below). Of these complexes, $SpaO_C$ and SpaO/$SpaO_C$ were analyzed in an associated collaborative investigation, which showed that $SpaO_C$ forms a dimer that stabilizes full-length SpaO in solution by stably binding to its N-terminal domain to form SpaO-2$SpaO_C$ complexes (Bernal, 2019; Bernal, Börnicke, et al., 2019). The larger complexes SpaO/$SpaO_C$/OrgB and SpaO/$SpaO_C$/OrgB/InvC were characterized in detail in this study.

3.3. Architecture of the SpaO/SpaO$_C$/OrgB complex

Co-expression with SpaO/SpaO$_C$ rendered the otherwise insoluble ATPase regulator protein OrgB soluble and allowed for the purification of a stable SpaO/SpaO$_C$/OrgB complex (Figure 2.1). This complex likely represents an important sorting platform sub-complex *in vivo*, since OrgB was the main sorting platform protein to co-purify with FLAG-tagged SpaO expressed in *Salmonella* (Lara-Tejero et al., 2011). *In vitro* characterization of SpaO/SpaO$_C$/OrgB by SEC-MALS and native MS showed molecular masses of 165 ± 8 kDa and 175.5 ± 0.5 kDa, respectively, while the SEC elution volume corresponded to a much higher mass of 369 ± 14 kDa (Figure 2.8, Figure 2.9, Figure 2.11). Because SEC is sensitive to both the mass and the shape of a molecule, this indicates that SpaO/SpaO$_C$/OrgB is in an elongated conformation in solution, compatible with previous findings that showed an extended configuration for the flagellar OrgB homolog FliH (Minamino, Gonzalez-Pedrajo, et al., 2002).

Native MS as a highly sensitive and accurate method for determining the mass of biomolecules was used to derive the precise number of subunits within the SpaO/SpaO$_C$/OrgB complex (Figure 2.11). This revealed a stoichiometry of 2(SpaO-2SpaO$_C$)-2OrgB and indicates that OrgB forms a dimer similar to its homologs MxiN and FliH (Minamino & MacNab, 2000; Case & Dickenson, 2018). The architecture of the complex was probed using tandem MS, in which single units of OrgB or SpaO dissociated from the complex without the loss of other subunits (Figure 2.12). This indicates an architecture in which two units of SpaO-2SpaO$_C$-OrgB, connected by an interaction of the OrgB N-terminus with the SpaO C-terminus (Notti et al., 2015), are arranged in a parallel fashion and held together by both the dimerization of OrgB and direct contacts between SpaO-2SpaO$_C$ heterotrimers (Figure 2.13). Moreover, the MS/MS results are in contrast to the dissociation of a SpaO$_C$ monomer from the SpaO-2SpaO$_C$ complex (Bernal, Börnicke, et al., 2019), showing that SpaO$_C$ becomes more buried or otherwise stabilized in OrgB-containing complexes. Although the exact reasons for this protection cannot be discerned with certainty, it also occurs in complexes containing only a single SpaO-2SpaO$_C$ (i.e. SpaO-2SpaO$_C$-2OrgB-InvC, Figure 2.28) and is therefore most likely not caused by the interaction between SpaO-2SpaO$_C$ heterotrimers. Similarly, it is unlikely that the binding of OrgB to SpaO results in a strengthening of SpaO-SpaO$_C$ interactions, given that no conformational changes are observed in the crystal structures of the C-terminal domains of SpaO upon association with the OrgB N-terminus (Notti et al., 2015). This leaves a direct interaction between SpaO$_C$ and OrgB as the most probable cause of SpaO$_C$ stabilization, although it could be questioned whether such interactions are possible considering that SpaO$_C$ stably associates with the N-terminal domain of SpaO while OrgB binds

to the distal part of the C-terminal domain (Bernal, Börnicke, et al., 2019; Notti et al., 2015). On the other hand, both cross-linking and native MS studies have indicated at least transient contacts between $SpaO_C$ and the SpaO C-terminal domain that could result in proximity between $SpaO_C$ and OrgB (Lara-Tejero et al., 2019; Bernal, Börnicke, et al., 2019), and small amounts of $SpaO_C$ have been pulled down by the OrgB N-terminus (Notti et al., 2015). However, these observations are not necessarily indicative of the situation within the complex containing all three proteins SpaO, $SpaO_C$ and OrgB, and future investigations into the precise subunit contacts through high-resolution structure determination or surface protection assays using e.g. hydrogen-deuterium exchange mass spectrometry could greatly contribute to answering this question.

3.4. Heterogeneity of SpaO/$SpaO_C$/OrgB/InvC complexes introduced by chromatographic purification limits *in vitro* characterization options

The largest sorting platform complex investigated in this study was formed by the binding of the ATPase InvC to SpaO/$SpaO_C$/OrgB (Figure 2.3). However, the characterization of SpaO/$SpaO_C$/OrgB/InvC was challenging due to the co-purification of the *E. coli* chaperone DnaK in Strep-Tactin affinity purification (Figure 2.15) and the formation of new molecular species with molecular masses in the megadalton range with each subsequent form of buffer exchange (Figure 2.17, Figure 2.21). This meant that no homogenous bulk sample of SpaO/$SpaO_C$/OrgB/InvC could be purified by chromatographic methods, which largely limited the available techniques of characterization to methods that can be coupled to SEC and those like native MS that can tolerate high sample heterogeneity. In addition, because the complex remained intact at the point of elution from the SEC column, rapid sample handling and freezing after SEC might in the future allow for the structural analysis of SpaO/$SpaO_C$/OrgB/InvC using cryo-EM.

While it could be speculated that the larger species created by buffer exchanges might be higher-order oligomerization products, they were only stable in low-salt buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl) and aggregated at near-physiological salt concentrations (Figure 2.19). Such aggregation upon increases in salt concentration, as opposed to a reversion to smaller complexes, indicates that their formation is an irreversible process and suggests that they might

not be the result of biologically relevant oligomerization, which would be expected to be reversible given the exchange of components in active T3SS (Diepold et al., 2015). Instead, it is more likely that an unknown stabilizing factor co-purifies with SpaO/SpaO$_C$/OrgB/InvC in affinity purification and is gradually removed from the complex in subsequent purification steps. Therefore, identification of this factor could help in obtaining more stable complexes that could be used in a wider variety of methods.

3.5. Molecular organization of the SpaO/SpaO$_C$/OrgB/InvC complex

SpaO/SpaO$_C$/OrgB/InvC samples consist of a mixture of different complexes of similar composition, with a molecular mass of 207 ± 1 kDa for the largest species in SEC-MALS (Figure 2.23). A similar mixture of complexes was also found by native MS, the most important of which are 2(SpaO-2SpaO$_C$)-2OrgB-InvC (215 kDa) and SpaO-2SpaO$_C$-2OrgB-InvC (159 kDa) (Figure 2.24). Also, significant amounts of SpaO-2SpaO$_C$ and 2(SpaO-2SpaO$_C$)-2OrgB were detected, indicating that the soluble sorting platform complexes are a dynamic system in which subunits may associate and dissociate to create a pool of different soluble sub-complexes. This conclusion is in line with pull-down and single-molecule tracking experiments in *Yersinia,* which showed that the SpaO and OrgB homologs YscQ and YscL can be found in a variety of different complexes whose compositions are affected by the activation of type III secretion (Rocha et al., 2018; Diepold et al., 2017). Interestingly, even though native MS showed OrgB to exist as a dimer, only a single InvC molecule was present in InvC-containing complexes. While it is conceivable that only one InvC associates with the OrgB dimer due to proximity of the InvC-binding domains, it is more likely that the two OrgB molecules adopt different structures and functions within the dimer. This conclusion is supported by tandem MS of 2(SpaO-2SpaO$_C$)-2OrgB-InvC and SpaO-2SpaO$_C$-2OrgB-InvC (Figure 2.28), which showed the dissociation of a single OrgB monomer and suggests that one OrgB simultaneously interacts with SpaO and InvC, while the other one is less tightly connected to the complexes. Furthermore, these results are in line with the asymmetry observed in the crystal structure of 2FliH$_C$-FliI, the flagellar homologs to OrgB-InvC (Imada et al., 2016). In this structure, the C-terminal globular domain of one FliH extends to interact with the ATPase FliI, while the same domain of the second FliH adopts a different fold and bends back onto its preceding helical region. This asymmetry in FliH is also of functional importance, because it exists even before the binding of FliI and is a prerequisite for the formation of this interaction (Minamino, Gonzalez-Pedrajo, et al., 2002).

The SpaO/SpaO$_C$/OrgB/InvC complex adopts an extended L-shaped conformation in solution as revealed by SEC-SAXS, with SpaO-2SpaO$_C$ located in the shorter arm of the structure (Figure 2.30). Combining these findings with information on the interacting domains obtained through co-purification experiments (Figure 2.14, also Notti et al., 2015), a schematic model of the architecture of the complex in solution can be derived (Figure 3.1). This model is further supported by SAXS hybrid modeling (Figure 2.31), which showed that this arrangement of sorting platform proteins is compatible with experimental SAXS data. The only difference between the schematic and the SAXS model is that while the hybrid model places the SpaO$_C$ dimer at the far end of the short arm of the L-shape, the stabilization of SpaO$_C$ in MS/MS experiments suggests that it should be located closer to the center of the complex forming contacts with both the N-terminal and C-terminal domains of SpaO and possibly with OrgB (see section 3.3). Even though MALS and native MS indicated a stoichiometry of 2(SpaO-2SpaO$_C$)-2OrgB-InvC, only a complex containing a single SpaO-2SpaO$_C$ heterotrimer could be fitted to the SAXS data. Although the reasons for this disagreement could not conclusively be determined, it might be caused by conformational flexibility of the system or the presence of smaller complexes in the measured sample, both of which would lead to a SAXS average model that appears smaller than the real complex in solution. Therefore, it might be possible to improve the SAXS model by further increasing the concentration of protein injected into the SEC column to increase the abundance of the largest 2(SpaO-2SpaO$_C$)-2OrgB-InvC complex while lowering the levels of smaller complexes, as well as by using an SEC column with a smaller fractionation range to achieve better separation of the largest complex from smaller species.

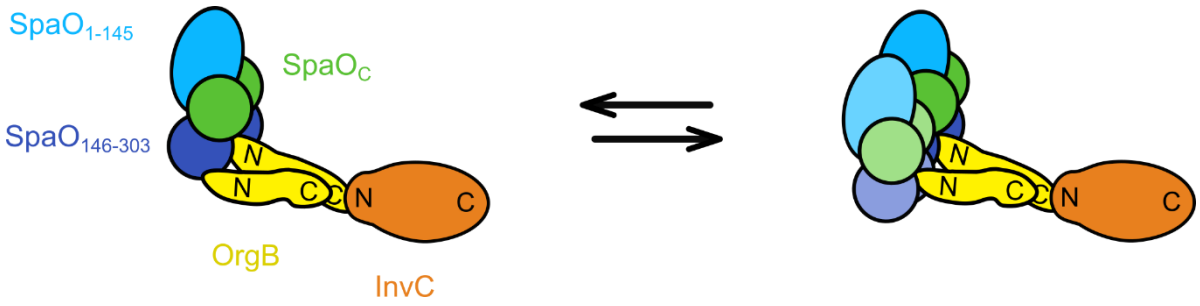


Figure 3.1. Schematic model of the SpaO/SpaO$_C$/OrgB/InvC complex architecture. While SAXS analysis indicated a stoichiometry of SpaO-2SpaO$_C$-2OrgB-InvC (left), native MS and MALS showed that this complex can recruit an additional SpaO-2SpaO$_C$ (right). The positions of N- and C-terminal domains are indicated.

3.6. SpaO/SpaO$_C$/OrgB/InvC is likely the soluble core building block of the T3SS sorting platform

Despite possible caveats associated with the SAXS analysis, the resulting bead model is in good agreement with the CET map of the assembled *Salmonella* sorting platform, both in terms of overall shape and the placement of SpaO/SpaO$_C$ in the outer pods, InvC in the central hub and OrgB in the spokes connecting the two (B. Hu et al., 2017). This strongly suggests that the SpaO/SpaO$_C$/OrgB/InvC complexes represent the major soluble building blocks from which the complete sorting platform could be assembled. Combining this with the other soluble sub-complexes described here and in an associated study (Bernal, Börnicke, et al., 2019; Bernal, 2019), a model can be proposed of how the individual protein components interact to form the soluble SpaO/SpaO$_C$/OrgB/InvC building block that subsequently binds to needle complexes through OrgA and InvI to assemble the complete sorting platform (Figure 3.2). Whether OrgA could also form a stable part of the soluble building blocks is uncertain, because although previous reports propose that OrgA exists in both a needle-associated and a cytosolic state (Zhang et al., 2017; Diepold et al., 2017), its cytosolic fraction in those experiments was lower than that of other sorting platform components (Zhang et al., 2017) and only low levels of OrgA-containing complexes could be purified in *in vitro* reconstitution due to its low solubility (Figure 2.2).

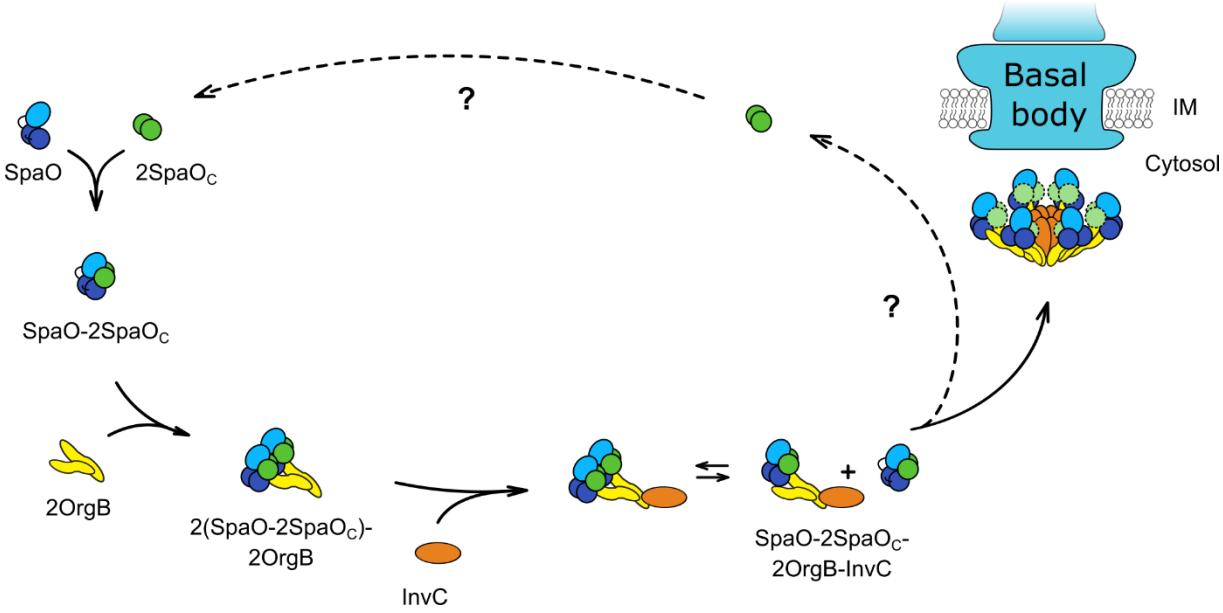


Figure 3.2. Assembly model of the sorting platform from its individual components. Due to conflicting evidence the inclusion of SpaO$_C$ in the assembled sorting platform is currently under debate, indicated here by a faded color and dotted outline of SpaO$_C$. A possible role of SpaO$_C$ as a recycling chaperone is indicated by dashed arrows (see section 3.8). IM = Inner membrane.

The comparison of the SpaO/SpaO$_C$/OrgB/InvC SAXS model and the CET map of the assembled sorting platform further suggests that a conformational change is required within the building block to form the complete sorting platform at the needle complex (Figure 2.33). This most likely involves a rotation of the ATPase InvC around its interaction site with OrgB, which would allow for the oligomerization of the T3SS ATPase required for full ATPase activity (Claret et al., 2003; Minamino et al., 2006; Burgess, Jones, et al., 2016). In fact, this proposed flexibility is in line with both the Kratky analysis of the SAXS data and the observation that the crystal structure of the flagellar homologs FliH-FliI could only be assembled into a FliI hexamer *in silico* when the ATPase domain of FliI was re-arranged relative to its FliH-interacting N-terminal domain (Imada et al., 2016). Additionally, flexibility of the ATPase N-terminal domain is also suggested by the fact that its removal was required for the successful crystallization of T3SS ATPases from several other species (Zarivach et al., 2007; Allison et al., 2014; Burgess, Burgess, et al., 2016).

The similarity between the SAXS model and the CET map indicates that six of the building blocks could assemble into the complete sorting platform. This means that the SpaO-2SpaO$_C$-2OrgB-InvC subunit stoichiometry of the SAXS hybrid model would translate into a 6SpaO-12SpaO$_C$-12OrgB-6InvC stoichiometry for the complete structure. However, it is unclear whether SpaO$_C$ is really present in the assembled platform (see section 3.8 below). While this stoichiometry is in agreement with the copy numbers of OrgB and InvC found by fluorescence microscopy, it is in contrast with the approximately 24 subunits of SpaO or its homolog YscQ that have been shown to be present at the T3SS needle base (Diepold et al., 2017; Zhang et al., 2017). Because it is not possible to fit additional units of SpaO or SpaO/SpaO$_C$ to the SAXS data and by extension to the CET pod densities, these results indicate a clear discrepancy between the six-pod model and the subunit stoichiometry determined *in vivo*. On the other hand, native MS showed that the SpaO-2SpaO$_C$-2OrgB-InvC complex can recruit an additional SpaO-2SpaO$_C$ to bring the total SpaO copy number in the sorting platform up to twelve, and although no more than two units of SpaO were found in any of the soluble complexes derived from the *Salmonella* sorting platform proteins in this study, results from the *Shigella* homolog Spa33 indicate that higher-order oligomeric states may be possible (McDowell et al., 2016). Therefore, it is conceivable that the pods of the CET map only display a stable core structure and that additional SpaO might associate with the platform in a dynamic fashion. Such dynamic subunits would likely not be visible in tomography maps due to sub-tomogram averaging and the application of symmetry in the calculations of the maps. Nevertheless, further experiments will be required to consolidate the model of a six-pod structure with the copy number of 24 SpaO proteins per sorting platform.

3.7. The complete sorting platform could not be assembled *in vitro*

The distribution of clusters of sorting platform components in previously reported *in vivo* fluorescence microscopy experiments indicates that assembled or partially assembled sorting platforms can exist both in association with T3SS needle complexes and in a soluble cytosolic state (Zhang et al., 2017). However, with the exception of possible dimers in some mass spectra, no higher-order oligomerization of the SpaO/SpaO$_C$/OrgB/InvC building blocks could be detected *in vitro* (Figure 2.23, Figure 2.24, Figure 2.25). Additionally, oligomerization of complexes could not be induced by addition of a non-hydrolyzable ATP analog (Figure 2.26), a method that was previously successful in promoting oligomerization of the flagellar ATPase FliI (Claret et al., 2003; Kazetani et al., 2009). This indicates that the assembly of complete sorting platforms *in vitro* might require additional factors like the adapter protein OrgA, which was shown by fluorescence microscopy to be necessary for the formation of SpaO clusters *in vivo* (Zhang et al., 2017). Whether InvI could also aid in this process is unclear, because while its homologs have been shown to induce the hexamerization of ATPases in the flagellar system and enteropathogenic *E. coli* (Ibuki et al., 2011; Majewski et al., 2019), it was not required for the formation of the *Salmonella* sorting platform *in vivo* (B. Hu et al., 2017). Nevertheless, the effects of these proteins on oligomerization could not be tested *in vitro* due to their low solubility (Figure 2.2, Figure 2.4). In addition to these factors, the formation of the sorting platform also appears to require a template to coordinate its assembly, as indicated by results showing that *in vivo* both cytosolic and membrane-localized sorting platforms only form in the presence of T3SS needle complexes (Zhang et al., 2017). Therefore, a possible strategy for the successful reconstitution of assembled, soluble sorting platforms *in vitro* might first involve steps to increase the solubility of OrgA and possibly InvI by e.g. fusion to solubility tags (Paraskevopoulou & Falcone, 2018; Nallamsetty & Waugh, 2006), and their subsequent combination with SpaO/SpaO$_C$/OrgB/InvC and purified needle complexes or basal bodies.

3.8. SpaO$_C$ likely acts as a chaperone in sorting platform assembly

The function of many T3SSs requires the homolog of both full-length SpaO and a shorter variant that is produced from an internal translation initiation site within the *spaO*-homologous gene (McDowell et al., 2016; Yu et al., 2011; Bzymek et al., 2012; Song et al., 2017). In *Salmonella*, the short isoform SpaO$_C$ is required for the solubility of the full-length SpaO and the two interact to

form SpaO-2SpaO$_C$ complexes (Bernal, Börnicke, et al., 2019, also Figure 2.11) similar to the 1:2 complexes also observed in other species (Bzymek et al., 2012; McDowell et al., 2016). However, the precise role of the short isoform SpaO$_C$ in type III secretion is currently uncertain. While YscQ$_C$ and Spa33$_C$ were found to be essential for the secretion of T3SS substrates in *Yersinia* and *Shigella*, respectively, SpaO$_C$ appears to be not as critical and different *Salmonella* Δ*spaO$_C$* mutants retained varying levels of substrate secretion and invasiveness (Song et al., 2017; Lara-Tejero et al., 2019; Bernal, Börnicke, et al., 2019). However, it is possible that the suppression of SpaO$_C$ production in these studies might have been incomplete (Bernal, Börnicke, et al., 2019). Nevertheless, very low levels of SpaO$_C$ appear to be sufficient for type III secretion, albeit at slightly reduced efficiency, which argues that SpaO$_C$ is likely not a structural component of the sorting platform. This hypothesis is supported by sorting platform CET maps of strains expressing fusions of SpaO$_C$ and a fluorescent protein, which did not display any additional densities compared to maps of wildtype strains (Lara-Tejero et al., 2019). On the other hand, it contrasts with evidence from *Yersinia*, in which YscQ$_C$ localized to needle complexes like other structural sorting platform components (Diepold et al., 2015). Given that a difference in SpaO$_C$/YscQ$_C$ function seems unlikely considering the conservation of T3SSs in the two species, additional experiments will be required to reconcile these opposing observations. Due to the low resolution of the structures involved, the superposition of the sorting platform CET map and the SAXS model of the SpaO/SpaO$_C$/OrgB/InvC complex (Figure 2.33) cannot be used to infer whether SpaO$_C$ might be present in the assembled sorting platform and is therefore not able to address this discrepancy.

While a role of SpaO$_C$ in the assembled sorting platform is under debate, SpaO$_C$ is clearly required for the stability of full-length SpaO in solution (Bernal, Börnicke, et al., 2019) and a reduction in SpaO$_C$ also decreased the yield and stability of soluble SpaO/OrgB complexes (section 2.3). Furthermore, native MS showed that it is part of all soluble sorting platform sub-complexes identified in this study and becomes stabilized in these complexes in the presence of OrgB (Figure 2.11, Figure 2.24, Figure 2.28). Together, these findings argue that SpaO$_C$ is a critical component of soluble sorting platform complexes, and combined with the apparent absence of SpaO$_C$ from the assembled sorting platform (Lara-Tejero et al., 2019) suggest that SpaO$_C$ acts as a chaperone during the sorting platform assembly process. By stabilizing the soluble building blocks in the cytosol, delivering them to the T3SS basal body and dissociating upon formation of the full structure (Figure 3.2), a relatively small amount of recycling SpaO$_C$ could be able to facilitate the formation of many functional sorting platforms and explain the observation that low levels of SpaO$_C$ appear to be sufficient for type III secretion function

(Bernal, Börnicke, et al., 2019). In addition, previous studies showed the exchange of sorting platform components between a needle-complex-bound and a cytosolic state, the rate of which increased upon activation of type III secretion (Diepold et al., 2015; Zhang et al., 2017; Bai et al., 2014). This suggests a role of $SpaO_C$ also in the substrate secretion stage of T3SS function, acting to re-solubilize the sorting platform building blocks and thereby allowing for the effective exchange of subunits at the needle complex. The reasons for the observed exchange are currently unknown, but it has been hypothesized that soluble sorting platform complexes act as dynamic substrate carriers that recruit T3SS substrate-chaperone complexes in the cytosol and deliver them to the needle complex for secretion (Bai et al., 2014).

3.9. Sorting platform sub-complexes of SpaO, $SpaO_C$ and OrgB do not interact with a chaperone-substrate complex *in vitro*

The sorting platform plays a critical role in type III secretion and is believed to ensure the hierarchy of secretion by selecting the appropriate substrates during each stage of T3SS assembly and effector secretion (Lara-Tejero et al., 2011). Because the mechanism of sorting and the exact proteins involved in this function are still unclear, different purified soluble sorting platform sub-complexes were tested for their ability to interact *in vitro* with a complex of the T3SS chaperone SicP and a fragment of the effector protein SptP that encompasses its secretion signal sequence and chaperone-binding domain. MST was used as a method of obtaining quantitative binding data but proved to be unreliable due to instability of the sorting platform complexes at the high protein concentrations required by this method (see section 2.6.1). Additional qualitative interaction analysis by native MS did not detect binding between the chaperone-substrate complex $SicP/SptP_{1\text{-}158}$ and $SpaO/SpaO_C$ or purified $SpaO/SpaO_C$/OrgB/InvC samples containing the larger $SpaO/SpaO_C$/OrgB species (Figure 2.39). These observations indicate that the sub-complexes of SpaO, $SpaO_C$ and OrgB do not interact with T3SS substrates and thus argue against a possible role of soluble sorting platform complexes as dynamic carriers that shuttle substrates from the cytosol to the T3SS. Because InvC-containing complexes were underrepresented in the $SpaO/SpaO_C$/OrgB/InvC sample used in this interaction assay, the possibility remains that soluble sorting platform complexes may interact with substrates through the ATPase, similar to the interactions seen in pull-down assays involving InvC in isolation (Akeda & Galán, 2005). Furthermore, substrate interactions might require the proteins OrgA or InvI, or only occur in the context of the completely assembled sorting platform.

Therefore, solubilization of OrgA and InvI and assembly of sorting platforms *in vitro* as suggested above are necessary to fully investigate interactions between the sorting platform and T3SS substrates *in vitro*. Furthermore, because different modes of binding may be employed for the recruitment of different substrates, future experiments should include a broader panel of substrates to cover early and middle substrates like needle proteins and translocators in addition to effectors.

3.10. Conclusion

Because the purification of the sorting platform from T3SS-expressing bacteria has not been successful to date, this work established a method of co-expression and co-purification to reconstitute soluble sub-complexes of the sorting platform *in vitro* and made them accessible to detailed analysis by biophysical methods. Characterization of these complexes by native mass spectrometry revealed their precise subunit stoichiometry and suggested a pathway of how the individual components assemble to form the SpaO/SpaO$_C$/OrgB/InvC complex, which most likely represents the soluble building block that further oligomerizes into the complete sorting platform at the T3SS needle base. Additional analysis by a combination of protein domain pull-down assays, tandem mass spectrometry and small-angle X-ray scattering granted insights into the arrangement of subunits within the building block and improves our understanding of the molecular architecture of the sorting platform. Utilizing the purified complexes in *in vitro* interaction studies to probe the mechanism of substrate binding by the sorting platform, no interactions with a T3SS substrate could be detected, suggesting that additional factors are required for the targeting and sorting of substrates prior to secretion. Thus, this work provides important information about the assembly, architecture and substrate binding ability of the T3SS sorting platform. Moreover, it establishes a foundation for the reconstitution of complete sorting platforms *in vitro* that in the future might serve as the basis for the discovery and development of drugs that combat infections by Gram-negative bacteria through targeting the T3SS sorting platform.

4. Materials & Methods

4.1. Chemicals, instruments and kits

Chemicals, instruments and commercially available kits used in this study are listed in Table 4.1, Table 4.2 and Table 4.3, respectively.

Table 4.1. List of chemicals used in this study

Chemical	Manufacturer
2-mercaptoethanol (2-ME)	Carl Roth
Acetic acid	Carl Roth
Acetone	Merck
Agarose	BLIRT, Serva
Ammonium acetate	Honeywell Research Chemicals
Ammonium persulfate	Carl Roth
Anhydrotetracycline hydrochloride (AHT)	Sigma-Aldrich
ATP	Sigma-Aldrich
ATPγS	Jena Bioscience
Bromophenol blue	Carl Roth
Carbenicillin	Carl Roth
Chloramphenicol	Sigma-Aldrich
Complete Ultra EDTA-free protease inhibitor cocktail	Roche
Coomassie Brilliant Blue G-250	Bio-Rad
Desthiobiotin	Sigma-Aldrich, IBA Lifesciences
DNase I	Roche
Ethidium bromide	Thermo Scientific
Ethylenediaminetetraacetic acid (EDTA)	SERVA
Gel loading dye (6x), purple, no SDS	New England Biolabs
GelGreen	Biotium
GelRed	Biotium
Glucose monohydrate	Merck
Glycerol	Carl Roth
Glycine	Carl Roth
HCl	Carl Roth
HEPES	Carl Roth
Imidazole	Carl Roth
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Fisher Scientific, Thermo Scientific
Kanamycin	Sigma-Aldrich

Chemical	Manufacturer
KCl	Carl Roth
KH_2PO_4	Carl Roth
LB (lysogeny broth) medium, (Luria/Miller)	Carl Roth
LB agar (Luria/Miller)	Carl Roth
Lysozyme	Sigma-Aldrich
Methanol	Acros Organics
$MgCl_2$	Carl Roth
Na_2HPO_4	Carl Roth
NaCl	Carl Roth
NaH_2PO_4	Carl Roth
NaOH	Carl Roth
Roti-Free stripping buffer	Carl Roth
Rotiphorese Gel 30 (37.5:1)	Carl Roth
Skim milk powder	Carl Roth
SOC medium	New England Biolabs
Sodium dodecyl sulfate (SDS)	Sigma-Aldrich
Streptomycin	Sigma-Aldrich
Tetramethylethylenediamine (TEMED)	Bio-Rad
Trichloroacetic acid (TCA)	Carl Roth
Tris base	Sigma-Aldrich
Tween 20	Sigma-Aldrich
Xylene cyanol FF	SERVA

Table 4.2. List of instruments used in this study

Instrument	Model	Manufacturer
Centrifuge	Avanti J-26 XP	Beckman Coulter
Centrifuge	Optima MAX-XP	Beckman Coulter
Centrifuge	Centrifuge 5804 R	Eppendorf
Centrifuge	Heraeus Fresco 21	Thermo Scientific
DNA Gel Documentation	AlphaImager	Alpha Innotec
DNA Gel Documentation	ChemiDoc XRS	Bio-Rad
Electroporation system	GenePulser Xcell	Bio-Rad
Homogenizer	EmulsiFlex-C3	Avestin
HPLC system	Äkta Pure	GE Healthcare
HPLC system	Äkta Explorer	GE Healthcare
HPLC system	Äkta Purifier	GE Healthcare
HPLC System	1260 Infinity	Agilent
Imaging system (chemiluminescence)	ImageQuant LAS4000 mini	GE Healthcare

Instrument	Model	Manufacturer
Imaging system (chemiluminescence)	ImageQuant LAS4000	GE Healthcare
Incubator	Multitron/Multitron Pro	Infors HT
Incubator	Innova 44R	New Brunswick Scientific
MALS detector	MiniDawn Tristar	Wyatt Technology
MALS detector	MiniDawn Treos	Wyatt Technology
Mass spectrometer	Q-Tof 2	Waters
Microscale thermophoresis	Monolith NT.115	Nanotemper Technologies
Refractive index detector	Optilab T-rEX	Wyatt Technology
Refractive index detector	RI-101	Shodex
SAXS detector	Pilatus 2M	Dectris
SDS-PAGE	SE250 Mighty Small II	Hoefer
SDS-PAGE	Mini Protean	Bio-Rad
Semi-dry blotter	V20-SDB	Scie-Plas
Spectrophotometer	NanoDrop 2000	Thermo Scientific
Spectrophotometer	BioPhotometer 6131	Eppendorf
Thermocycler	C1000	Bio-Rad
Thermocycler	T100	Bio-Rad
Thermomixer	Thermomixer Compact	Eppendorf
Thermomixer	Thermomixer F1.5	Eppendorf
Ultrasonic homogenizer	Sonopuls HD 2070	Bandelin
UV transilluminator	TS-40	Ultra-Violet Products

Table 4.3. List of commercially available kits used in this study

Kit type	Name	Manufacturer
DNA miniprep	Zyppy Plasmid Miniprep	Zymo Research
	QIAprep Spin Miniprep	Qiagen
	GeneJET Plasmid Miniprep Kit	Thermo Fisher Scientific
DNA gel extraction	Zymoclean Gel DNA Recovery	Zymo Research
	GeneJet Gel Extraction	Thermo Fisher Scientific
DNA clean-up	DNA Clean & Concentrator	Zymo Research
	GeneJet PCR Purification	Thermo Fisher Scientific

4.2. Standard buffers

Standard buffers used in this study are listed in Table 4.4. Buffers used in protein purification and specific experiments are described in the respective experimental procedures below.

Table 4.4. List of standard buffers used in this study

SDS-PAGE	
4x SDS sample buffer	40% (v/v) glycerol 240 mM Tris-HCl, pH 6.8 8% (w/v) SDS 5% (v/v) 2-mercaptoethanol 0.04% (w/v) bromophenol blue
Tris-glycine SDS running buffer	25 mM Tris base 192 mM glycine 0.1% (w/v) SDS
Coomassie staining solution	0.1 g/l Coomassie Brilliant Blue G-250 36 mM HCl
Western blot	
Transfer buffer	25 mM Tris base 192 mM glycine 20% (v/v) methanol
PBS-T	137 mM NaCl 2.7 mM KCl 10 mM Na_2HPO_4 1.8 mM KH_2PO_4 0.1% (v/v) Tween 20
Agarose gel electrophoresis	
6x DNA loading dye	30% (v/v) glycerol 0.25% (w/v) bromophenol blue 0.25% (w/v) xylene cyanol FF
TAE buffer	40 mM Tris base 20 mM glacial acetic acid 1 mM EDTA

4.3. Bacterial cell culture

S. Typhimurium and *E. coli* strains were grown on LB agar plates or in liquid LB medium with agitation at 180 rpm. When applicable, antibiotic selection was achieved using the concentrations shown in Table 4.5. If more than one antibiotic was used for selection in liquid cultures, half the indicated concentration was used for each antibiotic. Bacterial growth in liquid medium was monitored using absorption at 600 nm (OD_{600}).

Table 4.5. Antibiotic concentrations used in cell culture

Antibiotic	Final concentration
Carbenicillin	100 µg/ml
Chloramphenicol	34 µg/ml
Kanamycin	50 µg/ml
Streptomycin	50 µg/ml

4.4. Molecular biology methods

4.4.1. Molecular cloning

Salmonella Typhimurium genes of interest were cloned into expression vectors by directional cloning using the polymerase chain reaction (PCR) and restriction enzymes. Primer oligonucleotides were synthesized by Eurofins Genomics or Sigma-Aldrich and are listed in Supplementary Table 2.

For the amplification of DNA containing genes of interest, Phusion High-Fidelity DNA Polymerase (New England Biolabs) was used according to the manufacturer's instructions. Primers were used at a final concentration of 0.5 µM and 5-10 ng of plasmid DNA carrying the gene of interest or 1 µl of a single colony of *S.* Typhimurium resuspended in 20 µl of ultrapure water served as templates. Amplification of DNA was carried out using the cycling parameters in Table 4.6 with 34 cycles of denaturation, annealing and extension.

Table 4.6. PCR thermocycling parameters

Cycle step	Temperature (°C)	Time (s)
Initial denaturation	98	30
Denaturation	98	10
Annealing	Depending on primers	15
Extension	72	15-30 per 1 kb of amplicon length
Final extension	72	300
Hold	12	

Amplified DNA was separated and analyzed by gel electrophoresis using 1.0-1.3% agarose gels in TAE buffer at a constant field strength of 5.5 V/cm. DNA bands were visualized by UV illumination after staining with either ethidium bromide mixed into the agarose during gel preparation or incubation with GelGreen or GelRed according to the manufacturer's instructions. SmartLadder (Eurogentec) was used as a standard marker for the estimation of DNA sizes. Bands of the appropriate size were excised from the gel and purified using DNA gel extraction kits.

For the cloning of amplified genes into expression vectors, 1 µg of both gel-purified PCR products and the target plasmids were digested using restriction enzymes (New England Biolabs) according to the manufacturer's instructions. Restricted PCR products were purified using DNA clean-up kits, while restricted plasmids were treated with Antarctic phosphatase (New England Biolabs) and purified by electrophoresis in 0.7% agarose gels and gel extraction. Restricted PCR products and the target plasmids were subsequently mixed at a molar ratio of 3:1 and ligated for 16 h at 16 °C using T4 DNA ligase (New England Biolabs). Ligated plasmids were desalted either by dialysis against deionized water for 20 min using mixed cellulose ester filters with pore size 0.025 µm (Merck Millipore), or by use of DNA clean-up kits. Desalted ligated plasmids were used to transform *E. coli* DH5α or Top10 using electroporation. Following plasmid amplification, successful cloning was confirmed by DNA sequencing (Eurofins Genomics).

4.4.2. Bacterial transformation and plasmid amplification

Bacterial transformation was achieved by electroporation. To generate electrocompetent cells, bacterial cultures were grown to an OD_{600} of 0.4-0.5, cooled on ice and harvested by centrifugation at 3000 x g at 4 °C for 10 min. Cells were washed twice by resuspension in ice-cold ultrapure water and subsequent centrifugation. After a final wash with ice-cold 10% (v/v)

glycerol, cells were resuspended in ice-cold 10% glycerol to a calculated OD_{600} of 200 and frozen in liquid nitrogen for storage at -80 °C.

For electroporation, electrocompetent cells were thawed on ice and 20-45 µl were mixed with 25-50 ng of plasmid DNA or 5-8 µl of desalted ligation reactions. Mixtures were transferred to pre-cooled electroporation cuvettes (1 mm gap, Bio-Rad) and pulsed with 200 Ω, 25 µF, 1800 V using a GenePulser Xcell (Bio-Rad). Cells were resuspended in pre-warmed LB or SOC medium (New England Biolabs) without antibiotics and grown at 37 °C for 1 h before plating onto LB agar containing appropriate antibiotics for the selection of transformants.

For the amplification of plasmid DNA, a single colony of *E. coli* DH5α or Top10 carrying the plasmid of interest was inoculated into LB medium and grown with antibiotic selection at 37 °C for 16 h. Cells were harvested by centrifugation and plasmids purified using commercial DNA miniprep kits.

4.5. Biochemical methods

4.5.1. Recombinant gene expression

S. Typhimurium genes were heterologously expressed in *E. coli* BL21 (DE3). A list of the plasmid constructs used in this study can be found in Table 4.7.

For expression, a single colony carrying the indicated plasmids was inoculated into LB medium and grown for 16 h at 30-37 °C. Cells were sub-cultured by 1:50 dilution into fresh medium. If more than one antibiotic was used for selection in these sub-cultures, half the normal concentration was used for each antibiotic. Sub-cultures were grown at 37 °C to an OD_{600} of 0.4-0.6 and cooled to 20 °C before expression was induced by addition of 0.3-0.5 mM IPTG for pET vectors, 1 mM IPTG for Duet vectors, and 200 µg/l anhydrotetracycline (AHT) for pASK-IBA vectors. Cells were grown for another 18 h, harvested by centrifugation at 6000 x g at 4 °C and frozen in liquid nitrogen for subsequent storage at -80 °C. In the case of *spaO+orgB+invC*-expressing cells to be used for protein purification by high-pressure liquid chromatography, 100 µg/l AHT and 0.75-1 mM IPTG were used for induction, and glucose was initially added to the medium at a concentration of 1% (w/v). Since glucose was found to affect neither expression levels nor final cell density, it was omitted in later experiments. In the case of *sicP+sptP*$_{1\text{-}158}$, expression was induced with 1 mM IPTG and carried out for 5 h at 30 °C.

Table 4.7. Plasmids used in heterologous protein expression. All plasmids were generated for this study using the methods outlined in section 4.4.1, with the exception of pASK-IBA5+ *spaO*-Strep, pET21a *spaO*(-His), pET28a *orgB*(-His), pET28a *orgA*, pET28a His-*sicP* and pET21a *sptP*$_{1\text{-}158}$-His, which were provided by Nicola Barison. pASK-IBA5+ Strep-*spaO* and pASK-IBA5+ *spaO*$_{V203A}$-Strep were provided by Ivonne Bernal. pASK-IBA3C+RBS is a variant of pASK-IBA3C in which the ribosome binding site has been replaced by that of pASK-IBA3+.

A) Small-scale affinity purification and solubility analyses (Figure 2.1 to Figure 2.4, Figure 2.6, Figure 2.14)

Protein combination	Constructs (co-)transformed
SpaO/SpaO$_C$	pASK-IBA5+ *spaO*-Strep
OrgB	pET28a *orgB*-His
SpaO/SpaO$_C$ + OrgB	pCDFDuet-1 *orgB*-His+*spaO*-Strep
PrgH$_{1\text{-}140}$ + SpaO/SpaO$_C$	pASK-IBA3C+RBS *prgH*$_{1\text{-}140}$-His, pASK-IBA5+ *spaO*-Strep
	pASK-IBA3C+RBS *prgH*$_{1\text{-}140}$-Strep, pET21a *spaO*-His
OrgA	pCOLADuet-1 *orgA*-Strep
OrgA + PrgH$_{1\text{-}140}$	pCOLADuet-1 *orgA*+*prgH*$_{1\text{-}140}$
	pCOLADuet-1 *orgA*-Strep+*prgH*$_{1\text{-}140}$
	pCOLADuet-1 *orgA*+*prgH*$_{1\text{-}140}$-Strep
OrgA + SpaO/SpaO$_C$ + OrgB + PrgH$_{1\text{-}140}$	pASK-IBA3+ *orgA*-Strep, pCDFDuet-1 *orgB*+*spaO*, pASK-IBA3C+RBS *prgH*$_{1\text{-}140}$
SpaO/SpaO$_C$ + OrgB + PrgH$_{1\text{-}140}$	pCDFDuet-1 *orgB*+*spaO*, pASK-IBA3C+RBS *prgH*$_{1\text{-}140}$,
OrgA + SpaO/SpaO$_C$	pASK-IBA3+ *orgA*-Strep, pASK-IBA3C+RBS *spaO*
InvC	pASK-IBA3C+RBS *invC*-Strep
InvC + OrgB	pASK-IBA3C+RBS *invC*-Strep, pET28a *orgB*
InvC + OrgB + SpaO	pASK-IBA3C+RBS *invC*-Strep, pET28a *orgB*, pET21a *spaO*
InvC + OrgA	pASK-IBA3C+RBS *invC*-Strep, pET28a *orgA*
InvC + SpaO	pASK-IBA3C+RBS *invC*-Strep, pET21a *spaO*
SpaO	pET21a *spaO*
InvC	pACYCDuet-1 *invC*-Strep
InvI	pACYC-Duet-1 Strep-*invI*
InvC + InvI	pACYCDuet-1 *invC*+*invI*
	pACYCDuet-1 *invC*-Strep+*invI*
	pACYCDuet-1 *invC*+Strep-*invI*
InvI + InvA$_{357\text{–}685}$	pACYC-Duet-1 *invA*$_{357\text{–}685}$+Strep-*invI*
	pACYC-Duet-1 *invA*$_{357\text{–}685}$+*invI*
InvC + InvI + InvA$_{357\text{–}685}$	pACYCDuet-1 *invA*$_{357\text{–}685}$+*invI*, pASK-IBA3+ *invC*-Strep

Protein combination	Constructs (co-)transformed
OrgC + SpaO/SpaO$_C$	pASK-IBA3+ *orgC*-Strep, pASK-IBA3C+RBS *spaO*
SpaO/SpaO$_C$	pASK-IBA3C+RBS *spaO*
SpaO/SpaO$_C$ + OrgB	pCDFDuet-1 *orgB+spaO*
OrgC + SpaO/SpaO$_C$ + OrgB	pASK-IBA3+ *orgC*-Strep, pCDFDuet-1 *orgB+spaO*
OrgC + OrgA	pASK-IBA3+ *orgC*-Strep, pCOLADuet-1 *orgA*
OrgC + OrgA + PrgH$_{1-140}$	pASK-IBA3+ *orgC*-Strep, pCOLADuet-1 *orgA,* pASK-IBA3C+RBS *prgH$_{1-140}$*-His
OrgC + InvC + InvI	pASK-IBA3+ *orgC*-Strep, pACYCDuet-1 *invC+invI*
OrgC + OrgB + InvC + InvI	pASK-IBA3+ *orgC*-Strep, pCDFDuet-1 *orgB,* pACYCDuet-1 i*nvC+invI*
OrgB + InvC + InvI	pCDFDuet-1 *orgB*, pACYCDuet-1 i*nvC+invI*
OrgC + OrgB	pASK-IBA3+ *orgC*-Strep, pCDFDuet-1 *orgB*
OrgB	pCDFDuet-1 *orgB*
OrgB$_{1-105}$	pET28a *orgB$_{1-105}$*-His
OrgB$_{1-105}$ + InvC	pET28a *orgB$_{1-105}$*-His, pASK-IBA3+ *invC*-Strep
OrgB$_{1-105}$ + SpaO/SpaO$_C$	pET28a *orgB$_{1-105}$*-His, pASK-IBA5+ *spaO*-Strep
OrgB$_{106-226}$	pET28a His-*orgB$_{106-226}$*
OrgB$_{106-226}$+ InvC	pET28a His-*orgB$_{106-226}$,* pASK-IBA3+ *invC*-Strep
OrgB$_{106-226}$+ SpaO/SpaO$_C$	pET28a His-*orgB$_{106-226}$,* pASK-IBA5+ *spaO*-Strep
InvC + OrgB	pASK-IBA3C *invC*-Strep, pET28a *orgB*-His
	pACYCDuet-1 *invC*-Strep, pCDFDuet-1 *orgB*-His
InvC$_{80-431}$ + OrgB	pASK-IBA3+ *invC$_{80-431}$*-Strep, pCDFDuet-1 *orgB*-His
OrgB	pCDFDuet-1 *orgB*-His
InvC$_{1-79}$ + OrgB	pASK-IBA3C+RBS *invC$_{1-79}$*-Strep, pCDFDuet-1 *orgB*-His

B) Purification by high-pressure liquid chromatography (HPLC)

Protein combination	Constructs (co-)transformed
SpaO$_C$	pASK-IBA3+ *spaO$_C$*-Strep
SpaO/SpaO$_C$	pASK-IBA5+ *spaO*-Strep
	pASK-IBA5+ Strep-*spaO*
SpaO/SpaO$_C$/OrgB	pASK-IBA5+ *spaO*-Strep, pET28a *orgB*-His
SpaO$_{V203A}$/OrgB	pASK-IBA5+ *spaO$_{V203A}$*-Strep, pET28a *orgB*-His
SpaO/SpaO$_C$/OrgB/InvC	pCDFDuet-1 *orgB+spaO*, pASK-IBA3+ *invC*-Strep
SicP/SptP$_{1-158}$	pET28a His-*sicP*, pET21a *sptP$_{1-158}$*-His

C) Duet vector combinations tested to express different sorting platform sub-complexes (Figure 2.5)

pCOLADuet-1 *orgA*-Strep+*prgH*$_{1\text{-}140}$, pCDFDuet-1 *orgB+spaO*
pCOLADuet-1 *orgA+prgH*$_{1\text{-}140}$, pCDFDuet-1 *orgB+spaO*
pCDFDuet-1 *orgB+spaO*, pACYCDuet-1 *invC*+Strep-*invI*
pCDFDuet-1 *orgB+spaO,* pACYCDuet-1 *invC+invI*
pCOLADuet-1 *orgA*-Strep+*prgH*$_{1\text{-}140}$, pCDFDuet-1 *orgB+spaO,* pACYCDuet-1 *invC+invI*
pCOLADuet-1 *orgA+prgH*$_{1\text{-}140}$, pCDFDuet-1 *orgB+spaO,* pACYCDuet-1 *invC*+Strep-*invI*
pCOLADuet-1 *orgA+prgH*$_{1\text{-}140}$, pCDFDuet-1 *orgB+spaO,* pACYCDuet-1 *invC+invI*
pCDFDuet-1 *orgB,* pACYCDuet-1 *invC*-Strep
pCOLADuet-1 *orgA*-Strep, pETDuet-1 *spaO*
pCOLADuet-1 *orgA*-Strep, pCDFDuet-1 *spaO*
pCOLADuet-1 *orgA*-Strep, pASK-IBA3C *spaO*

4.5.2. Protein purification

Proteins used in this study were purified using immobilized metal ion affinity columns and resin, as well as ion-exchange and size-exclusion chromatography columns from GE Healthcare. Strep-Tactin columns and resins were purchased from IBA Lifesciences. All purification steps were performed on ice or at 4 °C. Buffers used can be found in Table 4.8. In all cases, buffers used to resuspend cells prior to lysis were supplemented with 1 mg/ml lysozyme, 5-10 µg/ml DNase I and cOmplete Ultra EDTA-free protease inhibitor cocktail (Roche).

Table 4.8. Buffers used in protein purification

Buffer	Ingredients
Buffer W	100 mM Tris-HCl pH 8.0 150 mM NaCl 2 mM 2-ME
Buffer E	100 mM Tris-HCl pH 8.0 150 mM NaCl 2 mM 2-ME 7.5 mM desthiobiotin (DTB)
Buffer A	20 mM sodium phosphate buffer pH 7.4 500 mM NaCl 2 mM 2-ME

Buffer	Ingredients
Buffer B	20 mM HEPES pH 7.5 350 mM NaCl 2 mM 2-ME
SEC buffer 1	20 mM HEPES pH 7.5 150 mM NaCl
SEC buffer 2	10 mM Tris-HCl pH 8.0 50 mM NaCl

4.5.2.1. Small-scale co-purification

4.5.2.1.1. Spin cup method

In most co-purification assays, purification of proteins was achieved by Strep-Tactin affinity chromatography using the spin cup method. Cells from 50 ml of culture were resuspended in 1.1-1.5 ml buffer W and lysed by sonication using a Bandelin Sonopuls HD 2070 with MS 72 sonotrode at 60% power with 8-10 repeats of 10 s sonication in pulsed mode (7x 10%) followed by 20 s of rest. In the case of InvI-InvC and $PrgH_{1-140}$-OrgA combinations expressed from Duet vectors (Figure 2.2B-C, Figure 2.4) cells from 100 ml culture were used and resuspended in 2 ml buffer W. Following lysis, soluble and insoluble fractions were separated by centrifugation at 125,000 x g for 1 h. The soluble fraction was applied 3 times to 90 µl of Strep-Tactin Superflow resin in a centrifuge column (Thermo Fisher Scientific) and loaded resins were washed three times with 350-400 µl buffer W. Loading and washing steps were performed using a mixture of gravity flow and centrifugation at 100 x g. Bound proteins were eluted from the resin with 100 µl buffer E supplemented with 1 mM EDTA using centrifugation at 100 x g for 1 min. In the case of *orgA+spaO+orgB+prgH* this was followed by a second elution step into additional 50 µl. For the purification of InvC/InvI, $InvC_{1-79}$ and $InvC_{80-431}$, as well as their combinations with other proteins, buffer W was supplemented with 1 mM $MgCl_2$ for both lysis and washing.

4.5.2.1.2. Batch method

For combinations of InvC with other proteins (Figure 2.3), purification was achieved using the batch method, in which cell lysates from 40 ml of culture were prepared as under 4.5.2.1.1 and cleared by centrifugation at 16,000 x g. Soluble fractions were filtered through 0.2 µm pore size cellulose acetate filters and added directly to Strep-Tactin beads in microcentrifuge tubes, incubated for 1 h at 4 °C and separated by centrifugation at 500 x g. Supernatants were removed and resins washed three times by resuspension with 900 µl buffer W followed by centrifugation at 500 x g. Bound proteins were eluted with 160 µl buffer E containing 5 mM DTB.

In the case of InvC-Strep+OrgB-His purification (Figure 2.14B), batch purification was performed using cells from 50 ml culture. For Strep-Tactin purification PBS was used as lysis buffer, loaded beads were washed three times with 500 µl PBS and bound proteins were eluted with 100 µl buffer E containing 5 mM DTB. For nickel immobilized metal ion affinity chromatography (IMAC) purification, buffer A + 40 mM imidazole was used as lysis and wash buffer and protein elution was with 150 µl buffer A + 400 mM imidazole.

4.5.2.2. Protein purification by high-pressure liquid chromatography

Protein complexes for biophysical analysis were purified by high-pressure liquid chromatography (HPLC) using Äkta chromatography systems. Cells were resuspended in the indicated buffers and lysed by French press or high-pressure homogenizer at pressures of 12,000-16,000 psi. Lysates were clarified by centrifugation at 48,000 x g for 30 min and the supernatant filtered through a 0.2 µm pore size cellulose acetate filter before loading onto the indicated affinity chromatography column.

SpaO/SpaO$_C$/OrgB complexes to be analyzed by native MS, analytical SEC and MALS were purified by a 2-step strategy of nickel IMAC and Strep-Tactin affinity chromatography using a C-terminal His-tag on OrgB and C-terminal Strep-tags on SpaO/SpaO$_C$, respectively. Cells were resuspended and lysed in buffer A + 40 mM imidazole and proteins were loaded onto a 5 ml HisTrap HP column. Bound protein was washed with buffer A + 40 mM imidazole and buffer A + 40 mM imidazole + 3 mM ATP + 10 mM $MgCl_2$, and eluted with buffer A + 400 mM imidazole. Eluted proteins were diluted 2-3-fold with buffer B and loaded onto a 5 ml Strep-Tactin Superflow high capacity column, washed with buffer B and eluted with buffer B + 5 mM DTB. For the analysis by SAXS, the IMAC protocol was modified to include an additional wash step

with buffer A + 80 mM imidazole prior to elution with buffer A + 320 mM imidazole, and buffer E was used for the elution from the Strep-Tactin column.

SpaO/SpaO$_C$/OrgB to be used in MST was purified by IMAC as above, with the modification that after the ATP-$MgCl_2$ wash protein was eluted with a linear gradient of 40-280 mM imidazole in buffer A. Elution fractions containing pure SpaO/SpaO$_C$/OrgB were pooled, concentrated using centrifugal filters (Merck Millipore) and further purified by SEC on a Superdex 200 16/60 column equilibrated with SEC buffer 1.

SpaO/SpaO$_C$ complexes were purified by Strep-Tactin affinity chromatography and SEC. Cells were lysed in buffer W, proteins loaded onto a 5 ml Strep-Tactin Superflow high capacity column, washed with buffer W and eluted with buffer E. Eluted protein was further purified by SEC on a Superdex 200 16/60 column equilibrated with SEC buffer 1. SpaO/SpaO$_C$ used in MST and native MS carried an N-terminal Strep-tag on SpaO.

SpaO$_C$ carrying a C-terminal Strep-tag was purified by Strep-Tactin affinity chromatography using a 1 ml Strep-Tactin Superflow high capacity column. Buffer B was used as lysis and wash buffer, and proteins were eluted with buffer B + 5 mM DTB.

For the purification of SicP/SptP$_{1-158}$ complexes (carrying an N-terminal His-tag on SicP and a C-terminal His-tag on SptP$_{1-158}$) to be used in MST and native MS, cells were lysed in buffer A + 16 mM imidazole. Samples were loaded onto a 1 ml HisTrap HP column, washed with buffer A + 45 mM imidazole and eluted with buffer A + 400 mM imidazole. Eluted protein was buffer-exchanged into SEC buffer 1 using a HiPrep 26/10 desalting column and the N-terminal His-tag of SicP was removed using the Thrombin CleanCleave kit (Sigma-Aldrich) according to the manufacturer's instructions. Protein complexes were subsequently purified by SEC on a Superdex 75 16/60 column in SEC buffer 1. SicP/SptP$_{1-158}$ used for MALS was purified using 20 mM HEPES pH 7.5 instead of sodium phosphate in the IMAC buffers and both thrombin digestion and preparative SEC were performed in 20 mM HEPES pH 7.5, 50 mM NaCl.

SpaO/SpaO$_C$/OrgB/InvC complexes (carrying a C-terminal Strep-tag on InvC) to be analyzed by SAXS were purified by Strep-Tactin affinity purification using a 5 ml Strep-Tactin Superflow high capacity column. Buffer W was supplemented with 1 mM EDTA for lysis and washing, and proteins were eluted from the column with buffer E. Samples used in the native MS of SpaO/SpaO$_C$/OrgB/InvC in combination with SicP/SptP$_{1-158}$ and ATPγS were purified analogously, but 7 mM 2-ME were included in buffer W and proteins were eluted using buffer E with 15 mM DTB. For MALS and native MS of SpaO/SpaO$_C$/OrgB/InvC without additives, cells

were lysed in buffer W using sonication (70% power, 7x 10% pulsed mode, 10 repeats of 30 s sonication followed by 30 s rest) and eluted using buffer E.

4.5.2.3. Establishing of SpaO/SpaO$_C$/OrgB/InvC purification strategy

For the development of a purification strategy for SpaO/SpaO$_C$/OrgB/InvC and the investigation of the complex's behavior under various purification conditions, different combinations and variations of Strep-Tactin affinity chromatography, dialysis, concentration, size-exclusion chromatography and ion-exchange chromatography were explored.

For Strep-Tactin affinity purification 5 ml Strep-Tactin high capacity columns were used. 7.5-10 g of cells (wet cell mass) were used in each preparation, buffer W with 1 mM EDTA served as lysis and washing buffer and proteins were eluted from the column using buffer E with 5 mM DTB. For higher protein yields in later experiments, 20-25 g of cells were lysed and protein eluted with buffer E with 7.5-15 mM DTB.

To remove contaminating DnaK, several modifications of the Strep-Tactin purification were tested: a) purification including wash steps with buffer W + 10 mM $MgCl_2$, followed by buffer W + 10 mM $MgCl_2$ + 3 mM ATP; b) purification using buffer W with 10% (v/v) glycerol, 7 mM 2-ME and 2 mM EDTA for lysis and initial washing, a washing step with the same buffer + 5 mM ATP + 15 mM $MgCl_2$, and elution with buffer E with 10% (v/v) glycerol and 7 mM 2-ME; c) purification using buffer W with 1 mM EDTA for lysis and initial washing, followed by washes with buffer W containing 10% (v/v) glycerol and 10% (v/v) glycerol + 3 mM ATP + 10 mM $MgCl_2$; d) lysis by sonication in buffer W with 1 mM EDTA and washing with buffer W + 3 mM ATP + 10 mM $MgCl_2$ and denatured proteins at an A_{280} of 0.3. Denatured proteins were derived by sonicating *E. coli* BL21 (DE3) in buffer W with no other additives except lysozyme, heating to 65 °C for 12 min and removal of insoluble components by centrifugation for 15 min at 12,000 x g and 4 °C (Rial & Ceccarelli, 2002).

Dialysis of Strep-Tactin-purified protein was achieved using Slide-A-Lyzer cassettes (Thermo Fisher Scientific) and samples were dialyzed for 16 h at 4 °C against SEC buffer 2 or buffer W without 2-ME. Where indicated, 1 mM EDTA, 2 mM $MgCl_2$, or 2 mM $MgCl_2$ + 0.5 mM ATP were included in the dialysis buffer.

SEC was performed at a flow rate of 0.3 ml/min on a Superose 6 10/300 GL column equilibrated with SEC buffer 2 using 100-200 µl sample. Where samples were concentrated prior to SEC, this was achieved using centrifugal filters (Merck Milipore).

Ion-exchange chromatography (IEX) was performed on a MonoQ 5/50 GL column. Strep-Tactin-purified samples were dialyzed into SEC buffer 2, loaded onto the column, washed with SEC buffer 2 and eluted with a linear gradient to 500 mM NaCl. Proteins that still bound to the column were eluted with 1000 mM NaCl.

4.5.3. Analytical size-exclusion chromatography (SEC)

Analytical SEC was performed at 4 °C on a Superdex 200 10/300 GL column (GE Healthcare) equilibrated with 20 mM HEPES pH 7.5 and 350 mM NaCl. 100-200 µl of SpaO/$SpaO_C$/OrgB (A_{280} of 3.3) or $SpaO_{V203A}$/OrgB (A_{280} of 2.0) that had been purified by nickel IMAC and Strep-Tactin affinity chromatography were injected into the column and resolved at a flow rate of 0.4 ml/min. Molecular masses of eluting proteins were estimated by comparison of their elution volumes to a calibration curve obtained using the Gel Filtration HMW Calibration Kit (GE Healthcare) under the same conditions (Supplementary Figure 1).

4.5.4. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed using the Tris-glycine buffer system. Samples were mixed with 4x-concentrated SDS sample buffer and resolved on self-cast 15% polyacrylamide gels at a constant voltage of 15 mA. Exceptions were samples containing $InvC_{1-79}$ (Figure 2.14C), which were resolved using 18% polyacrylamide gels, and InvC in combinations with SpaO, OrgB and OrgA (Figure 2.3), for which a Bio-Rad Criterion TGX AnyKd gel was used. Gels were stained in Coomassie staining solution for 16 h and destained in deionized water. PageRuler Unstained and PageRuler Plus Prestained protein ladders were used as size standards.

4.5.5. Western blot

Samples were resolved by SDS-PAGE and transferred to methanol-activated PVDF membranes (Amersham Hybond, GE Healthcare) at 1.7-2.4 mA/cm^2 for 60-80 min using a semi-dry blotting

system. Membranes were blocked by incubation in 5% (w/v) skim milk powder in PBS-T for 1 h, washed three times for 10 min with PBS-T and incubated with primary antibody solution for 16 h at 4 °C. Membranes were again washed three times with PBS-T and incubated for 30 min with secondary antibodies conjugated to horseradish peroxidase (HRP). After another three washes with PBS-T, enhanced chemiluminescence (ECL) substrate (SuperSignal West Dura, Thermo Fisher Scientific or Clarity Max, Bio-Rad) was added to the membranes and the chemiluminescent signal detected using ImageQuant LAS 4000 and 4000 mini imagers (GE Healthcare).

Antibodies used in this study were mouse monoclonal antibodies against Strep-tag (Qiagen) and 6xHis-tag (GE Healthcare and Thermo Fisher Scientific), as well as an HRP-coupled goat anti-mouse secondary antibody (Jackson ImmunoResearch). Antibodies were diluted in PBS-T and used at the concentrations recommended by the manufacturer.

Stripping of blots was achieved by incubation with Roti-Free stripping buffer for 30 min at 56 °C. Stripped blots were washed two times for 20 min with PBS-T and blocked with 5% (w/v) skim milk powder in PBS-T for 1 h before application of new primary antibodies at room temperature for 4 h.

4.5.6. TCA precipitation

Low-concentration elution fractions from SEC and IEX runs were concentrated by precipitation with trichloroacetic acid (TCA) prior to analysis by SDS-PAGE. TCA was added to a final concentration of 10% (v/v) and proteins precipitated for 30 min at -20 °C. Precipitate was pelleted by centrifugation at 16,000 x g for 1 h at 4 °C, washed with ice-cold acetone and centrifuged again for 1 h. Supernatants were discarded, samples air-dried and resuspended in SDS sample buffer with increased buffer capacity (i.e. 1x SDS sample buffer with 360 mM Tris-HCl pH 6.8).

4.5.7. Microscale thermophoresis (MST)

MST was performed in MST buffer, which contained 20 mM HEPES pH 7.5, 150 mM NaCl, 0.05% (v/v) Tween 20. SicP/SptP$_{1\text{-}158}$ was labeled with the fluorescent dye NT-647 using the Monolith Protein Labeling Kit RED-NHS (NanoTemper Technologies) according to the

manufacturer's instructions. Briefly, SicP/SptP$_{1-158}$ was diluted to 20 µM in the provided labeling buffer and incubated with a 3-fold excess of dye for 30 min in the dark. Unreacted free dye was removed using the supplied buffer-exchange SEC column pre-equilibrated with MST buffer. Unlabeled ligand proteins were concentrated to A_{280} of 20 (SicP/SptP$_{1-158}$), 61 or 80 (*S*paO/SpaO$_C$), and 45 or 71 (SpaO/SpaO$_C$/OrgB) using centrifugal filters (Merck Millipore). A dilution series of 16 1:1 dilutions of each ligand sample was prepared and each dilution was mixed 1:1 (v:v) with labeled SicP/SptP$_{1-158}$ and transferred to standard treated capillaries (NanoTemper Technologies) for analysis in a Monolith NT.115 (NanoTemper Technologies). Labeled SicP/SptP$_{1-158}$ was kept constant at a final concentration of 100 nM in each experiment and the excitation power of the instrument was set to 10-20 % to achieve raw fluorescence counts between 900 and 1100. MST was performed at 25 °C and medium MST power (40 %), with laser off and on times of 5 s and 30 s, respectively. The results were analyzed with the program MO.Affinity Analysis 2.2.4 using the "10 s On Time" preset. For the interaction of SicP/SptP$_{1-158}$ with itself, three series of measurements were included in the calculation of the binding curve: for two, the concentrated ligand was diluted in two separate dilution series and MST data was acquired, while for a third measurement one of the dilution series was extended by five additional steps and MST measured for the 16 most-diluted sample.

4.6. Biophysical methods

4.6.1. Native mass spectrometry

For the analysis by native mass spectrometry (native MS), proteins were exchanged into ammonium acetate buffer. In the case of SpaO/SpaO$_C$/OrgB/InvC this was achieved by concentrating Strep-Tactin-purified samples to an A_{280} of 23 and running SEC over a Superose 6 10/300 GL column in 60 mM ammonium acetate pH 8.0, followed by 1:1 dilution in 300 mM ammonium acetate for a final concentration of 180 mM. SpaO/SpaO$_C$ and SicP/SptP$_{1-158}$ were purified as described above and buffer-exchanged into 150 mM ammonium acetate pH 8.0 using Vivaspin 500 MWCO centrifugal concentrators (Sartorius), SpaO/SpaO$_C$/OrgB was exchanged into 300 mM buffer by the same method. In some cases, such as for the mixing with SicP/SptP$_{1-158}$ and ATPγS, SpaO/SpaO$_C$/OrgB/InvC was purified by Strep-Tactin affinity chromatography and exchanged into 300 mM ammonium acetate pH 8.0 using centrifugal concentrators. When indicated, stock solutions of 1 mM ATPγS and 10 mM magnesium acetate were prepared in water and added to a final concentration of 0.2 mM and 1 mM, respectively.

Native MS experiments were performed on a Quadrupole Time-of-Flight 2 (QTof2) mass spectrometer (Waters and MS Vision) adapted for high-mass experiments (van den Heuvel et al., 2006), using nano-electrospray ionization (ESI) with gold-coated glass capillaries prepared in-house (Dunne et al., 2016). Experiments were conducted in positive ion mode with 1.3 to 1.5 kV capillary voltage, 110 to 150 V cone voltage, 6 to 10 mbar source pressure and 1.7 to 1.9 $x10^{-2}$ mbar of argon in the collision cell. Collision voltages were optimized for resolution and minimal complex dissociation. For tandem MS experiments, specific precursor peaks were selected and the collision energies were ramped up to 400 V or until disappearance of the precursor signal. MassLynx (Waters) and Massign (Morgner & Robinson, 2012) software were used to assign peak series and raw data was calibrated using spectra of 25 mg/ml cesium acetate acquired on the day of each measurement. Average experimental masses of protein complexes, standard deviations and average full-width at half-maximum (FWHM) values as a measure of resolution and mass heterogeneity are given in Table 2.1 and Table 2.2. Native MS was performed by Johannes Heidemann from the laboratory of Charlotte Uetrecht.

4.6.2. Multi-angle light scattering (MALS)

For MALS analysis of SpaO/SpaO$_C$/OrgB/InvC, Strep-Tactin-purified samples were concentrated to an A_{280} of 22 and run over a Superose 6 10/300 column in 10 mM Tris-HCl pH 8.0, 50 mM NaCl with a flow rate of 0.3 ml/min. The column was followed by an in-line array of multi-angle light scattering (MiniDawn Treos, Wyatt) and refractive index detectors (Optilab T-rEX, Wyatt). Molecular masses of SpaO/SpaO$_C$/OrgB/InvC were calculated by Cy Jeffries of the European Molecular Biology Laboratory (EMBL) Hamburg using the ASTRA software (version 7, Wyatt).

In the case of SpaO/SpaO$_C$/OrgB, samples were purified as described in section 4.5.2.2 and run over a Superdex 200 10/300 column as described in 4.5.3. The column was followed by an in-line array of UV absorption, multi-angle light scattering (miniDawn Tristar, Wyatt) and refractive index detectors (RI-101; Shodex). Molecular masses were calculated with ASTRA software (version 5.3.4.20, Wyatt) using a dn/dc value of 0.126 based on calibration experiments using chymotrypsinogen A, conalbumin, ovalbumin and aldolase.

4.6.3. Small-angle X-ray scattering (SAXS)

Small-angle X-ray scattering data was collected at the European Molecular Biology Laboratory (EMBL) beamline P12 at the PETRA III storage ring (DESY, Hamburg, Germany) using a Pilatus 2M detector (Dectris). The scattering intensity was recorded in the momentum transfer (q) range of 0.008-0.47 $Å^{-1}$ ($q = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle and $\lambda = 1.24$ Å is the X-ray wavelength) with the camera set to a sample-detector distance of 3.1 m. Protein was concentrated to an A_{280} of 21 using centrifugal filters and 75 µl were run at room temperature at a flow rate of 0.3 ml/min over a Superose 6 10/300 column equilibrated with 10 mM Tris-HCl pH 8.0, 50 mM NaCl. Eluting protein was directed to the SAXS flow cell, where the scattering profile was measured with an exposure time of 1 s per frame. Sample and buffer regions of the elution profile were separately averaged and CHROMIXS (Franke et al., 2017) was used to subtract the buffer from the sample signal. The experimental SAXS pattern was used to calculate the radius of gyration Rg and the forward scattering intensity I(0) using both Guinier analysis (Guinier, 1939) and the indirect Fourier transformation approach of the program GNOM (Svergun, 1992). The distance distribution function *P(r)* and the maximum particle dimension D_{max} were also calculated using GNOM. SAXS data analysis was performed by Anne Tuukkanen of the EMBL Hamburg.

4.6.4. SAXS modeling

SAXS *ab initio* models of SpaO/$SpaO_C$/OrgB/InvC were reconstructed with the bead modeling program DAMMIF (Franke & Svergun, 2009). A representative model was obtained by averaging ten independent DAMMIF reconstructions using the program DAMAVER (Volkov & Svergun, 2003). This model was also used to derive the excluded volume of the particle V_{DAM} and an estimate of its molecular mass (empirically, $MM_{DAM} \sim V_{DAM}/2$). A Fourier shell correlation-based approach was used to calculate the resolution of the *ab initio* model (Tuukkanen et al., 2016). Multiphase *ab initio* modeling was performed using the program MONSA (Svergun, 1999).

For the generation of the SAXS hybrid model a combination of available high-resolution crystal structures and fragment-based structural models was used. Structural models for OrgB and InvC were obtained using I-TASSER fragment-based modeling (Yang et al., 2015) and dimeric $SpaO_C$ was based on the crystal structure of the SpaO SPOA2-SPOA2 domain dimer (PDB ID: 4YX1 (Notti et al., 2015)). For the model of full-length SpaO, an I-TASSER model of the N-terminal domain of SpaO ($SpaO_{1-145}$) was generated and validated using experimental SAXS data (Bernal,

2019). This was combined with the crystal structure of the SpaO C-terminal SPAO1-SPOA2 domain dimer (PDB ID: 4YX5 (Notti et al., 2015)) and the resulting composite model of full-length SpaO was optimized by nanosecond time-scale all-atom molecular dynamics (MD) simulation using the program NAMD (Phillips et al., 2005) with CHARMM27 force field for the protein (MacKerell et al., 1998), the TIP3P solvent model for water (Jorgensen et al., 1983), as well as constant particle number, pressure and temperature. The program VMD was used for simulation setup and trajectory analysis (Humphrey et al., 1996). When combined with the model of dimeric $SpaO_C$, the MD-optimized model of full-length SpaO was found to fit well with experimental SAXS data of purified SpaO-2$SpaO_C$ (Bernal, 2019). The interaction between SpaO and OrgB was modeled using the crystal structure of the SpaO SPOA1-SPOA2 dimer in complex with the OrgB N-terminus (PDB ID: 4YX7 (Notti et al., 2015)) and the interaction between InvC and OrgB is based on the interaction seen in the crystal structure of the homologous flagellar FliH-FliI complex (PDB ID: 5B0O (Imada et al., 2016)). $SpaO_C$ was constrained to be in the proximity of the SpaO N-terminal domain based on the stable interaction between these units observed in both SEC-MALS and native MS (Bernal, Börnicke, et al., 2019; Bernal, 2019), and OrgB dimerization was ensured by constraining proximity between OrgB units at residue R56, as well as at residue E99. SAXS-based hybrid modeling was done using the program CORAL (Petoukhov & Svergun, 2005). The structures and models were defined as rigid bodies and missing structural features and linkers were modeled as flexible polypeptide chains. A schematic overview of the used models and regions that were modeled as flexible linkers can be found in Supplementary Figure 3. SAXS-based modeling was performed by Anne Tuukkanen of the EMBL Hamburg.

4.6.5. Superposition of SAXS model and CET map

Superposition of the SAXS *ab initio* model with the cryo-electron tomography (CET) map of the *Salmonella* T3SS (EMDB ID: EMD-8544) was performed with the program Chimera (Pettersen et al., 2004). The CET map was used at the author-recommended contour level of 2.53. Superposition was performed by Michele Lunelli from the laboratory of Michael Kolbe.

5. References

Abby SS, Rocha EPC. 2012. “The Non-Flagellar Type III Secretion System Evolved from the Bacterial Flagellum and Diversified into Host-Cell Adapted Systems.” *PLOS Genetics* 8 (9): e1002983.

Abrusci P, Vergara-Irigaray M, Johnson S, Beeby MD, Hendrixson DR, Roversi P, Friede ME, Deane JE, Jensen GJ, Tang CM, et al. 2013. “Architecture of the Major Component of the Type III Secretion System Export Apparatus.” *Nature Structural & Molecular Biology* 20 (1): 99–104.

Akeda Y, Galán JE. 2005. “Chaperone Release and Unfolding of Substrates in Type III Secretion.” *Nature* 437 (7060): 911–15.

Allaoui A, Sansonetti PJ, Parsot C. 1992. “MxiJ, a Lipoprotein Involved in Secretion of Shigella Ipa Invasins, Is Homologous to YscJ, a Secretion Factor of the Yersinia Yop Proteins.” *Journal of Bacteriology* 174 (23): 7661–7669.

Allison SE, Tuinema BR, Everson ES, Sugiman-Marangos S, Zhang K, Junop MS, Coombes BK. 2014. “Identification of the Docking Site between a Type III Secretion System ATPase and a Chaperone for Effector Cargo.” *Journal of Biological Chemistry* 289 (34): 23734–44.

Armentrout EI, Rietsch A. 2016. “The Type III Secretion Translocation Pore Senses Host Cell Contact.” *PLOS Pathogens* 12 (3): e1005530.

Arnold R, Brandmaier S, Kleine F, Tischler P, Heinz E, Behrens S, Niinikoski A, Mewes H-W, Horn M, Rattei T. 2009. “Sequence-Based Prediction of Type III Secreted Proteins.” *PLOS Pathogens* 5 (4): e1000376.

Bai F, Morimoto YV., Yoshimura SDJ, Hara N, Kami-Ike N, Namba K, Minamino T. 2014. “Assembly Dynamics and the Roles of FliI ATPase of the Bacterial Flagellar Export Apparatus.” *Scientific Reports* 4: 6528.

Benesch JLP. 2009. “Collisional Activation of Protein Complexes: Picking Up the Pieces.” *Journal of the American Society for Mass Spectrometry* 20 (3): 341–48.

Bergeron JRC, Fernández L, Wasney GA, Vuckovic M, Reffuveille F, Hancock REW, Strynadka NCJ. 2016. “The Structure of a Type 3 Secretion System (T3SS) Ruler Protein Suggests a Molecular Mechanism for Needle Length Sensing.” *Journal of Biological Chemistry* 291 (4): 1676–91.

Bergeron JRC, Worrall LJ, Sgourakis NG, DiMaio F, Pfuetzner RA, Felise HB, Vuckovic M, Yu AC, Miller SI, Baker D, et al. 2013. “A Refined Model of the Prototypical Salmonella SPI-1 T3SS Basal Body Reveals the Molecular Basis for Its Assembly.” *PLOS Pathogens* 9 (4): e1003307.

Bernal I. 2019. “Structural-Functional Characterization of Soluble Components of the Salmonella Type III Secretion System Sorting Platform.” *Unpublished Doctoral Thesis*. Humboldt Universität Berlin.

Bernal I, Börnicke J, Heidemann J, Svergun D, Horstmann JA, Erhardt M, Tuukkanen A, Uetrecht C, Kolbe M. 2019. "Molecular Organization of Soluble Type III Secretion System Sorting Platform Complexes." *Journal of Molecular Biology* 431 (19): 3787–3803.

Bernal I, Römermann J, Flacht L, Lunelli M, Uetrecht C, Kolbe M. 2019. "Structural Analysis of Ligand-Bound States of the Salmonella Type III Secretion System ATPase InvC." *Protein Science* 28 (10): 1888–1901.

Blaylock B, Riordan KE, Missiakas DM, Schneewind O. 2006. "Characterization of the Yersinia Enterocolitica Type III Secretion ATPase YscN and Its Regulator, YscL." *Journal of Bacteriology* 188 (10): 3525–34.

Botteaux A, Sory MP, Biskri L, Parsot C, Allaoui A. 2009. "MxiC Is Secreted by and Controls the Substrate Specificity of the Shigella Flexneri Type III Secretion Apparatus." *Molecular Microbiology* 71 (2): 449–60.

Buchko GW, Niemann G, Baker ES, Belov ME, Smith RD, Heffron F, Adkins JN, McDermott JE. 2010. "A Multi-Pronged Search for a Common Structural Motif in the Secretion Signal of Salmonella Enterica Serovar Typhimurium Type III Effector Proteins." *Molecular BioSystems* 6 (12): 2448–58.

Buchmeier NA, Heffron F. 1991. "Inhibition of Macrophage Phagosome-Lysosome Fusion by Salmonella Typhimurium." *Infection and Immunity* 59 (7): 2232–2238.

Burgess JL, Burgess RA, Morales Y, Bouvang JM, Johnson SJ, Dickenson NE. 2016. "Structural and Biochemical Characterization of Spa47 Provides Mechanistic Insight into Type III Secretion System ATPase Activation and Shigella Virulence Regulation." *Journal of Biological Chemistry* 291 (50): 25837–52.

Burgess JL, Jones HB, Kumar P, Toth RT, Middaugh CR, Antony E, Dickenson NE. 2016. "Spa47 Is an Oligomerization-Activated Type Three Secretion System (T3SS) ATPase from Shigella Flexneri." *Protein Science* 25 (5): 1037–48.

Büttner D. 2012. "Protein Export According to Schedule: Architecture, Assembly, and Regulation of Type III Secretion Systems from Plant- and Animal-Pathogenic Bacteria." *Microbiology and Molecular Biology Reviews* 76: 262–310.

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 (8): 1669–77.

Carter PB, Collins FM. 1974. "The Route of Enteric Infection in Normal Mice." *Journal of Experimental Medicine* 139 (5): 1189–1203.

Case HB, Dickenson NE. 2018. "MxiN Differentially Regulates Monomeric and Oligomeric Species of the Shigella Type Three Secretion System ATPase Spa47." *Biochemistry* 57 (15): 2266–77.

Chen L, Ai X, Portaliou AG, Minetti CASA, Remeta DP, Economou A, Kalodimos CG. 2013. "Substrate-Activated Conformational Switch on Chaperones Encodes a Targeting Signal in Type III Secretion." *Cell Reports* 3 (3): 709–15.

Cheng LW, Anderson DM, Schneewind O. 1997. "Two Independent Type III Secretion Mechanisms for YopE in Yersinia Enterocolitica." *Molecular Microbiology* 24 (4): 757–65.

Cheng LW, Kay O, Schneewind O. 2001. "Regulated Secretion of YopN by the Type III Machinery of Yersinia Enterocolitica." *Journal of Bacteriology* 183 (18): 5293–5301.

Claret L, Calder SR, Higgins M, Hughes C. 2003. "Oligomerization and Activation of the FliI ATPase Central to Bacterial Flagellum Assembly." *Molecular Microbiology* 48: 1349–1355.

Coburn B, Sekirov I, Finlay BB. 2007. "Type III Secretion Systems and Disease." *Clinical Microbiology Reviews* 20 (4): 535–49.

Crago AM, Koronakis V. 1998. "Salmonella InvG Forms a Ring-like Multimer That Requires the InvH Lipoprotein for Outer Membrane Localization." *Molecular Microbiology* 30 (1): 47–56.

Criss AK, Casanova JE. 2003. "Coordinate Regulation of Salmonella Enterica Serovar Typhimurium Invasion of Epithelial Cells by the Arp2/3 Complex and Rho GTPases." *Infection and Immunity* 71 (5): 2885–91.

Crump JA, Wain J. 2017. "Salmonella." In *International Encyclopedia of Public Health, 2nd Edition*, 425–33. Elsevier Inc.

Daefler S, Russel M. 1998. "The Salmonella Typhimurium InvH Protein Is an Outer Membrane Lipoprotein Required for the Proper Localization of InvG." *Molecular Microbiology* 28 (6): 1367–80.

Dean P. 2011. "Functional Domains and Motifs of Bacterial Type III Effector Proteins and Their Roles in Infection." *FEMS Microbiology Reviews* 35 (6): 1100–1125.

Deane JE, Abrusci P, Johnson S, Lea SM. 2010. "Timing Is Everything: The Regulation of Type III Secretion." *Cellular and Molecular Life Sciences* 67 (7): 1065–75.

Deng W, Marshall NC, Rowland JL, McCoy JM, Worrall LJ, Santos AS, Strynadka NCJ, Finlay BB. 2017. "Assembly, Structure, Function and Regulation of Type III Secretion Systems." *Nature Reviews Microbiology* 15 (6): 323–37.

Diepold A, Amstutz M, Abel S, Sorg I, Jenal U, Cornelis GR. 2010. "Deciphering the Assembly of the Yersinia Type III Secretion Injectisome." *EMBO Journal* 29 (11): 1928–40.

Diepold A, Armitage JP. 2015. "Type III Secretion Systems: The Bacterial Flagellum and the Injectisome." *Philosophical Transactions of the Royal Society B: Biological Sciences* 370 (1679).

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 Biology* 13 (1): e1002039.

Diepold A, Sezgin E, Huseyin M, Mortimer T, Eggeling C, Armitage JP. 2017. "A Dynamic and Adaptive Network of Cytosolic Interactions Governs Protein Export by the T3SS Injectisome." *Nature Communications* 8: 15940.

Diepold A, Wiesand U, Cornelis GR. 2011. "The Assembly of the Export Apparatus (YscR,S,T,U,V) of the Yersinia Type III Secretion Apparatus Occurs Independently of Other Structural Components and Involves the Formation of an YscV Oligomer." *Molecular Microbiology* 82: 502–514.

Dietsche T, Tesfazgi Mebrhatu M, Brunner MJ, Abrusci P, Yan J, Franz-Wachtel M, Schärfe C, Zilkenat S, Grin I, Galán JE, et al. 2016. "Structural and Functional Characterization of the Bacterial Type III Secretion Export Apparatus." *PLOS Pathogens* 12 (12): e1006071.

Dohlich K, Zumsteg AB, Goosmann C, Kolbe M. 2014. "A Substrate-Fusion Protein Is Trapped inside the Type III Secretion System Channel in Shigella Flexneri." *PLOS Pathogens* 10 (1): e1003881.

Duhr S, Braun D. 2006. "Why Molecules Move along a Temperature Gradient." *Proceedings of the National Academy of Sciences of the United States of America* 103 (52): 19678–82.

Dunne M, Leicht S, Krichel B, Mertens HD, Thompson A, Krijgsveld J, Svergun DI, Gomez-Torres N, Garde S, Uetrecht C, et al. 2016. "Crystal Structure of the CTP1L Endolysin Reveals How Its Activity Is Regulated by a Secondary Translation Product." *Journal of Biological Chemistry* 291 (10): 4882–93.

Erhardt M, Mertens ME, Fabiani FD, Hughes KT. 2014. "ATPase-Independent Type-III Protein Secretion in Salmonella Enterica." *PLOS Genetics* 10 (11): e1004800.

Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM, et al. 2017. "ATSAS 2.8: A Comprehensive Data Analysis Suite for Small-Angle Scattering from Macromolecular Solutions." *Journal of Applied Crystallography* 50: 1212–25.

Franke D, Svergun DI. 2009. "DAMMIF, a Program for Rapid Ab-Initio Shape Determination in Small-Angle Scattering." *Journal of Applied Crystallography* 42: 342–46.

Fraser GM, Gonzalez-Pedrajo B, Tame JR, Macnab RM. 2003. "Interactions of FliJ with the Salmonella Type III Flagellar Export Apparatus." *Journal of Bacteriology* 185 (18): 5546–54.

Fu Y, Galan JE. 1998. "Identification of a Specific Chaperone for SptP, a Substrate of the Centisome 63 Type III Secretion System of Salmonella Typhimurium." *Journal of Bacteriology* 180 (13): 3393–99.

Galán JE. 2009. "Common Themes in the Design and Function of Bacterial Effectors." *Cell Host and Microbe* 5 (6): 571–79.

Gonzalez-Pedrajo B, Fraser GM, Minamino T, Macnab RM. 2002. "Molecular Dissection of Salmonella FliH, a Regulator of the ATPase FliI and the Type III Flagellar Protein Export Pathway." *Molecular Microbiology* 45 (4): 967–82.

Green ER, Mecsas J. 2016. "Bacterial Secretion Systems: An Overview." *Microbiology Spectrum* 4 (1).

Guinier A. 1939. "La Diffraction Des Rayons X Aux Très Petits Angles : Application à l'étude de Phénomènes Ultramicroscopiques." *Annales de Physique* 11 (12): 161–237.

Guo EZ, Desrosiers DC, Zalesak J, Tolchard J, Berbon M, Habenstein B, Marlovits T, Loquet A, Galán JE. 2019. "A Polymorphic Helix of a Salmonella Needle Protein Relays Signals Defining Distinct Steps in Type III Secretion." *PLOS Biology* 17 (7): e3000351.

Haraga A, Ohlson MB, Miller SI. 2008. "Salmonellae Interplay with Host Cells." *Nature Reviews Microbiology* 6 (1): 53–66.

Held D, Yaeger K, Novy R. 2003. "New Coexpression Vectors for Expanded Compatibilities in E. Coli." *Novagen InNovations Newsletter* 18: 4–6.

Held D, Yaeger K, Novy R. 2004. "PCOLADuet™-1, a Lower-Copy, KanR Vector Option for Coexpression in E. Coli." *Novagen InNovations Newsletter* 19: 17–19.

Ho O, Rogne P, Edgren T, Wolf-Watz H, Login FH, Wolf-Watz M. 2017. "Characterization of the Ruler Protein Interaction Interface on the Substrate Specificity Switch Protein in the Yersinia Type III Secretion System." *Journal of Biological Chemistry* 292 (8): 3299–3311.

Hodgkinson JL, Horsley A, Stabat D, Simon M, Johnson S, da Fonseca PC, Morris EP, Wall JS, Lea SM, Blocker AJ. 2009. "Three-Dimensional Reconstruction of the Shigella T3SS Transmembrane Regions Reveals 12-Fold Symmetry and Novel Features Throughout." *Nature Structural & Molecular Biology* 16 (5): 477–85.

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 (6): 1065-1074.

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." *Proceedings of the National Academy of Sciences of the United States of America* 112: 1047–52.

Hu J, Worrall LJ, Vuckovic M, Hong C, Deng W, Atkinson CE, Brett Finlay B, Yu Z, Strynadka NCJ. 2019. "T3S Injectisome Needle Complex Structures in Four Distinct States Reveal the Basis of Membrane Coupling and Assembly." *Nature Microbiology*.

Hueck CJ. 1998. "Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants." *Microbiology and Molecular Biology Reviews* 62 (2): 379–433.

Hume PJ, Singh V, Davidson AC, Koronakis V. 2017. "Swiss Army Pathogen: The Salmonella Entry Toolkit." *Frontiers in Cellular and Infection Microbiology* 7.

Humphrey W, Dalke A, Schulten K. 1996. "VMD: Visual Molecular Dynamics." *Journal of Molecular Graphics* 14 (1): 33-38.

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." *Nature Structural & Molecular Biology* 18: 277–82.

Ibuki T, Uchida Y, Hironaka Y, Namba K, Imada K, Minamino T. 2013. "Interaction between FliJ and FlhA, Components of the Bacterial Flagellar Type III Export Apparatus." *Journal of Bacteriology* 195 (3): 466–73.

Imada K, Minamino T, Tahara A, Namba K. 2007. "Structural Similarity between the Flagellar Type III ATPase FliI and F1-ATPase Subunits." *Proceedings of the National Academy of Sciences of the United States of America* 104 (2): 485–90.

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." *Proceedings of the National Academy of Sciences of the United States of America* 113 (13): 3633–38.

Ishibashi Y, Arai T. 1990. "Specific Inhibition of Phagosome-Lysosome Fusion in Murine Macrophages Mediated by Salmonella Typhimurium Infection." *FEMS Microbiology Letters* 64 (1): 35–43.

Jackson MW, Plano G V. 2000. “Interactions between Type III Secretion Apparatus Components from Yersinia Pestis Detected Using the Yeast Two-Hybrid System.” *FEMS Microbiology Letters* 186: 85–90.

James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Abbastabar H, Abd-Allah F, Abdela J, Abdelalim A, et al. 2018. “Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 354 Diseases and Injuries for 195 Countries and Territories, 1990-2017: A Systematic Analysis for the Global Burden of Disease Study 2017.” *The Lancet* 392 (10159): 1789–1858.

Johnson S, Kuhlen L, Deme JC, Abrusci P, Lea SM. 2019. “The Structure of an Injectisome Export Gate Demonstrates Conservation of Architecture in the Core Export Gate between Flagellar and Virulence Type III Secretion Systems.” *MBio* 10 (3): e00818-19.

Jones BD, Ghori N, Falkow S. 1994. “Salmonella Typhimurium Initiates Murine Infection by Penetrating and Destroying the Specialized Epithelial M Cells of the Peyer’s Patches.” *Journal of Experimental Medicine* 180 (1): 15–23.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. “Comparison of Simple Potential Functions for Simulating Liquid Water.” *Journal of Chemical Physics* 79 (2): 926–35.

Journet L, Agrain C, Broz P, Cornelis GR. 2003. “The Needle Length of Bacterial Injectisomes Is Determined by a Molecular Ruler.” *Science* 302 (5651): 1757–60.

Kato J, Dey S, Soto JE, Butan C, Wilkinson MC, De Guzman RN, Galan JE. 2018. “A Protein Secreted by the Salmonella Type III Secretion System Controls Needle Filament Assembly.” *eLife* 7: e35886.

Kazetani K, Minamino T, Miyata T, Kato T, Namba K. 2009. “ATP-Induced FliI Hexamerization Facilitates Bacterial Flagellar Protein Export.” *Biochemical and Biophysical Research Communications* 388 (2): 323–27.

Kebarle P, Verkerk UH. 2009. “Electrospray: From Ions in Solution to Ions in the Gas Phase, What We Know Now.” *Mass Spectrometry Reviews* 28 (6): 898–917.

Kenjale R, Wilson J, Zenk SF, Saurya S, Picking WL, Picking WD, Blocker A. 2005. “The Needle Component of the Type III Secreton of Shigella Regulates the Activity of the Secretion Apparatus.” *Journal of Biological Chemistry* 280 (52): 42929–37.

Kimbrough TG, Miller SI. 2000. “Contribution of Salmonella Typhimurium Type III Secretion Components to Needle Complex Formation.” *Proceedings of the National Academy of Sciences of the United States of America* 97 (20): 11008–13.

Klein JR, Fahlen TF, Jones BD. 2000. “Transcriptional Organization and Function of Invasion Genes within Salmonella Enterica Serovar Typhimurium Pathogenicity Island 1, Including the PrgH,PrgI, PrgJ, PrgK, OrgA,OrgB, and OrgC Genes.” *Infection and Immunity* 68: 3368–76.

Konermann L, Ahadi E, Rodriguez AD, Vahidi S. 2013. “Unraveling the Mechanism of Electrospray Ionization.” *Analytical Chemistry* 85 (1): 2–9.

Kowal J, Chami M, Ringler P, Müller SA, Kudryashev M, Castaño-Díez D, Amstutz M, Cornelis GR, Stahlberg H, Engel A. 2013. “Structure of the Dodecameric Yersinia Enterocolitica Secretin YscC and Its Trypsin-Resistant Core.” *Structure* 21 (12): 2152–61.

Kubori T, Galán JE. 2002. "Salmonella Type III Secretion-Associated Protein InvE Controls Translocation of Effector Proteins into Host Cells." *Journal of Bacteriology* 184: 4699–4708.

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 S. 2018. "Structure of the Core of the Type III Secretion System Export Apparatus." *Nature Structural & Molecular Biology* 25 (7): 583–90.

Lara-Tejero M, Galán JE. 2009. "Salmonella Enterica Serovar Typhimurium Pathogenicity Island 1-Encoded Type III Secretion System Translocases Mediate Intimate Attachment to Nonphagocytic Cells." *Infection and Immunity* 77: 2635–42.

Lara-Tejero M, Kato J, Wagner S, Liu X, Galán JE. 2011. "A Sorting Platform Determines the Order of Protein Secretion in Bacterial Type III Systems." *Science* 331 (6021): 1188–91.

Lara-Tejero M, Qin Z, Hu B, Butan C, Liu J, Galán JE. 2019. "Role of SpaO in the Assembly of the Sorting Platform of a Salmonella Type III Secretion System." *PLOS Pathogens* 15 (1): e1007565.

Larock DL, Chaudhary A, Miller SI. 2015. "Salmonellae Interactions with Host Processes." *Nature Reviews Microbiology* 13 (4): 191–205.

Lee SH, Galán JE. 2004. "Salmonella Type III Secretion-Associated Chaperones Confer Secretion-Pathway Specificity." *Molecular Microbiology* 51 (2): 483–95.

Lyons BJE, Strynadka NCJ. 2019. "On the Road to Structure-Based Development of Anti-Virulence Therapeutics Targeting the Type III Secretion System Injectisome." *MedChemComm* 10 (8): 1273–89.

MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, et al. 1998. "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins." *Journal of Physical Chemistry B* 102 (18): 3586–3616.

Majewski DD, Worrall LJ, Hong C, Atkinson CE, Vuckovic M, Watanabe N, Yu Z, Strynadka NCJ. 2019. "Cryo-EM Structure of the Homohexameric T3SS ATPase-Central Stalk Complex Reveals Rotary ATPase-like Asymmetry." *Nature Communications* 10 (1).

Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, Jones TF, Fazil A, Hoekstra RM, for the International Collaboration on Enteric Disease "Burden of Illness" Studies. 2010. "The Global Burden of Nontyphoidal Salmonella Gastroenteritis." *Clinical Infectious Diseases* 50 (6): 882–89.

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." *Scientific Reports* 6: 33341.

Marlovits TC, Kubori T, Lara-Tejero M, Thomas D, Unger VM, Galán JE. 2006. "Assembly of the Inner Rod Determines Needle Length in the Type III Secretion Injectisome." *Nature* 441 (7093): 637–40.

McDermott JE, Corrigan A, Peterson E, Oehmen C, Niemann G, Cambronne ED, Sharp D, Adkins JN, Samudrala R, Heffron F. 2011. "Computational Prediction of Type III and IV Secreted Effectors in Gram-Negative Bacteria." *Infection and Immunity* 79 (1): 23–32.

McDowell MA, Marcoux J, Mcvicker G, Johnson S, Fong YH, Stevens R, Bowman LAH, Degiacomi MT, Yan J, Wise A, et al. 2016. "Characterisation of Shigella Spa33 and Thermotoga FliM/N Reveals a New Model for C-Ring Assembly in T3SS." *Molecular Microbiology* 99 (4): 749–66.

McShan AC, De Guzman RN. 2015. "The Bacterial Type III Secretion System as a Target for Developing New Antibiotics." *Chemical Biology & Drug Design* 85 (1): 30–42.

Miki T, Okada N, Shimada Y, Danbara H. 2004. "Characterization of Salmonella Pathogenicity Island 1 Type III Secretion-Dependent Hemolytic Activity in Salmonella Enterica Serovar Typhimurium." *Microbial Pathogenesis* 37 (2): 65–72.

Minamino T, Gonzalez-Pedrajo B, Oosawa K, Namba K, Macnab RM. 2002. "Structural Properties of FliH, an ATPase Regulatory Component of the Salmonella Type III Flagellar Export Apparatus." *Journal of Molecular Biology* 322 (2): 281–90.

Minamino T, González-Pedrajo B, Oosawa K, Namba K, Macnab RM. 2002. "Structural Properties of FliH, an ATPase Regulatory Component of the Salmonella Type III Flagellar Export Apparatus." *Journal of Molecular Biology* 322 (2): 281–90.

Minamino T, Kazetani K, Tahara A, Suzuki H, Furukawa Y, Kihara M, Namba K. 2006. "Oligomerization of the Bacterial Flagellar ATPase FliI Is Controlled by Its Extreme N-Terminal Region." *Journal of Molecular Biology* 360 (2): 510–19.

Minamino T, Macnab RM. 2000. "Domain Structure of Salmonella FlhB, a Flagellar Export Component Responsible for Substrate Specificity Switching." *Journal of Bacteriology* 182 (17): 4906–14.

Minamino T, MacNab RM. 2000. "FliH, a Soluble Component of the Type III Flagellar Export Apparatus of Salmonella, Forms a Complex with FliI and Inhibits Its ATPase Activity." *Molecular Microbiology* 37: 1494–1503.

Minamino T, Namba K. 2008. "Distinct Roles of the FliI ATPase and Proton Motive Force in Bacterial Flagellar Protein Export." *Nature* 451: 485–88.

Monack DM, Hersh D, Ghori N, Bouley D, Zychlinsky A, Falkow S. 2000. "Salmonella Exploits Caspase-1 to Colonize Peyer's Patches in a Murine Typhoid Model." *Journal of Experimental Medicine* 192 (2): 249–58.

Monjarás Feria JV, Lefebre MD, Stierhof YD, Galán JE, Wagner S. 2015. "Role of Autocleavage in the Function of a Type III Secretion Specificity Switch Protein in Salmonella Enterica Serovar Typhimurium." *MBio* 6 (5): e01459-15.

Morgner N, Robinson CV. 2012. "Massign: An Assignment Strategy for Maximizing Information from the Mass Spectra of Heterogeneous Protein Assemblies." *Analytical Chemistry* 84 (6): 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." *Journal of Biological Chemistry* 281 (1): 599–607.

Myeni SK, Wang L, Zhou D. 2013. "SipB-SipC Complex Is Essential for Translocon Formation." *PLOS ONE* 8 (3): e60499.

Nallamsetty S, Waugh DS. 2006. "Solubility-Enhancing Proteins MBP and NusA Play a Passive Role in the Folding of Their Fusion Partners." *Protein Expression and Purification* 45 (1): 175–82.

Notti RQ, Bhattacharya S, Lilic M, Stebbins CE. 2015. "A Common Assembly Module in Injectisome and Flagellar Type III Secretion Sorting Platforms." *Nature Communications* 6: 7125.

Novy R, Held D, Yaeger K, Mierendorf R. 2002. "Coexpression of Multiple Target Proteins in E. Coli." *Novagen InNovations Newsletter* 15: 2–6.

Paraskevopoulou V, Falcone FH. 2018. "Polyionic Tags as Enhancers of Protein Solubility in Recombinant Protein Expression." *Microorganisms* 6 (2): 47.

Park D, Lara-Tejero M, Waxham MN, Li W, Hu B, Galán JE, Liu J. 2018. "Visualization of the Type III Secretion Mediated Salmonella–Host Cell Interface Using Cryo-Electron Tomography." *eLife* 7: e39514.

Paul K, Erhardt M, Hirano T, Blair DF, Hughes KT. 2008. "Energy Source of Flagellar Type III Secretion." *Nature* 451: 489.

Petoukhov MV, Svergun DI. 2005. "Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data." *Biophysical Journal* 89 (2): 1237–50.

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." *Journal of Computational Chemistry* 25 (13): 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." *Journal of Computational Chemistry* 26 (16): 1781–1802.

Pinaud L, Sansonetti PJ, Phalipon A. 2018. "Host Cell Targeting by Enteropathogenic Bacteria T3SS Effectors." *Trends in Microbiology* 26 (4): 266–83.

Portaliou AG, Tsolis KC, Loos MS, Balabanidou V, Rayo J, Tsirigotaki A, Crepin VF, Frankel G, Kalodimos CG, Karamanou S, Economou A. 2017. "Hierarchical Protein Targeting and Secretion Is Controlled by an Affinity Switch in the Type III Secretion System of Enteropathogenic Escherichia Coli." *EMBO Journal*.

Portaliou AG, Tsolis KC, Loos MS, Zorzini V, Economou A. 2016. "Type III Secretion: Building and Operating a Remarkable Nanomachine." *Trends in Biochemical Sciences* 41 (2): 175–89.

Poyraz Ö, Schmidt H, Seidel K, Delissen F, Ader C, Tenenboim H, Goosmann C, Laube B, Thünemann AF, Zychlinsky A, et al. 2010. "Protein Refolding Is Required for Assembly of the Type Three Secretion Needle." *Nature Structural & Molecular Biology* 17 (7): 788-792.

Radics J, Königsmaier L, Marlovits TC. 2014. "Structure of a Pathogenic Type 3 Secretion System in Action." *Nature Structural & Molecular Biology* 21 (1): 82–87.

Rathinavelan T, Lara-Tejero M, Lefebre M, Chatterjee S, McShan AC, Guo DC, Tang C, Galan JE, De Guzman RN. 2014. "NMR Model of PrgI–SipD Interaction and Its Implications in the Needle-Tip Assembly of the Salmonella Type III Secretion System." *Journal of Molecular Biology* 426 (16): 2958–69.

Rial DV, Ceccarelli EA. 2002. "Removal of DnaK Contamination during Fusion Protein Purifications." *Protein Expression and Purification* 25 (3): 503–7.

Rocha JM, Richardson CJ, Zhang M, Darch CM, Cai E, Diepold A, Gahlmann A. 2018. "Single-Molecule Tracking in Live Yersinia Enterocolitica Reveals Distinct Cytosolic Complexes of Injectisome Subunits." *Integrative Biology* 10 (9): 502–15.

Roehrich AD, Guillossou E, Blocker AJ, Martinez-Argudo I. 2013. "Shigella IpaD Has a Dual Role: Signal Transduction from the Type III Secretion System Needle Tip and Intracellular Secretion Regulation." *Molecular Microbiology* 87: 690–706.

Romano FB, Tang Y, Rossi KC, Monopoli KR, Ross JL, Heuck AP. 2016. "Type 3 Secretion Translocators Spontaneously Assemble a Hexadecameric Transmembrane Complex." *Journal of Biological Chemistry* 291 (12): 6304–15.

Roth GA, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Abbastabar H, Abd-Allah F, Abdela J, Abdelalim A, et al. 2018. "Global, Regional, and National Age-Sex-Specific Mortality for 282 Causes of Death in 195 Countries and Territories, 1980-2017: A Systematic Analysis for the Global Burden of Disease Study 2017." *The Lancet* 392 (10159): 1736–88.

Samudrala R, Heffron F, McDermott JE. 2009. "Accurate Prediction of Secreted Substrates and Identification of a Conserved Putative Secretion Signal for Type III Secretion Systems." *PLOS Pathogens* 5 (4): e1000375.

Sansonetti PJ, Phalipon A. 1999. "M Cells as Ports of Entry for Enteroinvasive Pathogens: Mechanisms of Interaction, Consequences for the Disease Process." *Seminars in Immunology* 11 (3): 193–203.

Santos RL, Zhang S, Tsolis RM, Kingsley RA, Garry Adams L, Bäumler AJ. 2001. "Animal Models of Salmonella Infections: Enteritis versus Typhoid Fever." *Microbes and Infection* 3 (14): 1335–44.

Schraidt O, Lefebre MD, Brunner MJ, Schmied WH, Schmidt A, Radics J, Mechtler K, Galán JE, Marlovits TC. 2010. "Topology and Organization of the Salmonella Typhimurium Type III Secretion Needle Complex Components." *PLOS Pathogens* 6 (4): e1000824.

Schraidt O, Marlovits TC. 2011. "Three-Dimensional Model of Salmonella's Needle Complex at Subnanometer Resolution." *Science* 331 (6021): 1192–95.

Seidel SAI, Dijkman PM, Lea WA, van den Bogaart G, Jerabek-Willemsen M, Lazic A, Joseph JS, Srinivasan P, Baaske P, Simeonov A, et al. 2013. "Microscale Thermophoresis Quantifies Biomolecular Interactions under Previously Challenging Conditions." *Methods* 59 (3): 301–15.

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." *Nature Communications* 8: 14737.

Sorg I, Wagner S, Amstutz M, Müller SA, Broz P, Lussi Y, Engel A, Cornelis GR. 2007. "YscU Recognizes Translocators as Export Substrates of the Yersinia Injectisome." *EMBO Journal* 26: 3015–24.

Soto E, Espinosa N, Díaz-Guerrero M, Gaytán MO, Puente JL, González-Pedrajo B. 2017. "Functional Characterization of EscK (Orf4), a Sorting Platform Component of the Enteropathogenic Escherichia Coli Injectisome." *Journal of Bacteriology* 199 (1): e00538-16.

Spaeth KE, Chen YS, Valdivia RH. 2009. "The Chlamydia Type III Secretion System C-Ring Engages a Chaperone-Effector Protein Complex." *PLOS Pathogens* 5 (9): e1000579.

Stebbins CE, Galán JE. 2001a. "Maintenance of an Unfolded Polypeptide by a Cognate Chaperone in Bacterial Type III Secretion." *Nature* 414 (6859): 77–81.

Stebbins CE, Galán JE. 2001b. "Structural Mimicry in Bacterial Virulence." *Nature* 412 (6848): 701–5.

Stebbins CE, Galán JE. 2003. "Priming Virulence Factors for Delivery into the Host." *Nature Reviews Molecular Cell Biology* 4 (9): 738–43.

Steele-Mortimer O. 2008. "The Salmonella-Containing Vacuole—Moving with the Times." *Current Opinion in Microbiology* 11 (1): 38–45.

Sukhan A, Kubori T, Wilson J, Galán JE. 2001. "Genetic Analysis of Assembly of the Salmonella Enterica Serovar Typhimurium Type III Secretion-Associated Needle Complex." *Journal of Bacteriology* 183 (4): 1159–67.

Svergun DI. 1992. "Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria." *Journal of Applied Crystallography* 25 (4): 495–503.

Svergun DI. 1999. "Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing." *Biophysical Journal* 76 (6): 2879–86.

Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, Pulcini C, Kahlmeter G, Kluytmans J, Carmeli Y, et al. 2018. "Discovery, Research, and Development of New Antibiotics: The WHO Priority List of Antibiotic-Resistant Bacteria and Tuberculosis." *The Lancet Infectious Diseases* 18 (3): 318–27.

Terashima H, Kawamoto A, Tatsumi C, Namba K, Minamino T, Imada K. 2018. "In Vitro Reconstitution of Functional Type III Protein Export and Insights into Flagellar Assembly." *MBio* 9 (3): e00988-18.

Torruellas J, Jackson MW, Pennock JW, Plano GV. 2005. "The Yersinia Pestis Type III Secretion Needle Plays a Role in the Regulation of Yop Secretion." *Molecular Microbiology* 57 (6): 1719–33.

Troisfontaines P, Cornelis GR. 2005. "Type III Secretion: More Systems Than You Think." *Physiology* 20: 326–39.

Tsirigotaki A, De Geyter J, Šoštarić N, Economou A, Karamanou S. 2016. "Protein Export through the Bacterial Sec Pathway." *Nature Reviews Microbiology* 15: 21-36.

Tuli A, Sharma M. 2019. "How to Do Business with Lysosomes: Salmonella Leads the Way." *Current Opinion in Microbiology* 47: 1–7.

Tuukkanen AT, Kleywegt GJ, Svergun DI. 2016. "Resolution of Ab Initio Shapes Determined from Small-Angle Scattering." *IUCrJ* 3: 440–47.

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, et al. 2006. "Improving the Performance of a Quadrupole Time-of-Flight Instrument for Macromolecular Mass Spectrometry." *Analytical Chemistry* 78 (21): 7473–83.

van der Velden AWM, Velasquez M, Starnbach MN. 2003. " Salmonella Rapidly Kill Dendritic Cells via a Caspase-1- Dependent Mechanism ." *The Journal of Immunology* 171 (12): 6742–49.

van Duijn E, Simmons DA, van den Heuvel RHH, Bakkes PJ, van Heerikhuizen H, Heeren RMA, Robinson CV, van der Vies SM, Heck AJR. 2006. "Tandem Mass Spectrometry of Intact GroEL–Substrate Complexes Reveals Substrate-Specific Conformational Changes in the Trans Ring." *Journal of the American Chemical Society* 128 (14): 4694–4702.

Vazquez-Torres A, Jones-Carson J, Bäumler AJ, Falkow S, Valdivia R, Brown W, Le M, Berggren R, Parks WT, Fang FC. 1999. "Extraintestinal Dissemination of Salmonella by CD18-Expressing Phagocytes." *Nature* 401 (6755): 804–8.

Veenendaal AKJ, Hodgkinson JL, Schwarzer L, Stabat D, Zenk SF, Blocker AJ. 2007. "The Type III Secretion System Needle Tip Complex Mediates Host Cell Sensing and Translocon Insertion." *Molecular Microbiology* 63 (6): 1719–30.

Volkov VV, Svergun DI. 2003. "Uniqueness of Ab Initio Shape Determination in Small-Angle Scattering." *Journal of Applied Crystallography* 36: 860–64.

Wagner S, Grin I, Malmsheimer S, Singh N, Torres-Vargas CE, Westerhausen S. 2018. "Bacterial Type III Secretion Systems: A Complex Device for the Delivery of Bacterial Effector Proteins into Eukaryotic Host Cells." *FEMS Microbiology Letters* 365 (19): 1–13.

Wagner S, Königsmaier L, Lara-Tejero M, Lefebre M, Marlovits TC, Galán JE. 2010. "Organization and Coordinated Assembly of the Type III Secretion Export Apparatus." *Proceedings of the National Academy of Sciences of the United States of America* 107 (41): 17745–50.

Wattiau P, Bernier B, Deslée P, Michiels T, Cornelis GR. 1994. "Individual Chaperones Required for Yop Secretion by Yersinia." *Proceedings of the National Academy of Sciences of the United States of America* 91 (22): 10493–97.

Wee DH, Hughes KT. 2015. "Molecular Ruler Determines Needle Length for the Salmonella Spi-1 Injectisome." *Proceedings of the National Academy of Sciences of the United States of America* 112 (13): 4098–4103.

Worrall LJ, Hong C, Vuckovic M, Deng W, Bergeron JRC, Majewski DD, Huang RK, Spreter T, Finlay BB, Yu Z, et al. 2016. "Near-Atomic-Resolution Cryo-EM Analysis of the Salmonella T3S Injectisome Basal Body." *Nature* 540 (7634): 597–601.

Worrall LJ, Vuckovic M, Strynadka NCJ. 2010. "Crystal Structure of the C-Terminal Domain of the Salmonella Type III Secretion System Export Apparatus Protein InvA." *Protein Science* 19: 1091–1096.

Xing Q, Shi K, Portaliou AG, Rossi P, Economou A, Kalodimos CG. 2018. "Structures of Chaperone-Substrate Complexes Docked onto the Export Gate in a Type III Secretion System." *Nature Communications* 9 (1): 1773.

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. "The I-TASSER Suite: Protein Structure and Function Prediction." *Nature Methods* 12 (1): 7–8.

Yoshida Y, Miki T, Ono S, Haneda T, Ito M, Okada N. 2014. "Functional Characterization of the Type III Secretion ATPase SsaN Encoded by Salmonella Pathogenicity Island 2." *PLOS ONE* 9 (4): e94347.

Yu XJ, Liu M, Matthews S, Holden DW. 2011. "Tandem Translation Generates a Chaperone for the Salmonella Type III Secretion System Protein SsaQ." *Journal of Biological Chemistry* 286 (41): 36098–107.

Yu XJ, McGourty K, Liu M, Unsworth KE, Holden DW. 2010. "PH Sensing by Intracellular Salmonella Induces Effector Translocation." *Science* 328 (5981): 1040–43.

Zarivach R, Deng W, Vuckovic M, Felise HB, Nguyen HV., Miller SI, Finlay BB, Strynadka NCJ. 2008. "Structural Analysis of the Essential Self-Cleaving Type III Secretion Proteins EscU and SpaS." *Nature* 453 (7191): 124–27.

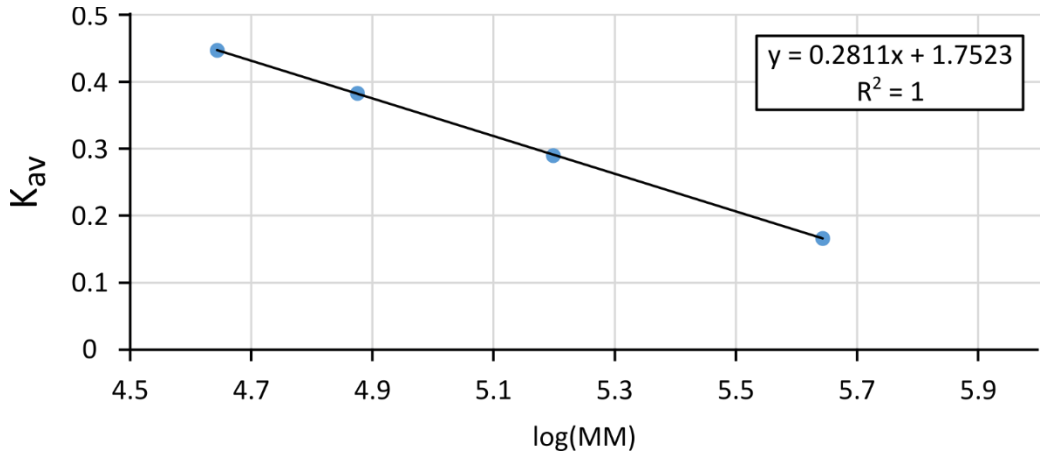
Zarivach R, Vuckovic M, Deng W, Finlay BB, Strynadka NCJ. 2007. "Structural Analysis of a Prototypical ATPase from the Type III Secretion System." *Nature Structural & Molecular Biology* 14 (2): 131–37.

Zhang Y, Lara-Tejero M, Bewersdorf J, Galán JE. 2017. "Visualization and Characterization of Individual Type III Protein Secretion Machines in Live Bacteria." *Proceedings of the National Academy of Sciences of the United States of America* 114 (23): 6098–6103.

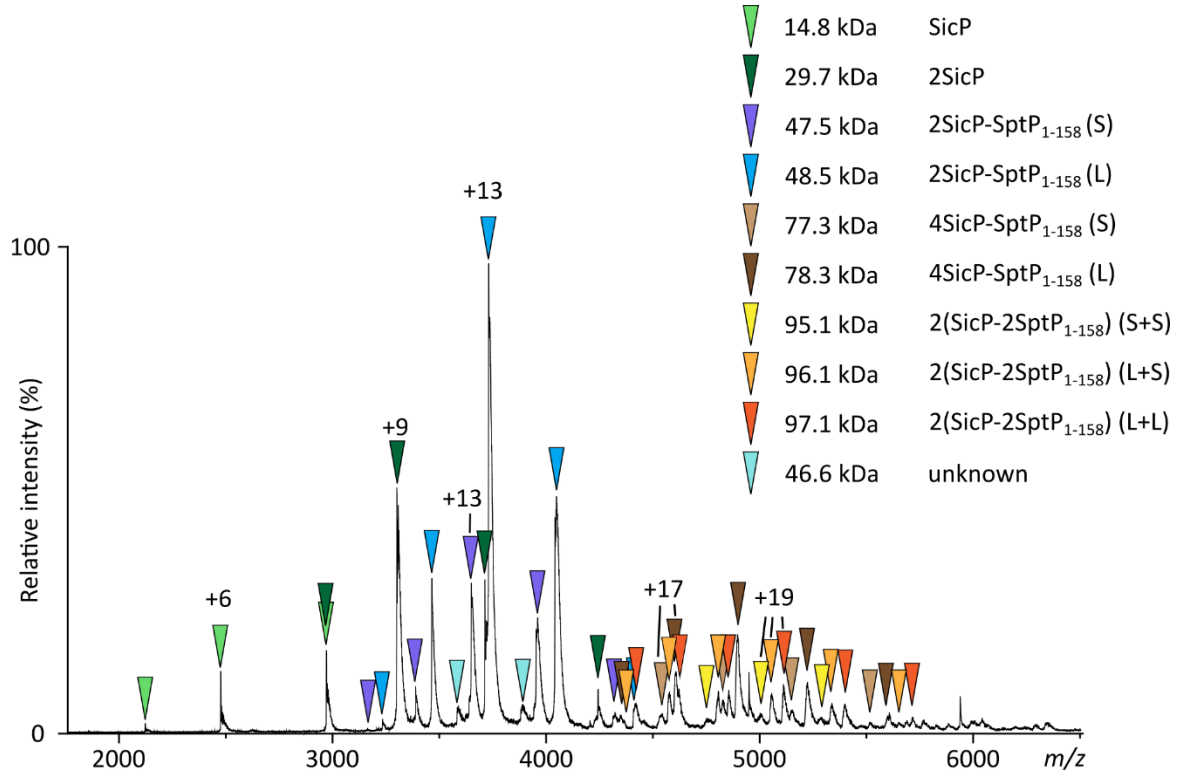
Zilkenat S, Franz-Wachtel M, Stierhof YD, Galán JE, Macek B, Wagner S. 2016. "Determination of the Stoichiometry of the Complete Bacterial Type III Secretion Needle Complex Using a Combined Quantitative Proteomic Approach." *Molecular and Cellular Proteomics* 15 (5): 1598–1609.

6. Appendix

6.1. Supplementary Figures



Supplementary Figure 1. Calibration curve for analytical size-exclusion chromatography on the Superdex 200 10/300 (GL) column. K_{av} is defined as $K_{av} = (V_e-V_0)/(V_C-V_0)$, where V_e is the elution volume of a protein, V_0 the void volume of the column and V_C the total column volume. MM is the molecular mass of a protein in Da. Ovalbumin (44 kDa), conalbumin (75 kDa), aldolase (158 kDa) and ferritin (440 kDa) were used in generating this calibration curve, V_0 was determined using Blue Dextran 2000. The equation of the regression curve and the R^2-value as quality-of-fit indicator are displayed in the top right.



Supplementary Figure 2. Native mass spectrometry of SicP/SptP$_{1-158}$. SptP$_{1-158}$ is present as both a long form (18.8 kDa, indicated as "L") and a shorter degradation product (17.8 kDa, indicated as "S"). The 46.6 kDa species could not be clearly identified but might indicate the existence of an even shorter form of SptP$_{1-158}$. Native MS analysis was performed by Johannes Heidemann.

SpaO$_C$:

```
        10         20         30         40         50         60
METLDIQHIEEENNTTETAETLPGLNQLP VKLEFVLYRKNVTLAELEAMGQQQLLSLPTN
----------flex.-------------- ----------PDB ID 4YX1----------

        70         80         90        100
AELNVEIMANGVLLGNGELVQMNDTLGVEIHEWL SESGNGE
---------------------------------- -flex.-
```

SpaO:

```
        10         20         30         40         50         60
MSLRVRQIDRREWLLAQTATECQRHGREATLEYPTRQGMWVRLSDAEKRWSAWIKPGDWL
------------------------------------------------------------

        70         80         90        100        110        120
EHVSPALAGAAVSAGAEHLVVPWLAATERPFELPVPHLSCRRLCVENPVPGSALPEGKLL
----I-TASSER model, validated by SAXS ----------------------

       130        140        150        160        170        180
HIMSDRGGLWFE HLPELPAVG GGRPKMLRWPLRFVIGSSDTQRSLLGRIGIGDVLLIRTS
------------ --flex.--     ----------------------PDB ID 4YX5----

       190        200        210        220        230        240
RAEVYCYAKKLGHFNRVEGGIIVETLDIQ HIEEENNTTETAETLPGLNQLP VKLEFVLYR
---------------------------- ---------flex.-------- ---------

       250        260        270        280        290        300
KNVTLAELEAMGQQQLLSLPTNAELNVEIMANGVLLGNGELVQMNDTLGVEIHEWL SESG
----------------PDB ID 4YX5-----------------------------

NGE
```

OrgB 1 (interacts with SpaO and InvC):

```
        10         20         30         40         50          60
MLKNIPIPSPLSPVEGILIKRKTLER YFSI ERLEQQAHQRAKRIL REAEEE AKTLRMYAY
-------PDB ID 4YX7-------- flex --------------- -flex- ---------

        70         80         90        100        110        120
QEGYEQGMIDALQQVA AYLTD NQTMAWKWMEKIQIYARELFSAA VDHPE TLLTVLDEWLR
---------------- -flex ----------------------- -flex -----------

       130        140        150        160        170        180
DFDKPEGQLFLTLPVNAKKDHQKLMVLLMENWPGTFNLKYHQEQRFIMSCGDQIAEF SPE
--all light blue: I-TASSER model------------------------- fl.

       190        200        210        220
QFVETAVGVIKHHL DEL PQDCRTISDNAINALIDEWKTKTQAEVIR
-------------- -fl. ----------------------------
```

OrgB 2:

```
        10         20         30         40         50         60
MLKNIPIPSPLSPVEGILIK RKTLERYFSIERLEQQAHQRAKRIL REAEEE AKTLRMYAY
-------flex--------- ------------------------ -flex- ---------

        70         80         90        100        110        120
QEGYEQGMIDALQQVA AYLTD NQTMAWKWMEKIQIYARELFSAA VDHPE TLLTVLDEWLR
---------------- -flex ----------------------- -flex -----------

       130        140        150        160        170        180
DFDKPEGQLFLTLPVNAKKDHQKLMVLLMENWPGTFNLKYHQEQRFIMSCGDQIAEF SPE
--all light blue: I-TASSER model------------------------- fl.

       190        200        210        220
QFVETAVGVIKHHL DEL PQDCRTISDNAINALIDEWKTKTQAEVIR
-------------- -fl. ----------------------------
```

InvC-Strep:

```
        10         20         30         40         50         60
MKTPRLLQYLAYPQKITGPIIEAELRDVAIGELCEIRRGWHQKQVVARAQVVGLQRERTV
------------------------------------------------------------

        70         80         90        100        110        120
LSLIGNAQGLSRDVVLYPTGRALSAWVGYSVLGAVLDPTGKIVERFTPEVAPISEERVID
------------------------------------------------------------

       130        140        150        160        170        180
VAPPSYASRVGVREPLITGVRAIDGLLTCGVGQRMGIFASAGCGKTMLMHMLIEQTEADV
------------------------------------------------------------

       190        200        210        220        230        240
FVIGLIGERGREVTEFVDMLRASHKKEKCVLVFATSDFPSVDRCNAAQLATTVAEYFRDQ
----------------------I-TASSER model-------------------------

       250        260        270        280        290        300
GKRVVLFIDSMTRYARALRDVALASGERPARRGYPASVFDNLPRLLERPGATSEGSITAF
------------------------------------------------------------

       310        320        330        340        350        360
YTVLLESEEEADPMADEIRSILDGHLYLSRKLAGQGHYPAIDVLKSVSRVFGQVTTPTHA
------------------------------------------------------------

       370        380        390        400        410        420
EQASAVRKLMTRLEELQLFIDLGEYRPGENIDNDRAMQMRDSLKAWLCQPVAQYSSFDDT
------------------------------------------------------------

       430        440
LSGMNAFADQNSAWSHPQFEK
---------------------
```

Supplementary Figure 3. Protein regions based on crystal structures and fragment-based models in the rigid-body modeling of the SpaO/SpaOc/OrgB/InvC complex against SAXS data (Figure 2.31). The amino acid sequence of each construct is shown. Dark blue indicates the use of an available crystal structure, light blue use of a fragment-based model (I-TASSER) and green regions were modeled as flexible linkers.

6.2. Supplementary Tables

Supplementary Table 1. SAXS data acquisition, data analysis, modeling and software used in the analysis of SpaO/SpaOC/OrgB/InvC. R_g: radius of gyration; D_{max}: maximum particle dimension.

A) Sample details	
Tags	C-terminal Strep-tag on InvC
Extinction coefficient ε (M^{-1} cm^{-1}, 280 nm)	186,670† & 256,110#
Molecular mass from chemical composition (kDa)	158† & 214#
Injection volume (µl)	75
Concentration, (mg ml^{-1})	17.6†/ 17.4#
Flow rate (ml min^{-1})	0.3
B) SAXS data collection parameters	
Source, instrument and description or reference	P12 (EMBL/DESY, storage ring PETRA III, Germany)
Wavelength (Å)	1.24
Beam geometry (size, sample-to-detector distance)	0.2 x 0.12 mm^2, 3.0 m
q-measurement range ($Å^{-1}$)	0.008 – 0.47
Method for monitoring radiation damage, X-ray dose	BECQUEREL software
Exposure time, number of exposures	1 s, 3600x
Flow rate	0.3 ml/min
Sample temperature (K)	283
C) Software employed for SAS data reduction, analysis and interpretation	
SAS data reduction to sample–solvent scattering, and extrapolation, merging, desmearing	PRIMUS
Calculation of ε from sequence	PROTPARAM
Basic analyses: Guinier, *P*(*r*), scattering particle volume (*e.g.* Porod volume V_P or volume of correlation V_c)	PRIMUS
Shape/bead modeling	DAMMIF, SASRES
Atomic structure modeling (homology, rigid body, ensemble)	CORAL
Modeling of missing sequence from PDB files	CORAL
Molecular graphics	PYMOL

D) Structural parameters	
Guinier Analysis	
I(0) (Arbitrary units)	4,149.32 ± 71.83
R_g (nm)	5.7 ± 0.6
q-range (Å^{-1})	0.01813 - 0.3
Fidelity of Primus Guinier analysis	0.63
***P*(*r*) analysis**	
I(0) (arbitrary units)	4,592.00 ± 459
R_g (nm)	7.0 ± 0.7
D_{max} (nm)	22.7 ± 2.3
q-range (Å^{-1})	0.01813 - 0.3
Volume Porod (nm^3)	302

E) Shape modeling results (DAMMIF)	
q-range for fitting (Å^{-1})	0.01813 - 0.3
Ambiguity measured by AMBIMETER	2.127
SASRES resolution (Å)	64 ± 5
Molecular mass estimate (kDa)	208

F) Atomistic modeling	
Method	CORAL
q-range for fitting (Å^{-1})	0.01813 - 0.3
χ^2 value	1.1231

G) Data Deposition	
SASBDB data and model ID	SASDEJ7

† based on the SpaO-2SpaO$_C$-2OrgB-InvC stoichiometry of the SAXS hybrid model
# based on the 2SpaO-4SpaO$_C$-2OrgB-InvC stoichiometry from native MS

Supplementary Table 2: List of oligonucleotides used in this study

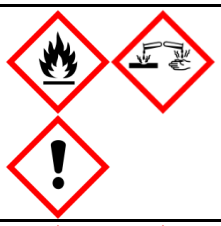
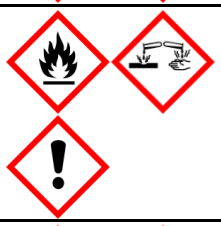
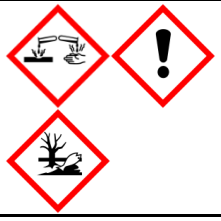
Construct/Purpose	Name	Sequence (5' – 3')
pASK-IBA3C+RBS (transfer of pASK-IBA3+ RBS to pASK-IBA3C)	IBA3 RBS Fw	AGCCTTCTTATTCGGCCTTG
	IBA3 RBS Re	GGGTACCGAGCTCGAATTC
orgB in pCDFDuet-1	Duet OrgB Fw	AATTAACCATGGTCAAAAATATCCCAATACCG TCC
	Duet OrgB Re	AATTAAGGATCCTCATCACCTTATAACCTCCG CTTGCG
	Duet OrgBHis Re	AATTAAGGATCCTCATCAGTGGTGGTGGTGGT G
spaO in pCDFDuet-1 or pETDuet-1	Duet SpaO Fw	AATTAACATATGTCATTGCGTGTGAGACAG
	Duet SpaO Re	AATTAAGGTACCTCATCATTCCCCATTACCAG ACTCG
invC in pASK-IBA3+, pASK-IBA3C or pASK-IBA3C+RBS	InvC IBA3 Fw	AATTAAGGTCTCTAATGAAAACACCTCGTTTA CTGCAATATC
	InvC IBA3 Re	AATTAAGGTCTCTGCGCTATTCTGGTCAGCGA ATGCATTC
invC in pACYCDuet-1	InvC Duet Fw	AATTAACCATGGGTATGAAAACACCTCGTTTA CTGCAATATC
	InvC-Strep Duet Re	AATTAAAAGCTTTTATTATTTTTCGAACTGCG GGTGG
	InvC-notag Duet Re	AATTAAAAGCTTTTATTAATTCTGGTCAGCGA ATGCATTC
invI in pACYCDuet-1	InvI Duet Fw	AATTAACATATGCATTCGCTGACCAGAATTAA AG
	Strep-InvI Duet Fw	AATTAACATATGGCTTGGAGCCACCCGCAGTT CGAAAAAGGCGCCATGCATTCGCTGACCAGA ATTAAAG
	InvI Duet Re	AATTAAGGTACCTTATTAAATTATCTCCTCTG ACTCGGCCTC
$invA_{357–685}$ in pACYCDuet-1	InvA Duet Fw	AATTAACCATGGGTACAGAGACCGTACCGTTG ATATTAC
	InvA Duet Re	AATTAAGAATTCTTATTATATTGTTTTTATAAC ATTCACTGACTTG
$invA_{357–685}$ in pASK-IBA5+	InvAnotag 5+ Fw	AATTAAGCTAGCACAGAGACCGTACCGTTGAT ATTAC
	InvA 5+ Re	AATTAAAAGCTTTTATTATATTGTTTTTATAAC ATTCACTGACTTG
orgA in pASK-IBA3+	OrgA IBA3 Fw	AATTAAGGTCTCGAATGATAAGGCGAAATCGT CAAATG
	OrgA-Strep IBA3 Re	AATTAAGGTCTCTGCGCTACAGGCGAAAGCG GGGAC

Construct/Purpose	Name	Sequence (5' – 3')
orgA in pCOLADuet-1	OrgA Duet Fw	AATTAACCATGGGTATGATAAGGCGAAATCGTCAAATG
	OrgA-Strep Duet Re	AATTAAGGATCCTTATTATTTTTCGAACTGCGGGTG
	OrgA-notag Duet Re	AATTAAGGATCCTTATTAACAGGCGAAAGCGGGGAC
$prgH_{1\text{-}140}$ in pCOLADuet-1	PrgH Duet Fw	AATTAACATATGGAAACATCAAAAGAGAAGACGATAAC
	PrgH-Strep Duet Re	AATTAAGGTACCTTATTATTTTTCGAACTGCGGGTG
	PrgH-notag Duet Re	AATTAAGGTACCTTATTAGTTTTTAAAACGCGGCTCGTTC
$prgH_{1\text{-}140}$ in pASK-IBA3C+RBS	PrgH140 IBA3 Fw	AATTAAGGTCTCGAATGGAAACATCAAAAGAGAAGACG
	PrgH140notag IBA3 Re	AATTAAGGTCTCTGCGCTTCATCAGTTTTTAAAACGCGGCTCG
	PrgH140Strep IBA3 Re	AATTAAGGTCTCTGCGCTGTTTTTAAAACGCGGCTCG
	PrgH140His IBA3 Re	AATTAAGGTCTCTGCGCTTCATCAGTGGTGGTGGTGGTG
$orgB_{1\text{-}105}$ in pET28a	OrgB 1-105 pET28a Fw	AATTAACCATGGGTATGCTCAAAAATATCCCAATACC
	OrgB 1-105-His pET28a Re	AATTAACTCGAGGACCGCAGCTGAAAATAACTC
$orgB_{106\text{-}226}$ in pET28a	OrgB His-106-226 pET28a Fw	AATTAACATATGGACCATCCCGAAACGCTTTTAAC
	OrgB 106-226 pET28a Re	AATTAACTCGAGTCATCACCTTATAACCTCCGCTTGCG
$invC_{1\text{-}79}$ in pASK-IBA3C+RBS	InvC1-79 IBA3 Fw	AATTAAGGTCTCTAATGAAAACACCTCGTTTACTGCAATATC
	InvC 1-79-Strep IBA3 Re	AATTAAGGTCTCTGCGCTAGTGGGATAAAGCACGACATC
$invC_{80\text{-}431}$ in pASK-IBA3+	InvC Delta79 Fw	AATTAAGGTCTCTAATGGGACGTGCGTTATCGGCGTG
	InvC IBA3 Re	AATTAAGGTCTCTGCGCTATTCTGGTCAGCGAATGCATTC
orgC in pASK-IBA3+	OrgC IBA3 Fw	AATTAAGGTCTCTAATGATACCGGGTACGATTCCGACTTC
	OrgC IBA3 Re	AATTAAGGTCTCTGCGCTCCAGTCAATTGCCTCTTTGTTTTC
$spaO_C$ in pASK-IBA3+	spaO′3C BsaI Fw	AAAAAAGGTCTCAAATGGAAACGTTAGATATTCA
	spaO′3C BsaI Rv	AAAAAAGGTCTCGGCGCTTTCCCCATTACCAGACTCGC

Construct/Purpose	Name	Sequence (5' – 3')
Sequencing of Duet vectors	Duet Seq MCS1 Fw	TTATGCGACTCCTGCATTAG
	Duet Seq MCS1 Re	GCCGTGTACAATACGATTAC
	Duet Seq MCS2 Fw	TTGTACACGGCCGCATAATC

6.3. Hazardous Substances

Substance	GHS symbols	Hazard statements	Precautionary statements
2-mercaptoethanol (2-ME)		H301+H331-H310-H315-H317-H318-H373-H410	P273-P280-P302+P352-P304+P340-P305+P351+P338-P310
Acetic acid		H226-H314	P210-P280-P301+P330+P331-P303+P361+P353-P305+P351+P338-P310
Acetone		H225-H319-H336	P210-P240-305+P351+P338-P403+P233
Ammonium persulfate		H334-H272-H302-H335-H315-H319-H317	P261-P280-P302+P352-P305+P351+P338-P332+P313-P337+P313
Carbenicillin		H317-H334	P272-P302+P352-P333+P313-P342+P311
Chloramphenicol		H351	P201-P308+P313
Complete Ultra EDTA-free protease inhibitor cocktail		H314-H412	P260-P273-P280-P303+P361+P353-P304+P340+P310-P305+P351+P338+P310
DNase I		H317-H334	P261-P280-P284-P304+P340-P333+P313-P342+P311
Ethidium bromide		H331-H341	P304+P340-P201-P311-P261-P308+P313-P271
Ethylenediaminetetraacetic acid (EDTA)		H332-H373	P260-P304+P340-P312
Gel loading dye (6x), purple, no SDS		H412	P273-P501

HCl		H290-H314-H335	P280-P303+P361+P353-P304+P340-P305+P351+P338-P310
Imidazole		H302-H314-H360D	P202-P270-P280-P305+P351+P338-P310
Kanamycin		H360D	P201-P308+P313
Methanol		H225-H301-H311-H331-H370	P210-P280-P301+P310-P302+P352-P304+P340-P240
NaOH		H290-H314	P233-P280-P303+P361+P353-P305+P351+P338-P310
Roti-Free stripping buffer		H317-H319-H412	P280-P302+P352-P305+P351+P338
Rotiphorese Gel 30 (37.5:1)		H302-H315-H317-H319-H340-H350-H361f-H372	P201-P280-P301+P312-P302+P352-P305+P351+P338-P308+P313
Sodium dodecyl sulfate (SDS)		H228-H302+H332-H315-H318-H335-H412	P210-P261-P280-P301+P312+P330-P305+P351+P338+P310
Streptomycin		H302-H361fd	P201-P301+P312+P330-P308+P313
Tetramethylethylenediamine (TEMED)		H225-H302-H332-H314	P210-P303+P361+P353-P305+P351+P338-P310-P405-P501
Trichloroacetic acid (TCA)		H314-H335-H410	P280-P301+P330+P331-P303+P361+P353-P305+P351+P338-P310

Xylene cyanol FF		H319-H335	P261-P264-P280-P304+P340-P305+P351+P338-P337+P313

6.4. List of Figures

Supplementary Figures

6.5. List of Tables

Supplementary Tables

6.6. Abbreviations

2-ME	2-mercaptoethanol
A_{280}	Absorbance measured at 600 nm
AHT	Anhydrotetracycline hydrochloride
ATP	Adenosine triphosphate
ATPγS	Adenosine 5'-(γ-thio)-triphosphate
CET	Cryo-electron tomography
CID	Collision-induced dissociation
DTB	Desthiobiotin
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
ESI	Electro-spray ionization
HEPES	2-[4-(2-Hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HPLC	High-pressure liquid chromatography
IEX	Ion-exchange chromatography
IMAC	Immobilized metal ion affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
MALS	Multi-angle light scattering
MS	Mass spectrometry
OD_{600}	Optical density/absorbance measured at 600 nm
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PVDF	Polyvinylidene difluoride
RBS	Ribosome binding site
SAXS	Small-angle X-ray scattering
SDS	Sodium dodecyl sulfate
SEC	Size-exclusion chromatography
SOC	Super optimal broth with catabolite repression
T3SS	Type III secretion system
TCA	Trichloroacetic acid
TEMED	Tetramethylethylenediamine
Tris	2-Amino-2-(hydroxymethyl)propane-1,3-diol

6.7. 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.

Hamburg, den 01.03.2020 Jonathan Börnicke